Iterative thresholding compressed sensing MRI based on contourlet transform

Xiaobo Qu , Weiru Zhang , Di Guo , Congbo Cai , Shuhui Cai & Zhong Chen

Published online: 09 Aug 2010.

Submit your article to this journal ⎘

View related articles ⎘

Taylor & Francis
Taylor & Francis Group

Iterative thresholding compressed sensing MRI based on contourlet transform

Xiaobo Qu, Weiru Zhang, Di Guo, Congbo Cai, Shuhui Cai and Zhong Chen*

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

Reducing the acquisition time is important for clinical magnetic resonance imaging (MRI). Compressed sensing has recently emerged as a theoretical foundation for the reconstruction of magnetic resonance images from under-sampled k-space measurements, assuming those images are sparse in a certain transform domain. However, most real-world signals are compressible rather than exactly sparse. For example, the commonly used two-dimensional wavelet for compressed sensing MRI (CS-MRI) does not sparsely represent curves and edges. In this article, we introduce a geometric image transform, the contourlet, to overcome this shortage. In addition, the improved redundancy provided by the contourlet can successfully suppress the pseudo-Gibbs phenomenon, a tiresome artefact produced by undersampling of k-space, around the singularities of images. For numerical calculation, a simple but effective iterative thresholding algorithm is employed to solve l_1 norm optimization for CS-MRI. Considering the recovered information and image features, we introduce three objective criteria, which are the peak signal-to-noise ratio (PSNR), mutual information and transferred edge information, to evaluate the performance of different image transforms. Simulation results demonstrate that contourlet-based CS-MRI can better reconstruct the curves and edges than traditional wavelet-based methods, especially at low k-space sampling rate.

Keywords: compressed sensing; MRI; data acquisition; geometric image transform; sparsity

AMS Subject Classifications: 92C55; 41A27; 41A29; 42C40; 65F10

1. Introduction

Undersampling the k-space is a good way to speed up magnetic resonance imaging (MRI). However, undersampling often violates the Nyquist sampling rule, resulting in artefacts in reconstructed magnetic resonance (MR) images. Consequently, improving the resolution and suppressing artefacts are interesting and significant topics in undersampled MRI.

Up to now, undersampled MRI methods can be roughly classified into two categories, which exploit the redundancy in either the k-space (e.g. using partial-Fourier [1] or

---

*Corresponding author. Email: chenz@xmu.edu.cn

following non-Cartesian trajectories to sample k-space [2–4]) or in the time domain. The latter is mainly focussed on three-dimensional MRI movies. These two types of methods are often used jointly, as the case of k-t sense [5,6].

MRI maps the spatial information, e.g. spin density, into the called k-space measurements with a finite Fourier integral, and the image is most commonly reconstructed by applying the inverse Fourier transform on these measurements. So, MR image reconstruction is linear because of the linearity of the Fourier transform. However, the reconstruction is considered to be an ill-posed problem because one can not get a unique solution owing to finite sampling on a practical MRI scanner. When undersampling is applied, more than one solution exists which fits the limited k-space measurements. This also enhances the uncertainty of finding the exact image. However, an appropriate image can be obtained with a known deviation from the true image function [7]. A traditional and fruitful way to reconstruct undersampled MR images is to constrain reconstruction by including various prior information of images. A comprehensive review can be found in reference [8].

The emergence of compressed sensing (CS) theory [9,10] excites great interest in signal processing, optimization, inverse problems and medical imaging. The theory provides a foundation to reconstruct signals from measurements fewer than what is specified by the Nyquist sampling rule with little or no information loss. Therefore, it is natural to introduce CS into undersampled MRI. Considering the specific requirements of MRI, CS has been applied to MRI with impressive reconstruction results in a few specific areas, e.g. two-dimensional (2D) MRI [11–13], dynamic MRI [14] and so on [15,16]. In this article, we call these sampling and reconstruction methods compressed sensing-MRI (CS-MRI).

According to the CS theory, two key points need further investigation for undersampled MRI with a high quality of reconstructed images and much fewer k-space measurements, which are:

(1) Coherence. CS requires the sampling matrix $\mathbf{\Phi}$ to be incoherent with the basis dictionary $\mathbf{\Psi}$, which sparsely represents the MR image. The lower the coherence, the less k-space measurements one needs. For MRI, $\mathbf{\Phi}$ is directly determined by the encoding scheme. For example, non-Fourier encoding is applied to reduce the coherence [15]. Parallel MRI is also evidenced with lower coherence than traditional imaging schemes [16]. However, reducing coherence in traditional Fourier encoding is still an open problem.
(2) Sparsity. The number of measurements required for exactly recovering signals is proportional to K, the number of non-zero entries for a signal $\mathbf{x} \in \mathbb{C}^N$, with respect to $\mathbf{\Psi}$. The required k-space measurements will be very few if $\mathbf{\Psi}$ can represent MR images sparsely. In CS-MRI, total variation (TV) [3,4,13] and wavelets [11,17] are commonly used as sparse transforms. TV penalizes local variation in the reconstructed image, while the wavelet transform enforces point singularities and isotropic features. Both of them can greatly suppress the artefacts generated by undersampling the k-space.

Most real-world signals are not exactly sparse in any orthogonal basis, however, most common signals are compressible [18,19]. For example, the MR images of brains often contain curves and edges. TV and wavelet may fail in recovering some of these features. TV penalizes oscillation of texture, thus resulting in loss of texture [17,20]. 2D wavelet fails in recovering these features from a small number of measurements due to the shorthand in

representing curves and edges [21,22]. Therefore, it is desirable to employ more effective sparse transform considering specific image features for CS-MRI.

In recent years, some geometric image transforms have been proposed for sparse image representation. In this article, focussing on recovering the curves and edges, we introduce one of these transforms, namely the contourlet transform [22], into CS-MRI. To further improve the sparsity, we increase the redundancy of the contourlet transform. One reason that we turn to redundant representation is the desire to have the shift variance property [23]. This property can overcome the pseudo-Gibbs phenomenon around singularities, e.g. blurring in edges, thus improving the image quality of undersampled MR images [24]. For numerical computation of reconstruction, a simple but effective iterative thresholding algorithm, which has not been used in CS-MRI before, is employed in this article. We analyse the convergence of iterative thresholding by introducing objective criteria to evaluate the performance of CS reconstruction.

In the rest of this article, a brief introduction of CS-MRI is given in Section 2. The theoretical foundation of contourlet-based CS-MRI and the iterative thresholding algorithm is presented in Section 3. In Section 4, objective criteria and simulation results are discussed. Finally, the conclusion is given in Section 5.

2. Compressed sensing for MRI

The CS proposed by Candès *et al.* [9] and Donoho [10] is a new sampling and compression theory. Its main idea is that a signal can be exactly or approximately recovered from highly incomplete measurements if the signal is sparse in a certain domain.

Suppose that signal $\mathbf{x} \in \mathbb{C}^N$ is sampled by a sensing matrix $\mathbf{\Phi}_{M \times N}$, then the measurements $\mathbf{y} \in \mathbb{C}^M$ of $\mathbf{x}$ are

$$\mathbf{y}_{M \times 1} = \mathbf{\Phi}_{M \times N} \mathbf{x}_{N \times 1}.$$

If $M < N$, it seems hard to recover $\mathbf{x}$ by solving the underdetermined system of equations. According to the CS theory, under the assumption that $\mathbf{x}$ can be sparsely represented in transform $\mathbf{\Psi}$ domain and $\boldsymbol{\alpha}$ is the coefficient with respect to $\mathbf{\Psi}$, $\mathbf{x}$ can be presented as

$$\mathbf{x} = \mathbf{\Psi}\boldsymbol{\alpha}.$$

The sparsity is often expressed as

$$\|\boldsymbol{\alpha}\|_0 = \|\mathbf{\Psi}^*\mathbf{x}\|_0 \ll N$$

where $\|\boldsymbol{\alpha}\|_0$ denotes the l_0 quasi-norm and is defined as $\|\boldsymbol{\alpha}\|_0 = \#\{\alpha_i \neq 0, i = 1, 2, \dots, I\}$ which counts the nonzero entries in $\boldsymbol{\alpha}$ with the length I. $\mathbf{\Psi}$ and $\mathbf{\Psi}^*$ mean the inverse and forward sparse transforms, respectively.

CS tries to reconstruct the signal from undersampled measurements by minimizing l_0 norm optimization. Letting $\hat{\boldsymbol{\alpha}}$ denote the estimation of $\boldsymbol{\alpha}$, the l_0 norm optimization is

$$\hat{\boldsymbol{\alpha}} = \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0, \quad \text{s.t.} \quad \mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}\boldsymbol{\alpha}.$$

However, the l_0 norm is known to be intractable and sensitive to noise [9,10], so l_1 norm convex optimization is commonly used in CS to recover $\mathbf{x}$

$$\hat{\boldsymbol{\alpha}} = \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1, \quad \text{s.t.} \quad \mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}\boldsymbol{\alpha} \tag{1}$$

where $\|\boldsymbol{\alpha}\|_1$ denotes the l_1 norm, i.e. the sum of absolute values of vector elements $\alpha_i (i = 1, 2, \ldots, I)$, and is defined as $\|\boldsymbol{\alpha}\|_1 = \sum_{i=1}^{I} |\alpha_i|$. The recovered signal $\hat{\mathbf{x}}$ is equal to $\boldsymbol{\Psi}\boldsymbol{\alpha}$.

For MRI, continuous spatial information $x(\mathbf{r})$, e.g. spin density, is often represented as the sum of N voxels in a discretized form [7,15]

$$x(\mathbf{r}) = \sum_{n=1}^{N} x_n \phi(\mathbf{r} - \mathbf{r}_n)$$

where $\phi(\mathbf{r})$ is the voxel basis function (typical choices include Dirac and box functions) and the $\mathbf{r}_n$ vectors specify the voxel grid. Therefore, the vector form of $x(\mathbf{r})$ can be expressed as

$$\mathbf{x}_{N \times 1} = [x_1, x_2, \ldots x_n \ldots, x_N]^{\mathrm{T}}.$$

Then the spatial information is then mapped into the so-called k-space measurements $y(\mathbf{k})$ with a finite Fourier integral as

$$y(\mathbf{k}_m) = \int_{\mathbb{R}^2} x(\mathbf{r}) e^{-i2\pi \mathbf{k}_m \mathbf{r}} \mathrm{d}\mathbf{r} = \int_{\mathbb{R}^2} \sum_{n=1}^{N} x_n \phi(\mathbf{r} - \mathbf{r}_n) e^{-i2\pi \mathbf{k}_m \mathbf{r}} \mathrm{d}\mathbf{r}, \quad m = 1, 2, \ldots, M \qquad (2)$$

where the integration is performed on the plane $\mathbb{R}^2$ for the 2D MR image and $y(\mathbf{k}_m)$ is the value at the mth k-space location $\mathbf{k}_m$. Thus, $y(\mathbf{k})$ can be expressed as the vector as

$$\mathbf{y}_{M \times 1} = [y(\mathbf{k}_1), y(\mathbf{k}_2), \ldots y(\mathbf{k}_m) \ldots, y(\mathbf{k}_M)]^{\mathrm{T}}.$$

If the $M \times N$ encoding matrix $\boldsymbol{\Phi}$ is defined as

$$\boldsymbol{\Phi}_{m,n} = \int \phi(\mathbf{r} - \mathbf{r}_n) e^{-i2\pi \mathbf{k}_m \mathbf{r}} \mathrm{d}\mathbf{r}$$

then Equation (2) can be written in matrix form as

$$\mathbf{y}_{M \times 1} = \boldsymbol{\Phi}_{M \times N} \mathbf{x}_{N \times 1} \qquad (3)$$

where $\mathbf{y} = [y_1, y_2, \ldots, y_M]^{\mathrm{T}}$ is the data vector $y \in \mathbb{C}^M$ and $\mathbf{x} = [x_1, x_2, \ldots, x_N]^{\mathrm{T}}$ is the vector of voxel coefficients $x \in \mathbb{C}^N$.

For CS-MRI, the encoding matrix $\boldsymbol{\Phi}$ is denoted as the undersampled Fourier transform operator $\mathbf{F}_{\mathrm{u}}$ to avoid the ambiguity expression with other applications of CS. $\mathbf{F}_{\mathrm{u}}$ directly relies on the sampling scheme. $M < N$ means that the acquired k-space samples is incomplete, therefore there exists more than one solution to fit for Equation (3).

According to Equation (1), the reconstructed MR image is estimated by solving the constrained optimization problem under the assumption that $\mathbf{x}$ can be sparsely represented in transform $\boldsymbol{\Psi}$ domain,

$$\hat{\boldsymbol{\alpha}} = \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1, \quad \text{s.t.} \quad \mathbf{y} = \mathbf{F}_{\mathrm{u}} \boldsymbol{\Psi} \boldsymbol{\alpha}. \qquad (4)$$

The reconstructed MR image is $\hat{\mathbf{x}} = \boldsymbol{\Psi}\hat{\boldsymbol{\alpha}}$. Since $\mathbf{y}$ is a complex vector, $\hat{\boldsymbol{\alpha}}$ and $\hat{\mathbf{x}}$ are complex too. However, since $\mathbf{y}$ is often contaminated by noise, one needs to estimate the minimal $\|\boldsymbol{\alpha}\|_1$ by solving the equation

$$\hat{\boldsymbol{\alpha}} = \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1, \quad \text{s.t.} \quad \left\| \mathbf{y} - \mathbf{F}_{\mathrm{u}} \boldsymbol{\Psi} \boldsymbol{\alpha} \right\|_2 < \varepsilon. \qquad (5)$$

The notation $\|\bullet\|_2$ stands for l_2 norm defined as $\|\mathbf{m}\|_2 = \left(\sum_{i=1}^{I} |m_i|^2\right)^{\frac{1}{2}}$ where I is the total number of elements of $\mathbf{m}$. ε controls the fidelity of the reconstruction to the measured k-space samples. The solution $\hat{\boldsymbol{\alpha}}$ has a small l_1 norm to enforce the sparsity and fits the data up to a tolerance ε.

Daubechies *et al.* [25] proved that generalizing Tikhonov's regularization method from the l_2 norm penalty case to the l_1 norm penalty provides a proper regularization method for an ill-posed problem $\mathbf{y} = \boldsymbol{\Phi}\mathbf{x}$ as

$$\hat{\boldsymbol{\alpha}} = \min_{\boldsymbol{\alpha}} \|\mathbf{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha}\|_2^2 + \sum_{i=1}^{I} \omega_i |\alpha_i|^p \tag{6}$$

where α_i is the ith element of vector $\boldsymbol{\alpha}$ with length I in orthonormal basis $\boldsymbol{\Psi}$, e.g. wavelet basis. ω_i is the weight for α_i. The notation $|\alpha_i|^p$ stands for pth power of the absolute value of α_i.

For the special case when $p = 2$ and ω_i is some positive constant $\omega_i = \beta$, Equation (6) reduces to the standard Tikhonov regularization equation

$$\hat{\boldsymbol{\alpha}} = \min_{\boldsymbol{\alpha}} \|\mathbf{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha}\|_2^2 + \beta \sum_{i=1}^{I} |\alpha_i|^2 \tag{7}$$

where β gives the trade off between the data fidelity and l_2 norm of vector $\boldsymbol{\alpha}$.

Compared with the standard Tikhonov regularization in Equation (7), the generalized Tikhonov regularization in Equation (6) puts a lesser penalty on only a few large entries of $\boldsymbol{\alpha}$, and a higher penalty on many small entries of $\boldsymbol{\alpha}$. Thus, the second term in Equation (6) promotes sparsity of the expansion of $\mathbf{x} = \boldsymbol{\Psi}\boldsymbol{\alpha}$ with respect to the orthonormal basis $\boldsymbol{\Psi}$. For the special case $p = 1$, the generalized Tikhonov regularization becomes the unconstrained form of l_1 norm optimization in CS [11].

3. Iterative thresholding CS-MRI based on the contourlet transform

3.1. *Compressible CS-MRI*

CS works well when the signal is exactly sparse. Suppose an ideal MR image (or MR image reconstructed from fully sampled k-space) is $\tilde{x} \in \mathbb{C}^N$, CS could work efficiently if the MR image is sparse with respect to the forward transform $\boldsymbol{\Psi}^*$, which means

$$\|\boldsymbol{\Psi}^*\tilde{\mathbf{x}}\|_0 \ll N.$$

However, most real-world signals are not exactly sparse in any orthogonal basis but *compressible* [18,19]. As a widely used image sparsifying transform, the wavelet transform has great success in representing point singularities, including for CS-MRI [11,18]. Unfortunately, the traditional 2D wavelet is a tensor-product of the one-dimensional (1D) wavelet. It is a separable extension from the 1D basis. Thus, it is not good at representing curves [21,22]. The representation of curves by the 2D wavelet is just compressible, i.e. satisfying the power law decay if the image is constructed from the k largest magnitude coefficients in the wavelet domain.

The contourlet transform was recently pioneered by Do and Vetterli [22]. Compared with the wavelet transform, it provides a flexible rather than fixed number of directions at each scale and thus can capture the intrinsic geometrical structure of images. Unlike the
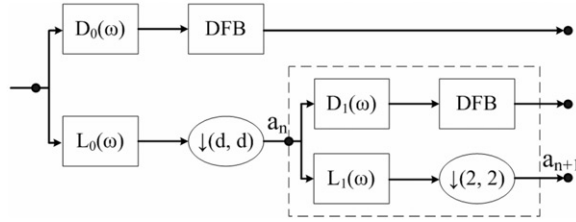
Figure 1. Decomposition structure of SFLCT.
Notes: DFB stands for directional filter banks, $D_i(\omega)$ and $L_i(\omega)$ are the high pass and low pass filter at the ith decomposition level, The DFB is short for directional filter banks, $\downarrow(d,d)$ is the down sampling factor for lowpass filter. a_n is the nth approximation subbands. Refer to reference [25] for the definitions of $D_i(\omega)$, $L_i(\omega)$ and a_n.

curvelet [21], which is another geometric image transform initially developed in continuous domain and then discretized for sampled data, the contourlet is constructed directly in a discrete domain and enjoys low computing complexity. It requires $O(N)$ operations for N-pixel images [22]. Thus, the contourlet can be easily implemented for MR images. Furthermore, the iterated filter banks algorithm in the contourlet transform allows one to choose different filters to represent the features of MR image. We adopt a sharp frequency localization contourlet (SFLCT) [26] that outperforms the original contourlet [22]. The new contourlet successfully cancels the aliasing components outside the desired trapezoid-shaped support. As is shown in Figure 1, it employs a new multiscale pyramid with different sets of lowpass and highpass filters for different levels.

In contrast with the nonsubsampled contourlet [27] to obtain fully shift-invariant property at the price of large time consumption, SFLCT is a semi-redundant contourlet transform. It only increases the redundancy in the lowpass filter because the pseudo-Gibbs phenomenon of original contourlet [22] is mainly induced by downsampling in the lowpass filter $L_0(\omega)$ [26,27]. Lu *et al.* [26] designed the lowpass filter $L_0(\omega)$ to cancel aliasing component in images. The 1D forms of $L_0(\omega)$ and $L_1(\omega)$ is

$$L_i(\omega) = \begin{cases} 1 & \text{for} \quad |w| \leq w_{p,i} \\ \frac{1}{2} + \frac{1}{2}\cos\frac{(|w|-w_{p,i})\pi}{w_{s,i}-w_{p,i}} & \text{for} \quad w_{p,i} \leq |w| \leq w_{s,i} \\ 0 & \text{for} \quad w_{s,i} \leq |w| \leq \pi \end{cases}$$

where $|\omega| \leq \pi$, $i = 0, 1$, $w_{p,i}$ is the passband frequency and $w_{s,i}$ is the stopband frequency.

In addition, the iterative decomposition structure of SFLCT, shown in Figure 1, provides an opportunity to easily handle the redundancy of the contourlet. For example, if we do not downsample the lowpass filter $L_0(\omega)$, the redundancy of contourlet is 2.33. If we downsample the lowpass filter $L_0(\omega)$ by setting $\downarrow(d,d) = \downarrow(2,2)$ where d is the down-sampling parameter that determines the redundancy of contourlet, the redundancy of the contourlet is 1.33 [26]. Consequently, SFLCT is a good choice for practical CS-MRI to easily control the redundancy and reduce the computing complexity. For simplicity, the contourlet mentioned in the rest of this article refer to SFLCT.

Figure 2 shows the basis elements of wavelets and contourlets. The wavelet basis has only three directions while the contourlet basis is anisotropic and has arbitrary directions (commonly $2^m(m \leq 5)$). Equipped with these bases, the contourlet requires fewer coefficients to represent curves of MR images than the wavelet does.
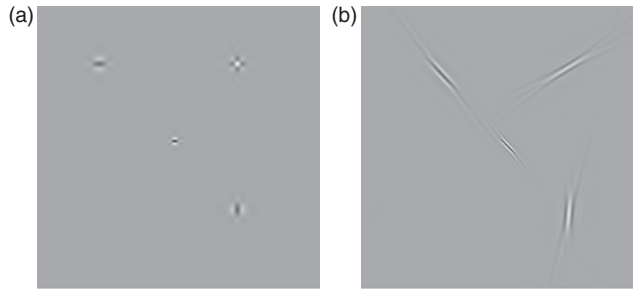
Figure 2. Basis elements of (a) wavelet and (b) contourlet.
Note: Contourlet offers richer set of directions and shapes, thus it is more effective in capturing curves and edges than wavelet.

For a 2D piece-wise smooth image $\mathbf{x}$ with twice continuously differentiable discontinuities (called C^2 singularities) [28], if we keep the k largest magnitude terms in the bases, the optimal approximation rate of the contourlet [22] achieves

$$\left\| \mathbf{x} - \hat{\mathbf{x}}_k \right\|_2^2 \propto O(k^{-2}).$$

For wavelets, the error decay rate [21,22] is

$$\left\| \mathbf{x} - \hat{\mathbf{x}}_k \right\|_2^2 \propto O(k^{-1}).$$

According to the CS theory, if we make the number of measurements $M \geq \mathrm{Const}\, k \log N$ for a compressible signal $x \in \mathbb{C}^N$, the solution $\hat{\mathbf{x}}$ to Equation (4) has the same error decay as $\hat{\mathbf{x}}_k$ [29]

$$\left\| \mathbf{x} - \hat{\mathbf{x}}_M \right\|_2 \propto O(k^{-r}) \tag{8}$$

which means that the reconstruction error is proportional to the power law decay k^{-r}. By using the contourlet as the sparse transform for an object with C^2 singularities, the reconstruction error (k^{-2}) attenuates faster than that of wavelets (k^{-1}) when $k > 1$. It implies that when an image contains more curves, the reconstruction error of the contourlet is lower than that of the wavelet from undersampled k-space. That is why we use the contourlet as the geometric transforms in this article.

Aiming to recover the curves with less error and to control the redundancy of transforms easily, we select the SFLCT as the sparse transform for CS-MRI in this article.

3.2. *Numerical calculation*

Many scientists seek simple and fast algorithms to solve Equation (1), such as the conjugate gradient method [11], Bregman iteration [17], the interior point method [30] and other methods. However, these methods are not straightforward in solving the problem of constraint l_1 norm optimization of CS [11,30]. In this article, we employ iterative thresholding [25,31–36], a simple but effective algorithm, which has never been reported in CS-MRI to solve Equation (5). The iterative thresholding algorithm directly cancels the

interference caused by undersampling the k-space [11], and has been recommended to solve CS problems [34].

The classic interpretations of iterative thresholding for solving constraint l_1 norm optimization were reported previously [25,32]. For theoretic analysis, Herrity *et al.* employed hard iterative thresholding to demonstrate that one could recover the k-term representation of the original signal up to any prescribed error tolerance under certain conditions [33]. Their simulation results showed that soft iterative thresholding performed better than hard thresholding. *Bredies* and *Lorenz* proved that soft iterative thresholding converged with a linear rate once the underlying operator satisfied the finite basis infectivity property or the minimizer possessed a strict sparsity pattern [31]. In addition, only two free parameters, which are independent of transforms, must be set in iterative thresholding algorithm. It allows convenient comparison among different transforms.

Inspired by the work of those authors and the convenience of iterative thresholding, we apply soft iterative thresholding to solve CS-MRI. It is defined as follows:

$$S_\theta(\alpha_i) = \begin{cases} \alpha_i + \theta & \alpha_i \leq -\theta \\ 0 & |\alpha_i| < \theta \\ \alpha_i - \theta & \alpha_i \geq \theta \end{cases} \tag{9}$$

where $\alpha_i (i = 1, 2, \ldots, I)$ is the ith entry in $\boldsymbol{\alpha}$. Since $\boldsymbol{\alpha}$ is a complex vector for MR images, we use the complex thresholding operator defined as $S_\theta(|\alpha_i|e^{jw}) = S_\theta(|\alpha_i|)e^{jw}$. For simplicity, let $S_\theta(\boldsymbol{\alpha})$ denote the soft thresholding operation that performs on all the entries of $\boldsymbol{\alpha}$ with the same threshold θ according to Equation (9).

Different from the sparsity constraint denoising problem presented in [32], we aim to reconstruct MR images from undersampled k-space, so we need to modify the iterative thresholding algorithm. For simplicity, we use $\mathbf{A}$ to stand for $\mathbf{F_u\Psi}$ and $\mathbf{A}^* = \mathbf{\Psi}^*\mathbf{F_u}^*$ denotes the adjoint transpose of $\mathbf{A}$. Because the contourlet is established by frame theory, $\mathbf{\Psi}^*$ means the forward contourlet transform, while $\mathbf{\Psi}$ means the inverse transform for CS [36].

The soft iterative thresholding algorithm for CS-MRI can be implemented according to the following steps:

Step 1 Initialization. The coefficients vector $\boldsymbol{\alpha}_0 = [0, 0, \ldots, 0]^\mathrm{T}$, the reconstructed image $\mathbf{x}_0 = \mathbf{\Psi}\boldsymbol{\alpha}_0$, the residual $\mathbf{r}_0 = \mathbf{y}$, and the initial threshold $\theta_0 = \max(\mathbf{A}^*\mathbf{r}_0)$ where $\mathbf{A}^* = \mathbf{\Psi}^*\mathbf{F_u}^*$ denotes the adjoint transpose of $\mathbf{A}$. The notation $\mathbf{F_u}^*$ means to do inverse Fourier transform on undersampled the Fourier coefficients $\mathbf{r}_0$ while $\mathbf{\Psi}^*$ means the forward contourlet transform.

Step 2 Update the coefficients vector $\boldsymbol{\alpha}_{t+1} = \boldsymbol{\alpha}_t + S_{\theta_t}(\mathbf{A}^*\mathbf{r}_t)$ by applying soft iterative thresholding according to Equation (9), and the residual $\mathbf{r}_{t+1} = \mathbf{y} - \mathbf{A}\boldsymbol{\alpha}_{t+1}$. Decrease the threshold $\theta_{t+1} = \rho\theta_t$, in which ρ is a decrease factor and typically defined as $0 < \rho < 1$.

Step 3 Repeat Step 2 until the residual $\mathbf{r}$ satisfies the given stop criteria. Output the estimated coefficients $\hat{\boldsymbol{\alpha}}$ and the reconstructed image $\hat{\mathbf{x}} = \mathbf{\Psi}\hat{\boldsymbol{\alpha}}$.

Since the acquired k-space measurements are often contaminated by noise, the stop criteria of iterative thresholding for CS-MRI is directly related to ε of Equation (5). In order to eliminate the effect of the magnitude of MR images on stop criteria, we change the traditional absolute stop criteria into a relative form as

$$\eta = \frac{\|\mathbf{A}\hat{\boldsymbol{\alpha}} - \mathbf{y}\|_2}{\|\mathbf{y}\|_2}. \tag{10}$$

Accordingly, relative residual R_t is defined in a relative form as

$$R_t = \frac{\left\| \mathbf{A}\boldsymbol{\alpha}_t - \mathbf{y} \right\|_2}{\left\| \mathbf{y} \right\|_2}. \tag{11}$$

Suppose that at the tth iterative time, the final estimated coefficient $\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha}_t$, if $R_t \leq \eta$.

With the same meaning as ε in CS in Equation (5), the stop criteria η is set as large if the k-space is corrupted by heavy noise, while η is small for light noise. It is worth noting that estimating the noise contained in k-space is very significant. If η is smaller than the true noise level, more noise will be presented in the reconstructed image. On the other hand, if η is larger than the true noise level, image details will be lost since they are taken as noise and thrown away. More details will be discussed in Section 4.2.3.

In the algorithm, ρ is adopted to decrease the threshold θ_t in each iteration, typically defined as $0 < \rho < 1$. The smaller ρ is, the faster α comes to convergence.

The two parameters η and ρ presented in the algorithm are constants, and we set them to be the same in all simulations except when we discuss the convergence of soft iterative thresholding in Section 4.2.3. From empirical analysis, $\eta = 10^{-6}$ and $\rho = 0.8$ assure promising results.

Two factors affecting the computational complexity of the algorithm are iteration times and multiplications involving $\mathbf{A}$ and $\mathbf{A}^*$. For the former, no nested loop is involved in this algorithm, and sufficient tests show that the algorithm approaches the stop criteria after only tens of iterations. For the latter, only one multiplication by $\mathbf{A}$ and one by $\mathbf{A}^*$ are performed per iteration. Simulations in Section 4.2.3 show that different transforms reach the stop criteria after a similar number of iterations, implying that the speeds of different transforms are mainly related to the computational complexity of their forward and inverse transforms.

4. Simulation results and analysis

4.1. *Objective performance evaluation*

Objective evaluation of performance of different methods is an important issue. Focussing on the influence of sparse transforms on the curve features and the information of reconstructed images, besides the visual appearance, we introduce three objective criteria to evaluate the correlation between reconstructed MR images from fully sampled and undersampled k-space. They are the peak signal-to-noise ratio (PSNR), mutual information (MI) [37] and transferred edge information (TEI) [38].

Suppose $\tilde{x}$ is the fully sampled k-space MR image, $\hat{x}$ is the solution of Equation (5) and the gray value of $\hat{x}$ is 0–255.

PSNR is defined as

$$\text{PSNR} = 20 \log_{10} \left(\frac{255}{\sqrt{\text{MSE}}} \right) \tag{12}$$

where $\text{MSE} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (\tilde{x}(i,j) - \hat{x}(i,j))^2$.

MI is defined as

$$\text{MI} = \sum p(\tilde{x}, \hat{x}) \log \frac{p(\tilde{x}, \hat{x})}{p(\tilde{x}) \, p(\hat{x})} \tag{13}$$

where $p(\tilde{x}, \hat{x})$ is the joint probability distribution and $p(\bullet)$ is the marginal probability distribution.

TEI is defined as

$$\text{TEI} = Q_g^{\tilde{x}\hat{x}} Q_\alpha^{\tilde{x}\hat{x}} \qquad (14)$$

where $Q_g^{\tilde{x}\hat{x}}$ and $Q_\alpha^{\tilde{x}\hat{x}}$ [38] stand for the edge strength and orientation preservation values.

PSNR evaluates the difference between the gray values of fully sampled k-space MR image and CS-based reconstructed image. MI essentially computes how much information from fully sampled k-space MR image is transferred to CS-based reconstructed image. TEI measures the amount of edge information that CS reconstructs using a Sobel edge detector which computes the horizontal derivative approximation G_x and vertical derivative approximation G_y with a 2D convolution operation on image f as follows:

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * f \quad \text{and} \quad G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * f$$

4.2. *Simulation results*

In the simulation, we use a binary mask with equal size of k-space to determine which k-space measurements are acquired. If the value of mask at location (x, y) is equal to 1, the measurement at location (x, y) is acquired. The ratio of k-space measurements acquired in undersampling, named as sampling rate, is between 0 and 1. We call the binary mask the sampling pattern because it directly determines the way to sample the k-space.

To demonstrate the performance of the contourlet transform and the efficiency of soft iterative thresholding, first we will show the weakness of the traditional 2D wavelet on representing curves in CS-MRI. This shortcoming can be successfully overcome by contourlets. The pseudo-Gibbs phenomenon around edges can be suppressed as well by improving the redundancy of the contourlet. We will then show the objective criteria at different sampling rates with two sampling patterns. Finally, the convergence of soft iterative thresholding will be presented.

We use the Daubechies wavelet with four vanishing moments and four decomposition levels. SFLCT is with decomposition level [5, 4, 4, 3], which means four decomposition levels and 2^5, 2^4, 2^4, 2^3 directional subbands from coarse to fine scales. One type of quincunx/fan filters named pkva [39], which can effectively localize edge direction [40], is employed as a decomposition filter in SFLCT. The non-redundant contourlet downsamples the first lowpass filter with factor 2, while the redundant contourlet does not downsample first lowpass filter. Both forms of contourlets use critical downsampled directional filter banks.

4.2.1. *Overcoming the shortage of the wavelet transform*

In this subsection, parameters of iterative thresholding are $\varepsilon = 10^{-4}$, $\eta = 10^{-6}$ and $\rho = 0.8$.

Figure 3(a) is the Cartesian sampling pattern for reducing the phase encoding time. The k-space data are acquired at the locations of white pixels as shown in Figure 3(a) and the
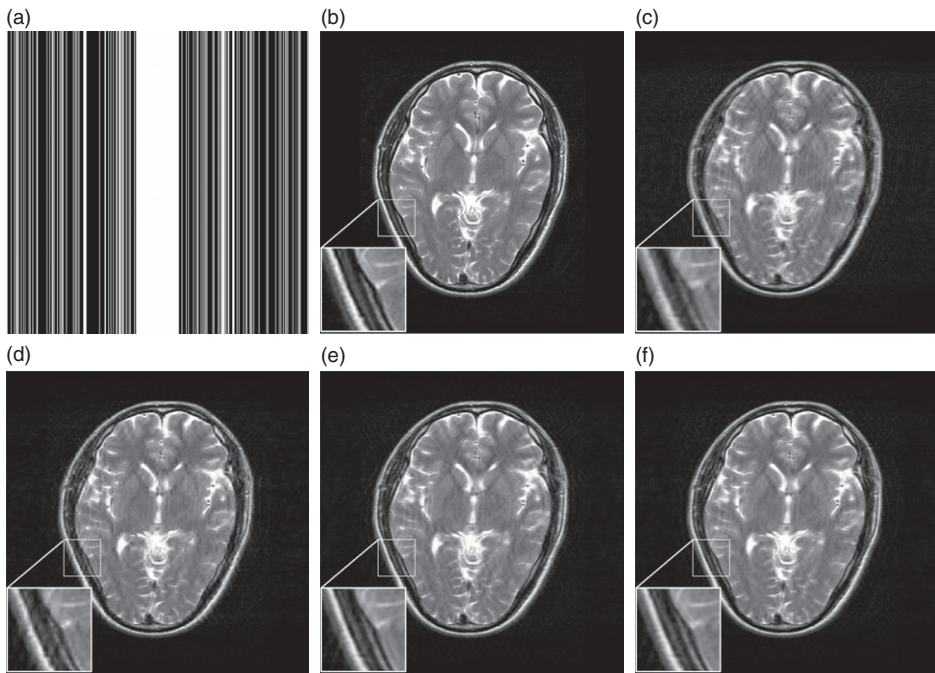
Figure 3. Comparison of wavelet and contourlet reconstruction of undersampled *k*-space using Cartesian sampling pattern.
Notes: (a) Cartesian sampling pattern with sampling rate 0.36, (b)–(f) are reconstructed images from fully sampled *k*-space, zero-filling *k*-space, wavelet-based, non-redundant contourlet-based and redundant contourlet-based CS, respectively.

Table 1. Comparison on objective criteria of wavelet-based and contourlet-based CS-MRI using Cartesian sampling pattern with sampling rate 0.36.

Reconstruction method	PSNR	TEI	MI
Zero-filling	35.2	0.41	1.75
Wavelet	35.8	0.52	1.90
Non-redundant contourlet	36.2	0.58	1.91
Redundant contourlet	36.6	0.62	2.01

other *k*-space data are filled with zeros. The sampling rate is 0.36, which means only 36% of *k*-space is acquired.

Filling zeros into *k*-space produces curve-like artefacts in Figure 3(c). The Cartesian sampling pattern makes the artefacts, commonly evaluated by the point spread function [10,14], spread in the zero-filling image. Figure 3(d)–(f) show that CS can suppress these artefacts and obtain clearer images.

The zoom in parts in Figure 3(d) and (e) show that contourlet outperforms wavelet in reconstructing the curves of MR image. This conclusion is also demonstrated by the objective criteria in Table 1 which implies that the contourlet can obtain a higher TEI than the wavelet.
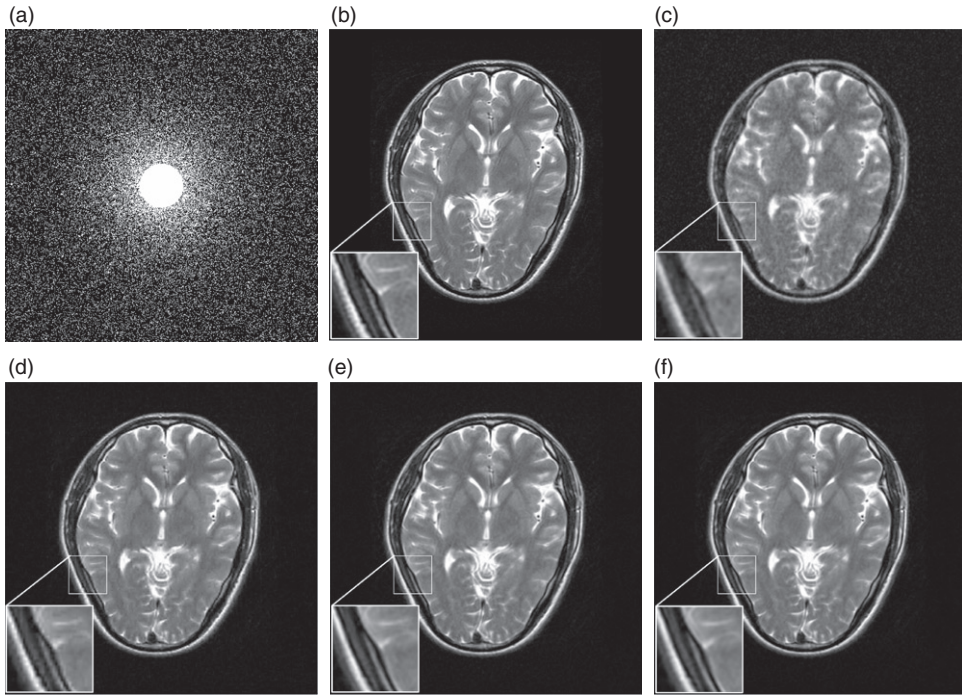
Figure 4. Comparison of wavelet and contourlet reconstruction of undersampled *k*-space using variable density sampling pattern.
Notes: (a) Variable density sampling pattern with sampling rate 0.20, (b)–(f) are reconstructed image from fully sampled *k*-space, zero-filling *k*-space, wavelet-based, non-redundant contourlet-based and redundant contourlet-based CS, respectively.

Just as in image denoising, thresholding coefficients of the image transform would induce a pseudo-Gibbs phenomenon around the edges. Improving the redundancy of the transform is a good way to suppress these effects. Figure 3(e) and (f) indicate that redundant contourlets can suppress the artefacts better than non-redundant contourlets. The objective criteria in Table 1 also demonstrate this conclusion.

The aforementioned example shows that undersampling the *k*-space with a Cartesian sampling pattern will introduce curve-like artefacts. These artefacts tend to be considered as meaningful image features and cannot be sufficiently suppressed by CS. In the view of point spread function [11,15] for CS-MRI, a good sampling pattern should be able to make the artefacts look like random noise. From this aspect, the variable density sampling pattern, shown in Figure 4(a), works well in our empirical simulation. This pattern is also recommended in literature [41]. Figure 4(c) shows variable density sampling pattern makes the artefacts much like the noise. Though the sampling rate is half of that for Cartesian sampling pattern of Figure 3(a), the reconstructed images in Figure 4(d)–(f) are much better than those in Figure 3(d)–(f). The zoom in parts in Figure 4(d) and (e) show that contourlet can

Table 2. Comparison on objective criteria of wavelet-based and contourlet-based CS-MRI using variable density sampling pattern with sampling rate 0.20.

Reconstruction method	PSNR	TEI	MI
Zero-filling	35.4	0.41	1.68
Wavelet	37.1	0.58	1.99
Non-redundant contourlet	38.1	0.64	2.11
Redundant contourlet	38.8	0.67	2.23

better preserve the edges than wavelet. This conclusion is also demonstrated by the objective criteria in Table 2.

4.2.2. *Contourlet-based CS at different sampling rates*

In this subsection, we will show the performance of different transforms at different sampling rates. Parameters of iterative thresholding are $\varepsilon = 10^{-4}$, $\eta = 10^{-6}$ and $\rho = 0.8$.

Figure 5 shows the variations of the objective evaluation criteria of CS-MRI versus sampling rates for the wavelet and the contourlet in reconstructing the MR image in Figure 3(b). The curves in the left column of Figure 5 show the criteria when a Cartesian sampling pattern is applied for undersampling and the curves in the right column of Figure 5 show the criteria when a variable density sampling pattern is applied for undersampling.

All criteria indicate that the contourlet can better preserve edges and recover more information than the wavelet does, especially at low sampling rates. For example, compared with wavelets, the non-redundant contourlet improves the TEI by 20% at the sampling rates of 0.4 and 0.2 with the Cartesian sampling pattern. The advantage of contourlet holds true for both Cartesian and variable density sampling patterns.

Figure 5 also indicates that, with the increase of sampling rate, the superiority of the contourlet becomes less obvious. For example, when the sampling rate of the Cartesian pattern is above 0.8 or that of the variable density sampling pattern is above 0.6, the criteria of contourlet and wavelet are nearly the same. According to the CS theory, if the sampling rate is sufficiently high and the transform can make the image sparse enough, the difference between transforms will only have slight effect on the reconstruction performance. However, most real-world signals are compressible but not sparse. The reconstruction error is proportional to the decay rate of k^{-r}, as shown in Equation (8). For edges and curves, the contourlet has $r = 2$ while the wavelet has $r = 1$. This is the reason why we can better reconstruct the edges and curves using contourlet than using wavelet in CS-MRI. In addition, a high sampling rate is meaningless for practical MRI application since a high sampling rate means the reduction of acquired data is insignificant. Therefore, it is worth seeking geometric image transforms to better reconstruct MR images from highly undersampled k-space. A good geometric image
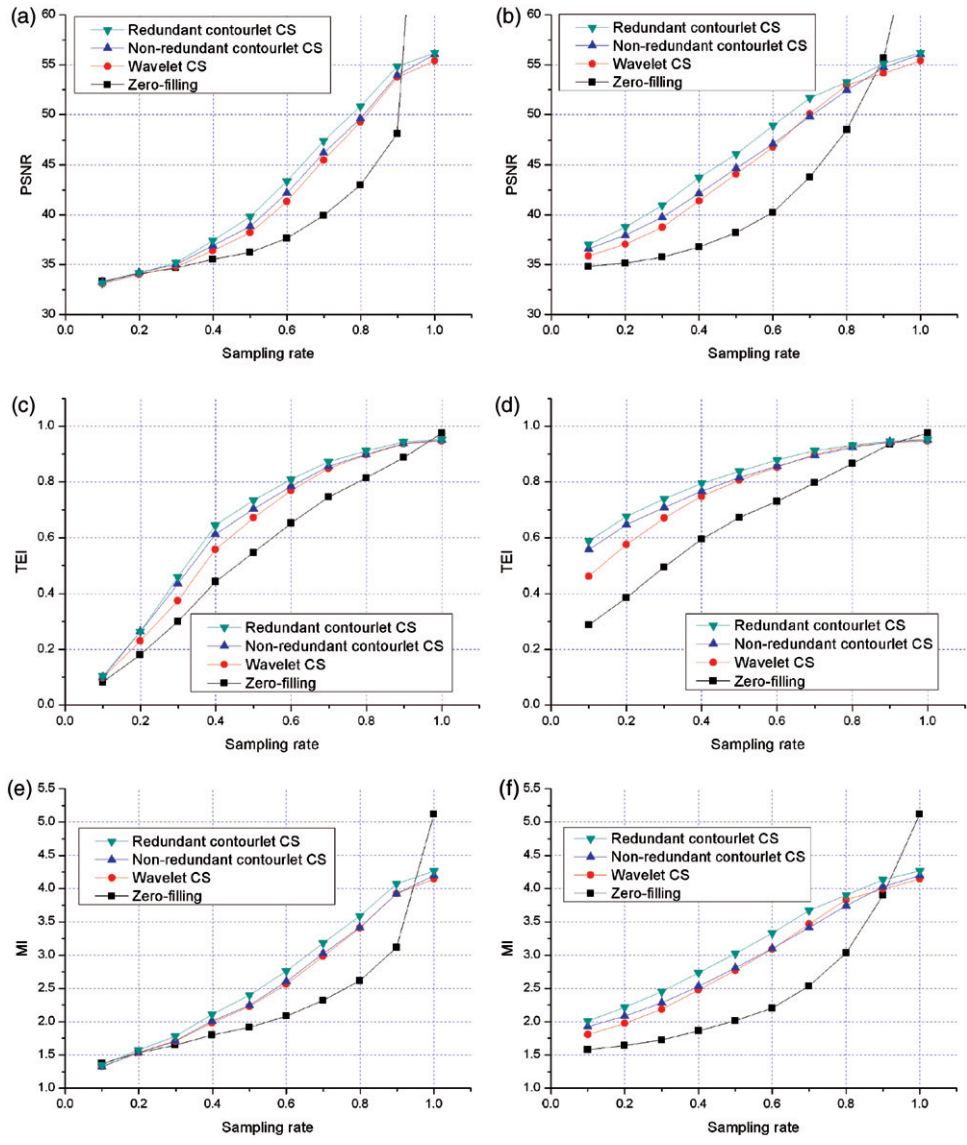
Figure 5. Objective evaluation criteria of wavelet and contourlet for CS-MRI with Cartesian sampling pattern (left column) and variable density sampling pattern (right column).
Note: PSNR, TEI and MI curves are shown in rows from top to bottom.

transform we pursue should at least enjoy some advantage in recovering certain image features over the others.

The comparison between the curves in the left and right columns of Figure 5 indicates that higher criteria are achieved by variable density sampling pattern relative to the Cartesian sampling pattern at the same sampling rate. The reason is that variable
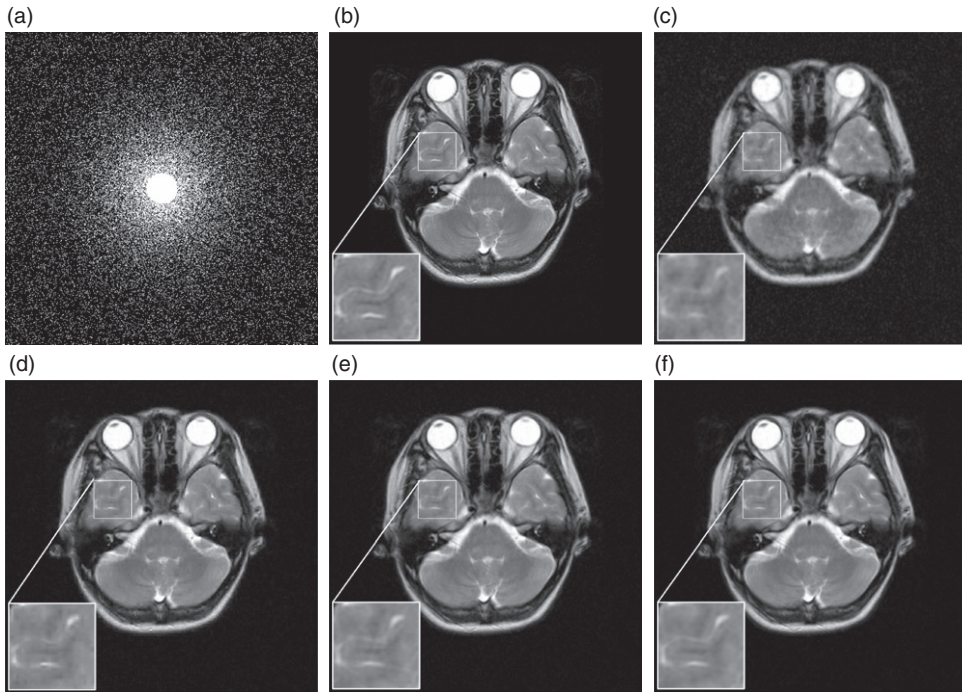
Figure 6. Comparison of wavelet and contourlet reconstruction of undersampled *k*-space using variable density sampling pattern with low sampling rate 0.15.
Notes: (a) The variable density sampling pattern, (b)–(f) are reconstructed images from fully sampled *k*-space, zero-filling *k*-space, wavelet-based, non-redundant contourlet-based and redundant contourlet-based CS, respectively.

density sampling pattern makes the artefacts look like random noise. This implies that the sampling pattern plays an important role in further reducing the measurements in *k*-space.

4.2.3. *Numerical convergence of iterative thresholding*

First, we use the objective criteria PSNR, TEI and MI to show the convergence of iterative thresholding at different values of the stopping criterion η. Parameters of iterative thresholding are $\varepsilon = 10^{-4}$, $\rho = 0.8$ and $\eta = 10^{-l}(l = 1, 2, \ldots, 9)$.

At the beginning, the relative noise level η in Equation (10) is set as a constant of 10^{-6}. We use the variable density sampling pattern with sampling rate 0.15 in Figure 6(a). The reconstructed image using the wavelet, the non-redundant contourlet and the redundant contourlet are shown in Figure 6(c)–(f). Figure 7(a) shows that with the increased iterations, the relative residual R_t in Equation (7) of all the transforms reaches η. Though the image is compressible, not rigorously sparse in transform domain,
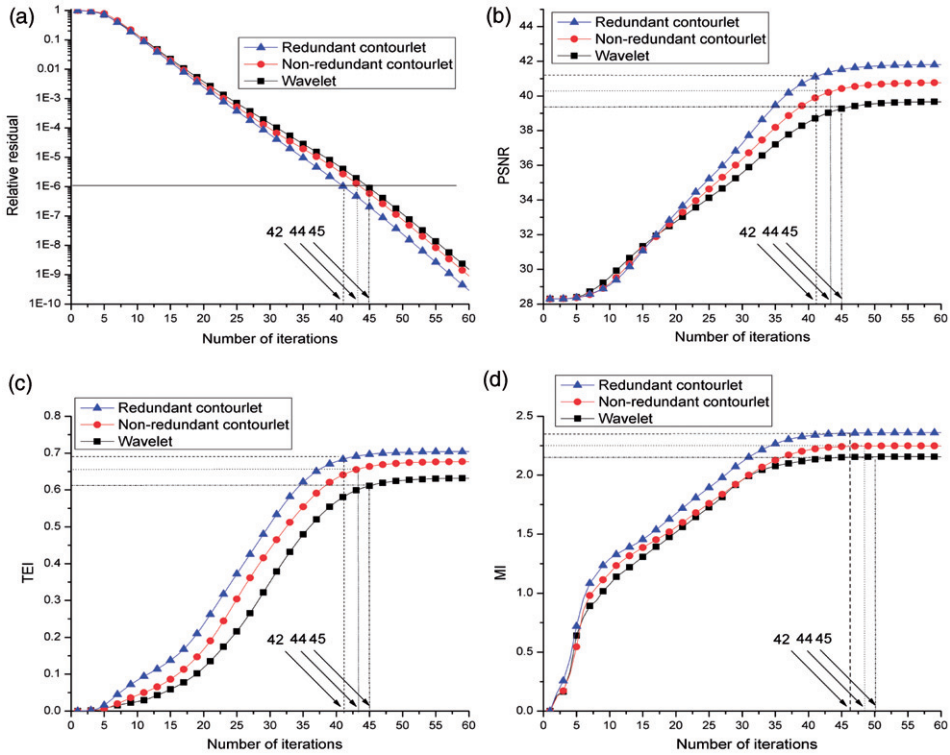
Figure 7. Objective evaluation criteria of wavelet and contourlet for CS-MRI *vs* number of iterations.
Notes: (a)–(d) are curves of relative residual, PSNR, TEI and MI, respectively. When $\eta = 10^{-6}$, iterative thresholding stops at 42nd, 44th and 45th iterations for redundant contourlet, non-redundant contourlet and wavelet, respectively.

iterative thresholding can recover the representation of a fully-sampled image up to any prescribed error tolerance, which is consistent with the results for k-term sparse signals [33]. If the number of iterations increases, the relative residual R_t will decrease, but the objective criteria, shown in Figure 7(b)–(d), tend to go stable and will eventually not increase.

Each transform comes to converge with almost the same speed, though the contourlet comes to η slightly faster than the wavelet. Values of PSNR, TEI and MI in Figure 7(b)–(d) show that the redundant contourlet obtains the highest objective criteria after the initial few iterations. This also demonstrates that the redundant contourlet can give the best reconstructed image, shown in Figure 6(f), at low sampling rate in k-space measurements.

Given a different stop criterion η of iterative thresholding, a large η corresponds to a sampled k-space highly corrupted by noise, while a small η corresponds to relatively low noise. Figure 8(d) shows that a too large η makes the image too smooth and loses texture. $\eta \leq 10^{-6}$ is sufficient to preserve image details and suppress the noise very well. Smaller η, e.g. 10^{-8} or 10^{-9}, also suppress the noise well because the l_1 norm enforces the sparsity to
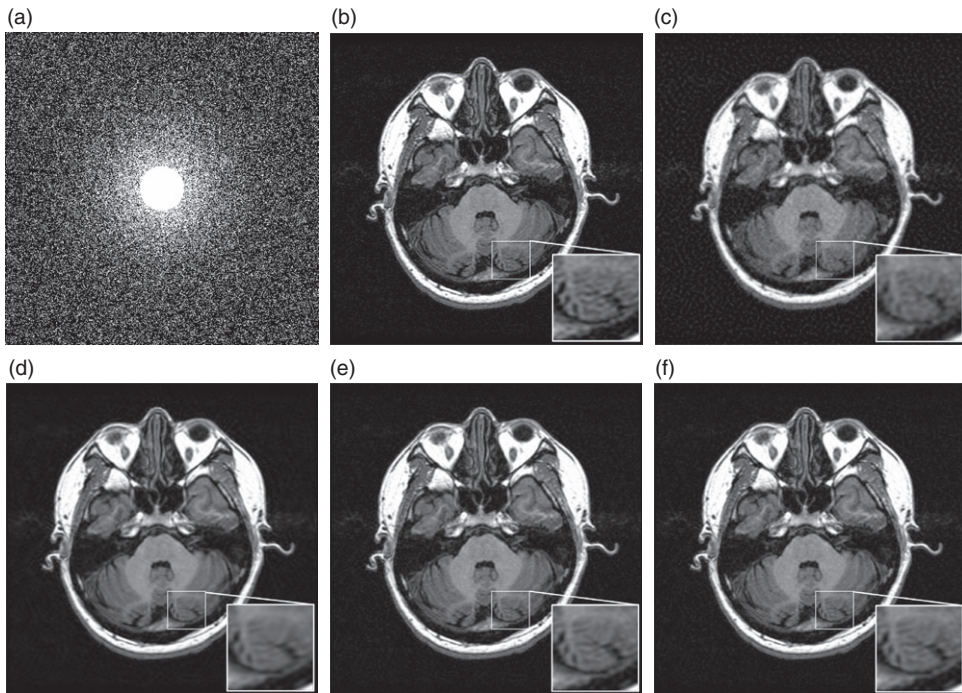
Figure 8. Reconstructed images at different stop criteria using redundant contourlet.
Notes: (a) Variable density sampling pattern with sampling rate 0.2, (b) and (c) are reconstructed images from fully sampled and zero-filling k-space, (d)–(f) are CS reconstructed images at stop criteria of 10^{-4}, 10^{-6} and 10^{-8}, respectively.

recover image details and noise is not sparse in contourlet domain. However, a smaller η requires more iterations. $\eta = 10^{-6}$ is empirically feasible. The objective criteria shown in Figure 9(a)–(c) prove that this conclusion holds since criteria do not improve with the decreasing of η when $\eta \leq 10^{-6}$. It is obvious that estimating the noise in k-space is very significant for CS-MRI. A precise estimation of noise can suppress the noise and preserve image details very well.

Second, we investigate and report the numerical results with various values of ρ. The sampling pattern in Figure 6(a) and the MR image in Figure 6(b) are employed in simulation. η is set as $\eta = 10^{-6}$. ρ varies from 0.1 to 0.95. The ith value of ρ is $\rho_i \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}(i = 1, 2, \ldots, 10)$.

Figures 10(a)–(c) show that increasing ρ leads to higher objective criteria, which means an improvement on quality of reconstructed image. However, the improvement is at the expense of more iterations in reconstruction, as shown in Figure 10(d). This expense is serious, especially when $\rho > 0.8$. For example, the number of iterations, when $\rho = 0.9$, is nearly twice that when $\rho = 0.8$. So, there is a tradeoff between the number of iterations and the objective criteria. Figure 10(e) shows the PSNR versus number of iterations. Each square point corresponds to one ρ as shown in Figure 10(d). Figure 10(e) indicates that with the increase of ρ, the gain in image quality brought by increasing the iterations becomes slight when $\rho > 0.8$. This is also the case in TEI and MI because
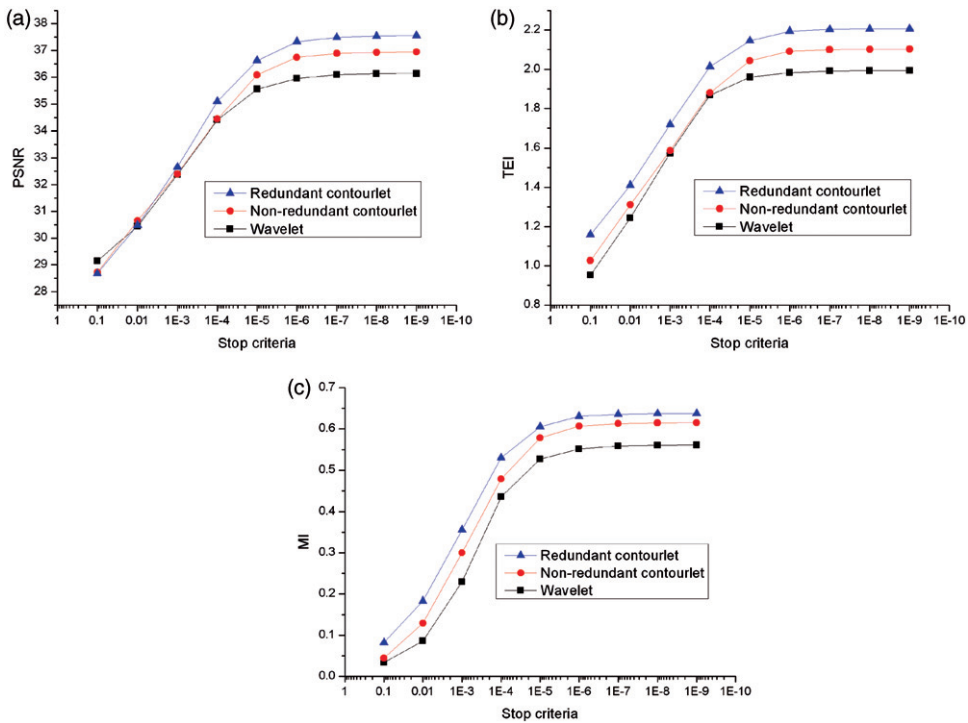
Figure 9. Stop criteria of iterative thresholding for CS-MRI.
Notes: (a)–(c) are PSNR, TEI and MI curves using variable density sampling pattern with sampling rate 0.2 for the image in Figure 8(b).

their similar relationship with ρ. So, $\rho > 0.8$ is a good choice regarding both objective criteria and number of iterations. For the sake of very good criteria, $0.8 \leq \rho \leq 0.9$ is recommended.

5. Discussion and conclusions

In this article, a geometric image transform, the contourlet, is introduced into CS-MRI and a simple but efficient iterative thresholding algorithm is applied to solve l_1 optimization of CS-MRI. Considering the transferred information and edge features, we introduce objective criteria to evaluate the performance of different geometric image transforms in CS-MRI. Simulation results demonstrate that enforcing the sparsity of the image in the contourlet transform domain can better reconstruct the curves of MR images than the traditional wavelet, especially at low sampling rates of k-space. Furthermore, improving the redundancy of the contourlet can better suppress the blurring artefact in reconstructed MR images. The convergence of iterative thresholding algorithm is analysed in simulations.
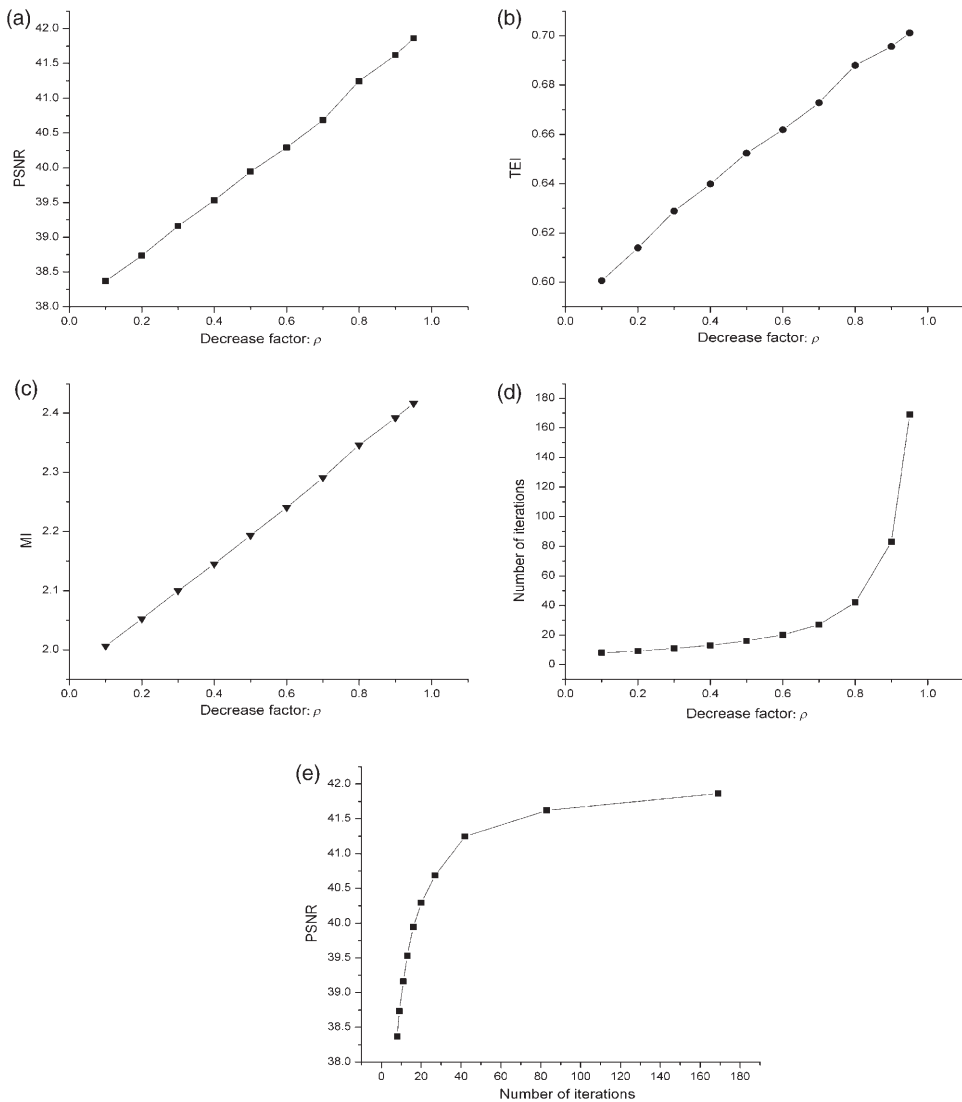
Figure 10. Decrease factor of iterative thresholding for CS-MRI.
Notes: (a)–(c) are PSNR, TEI and MI curves when decrease factor ρ varies from 0.1 to 0.95, (d) is the number of iterations for different decrease factor, (e) PSNR versus number of iterations for corresponding decrease factor in (d).

However, most MR images are not sparse in one transform domain. This article only presents how to better reconstruct the curves and edges by applying a geometric image transform. For further improvements in reconstructing MR images, combination of multiple transforms, which sparsely represents different image structures, i.e. constructing redundant basis, is a wise choice for CS-MRI. This is one of our future works and the primary results were published in [42] in the mean time of revising this article. Nevertheless, as an emergent sampling and

reconstruction method for MRI, CS will have further applications to improve the imaging speed.

References

[1] G. McGibney, M.R. Smith, S.T. Nichols, and A. Crawley, *Quantitative evaluation of several partial Fourier reconstruction algorithms used in MRI*, Magn. Reson. Med. 43 (1993), pp. 335–341.

[2] F. Wajer, *Non-cartesian MRI scan time reduction through sparse sampling*, Ph.D. thesis, Delft University of Technology, 2001.

[3] J. Song, Y. Liu, S. Gewalt, G. Cofer, G. Johnson, and Q. Liu, *Least-square NUFFT methods applied to 2D and 3D radially encoded MR image reconstruction*, IEEE Trans. Biomed. Eng. 56 (2009), pp. 1134–1142.

[4] K.T. Block, M. Uecker, and J. Frahm, *Undersampled radial MRI with multiple coils, iterative image reconstruction using total variation constraint*, Magn. Reson. Med. 57 (2007), pp. 1086–1098.

[5] D. Xu, K.F. King, and Z.P. Liang, *Improving k-t SENSE by adaptive regularization*, Magn. Reson. Med. 57 (2007), pp. 918–930.

[6] H. Jung, K. Sung, K.S. Nayak, E.Y. Kim, and J.C. Ye, *k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI*, Magn. Reson. Med. 61 (2009), pp. 103–116.

[7] Z.P. Liang and P.C. Lauterbur, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*, Wiley-IEEE Press, New York, 1999.

[8] Z.P. Liang, F. Boada, T. Constable, E.M. Haacke, P.C. Lauterbur, and M.R. Smith, *Constrained reconstruction methods in MR imaging*, Rev. Magn. Reson. Med. 4 (1992), pp. 67–185.

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

[10] D.L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory 52 (2006), pp. 1289–1306.

[11] M. Lustig, D.L. Donoho, and J.M. Pauly, *Sparse MRI: The application of compressed sensing for rapid MR imaging*, Magn. Reson. Med. 58 (2007), pp. 1182–1195.

[12] J.C. Ye, S. Tak, Y. Han, and H.W. Park, *Projection reconstruction MR imaging using FOCUSS*, Magn. Reson. Med. 57 (2007), pp. 764–775.

[13] J. Trzasko and A. Manduca, *Highly undersampled magnetic resonance image reconstruction via homotopic l(0)-minimization*, IEEE Trans. Med. Imaging 28 (2009), pp. 106–121.

[14] U. Gamper, P. Boesiger, and S. Kozerke, *Compressed sensing in dynamic MRI*, Magn. Reson. Med. 59 (2008), pp. 365–373.

[15] J.P. Haldar, D. Hernando, and Z.-P. Liang, *Compressed sensing in MRI with random encoding*, Preprint (2009). Available at http://coil.ifp.uiuc.edu

[16] A. Bilgin, Y. Kim, H.G. Lalgudi, T.P. Trouard, and M.I. Altbach, *Parallel magnetic resonance imaging using compressed sensing*, Proceedings of 2008 applications of digital image processing XXXI, SPIE, San Diego, CA, 2008, p. 70731G.

[17] 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*, UCLA CAM report 06-35, University of California, Los Angeles, 2006.

[18] B. Liu, K. King, M. Steckner, J. Xie, J. Sheng, and L. Ying, *Regularized sensitivity encoding (SENSE) reconstruction using Bregman iterations*, Magn. Reson. Med. 61 (2009), pp. 145–152.

[19] D.L. Donoho, M. Vetterli, R.A. DeVore, and I. Daubechies, *Data compression and harmonic analysis*, IEEE Trans. Inform. Theory 44 (1998), pp. 2435–2476.

[20] D.L. Donoho, *Unconditional bases are optimal bases for data compression and for statistical estimation*, Appl. Comput. Harmon. Anal. 1 (1993), pp. 100–115.

[21] E. Candès and D.L. Donoho, *Curvelets – A surprisingly effective nonadaptive representation for objects with edges*, in *Curves and Surface Fitting*, S.A. Cohen, C. Rabut, and L. Schumaker, eds., Vanderbilt University Press, Nashville, 2000, pp. 105–120.

[22] M.N. Do and M. Vetterli, *The contourlet transform: An efficient directional multiresolution image representation*, IEEE Trans. Image Process. 14 (2005), pp. 2091–2106.

[23] R.R. Coifman and D.L. Donoho, *Translation Invariant De-noising, Wavelets and Statistics*, Springer-Verlag, New York, 1995, pp. 125–150.

[24] X.B. Qu, D. Guo, Z. Chen, and C.B. Cai, *Compressed sensing MRI based on nonsubsampled contourlet transform*. Proceedings of 2008 IEEE International Symposium on IT in Medicine & Education-ITME'08, Xiamen, China, 2008, pp. 693–696.

[25] I. Daubechies, M. Defrise, and C. de Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Commun. Pure Appl. Math. 57 (2004), pp. 1413–1457.

[26] Y. Lu and M.N. Do, *A new contourlet transform with sharp frequency localization*, Proceedings of 2006 International Conference on Image Processing-ICIP'06, Atlanta, USA, 2006, pp. 1629–1632.

[27] D.C. Arthur, J.P. Zhou, and M.N. Do, *The nonsubsampled contourlet transform: Theory, design, and applications*, IEEE Trans. Image Process. 15 (2006), pp. 3089–3101.

[28] E. Candès and D.L. Donoho, *New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities*, Comm. Pure Appl. Math. 52 (2004), pp. 219–266.

[29] E. Candès and J. Romberg, *Practical signal recovery form random projections*. Available at http://www.acm.caltech.edu/~emmanuel

[30] S. Chen, D.L. Donoho, and M.A. Saunders, *Atomic decomposition by basis pursuit*, SIAM J. Sci. Comp. 20 (1999), pp. 33–61.

[31] K. Bredies and D.A. Lorenz, *Linear convergence of iterative soft-thresholding*, J. Fourier Anal. Appl. 14 (2008), pp. 813–837.

[32] D.L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, *Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit*, Tech. Rep. 2006-02, Department of Statistics, Stanford University, Stanford, California, USA, 2006.

[33] K.K. Herrity, A.C. Gilbert, and J.A. Tropp, *Sparse approximation via iterative thresholding*, Proceedings of 2006 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'06, Toulouse, France, 2006, pp. 624–627.

[34] I. Drori, *Fast l_1 minimization by iterative thresholding for multidimensional NMR spectroscopy*, Eurasip. J. Adv. Signal Process 2007, Article ID 20248. doi:10.1155/2007/20248.

[35] V. Agarwal, A.V. Gribok, and M.A. Abidi, *Image restoration using L-1 norm penalty function*, Inverse Probl. Sci. Eng. 15 (2007), pp. 559–583.

[36] J.W. Ma, *Compressed sensing by inverse scale space and curvelet thresholding*, Appl. Math. Comput. 206 (2008), pp. 980–988.

[37] G.H. Qu, D.L. Zhang, and P.F. Yan, *Information measure for performance of image fusion*, Electron. Lett. 38 (2002), pp. 313–315.

[38] C.S. Xydeas and V. Petrovic, *Objective image fusion performance measure*, Electron. Lett. 36 (2000), pp. 308–309.

[39] S.M. Phoong, C.W. Kim, P.P. Vaidyanathan, and R. Ansari, *A new class of two-channel biorthogonal filter banks and wavelet bases*, IEEE Signal Process. Mag. 43 (1995), pp. 649–665.

[40] D.D.-Y. Po and M.N. Do, *Directional multiscale modeling of images using the contourlet transform*, IEEE Trans. Image Process. 15 (2006), pp. 1610–1620.

[41] M. Lustig, *Sparse MRI*, Ph.D. thesis, Stanford University, 2008.

[42] X. Qu, X. Cao, D. Guo, C. Hu, and Z. Chen, *Combined sparsifying transforms for compressed sensing MRI*, Electron. Lett. 46 (2010), pp. 121–122.