Iterative feature refinement for accurate undersampled MR image reconstruction

To cite this article: Shanshan Wang et al 2016 Phys. Med. Biol. 61 3291

View the article online for updates and enhancements.

Related content
- Sparsity-promoting orthogonal dictionary updating for image reconstruction from highly undersampled magnetic resonance data
  Jinhong Huang, Li Guo, Qianjin Feng et al.
- Dictionary learning for data recovery in positron emission tomography
  SeyyedMajid Valiollahzadeh, John W Clark Jr and Osama Mawlawi
- Tensor-based dictionary learning for dynamic tomographic reconstruction
  Shengqi Tan, Yanbo Zhang, Ge Wang et al.

Recent citations
- Improved parallel image reconstruction using feature refinement
  Jing Cheng et al.
Iterative feature refinement for accurate undersampled MR image reconstruction

Shanshan Wang\textsuperscript{1,3}, Jianbo Liu\textsuperscript{1,5}, Qiegen Liu\textsuperscript{2,3}, Leslie Ying\textsuperscript{4}, Xin Liu\textsuperscript{1}, Hairong Zheng\textsuperscript{1} and Dong Liang\textsuperscript{1,6}

\textsuperscript{1} Paul C Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518055, People’s Republic of China
\textsuperscript{2} Department of Electronic Information Engineering, Nanchang University, Nanchang 330031, People’s Republic of China
\textsuperscript{3} Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL, USA
\textsuperscript{4} Department of Biomedical Engineering and Department of Electrical Engineering, University at Buffalo, The State University of New York, Buffalo, NY 14260, USA

E-mail: dong.liang@siat.ac.cn

Received 15 November 2015, revised 11 January 2016
Accepted for publication 4 February 2016
Published 1 April 2016

Abstract

Accelerating MR scan is of great significance for clinical, research and advanced applications, and one main effort to achieve this is the utilization of compressed sensing (CS) theory. Nevertheless, the existing CSMRI approaches still have limitations such as fine structure loss or high computational complexity. This paper proposes a novel iterative feature refinement (IFR) module for accurate MR image reconstruction from undersampled $K$-space data. Integrating IFR with CSMRI which is equipped with fixed transforms, we develop an IFR-CS method to restore meaningful structures and details that are originally discarded without introducing too much additional complexity. Specifically, the proposed IFR-CS is realized with three iterative steps, namely sparsity-promoting denoising, feature refinement and Tikhonov regularization. Experimental results on both simulated and \textit{in vivo} MR datasets have shown that the proposed module has a strong capability to capture image details, and that IFR-CS is comparable and even superior to other state-of-the-art reconstruction approaches.

Keywords: magnetic resonance imaging, undersampled image reconstruction, iterative feature refinement

(Some figures may appear in colour only in the online journal)

\textsuperscript{5} Shanshan Wang and Jianbo Liu contributed equally to this paper.
\textsuperscript{6} Author to whom any correspondence should be addressed.
1. Introduction

With the emerging popularity of compressed sensing (CS) theory, fast magnetic resonance imaging (MRI) exploiting sparsity has attracted much attention from researchers to fulfill the need for effective and efficient clinical uses, research and advanced applications (Lustig et al. 2006, 2007, Lustig et al. 2008, Gamper et al. 2008, Liang et al. 2009, Otazo et al. 2010, Liu et al. 2012, Liang et al. 2012, Hollingsworth 2015). In most existing utilizations of CS in MRI (CSMRI), the classical CS formulation is commonly adopted, which is equipped with fixed sparsifying transforms (e.g. finite-difference and wavelet) and L1 norm on sparse coefficients. Nevertheless, although the classical formulation can be solved very efficiently, losses of fine structures or details are usually observed and become especially severe at high acceleration factors (Ravishankar and Bresler 2011, Liang et al. 2011, Patel et al. 2012, Ning et al. 2013, Liu et al. 2013b). To improve the reconstruction accuracy, many efforts have been made (Buades et al. 2005, Aharon et al. 2006, Elad and Aharon 2006, Ravishankar and Bresler 2011, Liu et al. 2013b, Zhu et al. 2014). These endeavours can be roughly categorized into two directions, namely, avoid losing details from the very beginning (Ravishankar and Bresler 2011, Liang et al. 2011, Qu et al. 2012, Liu et al. 2013b, Zhu et al. 2014, Liu et al. 2015) and recover the useful information discarded by the classical CS approach (Bregman 1967, Osher et al. 2005, Chang et al. 2006, Liu et al. 2009, 2013a, Ramani and Fessler 2011).

Since fixed transforms usually prefer certain type of structures, the first direction methods are targeted to develop more powerful sparsifying transforms or exploit nonlocal operations to accurately represent the MR image (Ravishankar and Bresler 2011, Liu et al. 2013b, Zhu et al. 2014). Typical examples include data-driven tight frame (Liu et al. 2015), nonlocal total variation regularization (Liang et al. 2011), patch-based directional wavelet (Qu et al. 2012) and dictionary learning (Ravishankar and Bresler 2011, Liu et al. 2013b, Zhu et al. 2014). In particular, dictionary-learning based magnetic resonance imaging (DLMRI) adaptively capture image structures while promoting sparsity and has shown outstanding performance (Ravishankar and Bresler 2011). Nevertheless, these methods improve the accuracy of MR image reconstruction at the expense of computational efficiency due to either time-consuming online training or huge non-local operations. As a consequence, the computational complexity limits their practical application to MR imaging.

The second group methods, on the other hand, try to restore fine structures which are thrown away by the basic CSMRI approach. To facilitate readers’ understanding of why useful structure information gets discarded, in figure 1 we have provided an overview of the basic CSMRI solved by the alternating minimization (AM) technique, in which context, each iteration is split into two steps: sparsity-promoting denoising and Tikhonov regularization (Huang et al. 2010). As pointed out by Osher et al. (2005), ideally the denoising procedure would decompose the input image into a noise-free part and a residual image which only contains the additive noise. However, it is not fully attainable practically. Some useful information is also included in the residual image. Figure 2 shows an iteration of AM-based CSMRI using finite-difference as the sparsifying transform, where (a) is the reconstruction of the previous CS iteration and the input of the current denoising step, (b) is the denoised image and (c) is the residual image between the before- and after- denoised images. We can observe that figure 2(c) contains not only noise and noise-like artifacts, but also some fine structure information. However, basic CSMRI throws away the residual image as noise and only sends the denoised image to the next regularization step. It is expected that the $k$-space updating for solving the Tikhonov regularization may bring back some details from the acquired $k$-space data, but some would never come back due to undersampling. Bregman distance (Bregman 1967) based iterative regularization methods (IRM) (Osher et al. 2005) address this issue in image
restoration by iteratively adding back the residual image. IRM’s equivalence-Bregman iterative method has also been applied in regularized MRI reconstruction (Chang et al. 2006, Liu et al. 2009, 2013a, Ramani and Fessler 2011). However, IRM has a limitation that it may enforce the useful structures and noise simultaneously and may converge to a noisy image due to the semi-convergence property of the Bregman iteration (Liu et al. 2009).

In this paper, we propose a novel iterative feature refinement module for CS (IFR-CS) based undersampled MR image reconstruction. Without introducing too much additional complexity, we integrate our proposed module into the AM-based CSMRI framework equipped with a fixed transform. Specifically, a feature descriptor is designed to extract fine structure and details from the residual image, which is then added back to the denoised image as the input of the Tikhonov regularization step. This strategy is different from the AM-based CS where the residual image is thrown away, and from the IRM method where the features and noise are both added back. In the proposed framework, the CS-based MR
image reconstruction problem is realized by solving three subproblems: sparsity-promoting update, feature-refined update and Tikhonov-regularized update. Figure 2(d) shows the image with feature refining using our proposed strategy. The difference between the input image (a) and the feature-refined image is shown in figure 2(e). We can see that our denoised result is clearly devoid of the structure and texture lost by the AM-based CS reconstruction. Experimental results on both simulated and in vivo MR datasets have shown that the proposed module has a strong capability to capture image details, and its integration with CSMRI is comparable and even superior to the AM-based CSMRI, DL-based CSMRI and Bregman-iteration based CSMRI.

The rest of this paper is organized as follows. Section 2 introduces the problem formulation and prior work in CSMRI. Our proposed module for accelerating MR imaging is detailed in section 3. Section 4 demonstrates the performance of our method using numerous experiments under different sampling schemes and undersampling ratios. Section 5 discusses the relevant properties of the proposed method. Conclusions and future work are presented in section 6.

2. Problem formulation and related work

In this section, we first describe the formulation of the basic CSMRI model. Then, some typical techniques for solving this formulation are introduced, including AM-based CSMRI, DL-based CSMRI and IRM-based CSMRI. The following notational conventions are used throughout the paper. Let \( I \) represent an image to be reconstructed, and \( f \) denote the raw measurement data in \( k \)-space. In the noiseless case, we have \( F_p I = f \), where \( F_p \) is the undersampled Fourier encoding matrix.

### 2.1. Problem formulation of basic CSMRI

The basic CSMRI model can be formulated as

\[
\min_{I} \| I \|_{L_1} \quad \text{s.t.} \quad F_p I = f
\]

(1)

where \( \| I \|_{L_1} \) is a general \( \ell_1 \) sparsity-promoting functional.

If the data are contaminated with white Gaussian noise, the minimization problem in (1) becomes

\[
\min_{I} \| I \|_{L_1} \quad \text{s.t.} \quad \| F_p I - f \|_2 \leq \sigma
\]

(2)

where \( \sigma \) denotes the standard deviation of the zero-mean complex Gaussian noise. This constrained optimization problem can be described in the following Lagrangian setup

\[
(P1) \quad \min_{I} \frac{1}{2} \| F_p I - f \|_2^2 + \lambda \| I \|_{L_1}
\]

(3)

where the first term \( \| F_p I - f \|_2^2 \) is the data fidelity term in the \( k \)-space domain, and the second term is the regularization term. The parameter \( \lambda \) determines the trade-off between these two terms. There are many ways to solve this problem and some typical techniques are introduced as follows.

### 2.2. AM-based CSMRI

If we introduce an auxiliary variable \( u \), equation (3) can be rewritten as

\[
\min_{I, u} \frac{1}{2} \| F_p I - f \|_2^2 + \lambda \| u \|_{L_1} \quad \text{s.t.} \quad I = u
\]

(4)
Using the strategy of quadratic penalty to relax the equality constraint, it can be approximated by

$$\min_{I,u} \frac{1}{2} \|F_p I - f\|_2^2 + \alpha \|I - u\|_2^2 + \lambda_1 \|u\|_{L1}. \quad (5)$$

According to the variable splitting technique, we apply alternating minimization to solve this unconstrained minimization problem. Firstly, for fixed \(I\), the minimizer \(u\) is obtained by

$$\min_u \|I - u\|_2^2 + \lambda_2 \|u\|_{L1} \quad \text{(6)}$$

where \(\lambda_2 = \lambda_1/\alpha\).

Secondly, for fixed \(u\), the minimization with respect to \(I\) becomes

$$\min_I \|F_p I - f\|_2^2 + \mu \|I - u\|_2^2 \quad \text{(7)}$$

where \(\mu = 2 \times \alpha\).

Equation (6) is the classical image denoising model using the sparsity-promoting regularization term, and equation (7) is the well-known Tikhonov regularization problem. As shown in figure 1, the basic CSMRI problem (3) can be solved iteratively by alternatingly minimizing the sub-problem (6) and (7) individually while keeping the other one fixed. Moreover, we can find that subproblem (6) plays the role of eliminating the noise-like artifacts caused by the random undersampling operation and subproblem (7) can recover some structure information from the acquired data.

Although the AM-based CSMRI model can be computed efficiently (Lustig et al. 2007, Ma et al. 2008, Yang et al. 2010), it may fail to preserve the detailed structure information. The reason is that the fixed and global sparsifying transform has limited representation capability and thus has difficulty representing the rich and different local structural patterns in MR images. The basic CSMRI may lead to staircase artifacts for TV regularization or fail to capture smooth contours for the 2D separable wavelet transform when the measurements are highly undersampled (Ravishankar and Bresler 2011, Liu et al. 2013a).

2.3. DLMRI

Exploring the sparsity and similarity in image patches by learning an adaptive dictionary (Aharon et al. 2006) has drawn considerable attention recently. It has been proven that patch-based dictionaries can capture local structure information effectively, and potentially remove noise and aliasing artifacts in CSMRI without sacrificing resolution. The model of applying the dictionary learning technique to MR image reconstruction (Ravishankar and Bresler 2011) from undersampled \(k\)-space data can be written as

$$\min_{I,D,\Gamma} \sum_i \|R_i I - D \alpha_i\|_2^2 + \chi \|F_p I - f\|_2^2 \quad \text{s.t.} \quad \|\alpha_i\|_0 \leq T_0, \forall \ i$$

where \(D\) represents the dictionary, \(R_i\) denotes the operator that extracts the \(i\)th patch from the image \(I\), \(T_0\) controls the sparsity of the patch representation, and \(\Gamma\) denotes the set \(\{\alpha_i\}\) of sparse representations of all patches. The first term enforces sparsity of the image patches with respect to an adaptive dictionary, while the second term enforces data fidelity in \(k\)-space. To solve this problem, a common strategy is to use an alternating minimization procedure which iterates dictionary learning step (S1) and reconstruction updating step (S2) (Ravishankar and Bresler 2011):
\[
\begin{align*}
\min_{D, I} \sum \| R_I - D \alpha_i \|_2^2 \quad \text{s.t.} \quad \| \alpha_i \|_0 \leq T_0, \; \forall \; i \\
\min_i \sum \| R_I - D \alpha_i \|_2^2 + \zeta \| F_p I - f \|_2^2
\end{align*}
\]

where (S1) and (S2) can be roughly regarded as the special case of equations (6) and (7) with adaptive dictionary.

Although patch-based methodology can capture local structure information effectively and recover MR images more accurately than the basic CSMRI method, the high computational complexity presents an issue in practical application.

2.4. Iterative regularization method via Bregman distance

The iterative regularization method using the Bregman distance was firstly applied by Osher et al. (2005) to the TV based image denoising problem. It was then extended to the wavelet domain (Xu and Osher 2007), and nonlinear inverse scale space (Burger et al. 2006). For any convex function \( E(x) \), the Bregman distance (Bregman 1967) at point \( y \) is defined as:

\[
D_E^p(x, y) = E(x) - E(y) - \langle p, x - y \rangle
\]

where \( p \in \partial E(y) \) is the subgradient of \( E \) at \( y \). Consider the following general problem:

\[
\min_x E(x) + \lambda H(x)
\]

where \( E \) and \( H \) are two convex energy functionals. By exploiting the Bregman distance, it can be solved by the following iterative regularization scheme (Yin et al. 2008, Goldstein and Osher 2009)

\[
\begin{align*}
x^{k+1} &= \arg \min_x D_E^p(x, x^k) + \lambda H(x) \\
p^{k+1} &= p^k - \nabla H(x^{k+1}).
\end{align*}
\]

The Bregman iterative method and its equivalence augmented Lagrangian (AL) method have been used to solve the CSMRI problem. The model of applying Bregman iteration to solve equation (3) can be described as (Chang et al. 2006, Liu et al. 2009):

\[
\begin{align*}
I^{k+1} &= \arg \min_f \frac{1}{2} \| F_p I - f \|_2^2 + \lambda \| f \|_1 \\
f^{k+1} &= f^k + f - F_p I^{k+1}.
\end{align*}
\]

As the operator \( F_p \) is the sub-sampling operation in the \( k \)-space domain, \( f \) denotes the undersampled data, and \( f^0 = 0 \), it is clearly seen that the refinement on the \( k \)-space data only works on the data points that are already acquired and ignores the locations where the data are not sampled. Considering that each \( k \)-space data contains the global information of the image, recovering the spatial structure information corresponding to the unacquired \( k \)-space data is a difficult task. Additionally, the reconstructions will converge to a noisy image as iteration continues due to the semi-convergence property as shown in Liu et al. (2009).
3. Proposed method

Without introducing too much additional complexity, we design a linear feature-refining module to recover useful image details while removing the noise and noise-like artifacts. The proposed IFR-CS consists of three main steps: sparsity-promoting denoising, feature refinement and Tikhonov regularization, whose formulations are

$$\begin{align}
\lambda &= \frac{\mu}{u \cdot I} \\
\mu &= \frac{1}{u \cdot I \cdot T} \\
\nu &= \frac{1}{u \cdot I \cdot (\mu + \mu)}
\end{align}$$

where $v = I - u$, $T$ is the feature descriptor and $I$ is the feature-refined denoised image.

In the following sections, each module in our framework will be introduced in detail individually. Figure 4 illustrates a visual example of one iteration using IFR-CS with TV seminorm as the representative regularization term. Please note that other sparsifying transforms can also be used and the examples with wavelet transform are shown in figure 17.

3.1. Sparsity-promoting denoising module

This module is classical in image denoising and many algorithms have been developed to solve equation (12a). In this paper, if the TV regularization is used, the image denoising problem (6) is solved by the projected gradient method in Chambolle (2004) and Peyré (2011). Specifically, the following two computations are iterated:

$$\begin{align}
u^{i+1} &= I - \lambda_{TV} \cdot \xi^i \\
\xi^{i+1} &= P(\xi^i + \tau \nabla u^{k+1})
\end{align}$$

where $\lambda_{TV}$ denotes the regularization parameter and $\xi^i$ is the vector field, and the projection $P(\xi) = \xi \max(1, |\xi|)$ is defined in a point-wise way. For details on the derivation of this method please refer to Chambolle (2004).

Since the fixed and global sparsifying transform used in this model cannot be optimal in sparsely representing all MR images, the residual image $v = I - u$ after denoising usually contains not only noise, but also the noise-like artifacts caused by undersampling and some useful structure information, as shown in figure 2. Therefore, it is highly desirable to extract this useful information from the residual image $v$ to refine the input of the next module for improved reconstruction.
3.2. Feature refinement

To preserve more details and filter out the noise and noise-like artifacts in reconstruction, we propose to use a feature descriptor to pick the structure information from the residual image $v$. There are many models to represent the feature-related map, and most of them only involve local contrast comparison (Tao et al 2010, Buades et al 2010). Motivated by the

Figure 4. One iteration of our proposed framework. The red dashed rectangle denotes the sparsity regularization module and the blue dashed rectangle is the feature refinement module in our proposed framework.
idea of structure SIMilarity (SSIM) (Wang et al. 2004) which is an image quality assessment criterion consisting of luminance comparison, local contrast comparison and structure correlation, we developed a new feature descriptor exploiting the properties of contrast variation and structure correlation (Liu et al. 2011). The presented image decomposition framework incorporating this feature descriptor shows better separation results of structure from texture when compared to the state-of-the-art decomposition approaches (Liu et al. 2011), which indicates the effectiveness of our feature descriptor.

Specifically, a Gaussian filter is used first to blur the denoised image \( u \). Let \( G_\sigma \ast u \) be the degraded image filtered by the linear Gaussian filter, where \( \sigma \) represents the standard deviation of the Gaussian filter, \( p \) and \( q \) denote two local image patches, whose associated central pixel is \( x \), extracted from \( u \) and the degraded images respectively. The feature descriptor \( T(u) \) can be calculated by the following equation:

\[
T(u) = 1 - |c(p, q)s(p, q)|
\]

\[
= 1 - \frac{2\sigma_p \sigma_q + C_1}{\sigma_p^2 + \sigma_q^2 + C_1} \frac{\sigma_{pq} + C_2}{\sigma_p \sigma_q + C_2}
\]

\[
= 1 - \frac{2\sigma_{pq} + C_1}{\sigma_p^2 + \sigma_q^2 + C_1}
\]

where local statistics \( \mu_p, \sigma_p \) and \( \sigma_{pq} \) are defined as \( \mu_p = \sum_{i=1}^{N} p_i \), \( \sigma_p = \sqrt{\sum_{i=1}^{N} (p_i - \mu_p)^2} \) and \( \sigma_{pq} = \sum_{i=1}^{N} (p_i - \mu_p)(q_i - \mu_q) \). The constants \( C_1 \) and \( C_2 = C/2 \) are introduced for numerical stability and their effect on the final reconstruction will be discussed later. This descriptor contains two components: the contrast variation \( c(p, q) \) which calculates the reduction of contrast variation due to the degraded operation, and the correlation \( s(p, q) \) which quantifies the structural correction between the original and degraded images. For details of the descriptor, please refer to Liu et al. (2011).

Please note that the descriptor can be calculated either from magnitude image or from complex-valued image because the operations in equation (13) work on both cases. However, since the term ‘feature’ is more likely a visual attribute and usually defined based on the magnitude image, we chose to compute on magnitude data and then apply on complex data. The effect of phase variation on the performance of the descriptor is discussed in the discussion section. The value of each element in the feature descriptor image \( T \) is in the interval [0, 1]. As shown in figure 4, for each pixel, the more its value in \( T \) is close to 1, the higher probability it belongs to the structure part. This descriptor can be seen as a filter or mask for preserving the meaningful structure information and discarding the noise-like artifacts. Given knowledge of \( T \), we can have the simple feature-refining model as

\[
I_t = u + v_t = u + T \otimes v
\]

where \( I_t \) is the feature-refined denoised image, \( v_t \) is the detected feature image and the symbol \( \otimes \) represents the point-wise multiplication.

### 3.3. Tikhonov regularization

To satisfy the data consistency in the \( k \)-space, we have

\[
\min I \|F_p I - f\|^2 + \mu \|I - L\|^2
\]  

(15)
Obviously, this problem is a classical Tikhonov regularization problem with the feature-refined image $I_r$ as the prior image. We can update $I$ with an analytic solution, which satisfies the normal equation

$$(F_p^T F_p + \mu I) I = F_p^T f + \mu I.$$  

(16)

Under the periodic boundary condition for $I$, $F_p^T F_p$ is a block circulant matrix (Ravishankar and Bresler 2011), which can be diagonalized by the 2D discrete Fourier transform. Let $F$ denote the normalized full Fourier encoding matrix satisfying $F F^T = I_N$, we obtain

$$I = F^{-1} \begin{pmatrix} F[F_p^T f + \mu I] \\ F F_p F_p^T + \mu \end{pmatrix}$$  

(17)

where the matrix $F F_p^T F_p F$ is a diagonal matrix consisting of ones and zeros with the ones corresponding to the sampled locations in $k$-space. Equation (17) then becomes

$$\hat{I}(i) = \begin{cases} \hat{I}_r(i), & i \notin \Omega \\ f(i) + \mu \hat{I}_r(i) + \mu, & i \in \Omega \end{cases}$$  

(18)

where $\hat{I} = F \hat{f}$, $\hat{I}_r = F I_r$, and $\Omega$ stands for the subset of $k$-space that has been sampled.

As in Ravishankar and Bresler (2011), considering the noise standard deviation $\sigma$, the weighting parameter $\mu$ is also defined as $\mu = \gamma / \sigma$, where $\gamma$ is a constant that can be set empirically. In the noiseless case, the parameter $\mu \rightarrow \infty$.

In summary, the framework we proposed is shown in IFR-CS.

<table>
<thead>
<tr>
<th>Proposed Framework IFR-CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: while the stop criterion is not met do</td>
</tr>
<tr>
<td>2: Sparsity-promoting denoising problem equation (12a)</td>
</tr>
<tr>
<td>3: Feature refinement problem equation (12b)</td>
</tr>
<tr>
<td>4: Tikhonov regularization problem equation (12c)</td>
</tr>
<tr>
<td>5: end while</td>
</tr>
</tbody>
</table>

### 3.4. Comparison with the IRM method

Since the iterative regularization method (IRM) exploiting the Bregman distance is also able to recover some structure information from the residual image and has been successfully applied to CSMRI (Chang et al 2006, Liu et al 2009), here we compare the IRM-based CS method (IRM-CS) with our proposed framework. In IRM-CS, each iteration is updated by

$$\begin{cases} I^{k+1} = \arg \min_I \frac{1}{2} \| F_p I - f^k \|^2 + \lambda \| I \|_1 \\ f^{k+1} = f^k + F_p I^{k+1}. \end{cases}$$  

(19)

By using an auxiliary variable and the strategy of quadratic penalty, IRM-CS becomes
While the corresponding update of our proposed method IFR-CS is

$$
\begin{align*}
\mathbf{u}^{k+1} &= \arg \min_u \| \mathbf{I}^k - \mathbf{u} \|_2^2 + \lambda \| \mathbf{u} \|_1 \\
\mathbf{I}^{k+1} &= \arg \min_I \| F_p I - f^k \|_2^2 + \mu \| I - \mathbf{u}^{k+1} \| \\
f^{k+1} &= f^k + F_p \mathbf{I}^{k+1}.
\end{align*}
$$

Comparing equation (20) with equation (21), the significant difference is the refinement step in equations (20) and (21). First, IRM-CS emphasizes the refinement in $k$-space and our framework performs the refinement in image domain. Second, in the IRM-CS method, the iteration equation (20) only refines the $k$-space data that have already been acquired by adding back the data-inconsistency from previous iteration, which may contain not only structure information but also noise. In our proposed framework, the refinement step equation (21) picks out only the meaningful features from the residual image. Third, the refinement affects the Tikhonov model differently for two methods. Specifically, our method modifies the regularization term by changing the prior image to a feature-enhanced one and therefore preserves more details, while IRM-CS works on the data consistency term by changing the acquired data to a feature-and noise-enhanced one and thus converges to a noisy image.

4. Experimental results

4.1. Experimental setup

In this section, the performance of our proposed IFR-CS is evaluated on two simulated and four in vivo datasets. Specifically, TV semi-norm is used as the representative regularization term due to its popularity (IFR-TV). To simulate noiseless data acquisition, we undersampled the 2D discrete Fourier transform of the real MR images that were from in vivo MR scans of size $512 \times 512$ (many of which are courtesy of American Radiology Services and used in Ravishankar and Bresler (2011), as shown in figures 5(a) and (b)). Sampling schemes include the Cartesian sampling with random phase encodings (1D random) and pseudo radial sampling which are sampling radial on a Cartesian grid, as shown in figures 5(c) and (d). The first and second scanned data (Chang et al 2006, Ning et al 2013) were acquired from a 3T MAGNETOM Trio scanner using the T2-weighted turbo spin echo sequence (TR/TE $= 6100/99$ ms, $220 \times 220$ mm field of view, $3$ mm slice thickness). The third scanned data set was acquired on a 1.5 T Philips scanner using the steady-state free precession (SSFP) sequence with a flip angle of 50 degree and TR $= 3.45$ ms. The field of view (FOV) was $345 \times 270$ mm and the slice thickness was $10$ mm. Retrospective cardiac gating was used with a heart rate of 66 bpm. The fourth scanned data were acquired from a GE 3T scanner (GE Healthcare, Waukesha, WI) with 32-channel head coil and 3D T1-weighted spoiled gradient echo sequence, TE $= \text{minimum full}$, TR $= 7.5$ ms, FOV $= 24 \times 24$ cm, matrix $= 256 \times 256$, slice thickness $= 1.7$ mm. The adaptive combination method (Walsh et al 2000) is applied.

\[\text{www.quxiaobo.org/index_software.html}\]
to combine multi-channel data to a single-channel data corresponding to a complex-valued image. Our IFR-TV method\(^8\) was compared with the basic CSMRI using TV regularization (CS-TV)\(^9\) (Yang et al 2010), and the representative methods in two groups for preserving details: patch-based DLMRI\(^10\) (Ravishankar and Bresler 2011) and IRM-TV (Liu et al 2009). All algorithms were implemented in MATLAB on a PC equipped with Intel 2.6 GHz CPU and 4 G RAM. The quality of the reconstruction was quantified using the peak signal-to-noise ratio (PSNR) and high-frequency error norm (HFEN) (Ravishankar and Bresler 2011).

Specifically, the parameters for DLMRI method were set as: patch size $M^{1/2} = 6$, over-completeness of the dictionary $K = 1$, patch overlap $r = 1$, $\beta = 36$, $\lambda = 140$, learning stage (K-SVD) employing 10 iterations, and $T_0 = 5$. This ‘nominal’ parameter combination was suggested in Ravishankar and Bresler (2011) and it works well in our experiments. For our IFR-CS, we tried different parameter combinations for all experiments and select one robust combination which performs well for all experiments with $\lambda_{TV} = 0.005$, $\sigma_T = 50$ and $\gamma = 3$. The same strategy was conducted for CS-TV and IRM-TV. Therefore, for each method in our work, only one set of ‘nominal’ parameters were used for all experiments. The stopping criterion of our IFR-CS is that the relative difference between successive iterations for the reconstructed image should satisfy:

$$\frac{\|f^k - f^{k-1}\|_2}{\|f^k\|_2} < 5 \times 10^{-4}. \quad (22)$$

4.2. Visual details on noiseless case

Figure 6 shows an example on the brain image using CS-TV, and IFR-TV from the simulated $k$-space data with a 80% Cartesian undersampling mask. The reference image and sampling trajectory are shown in figures 5(a) and (d). The reconstructed images obtained by CS-TV and IFR-TV and their corresponding error magnitude images are shown in figures 6(c)-(f). We can see that our proposed framework can suppress undesirable artifacts and recover more features than the basic CS model. The PSNRs of the reconstructions are 35.67 dB and 40.26 dB for CS-TV and our proposed method, respectively.

To highlight the feature-refining property of the proposed framework, we used the feature descriptor to extract the feature part from the difference image between CS-TV and IFR-TV, as shown in figure 6(g), and also from the difference image between CS-TV and the original

---

8 Our demo can be provided based on request.
9 www.caam.rice.edu/optimization/L1/RecPF/RecPF_v1.1.zip
10 www.ifp.illinois.edu/ yoram/DLMRI-Lab/DLMRI.html
According to the high similarity between (g) and (h), it is evident that the feature loss of CS-TV in figure 6(h) can be recovered by the proposed IFR-TV very well. The result illustrates that our proposed method is superior to the basic CS method on feature recovery.

For a detailed description of the convergence of our proposed framework, the PSNR curve of the reconstructions versus iteration is depicted in figure 6(b). We can see that the PSNR curve quickly increases in the first 20 iterations and then changes slowly in following iterations. Figure 7 shows the intermediate images $I^k$, the denoised images $u^k$, the residual images $v^k$, the feature descriptors $T^k$ and the detected feature images $v^k = v^k \otimes T^k$ in the first, fifth and...
last iterations. It can be observed that in the first iteration, the zero-padding image $I^0$ exhibits many noise-like artifacts, most of which are filtered out in the residual image $v^0$ thanks to the sparsity-promoting denoising module. However, we can see that the denoised image $u^0$ still has some artifacts in the first iteration due to heavy undersampling. Therefore, the feature descriptor $T^0$ estimated from $u^0$ could recognize some artifacts as structure components by mistake. Even though, after being filtered by $T^0$, the detected feature image $v^0_t$ can filter out many noise-like artifacts and preserve the main structure information from the residual image $v^0$.

As the iteration continues, the feature descriptor $T^k$ contains more and more true structure gradually. Finally, with a precise feature descriptor, more structural information is recovered and a better reconstruction is obtained by our model than the basic CS model.

This phenomenon can be explained as follows: the denoised image $u^0$ contains less artifacts than the input $I^0$ due to denoising, the feature descriptor $T^0$ is estimated from $u^0$ but not $I^0$ that makes it extract more true features. Moreover, only partial artifacts in $u^0$ may be regarded as features and their values are usually smaller than those corresponding to true features. Considering that the extracted feature image $v^0_t$ is the point-wise product of the descriptor $T^0$ and the residual image $v^0$, it can be expected that the extracted feature image $v^0_t$ contains only few artifacts which is much less than in $I^0$. After being added back to the denoised image $u^0$, the output of the feature refinement step contains less artifacts than $I^0$. As iteration continues, the reconstruction $I^k$ exhibits better and better quality.

Figure 7. From left to right, the input image, the denoised image, the residual image, the feature descriptor and the detected feature $v_t = v \otimes T$ by IFR-TV at 1st (a), 5th (b) and the last iteration (c), respectively.
4.3. Different sampling schemes and ratios

To investigate the reconstruction performance under different sampling masks and ratios, we applied the combinations of two sampling schemes (Cartesian and radial) and five under-sampling ratios (0.43, 0.53, 0.60, 0.70 and 0.80), which correspond to acceleration factors of 1.79, 2.13, 2.5, 3.33 and 5.0, to obtain the partial Fourier measurements respectively. The tested images and the sampling masks are shown in figure 5. CS-TV, IRM-TV, DLMRI and our proposed IFR-TV were used to perform reconstruction. The corresponding PSNRs and HFENs of the reconstructed images are given in tables 1 and 2, respectively. First, the results of our proposed method IFR-TV are always better than those of CS-TV, which indicates the effectiveness of the scheme of iterative feature refinement. Second, IFR-TV provides larger PSNR values and smaller HFEN values than DLMRI for most specified undersampling ratios and sampling masks. Meanwhile, compared with IRM-TV in most cases, our IFR-TV method exhibits an advantage, which is significant at the case with high undersampling ratios. It is probably because the number of $k$-space locations where the data were acquired is small at the case of heavy undersampling and IRM-TV only refines the data at these locations. While our method can refine global structure information by using a specifically designed feature descriptor.

The PSNR values of the four methods versus iteration number are shown in figure 8. In this example, the numbers of iteration for CS-TV, IRM-TV, DLMRI and our IFR-TV are 10, 50, 25 and 33. The computational times for CS-TV, IRM-TV, DLMRI and our IFR-TV in this experiment are 1.2 s, 109.1 s, 1146.7 s and 12.9 s. For a fair comparison, all four methods were implemented in Matlab with no compiled MEX code or any CPU parallelization used for acceleration.

### Table 1. The reconstruction PSNR values using different undersampling ratios.

<table>
<thead>
<tr>
<th>Sampling</th>
<th>Image</th>
<th>Undersampling ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radial</td>
<td>Spine</td>
<td>43%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>51.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>51.82</strong></td>
</tr>
<tr>
<td>Brain</td>
<td></td>
<td>48.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>52.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>53.81</strong></td>
</tr>
<tr>
<td>Cartesian</td>
<td>Spine</td>
<td>42.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>52.53</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>45.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50.84</td>
</tr>
<tr>
<td>Brain</td>
<td></td>
<td>44.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>51.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>45.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>51.40</strong></td>
</tr>
</tbody>
</table>

**Note:** For each test setting, four results are provided: CS-TV (the first row), IRM-TV (the second row), DLMRI (the third row) and our proposed IFR-TV (the last row). The best among each result is highlighted in bold face.
Table 2. The reconstruction HFEN values using different undersampling ratios.

<table>
<thead>
<tr>
<th>Sampling Image</th>
<th>Undersampling ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>43%</td>
</tr>
<tr>
<td>Radial Spine</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td><strong>0.11</strong></td>
</tr>
<tr>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>0.13</td>
</tr>
<tr>
<td>Brain</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td><strong>0.05</strong></td>
</tr>
<tr>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
</tr>
<tr>
<td>Cartesian Spine</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td><strong>0.27</strong></td>
</tr>
<tr>
<td></td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>0.33</td>
</tr>
<tr>
<td>Brain</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td><strong>0.27</strong></td>
</tr>
</tbody>
</table>

Note. For each test setting, four results are provided: CS-TV (the first row), IRM-TV (the second row), DLMRI (the third row) and our proposed IFR-TV (the last row). The best among each result is highlighted in bold face.

Figure 8. The plot of PSNR values of reconstructions using four different methods versus iteration number on the figure 5(a) with 70% Cartesian undersampling.
4.4. Noise sensitivity

In this part, we explore the noise sensitivity of the proposed method when compared with the other three methods. The simulated noisy data can be obtained by adding the complex zero-mean Gaussian noise with different variances to the original $k$-space data. Figure 9 shows the plots of PSNR and HFEN values against the noise variance ($\sigma = 2, 5, 8, 10, 13, 15$) for CS-TV, IRM-TV, DLMRI, and IFR-TV on the image of figure 5(a) with 60% radial undersampling. We can see that our proposed IFR-TV exhibits the best performance for all noise levels, due to the capability of enhancing only structure information in our feature refinement scheme. The operation of averaging and capture of local features using adaptive dictionary make DLMRI also have a superior performance compared with the other two methods. IRM-TV performs better than the basic CS-TV only when the noise level is low, but its PSNR curve shows a sharp drop with increasing noise level. It is probably because the noise and structure information are both brought back and may contaminate the reconstruction in the heavy noise case.

4.5. Results on complex-valued data

In this part, we demonstrate the property of our proposed method on four sets of actual scanned MR data. Figure 10 shows an example of the reconstructions on the first scanned data from its radial undersampled measurements with undersampling ratio 60%. For the complex-valued MR image, the feature descriptor is estimated from the magnitude of the complex-valued denoised image. Then, this feature descriptor is applied to the real part and the imaginary part of the complex-valued residual image, respectively. Figures 10(b)–(d) show the reconstructed results obtained by IRM-TV, DLMRI and IFR-TV. Figure 10(e) illustrates the estimated feature descriptor. Figures 10(f)–(h) show the corresponding error images, where the error images of IRM-TV and DLMRI are seen to have more severe structure errors indicating loss of features. The improvement of IFR-TV over IRM-TV and DLMRI can be better visualized in figures 10(j)–(l), where a portion of the images are zoomed. It is seen that our proposed method can preserve more details.
Figure 11 shows the overall performance of IRM-TV, DLMRI and IFR-TV on the second complex-valued scanned MR data. The full $k$-space was retrospectively undersampled in Cartesian trajectory to simulate undersampling ratio of 37\%, 45\%, 53\%, 60\%, 70\% and 80\%. Figure 11(b) plots the curve of the PSNR values versus different undersampling ratios, which shows that our proposed method provides better reconstructions than the methods of IRM-TV and DLMRI on this dataset. Figures 11(c)–(e) show the reconstructions by these methods from the 53\% undersampled measurements. From the magnitudes of the reconstruction error shown in figures 11(f)–(h), we can see that IFR-TV leads to fewer errors and less structure loss than IRM-TV and DLMRI.

Figures 12 and 13 show the results from other two scanned datasets on cardiac imaging and sagittal brain imaging, respectively. From the comparison on the zoom-in images, our proposed method performs better than IRM-TV and DLMRI. Specifically, the ROIs labeled by red arrows indicate the superior capability of our proposed method in recovering small features to other methods. It is because the small features usually smoothed out in the sparsity-promoting denoising operation are gotten back from the residual image through the feature refinement step proposed in our framework.
5. Discussion

5.1. Parameter evaluation

In this part, we evaluated the sensitivity of IFR-TV to parameter settings by varying one parameter at a time while keeping the rest fixed at their nominal values. The main parameters evaluated were the regularized parameter $\lambda_{TV}$ in the sparsity-promoting denoising module, the blurring parameter $\sigma_t$ and $C_1$ in estimating the feature descriptor, and the weighting parameter $\gamma$ in the noise case.

![Figure 11](image-url)

Figure 11. (a) Reference image. (b) The PSNR values of IRM-TV, DLMRI and IFR-TV at different sampling ratios with the Cartesian trajectory. (c)–(e) The images reconstructed by IRM-TV, DLMRI and our proposed IFR-TV from its Cartesian undersampled measurements with undersampling ratio 53%, respectively. (f)–(h) The magnitudes of reconstruction error with respect to (c)–(e).
Figures 14(a) and (b) show the influence of the regularization parameter $\lambda_{TV}$ on the reference images in figures 5(a) and (b) with Cartesian and radial sampling, respectively. Generally speaking, small $\lambda_{TV}$ gives better reconstructed result at the case of low undersampling ratio, while large $\lambda_{TV}$ works well at heavy undersampling cases. In this work, we chose $\lambda_{TV}$ to be 0.005.

The influence of the blurring parameter $\sigma_t$ in the step of estimating the feature descriptor is shown in figure 14(c) on the reference image figure 5(b) with 80% radial undersampling. As $\sigma_t$ changes from 5 to 75, the value of PSNR increases. This improvement can be explained by figure 15 which shows different feature descriptors estimated using different $\sigma_t$ values. The red rectangle regions show that the feature descriptor estimated with larger values like $\sigma_t = 50$ can extract more meaningful structure information than that with $\sigma_t = 10$ and $\sigma_t = 5$. However, with the feature refinement scheme, our reconstruction result at the case of $\sigma_t = 5$ is still better than that obtained by the basic CSMRI method shown in table 1. The values of PSNR versus the weighting parameter $\gamma$ are plotted in figure 14(d) for two cases of $\sigma = 0.1$ and $\sigma = 20$. Finally, the values of PSNR versus the parameter $C_1$ in equation (13) are plotted in figures 14(e) and (f). In our experiments, we empirically set $\sigma_t = 50$ and $C_1 = 2e^{-4}$. 

Figure 12. (a) Reference image. (b) The estimated feature descriptor. (c) The reference zoom-in image. (d)-(f) The zoom-in images reconstructed by IRM-TV, DLMRI and IFR-TV from its Cartesian undersampled measurements with undersampling ratio 60%, respectively.
5.2. Feature descriptor

The feature descriptor plays a very important role in our proposed IFR-CS framework. The feature descriptor used in this paper is regarded as a map or mask which defines whether the pixel is located on the meaningful structure part from the respective of vision information. For MRI, we can also estimate the descriptor from the complex image by following equation (13). Figure 16 shows the comparison on estimating the feature descriptor from magnitude and complex-valued images. From this figure, we can see that the descriptor estimated from the magnitude image has a smaller error on the non-structure positions and a larger value on the structure part than the descriptor estimated from the complex-valued image. Therefore, the feature descriptor is estimated according to the magnitude difference between the original image and its blurred image in our work, and the variations in phase do not affect the performance of the descriptor.

5.3. Extension and limitation

In the current work, we only apply our proposed framework on the single-channel data acquired on a grid (radial sampling is the pseudo-radial trajectory in experiments) with TV semi-norm as the regularization term. Realizing that MR data are now almost exclusively collected using phased array receive coils with a number of channels, in the future we will extend our proposed IFR to parallel imaging for more practical applications. In that case, we
Figure 14. Parameter evaluation. PSNR versus regularization parameter $\lambda_{TV}$: (a) The results of figure 5(a) under the Cartesian sampling trajectory with different undersampling ratios and (b) the results of figure 5(b) under the radial sampling trajectory with different undersampling ratios. (c) PSNR versus blurring parameter $\sigma_t$ and (d) PSNR versus weighting parameter $\gamma$ from figure 5(b) with 80% radial undersampled measurements. (e)/(f) PSNR versus parameter $C_1$ in the equation (13) under Cartesian and radial sampling trajectories respectively.

Figure 15. (a), (b), (c) The feature descriptor estimated from the reference image figure 5(b) with blurring parameter $\sigma_t = 5, 10, 50$, respectively.
need to consider how to incorporate the coil sensitivity information and how to handle the additional computational load that might be incurred due to the more complex multi-channel data. IFR-CS can also be applied to more general MRI scenarios by assigning the variables in original optimization problem (P1) new meanings. For example, we can change the regularization term by using the sparse prior in the wavelet domain. The basic CS-Wavelet can be formulated in the following form:

\[
\min_{f} \frac{1}{2} \|F_{i} I - f\|_{2}^{2} + \lambda_{2} \|\Phi f\|_{1}
\]  

(23)

where \(\Phi\) is the wavelet transform. Similarly, the IFR-Wavelet method can be derived by introducing the scheme of iterative feature refinement:

\[
\begin{align*}
\lambda^{k+1} &= \arg\min_{\lambda} \|f^{k} - \lambda I\|_{2}^{2} + \lambda_{w} \|\Phi f\|_{1} \\
\lambda^{k+1} &= \lambda^{k+1} + \nu^{k+1} \\
I^{k+1} &= \min_{I} \|F_{i} I - f^{k}\|_{2}^{2} + \mu \|I - I^{k+1}\|_{2}^{2}
\end{align*}
\]  

(24)

Specifically, the iterative soft-thresholding method (Figueiredo and Nowak 2003) was used to solve problem (24). Table 3 shows the PSNRs of the reconstruction results for two real-valued images and one complex-valued image with radial sampling trajectory at different undersampling ratios.

It can be observed that with the feature refinement scheme, IFR-Wavelet improves the reconstruction quality at all undersampling ratios. Figure 17 shows one example on the...
complex-valued image (figure 10(a)) with 60% radial undersampling. The magnitude image of the reconstruction error shows lower errors in structure for IFR-Wavelet (right column) than for CS-Wavelet (left column), which demonstrates the effectiveness of our proposed feature-refinement framework.

We can also replace the fixed transform (finite-difference/wavelet) to adaptive dictionary or using finite-difference and wavelet simultaneously. If the encoding matrix is not just the Fourier encoding but the sensitivity encoding (comprising Fourier encoding and sensitivity weighting) or non-Cartesian encoding (comprising Fourier encoding and gridding), the proposed framework can be applied to parallel imaging/ non-Cartesian MRI with the problem formulation:

$$\min_i \frac{1}{2} ||E f - I||_2^2 + \lambda ||I||_{L1}. \tag{25}$$

The detailed description is beyond the scope of this paper due to limited space. Since the feature descriptor plays a very important role in our framework, it can be expected that employing a novel method or designing a more powerful descriptor for identifying true structure information from the residual image can further improve the performance of IFR-CS.

There are two limitations in our current framework. First, for images with simple or low contrast structures, the estimated feature descriptor may fail to extract sufficient structure information from the residual image, and thus only present tiny improvement. Second, as the performance of basic CSMRI with fixed transform is limited in the case of heavy undersampling, our method usually does not exhibit an advantage to DLMRI with more than 80% undersampling in 2D static MR imaging, and this can be fixed by replacing the fixed transform to the adaptive dictionary as mentioned before.

6. Conclusion

In this work, a novel feature refinement module has been introduced for accurate MR image reconstruction from undersampled $K$-space data. The proposed module is compatible with the CSMRI framework equipped with any sparsifying transforms. We name the integration of our proposed module with CSMRI as IFR-CS. To evaluate the performance of IFR-CS, we have tested it on two simulated MR images and four in vivo MR datasets. Experimental results have shown that the proposed descriptor has a strong capability to restore meaningful structures and details and IFR-CS is comparable and even superior to other three state-of-the-art MR image reconstruction methods.
Acknowledgment

We gratefully acknowledge the grant support: China NSFC 61471350, 11301508, 61401449, the Natural Science Foundation of Guangdong 2015A020214019, 2015A030310314, 2015A030313740 the Basic Research Program of Shenzhen JCYJ20140610152828678, JCYJ20140610151856736, JCYJ20150630114942318 and in part by the SIAT Innovation Program for Excellent Young Researchers 201403 and 201313. The authors would like to express their gratitude to the reviewers for their positive comments and valuable advice on this paper.

References


American Radiology Services www3.americanradiology.com/pls/web1/wwwmain.home


Chambolle A 2004 An algorithm for total variation minimization and applications J. Math. Imaging Vis. 20 89–97

Chang T C, He L and Fang T 2006 MR image reconstruction from sparse radial samples using bregman iteration Proc. 13th Annual Meeting ISMRM (Seattle) p 696


Goldstein T and Osher S 2009 The split bregman method for $\ell_1$-regularized problems SIAM J. Imaging Sci. 2 323–43


Otazo R, Kim D, Axel L and Sodickson D K 2010 Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion mri Magn. Reson. Med. 64 767–76
Tao M W, Johnson M K and Paris S 2010 Error-tolerant image compositing European Conf. on Computer Vision
Yin W, Osher S, Goldfarb D and Darbon J 2008 Bregman iterative algorithms for \(\ell_1\)-minimization with applications to compressed sensing SIAM J. Imaging Sci. 1 143–68