Robust Kernel Methods for Sparse MR Image Reconstruction

Joshua Trzasko[1], Armando Manduca[1], and Eric Borisch[2]

[1] Department of Physiology and Biomedical Engineering
[2] Magnetic Resonance Research Lab
Mayo Clinic College of Medicine, Rochester, MN, USA
{trzasko.joshua,manduca,borisch.eric}@mayo.edu

Abstract. A major challenge in contemporary magnetic resonance imaging (MRI) lies in providing the highest resolution exam possible in the shortest acquisition period. Recently, several authors have proposed the use of L_1-norm minimization for the reconstruction of sparse MR images from highly-undersampled k-space data. Despite promising results demonstrating the ability to accurately reconstruct images sampled at rates significantly below the Nyquist criterion, the extensive computational complexity associated with the existing framework limits its clinical practicality. In this work, we propose an alternative recovery framework based on homotopic approximation of the L_0-norm and extend the reconstruction problem to a multiscale formulation. In addition to several interesting theoretical properties, practical implementation of this technique effectively resorts to a simple iterative alternation between bilteral filtering and projection of the measured k-space sample set that can be computed in a matter of seconds on a standard PC.

1   Introduction

One of the fundamental limitations of MRI is the linear relation between scan time and the number of measured data samples. With the recent drive towards dynamic imaging as well as high-resolution scans and those covering an extended field-of-view (FOV), the need for shorter examination times to improve patient comfort and improve clinical throughput is substantial. Barring the introduction of additional hardware such as is used in parallel imaging techniques, conventional undersampled MRI reconstruction techniques such as homodyne detection [1] and projections onto convex sets (POCS) [2] are often inherently limited by the Nyquist criterion to a maximum theoretical sampling reduction of 50%. In practice, even less undersampling is typically employed when using these approaches.

Recently, there has been great interest in "compressed sensing" methods for reconstructing MR images from only a small fraction of the complete k-space sample set [3,4,5]. For naturally sparse scenarios such as MR angiography, high-quality recovery has been demonstrated even at up to 80% undersampling,

offering the potential for dramatically reducing clinical scan times. The main limitation of the existing compressed sensing framework lies in the extensive computation that is required to generate a solution, e.g. the reconstruction of a single 256x256 image can take on the order of several hours [6] using standard descent methods and even several minutes using state- of-the-art matrix-based solvers [7]. The considerable computational burden of these techniques precludes their use in clinical practice, especially when considering the extension to 3D data.

In this paper, we consider an alternative formulation of the sparse reconstruction problem that is both theoretically alluring and computationally practical. While compressed sensing methods typically deal with L_1-norm minimization as they are the closest convex approximation to the ideal L_0-minimization problem, we attack the L_0 problem directly using a quasiconvex homotopy scheme. This approximation is closely related to work on robust anisotropic diffusion [8] and, when considering image gradients across multiple scales, kernel regression methods such as the bilateral filter [9,10]. Following these developments, we address the handling of complex MR image data and discuss practical and simple numerical implementation of the technique.

2   Methods

For many MR images, the underlying image structure is piecewise smooth and thus the signal is sparse in the gradient domain. Let Φ represent a k-space measurement matrix such as that defined by the trajectory of a projection reconstruction or spiral-type acquisition after gridding. The goal of the sparse MRI reconstruction problem is to recover an image, f, from only a small subset of Fourier transform samples, $\Phi \hat{f}$. The ideal approach to recovering a signal with limited support involves solving the following combinatorial optimization problem:

$$\min_u \|\nabla u\|_0 \quad s.t. \quad \Phi \hat{u} = \Phi \hat{f}, \tag{1}$$

where $\hat{f}$ denotes the Fourier transform of f and u is the recovered image; however, as (1) is NP-complete, it is computationally intractable except for very small problems. In the recent work by Candès et al. [11] and Donoho [12], it has been shown that, given sufficient gradient sparsity, f can be almost exactly recovered with overwhelming probability by solving the L_1 analog of (1),

$$\min_u \|\nabla u\|_1 \quad s.t. \quad \Phi \hat{u} = \Phi \hat{f}. \tag{2}$$

While few signals are truly sparse in practice, most are compressible within some transform domain, e.g. the spatial gradient of a piecewise-smooth image exhibits exponential decay upon enumeration. In this scenario, the "compressed sensing" paradigm offers reconstruction whose error is comparable to the best possible K-term approximation within the sparsity basis [13].

Although (2) is a convex optimization problem closely related to basis pursuit [14] that can be solved using standard Interior Point methods [14,11,7], many of

these modern numerical solvers rely heavily on matrix-based operators that are not only computationally expensive but also require extensive parameterization. For example, Kim et al. [7] have recently proposed a method for solving (2) based on preconditioned conjugate gradients (PCG) which, while converging in only around 100 iterations, requires several minutes to reconstruct a single 256x256 image. In effect, this class of approaches is neither readily implimentable nor computationally practical and this may preclude their widespread usage in clinical application.

2.1   Robust Error Norms and L_0-Continuation

Although (1) is generally impossible to solve directly, it is the ideal formulation to address for the sparse reconstruction problem whether for MRI or any other application. Begin by realizing that the zero semi-norm can be defined as

$$\|\nabla u\|_0 = \sum_{\Omega} \mathbf{1}\left(|\nabla u| > 0\right), \tag{3}$$

where Ω is the image domain and $\mathbf{1}$ is the indicator function. Consequently,

$$\{\mathbf{1}\left(|\nabla u| > 0\right) = 1\} \Longleftrightarrow \{\exists n \in [1, N] \,|\, |u_{x_n}| > 0\}, \tag{4}$$

where $N = \dim\{\Omega\}$ and u_{x_n} is the partial derivative of u along the n-th dimension. Given (4),

$$\mathbf{1}\left(|\nabla u| > 0\right) \leq \sum_{n=1}^{N} \mathbf{1}\left(|u_{x_n}| > 0\right) \tag{5}$$

follows trivially and a new sparsity semi-norm can be defined as

$$\|\nabla u\|_{0*} = \sum_{\Omega} \sum_{n=1}^{N} \mathbf{1}\left(|u_{x_n}| > 0\right); \tag{6}$$

note that this is essentially just a migration of u to Markovian form.

At this point, (6) is still combinatorial much like $\|\nabla u\|_0$ and thus of little practical use. Suppose a continuous function, ρ, can be defined such that it is homotopic with $\mathbf{1}$ through the limit function, i.e.

$$\lim_{\sigma \to 0} \rho\left(x, \sigma\right) = \mathbf{1}\left(|x| > 0\right). \tag{7}$$

Consequently, (6) can be redefined as

$$\|\nabla u\|_{0*} = \lim_{\sigma \to 0} \sum_{\Omega} \sum_{n=1}^{N} \rho\left(u_{x_n}, \sigma\right) \tag{8}$$

yielding the new reconstruction problem,

$$\min_{u} \|\nabla u\|_{0*} \quad s.t. \quad \Phi\hat{u} = \Phi\hat{f}. \tag{9}$$

Although there is no guarantee of achieving a global minima when using non-convex priors, standard continuation schemes simliar to those developed to numerically handle the discontinuity of the total variation (TV) semi-norm at the origin [15] yield local minima which are more than acceptable in practice.

While one approach to defining ρ includes using a $p < 1$ semi-norm as utilized in [6], the non-differentiability of these functionals requisites need for additional continuation, limiting their practicality within our approach. Alternatively, consider the class of robust error functions known as redescending M-estimators [16]. Two of the more common examples, the Gaussian and Tukey Biweight error functions, are respectively described by

$$\rho(x,\sigma) = 1 - \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \qquad \rho(x,\sigma) = \begin{cases} \frac{3x^2}{\sigma^2} - \frac{3x^4}{\sigma^4} + \frac{x^6}{\sigma^6}, & |x| \leq \sigma \\ 1, & \text{else} \end{cases}. \qquad (10)$$

For both measures in (10), σ is a scale parameter that controls the dilation of the error functions. Unlike traditional error norms such as the various p-norms, the influence of outlier values beyond a threshold determined by σ is reduced. More interestingly, as $\sigma \to 0$, these redescending functions naturally approach the indicator function, and they remain continuous until that limit is reached. We note that the use of such non-convex estimators have been previously employed in the imaging community for denoising [17,8] and deconvolution [18]; however, to our knowledge, this approach has not been utilized for the highly-undersampled reconstruction problem.

2.2   Multiscale Image Sparsity

One approach to discretizing the partial derivatives of u in (9) involves computing the finite differences between u at a point x and its immediate neighbors. If the set of all immediate neighbors is denoted as η, then (8) can be approximated by

$$\|\nabla u\|_{0^*} = \lim_{\sigma \to 0} \sum_{x \in \Omega} \sum_{n \in \eta} \rho\left(u\left(x + \xi_n\right) - u\left(x\right), \sigma\right), \qquad (11)$$

where the vector $\xi_n = n - x$.

Enforcement of image sparsity can subsequently be generalized to address multiple scales of image gradients by simply extending the neighborhood over which finite differences are computed. Spatial proximity can easily be incorporated into (11) through the addition of an auxiliary influence function, namely

$$\|\nabla u\|_{0^*} = \lim_{\sigma \to 0} \sum_{x \in \Omega} \sum_{n \in \eta} \rho\left(u\left(x + \xi_n\right) - u\left(x\right), \sigma\right) \phi\left(|\xi_n|, \kappa\right), \qquad (12)$$

where ϕ is commonly defined by a Gaussian function with scale κ.

2.3   Practical Implementation and Numerical Considerations

For a fixed value of σ, a minima of (12) is given when,

$$\sum_{n \in \eta} \psi\left(u\left(x + \xi_n\right) - u\left(x\right), \sigma\right) \phi\left(|\xi_n|, \kappa\right) = 0, \qquad \forall x \in \Omega \qquad (13)$$

where the influence function $\psi = \rho'$. Letting $g(x) = \psi(x)/x$, (13) can be written in operator form as

$$\left[\sum_{n \in \eta} g\left(u\left(x + \xi_n\right) - u\left(x\right), \sigma\right) \phi\left(|\xi_n|, \kappa\right) \left(\delta\left(x + \xi_n\right) - \delta\left(x\right)\right) \otimes \right] u = 0, \quad (14)$$

where δ is the Kronecker delta function and $\otimes$ is the convolution operator. When considering (14) in the homogeneous form $A(u)u = 0$, and noting that $A(u)$ is non-expansive and strongly connected, (14) can be solved for iteratively using a nonlinear Jacobi or Fixed-Point iteration, resulting in

$$u^{t+1}(x) = \frac{\sum_{n \in \eta} g\left(u^t\left(x + \xi_n\right) - u^t\left(x\right), \sigma\right) \phi\left(|\xi_n|, \kappa\right) u^t\left(x + \xi_n\right)}{\sum_{n \in \eta} g\left(u^t\left(x + \xi_n\right) - u^t\left(x\right), \sigma\right) \phi\left(|\xi_n|, \kappa\right)}. \quad (15)$$

(15) can be interpreted as an iterative zero-order or Nadaraya-Watson type kernel regression estimator, or more familiarly as a bilateral filter [9]. Recently, kernel regression estimators [10] (including higher-order) have been applied to the related problems of super-resolution and deconvolution and shown excellent results. In effect, the derived form in (15) is very promising as there has been extensive work on developing fast implementations and approximations of the bilateral filter. In particular, it has been shown that a separable version of (15) [19], where 1D bilateral filtering is performed sequentially along each data dimension, can provide a dramatic computational speed-up with no substantial degradation of the result.

When dealing with complex image data such as in MRI, addressing $|\nabla u|$ is not as straighforward as for the strictly real case. For our application, we choose to assess sparsity in the real and imaginary data channels separately as mentioned in [4] and is commonly used for MR image denoising. Given (15), the complex multiscale extension of (9) can be solved using the following iterated projection procedure with continuation:

```
Let  û⁰ = Φf̂,  σ >> 0
1.  ℜ{v^{t+1}} = BilateralFilter [ℜ{u^t}, σ]
2.  ℑ{v^{t+1}} = BilateralFilter [ℑ{u^t}, σ]
3.  Φv̂^{t+1} = Φf̂
4.  if ‖v^{t+1} − u^t‖ < tol,  σ = σ * β
    else u^{t+1} = v^{t+1}, go to Step 1.
```

In the above algorithm, $\Re$ and $\Im$ denote the real and imaginary operators, respectively, tol is a threshold indicating when filtering at a given σ-level has numerically reached steady-state, and $\beta \in (0, 1)$ controls the reduction rate of σ in the continuation procedure. The number of iterations can either be specified *a priori* or an intelligent termination scheme can be incorporated when some target level for σ has been achieved.

Note that the presented algorithm is significantly less intimidating than its matrix-based counterparts for the sparse reconstruction problem as it requires little more than a simple filtering and Fourier tranform operation. Consequently, we hope the inherent simplicity and resultant speed of this method will promote its use in practical application.
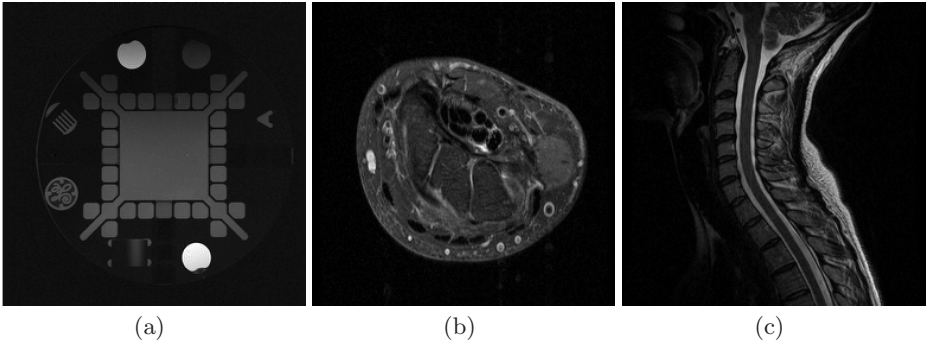
Fig. 1. Example MR images: standard phantom (a), wrist (b), and spine (c)

3   Results

Examples of the presented reconstruction algorithm are given for the three images in Fig. 1. For each case, image intensities were normalized to unity and the following parameterizations were used: $tol = 1e{-}4$, $\beta = 0.5$, $\sigma_0 = 1$, and $iter = 80$. Additionally, bilateral filters were implemented in separable form as described in [19]. On a 3.4 GHz Pentium IV machine with 4GB RAM, a C++ implementation of the reconstruction algorithm using the FFTW library runs at roughly 80ms/iteration for a 256x256 image, yielding a total reconstruction time of around 7s.

The k-space patterns in Fig. 2 imposed 82% (2a), 77% (2b), and 75% (2c) undersampling, respectively. Additionally, the difference between the fully-sampled and reconstructed images, shown in Figs. 2j-l, was quantified using a standard root-mean-square (RMS) measure to yield a per-pixel average intensity error of only 1.687e-5, 2.256e-5, and 1.989e-5, respectively. While some expected textural loss is present in the anatomical image examples, notice that all of the main structural components of the images are accurately recovered as are many of the smaller objects. Some prominent areas of focus showcasing this capability include the comb object inside the physical resolution phantom (Fig. 2j), the carpal tunnel region of the wrist (Fig. 2k), and the walls of the spinal column (Fig. 2l). We note that these results are similar in quality to those obtained by L_1 methods but we do not show a comparison here for sake of brevity.

4   Summary

In this work, we have developed a novel approach to the sparse image reconstruction problem and shown the application of our methods to phantom and clinical MR images. As detailed in Section 2, we offer a new approach at directly handling the L_0-minimization problem and extend the formulation to incorporate multiscale information. Numerically, solving the proposed formulation of the sparse reconstruction problem resorts to a simple iterative scheme based on
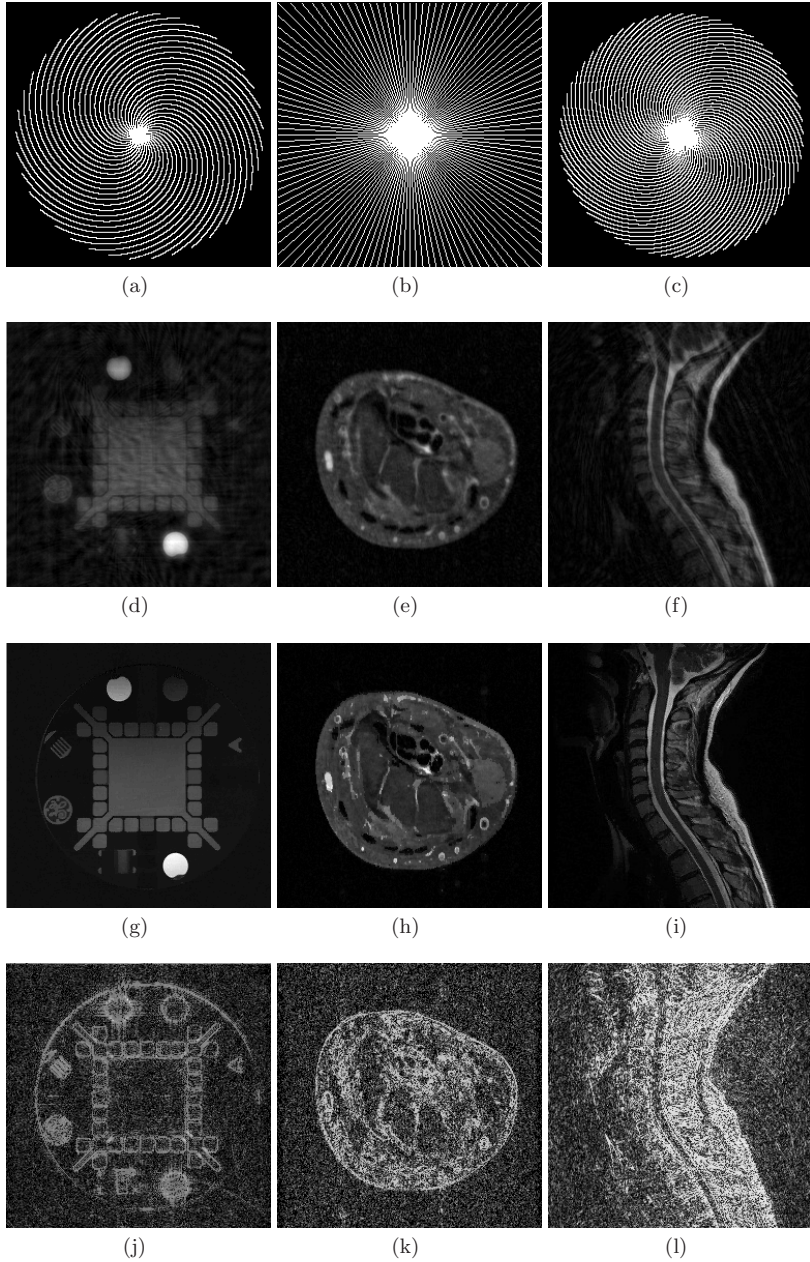
Fig. 2. Example reconstruction results: k-space sampling patterns (a-c), zero-filled reconstructions (d-f), proposed reconstructions (g-i), and reconstruction errors (j-l). Note that images (a-i) are scaled uniformly; images (j-l) have been amplified by 10x. For more details, see Section 3.

alternating between bilateral filtering in image space and reinforcement of the measured Fourier samples in frequency space, two operations which are very computationally efficient and relatively trivial to implement. Consequently, we are able to achieve reconstruction results which are comparable to those of L_1-based methods in a clinically practical amount of time and with relatively little effort spent in the development of the numerical sovler.

References

1. Noll, D., Nishimura, D., Macovski, A.: Homodyne detection in magnetic resonance imaging. IEEE Trans. Med. Imag. 10(2), 154–163 (1991)
2. Haacke, E., Lindskog, E., Lin, W.: A fast, iterative, partial-fourier technique capable of local phase recovery. J. Mag. Res. 92, 125–146 (1991)
3. Lustig, M., Donoho, D., Pauly, J.: Sparse MRI: the application of compressed sensing for rapid MR imaging. Manuscript (2007)
4. He, L., Chang, T., Osher, S., Fang, T., Speier, P.: MR image reconstruction by using the iterative refinement method and nonlinear inverse scale space methods. UCLA CAM Reports 06-35 (2006)
5. Boubertakh, R., Giovanelli, J., Cesare, A.D., Herment, A.: Non-quadratic convex regularized reconstruction of MR images from spiral acquisitions. Sig. Proc. 86, 2479–2494 (2006)
6. Chartrand, R.: Exact reconstruction of sparse signal via nonconvex minimization. Manuscript (2007)
7. Kim, S., Koh, K., Lustig, M., Boyd, S., Gorinevsky, D.: A method for large-scale l1-regularized least squares problems with applications in signal processing and statistics. Manuscript (2007)
8. Black, M., Sapiro, G., Marimon, D., Heeger, D.: Robust anisotropic diffusion. IEEE Trans. Imag. Proc. 7(3), 421–432 (1998)
9. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Proc. IEEE ICIP (1998)
10. Takeda, H., Farsiu, S., Milanfar, P.: Kernel regression for image processing and reconstruction. IEEE Trans. Imag. Proc. 16(2), 349–366 (2007)
11. Candés, E., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Info. Theory 52(2), 489–509 (2006)
12. Donoho, D.: Compressed sensing. IEEE Trans. Info. Theory 52(4), 1289–1306 (2006)
13. Candés, E., Romberg, J., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Comm. Pure Appl. Math. 59, 1207–1223 (2006)
14. Chen, S., Donoho, D., Saunders, M.: Atomic decomposition by basis pursuit. SIAM J.Sci. Comp. 20(1), 33–61 (1998)
15. Chan, T., Zhou, H., Chan, R.: Continuation method for total variation denoising. UCLA CAM Reports 95-28 (1995)
16. Huber, P.: Robust Statistics. Wiley, New York (1981)
17. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Patt. Anal. Mach. Intel. 12(7), 629–639 (1990)
18. Geman, D., Reynolds, G.: Constrained restoration and the recovery of discontinuities. IEEE Trans. Patt. Anal. Mach. Intel. 14(3), 367–383 (1990)
19. Pham, T., van Vliet, L.: Separable bilateral filtering for fast video preprocessing. In: Proc. IEEE ICME (2005)