MR IMAGE RECONSTRUCTION BY USING THE ITERATIVE REFINEMENT METHOD AND NONLINEAR INVERSE SCALE SPACE METHODS

LIN HE*, TI-CHIUN CHANG†, STANLEY OSHER *, TONG FANG †, AND PETER SPEIER ‡

Abstract. Magnetic resonance imaging (MRI) reconstruction from sparsely sampled data has been a difficult problem in medical imaging field. We approach this problem by formulating a cost functional that includes a constraint term that is imposed by the raw measurement data in k-space and the L^1 norm of a sparse representation of the reconstructed image. The sparse representation is usually realized by total variational regularization and/or wavelet transform. We have applied the Bregman iteration to minimize this functional to recover finer scales in our recent work. Here we propose nonlinear inverse scale space methods in addition to the iterative refinement procedure. Numerical results from the two methods are presented and it shows that the nonlinear inverse scale space method is a more efficient algorithm than the iterated refinement method.

Key words. Magnetic Resonance Image Reconstruction, Total Variation, Wavelet Transform, Iterated Refinement Techniques, Inverse Scale Space Methods, Sparse Representation.

1. Introduction. Magnetic resonance imaging (MRI) has been very successful for imaging parts of the body that are stationary. However, the major drawback of MRI is imaging speed. Many of the new applications of MRI require fast imaging, i.e., short scan time. To this end, during the scanning process, it is necessary to sample the frequency plane (or k-space) very sparsely. This usually causes difficulties to image reconstruction. Image artifacts and/or low signal to noise ratio (SNR) are often seen in images reconstructed from the sparsely sampled data.

As a result, a general problem was proposed in [11] by Candes et al.. They consider for a discrete complex signal n of length N and a randomly chosen set of frequencies Ω of mean size τN with $\tau < 1$, whether it is possible to reconstruct n from the partial knowledge of its Fourier coefficients on the set Ω. The answer is yes,

PROPOSITION 1.1. *if the signal n obeys*

$$\#\{t,\, n(t) \neq 0\} \leq \alpha(M) \cdot (\log N)^{-1} \cdot \#\Omega,\ \forall M \in \mathbb{N}$$

then with the probability at least $1 - O(N^{-M})$, n can be reconstructed exactly as the solution to the L^1 minimization problem

$$\min_m \sum_{t=0}^{N-1} |m(t)|,\ \ s.t.\ \hat{m}(\omega) = \hat{n}(\omega)\ \textit{for all}\ \omega \in \Omega,$$

where m is the reconstructed signal.

The assumption of the signal n being sparse is the key point in the proof of the above proposition (cf. [11]). However, in practice many applications deal with more complicated signals rather than sparse ones. Therefore, Candes and etc choose to solve a extended L^1 minimization problem by assuming that the signal n has a sparse

*UCLA Mathematics Department, Box 951555, Los Angeles, CA 90095-1555, U.S.A. ({helin, sjo }@math.ucla.edu)

†Siemens Corporate Research, Princeton, NJ 08536, U.S.A. ({ Ti-Chiun.chang, Tong.fang }@siemens.com)

‡Siemens AG Med, Erlangen, Germany.

representation. The L^1 minimization problem is reformulated as

$$\min_m ||\psi(m)||_1 \text{ s.t. } Fm = y, \tag{1.1}$$

where m is the reconstructed signal/image, ψ transforms the image m into a sparse representation, F is an undersampled Fourier matrix which satisfies $Fm = \hat{m}|_\Omega$, and $y = \hat{n}|_\Omega$ is the under-sampled measured k-space data which satisfies. There is a series of theoretical results on the extended L^1 minimization problem involving the uniform uncertainty principle (cf. [6]). Furthermore, the numerical solutions of (1.1) solved from numerical phantoms (cf. [11]) or from practical signals (cf. [10, 20]) have shown that the reconstructed MR image quality surpasses the conventional techniques.

Based on their work, we proposed in [14] to solve (1.1) with an iterated refinement method or the so called Bregman iteration (cf. [24, 18, 17, 27]). It was recently introduced to image denoising problems by Osher et. al. in [24]. The main idea is that given a noisy image f, by denoting the residual for the first time iteration $v_1 = 0$, we obtain the reconstructed image u_1 by solving the well known Rudin, Osher and Fatemi (ROF) minimization (cf. [25]) from the observed image $f/f + v_1$; at the kth iteration, we denote $v_k = f + v_{k-1} - u_{k-1}$, then we obtain the reconstructed image u_k by simply performing the ROF minimization again from $f + v_k$. This resulted sequence of $\{u_k\}$ has been proved to converge monotonically in L^2 to the noisy image f. If f is noise free, then the reconstructed image sequence $\{u_k\}$ will converge to the true image f. In addition the sequence $\{u_k\}$ monotonically gets closer to the noise free image f in the sense of Bregman distance (cf. [24]). This fact inspired us to apply the Bregman iteration [14] to solve our modified formulation of (1.1) to obtain a sequence of reconstructed MR image $\{m_k\}$ where Fm_k converges to y. In the numerical demonstrations of this paper and [14], finer scales have been recovered. The details of the procedure is described in Subsection 4.1.

Furthermore, the new idea of this paper is that a time-continuous inverse scale-space formulation (cf. [4]) is applied to solve (1.1). In the context of processing noisy images, scale space methods which start with the noisy image can separate the noise or so called small scales from large scales by smoothing small scale features faster than large ones. However, inverse scale space methods start with the image $u_0 = 0$ and approach the noisy image f as time increases, with large scales converge faster than small ones. Thus, if the method is stopped at a suitable time, large scale features may already be incorporated into the reconstruction, while small scale features (including the noise) are still missing. Therefore, the inverse scale space method may be treated as the time continuous version of the iterated refinement method. But to apply the Bregman iteration, the minimization problem has to be solved several times before a stopping criterion is satisfied. This could be very time-consuming. The inverse scale-space method is shown to be more straightforward and more efficient than Bregman iteration. We will discuss it in Subsection 4.2 and demonstrate it in our numerical results.

The organization of this paper is as follows. Section 2 describes prerequisites to implementing the model (1.1). Section 3 is dedicated to the sparse representation of an image, particularly a piece-wise smooth one. As it has been mentioned above, Section 4 focuses on the iterated refinement procedure and the time continuous inverse scale space methods. Finally, the numerical results are performed in Section 5 to show the good quality of reconstructed images from (1.1) and the benefits of using the Bregman iteration and the nonlinear inverse scale space method.

2. Prerequisite to Solve the Model (1.1)**.** In this section, we describe related tools in more details to solve the given problem (1.1). This includes the use of sparse transformations and the implementation of maintaining the constraint condition. Following the work in [10, 20], the total variation (TV) regularization or the wavelet transform are used for the MR image reconstruction from raw k-space data of a numerical phantom.

2.1. The Total Variation Regularization . The total variation regularization or the so-called bounded variation (BV) norm has been first introduced by Rudin et al. in [25] in the context of image denoising. Afterwards this approach was generalized to other image processing problems, such as deconvolution (cf. [13]), decomposition (cf. [12, 17, 28]), computer tomography (cf. [29]), etc. It is proved to be particularly relevant in recovering piecewise smooth functions without smoothing sharp discontinuities while also being a sparse transformation for piecewise constant functions.

Normally, the bounded variation norm of an two dimensional real image m is defined as

$$||\psi(m)||_1 = ||m||_{BV} := \sum_{i,j} \sqrt{|\nabla_x m(i,j)|^2 + |\nabla_y m(i,j)|^2}. \tag{2.1}$$

However, our reconstructed image m from (1.1) equals the magnitude of a complex image (m_r, m_i) from the inverse Fourier transform of the under-sampled data y, i.e., $m = \sqrt{m_r^2 + m_i^2}$. We can not just simply define the BV norm of image m by (2.1) due to the non-convexity of this formulation. Instead, we define the BV norm of m as

$$||m||_{BV} = \sum_{i,j} \sqrt{|\nabla_x m_r(i,j)|^2 + |\nabla_x m_i(i,j)|^2 + |\nabla_y m_r(i,j)|^2 + |\nabla_y m_i(i,j)|^2}. \tag{2.2}$$

Another alternative is also used. The basic idea is similar to (2.2), that is, to separate the real and imaginary part,

$$||m||_{BV} = \sum_{i,j} \sqrt{|\nabla_x m_r(i,j)|^2 + |\nabla_y m_r(i,j)|^2} + \sqrt{|\nabla_x m_i(i,j)|^2 + |\nabla_y m_i(i,j)|^2}. \tag{2.3}$$

The minimization of (2.2) and (2.3) subject to the constraint condition $Fm = y$ is guaranteed to converge to a point according to [2]. Indeed, our experiments show that they both produce similar and reasonable results.

2.2. The Wavelet Transform. The wavelet transform is tremendously popular in the signal and image processing communities, due in large part to its ability to provide parsimonious representations for signals that are smooth away from isolated discontinuities. The wavelet transform is widely used in image processing, and it generally represents a sparse representation of the signal/image g. The result from the wavelet transform is compared with the result from the total variation regularization subject to the constraint condition in [20]. The wavelet transform does a fairly good job although not as good as the BV norm.

In the following we will first define the L^1/wavelet norm $||\psi(g)||_1$ and then derive the sub-gradient of the L^1/wavelet norm. For the reason of simplicity, we consider a 1D signal g of length $N = 2^J$. At the level L of the wavelet transform, the signal g is

decomposed as

$$g = \sum_{p=0}^{2^L-1} \alpha_{L,p}\varphi_{L,p} + \sum_{q=L+1}^{J} \sum_{p=0}^{2^{q-1}-1} \alpha_{q,p}\phi_{q,p}, \tag{2.4}$$

where the $\varphi_{L,p}$ are scaling functions, the $\phi_{q,p}$ are oscillatory wavelets, see [22, 5, 21]. Then the L^1/wavelet norm $||\psi(g)||_1$ is defined as

$$||\psi(g)||_1 = ||g||_W := \sum_{p=0}^{2^L-1} |\alpha_{L,p}| + \sum_{q=L+1}^{J} \sum_{p=0}^{2^{q-1}-1} |\alpha_{q,p}|. \tag{2.5}$$

From (2.4), for the signal g and a variable signal h, we assume that

$$\begin{cases} g &= \sum_{p=0}^{2^L-1} \alpha_{L,p}\varphi_{L,p} + \sum_{q=L+1}^{J} \sum_{p=0}^{2^{q-1}-1} \alpha_{q,p}\phi_{q,p}, \\ h &= \sum_{p=0}^{2^L-1} \beta_{L,p}\varphi_{L,p} + \sum_{q=L+1}^{J} \sum_{p=0}^{2^{q-1}-1} \beta_{q,p}\phi_{q,p}. \end{cases}$$

The sub-gradient of $||g||_W$ is defined as:

$$\begin{aligned} & \lim_{s\to 0} \frac{1}{s}(||g+sh||_W - ||g||_W) \\ = \; & \lim_{s\to 0} \frac{1}{s}(\sum_{p,q} |\alpha_{p,q} + s\beta_{p,q}| - \sum_{p,q} |\alpha_{p,q}|) \\ = \; & \textstyle\sum_{p,q} \operatorname{sgn}(\alpha_{p,q})\beta_{p,q} \\ = \; & (\textstyle\sum_{p=0}^{2^L-1} \operatorname{sgn}(\alpha_{L,p})\varphi_{L,p}, \sum_{p=0}^{2^L-1} \beta_{L,p}\varphi_{L,p}) \\ & +(\textstyle\sum_{q=L+1}^{J} \sum_{p=0}^{2^{q-1}-1} \operatorname{sgn}(\alpha_{q,p})\phi_{q,p}, \sum_{q=L+1}^{J} \sum_{p=0}^{2^{q-1}-1} \beta_{q,p}\phi_{q,p}) \\ = \; & (\textstyle\sum_{p=0}^{2^L-1} \operatorname{sgn}(\alpha_{L,p})\varphi_{L,p} + \sum_{q=L+1}^{J} \sum_{p=0}^{2^{q-1}-1} \operatorname{sgn}(\alpha_{q,p})\phi_{q,p},\ h). \end{aligned} \tag{2.6}$$

The last two steps are valid because the scaling functions φ and the wavelet functions ϕ are an orthonormal basis. Thus, the sub-gradient of $||g||_W$ is

$$\sum_{p=0}^{2^L-1} \operatorname{sgn}(\alpha_{L,p})\varphi_{L,p} + \sum_{q=L+1}^{J} \sum_{p=0}^{2^{q-1}-1} \operatorname{sgn}(\alpha_{q,p})\phi_{q,p},$$

which is implemented by the inverse wavelet transformation.

2.3. The Constraint Condition. Depending on the accuracy for which the locations of the original k-space samples are maintained, there are two ways to deal with the constraint condition.

After we have the measurement data from the MR scanner which for example consists of 63 radial lines with 512 samples each, we can round off the sample locations to the integer grids. Then the nonzero integer grids become the constraint set Ω. This corresponds to 0 order interpolation which is not accurate. But we have the advantage of imposing the condition $Fm = y$ directly by an inverse fast Fourier transform (FFT). Combining with the minimization of $||\psi(m)||_{L^1}$, m converges to the solution of (1.1) according to the well-known algorithm of projection onto convex set (POCS) (cf. [2]).

A better interpolation is to retain the locations of the sampled data and use a non-uniform FFT operator to approach y. We achieve this by minimizing $||Am-y||_2^2$, where A denotes the non-uniform FFT operator, which we use NFFT from [19] or NUFFT from [16]. We have found great convenience in modifying NFFT package that is written in C into C++. However, it is faster to operate the interpolation by NUFFT package written in matlab because the NUFFT is designed under the optimal min-max distance. To clarify it, the interpolation coefficients of NUFFT depend on the location of the data in the spatial domain only while the interpolation coefficients of NFFT also depend on the value of the data in the spatial domain. Thus for NUFFT only one time pre-computation of the interpolation coefficients is needed, while for NFFT this computation needs to be repeated because the value of the data Am is changing.

By adding a parameter λ to control the L^2 distance between Am and y, our problem is reformulated as

$$\min_m \left\{ ||\psi(m)||_1 + \frac{\lambda}{2}||Am-y||_2^2 \right\}. \tag{2.7}$$

The sub-gradient of the second term in (2.7) is $\widehat{A}(Am-y)$, where $\widehat{A}$ is the adjoint of A.

3. The Sparse Representation for a Piecewise Smooth Image. As it was proven in [11], the reason that we can reconstruct the exact signal from incomplete frequency information is based on the assumption that the image has a sparse representation ψ. For a piecewise constant image, both the BV norm and the wavelet transform are good sparse representations. However when we are given raw measurement k-space data scanned from a real object, which is often piecewise smooth, the TV regularization or the L^1/wavelet alone will not be a good sparse transformation. We have to look for other means.

3.1. Wavelet + TV. In [15], Chen et al. proposed a dictionary merger *jump+wavelet*. As they claim, the *jump+wavelet* dictionary based on a merger of wavelets with tapered Heavisides (the Heaviside is equivalent to total variation norm in one dimension) will lead to a sparse representation for a piecewise smooth image. Based on such an over-complete dictionary we will decompose the image m into two components where h is composed of a Heaviside orthonormal basis and w is composed of a wavelet orthonormal basis. To be more general, when $m = \mu h + \nu w$, the following problem will be considered:

$$\min_{h,w,m=\mu h+\nu w} \left\{ ||m-\nu w||_{BV} + ||\nu w||_W + \frac{\lambda}{2}||Am-y||_2^2 \right\}, \tag{3.1}$$

where μ and ν are nonnegative parameters. (3.1) is still a convex formulation due to the linear relationship between m, h and w. Furthermore, when $\mu = 0$ or $\nu = 0$, (3.1) is simplified to the case of using the wavelet transformation or the total variation regularization alone.

We also want to draw a connection to the work [10], in which

$$\min_m \left\{ \mu||m||_{BV} + \nu||m||_W + \frac{\lambda}{2}||Am-y||_2^2 \right\} \tag{3.2}$$

is considered, where the parameter μ and ν prescribes the importance of the solution having small L^1 norm in the wavelet domain versus having small BV norm in the spatial domain. The formulation (3.1) is more faithful to finding a sparse representation

over an over-complete dictionary *jump+wavelet* for the piecewise smooth functions. However similar numerical results from both formulations are observed.

3.2. Curvelet + TV. The curvelet transformation (cf. [8, 7, 9]) was developed by Candes et al. in the last few years in an attempt to overcome inherent limitations of traditional multiscale representations such as wavelets. Conceptually, the curvelet transform is a multiscale pyramid with many directions and positions at each length scale, and needle-shaped elements at fine scales. Thus those curvelets including new tight frames of curvelets (cf. [8]) are able to address the problem of finding optimally sparse representations of objects with discontinuities along C^2 edges, which the wavelet is far from ideal. Based on the good behavior of curvelet+TV algorithm in denoising images (cf. [9]) we believe that it is worth implementing curvelet+TV as the sparse representation.

4. The Iterated Refinement Method and the Inverse Scale Space Method.

4.1. Bregman Iteration. Consider the minimization of a general model:

$$\min_u \{J(u) + \lambda H(u, f)\},$$

where u and f denote, respectively, the reconstructed image and the known data or the observed image. The regularization term $J(u)$ and the fidelity term $H(u, f)$ are required to be convex functionals of u. Moreover, $J(u)$ has to be non-negative.

Given a differentiable function J, the Bregman distance is defined by

$$D_J(x, z) = J(x) - J(z) - < x - z, \partial J(z) >,$$

where $< \cdot, \cdot >$ denotes the inner product in R^n and $\partial J(z)$ is an element of the sub-gradient of J at point z. See [3].

Since $J(u)$ is convex, $D_J(u, v)$ is convex in u for each v. The quantity $D_J(u, v)$ is not a distance in the usual sense; e.g., in general, $D_J(u, v) \neq D_J(v, u)$ and also the triangle inequality does not hold. However, it is a measure of closeness in that $D_J(u, v) \geq 0$ and $D_J(u, v) = 0$ if $u = v$ (if and only if for strictly convex functionals).

Suppose m_1^λ is the minimizer of $J(m) + \lambda H(m, y)$. An iterative procedure is given by the sequence of variational problems for $k \geq 2$,

$$m_k^\lambda = \arg\min_m \left\{ D_J(m, m_{k-1}^\lambda) + \lambda H(m, y) \right\}. \tag{4.1}$$

In this paper, we consider:

$$\begin{aligned} H(m, y) &= \frac{||Am - y||_2^2}{2}, \\ J(m) &= ||\psi(m)||_1. \end{aligned}$$

In [24], it is shown that the sequence Am_k^λ monotonically converges in the L^2 norm to the noisy image f, here to the Fourier coefficients y on the constraint set Ω. Another benefit of applying the Bregman iteration is that in the context of image denoising, we observe a smooth image at the beginning when the parameter λ is chosen small; and the fine scales are recovered along the Bregman iteration and they are recovered before the noise. Thus we just need to stop at the moment when the noise is still missing. During our experiments on MR image reconstruction, we observe similar

results, see Figure 5.2 and Figure 5.6. As surprised as we are by this observation that the sequence of the reconstructed images "know" what to recover along the Bregman iteration and recover the noise last, we will include this as a future work.

In the following we will see how to apply the Bregman distance.

Take the sub-gradient of (4.1) for our problem. For $k \geq 2$, we have :

$$\lambda H_m(m_k^\lambda, y) + \partial J(m_k^\lambda) - \partial J(m_{k-1}^\lambda) = 0. \tag{4.2}$$

While for $k = 1$,

$$\lambda \widehat{A}(Am_1^\lambda - y) + \partial J(m_1^\lambda) = 0.$$

Write $y - Am_1^\lambda = v_1^\lambda$, i.e.

$$\partial J(m_1^\lambda) = \lambda \widehat{A} v_1^\lambda,$$

and plug the above equation into (4.2) for $k = 2$, it becomes

$$\lambda \widehat{A}(Am_2^\lambda - y) + \partial J(m_2^\lambda) - \lambda \widehat{A} v_1^\lambda = 0.$$

It can be simplified as

$$\lambda \widehat{A}(Am_2^\lambda - y - v_1^\lambda) + \partial J(m_2^\lambda) = 0.$$

Write $y + v_1^\lambda - Am_2^\lambda = v_2^\lambda$, we then have

$$\partial J(m_2^\lambda) = \lambda \widehat{A} v_2^\lambda.$$

From above we can see, if we denote $v_0^\lambda = 0$ and

$$v_k^\lambda = y + v_{k-1}^\lambda - Am_k^\lambda \quad \forall k \geq 1,$$

we will have

$$\partial J(m_k^\lambda) = \lambda \widehat{A} v_k^\lambda.$$

In other words, to apply the Bregman iteration we only have to change the constraint set of Fourier coefficients from y to $y + v_{k-1}^\lambda$ at the kth iteration. Therefore, solving the sequence minimizing problem of (4.1) is equivalent to solving the following sequence minimizing problem

$$m_k^\lambda = \arg\min_m \left\{ J(m) + \lambda H(m, y + v_{k-1}^\lambda) \right\}. \tag{4.3}$$

There is an intuitive perspective to explain the reason why we apply the Bregman iteration. We consider a much simpler case where we full sample the Fourier coefficients of a disk on the integer grids. That means we have the complete information of $y = F(\alpha \chi^R_{(x_0,y_0)})$, where F denotes the uniform Fourier transform, α is the grey value of the image and

$$\chi^R_{x_0,y_0}(x,y) = \begin{cases} 1, & \text{if } (x-x_0)^2 + (y-y_0)^2 \leq R^2; \\ 0, & \text{otherwise.} \end{cases}$$

To study the link between the Bregman iteration and our model (2.7) we consider the following specific minimization problem:

$$\min_{m \in BV} \left\{ ||m||_{BV} + \frac{\lambda}{2} ||Fm - y||_2^2 \right\}. \tag{4.4}$$

i.e. $||\psi(m)||_1 = ||m||_{BV}$ and non-uniform fast Fourier transform A becomes uniform Fourier transform F. Based on a simple fact that $\hat{F}Fm = m$ and denote $f = \alpha\chi^R_{(x_0,y_0)}$, the simplified minimization problem (4.4) will become

$$\min_{m\in BV}\left\{||m||_{BV} + \frac{\lambda}{2}||m-f||_2^2\right\}.$$

This actually is the well-known total variation model, the so called ROF model. From the analysis of the ROF model by Meyer in [23], we obtain the solution of (4.4) $m_1^\lambda = (\alpha - \frac{2}{\lambda R})\chi^R_{(x_0,y_0)}$ when $\alpha\lambda R \geq 1$. Thus the residual $v_1^\lambda = y - Fm_1^\lambda = F\frac{2}{\lambda R}\chi^R_{(x_0,y_0)}$. Therefore, at the second Bregman iteration, we end up minimizing (4.4) with $y = F(\alpha + \frac{2}{\lambda R})\chi^R_{(x_0,y_0)}$, see (4.3). This will reconstruct the image m exactly, i.e. $m_2^\lambda = \alpha\chi^R_{(x_0,y_0)}$. And if $\alpha\lambda R < 1$, we denote the smallest integer n that satisfies $n\alpha\lambda R > 1$, then at the $n+1$th Bregman iteration, an exact reconstruction $m = \alpha\chi^R_{x_0,y_0}$ would be achieved, see [24].

4.2. Nonlinear Inverse Scale Space Methods. The nonlinear inverse scale space method described in [4] is derived as a limit of the iterated refinement procedure for $\lambda \to 0$. The key point is reinterpreting the fidelity parameter $\lambda = \Delta t$ as a time step and divide by Δt on both sides of the Euler-Lagrange equation (4.2), then

$$\frac{\partial J(m_k^{\Delta t}) - \partial J(m_{k-1}^{\Delta t})}{\Delta t} = -\widehat{A}(Am_k^{\Delta t} - y).$$

Setting $t_k = k\Delta t$, $p^{\Delta t}(t_k) = \partial J(m_k^{\Delta t})$ and $m^{\Delta t}(t_k) = m_k^{\Delta t}$, then we have

$$\frac{p^{\Delta t}(t_k) - p^{\Delta t}(t_k - \Delta t)}{\Delta t} = -\widehat{A}(Am^{\Delta t}(t_k) - y).$$

Letting $\Delta t \to 0$ and dropping the subindex k we arrive at the differential equation

$$\partial_t p(t) = -\widehat{A}(Am(t) - y),\ p(t) \in \partial J(m(t)), \tag{4.5}$$

with initial values given by $m(0) = p(0) = 0$.

The concise formulation of (4.5) is not straightforward to compute, as the relations between p and m are quite complicated in nonlinear cases. Here we resort to a relaxed version which aims at having a flow with qualitatively similar properties to that of (4.5) by using standard variational formulations, which are simple to compute. We compute the following coupled equations :

$$\begin{cases} \partial_t m(t) &= -p(t) - \lambda(\widehat{A}Am(t) - \widehat{A}y - v(t)) \\ \partial_t v(t) &= -\alpha(\widehat{A}Am(t) - \widehat{A}y), \end{cases} \tag{4.6}$$

where $m(0) = v(0) = 0$ and $\alpha > 0$ is a constant. It is easy to see that the steady state of these equations is $u = f$, $v = p/\lambda$. In order for the solution of (4.6) to converge to the steady state, $\alpha \leq \lambda/4$ is required. Please see [4] for detailed analysis. Our numerical experiments indicate that the nonlinear inverse scale space method works much faster than the iterated refinement method with similar results.

5. Experimental Results. The raw MR measurement data were obtained from a Siemens Magnetom Avanto 1.5T scanner. These data samples have a radial trajectory in k-space. Figure 5.1, 5.2 and 5.4 are reconstructed from the raw measurement

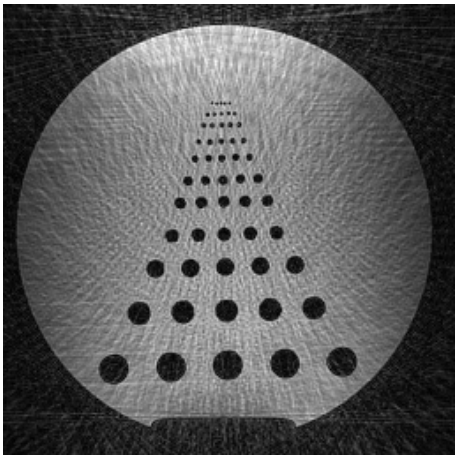

FIG. 5.1. *Conventional gridding algorithm*

k-space data scanned from a numeric phantom. The k-space data is composed of a total of 63 radial lines with 512 samples each. During the MR scanning, three coils/channels were used, the pulse sequence is trueFISP, the scanning parameters are TR=4.8ms, TE=2.4ms, flip angle a=60o, FOV=206mm with a resolution of 256 pixels; and the final image is obtained by taking the square root of the sum of each channel, which separately went through the proposed iterative procedure or the non-linear inverse scale space method.

We solve the minimization problem (2.7) by conjugate gradient descent method (cf. [26]) and back-tracking line search (cf. [1]). And we choose the sparse representation and the parameter λ to reach the best visual effects.

Figure 5.1 is obtained from conventional gridding algorithm. We see a lot noise and image artifacts due to the reason that the data is sampled sparsely.

Compared with Figure 5.1, the top left image in Figure 5.2 obtained from solving (2.7) with $||\phi(m)||_1 = ||m||_{BV}$ and $\lambda = 100.0$ demonstrates superior noise deduction. In order to enforce the constrained condition $Am = y$ we require a small value of the fidelity term $||Am - y||_2^2$. We could use a large λ, however this would result in a noisy image with a large bounded variation norm. Therefore, we choose to apply the Bregman iteration in solving the minimization problem (2.7). The sequence images of the 1st, 2nd, 4th and 6th Bregman iteration are shown in figure 5.2. Each Bregman iteration takes 200-500 steps to converge. As the iteration k increases, we observe a recovery of fine details, particularly including the black dots in the top row, see Figure 5.3.

In Figure 5.4, we present another set of four images corresponding to the 100th, 200th, 400th and 600th iteration results obtained from the time continuous flow of the relaxed inverse scale space method with $||\phi(m)||_1 = ||m||_{BV}$ and $\lambda = 100$. As mentioned in equation (4.6), it starts from $m(0) = 0$ and $v(0) = 0$. Thus we notice image artifacts at the early step, see the top two images of Figure 5.4. The artifacts disappear at the later step and the image of the 600th steps from the non-linear inverse scale space flow (see the bottom right of Figure 5.4) is comparable with the result of the 6th Bregman iteration (see the bottom right of Figure 5.2). However, the iterated refinement procedure is around 3-4 times more time-consuming.

Figure 5.5, 5.6 and 5.7 are reconstructed from 2 out of 8 interleaves of the k-

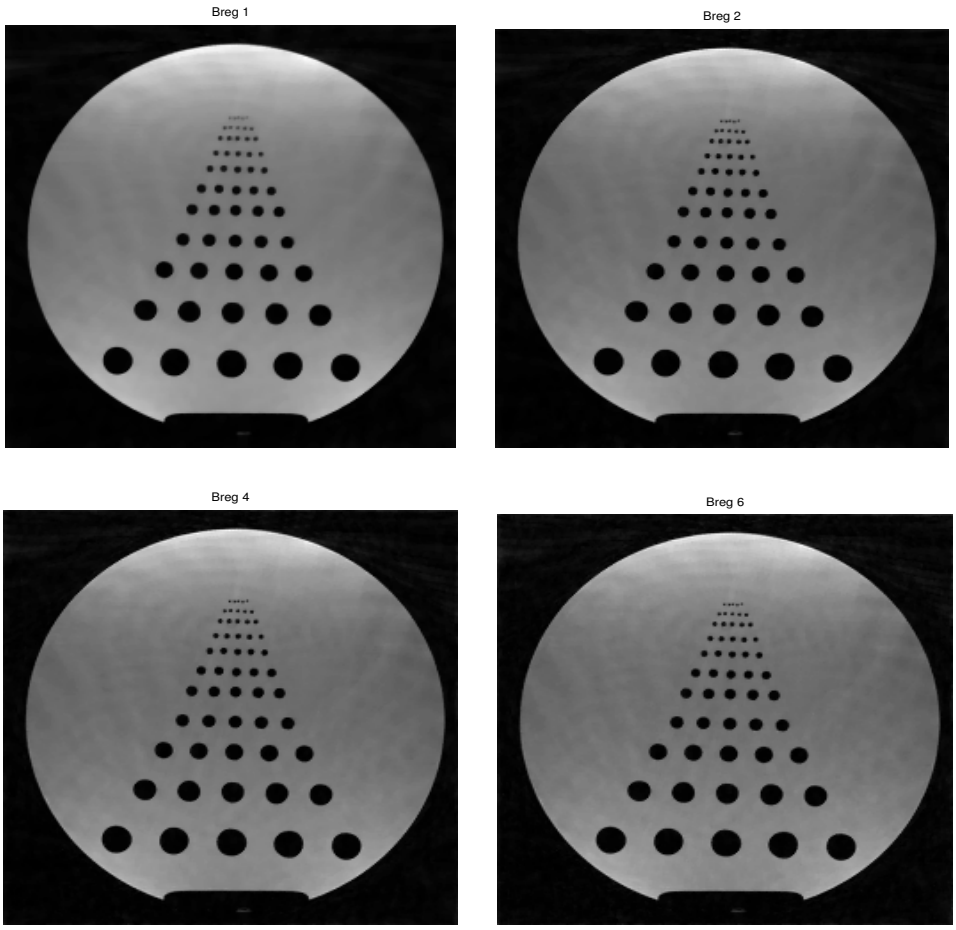

FIG. 5.2. *A sequence images obtained from the iterated refinement method : top left, 1st Bregman iteration; top right, 2nd Bregman iteration; bottom left, 4th Bregman iteration; bottom right, 6th Bregman iteration.*

space data scanned from a head. Each interleave has 31 radial lines with 512 samples each and 4 channels are used during the scan. Scanning parameters are TR=4.46ms, TE=2.23ms, flip angle a=50o, FOV=250mm with a resolution of 256 pixels.

Figure 5.5 is obtained from the conventional gridding algorithm. Again we see a lot of noise and image artifacts.

The minimization problem (3.2) is solved for Figure 5.6 and 5.7 with $\mu = 1.0$, $\nu = 0.1$ and $\lambda = 200.0$. This is only because, the computation of solving (3.2) is easier than the computation of solving (3.1). The Bregman iteration (see Figure 5.6) and the nonlinear inverse scale space method (see Figure 5.7) are carried on following the above section. We observe that Figure 5.6 and Figure 5.7 have less noise and are smoother than Figure 5.5. However, in terms of sharpness, Figure 5.5 is the best. This might be resulted from the ability of total variation to suppress oscillations. The nonlinear inverse scale space method is two times faster than the iterated refinement method.

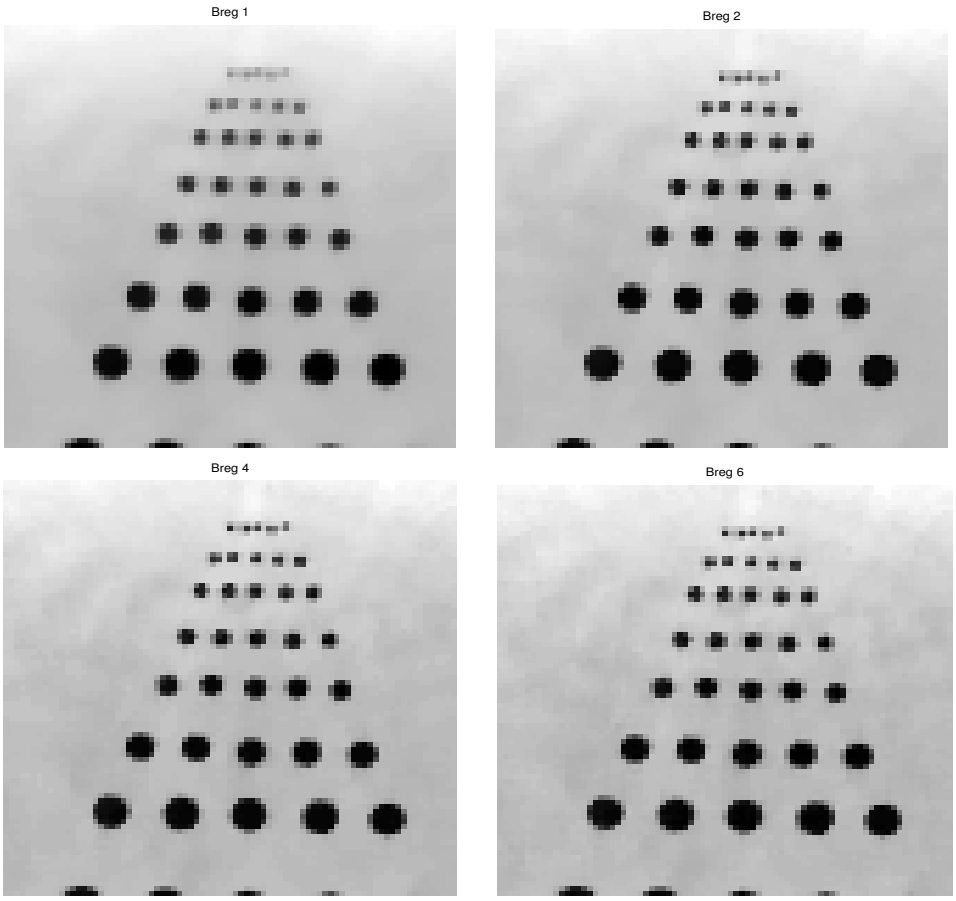

FIG. 5.3. *Zoom in images of the top row: top left, 1st Bregman iteration; top right, 2nd Bregman iteration; bottom left, 4th Bregman iteration; bottom right, 6th Bregman iteration.*

6. Conclusion. In this paper, we have applied the iterative refinement procedure and the inverse scale space method on MR image reconstruction. The reconstruction is realized by minimizing the summation of a regularization term and a constraint term (2.7). By simplify the formulation (2.7) to (4.4), we can actually obtain the ROF model. By following the same analysis for the ROF model by Meyer [23], we obtain an exact recovery for $y = F(\alpha\chi^R_{(x_0,y_0)})$ at the $(n+1)$th Bregman iteration where n is the smallest integer that satisfies $n\alpha\lambda R > 1$.

The numerical experiments from the iterative refinement method demonstrate similar behaviors as we have seen in applying the iterative refinement procedure to denoising images by solving the ROF model. The behavior is described as: by choosing a small λ, we obtain a smooth image at the first iteration; and the details of the image come back before the noise comes back as the Bregman iteration number k increases. Furthermore, the experimental results from inverse scale space methods have shown that it is a more efficient method than the iterative refinement method while being able to obtain similar results.

Compared with the conventional gridding method, our results demonstrate superior noise deduction while keeping the necessary features. However for the head image

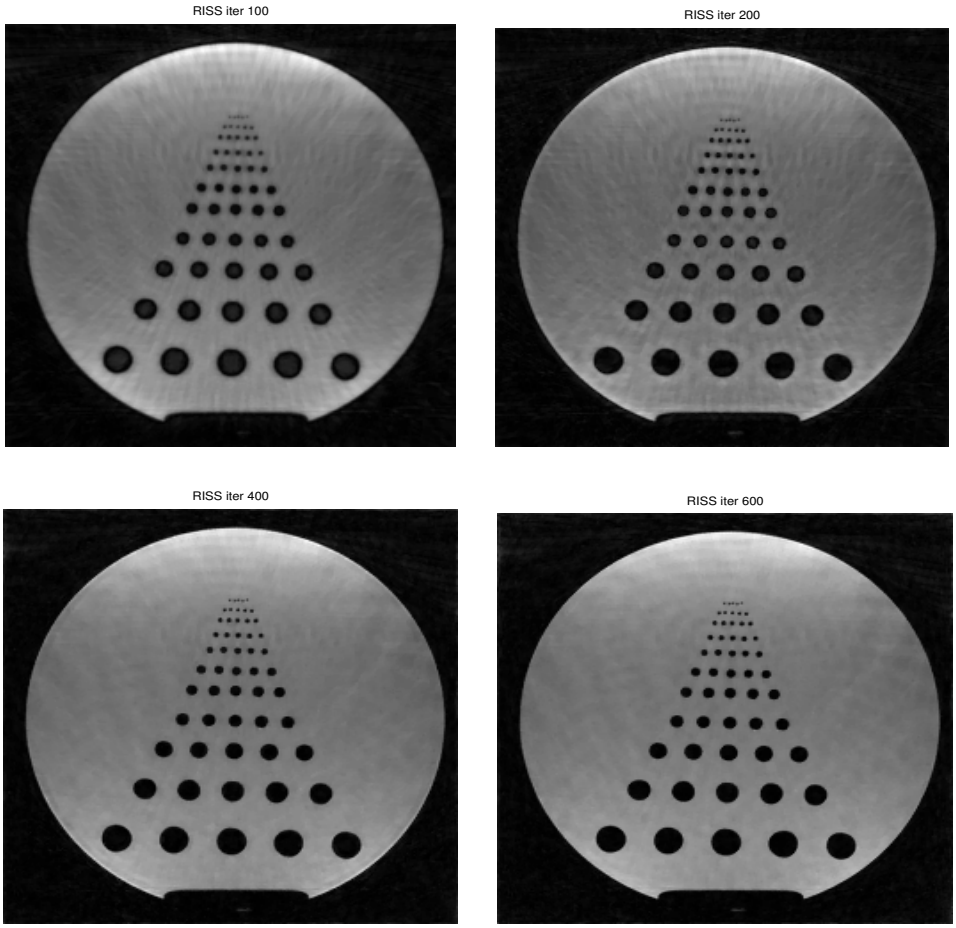

FIG. 5.4. *A sequence images obtained from the relaxed inverse scale space flow : top left, iteration 100; top right, iteration 200; bottom left, iteration 400; bottom right, iteration 600.*

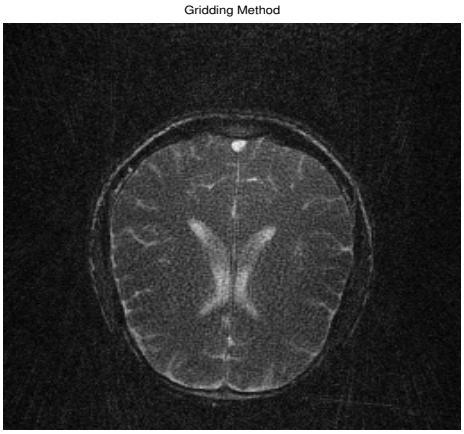

FIG. 5.5. *Conventional gridding algorithm*

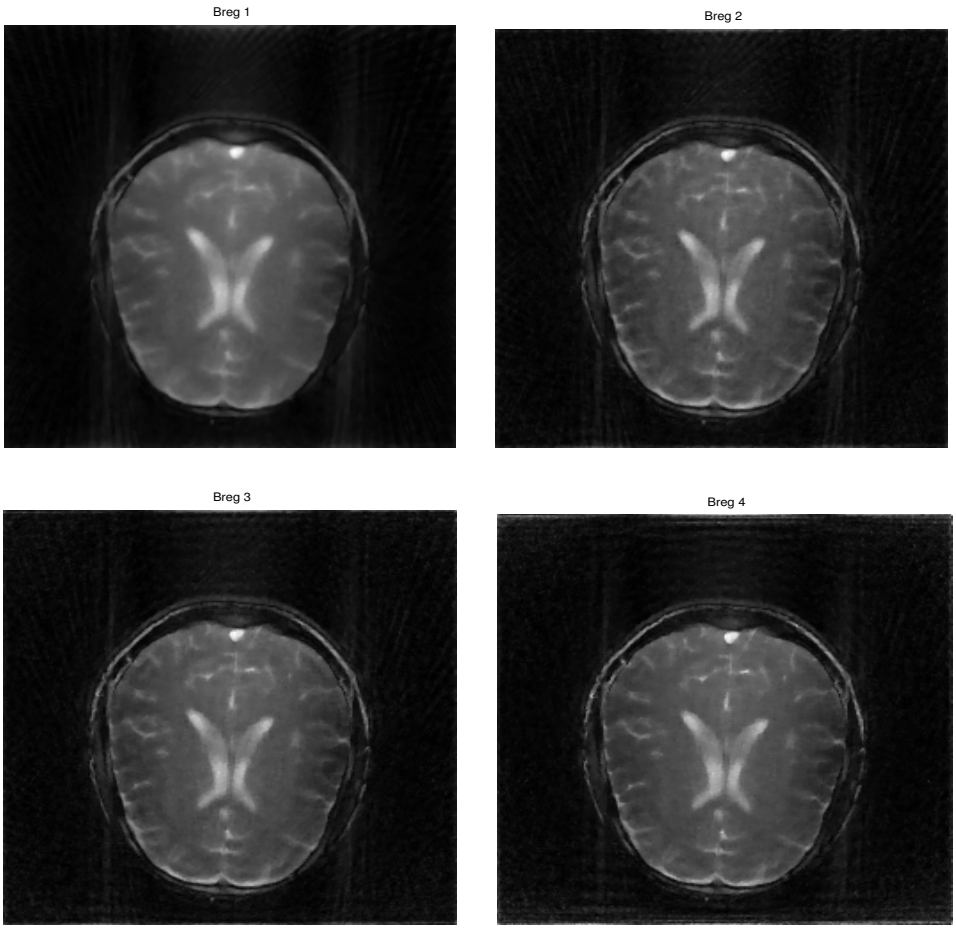

FIG. 5.6. *A sequence images obtained from the iterated refinement method with* (3.2) *where* $\mu = 1.0$, $\nu = 0.1$, *and* $\lambda = 200$*: top left, 1st Bregman iteration; top right, 2nd Bregman iteration; bottom left, 3rd Bregman iteration; bottom right, 4th Bregman iteration.*

which is piecewise smooth, we'd like to find better sparse representations to avoid the over-smooth visual effects brought by the total variation term. Thus, the curvelet + TV could be our future work.

REFERENCES

[1] S. BOYD AND L. VANDENBERGHE, *Convex Optimization*, Cambridge University Press, 2004.
[2] L. M. BREGMAN, *Finding the common point of convex sets by the method of successive projection*, Dokl. Akad. Nauk. USSR, 162 (1965), pp. 487–490.
[3] L. M. BREGMAN, *The relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming*, USSR Comp. Math. and Math. Phys., 7 (1967), pp. 200–217.
[4] M. BURGER, G. GILBOA, S. OSHER, AND J.-J. XU, *Nonlinear inverse scale space methods*, Comm. Math. Sci., 4 (2006), pp. 175–208.
[5] C. BURRUS, R. GOPINATH, AND H. GUO, *Introduction to wavelets and wavelet transformations*, (1998). A Primer, Prentice-Hall.
[6] E. CANDES AND T. TAO, *Near-optimal signal recovery from random projections and universal*

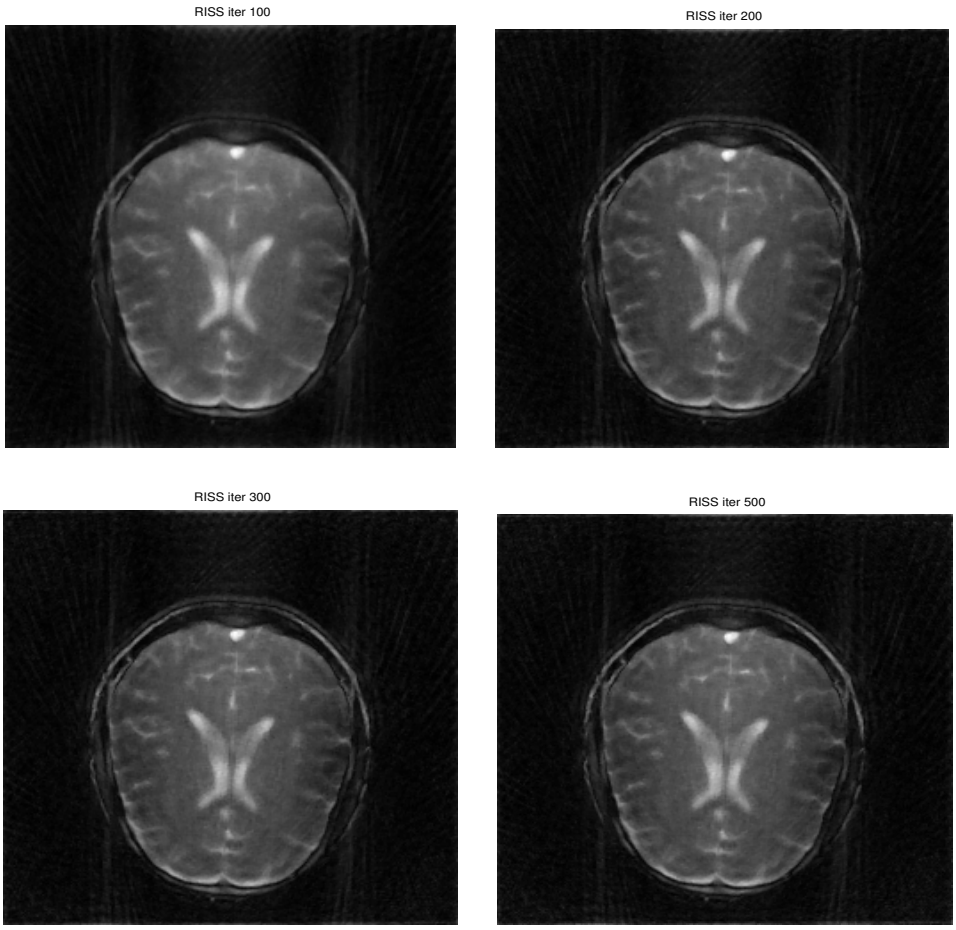

FIG. 5.7. *A sequence images obtained from the relaxed inverse scale space flow with* (3.2) *where* $\mu = 1.0$, $\nu = 0.1$, *and* $\lambda = 200$*: top left, iteration 100; top right, iteration 200; bottom left, iteration 300; bottom right, iteration 500.*

encoding strategies, submitted to IEEE Trans. Inform., (2004).

[7] E. J. CANDES, L. DEMANET, D. L. DONOHO, AND L. YING, *Fast discrete curvelet transforms*. 2005.

[8] E. J. CANDES AND D. L. DONOHO, *New tight frames of curvelets and optimal representations of objects with smooth singularities*. Technical Report, Stanford University, Submitted, 2002.

[9] E. J. CANDES AND F. GUO, *New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction*, Signal Processing, 82 (2002), pp. 1519–1543.

[10] E. J. CANDES AND J. ROMBERG, *Practical signal recovery from random projections*. 2005.

[11] E. J. CANDES, J. ROMBERG, AND T. TAO, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, submitted to IEEE Trans. Inform. Theory, (June 2004).

[12] T. F. CHAN AND S. ESEDOGLU, *Aspects of total variation regularized* L^1 *function approximation*, SIAM J. Appl. Math, 65 (2005), pp. 1817–1837.

[13] T. F. CHAN AND C.-K. WONG, *Total variation blind deconvolution*, IEEE Trans. Image Process., 7 (1998), pp. 370–375.

[14] T.-C. CHANG, L. HE, AND T. FANG, *MR image reconstruction from sparse radial samples using Bregman iteration*, ISMRM, 2006.

[15] S. S. Chen, D. L. Donoho, and M. A. Saunders, *Atomic decomposition by basis pursuit*, SIAM J. Scientific Computing, 20 (1998), pp. 33–61.
[16] J. A. Fessler and B. P. Sutton, *Nonuniform fast fourier transforms using min-max interpolation*, IEEE Trans. Signal Process., 51 (2003).
[17] L. He, M. Burger, and S. Osher, *Iterative total variation regularization with non-quadratic fidelity*, J. of Mathematical Imaging and Vision, (2006).
[18] L. He, A. Marquina, and S. Osher, *Blind deconvolution using TV regularization and Bregman iteration*, International Journal of Imaging Systems and Technology, 15 (2005), pp. 74–83.
[19] S. Kunis and D. Potts, *NFFT2.0.beta.* technical report, University of Lubeck, Germany, 2004.
[20] M. Lustig, J. H. Lee, D. L. Donoho, and J. M. Pauly, *Faster imaging with randomly perturbed undersampled spirals and l1 reconstruction*, in Proc. Of the ISMRM '05, 2005.
[21] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.
[22] Y. Meyer, *Wavelets Algorithms and Applications*, SIAM, Philadelphia, 1993.
[23] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, AMS, Providence, RI, 2001.
[24] S. Osher, M. Burger, D. Goldfarb, J.-J. Xu, and W. Yin, *An iterative regularization method for total variation based image restoration*, Multiscale Modeling and Simulation, 4 (2005), pp. 460–489.
[25] L. I. Rudin, S. J. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
[26] J. R. Shewchuk, *An introduction to the conjugate gradient method without the agonizing pain.* August 1994.
[27] J.-J. Xu and S. Osher, *Iterative regularization and nonlinear inverse scale space applied to wavelet based denoising*, (2006). submitted to IEEE Trans. Image Proc.
[28] W. Yin, D. Goldfarb, and S. Osher, *Image cartoon-texture decomposition and feature selection using the total variation regularized l1 functional*, in Lecture Notes in Computer Science 3752, Variational, Geometric, and Level Set Methods in Computer Vision, Springer, 2005, pp. 73–84.
[29] X.-Q. Zhang and J. Froment, *Total variation based fourier reconstruction and regularization for computerized tomography*, Nuclear Science Symposium and Medical Imaging Conference, 2005.