Inverse Problems and Imaging
Volume 4, No. 2, 2010, 223–240
doi:10.3934/ipi.2010.4.223

# A NOVEL METHOD AND FAST ALGORITHM FOR MR IMAGE RECONSTRUCTION WITH SIGNIFICANTLY UNDER-SAMPLED DATA

Yunmei Chen and Xiaojing Ye

Department of Mathematics, University of Florida, Gainesville, FL 32611, USA
358 Little Hall, Gainesville, FL 32611, USA

Feng Huang

Advanced Concept Development, Invivo Corporation
3545 SW 47th Avenue, Gainesville, FL 32608, USA

Abstract. The aim of this work is to improve the accuracy, robustness and efficiency of the compressed sensing reconstruction technique in magnetic resonance imaging. We propose a novel variational model that enforces the sparsity of the underlying image in terms of its spatial finite differences and representation with respect to a dictionary. The dictionary is trained using prior information to improve accuracy in reconstruction. In the meantime the proposed model enforces the consistency of the underlying image with acquired data by using the maximum likelihood estimator of the reconstruction error in partial $k$-space to improve the robustness to parameter selection. Moreover, a simple and fast numerical scheme is provided to solve this model. The experimental results on both synthetic and in vivo data indicate the improvement of the proposed model in preservation of fine structures, flexibility of parameter decision, and reduction of computational cost.

1. **Introduction.** Magnetic resonance (MR) imaging is a technique that allows visualization of structures and functions of a body by non-invasive and non-ionizing means. It provides better contrast between the different soft tissues than most other modalities. However, MR imaging takes much longer acquisition time than many other imaging modalities, which limits its applications. To reduce acquisition time, the most common and feasible approach is by acquiring only partial $k$-space data, followed by adequate reconstruction techniques to obtain images with well-preserved quality.

The idea of reconstructing images from partial data coincides with the compressed sensing (CS), a technique used in signal/image processing. CS can accurately recover a signal/image using data with significantly less measurements than regular, provided the sparsity of the underlying signal/image and a sophisticated reconstruction procedure. Recently, the application of this technique in medical imaging has become a hot research topic, and shown promising results [6, 19, 8, 26, 4, 5, 9, 16, 11, 18]. In particular, the redundancy of the MR data acquired in the frequency domain, i.e. the $k$-space, and implicit sparsity in MR images have motivated many researchers to study the application of CS to fast MR imaging (CS-MRI).

2000 *Mathematics Subject Classification.* Primary: 62H35; Secondary: 65K10.
*Key words and phrases.* compressed sensing, dictionary, sparse representation, convex optimization.

CS-MRI has the advantage of producing high quality reconstruction of MR images from partial Fourier data. Recent study has shown that the key to the success of CS-MRI is a combination of the sparsity of the underlying image under an appropriate domain, the $k$-space trajectory that provides incoherent undersampling artifacts, and an adequate nonlinear reconstruction method that enforces both the sparsity and data consistency of the underlying image [19, 10, 15]. A great progress of researches on CS-MRI has been made. However, for clinical applications, radiologists often demand improvements on accuracy, robustness, and efficiency of the current CS-MRI algorithms. The desired improvements include the ability of removing artifacts while preserving important diagnostic information (in particular, sharp edges and fine structures), the robustness to the choice of parameters, and the speed of reconstructions.

In this paper, we propose a novel variational model and a fast numerical algorithm for MR image reconstruction with highly undersampled data, which tackles the three problems mentioned above as follows.

- *Accuracy*
  The proposed model enforces the sparsity of the underlying image in terms of its spatial finite differences and representation by a dictionary trained using prior information. Thus, improvement on accuracy of reconstruction can be achieved.
- *Robustness*
  The proposed model enhances the data consistency by the approach of maximum likelihood estimation for the discrepancy between the reconstruction and acquired data in $k$-space. This leads to an automatically optimized weighting parameter which makes the parameter selection more flexible.
- *Efficiency*
  To make the proposed model clinically applicable, we also provide a simple and fast numerical algorithm to solve the model. The main computations involve only shrinkage, matrix-vector multiplication and fast Fourier transform (FFT).

The background and brief introduction of our contributions to these issues are provided in the following three subsections.

1.1. **Trained dictionaries as sparsifying transforms.** Since sparsity is the key to the success of CS and consequent reconstructions, many researches exploited the transforms under which images have their sparse representations [19]. The theory of CS indicates that once such transforms were found, an image can be accurately recovered using a set of random measurements with cardinality much less than the original resolution of the image [6, 10].

In recent years, finite difference operator and wavelet transforms have been widely used as such sparsifying transforms for MR images [13, 19]. In [13], the authors proposed a total variation (TV) based model to reconstruct MR images from partially acquired $k$-space data. Their model works well for piecewise constant or very homogeneous images [21]. For images with inhomogeneous intensity, TV based models may not work well when the undersampling rate is high. In [19], Lustig *et al.* proposed a model that minimizes the Besov together with TV norms of the underlying image, subjected to a data consistency constraint:

$$\min_u TV(u) + \mu\|\Psi^\top u\|_1, \quad \text{s.t. } \|\mathcal{F}_p u - f_p\|_2 < \sigma, \tag{1}$$

where $\|\cdot\|_1$ is the $\ell_1$ norm, $TV(u) \triangleq \|Du\|_1$ is the (anisotropic) TV semi-norm of $u$, $\Psi$ is the Haar wavelet transform, the superscript $^\top$ denotes (conjugate) transpose of matrices. In the constraint, $\mathcal{F}_p$ denotes the undersampled Fourier operator corresponding to the customized $k$-space sampling pattern, $f_p$ is the partially acquired $k$-space data, and $\sigma$ is an estimate of acquisition error. As proved in [6, 10], minimizing $\ell_1$ norm subjected to data consistency yields sparse solutions. Therefore, model (1) in fact leads to a reconstructed image that has sparse gradient and wavelet transform coefficients. Simulations in [19] showed very promising results using model (1). However, image quality degrading and the loss of diagnostic information may happen in reconstructions using TV and wavelet transforms as they may eliminate some fine structures and/or useful local information in the recovered images. As an alternate, we propose to use the dictionaries trained using prior information as sparsifying transforms to tackle this problem.

A recent boost of the study on dictionary design shows the great potential of using dictionaries in signal/image processing. Dictionary is usually formed as a set of overcomplete bases and its elements/atoms have much smaller sizes than the image size. On the contrary, wavelet transform has a set of complete bases with elements of the same size as image itself, and therefore can be treated as a special case of dictionary. Furthermore, a dictionary can be properly trained such that its prototype signal-atoms are more adequate to sparsely represent objective signals than wavelet. A number of researches have shown the benefits of using dictionaries for sparse representation; see, e.g. [22, 18].

In this work, we train a dictionary by applying $K$-SVD algorithm to a database consisting of patches extracted from images acquired from the same sequence but perhaps different subjects. Then the trained dictionary $A$, as shown in Fig. 1, is used to reconstruct other MR images under the same acquisition sequence with similar structures. Details on the training and reconstruction processes are provided in the following sections. Comparison of the accuracy of sparse representations using wavelet transform and the trained dictionary $A$ is shown in Fig. 2. Results by using dictionary has has better preserved edges and fine structures because dictionary absorbed prior knowledge by learning features of this type of images during the training process [1]. Moreover, the dictionary training and representation processes are stable and robust as shown in our experimental results in section 4.

1.2. **Likelihood estimate as data fidelity measure.** To improve the robustness to the parameter selection, we use the likelihood estimation of the reconstruction error as the data fidelity measure. The reconstruction error is the difference between the partially acquired data and the Fourier transform of the reconstruction at sampled $k$-space locations, i.e. $\mathcal{F}_p u - f_p$ in (1). In previously proposed CS-MRI algorithms, least squares, i.e. the sum of squared difference (SSD) $\|\mathcal{F}_p u - f_p\|_2^2$, is a commonly used data consistency measure. For instance, the unconstrained version of model (1) solves for the reconstruction by

$$\min_u TV(u) + \mu\|\Psi^\top u\|_1 + \frac{\lambda}{2}\|\mathcal{F}_p u - f_p\|_2^2, \tag{2}$$

where the parameter $\lambda$ is crucial to the reconstruction results: an improperly large weight for the data fidelity term results in serious residual artifacts, whereas an improperly small weight results in damaged edges and/or fine structures. In this work, we tackle this problem by treating the reconstruction errors at all pixels as samples independently drawn from a Gaussian distribution with mean zero and

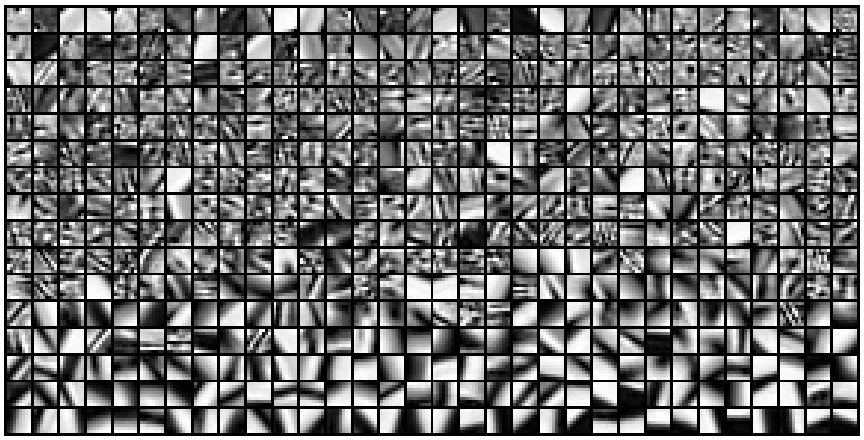


FIGURE 1. Dictionary trained by $K$-SVD algorithm. The database used for training consists of 4096 patches extracted from four MR brain images (but excludes the image to be reconstructed). Each block represents an atom of size $8 \times 8$. Atoms are sorted by ascending the variances of intensities.

variance $\sigma^2$ to be optimized. By maximum likelihood estimate (MLE) approach, the weight on $\|\mathcal{F}_p u - f_p\|_2^2/2$ becomes $\lambda/\sigma^2$ rather than a prescribed $\lambda$. Since $\sigma$ is to be optimized, it is updated during iterations (in fact, it is the standard deviation of the reconstruction error, see (17) below). When the reconstruction error reduces, the weight $\lambda/\sigma^2$ on $\|\mathcal{F}_p u - f_p\|_2^2$ increases, and hence the accuracy is improved. This automatically optimized weighting feature makes the choice of $\lambda$ much more flexible.

1.3. **Fast numerical algorithms for solving CS-MRI models.** Despite that dictionaries are more adequate in signal/image reconstructions, the computational cost is higher than that using wavelet transform due to the redundancy of dictionaries and overlapping of patches to be represented. Also, the non-differentiability of TV and $\ell_1$ norms brings difficulties to fast solutions of CS-MRI models. There have been many numerical algorithms developed to solve TV and $\ell_1$ regularized minimization problems, more recent developments can be found in [7, 27, 28, 23, 25, 13] and [17, 20, 14, 3, 2, 24] and references therein. Our approach in this work is closely related to the algorithm developed in [25], in which Yang *et al.* introduced a simple and fast method to solve model (2) with isotropic TV norm of $u$ defined by

$$TV(u) \triangleq \sum_{i=1}^{N} \|D_i u\|_2. \tag{3}$$

In (3) $u \in \mathbb{R}^N$ is the vector formed by stacking all columns of the image vertically, $N$ is the total number of pixels in the image, and $D_i \in \mathbb{R}^{2\times N}$ represents the gradient operator at the $i$-th element in the vector of $u$. To overcome the non-differentiability of TV and $\ell_1$ norms, they introduced auxiliary variables and used a classical quadratic penalty method which yields an alternating minimization scheme. By diagonalizing the gradient operator using Fourier transform, they made the main computation of the algorithm involving only soft shrinkage and fast

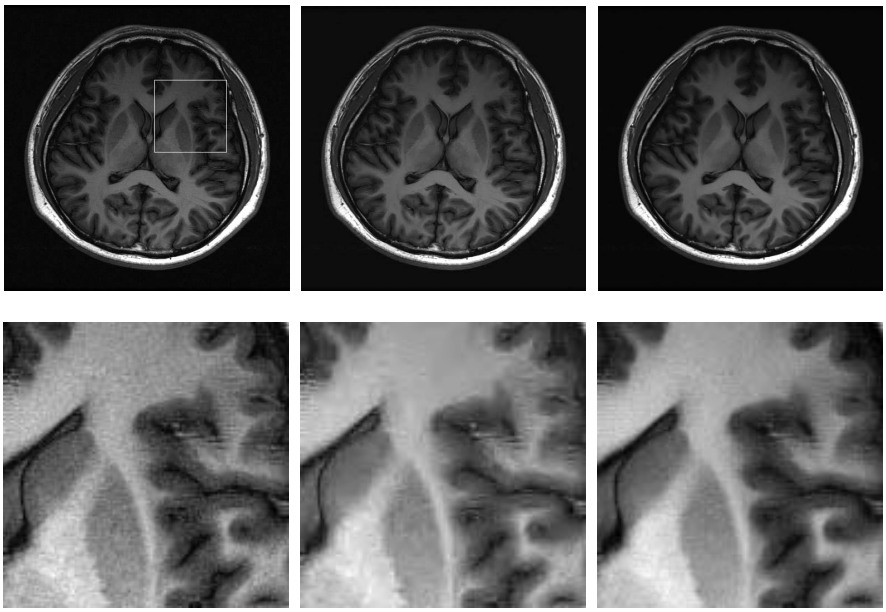


Figure 2. Compare the accuracy of sparse representations by wavelet and trained dictionary. In both cases images are represented by picking up the largest 12.5% transform coefficients. Bottom images are corresponding zoomed-in square area shown on the top left image. Left column: reference. Middle column: representation by wavelet, with RMSE 5.74% and SNR 22.3. Right column: representation by trained dictionary $A$ shown in Fig. 1, with RMSE 4.36% and SNR 24.7.

Fourier transform. However, their algorithm cannot be directly applied to the models using dictionary $A$ since it requires the orthogonality of $\Psi$ in (2). Therefore the development of efficient algorithms involving the use of a dictionary is still a remaining problem. In this paper, we show that a simple trick on selecting patches from the underlying image can be used to overcome this difficulty. Based on the method in [25] and this trick, we provide a simple and fast numerical algorithm that can be applied to reconstruction models involving dictionaries.

1.4. **Organization.** The rest of this paper is organized as follows. A detailed description of the proposed model is given in 2. In section 3, a fast algorithm to solve the proposed model and its derivation are provided. Experimental results on synthetic and in vivo data are presented in section 4. The last section concludes the paper.

2. **Proposed model.** Before going into details of the proposed model, we address the notations used throughout the rest of the paper. First of all, all vectors in this paper are column vectors. Let $u \in \mathbb{R}^N$ be the underlying reconstruction as in (3), and $\mathcal{F}$ be the discrete Fourier transform, which can be treated as an $N \times N$ unitary matrix. Let $P \in \mathbb{R}^{p\times N}$ denote the binary matrix that selects certain rows of $\mathcal{F}$ according to the $k$-space sampling pattern. Then $\mathcal{F}_p \triangleq P\mathcal{F}$ is the undersampled

Fourier transform. Let $f_p \in \mathbb{C}^p$ be the partially acquired $k$-space data and use $\|\cdot\|$ to denote Euclidean norm $\|\cdot\|_2$ of vectors henceforth. The notation $(\cdot;\cdot)$ represents a matrix formed by stacking its arguments vertically.

In this paper, all patches have size $\sqrt{n} \times \sqrt{n}$ and are often to be treated as $n$-vectors unless otherwise noted (in our experiments $n = 64$). The dictionary $A \in \mathbb{R}^{n\times K}$ consists of $K$ $n$-vectors as atoms. Binary matrix $R_j \in \mathbb{R}^{n\times N}$ extracts the $j$-th patch of $u$, and forms the patch $R_j u$ as an $n$-vector. All patches $\{R_j u\}_{j=1}^J$ cover the entire image $u$ and may overlap.

2.1. **Sparse representation using trained dictionary.** To improve the accuracy of reconstruction, especially the ability of preserving diagnostic information, we here propose to use a trained dictionary instead of wavelet as the sparsifying transform in MR image reconstructions. We chose the recently proposed dictionary design method, termed as $K$-SVD algorithm, to perform the dictionary training. $K$-SVD is an iterative method that alternates between sparse coding of the examples based on the current dictionary and a process of updating the dictionary atoms to better fit the given database. The output is a trained dictionary that can represent all signals in the database under strict sparsity constraints and error tolerance. Interested readers are referred to [1, 12] for details.

Our procedure of forming a database and applying $K$-SVD algorithm to train an adaptive dictionary for MR image reconstruction is depicted as follows.

1. Collect a number of MR images acquired using the same sequence as that for the image to be reconstructed, but from different subjects. The training images and the image to be reconstructed are preferred to be the same body parts to get a better sparse representation. Using the same acquisition sequence ensures that they have similar contrast.
2. Decompose the training images to $\sqrt{n}\times\sqrt{n}$ patches, and discard those patches with constant intensity. Then randomly choose $8K$ patches from the remaining patches, where $K$ is the number of atoms in the dictionary to be trained.
3. Train a dictionary $A$ by applying $K$-SVD algorithm to that $8K$ patches with the overcomplete DCT matrix as the initial. The resulting trained dictionary has $K$ elements, i.e. $A \in \mathbb{R}^{n\times K}$.

In our experiments, we set $n$ to 64 and $K$ to 512/256 for brain/chest MR data. The dictionary trained for brain image reconstruction is illustrated in Fig. 1. The dictionary $A$ we obtain from this training procedure can adequately represent any patches of brain MR images (e.g. from different subjects or the same subject in different periods) acquired under the same sequence. In particular, each patch $R_j u$ of $u$ can be sparsely represented by $A$. Namely, there exist representation coefficients $\alpha_j \in \mathbb{R}^K$ such that

$$\|\alpha_j\|_0 < n < K \quad \text{s.t.} \quad A\alpha_j \approx R_j u, \quad j = 1, \cdots, J,$$

where $\|\cdot\|_0$ counts the number of nonzero components of its argument. Therefore, the sparsity of $u$ under the representation of $A$ can be used as a regularization in the reconstruction. That is, we enforce the following into our model

$$\min_{\alpha} \sum_{j=1}^{J} \left( \|\alpha_j\|_1 + \frac{\nu}{2}\|A\alpha_j - R_j u\|^2 \right), \tag{4}$$

where $\alpha = (\alpha_1; \cdots; \alpha_J) \in \mathbb{R}^{KJ}$. This is in fact the relaxed form of sparse-land problem with $\ell_0$-norm substituted by $\ell_1$-norm. The reason why we use $\ell_1$ instead of

$\ell_0$ is that minimizing the non-convex $\ell_0$ is generally a NP-hard problem and hence is not tractable in practice. Moreover, it has been proved that minimization problems with $\ell_1$ and $\ell_0$ share the same solution under certain conditions [6, 10].

Note that if $J = 1$, $R_1 = I$ (the identity matrix) and $A = \Psi$ (the wavelet transform), then (4) reduces to $\|\Psi^\top u\|_1$ as in (2) when the difference in the quadratic is exactly 0. Namely, wavelet is a special case of dictionary.

2.2. **Likelihood estimate for the data consistency.** One difficulty of applying the unconstrained energy minimization problem (2) for MR image reconstruction is in determining the weighting parameter that balances the data consistency and image sparsity. The reconstruction results are sensitive to the choice of this parameter. To tackle this problem, we derive the data consistency measure, the so-called data fidelity, from maximum likelihood estimate (MLE) approach.

Let $\omega = (\omega_1, \cdots, \omega_p)^\top \in \mathbb{C}^p$ be the reconstruction error in $k$-space, which is the difference between the Fourier transform of the reconstruction $u$ and the acquired data $f_p$ at the sampled $k$-space locations:

$$f_p = \mathcal{F}_p u + \omega.$$

Consider $\{\omega_l\}_{l=1}^p$ as independent random samples drawn from a normal distribution of mean zero and variance $\sigma^2$ to be determined. Therefore, the joint probability density function (pdf) of $\{\omega_l\}_{l=1}^p$, which is also the likelihood of $\sigma$, becomes

$$\mathcal{L}(\sigma|\omega) = \prod_{l=1}^{p} \left( \frac{1}{\sqrt{2\pi}\sigma} e^{-\omega_l^2/2\sigma^2} \right) = \left(2\pi\sigma^2\right)^{-p/2} e^{-\|\omega\|^2/2\sigma^2}.$$

Thus, the negative log-likelihood is

$$-\log \mathcal{L}(\sigma|\omega) = \|\omega\|^2/2\sigma^2 + p\log\sqrt{2\pi}\sigma. \tag{5}$$

Substituting $\omega$ by $\mathcal{F}_p u - f_p$, and omitting the constant $p\log\sqrt{2\pi}$, we obtain a MLE based consistency estimation with the partially acquired data:

$$F(u, \sigma, f_p) = \|\mathcal{F}_p u - f_p\|^2/2\sigma^2 + p\log\sigma. \tag{6}$$

This is a generalization of the least square estimation, which is just the case where $\sigma \equiv 1$. We will use (6) as data fidelity term in our energy functional.

2.3. **Variational model for MR image reconstruction from undersampled data.** Now we are ready to present our model. We propose to use TV and sparse representation by trained dictionary as regularization and MLE (6) as data consistency measure. Our model is formulated as an unconstrained minimization problem

$$\min_{u,\alpha,\sigma} TV(u) + \mu \sum_{j=1}^{J} \left( \|\alpha_j\|_1 + \frac{\nu}{2}\|A\alpha_j - R_j u\|^2 \right) + \lambda F(u, \sigma, f_p), \tag{7}$$

where $TV(u)$ is the TV norm of $u$ defined as in (3), the summation over $\alpha$ is the regularization of $u$ using the sparsity under representation by dictionary $A$, and $F(u, \sigma, f_p)$ is MLE data consistency measure (6). By using MLE based approach, $\sigma$ is also optimized along with $u$. In (7) the weight on $\|\mathcal{F}_p u - f_p\|^2$ versus the sparsity of the underlying image is $\lambda/\sigma^2$ rather than $\lambda$ only. In the Euler-Lagrange (EL) equations associated with the proposed energy function below, one can see that $\sigma$ is the standard deviation of the reconstruction error $\omega$. Hence, when the construction error $\omega$ decreases, the weight $\lambda/\sigma^2$ on minimizing $\ell_2$ norm of $\omega$ increases automatically, which makes the choice of the initial weighting parameter $\lambda$ more

flexible. This flexibility dramatically reduces the difficulty of parameter decision and improves the applicability of CS. Moreover, our experimental results below show that this automatically updated weighting makes faster convergence, and better accuracy in reconstruction.

3. **Algorithm.** There minimization problem (7) is closely related to the well-known TV and $\ell_1$ based signal/image reconstruction problems. Since the non-differentiability of TV and $\ell_1$ terms bring a difficulty in computations, there have been a number of numerical algorithms developed to efficiently solve this type of problems. The algorithm provided in this section is inspired by the work in [25, 23], which uses the variable splitting and classical quadratic penalty technique in optimization to make the computation fast and stable.

3.1. **A fast algorithm for solving the proposed model.** We first introduce two auxiliary variables $w = (w_1^\top; w_2^\top; \cdots; w_N^\top) \in \mathbb{R}^{N\times 2}$ and $\beta = (\beta_1; \beta_2; \cdots; \beta_J) \in \mathbb{R}^{KJ}$ where $w_i \in \mathbb{R}^2$ and $\beta_j \in \mathbb{R}^K$ for all $i = 1, \cdots, N$ and $j = 1, \cdots, J$. Then we consider the minimization problem equivalent to (7):

$$\begin{aligned} \min_{u,w,\alpha,\beta,\sigma} \quad & \sum_{i=1}^{N} \|w_i\| + \mu \sum_{j=1}^{J} \left( \|\beta_j\|_1 + \frac{\nu}{2} \|A\alpha_j - R_j u\|^2 \right) + \lambda F(u, \sigma, f_p) \\ \text{s.t.} \quad & w_i = D_i u, \ \beta_j = \alpha_j, \ \forall\, i = 1, \cdots, N, j = 1, \cdots, J. \end{aligned} \tag{8}$$

Relaxing the equality constraint and penalizing their violations by quadratic functions, we obtain an unconstrained version of (8):

$$\min_{u,w,\alpha,\beta,\sigma} \sum_{i=1}^{N} \phi(w_i, D_i u) + \mu\psi(\beta, \alpha) + \sum_{j=1}^{J} \frac{\mu\nu}{2} \|A\alpha_j - R_j u\|^2 + \lambda F(u, \sigma, f_p) \tag{9}$$

where functions $\phi$ and $\psi$ are defined as

$$\phi(s,t) = \|s\| + \frac{\eta}{2}\|s - t\|^2, \quad s, t \in \mathbb{R}^2$$

and

$$\psi(s,t) = \|s\|_1 + \frac{\theta}{2}\|s - t\|^2, \quad s, t \in \mathbb{R}^{KJ}$$

for given $\eta, \theta > 0$. With $\eta, \theta$ gradually increasing, solving (9) lead to approximations to the solution of (8).

The minimization (9) can be carried out in a much faster and more stable manner than (7): first, for fixed $u$ and $\alpha$, the minimization with respect to $w$ and $\beta$ can be carried out in parallel:

$$w_i = \mathcal{S}_2(D_i u), \quad i = 1, \cdots, N, \tag{10}$$

where $\mathcal{S}_2(t)$ is the two-dimensional (2D) shrinkage that minimizes $\phi(s,t)$ for fixed $t$:

$$\mathcal{S}_2(t) \triangleq \max\left\{ \|t\| - \frac{1}{\eta}, 0 \right\} \cdot \frac{t}{\|t\|}, \quad t \in \mathbb{R}^2.$$

Moreover, we have

$$\beta = \mathcal{S}_c(\alpha) \tag{11}$$

where $\mathcal{S}_c(t)$ is the componentwise shrinkage that minimizes $\psi(s,t)$ for fixed $t = (t_1, \cdots, t_{KJ})^\top \in \mathbb{R}^{KJ}$:

$$\mathcal{S}_c(t) = (\mathcal{S}(t_1), \cdots, \mathcal{S}(t_{KJ}))^\top$$

and

$$\mathcal{S}(x) = \max\{x - 1/\theta, 0\} \cdot \operatorname{sign}(x), \quad x \in \mathbb{R},$$

with assumption $0 \cdot (0/0) = 0$. Both computational costs for $w$ and $\beta$ are linear in $N$.

Secondly, for fixed $u$ and $\beta$, we can have $\alpha = (\alpha_1; \cdots; \alpha_J)$ by solving the following minimization problem:

$$\min_{\alpha} \sum_{j=1}^{J} \left( \theta \|\alpha_j - \beta_j\|^2 + \nu \|A\alpha_j - R_j u\|^2 \right). \tag{12}$$

The solution can be obtained by setting $\alpha_j$ as

$$\alpha_j = V(\theta I + \nu\Lambda)^{-1} V^\top \left( \theta\beta_j + \nu A^\top R_j u \right) \tag{13}$$

where the diagonal matrix $\Lambda$ and orthogonal matrix $V$ come from the eigendecomposition $A^\top A = V\Lambda V^\top$. This decomposition does not drag the computation since the dictionary $A$ is prepared before any experiments, and hence, $V$ and $\Lambda$ can be pre-computed. Also the largest dimension $K$ of $A$ is usually much less than $N$, and it can be seen that the number of nonzero eigenvalues can never exceed $n$. As a result, the computations of $\alpha_j$'s can be carried out in parallel, and each of them only involves matrix-vector multiplication.

Thirdly, for fixed $w$, $\alpha$ and $\sigma$, the minimization of $u$ is

$$\min_u \|w_x - D_x u\|^2 + \|w_y - D_y u\|^2 + \sum_{j=1}^{J} \gamma \|A\alpha_j - R_j u\|^2 + \xi \|\mathcal{F}_p u - f_p\|^2, \tag{14}$$

Here $D_x, D_y$ are $N$-square matrices formed by the top and bottom rows of $D_i \in \mathbb{R}^{2\times N}$, $i = 1, \cdots, N$, and hence $D_x u, D_y u$ represent the gradient of $u$ along the $x$ and $y$ directions, respectively. $w_x$ and $w_y$ are the first and second column of $w$, respectively, and $\gamma = \mu\nu/\eta$, $\xi = \xi(\sigma) = \lambda/\eta\sigma^2$. Thus the Euler-Lagrange equation of (14) yields

$$Lu = r, \tag{15}$$

where

$$L = D_x^\top D_x + D_y^\top D_y + \sum_{j=1}^{J} \gamma R_j^\top R_j + \xi \mathcal{F}_p^\top \mathcal{F}_p$$

and

$$r = D_x^\top w_x + D_y^\top w_y + \sum_{j=1}^{J} \gamma R_j^\top A\alpha_j + \xi \mathcal{F}_p^\top f_p.$$

Under the periodic boundary condition for $u$, the finite difference operators $D_x$ and $D_y$ are block circulant matrices with circulant blocks and hence can be diagonalized by Fourier matrix $\mathcal{F}$. Thus, $\hat{D}_x = \mathcal{F} D_x \mathcal{F}^\top$ and $\hat{D}_y = \mathcal{F} D_y \mathcal{F}^\top$ are diagonal. Also, periodic boundary condition enables us to extract patches that cover each pixel $m$ times, where $m = n/d^2$ and $d$ is the sliding distance between all concatenated patches. Usually, we fix the patch size as $8\times 8$, i.e. $n = 64$, and choose $(d, m) = (8, 1)$

or $(4,4)$. Since $\sum_j R_j^\top R_j$ is a diagonal matrix with the $i$-th diagonal entry counting the number of times the $i$-th pixel was covered by patches, we have $\sum_j R_j^\top R_j = mI$, where $I$ is the identity matrix. So multiplying $\mathcal{F}$ on both sides of (15) gives

$$\hat{L}\mathcal{F}(u) = \hat{r}, \tag{16}$$

where

$$\hat{L} = \hat{D}_x^\top \hat{D}_x + \hat{D}_y^\top \hat{D}_y + m\gamma I + \xi P^\top P$$

is a diagonal matrix since $P^\top P$ is diagonal, and

$$\hat{r} = \hat{D}_x^\top \mathcal{F}(w_x) + \hat{D}_y^\top \mathcal{F}(w_y) + \gamma \mathcal{F}(u_\alpha) + \xi P^\top f_p$$

where $u_\alpha = \sum_j R_j^\top A\alpha_j$ is an "image" assembled using patches that are represented by dictionary $A$ and $\alpha$.

Finally, the computation of first variation of $F(u, \sigma, f_p)$ gives an update of $\sigma$ in each iteration:

$$\sigma = \sqrt{\|\mathcal{F}_p u - f_p\|^2/p}. \tag{17}$$

Therefore, similar to the algorithm developed in [25], the main computation of our algorithm for the CS model using a dictionary also involves only shrinkage, matrix-vector multiplication and fast Fourier transform.

3.2. **Numerical algorithm and convergence analysis.** Based on the discussion above, we are ready to propose the fast algorithm used for solving model (7). Note that the solution to (7) can be approximated by solving (9) with continuation on the penalty parameters $\eta$ and $\theta$ [14]. For stopping criterion, we let "res" be the maximum absolute/norm value of increments of $w, \alpha, \beta, u$, and terminate each inner loop once res $< \epsilon$ for a prescribed error tolerance $\epsilon$, then update $u$ and start a new loop with doubled $\eta$ and $\theta$. The upper bound $2^{12}$ for $\eta, \theta$ is chosen empirically so it is sufficiently large that solutions to (9) is a close approximation of (7). Based on derivations above, we summarize the algorithm for our model as Algorithm 1.

**Algorithm 1**

MR Image Reconstruction via Sparse Representation (**recMRI**)

- Input $P$, $f_p$, and $\mu, \nu, \lambda, \epsilon > 0$. Initialize $u = \mathcal{F}_p^\top f_p$, $\eta = \theta = 2^6$ and $\alpha = 0$.
- While $\eta, \theta < 2^{12}$, do
  1. Given $u$ and $\alpha$, compute $w$ and $\beta$ using (10) and (11).
  2. For $j = 1$ to $J$, do
     - Given $u$ and $\beta$, compute $\alpha_j$ using (13).
  3. Given $w$ and $\alpha$, compute $u$ by solving (16) and update $\sigma$ by (17).
  4. If res $< \epsilon$ go to 5, otherwise go to 1.
  5. Return $u^{\eta,\theta}$
  6. $u \leftarrow u^{\eta,\theta}$, $(\eta, \theta) \leftarrow (2\eta, 2\theta)$

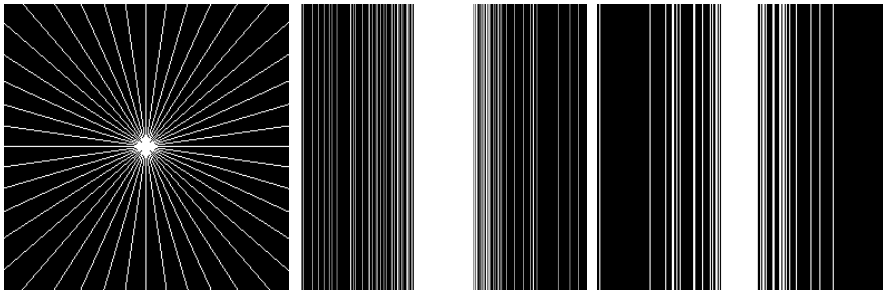
The proof of the convergence of the proposed algorithm 1 is similar to the one given in [23] with slight modifications, and thus is omitted here.

4. **Experimental results.** In this section, we present the experimental results of the proposed model (7) using Algorithm 1 and the comparisons with that resulting from using wavelet transform on both synthetic and in vivo MR data. All implementations involved in the experiments were coded in Matlab v7.3 (R2006b), except

the shrinkage and wavelet transform operators, which were coded in C++. Computations were performed on a Linux (version 2.6.16) workstation with Intel Core 2 CPU at 1.86GHz and 2GB memory.

4.1. **Improvement on accuracy by using dictionaries.** To show the improvement on the accuracy of reconstructions by using dictionaries as sparsifying transforms, we applied model (2), which uses wavelet as the sparsifying transform, and the proposed model (7) to three data sets. The three data sets are: the default Shepp-Logan phantom image provided by Matlab, a 2D axial brain MR image and a 2D chest MR image. The sampling masks used for these three data sets are depicted in Fig. 3, where white pixels indicate sampled locations in $k$-space. We used pseudo radial mask for phantom, and Cartesian mask undersampling phase encoding (PE) lines for in vivo data to simulate random acquisition. All of the $k$-space data in the simulated pseudo-radial trajectory is located on Cartesian grid.

In practice, CS-MRI algorithm prefers random acquisition trajectory that can lead to incoherent artifacts aliasing. However, the trajectory of the acquired data in each echo time (TE) is limited by the MR system, and hence true random acquisition is not possible. In recent years, acquisition schemes that are feasible and can produce incoherent aliasing artifacts are developed, e.g. radial and spiral trajectories. In our experiments, for simplicity, we used pseudo sampling masks which can simulate randomness in acquisition for demonstration purpose.

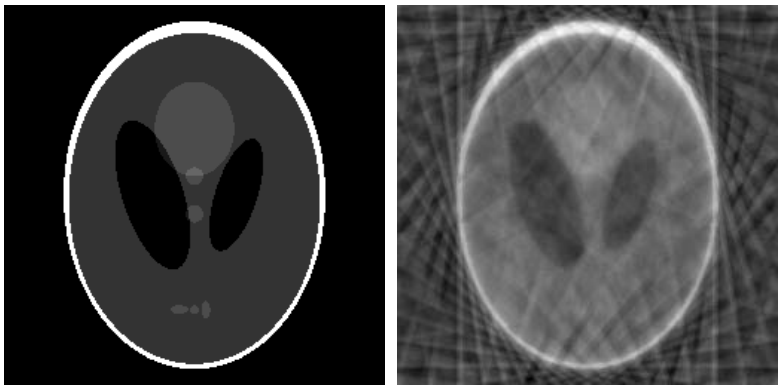


(a) SR=8.4% (b) SR=34.0% (c) SR=25.0%

Figure 3. Sampling masks used for (a) phantom (b) brain image and (c) chest image with sampling ratios (SR).

4.1.1. *Results of Phantom Reconstruction.* The default Shepp-Logan phantom of size $256 \times 256$ is shown in Fig. 4(a). Then a full $k$-space data was simulated by the 2D Fast Fourier transform (`fft2` in Matlab) of the phantom. We used the pseudo radial mask (shown in Fig. 3(a)) to the full $k$-space data and added complex valued Gaussian random noise with standard deviation (of magnitude) 3.2 to simulate the partially acquired data $f_p$. Direct using FFT of zero filling unscanned $k$-space locations results in notorious artifact aliasing, as shown in Fig. 4(b).

Then we applied model (2) with Haar wavelet and model (7) with an overcomplete discrete cosine transform (DCT) consisting of 256 atoms of size $8 \times 8$ [1] as the dictionary to the partial data $f_p$. The parameters we used for both models were $(\mu, \lambda, \epsilon) = (1, 10^3, 10^{-3})$ and the parameter $\nu$ was set to 1. The results by using wavelet and dictionary are shown in Fig. 4(c) and 4(d), and the corresponding RMSEs are 2.47% and 2.18%, and SNR are 32.9 and 34.5, respectively.



(a) Reference (b) Zero-Filling

(c) Wavelet Rec. (d) Dictionary Rec.

FIGURE 4. Reconstructed phantom image from simulated partial $k$-space data. (a) Reference image. (b) Zero-filling. (c) Reconstruction by model (2) with Haar wavelet (RMSE=2.47%, SNR=32.9) (d) Reconstruction using overcomplete DCT dictionary (RMSE=2.18%, SNR=34.5).

4.1.2. *Results of Brain Image Reconstruction.* The second test is on an axial brain MR image. The 2-dimentional multi-slice data set was collected on a 3T GE system (GE Healthcare, Waukesha, Wisconsin, USA) using the T1 FLAIR sequence (FOV 220mm, matrix size $512 \times 512$, TR 3060ms, TE 126ms, flip angle 90°, slice thickness 5mm, number of averages 1) with an 8-channel head coil (Invivo Corporation, Gainesville, FL, USA). Phase encoding direction was anterior-posterior. The imaging process output a high resolution and SNR image of size $512 \times 512$, which were used as the reference image in our experiment, as shown in Fig. 5(a).

The simulated full $k$-space data in our experiment was obtained by the Fourier transform of this reference image, and then was artificially undersampled using the Cartesian mask shown in Fig. 3(b), which led to the partial $k$-space data $f_p$. The zoomed-in of the square area in Fig. 5(a) is shown in Fig. 5(b). With only 34.0% data for reconstruction, strong aliasing artifacts can be observed in the image

reconstructed by zero-filling unscanned $k$-space locations, and the zoomed-in is shown in Fig. 5(c).

In this experiment the database used for training a dictionary consists of 4096 $8\times 8$ patches extracted randomly from four 2D brain MR images of different normal subjects (excluding the one to be reconstructed one) with the same acquisition sequence. The trained dictionary $A \in \mathbb{R}^{64\times 512}$, as shown in Fig.1, consists of 512 atoms of size $8 \times 8$.

We applied model (2) with Haar wavelet and model (7) with this trained dictionary to the partial Fourier data $f_p$ for brain MR image reconstruction. The parameters $\lambda$, $\mu$ and $\epsilon$ were set to 2e+3, 1 and 5e-4 in model (2) and (7), respectively. The parameter $\nu$ in model (7) was set to $10^6$. The zoomed-in area of the reconstructed images by model (2) and proposed model (7) are shown in Fig. 5(e) and 5(f), respectively. The RMSEs of reconstructions are 8.52% for model (2) and 7.74% for proposed model (7), and SNRs are 20.7 and 22.0, respectively. It can be seen that the image reconstructed by model (2) has oil-painting effect. On the contrary, the image reconstructed by the proposed model has better preserved fine structures. This further confirms the higher accuracy obtained by the proposed method.

We also simulated a low resolution (LR) image by using the 34.0% central PE lines (i.e. all white vertically lines in the middle in Fig. 3(b)), which has RMSE 10.32% and SNR 17.2.

4.1.3. *Results of Chest Image Reconstruction.* We also validate the proposed method on chest MR images. In this experiment the dictionary was trained by slices extracted from a three-dimensional (3D) MR chest data set, that consists of 19 adjacent 2D image slices of size $256 \times 256$ near the thorax region. Our procedure of training a dictionary is as follows: we randomly chose four slices and decomposed them into non-overlapping patches of size $8 \times 8$, and discarded those patches with homogeneous intensities, and then use $K$-SVD on the remaining patches to train a dictionary $A \in \mathbb{R}^{8^2\times 256}$ with 256 atoms. It is worth noting that if additional data (e.g. chest MR data scanned from different subjects but the same sequence, which are usually available in clinical applications) were available, one can readily construct a dictionary that is comparable to the one we trained using adjacent slices, and obtain similar reconstruction results as we showed below.

To demonstrate the improved accuracy by using dictionary, we randomly chose a 2D slice shown in Fig. 6(a), which is different from the four used as training slices. Then we artificially downsampled its $k$-space data using a Cartesian mask with 25% sampling ratio, as shown in Fig. 3(c). Zero-filling the unscanned $k$-space locations results in severe artifacts as shown in Fig. 6(b) with RMSE 15.59%. We again simulated a low resolution (LR) image, as shown in Fig. 6(d), by using the 25.0% central PE lines, which has RMSE 14.44% and SNR 15.4. From the corresponding error map Fig. 6(g), i.e. the absolute difference to the reference image, we can see the potential loss of edges and diagnostic information.

The reconstructions performed by using model (2) with Haar wavelet and model (7) with the trained dictionary are shown in Fig. 6(e) and 6(f), and the corresponding RMSEs are 12.04% and 8.48%, and SNRs are 17.3 and 20.1. In both cases $\lambda$, $\mu$ and $\epsilon$ were set to 1e+4, 2.5 and 1e-5, respectively. Parameter $\nu$ in model (7) was set to $10^5$. The error maps of these two reconstructions are shown in Fig. 6(h) and 6(i), respectively. It can be seen that the image reconstructed by the proposed model

(7) has lower artifacts level and better preserved edges. This experiment demonstrates again the advantages of using prior information to define the sparsifying transformation, which results in higher accuracy of reconstructions.

4.2. **Improvement on robustness (to parameter selection) and efficiency.** To demonstrate the improvement on the robustness of the proposed model (7) with respect to the choice of parameter $\lambda$, we tested the reconstruction of the Shepp-Logan phantom using model (2) with SSD and the proposed model (7) using MLE as as data consistency measures on various choices of $\lambda$. The resulting RMSEs are shown in table 1. From the changes of RMSEs, we can see the model with MLE as data consistency measure generated similarly good results whereas the model with SSD failed when $\lambda$ went improperly large. This result shows that the proposed model with MLE data consistency measure is much less sensitive to the choice of $\lambda$ and hence makes the reconstruction more robust.

Table 1 also shows the CPU time (in seconds) for phantom reconstruction using three different algorithms: nonlinear conjugate gradient (CG) algorithm for model (2), algorithm 1 **recMRI** for model (7) with the term involving dictionary

$$\sum_{j=1}^{J}\left(\|\alpha_j\|_1+\frac{\nu}{2}\|A\alpha_j-R_ju\|^2\right)$$

replaced by Haar wavelet term $\|\Psi^\top u\|_1$, and algorithm 1 **recMRI** for model (7) with overcomplete DCT as dictionary. It can be seen that the proposed numerical method was over 2.6 times faster than conjugate gradient based method. The dictionary based sparse representation consistently produced images with lower RMSE than wavelet based method, but it takes longer reconstruction time due to the redundancy of dictionaries. Moreover, using optimized discrete wavelet transform (DWT) package in Matlab makes the computation for wavelet based model even faster. Therefore, we expect an improvement on speed by code optimization when using dictionaries.

TABLE 1. Comparison of results of phantom reconstructions using nonlinear conjugate gradient (CG) for model (2) with Haar wavelet, algorithm **recMRI** with Haar wavelet, and **recMRI** for model (7) with overcomplete DCT dictionary. *Wavelet transforms were generated using optimized DWT package for Matlab

| **Method** | **CG**(Wavelet*) | | | **recMRI**(Wavelet*) | | | **recMRI**(Dictionary) | | |
|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | RMSE | Obj | CPU | RMSE | Obj | CPU | RMSE | Obj | CPU |
| 1e+2 | 5.93% | 12.13 | 86.3 | 7.93% | 13.98 | 28.2 | 5.21% | 11.79 | 211 |
| 1e+3 | 2.47% | 11.92 | 71.6 | 2.52% | 12.11 | 27.7 | 2.18% | 10.92 | 199 |
| 1e+4 | 5.05% | 3.147 | 71.4 | 4.98% | 3.025 | 26.9 | 3.47% | 2.540 | 198 |
| 1e+5 | 25.9% | 2.271 | 87.1 | 5.93% | 1.375 | 27.0 | 3.67% | 1.116 | 201 |
| 1e+6 | 37.0% | 2.165 | 81.2 | 6.16% | 1.129 | 28.7 | 5.52% | 1.091 | 212 |

4.3. **Robustness of dictionary training and reconstruction.** In the experiment on brain MR image reconstruction in section 4.1.2, the patches used by $K$-SVD algorithm were randomly extracted from the four training images. Different patches may lead to different trained dictionary using $K$-SVD algorithm, and hence may impact the consequent reconstructions. Therefore it is important to verify that the dictionary training and reconstruction process are robust to certain level

of variability on the training database used in $K$-SVD algorithm. In this experiment, we repeated 10 times of the entire process from forming a data set of training image to using trained dictionary in brain MR image reconstruction as described in section 4.1.2. The difference is that, in each run, the 4096 training images are randomly chosen from a pool of patches (around $50,000$) extracted from images acquired under the same sequence as that used for the image to be reconstructed. Therefore, the training patches used in one run are different from those in another. The RMSEs and SNRs of reconstruction results are shown in table 2. Meanwhile,

Table 2. Experimental results of 10 runs of dictionary training and brain image reconstruction as in 4.1.2.

| Runs | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| RMSE(%) | 7.74 | 7.77 | 7.65 | 7.81 | 7.83 | 7.78 | 7.79 | 7.79 | 7.72 | 7.68 |
| SNR | 22.0 | 21.8 | 22.6 | 21.6 | 21.6 | 21.8 | 21.7 | 21.7 | 22.1 | 22.5 |

when we directly used the overcomplete DCT instead of the trained dictionary in the reconstruction (7), the reconstruction had RMSE 8.27% and SNR 21.1. Table 2 indicates that the $K$-SVD algorithm and the consequent reconstructions using trained dictionaries can consistently generate good results despite that the training patches may vary. Therefore, the proposed scheme using dictionaries trained by $K$-SVD algorithm in MR image reconstruction is stable and robust, and hence has great practical potential.

## 5. Conclusion.

A novel variational model and a fast numerical algorithm are introduced in this paper to improve the accuracy, robustness and efficiency of CS-MRI reconstructions. A trained dictionary is used to better preserve edges and fine structures by taking advantage of prior information; negative log-likelihood estimation of the recovering error is used to automatically adjust the weighting on the data consistency to improve the robustness of the model to the choice of parameters; a quadratic penalty approach is adopted to speed up the reconstruction. The qualitative and quantitative comparisons on both synthetic and in vivo data demonstrate the improvements by the proposed method.

## REFERENCES


[1] M. Aharon, M. Elad and A. Bruckstein, *The k-svd: An algorithm for designing of overcomplete dictionaries for sparse representation*, IEEE Trans. Signal Process., **54** (2006), 4311–4322.

[2] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imag. Sci., **2** (2009), 4311–4322.

[3] J. Bioucas-Dias and M. Figueiredo, *A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration*, IEEE Trans. Image Process., **16** (2007), 2992–3004.

[4] K. Block, M. Uecker and J. Frahm, *Undersampled radial mri with multiple coils. Iterative image reconstruction using a total variation consraint*, Magn. Reson. Med., **57** (2007), 1086–1098.

[5] E. J. Candes and J. K. Romberg, *Signal recovery from random projections*, Proceedings of SPIE Computational Imaging III, **5674** (2005), 76–86.

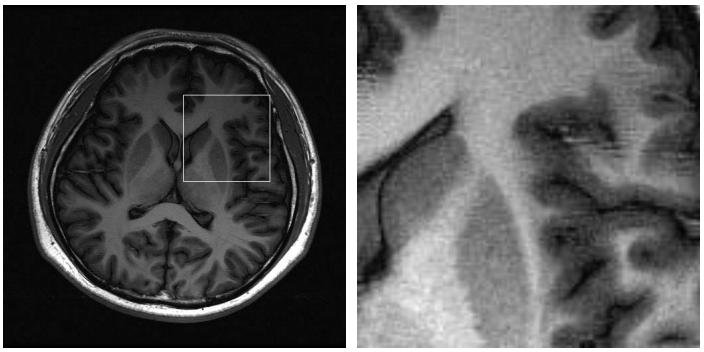
[6] E. J. Candes, J. K. Romberg and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inf. Theory, **52** (2006), 489–509.

[7] A. Chambolle, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vis., **20** (2004), 89–97.

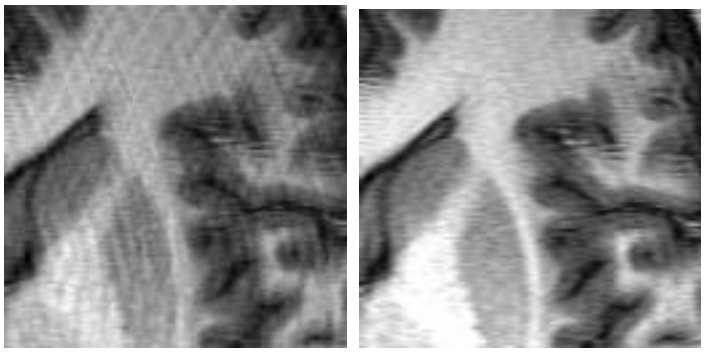
[8] T. Chang, L. He and T. Fang, *Mr image reconstruction from sparse radial samples using bregman iteration*, Proc. Intl. Soc. Mag. Reson. Med., (2006), 696.
[9] S. Chen, D. Donoho and M. Saunders, *Atomic decomposition by basis pursuit*, SIAM J. Sci. Comput., **20** (1999), 33–61.
[10] D. L. Donoho, *For most large underdetermined systems of linear equations, the minimal l1-norm solution is also the sparsest solution*, Commun. Pure Appl. Math., **59** (2006), 907–934.
[11] D. L. Donoho, M. Elad and V. Temlyakov, *Stable recovery of sparse overcomplete represenation in the presence of noise,* IEEE Trans. Inf. Theory, **52** (2006), 6–18.
[12] M. Elad and M. Aharon, *Image denoising via sparse and redundant representations over learned dictionaries*, IEEE Trans. Image Process., **15** (2006), 3736–3745.
[13] T. Goldstein and S. Osher, *The split bregman method for l1 regularized problems*, SIAM J. Imag. Sci., **2** (2009), 323–343.
[14] E. Hale, W. Yin and Y. Zhang, *Fixed-point continuation method for $\ell_1$-minimization with applications to compressed sensing*, SIAM J. Optim., **19** (2008), 1107–1130.
[15] L. He, T.-C. Chang, S. Osher, T. Fang and P. Speier, "Mr Image Reconstruction by Using the Iterative Refinement Method and Nonlinear Inverse Scale Space Methods," CAM Report 06-35, UCLA, 2006.
[16] S. Kim, K. Koh, M. Lustig and S. Boyd, *An efficient method for compressed sensing*, Proceedings of IEEE Intl. Conf. Image Process.(ICIP), **3** (2007), 117–120.
[17] S.-J. Kim, K. Koh, M. Lustig, S. Boyd and D. Gorinevsky, *An interior-point method for large-scale $\ell_1$-regularized least squares*, SIAM J. Optim., **1** (2007), 606–617.
[18] H. Liao and G. Sapiro, *Sparse representation for limited data tomography*, Proceedings of IEEE Intl. Symp. Biomed. Imag. (ISBI), (2008), 1375–1378.
[19] M. Lustig, D. Donoho and J. M. Pauly, *Sparse mri: The application of compressed sensing for rapid mr imaging*, Magn. Reson. Med., **58** (2007), 1182–1195.
[20] Y. Nesterov, *Gradient methods for minimizing composite objective function*, Center for Operations Research and Econometrics (CORE), Catholic Univ. Louvain, Louvain-la-Neuve, Belgium, Discussion Paper, **76** (2007).
[21] L. Rudin, S. Osher and E. Fatemi, *Non-linear total variation noise removal algorithm*, Physics D., **60** (1992), 259–268.
[22] J. Starck, M. Elad and D. Donoho, *Image decomposition via the combination of sparse representations and a variational approach*, IEEE Trans. Image Process., **14** (2005), 1570–1582.
[23] Y. Wang, J. Yang, W. Yin and Y. Zhang, *A new alternating minimization algorithm for total variation image reconstruction*, SIAM J. Imag. Sci., **1** (2008), 248–272.
[24] S. J. Wright, R. D. Nowak and M. Figueiredo, *Sparse reconstruction by separable approximation,* IEEE Trans. Signal Process., **57** (2009), 2479–2493.
[25] J. Yang, Y. Zhang and W. Yin, "A Fast tvl1-l2 Minimization Algorithm for Signal Reconstruction from Partial Fourier Data," Technical Report 08-29, CAAM Rice Univ., (2008).
[26] J. Ye, S. Tak, Y. Han and H. Park, *Projection reconstruction mr imaging using focuss*, Magn. Reson. Med., **57** (2007), 764–775.
[27] M. Zhu and T. Chan, "An Efficient Primal-Dual Hybrid Gradient Algorithm for Total Variation Image Restoration," Technical Report 08-34, CAM UCLA, (2008).
[28] M. Zhu, S. Wright and T. Chan, *Duality-based algorithms for total-variation-regularized image restoration*, Comput. Optim. Appl., (2008).

Received March 2009; revised January 2010.

*E-mail address:* yun@ufl.edu
*E-mail address:* xye@ufl.edu
*E-mail address:* f.huang@philips.com



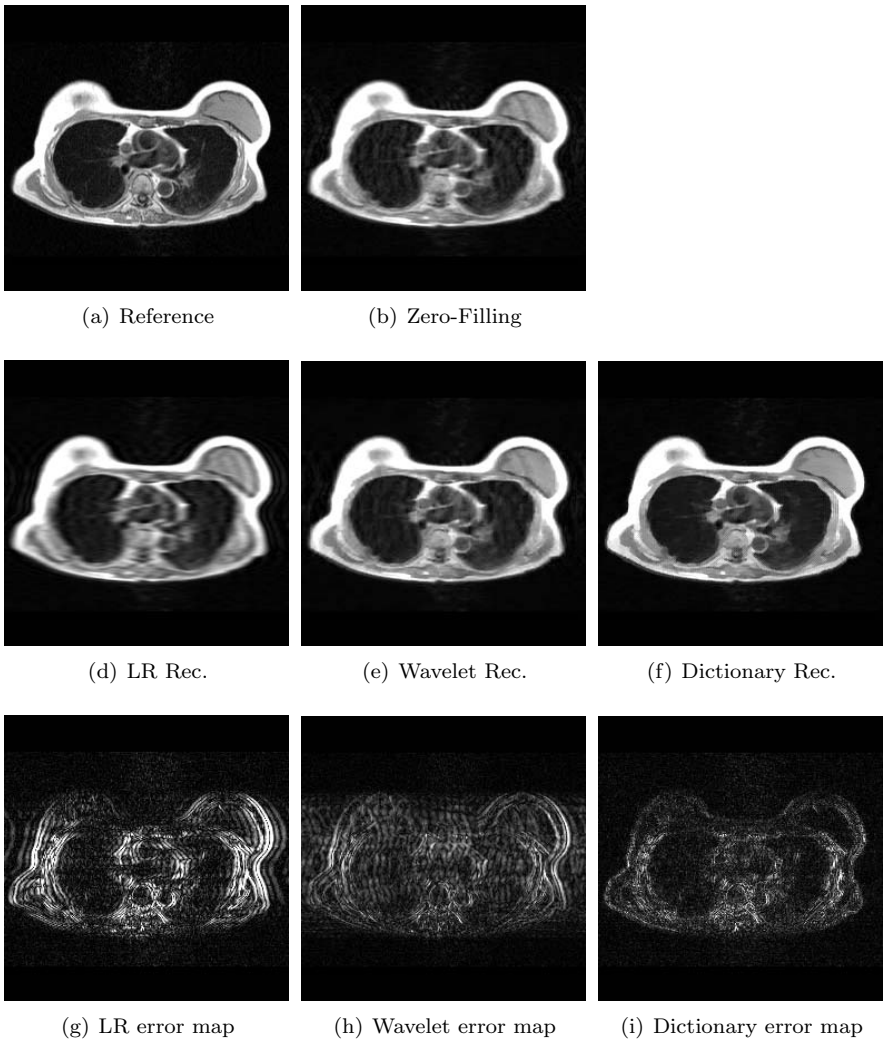
(a) Reference (b) Zoomed-In Reference



(c) Zoomed-In Zero-Filling (d) Zoomed-In LR

(e) Zoomed-In Wavelet (f) Zoomed-In Dictionary

Figure 5. Reconstruction of brain MR Image. (a) Reference. (b) Zoomed-in of square area in the reference image. (c) Zoomed-in of reconstruction by zero-filling unscanned $k$-space locations. (d) Zoomed-in of low resolution (LR) image reconstructed by 34.0% central PE lines, with RMSE 10.32% and SNR 17.2. (e) Zoomed-in of reconstruction obtained using wavelet as sparsifying transform, RMSE is 8.52% and SNR is 20.7. (f) Zoomed-in of reconstruction obtained using trained dictionary shown in Fig. 1, RMSE is 7.74% and SNR is 22.0.



(a) Reference (b) Zero-Filling

(d) LR Rec. (e) Wavelet Rec. (f) Dictionary Rec.

(g) LR error map (h) Wavelet error map (i) Dictionary error map

FIGURE 6. Reconstructions of chest MR image using model (2) and (7). (a) Reference image. (b) Zero-filling unscanned locations, RMSE is 15.59% and SNR is 14.2. (d) LR image obtained by using central PE lines, RMSE is 14.44% and SNR is 15.4. (e) Reconstruction by model (2) with Haar wavelet, RMSE is 12.09% and SNR is 17.1. (f) Reconstruction by proposed model (7) with trained dictionary, RMSE is 8.48% and SNR is 20.1. Figures (g), (h) and (i) are corresponding error maps of (d), (e) and (f) to the reference image (a).