Available online at www.sciencedirect.com

ScienceDirect

Magnetic Resonance Imaging 29 (2011) 907 – 915

MAGNETIC RESONANCE IMAGING

# Wavelet-based edge correlation incorporated iterative reconstruction for undersampled MRI ☆

Changwei Hu [a], Xiaobo Qu [a,b], Di Guo [b], Lijun Bao [a], Zhong Chen [a,b,]*

[a]Department of Electronic Science, Fujian Key Laboratory of Plasma and Magnetic Resonance, Xiamen University, Xiamen 361005, China

[b]Department of Communication Engineering, Xiamen University, Xiamen 361005, China

Received 19 January 2011; revised 12 April 2011; accepted 22 April 2011

**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 $\ell_2$ data fidelity term and $\ell_1$ sparsity regularization term. Rather than manually setting the regularization parameter for the $\ell_1$ 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 nonlinear 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. [5] and Donoho [6] is a new sampling and compression theory. CS reconstructs the $N\times1$ signal $\mathbf{x}$ from far fewer $M(M\ll N)$ measurements $\mathbf{y}$ ($\mathbf{y}=\mathbf{\Phi x}$, $\mathbf{\Phi}$ is an $M\times N$ 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

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

* Corresponding author. Tel.: +86 592 2181712; fax: +86 592 2189426.
*E-mail address:* chenz@xmu.edu.cn (Z. Chen).

0730-725X/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.mri.2011.04.016

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 $\ell_0$ 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 an $M\times N$ undersampling matrix and $\mathbf{F}\in\mathbb{C}^{N\times N}$ represents the forward Fourier transform.

Unfortunately, the $\ell_0$ norm is not convex and the computational complexity of the optimization is nonpolynomial hard [7]. To overcome this difficulty, one option is to optimize with the $\ell_1$, $\ell_p$ (0<$p$<1) [8,9] or smoothed $\ell_0$ norm [10–12] instead. However, the $\ell_p$ (0<$p$<1) and smoothed $\ell_0$ norm minimization could sink into local minima, and $\ell_1$ norm minimization requires more measurements for exact reconstruction [5,6]. The detailed discussion is beyond the scope of this work. In this article, we utilize the widely used $\ell_1$ norm minimization to enforce the sparsity of solutions by replacing the $\ell_0$ 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 $\varepsilon$ is the error tolerance and controls the reconstruction fidelity [13].

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

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

where $\lambda$ 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 Ref. [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 article, we present a WT-based edge correlation incorporated algorithm (ECIA) for undersampled MRI reconstruction. A prior matrix, which incorporates the inter- and intrascale 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 $\lambda$ in Eq. (4). In ECIA, the value of the regularization parameter is automatically assigned according to an estimated lowest threshold, which is calculated according to the noise intensity.

This article is organized as follows. In Section 2, we first give an introduction to the undecimated WT and IT-EDT 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 nonlinear 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 to 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 intrascales. What is more, the redundancy of the sparsifying transform has the potential benefit to improve the reconstruction quality [15]. Thus à trous WT [23], 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 $\ell_1$ norm minimization were reported previously [21]. For theoretical analysis, Herrity et al. [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 et al. [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

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

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

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 $t_{max}$, $\mathbf{w}_0=[0,0,L,0]^T$, $\mathbf{r}=\mathbf{y}$, $\rho(0<\rho<1)$, $\theta_0=\max((\mathbf{F}_u\boldsymbol{\Psi})^H\mathbf{r}_0)$;
(2) **While** $\|\mathbf{y}-\mathbf{F}_u\boldsymbol{\Psi}\mathbf{w}_t\|_2/\|\mathbf{y}\|_2>R_e$ **and** $t<t_{max}$
(3) $\mathbf{w}_{t+1}=\mathbf{w}_t+S_{\theta_t}((\mathbf{F}_u\boldsymbol{\Psi})^H\mathbf{r})$;
(4) $\mathbf{r}=\mathbf{y}-\mathbf{F}_u\boldsymbol{\Psi}\mathbf{w}_{t+1}$;
(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 $\theta_t$ as the threshold, $(\mathbf{F}_u\boldsymbol{\Psi})^H$ is the adjoint operator of $\mathbf{F}_u\boldsymbol{\Psi}$.

We find that IT-EDT employs a threshold $\theta_t$ decreasing exponentially with the iteration count, which is somewhat similar to the continuation strategy adopted in the methods of gradient projection for sparse reconstruction [27], fixed-point continuation [28] and sparse reconstruction by separable approximation with continuation [29]. Instead of solving Eq. (4) directly with $\lambda$, the continuation strategy obtains the final solution using a decreasing sequence $\{\lambda_1,\lambda_2,\ldots\lambda_t,\lambda_{t+1},\ldots\}$ $(\lambda_1<\lambda_2<\ldots<\lambda_t,\lambda_{t+1}<\ldots)$ 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 $\theta_t$ to $\theta_{t+1}$, 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 $\theta_t=\frac{\lambda_t}{c}$, $\{\lambda_1,\lambda_2,\ldots\lambda_t,\lambda_{t+1},\ldots\}$ is therefore equivalent to a threshold incorporated sequence $\{c\theta_1,c\theta_2,\ldots c\theta_t,c\theta_{t+1},\ldots\}$. The modified IT-EDT is referred to 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; $\mathbf{w}_1=[0,0,\ldots,0]^T$; $\mathbf{r}=\mathbf{y}$; $\rho(0<\rho<1)$; $\theta_1=\max((\mathbf{F}_u\boldsymbol{\Psi})^H\mathbf{r})$;
(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) $\mathbf{w}_{t+1} = S_{\theta_{t_{out}}}\left(\mathbf{w}_t + \frac{1}{c}(\mathbf{F}_u\boldsymbol{\Psi})^H\mathbf{r}\right)$;
(5) $\mathbf{r}=\mathbf{y}-\mathbf{F}_u\boldsymbol{\Psi}\mathbf{w}_{t+1}$;
(6) $t=t+1$;
(7) **End While (inner-loop ends)**;
(8) $\mathbf{w}_1=\mathbf{w}_t$;
(9) $t$=1;
(10) $\theta_{tout+1}=\rho\theta_{tout}$;
(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 $\lambda$ needs to be set (usually manually) in advance. Some algorithms, such as the 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 $\theta_{low}$ based on the noise estimation in the WT subbands, and the algorithm stops when the threshold reaches $\theta_{low}$. 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 $\theta_{low}$.

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 intrascale 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 research studies have addressed the development of statistical models for image denoising. An 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 $\theta_{low}$ 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

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

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

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

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

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

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

$$\theta_{low} = \min\left(\sqrt{3}\sigma^2/\left(\upsilon\sqrt{1+\tau^2}\right)\right) \quad (10)$$

### 2.3.2. Correlation matrix design

The design of matrix **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 intrascale dependencies of edge features, respectively. If we can discriminate edges from non-edge WT coefficients in matrix **B** using the inter- and intrascale dependencies, reconstructions with better edge recoveries can be expected.

Correlation matrix **B** is designed in the context of IT-EDTC algorithm, as shown in Fig. 2. 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 peak is therefore labeled in binary matrix $\mathbf{L}_s^k$ *by*

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

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

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

As for the intrascale 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 three eight-connected regions with one, two and nine nonzero entries, respectively, as labeled in matrix $\mathbf{E}_s^k$. To select the regions with better spatial continuity, we design a guiding map $\mathbf{B}_{s,\text{intra}}^k$ by

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

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

Taking both the inter- and intrascale dependencies into account, the correlation matrix $\mathbf{B}_s^k$ is therefore defined as

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

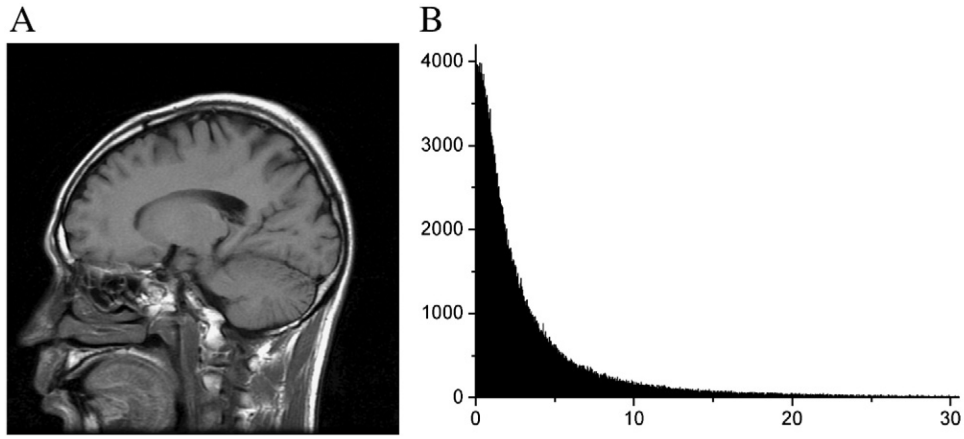


Fig. 1. (A) Original image. (B) Histogram of $|\mathbf{W}_{s+1}^{k}(m,n)/\mathbf{W}_{s}^{k}(m,n)|$.

Then the correlation matrix **B** is built by

$$\mathbf{B} = \mathrm{diag}\left(\mathrm{vec}\left(\left\{\mathbf{B}_1^1, \mathbf{B}_1^2, \mathbf{B}_1^3, \cdots, \mathbf{B}_s^{k-1}, \mathbf{B}_s^k\right\}\right)\right), \tag{15}$$

where vec(·) is an operator stacking a matrix into a column vector; diag(·) 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.

### 2.3.3. The proposed algorithm

Once obtaining the estimated lowest threshold $\theta_{\mathrm{low}}$ and the correlation matrix **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 $t_{\mathrm{maxout}}$, $t_{\mathrm{maxin}}$; inner-loop iteration count $t$=1; outer-loop iteration count $t_{out}$=1; $\mathbf{w}_1=[0,0,\ldots,0]^{\mathrm{T}}$; $\mathbf{r}=\mathbf{y}$, $\rho(0<\rho<1)$; $\theta_1=[\theta(1),\theta(2),\ldots\theta(3L+1)]$; where $\theta(i)$ is the maximal amplitude of different subbands of $(\mathbf{F}_u\boldsymbol{\Psi})^H\mathbf{r}$ and $L$ is the WT decomposition levels;
(2) **Outer-loop: While** $\theta(1),\theta(2),\ldots,\theta(3L+1)>\theta_{\mathrm{low}}$ and $t_{\mathrm{out}}<t_{\mathrm{maxout}}$
(3) estimate $\theta_{\mathrm{low}}$ according to Eq. (10);
(4) **Inner-loop: While** $t<t_{\mathrm{maxin}}$
(5) $\mathbf{w}_{t+1}=S_{\theta_{t_{\mathrm{out}}}}\left(\mathbf{w}_t+\frac{1}{c}(\mathbf{F}_u\boldsymbol{\Psi})^H\mathbf{r}\right)$;
(6) update **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) $\theta_{t+1}=\max(\rho\theta_t,\theta_{\mathrm{low}})$;
(14) $t_{\mathrm{out}}=t_{\mathrm{out}}+1$;
(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 $\sigma_2$ 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.

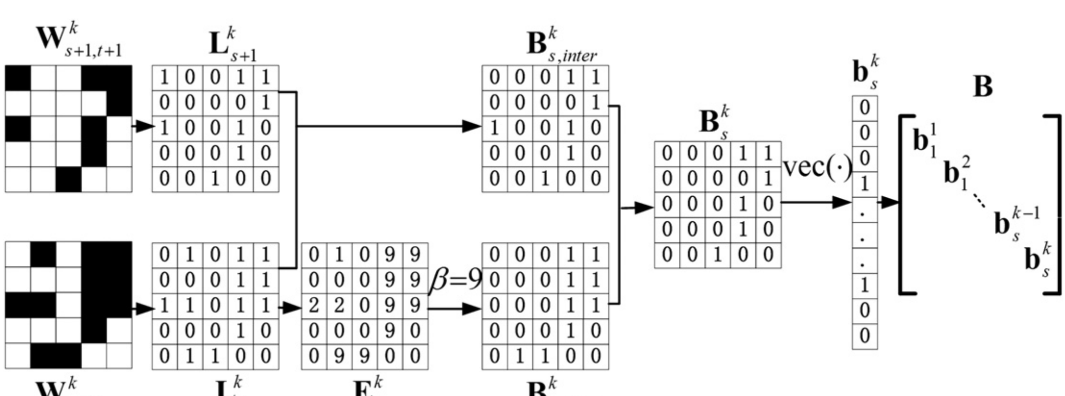


Fig. 2. The design of correlation matrix **B**.

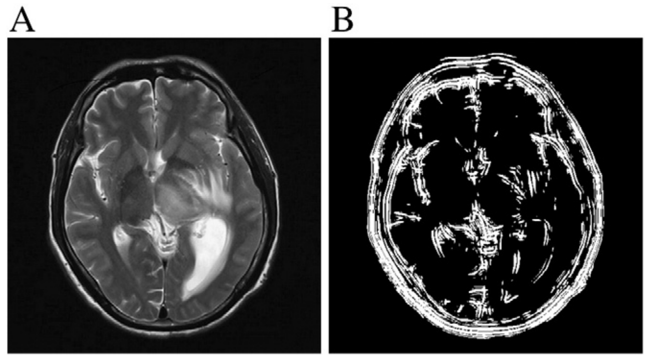


Fig. 3. (A) Original image. (B) The guiding maps for the third scale of à trous WT. The decomposition level is 4.

In Ref. [31], the noise variance $\sigma^2$ was only calculated at the finest wavelet scale. Similarly, we estimate $\theta_{\text{low}}$ only from the diagonal subband of the finest scale to reduce the time for coefficient 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 Refs. [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 a 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 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 of 0.02 is added to both the real and imaginary parts of *k*-space data, respectively.

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 ρ=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.

### 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 $\text{SNR}=10\times\log_{10}(\|\mathbf{x}-\bar{\mathbf{x}}\|_2^2/\|\mathbf{x}-\mathbf{x}_{\text{rec}}\|_2^2)$, where $\mathbf{x}$ is the fully sampled MR image, $\bar{\mathbf{x}}$ is the mean value of $\mathbf{x}$ and $\mathbf{x}_{\text{rec}}$ is the reconstruction result. The curves of SNR vs. noise variance are given in Fig. 7 for the same MR image in Fig. 5. It indicates that ECIA yields an SNR with 2–6 dB higher than those of NLCG, IST, FISTA and IT-EDTC.

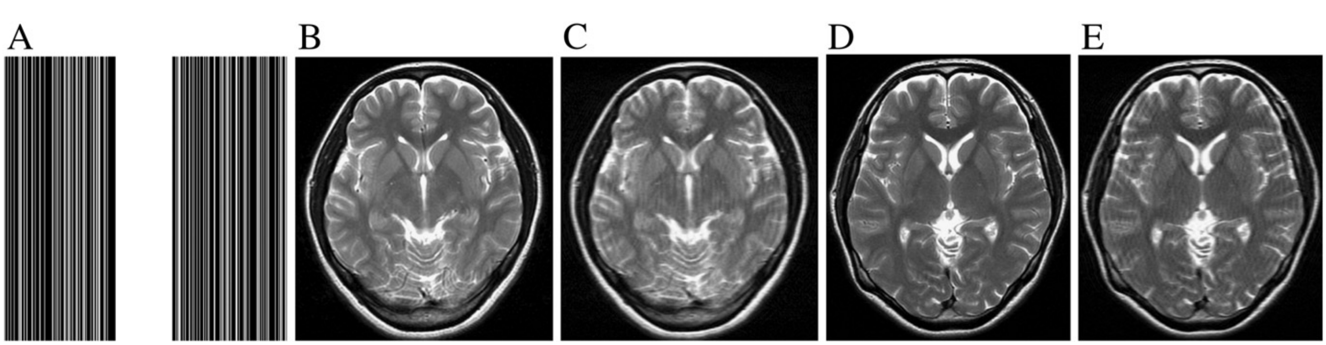


Fig. 4. (A) Variable density Cartesian sampling pattern with 0.4 sampling rate. (B) and (D) are fully sampled MR images. (C) and (E) are the zero-filling undersampled MR images of (B) and (D) using the sampling pattern in (A).

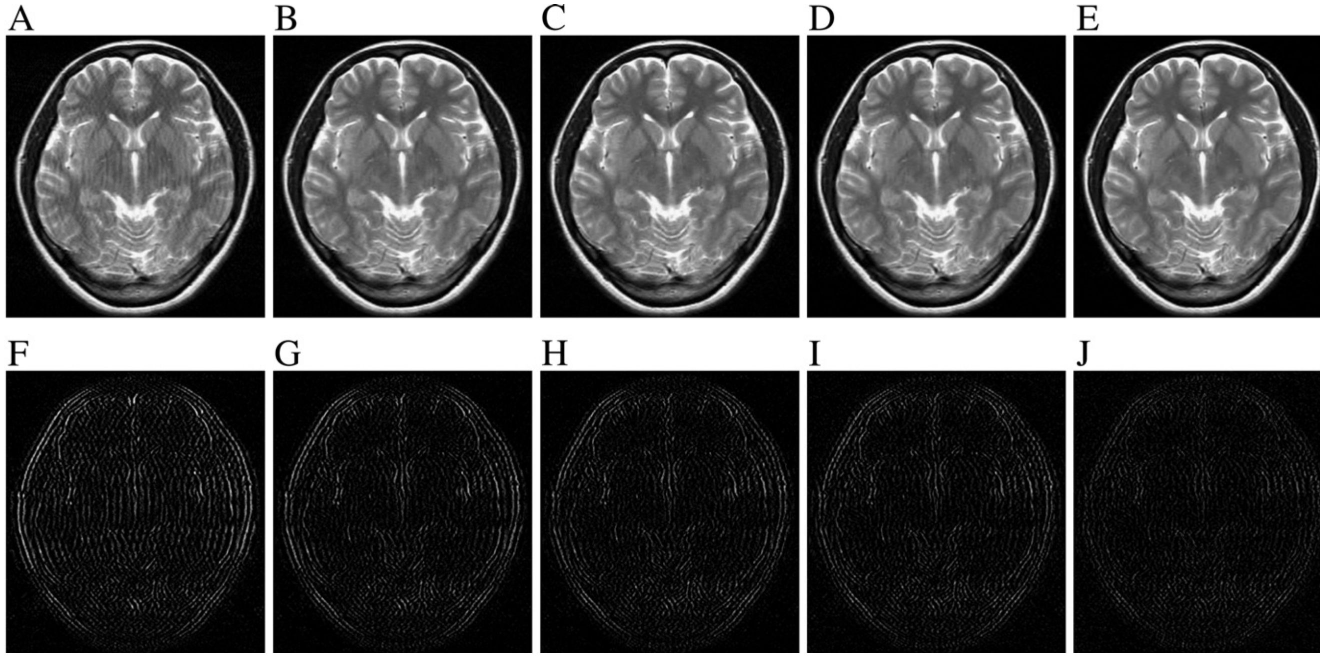


Fig. 5. (A)–(E) are the reconstructions of Fig. 4B by NLCG, IST, FISTA, IT-EDTC and ECIA, respectively. (F)–(J) are the difference images between fully sampled MR image and (A)–(E) with the gray scale of [0,50].

To show the value of $\theta_{low}$ under different noise levels, the curve of $\theta_{low}$ vs. noise variance is presented in Fig. 8A. It indicates that $\theta_{low}$ 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 $\theta_{low}$ will decrease the introduction of significant noise. To show the accuracy of $\theta_{low}$ estimation for noise suppression, Fig. 8B and C gives the curves of SNR between the reconstruction results of ECIA and fully sampled MR image under different $\theta_{low}$. In Fig. 8B and C, Gaussian noise with variances of 0.02 and 0.05 is added to *k*-space, respectively, and we find that the highest SNRs are both

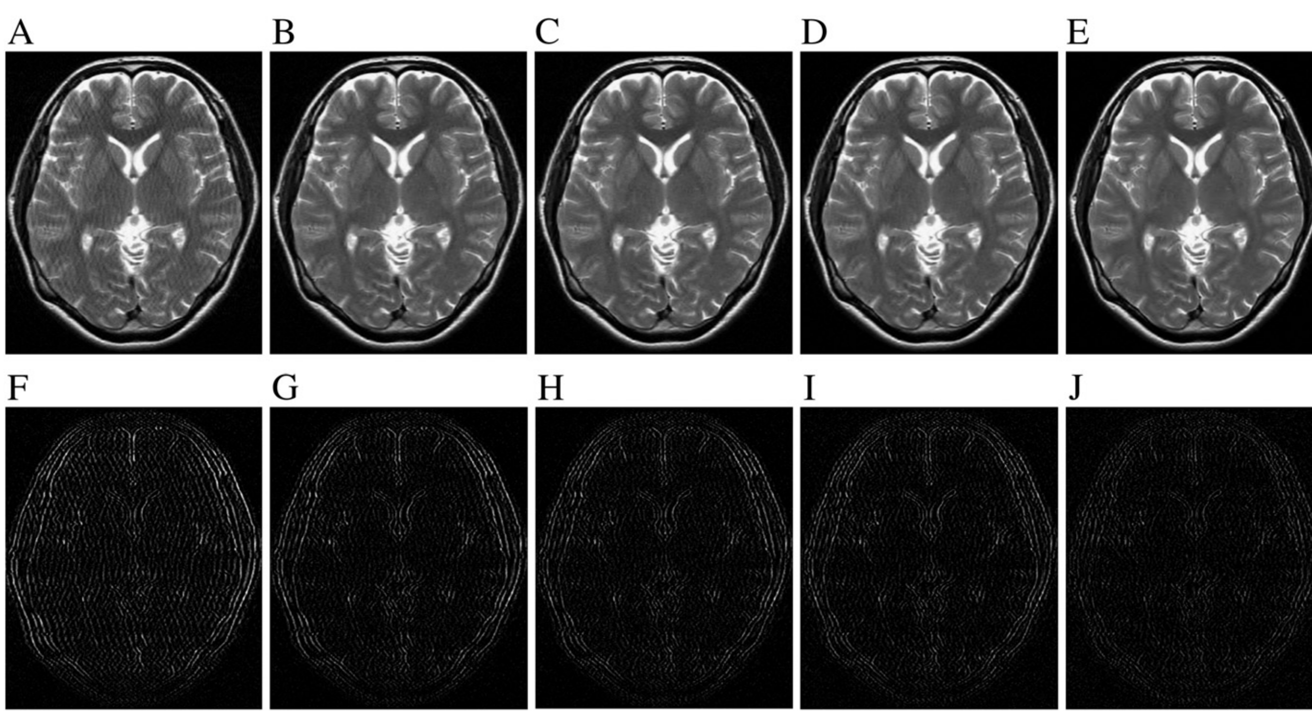


Fig. 6. (A)–(E) are the reconstructions of Fig. 4D by NLCG, IST, FISTA, IT-EDTC and ECIA, respectively. (F)–(J) are the difference images between fully sampled MR image and (A)–(E) with the gray scale of [0,50].

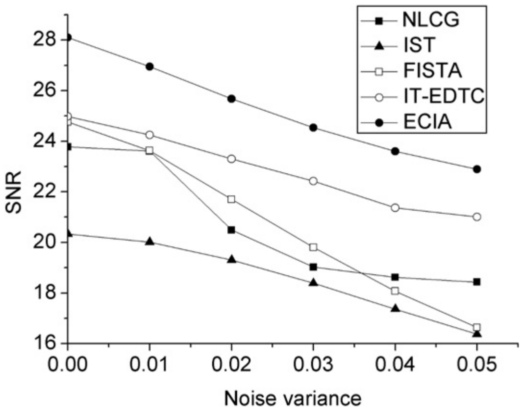


Fig. 7. Curves of SNR vs. noise variance of different algorithms.

achieved near the corresponding $\theta_{\text{low}}$ estimated under the two different noise levels in Fig. 8A.

### 3.3. Reconstruction time comparison

In this section, we report the results of experiments aiming at comparing the speed of ECIA with other algorithms. Fig. 4B and D is used for the 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 a Windows XP operating system. All the algorithms are carried out until their SNRs stabilize. Table 1 gives the average CPU time of five instances for each experiment.

We can observe that the computations of IT-EDTC and ECIA are fast, while that of IST takes more time than other algorithms. Due to the estimation of $\theta_{\text{low}}$ and to the search for 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- to 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) vs. 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 $\theta_{\text{low}}$ is unknown beforehand, we first run the algorithm and record the final estimated $\theta_{\text{low}}$; the curve is then obtained with the recorded $\theta_{\text{low}}$ 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.

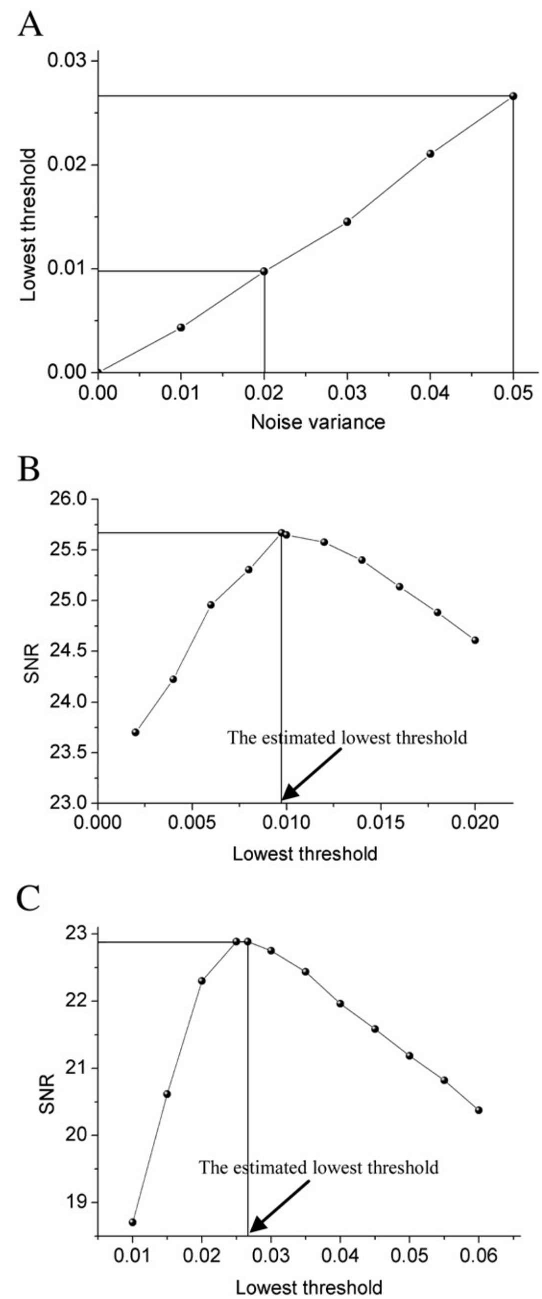


Fig. 8. (A) The curve of estimated lowest threshold vs. noise variance. (B) The curve of SNR vs. lowest threshold; Gaussian noise with variance of 0.02 is added. (C) The curve of SNR vs. lowest threshold; Gaussian noise with variance of 0.05 is added.

## 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

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 |

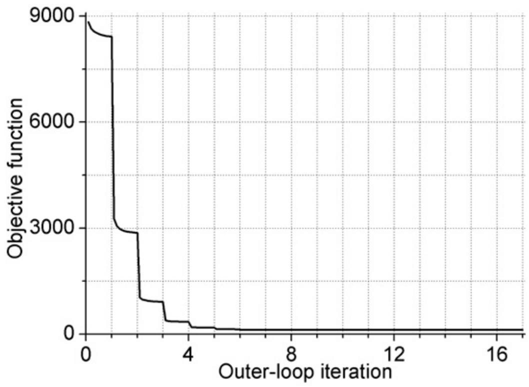


Fig. 9. The evolution of objective function vs. outer-loop iteration number.

MR images with better noise suppression and edge recovery compared with NLCG, IST, FISTA and IT-EDTC algorithms; a 2- to 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 intrascale 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 to be integrated with nonconvex optimization, e.g., replacing the $\ell_1$ norm with the $\ell_p$ ($0<p<1$) norm or smoothed $\ell_0$ norm.

## References


[1] Redpath TW. Signal-to-noise ratio in MRI. Br J Radiol 1998;71: 704–7.

[2] Liu YX, Teverovskiy L, Carmichael O, Kikinis R, Shenton M, Carter CS, et al. Discriminative MR image feature analysis for automatic schizophrenia and Alzheimer's disease classification. In: Barillot C, Haynor DR, Hellier P, editors. Proceedings of Medical Image Computing and Computer-Assisted Intervention-MICCAI 2004, Saint-Malo, France; 2004. p. 393–401.

[3] Kato H, Kanematsu M, Zhang XJ, Saio M, Kondo H, Goshima S, et al. 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;189:117–22.

[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;191:1517–22.

[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: 1289–306.

[7] Natarajan BK. Sparse approximate solutions to linear systems. SIAM J Comput 1995;24:227–34.

[8] Chartrand R. Exact reconstruction of sparse signals via nonconvex minimization. IEEE Signal Proc Let 2007;14:707–10.

[9] Majumdar A, Ward RK. An algorithm for sparse MRI reconstruction by Schatten $p$-norm minimization. Magn Reson Imaging 2011;29:408–17.

[10] Qu XB, Cao X, Guo D, Hu CW, Chen Z. Compressed sensing MRI with combined sparsifying transforms and smoothed $\ell_0$ norm minimization. Proceedings of the 35th International Conference on Acoustics, Speech and Signal Processing-ICASSP'10, Dallas, TX, USA: IEEE; 2010. p. 626–9.

[11] Mohimani H, Babaie-Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed $\ell_0$ norm. IEEE T Signal Proces 2009;57:289–301.

[12] Trzasko J, Manduca A. Highly undersampled magnetic resonance image reconstruction via homotopic $\ell_0$-minimization. IEEE T Med Imaging 2009;28:106–21.

[13] Lustig M, Donoho DL, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn Reson Med 2007;58: 1182–95.

[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;28:637–45.

[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;18:737–58.

[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;28:1395–402.

[17] Qu XB, Cao X, Guo D, Hu CW, Chen Z. Combined sparsifying transforms for compressed sensing MRI. Electron Lett 2010;46:121–3.

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

[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;3:747–58.

[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;57:1413–57.

[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. Proceedings of the 31st International Conference on Acoustics, Speech and Signal Processing-ICASSP'06, Toulouse, France: IEEE; 2006. p. 624–7.

[25] Bredies K, Lorenz DA. Linear convergence of iterative soft-thresholding. J Fourier Anal Appl 2008;14:813–37.

[26] Drori I. Fast $\ell_1$ minimization by iterative thresholding for multidimensional NMR spectroscopy. EURASIP Journal on Advances in Signal Processing Volume 2007, Article ID 20248.

[27] Figueiredo M, Nowak RD, Wright SJ. Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. IEEE J–STSP 2007;4:586–97.

[28] Hale ET, Yin W, Zhang Y. Fixed-point continuation for $\ell_1$-minimization: methodology and convergence. SIAM J Optimiz 2008;19:1107–30.

[29] Wright SJ, Nowak RD, Figueiredo M. Sparse reconstruction by separable approximation. IEEE T Signal Proces 2009;57:2479–93.

[30] Hansen PC, Oleary DP. The use of the L-curve in the regularization of discrete ill-posed problems. SIAM J Sci Comput 1993;14:1487–503.

[31] Sendur L, Selesnick IW. Bivariate shrinkage with local variance estimation. IEEE Signal Proc Let 2002;9:438–41.

[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;56:1134–42.

[33] Kashyap S, Yang ZL, Jacob M. Non-iterative regularized reconstruction algorithm for non-Cartesian MRI: NIRVANA. Magn Reson Imaging 2011;29:222–9.