Wavelet-based Edge Correlation Incorporated Iterative Reconstruction for Undersampled MRI

Changwei Hu, Xiaobo Qu, Di Guo, Lijun Bao, Zhong Chen *

Departments of Electronic Science and Communication Engineering, Fujian Key Laboratory of Plasma and Magnetic Resonance, Xiamen University, Xiamen 361005, China

Submitted to ***Magnetic Resonance Imaging***

* Corresponding author. Department of Electronic Science, Xiamen University, Xiamen, 361005, China. Tel.: +86-592-2181712; fax: +86-592-2189426.
E-mail: chenz@xmu.edu.cn

Abstract

Undersampling k-space is an effective way to decrease acquisition time for MRI. However, aliasing artifacts introduced by undersampling may blur the edges of magnetic resonance images, which often contain important information for clinical diagnosis. Moreover, k-space data is often contaminated by the noise signals of unknown intensity. To better preserve the edge features while suppressing the aliasing artifacts and noises, we present a new wavelet-based algorithm for undersampled MRI reconstruction. The algorithm solves the image reconstruction as a standard optimization problem including a ℓ_2 data fidelity term and ℓ_1 sparsity regularization term. Rather than manually setting the regularization parameter for the ℓ_1 term, which is directly related to the threshold, an automatic estimated threshold adaptive to noise intensity is introduced in our proposed algorithm. In addition, a prior matrix based on edge correlation in wavelet domain is incorporated into the regularization term. Compared with non-linear conjugate gradient descent algorithm, iterative shrinkage/thresholding algorithm, fast iterative soft-thresholding algorithm, and the iterative thresholding algorithm using exponentially decreasing threshold, the proposed algorithm yields reconstructions with better edge recovery and noise suppression.

Keywords: Compressed sensing; Edge correlation; Iterative thresholding; MRI reconstruction; Continuation scheme

1. Introduction

MRI, a widely used analytical tool for medical diagnosis, is burdened by slow data acquisition. An effective way to speed up MRI is to undersample k-space. However, undersampling violates the Nyquist-Shannon sampling theorem, resulting in aliasing artifacts in reconstructed magnetic resonance (MR) images. In addition, k-space is often contaminated by signals due to the coils and eddy currents in the patient [1]. Both the artifacts and the noise signals will affect the clarity of the MR image edges, which usually contain significant information for pathological diagnosis. For instance, the edges and textures in brain images are useful for diagnosis and research of schizophrenia and Alzheimer's disease [2]. The degree of liver fibrosis, which can be measured by MR image texture analysis, is a useful predictive factor for the occurrence of hepatocellular carcinoma [3]. The tumor margin, caliber of vessel, and the vessel border are suggestive of extramural vascular invasion, which is a pathologic feature predictive of distant relapse and poor survival among patients with colorectal cancer [4]. Therefore, undersampled MRI reconstruction with good edge recovery is important for some clinical applications, such as the applications mentioned above.

Compressed sensing (CS) proposed by Candès et al. [2] and Donoho [3] is a new sampling and compression theory. CS reconstructs the $N \times 1$ signal $\mathbf{x}$ from far fewer M ($M \ll N$) measurements $\mathbf{y}$ ($\mathbf{y} = \mathbf{\Phi x}$, $\mathbf{\Phi}$ is a $M \times N$ measurement matrix) than Nyquist sampling rule by exploiting the sparsity of signal $\mathbf{x}$ in a certain

transform domain.

Undersampled MRI reconstruction is a special case of CS where the measurements are Fourier coefficients (k-space samples) for the Fourier encoding scheme. If the MR image vector $\mathbf{x}$ can be sparsely represented by a transform $\mathbf{\Psi}$ with coefficient vector $\mathbf{w}$ ($\mathbf{x}=\mathbf{\Psi w}$), then $\mathbf{x}$ can be accurately reconstructed from a small subset of k-space data by solving the ℓ_0 norm minimization problem

$$\min_{\mathbf{w}} \|\mathbf{w}\|_0 , \ s.t. \ \mathbf{y}=\mathbf{F}_u\mathbf{\Psi w} , \tag{1}$$

where $\mathbf{F}_u=\mathbf{UF}$, and $\mathbf{U}$ is a $M\times N$ undersampling matrix, $\mathbf{F}\in\mathbb{C}^{N\times N}$ represents the forward Fourier transform.

Unfortunately, the ℓ_0 norm is not convex, and the computational complexity of the optimization is non-polynomial (NP) hard [7]. To overcome this difficulty, one option is to optimize with the ℓ_1, ℓ_p $(0<p<1)$ [8, 9], or smoothed ℓ_0 norm [10-12] instead. However, the ℓ_p $(0<p<1)$ and smoothed ℓ_0 norm minimization could sink into local minima, and ℓ_1 norm minimization requires more measurements for exact reconstruction [5, 6]. The detailed discussion is beyond the scope of this work. In this paper, we utilize the widely used ℓ_1 norm minimization to enforce the sparsity of solutions by replacing the ℓ_0 norm

$$\min_{\mathbf{w}} \|\mathbf{w}\|_1 , \ s.t. \ \mathbf{y}=\mathbf{F}_u\mathbf{\Psi w} . \tag{2}$$

As measured k-space data $\mathbf{y}$ is often contaminated by noise, the data consistency in Eq. (2) is violated. The reconstruction is then obtained by solving

$$\min_{\mathbf{w}} \|\mathbf{w}\|_1 , \ s.t. \ \|\mathbf{y}-\mathbf{F}_u\mathbf{\Psi w}\|_2<\varepsilon, \tag{3}$$

where ε is the error tolerance and controls the reconstruction fidelity [13].

The constrained optimization problem in Eq. (3) can be written in the Lagrangian form

$$\min_{\mathbf{w}} \ \frac{1}{2}\|\mathbf{y}-\mathbf{F}_u\boldsymbol{\Psi}\mathbf{w}\|_2^2+\lambda\|\mathbf{w}\|_1, \tag{4}$$

where λ is a regularization parameter governing the tradeoff between the reconstruction error and its sparsity.

A successful application of CS requires the sparsity of the desired MR image. Most MR images do show sparsity in certain transform domains. Angiograms, for instance, are structurally simple and sparse in identity transform domain [13]. More complicated MR images can be sparsified by total variation [14], wavelet (WT) [13], contourlet [15, 16], or some more complicated transform, such as combined sparsifying transforms [17], and dictionary with more than one basis function [18].

Image edges exhibit high spatial correlation in the WT domain, both within and across scales, and therefore can be located very effectively [19]. According to CS MRI requirements in [13], aliasing artifacts introduced by ideal sampling patterns for k-space undersampling should be incoherent (noise like, rather than edge feature-like) in the sparsifying transform domain. If sampling patterns meet this requirement, the correlation can be used as a helpful tool to discriminate edges from aliasing artifacts. Thus, good reconstruction of edges and suppression of aliasing artifacts can be expected.

In this paper, we present a WT-based edge correlation incorporated algorithm (ECIA) for undersampled MRI reconstruction. A prior matrix, which incorporates the

inter- and intra-scale edge correlation in WT domain into Eq. (4), is designed to modulate the wavelet coefficients. ECIA modifies the iterative thresholding algorithm using exponentially decreasing threshold (IT-EDT) [15] to make use of the prior matrix. In addition, as the k-space data is often contaminated by noise of unknown intensity, it is sometimes difficult to set the appropriate value of the regularization parameter λ in Eq. (4). In ECIA, the value of regularization parameter is automatically assigned according to an estimated lowest threshold, which is calculated according to the noise intensity.

This paper is organized as follows. In Section 2, we first give an introduction of undecimated WT and IT-EDTC algorithm. Then the proposed algorithm is presented, including the calculation of the estimated lowest threshold and the prior matrix design. In Section 3, we use the ECIA for undersampled MRI reconstructions. The performance of ECIA is compared with non-linear conjugate gradient descent algorithm (NLCG) [13], iterative shrinkage/thresholding algorithm (IST) [20, 21], fast iterative soft-thresholding algorithm (FISTA) [22], and IT-EDT [15]. The effect of estimated lowest threshold on noise suppression, the reconstruction time, and the empirical convergence of the algorithm are also reported. The discussion and conclusions part are given in Section 4.

2. WT-based ECIA algorithm for MRI reconstruction

2.1. Undecimated WT

Traditional orthogonal WT reduces resolution by one-half at each level via subsampling data. It is not easy to follow the evolution of edges through scales using orthogonal WT. In addition, as the orthogonal WT produces fewer coefficients at coarse scale, edges within coarser scales are difficult to track.

An alternative referred as undecimated WT has been developed. Undecimated WT eliminates the decimation step in the orthogonal WT transform. It is redundant and has the same number of coefficients at all scales, which allows edge analysis pixel by pixel. This property is convenient for investigation of the edge correlation in inter- and intra-scales. What is more, the redundancy of the sparsifying transform has the potential benefit to improve the reconstruction quality [15]. Thus à trous WT [24], a computationally efficient and widely used undecimated WT, is employed to sparsify the MR image in this work.

2.2. WT-based IT-EDT for CS MRI

The classic interpretations of iterative thresholding (IT) for solving ℓ_1 norm minimization were reported previously [21]. For theoretic analysis, Herrity and co-workers [24] employed hard IT to demonstrate that one could recover the k-term representation of the original signal up to any prescribed error tolerance under certain conditions. Bredies and Lorenz [25] proved that soft IT converged with a linear rate once the underlying operator satisfied the finite basis infectivity property or the minimizer possessed a strict sparsity pattern. Inspired by these works, Qu and co-workers [15] applied IT-EDT, which was originally used for NMR spectra reconstruction [26], to solve Eq. (4). In addition to the IT-EDT, there are also some

other soft-thresholding algorithms, such as IST [20, 21] and FISTA [22]. IST and FISTA seek the solution to Eq. (4) by applying the iteration step

$$\mathbf{w}_{t+1} = S_{\theta_t}(\mathbf{w}_t + \frac{1}{c}(\mathbf{F}_u\mathbf{\Psi})^H\mathbf{r}), \tag{5}$$

where $\mathbf{r}$ is the residual in k-space, $c > \left\|(\mathbf{F}_u\mathbf{\Psi})^H\mathbf{F}_u\mathbf{\Psi}\right\|_2$, $\theta_t = \frac{\lambda}{c}$.

As our proposed algorithm is modified on the basis of IT-EDT, we will give the pseudo-code of IT-EDT, which is implemented by the following steps.

Algorithm IT-EDT

(1) Initialize the relative error tolerance R_e, $t = 0$, maximal iteration times t_{max}, $\mathbf{w}_0 = [0, 0, \cdots, 0]^T$, $\mathbf{r} = \mathbf{y}$, ρ ($0 < \rho < 1$), $\theta_0 = \max((\mathbf{F}_u\mathbf{\Psi})^H\mathbf{r_0})$;

(2) **While** $\|\mathbf{y} - \mathbf{F}_u\mathbf{\Psi}\mathbf{w}_t\|_2 / \|\mathbf{y}\|_2 > R_e$ **and** $t < t_{max}$

(3) $\mathbf{w}_{t+1} = \mathbf{w}_t + S_{\theta_t}((\mathbf{F}_u\mathbf{\Psi})^H\mathbf{r})$;

(4) $\mathbf{r} = \mathbf{y} - \mathbf{F}_u\mathbf{\Psi}\mathbf{w}_{t+1}$;

(5) $\theta_{t+1} = \rho\theta_t$;

(6) $t = t + 1$;

(7) **End While.**

In line (3), $S_{\theta_t}(\cdot)$ is a soft-thresholding operator with θ_t as the threshold, $(\mathbf{F}_u\mathbf{\Psi})^H$ is the adjoint operator of $\mathbf{F}_u\mathbf{\Psi}$.

We find that IT-EDT employs a threshold θ_t decreasing exponentially with iteration count, which is somewhat similar to the continuation strategy adopted in the methods of gradient projection for sparse reconstruction (GPSR) [27], fixed-point

continuation (FPC) [28], and sparse reconstruction by separable approximation (SpaRSA) with continuation [29]. Instead of solving Eq. (4) directly with λ, the continuation strategy obtains the final solution using a decreasing sequence $\{\lambda_1, \lambda_2, \cdots \lambda_t, \lambda_{t+1}, \cdots\}$ ($\lambda_1 < \lambda_2 < \cdots < \lambda_t < \lambda_{t+1} < \cdots$) as the regularization parameter. It was proved that continuation strategy yielded a fast convergence [28, 29]. In IT-EDT, however, when the threshold decreases from θ_t to θ_{t+1}, one iteration in IT-EDT may not obtain an optimal solution to the current problem in the problem sequence. Inspired by the works in Refs. [28] and [29], we embed IST as an inner-loop in IT-EDT. In each IST inner-loop, as the threshold is determined by $\theta_t = \frac{\lambda_t}{c}$, $\{\lambda_1, \lambda_2, \cdots \lambda_t, \lambda_{t+1}, \cdots\}$ is therefore equivalent to a threshold incorporated sequence $\{c\theta_1, c\theta_2, \cdots c\theta_t, c\theta_{t+1}, \cdots\}$. The modified IT-EDT is referred as IT-EDT with continuation (IT-EDTC), and is given by the following pseudo-code.

Algorithm IT-EDTC

(1) initialize the relative error tolerance R_e, maximal outer-loop and inner-loop iteration times t_{maxout}, t_{maxin}, inner-loop iteration count $t = 1$, outer-loop iteration count $t_{out} = 1$, $\mathbf{w}_1 = [0, 0, \cdots, 0]^T$, $\mathbf{r} = \mathbf{y}$, ρ ($0 < \rho < 1$), $\theta_1 = \max((\mathbf{F}_u \boldsymbol{\Psi})^H \mathbf{r})$;

(2) **Outer-loop: While** $\|\mathbf{y} - \mathbf{F}_u \boldsymbol{\Psi} \mathbf{w}_t\|_2 / \|\mathbf{y}\|_2 > R_e$ **and** $t_{out} < t_{maxout}$

(3) **Inner-loop: While** $t < t_{maxin}$

(4) $$\mathbf{w}_{t+1} = S_{\theta_{t_{out}}}(\mathbf{w}_t + \frac{1}{c}(\mathbf{F}_u \boldsymbol{\Psi})^H \mathbf{r});$$

(5) $\mathbf{r} = \mathbf{y} - \mathbf{F}_u \mathbf{\Psi} \mathbf{w}_{t+1}$;

(6) $t = t + 1$;

(7) **End While (inner-loop ends);**

(8) $\mathbf{w}_1 = \mathbf{w}_t$ **;**

(9) $t = 1$;

(10) $\theta_{t_{out}+1} = \rho \theta_{t_{out}}$;

(11) $t_{out} = t_{out} + 1$;

(12) **End While (outer-loop ends).**

2.3. The Proposed Algorithm

For the traditional CS-MRI algorithms using soft-thresholding to solve Eq. (4) with noisy measurements, such as IST and FISTA, the regularization parameter λ needs to be set (usually manually) in advance. Some algorithms, such as L-curve [30], can be used to set this parameter, but need to solve the problem several times and then it is possible to find a good regularization parameter. The proposed ECIA presents one way to stop the iteration by estimating a lowest threshold θ_{low} based on the noise estimation in the WT subbands, and the algorithm stops when the threshold reaches θ_{low}. Similar to IT-EDTC, ECIA also uses IST as the inner-loop, therefore the regularization parameter $\lambda = c\theta_{low}$ is automatically obtained according to θ_{low}.

In addition, to obtain better edge recovery, ECIA uses the following model by plugging a correlation matrix $\mathbf{B}$ into Eq. (4),

$$\min_{\mathbf{w}} \ \frac{1}{2}\|\mathbf{y}-\mathbf{F}_u\boldsymbol{\Psi}\mathbf{w}\|_2^2+\lambda\|\mathbf{B}\mathbf{w}\|_1, \tag{6}$$

where $\mathbf{B}$ is a binary diagonal matrix containing the inter- and intra-scale correlation of edges in the WT domain. Below we will explain the estimation of lowest threshold and the design of correlation matrix $\mathbf{B}$ in detail.

2.3.1 Estimation of Lowest Threshold

In the recent years, plenty of researches have addressed the development of statistical models for image denoising. Accurate statistical model, designed directly on images or their transform coefficients, is critical for the denoising results. Sendur and Selesnick [31] proposed a WT-based bivariate shrinkage algorithm with local variance estimation for image denoising. The algorithm models the statistical dependency of the wavelet coefficients, and defines a nonlinear thresholding function (shrinkage function) using Bayesian estimation theory. Inspired by this work, we derive the estimated lowest threshold θ_{low} from the shrinkage function.

According to the algorithm in Ref. [31], suppose $\mathbf{W}_s^k(m,n)$ is a noise-corrupted WT coefficient in the k th subband at scale s with the spatial location (m,n), $[\upsilon(m,n)]^2$ is the marginal variance of coefficient $\mathbf{W}_s^k(m,n)$ in a local neighborhood, $\sigma^2 = median(|\mathbf{W}_s|)/0.6745$ is the noise variance estimated from the wavelet coefficients, then the estimation from $\mathbf{W}_s^k(m,n)$ is calculated by the following bivariate shrinkage function

$$\hat{\mathbf{W}}_s^k(m,n)=\frac{f(\sqrt{\mathbf{W}_s^k(m,n)^2+\mathbf{W}_{s+1}^k(m,n)^2}-\frac{\sqrt{3}\sigma^2}{\upsilon(m,n)})}{\sqrt{\mathbf{W}_s^k(m,n)^2+\mathbf{W}_{s+1}^k(m,n)^2}}\cdot\mathbf{W}_s^k(m,n), \tag{7}$$

where function $f(\cdot)$ is defined as

$$f(x)=\begin{cases}0, & \text{if } x<0\\ x, & \text{otherwise}\end{cases} \tag{8}$$

Empirically, we find that for most WT coefficients, $\left|\mathbf{W}_{s+1}^k(m,n)/\mathbf{W}_s^k(m,n)\right|$ varies within a small range, as shown in Fig. 1B, and therefore can be approximately viewed as a constant $\tau=\left|\mathbf{W}_{s+1}^k(m,n)/\mathbf{W}_s^k(m,n)\right|$. After the following manipulations, Eq. (7) can be rewritten as a soft-thresholding function

$$\begin{aligned}\hat{\mathbf{W}}_s^k(m,n)&=\frac{f(\sqrt{\mathbf{W}_s^k(m,n)^2+\mathbf{W}_{s+1}^k(m,n)^2}-\frac{\sqrt{3}\sigma^2}{\upsilon(m,n)})}{\sqrt{\mathbf{W}_s^k(m,n)^2+\mathbf{W}_{s+1}^k(m,n)^2}}\cdot\mathbf{W}_s^k(m,n)\\ &=\frac{f(\sqrt{1+\tau^2}\left|\mathbf{W}_s^k(m,n)\right|-\frac{\sqrt{3}\sigma^2}{\upsilon(m,n)})}{\sqrt{1+\tau^2}\left|\mathbf{W}_s^k(m,n)\right|}\cdot\mathbf{W}_s^k(m,n)\\ &=\begin{cases}0, & \text{if } -\frac{\sqrt{3}\sigma^2}{\upsilon(m,n)\sqrt{1+\tau^2}}\le\mathbf{W}_s^k(m,n)\le\frac{\sqrt{3}\sigma^2}{\upsilon(m,n)\sqrt{1+\tau^2}}\\ \mathbf{W}_s^k(m,n)-\frac{\sqrt{3}\sigma^2}{\upsilon(m,n)\sqrt{1+\tau^2}}\cdot\frac{\mathbf{W}_s^k(m,n)}{\left|\mathbf{W}_s^k(m,n)\right|}, & \text{otherwise}\end{cases}\\ &=S_{\sqrt{3}\sigma^2/(\upsilon(m,n)\sqrt{1+\tau^2})}(\mathbf{W}_s^k(m,n))\end{aligned} \tag{9}$$

where with $\sqrt{3}\sigma^2/(\upsilon(m,n)\sqrt{1+\tau^2})$ as the threshold, therefore the lowest threshold θ_{low} is then estimated with

$$\theta_{low}=\min(\sqrt{3}\sigma^2/(\upsilon\sqrt{1+\tau^2}))\,. \tag{10}$$

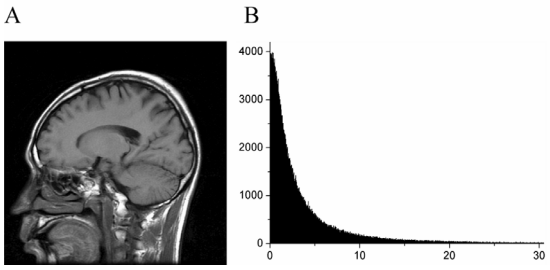

Fig. 1. (A) Original image. (B) Histogram of $\left|\mathbf{W}_{s+1}^{k}(m,n)/\mathbf{W}_{s}^{k}(m,n)\right|$.

2.3.2. Correlation Matrix Design

The design of matrix $\mathbf{B}$ implies the following beliefs. (a) The edge features usually have signal peaks across different WT scales. (b) Within each WT scale, the coefficients corresponding to edge features tend to cluster together and show a spatial continuity [19]. The former and the latter reflect the inter- and intra-scale dependencies of edge features respectively. If we can discriminate edges from non-edge WT coefficients in matrix $\mathbf{B}$ using the inter- and intra-scale dependencies, reconstructions with better edge recoveries can be expected. Correlation matrix $\mathbf{B}$ is designed in the context of IT-EDTC algorithm, as shown in Fig. 2.

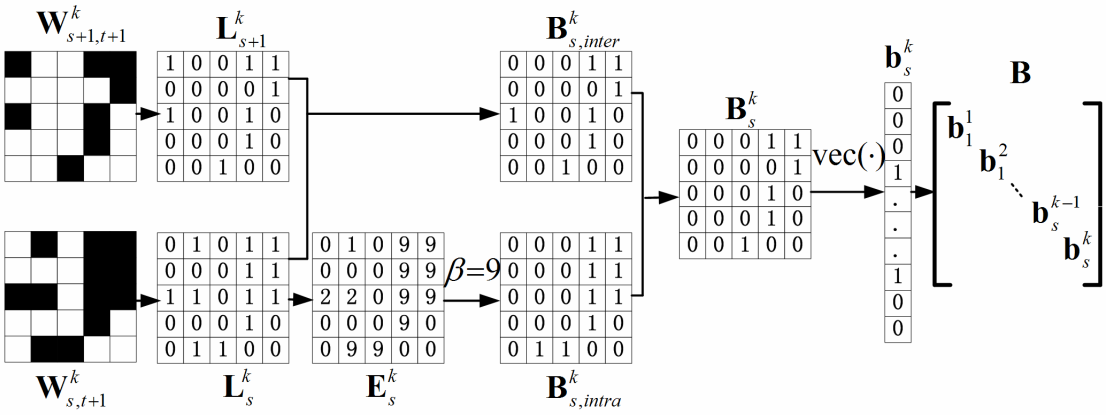

Fig. 2. The design of correlation matrix $\mathbf{B}$.

Suppose $\mathbf{W}_{s,t+1}^{k}$ is WT coefficients obtained from the soft-thresholding (see line (4) of IT-EDTC pseudo-code), with k and s being the subband and scale index. If we regard the WT coefficients with amplitude larger than the threshold as the signal peaks of edges, then these peaks can be labeled out using the nonzero entries in $\mathbf{W}_{s,t+1}^{k}$. In Fig. 2, let the white squares in $\mathbf{W}_{s,t+1}^{k}$ denote the zero entries, and the black squares the nonzero entries. The signal peaks is therefore labeled in binary matrix $\mathbf{L}_{s}^{k}$ by

$$\mathbf{L}_{s}^{k}(m,n)=\begin{cases}1, & \text{if } \mathbf{W}_{s,t+1}^{k}(m,n)\neq 0\\ 0, & \text{otherwise}\end{cases} \tag{11}$$

As the inter-scale edge dependency usually involves in the two adjacent WT scales [19], we therefore build a binary inter-scale correlation matrix $\mathbf{B}_{s,inter}^{k}$ labeling the edges by

$$\mathbf{B}_{s,inter}^{k}(m,n)=\begin{cases}1, & \text{if } \mathbf{L}_{s}^{k}(m,n)=1 \text{ and } \mathbf{L}_{s+1}^{k}(m,n)=1\\ 0, & \text{otherwise}\end{cases} \tag{12}$$

As for the intra-scale dependency, we use the number of entries in each eight-connected nonzero regions of $\mathbf{W}_{s,t+1}^{k}$ to measure the spatial continuity of the edges. The eight-connected nonzero region is a group of nonzero entries in which each member can touch at least one member at its adjacent vertical, horizontal, or diagonal positions. The more entries the eight-connected nonzero regions contain, the better they exhibit spatial continuity. For instance, in $\mathbf{L}_{s}^{k}$ of Fig. 2, there are 3 eight-connected regions with 1, 2, and 9 nonzero entries respectively, as labeled in matrix $\mathbf{E}_{s}^{k}$. To select the regions with better spatial continuity, we design a guiding

map $\mathbf{B}_{s,intra}^{k}$ by

$$\mathbf{B}_{s,intra}^{k}(m,n)=\begin{cases}1, & \text{if } \mathbf{E}_{s}^{k}(m,n)\geq\beta\\ 0, & \text{if } \mathbf{E}_{s}^{k}(m,n)<\beta\end{cases} \tag{13}$$

where β is a positive integer controlling the spatial continuity of the selected edges. For instance, when β is set to 9 in Fig. 2, only the eight-connected region with strongest spatial continuity is labeled out in $\mathbf{B}_{s,intra}^{k}$.

To take both inter- and intra-scale dependencies into account, the correlation matrix $\mathbf{B}_{s}^{k}$ is therefore defined as

$$\mathbf{B}_{s}^{k}(m,n)=\begin{cases}1, & \text{if } \mathbf{B}_{s,inter}^{k}(m,n)=1 \text{ and } \mathbf{B}_{s,intra}^{k}(m,n)=1\\ 0, & \text{otherwise}\end{cases} \tag{14}$$

Then the correlation matrix $\mathbf{B}$ is built by

$$\mathbf{B}=diag(vec(\{\mathbf{B}_{1}^{1},\mathbf{B}_{1}^{2},\mathbf{B}_{1}^{3},\cdots,\mathbf{B}_{s}^{k-1},\mathbf{B}_{s}^{k}\})), \tag{15}$$

where $vec(\cdot)$ is an operator stacking a matrix into a column vector, $diag(\cdot)$ creates a diagonal matrix with a vector down the diagonal. Fig. 3 shows an example of $\mathbf{B}_{s}^{k}$ for a fully sampled image, where image edges are effectively located in the guiding matrix.

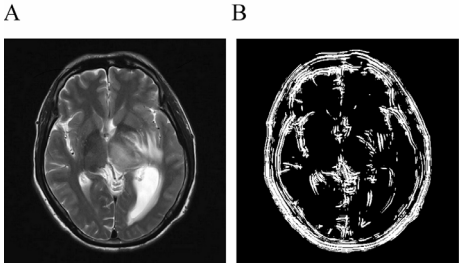

Fig. 3. (A) Original image. (B) The guiding maps for the 3rd scale of à trous WT. The decomposition level is 4.

2.3.3. The Proposed Algorithm

Once obtaining the estimated lowest threshold θ_{low} and the correlation matrix $\mathbf{B}$, the proposed ECIA algorithm is implemented according to the following steps.

Algorithm ECIA

(1) Initialize the maximal outer-loop and inner-loop iteration times t_{maxout}, t_{maxin}, inner-loop iteration count $t=1$, outer-loop iteration count $t_{out}=1$, $\mathbf{w}_1=[0,0,\cdots,0]^T$, $\mathbf{r}=\mathbf{y}$, ρ ($0<\rho<1$), $\boldsymbol{\theta}_1=[\theta(1),\theta(2),\cdots,\theta(3L+1)]$, where $\theta(i)$ is the maximal amplitude of different subbands of $(\mathbf{F}_u\boldsymbol{\Psi})^H\mathbf{r}$, L is WT decomposition levels;

(2) **Outer-loop: While** $\theta(1),\theta(2),\cdots,\theta(3L+1)>\theta_{low}$ and $t_{out}<t_{maxout}$

(3) estimate θ_{low} according to Eq. (10);

(4) **Inner-loop: While** $t<t_{maxin}$

(5) $\mathbf{w}_{t+1}=S_{\theta_{t_{out}}}(\mathbf{w}_t+\frac{1}{c}(\mathbf{F}_u\boldsymbol{\Psi})^H\mathbf{r})$;

(6) update $\mathbf{B}$ according to Eqs. (11)-(15);

(7) $\mathbf{w}_{t+1}=\mathbf{B}\mathbf{w}_{t+1}$;

(8) $\mathbf{r}=\mathbf{y}-\mathbf{F}_u\boldsymbol{\Psi}\mathbf{w}_{t+1}$;

(9) $t=t+1$;

(10) **End While (inner-loop ends);**

(11) $\mathbf{w}_1=\mathbf{w}_t$**;**

(12) $t=1$;

(13) $\theta_{t+1}=\max(\rho\theta_t,\theta_{low})$;

(14) $t_{out} = t_{out} + 1$;

(15) **End While (outer-loop ends).**

2.3.4. Improvement on the Speed of ECIA

Compared with IT-EDTC, two factors will slow down the ECIA. (a) During lowest threshold estimation, the calculation of noise variance σ^2 requires sorting the coefficients of undecimated WT. (b) As for $\mathbf{L}_s^k$ of large size, searching the eight-connected nonzero entries is time-consuming.

In Ref. [31], the noise variance σ^2 was only calculated at the finest wavelet scale. Similarly, we estimate θ_{low} only from the diagonal subband of the finest scale to reduce the time for coefficients sorting. On the other hand, we divide matrix $\mathbf{L}_s^k$ into small patches and search the eight-connected regions in each patch.

3. Simulation results

3.1. Edge recovery

Many sampling patterns are proposed for CS MRI. Non-Cartesian sampling patterns, such as radial and spiral sampling patterns, were used for MRI reconstruction in [32, 33]. However, Cartesian sampling pattern [13] is the most popular trajectory for k-space data acquisition. In simulations, variable density Cartesian sampling pattern [13] with the rate of 0.4 (40% measurements), as shown in Fig. 4A, is used for k-space sampling. The fully sampled MR images for simulations

are obtained from a 1.5 T GE MRI scanner with a fast-recovery fast spin echo (FRFSE) T2 weighted sequence, as shown in Fig. 4B and D (Fig. 4B: TR/TE = 4020/103 ms, 24×24 cm field of view, 7 mm slice thickness; Fig. 4D: TR/TE = 4000/102 ms, 24×24 cm field of view, 6 mm slice thickness). Gaussian white noise with variance 0.02 is added to both the real and imaginary parts of k-space data respectively.

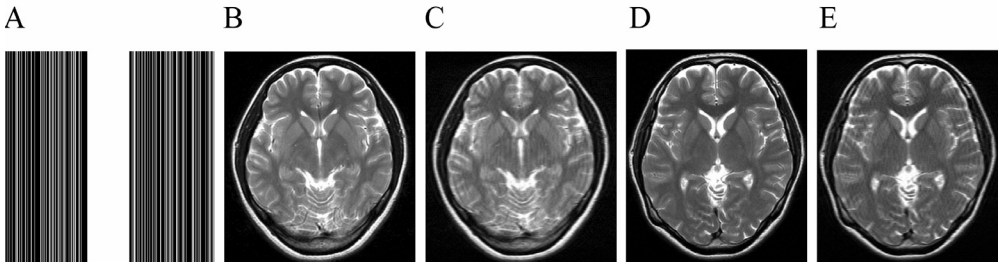

Fig. 4. (A) Variable density Cartesian sampling pattern with 0.4 sampling rate. (B) and (D) are fully sampled MR images. (C) and (E) are the zero filling undersampled MR images of (B) and (D) using the sampling pattern in (A).

The performance of the proposed ECIA is compared with NLCG, IST, FISTA, and IT-EDTC algorithms. Reconstruction results by different algorithms are given in Figs. 5 and 6. À trous WT with spline biorthogonal filters and four decomposition levels is applied. The decreasing factor for IT-EDTC and ECIA is $\rho = 0.5$. Fig. 5F-J and Fig. 6F-J indicate that ECIA has the weakest edge features left in the difference image. Compared with other algorithms, ECIA achieves the most precise edge reconstruction.

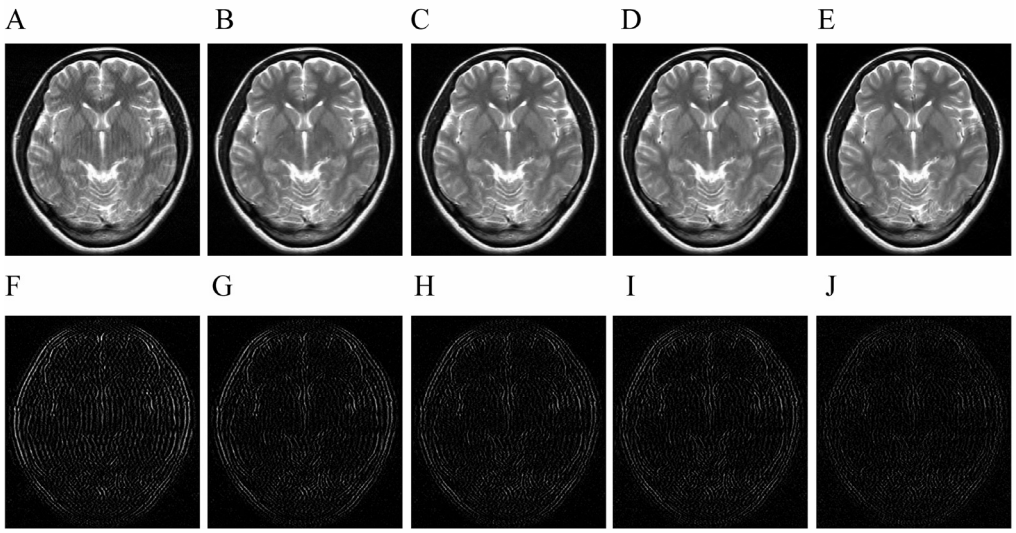

Fig. 5. (A)-(E) are the reconstructions of Fig. 4B by NLCG, IST, FISTA, IT-EDTC, and ECIA, respectively. (F)-(J) are the difference images between fully sampled MR image and (A)-(E) with the gray scale of [0,50].

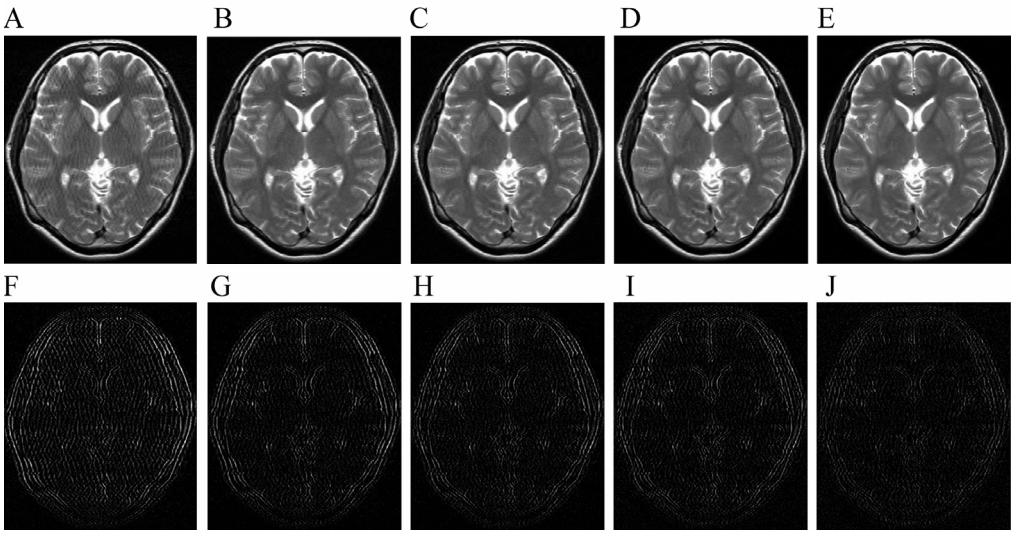

Fig. 6. (A)-(E) are the reconstructions of Fig. 4D by NLCG, IST, FISTA, IT-EDTC, and ECIA, respectively. (F)-(J) are the difference images between fully sampled MR image and (A)-(E) with the gray scale of [0,50].

3.2. Noise suppression

For quantitative comparisons of noise suppressions between ECIA and other algorithms under different noise levels, signal-to-noise ratio (SNR) between reconstructions and fully sampled MR images are computed. SNR is defined as $SNR = 10 \times \log_{10}(\left\|\mathbf{x} - \bar{\mathbf{x}}\right\|_2^2 / \left\|\mathbf{x} - \mathbf{x}_{rec}\right\|_2^2)$, where $\mathbf{x}$ is the fully sampled MR image, $\bar{\mathbf{x}}$ is the mean value of $\mathbf{x}$, and $\mathbf{x}_{rec}$ is the reconstruction result. The curves of SNR versus noise variance are given in Fig. 7 for the same MR image in Fig. 5. It indicates that ECIA yields SNR with 2-6 dB higher than those of NLCG, IST, FISTA, and IT-EDTC.

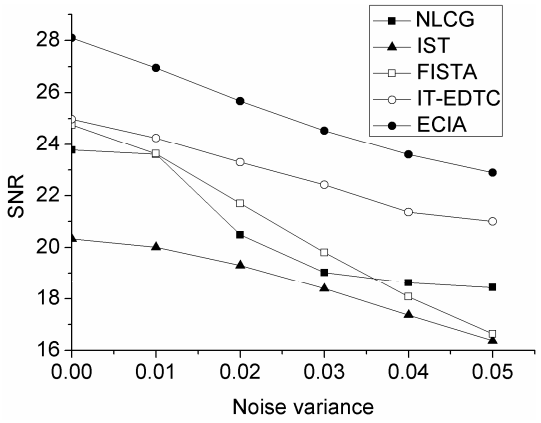

Fig. 7. Curves of SNR versus noise variance of different algorithms.

To show the value of θ_{low} under different noise levels, the curve of θ_{low} versus noise variance are presented in Fig. 8A. It indicates that θ_{low} increases with the growth of noise variance. When heavy noise is added to the k-space, more WT coefficients are submerged into the noise, in which case a higher θ_{low} will decrease the introduction of significant noise. To show the accuracy of θ_{low} estimation for

noise suppression, Fig. 8B and C give curves of SNR between reconstruction results of ECIA and fully sampled MR image under different θ_{low}. In Fig. 8B and C, Gaussian noise with variance 0.02 and 0.05 are added to k-space respectively, and we find that highest SNRs are both achieved near the corresponding θ_{low} estimated under the two different noise levels in Fig. 8A.

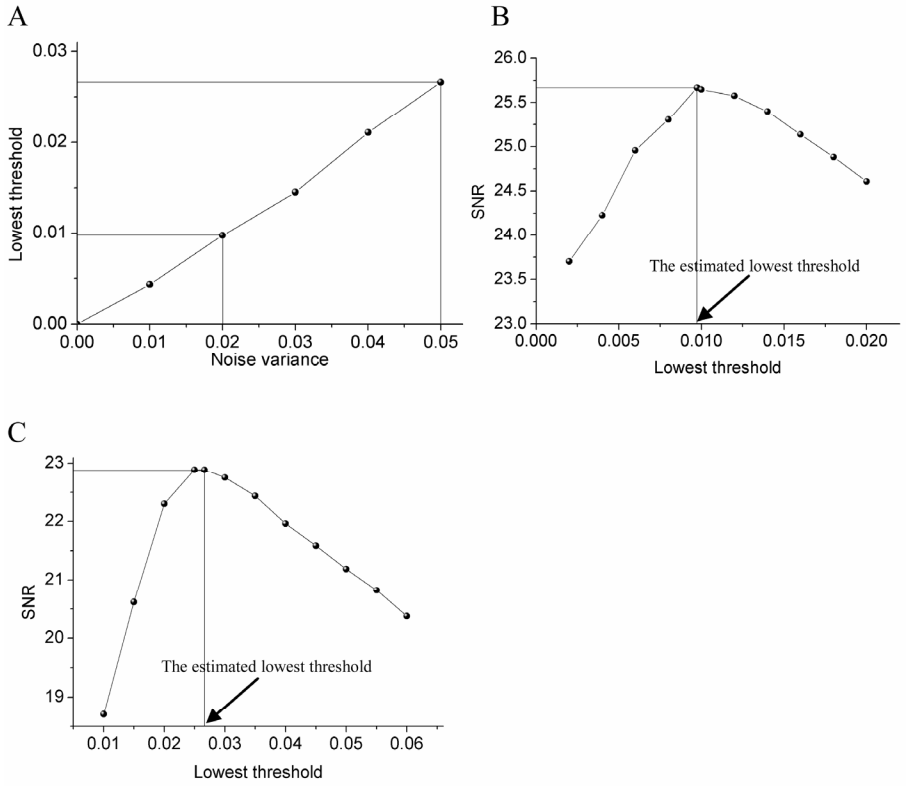

Fig. 8. (A) The curve of estimated lowest threshold versus noise variance. (B) The curve of SNR versus lowest threshold, Gaussian noise with variance 0.02 is added. (C) The curve of SNR versus lowest threshold, Gaussian noise with variance 0.05 is added.

3.3. Reconstruction time comparison

In this section, we report results of experiments aiming at comparing the speed of ECIA with other algorithms. Fig. 4B and D are used for experiments. All the experiments are performed using MATLAB, on a computer equipped with an Intel 2.4 GHz processor, with 2.0 GB of RAM, and Windows XP system. All the algorithms are carried out until their SNRs stabilize. Table 1 gives the average CPU time of 5 instances for each experiment.

Table 1 Reconstruction time comparison between different algorithms

Algorithms		NLCG	IST	FISTA	IT-EDTC	ECIA
CPU times (s)	Fig. 4B	948	5120	690	334	402
	Fig. 4D	894	5085	667	329	403

We can observe that the computations of IT-EDTC and ECIA are fast, while IST takes more time than other algorithms. Due to the estimation of θ_{low} and searching the eight-connected nonzero entries, the reconstruction time of ECIA is about 20% more than that of IT-EDTC, which is acceptable considering the 2-3 dB SNR improvement compared with IT-EDTC in Fig. 7.

3.4. Empirical convergence of the objective function

In Fig. 9, we plot the evolution of objective function in Eq. (6) versus outer-loop iteration number. Fig. 4B is used for the experiment, and the maximal IST inner-loop iteration number t_{maxin} is set as 10. As the estimation of lowest threshold θ_{low} is

unknown beforehand, we first run the algorithm and record final estimated θ_{low}, the curve is then obtained with the recorded θ_{low} when we run the algorithm for the second time. From Fig. 9, we observe that the objective function decreases and gradually stabilizes when a threshold is fixed within the inner-loop iterations.

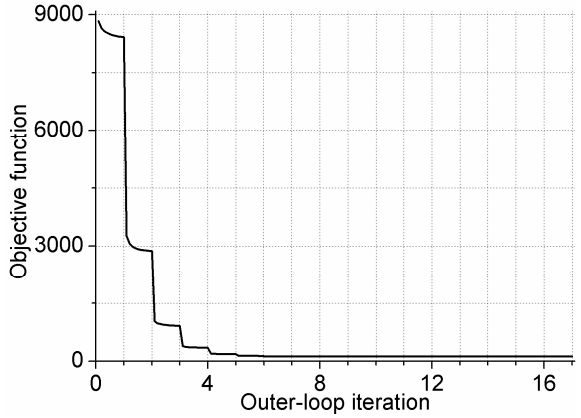

Fig. 9. The evolution of objective function versus outer-loop iteration number.

4. Discussion and conclusions

In this work, we propose an algorithm named ECIA. It automatically assigns the value of regularization parameter according to an estimated lowest threshold adaptive to the noise intensity, and incorporates a prior matrix based on edge correlation in the WT domain into the objective function. Simulations demonstrate that ECIA reconstructs MR images with better noise suppression and edge recovery compared with NLCG, IST, FISTA, and IT-EDTC algorithms, 2-6 dB improvement on SNR is achieved for the given MR images.

In addition, CS assumes that the signal of interest is sparse in a particular

transform domain. We only consider the prior information of edge correlation in the inter- and intra-scale for the WT transform. One possible extension may include designing models to make use of different prior information of other popular sparsifying transforms, such as contourlet, and discrete cosine transforms. In addition, we also expect our model be integrated with the non-convex optimization, e.g. replacing ℓ_1 norm with ℓ_p $(0 < p < 1)$ norm or smoothed ℓ_0 norm.

Acknowledgments

This work was supported by NNSF of China under Grants (10774125 and 10974164) and the Research Fund for the Doctoral Program of Higher Education of China under Grant (200803840019). Xiaobo Qu and Di Guo were supported by Postgraduates' Oversea Study Program for Building High-Level Universities from the China Scholarship Council.

Reference

[1] Redpath TW. Signal-to-noise ratio in MRI. The British J of Radiology 1998;71:704–707.

[2] Liu YX, Teverovskiy L, Carmichael O, Kikinis R, Shenton M, Carter CS, Stenger VA, Davis S, Aizenstein H, Becker J, Lopez O, Meltzer C. Discriminative MR image feature analysis for automatic schizophrenia and Alzheimer's disease classification. 7th International Conference on Medical Image Computing and Computer Aided Intervention. Saint-Malo, France, 2004. p. 393.

[3] Kato H, Kanematsu M, Zhang XJ, Saio M, Kondo H, Goshima S, Fujita H. Computer-aided diagnosis of hepatic fibrosis: Preliminary evaluation of MRI texture analysis using the finite difference method and an artificial neural network. Am J Roentgenol 2007;189:117–122.

[4] Smith NJ, Shihab O, Arnaout A, Swift RI, Brown G. MRI for detection of extramural vascular invasion in rectal cancer. Am J Roentgenol 2008;191:1517-1522.

[5] Candès EJ, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE T Inform Theory 2006;52:489–509.

[6] Donoho DL. Compressed sensing. IEEE T Inform Theory 2006; 2:1289–1306.

[7] Natarajan BK. Sparse approximate solutions to linear systems. SIAM J Comput 1995;24:227–234.

[8] Chartrand R. Exact reconstruction of sparse signals via nonconvex minimization. IEEE Signal Proc Let 2007;14:707–710.

[9] Majumdar A, Ward RK. An algorithm for sparse MRI reconstruction by Schatten *p*-norm minimization. Magn Reson Imaging 2011;29:408-417.

[10] Qu XB, Cao X, Guo D, Hu CW, Chen Z. Compressed sensing MRI with combined sparsifying transforms and smoothed ℓ_0 norm minimization. 35th International Conference on Acoustics, Speech, and Signal Processing. Dallas,

USA: IEEE; 2010. p. 626.

[11] Mohimani H, Babaie-Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed ℓ_0 norm. IEEE T Signal Proces 2009;57:289–301.

[12] Trzasko J, Manduca A. Highly undersampled magnetic resonance image reconstruction via homotopic ℓ_0 -minimization. IEEE T Med Imaging 2009;28:106–121.

[13] Lustig M, Donoho DL, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn Reson Med 2007;58:1182–1195.

[14] Chen LY, Schabel MC, DiBella EVR. Reconstruction of dynamic contrast enhanced magnetic resonance imaging of the breast with temporal constraints. Magn Reson Imaging 2010;28:637–645.

[15] Qu XB, Zhang WR, Guo D, Cai CB, Cai SH, Chen Z. Iterative thresholding compressed sensing MRI based on contourlet transform. Inverse Probl Sci En 2010;18:737–758.

[16] Gho SM, Nam Y, Zho SY, Kim EY, Kim DH. Three dimension double inversion recovery gray matter imaging using compressed sensing. Magn Reson Imaging 2010;28:1395–1402.

[17] Qu XB, Cao X, Guo D, Hu CW, Chen Z. Combined sparsifying transforms for compressed sensing MRI. Electron Lett 2010;46:121–123.

[18] Bao LJ, Zhu YM, Liu WY, Croisille P, Pu ZB, Robini M, Magnin IE. Denoising human cardiac diffusion tensor magnetic resonance images using sparse representation combined with segmentation. Phys Med Biol 2009;56:1435–1456.

[19] Xu YS, Weaver JB, Healy DM, Lu J. Wavelet transform domain filters: A spatially selective noise filtration technique interscale significant location. IEEE T Image Process 1994;3:747–758.

[20] Zibulevsky M, Elad M. L1-L2 optimization in signal and image processing: Iterative shrinkage and beyond. IEEE Signal Proc Mag 2010;27:76–88.

[21] Daubechies I, Defries M, De Mol C. An iterative thresholding algorithm for

linear inverse problems with a sparsity constraint. Commun Pur Appl Math 2004;57:1413–1457.

[22] Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci 2009;2:183–202.

[23] Mallat M. A Wavelet Tour of Signal Processing. New York: Academic Press; 1999.

[24] Herrity KK, Gilbert AC, Tropp JA. Sparse approximation via iterative thresholding. 31st International Conference on Acoustics, Speech, and Signal Processing. Toulouse, France: IEEE; 2006. p. 624.

[25] Bredies K, Lorenz DA. Linear convergence of iterative soft-thresholding. J Fourier Anal Appl 2008;14:813–837.

[26] Drori I. Fast ℓ_1 minimization by iterative thresholding for multidimensional NMR spectroscopy. EURASIP J Adv Sig Pr 2007.

[27] Figueiredo M, Nowak RD, Wright SJ. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. IEEE J–STSP 2007;4:586–597.

[28] Hale ET, Yin W, Zhang Y. Fixed-point continuation for ℓ_1-minimization: Methodology and convergence, SIAM J Optimiz 2008;19:1107–1130.

[29] Wright SJ, Nowak RD, Figueiredo M. Sparse reconstruction by separable approximation. IEEE T Signal Proces 2009;57:2479–2493.

[30] Hansen PC, Oleary DP. The use of the L-curve in the regularization of discrete ill-posed problems. SIAM J Sci Comput 1993;14:1487–1503.

[31] Sendur L, Selesnick IW. Bivariate shrinkage with local variance estimation. IEEE Signal Proc Let 2002;9:438–441.

[32] Song JY, Liu Y, Gewalt SL, Cofer G, Johnson GA, Liu QH. Least square NUFFT methods applied to 2D and 3D radially encoded MR image reconstruction. IEEE T Bio-med Eng 2009;56:1134–1142.

[33] Kashyap S; Yang ZL; Jacob M. Non-iterative regularized reconstruction algorithm for non-Cartesian MRI: NIRVANA. Magn Reson Imaging 2011;29:222–229.