

SPARSE MRI RECONSTRUCTION VIA MULTISCALE L_0 -CONTINUATION

Joshua Trzasko¹, Armando Manduca¹, and Eric Borisch²

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

ABSTRACT

“Compressed Sensing” and related L_1 -minimization methods for reconstructing sparse magnetic resonance images (MRI) acquired at sub-Nyquist rates have shown great potential for dramatically reducing exam duration. Nonetheless, the non-triviality of numerical implementation and computational intensity of these reconstruction algorithms has thus far precluded their widespread use in clinical practice. In this work, we propose a novel MRI reconstruction framework based on homotopy continuation of the L_0 semi-norm using redescending M-estimator functions. Following analysis of the continuation scheme, the sparsity measure is extended to multiscale form and a simple numerical solver that can achieve accurate reconstructions in a matter of seconds on a standard desktop computer is presented.

Index Terms— Magnetic Resonance Imaging, Sparse Reconstruction, L_0 -minimization, Homotopy

1. INTRODUCTION

The data-intensive nature of many contemporary MRI applications inherently prescribes a lengthening of scan duration which can decrease patient comfort, increase the risk of physiological artifacts within measurements, and reduce clinical throughput. As many MR images are piecewise smooth and thus naturally exhibit sparsity in the gradient domain, it is now accepted that accurate reconstruction of the constituent image structures can be achieved using only a small subset of their Fourier or k-space measurement ensemble [1, 2].

Let Φ denote the Fourier undersampling operator arising from the MRI acquisition process such as a randomized Cartesian or spiral trajectory. Ideally, one wishes to exactly reconstruct a image, f , from only a small subset of Fourier transform samples, $\Phi\hat{f}$. When f is noise-free and $|\text{supp}(\Phi)| > 2 \cdot |\text{supp}(f)|$, it can be recovered exactly by solving the following combinatorial optimization problem:

$$\min_u \|\nabla u\|_0 \quad \text{s.t.} \quad \Phi\hat{u} = \Phi\hat{f}. \quad (1)$$

In the event $\Phi\hat{f}$ is noisy, one can instead solve

$$\min_u \|\nabla u\|_0 \quad \text{s.t.} \quad \|\Phi\hat{u} - \Phi\hat{f}\|_2^2 \leq \epsilon \quad (2)$$

to obtain a reconstruction whose error is proportional to the degree of contamination [3]. Similarly, (2) can be used to estimate an image that is only approximately sparse. In practice, L_0 -minimization is NP-hard and thus computationally intractable except for very small problems. Recently, Candès et al. [1] and Donoho [2] have shown in their groundbreaking work that, at the expense of mild oversampling, exact reconstruction can still be achieved with overwhelming probability by solving the L_1 analog of both (1) and (2) as the L_1 -norm is the closest convex approximation of the L_0 semi-norm albeit a weaker and indirect measure of signal sparsity.

One of the greatest areas of success for “Compressed Sensing” has been in medical imaging, particularly MRI. Quality reconstructions of clinical MR images with up to 80% undersampling have been demonstrated using this theory [4] and related L_1 -minimization techniques [5, 6]. Unfortunately, the computational burden associated with solving the L_1 problems limits their clinical practicality - reconstruction of a single 256×256 image can take several minutes using state-of-the-art Interior Point methods [7, 1, 8] and up to several hours using conventional descent approaches [9]. Resultantly, the extension to 3D or real-time imaging applications with such techniques is clinically impractical.

2. METHODS

2.1. Robust Error Norms and L_0 -Continuation

Let the zero semi-norm of a signal, u , be defined as follows:

$$\|u\|_0 = \sum_{\Omega} \mathbf{1}(|u| > 0), \quad (3)$$

where Ω is the image domain and $\mathbf{1}$ is the indicator function. Suppose there exists a continuous function, ρ , that is homotopic with $\mathbf{1}$ such that

$$\lim_{\sigma \rightarrow 0} \rho(u, \sigma) = \mathbf{1}(|u| > 0), \quad (4)$$

where σ is the deformation variable. While both (1) and (2) are non-convex and thus any non-exhaustive search cannot guarantee achievement of a global minima, if ρ is at least quasiconvex then these problems can be practically addressed

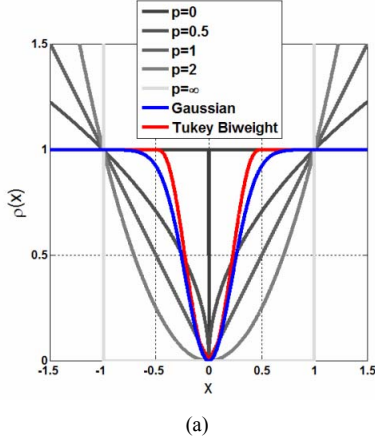


Fig. 1. A visual comparison of P-norms vs robust error norms.

with a basic continuation scheme similar to that in [10] used to address the discontinuity of the L_1 -norm at the origin.

Consider the class of robust error functions known as re-descending M-estimators [11] which includes the Gauss and Tukey Biweight error functions, given by

$$\rho(u, \sigma) = 1 - e^{-\frac{u^2}{2\sigma^2}} \quad (5)$$

and

$$\rho(u, \sigma) = \begin{cases} \frac{3u^2}{\sigma^2} - \frac{3u^4}{\sigma^4} + \frac{u^6}{\sigma^6}, & |u| \leq \sigma \\ 1, & |u| > \sigma \end{cases}, \quad (6)$$

respectively (see Figure 1). Not only are these functions continuous when $\sigma > 0$ but they readily satisfy (4). We point to the success of applying non-convex priors, and specifically re-descending M-estimators, to image processing problems such as denoising [12, 13] and deconvolution [14]. To our knowledge, this is the first application of this class of priors as well as addition of the continuation scheme to the ‘‘compressed sensing’’ image reconstruction problem. We do note that the $p < 1$ semi-norm [9] has been addressed in this context and seemingly might be an intuitive choice for ρ ; however, this functional is not continuous and thus would require intermediary continuation, limiting its practicality within our proposed approach.

2.2. Multiscale Image Sparsity

Consider the Markovian analog of the L_0 variational semi-norm discussed in Section 2.1. Letting $N = \dim\{\Omega\}$ and u_{x_n} be the partial derivative of u along the n -th dimension,

$$\{\mathbf{1}(|\nabla u| > 0) = 1\} \iff \{\exists n \in [1, N] \mid |u_{x_n}| > 0\}. \quad (7)$$

Subsequently, a new measure of image gradient sparsity which penalizes the non-zero partial components of the gradient can

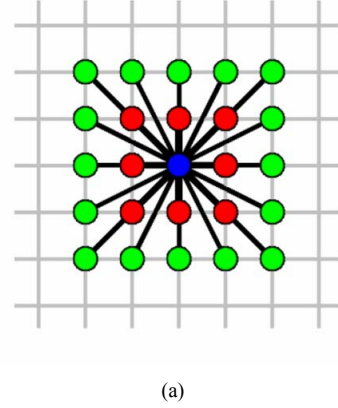


Fig. 2. For a given pixel of interest (blue), gradient estimation is typically performed using the immediate local neighbors (red). In applications such as the bilateral filter, non-immediate neighbors (green) are also assessed.

be defined by

$$\|\nabla u\|_{0^*} = \sum_{\Omega} \sum_{n=1}^N \mathbf{1}(|u_{x_n}| > 0). \quad (8)$$

While we note that (8) is not rotationally invariant, this formulation allows a trivial extension to multiscale form.

Let η be the set of all immediate neighbors of a point $x \in \Omega$. Discretization of u_{x_n} via finite differences and incorporation of (4) yields

$$\|\nabla u\|_{0^*} = \lim_{\sigma \rightarrow 0} \sum_{x \in \Omega} \sum_{n \in \eta} \rho(u(x + \xi_n) - u(x), \sigma), \quad (9)$$

where the vector $\xi_n = n - x$. The generalization to multiscale form can be achieved by simply extending the neighborhood, η , over which finite differences are computed (see Figure 2). Neighbor spatial proximity can also be incorporated into (9) via an auxiliary influence function, ϕ , to yield

$$\|\nabla u\|_{0^*} = \lim_{\sigma \rightarrow 0} \sum_{x \in \Omega} \sum_{n \in \eta} \rho(u(x + \xi_n) - u(x), \sigma) \phi(|\xi_n|, \kappa); \quad (10)$$

in practice, ϕ is usually defined as a Gaussian function with scale κ .

2.3. Practical Numerical Considerations

For a fixed σ ,

$$\sum_{n \in \eta} \psi(u(x + \xi_n) - u(x), \sigma) \phi(|\xi_n|, \kappa) = 0, \quad \forall x \in \Omega \quad (11)$$

yields a minima of (10), where $\psi = \rho'$. Letting $g(x) = \psi(x)/x$, (11) resorts to a stable homogeneous operator form

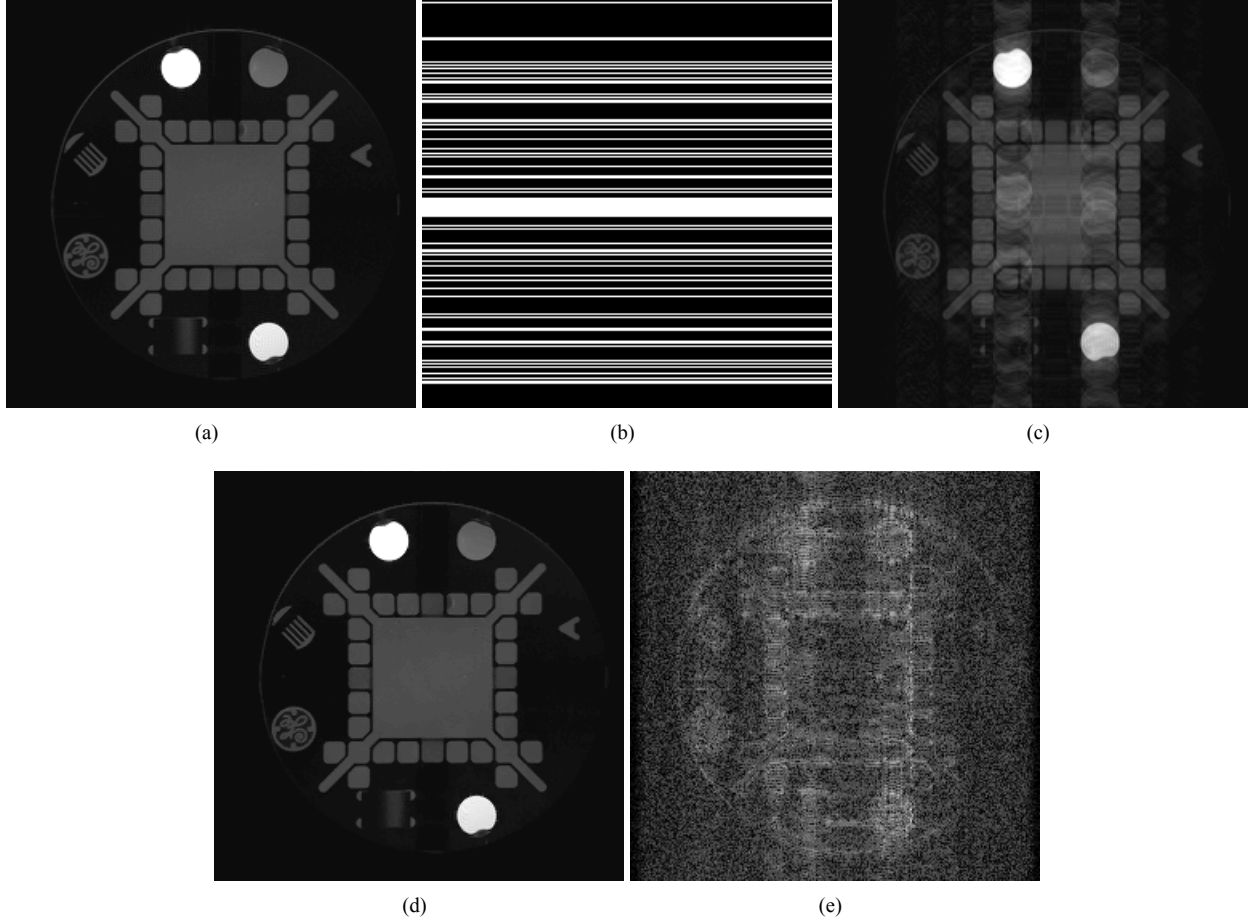


Fig. 3. Example: (a) standard MRI resolution phantom , (b) simulated K-space trajectory (radial, 78% undersampling), (c) minimum-energy reconstruction, (d) proposed reconstruction , and (e) reconstruction error (amplified 10x, RMS = 1.14e-4).

and u can be solved for via nonlinear Jacobi or Fixed-Point iteration to yield

$$u^{t+1}(x) = \frac{\sum_{n \in \eta} g(u^t(x+\xi_n) - u^t(x), \sigma) \phi(|\xi_n|, \kappa) u^t(x+\xi_n)}{\sum_{n \in \eta} g(u^t(x+\xi_n) - u^t(x), \sigma) \phi(|\xi_n|, \kappa)}. \quad (12)$$

which can be interpreted as a zero-order or Nadaraya-Watson type kernel regression estimator, or more familiarly as a bilateral filter [15, 16, 17]. Conveniently, several fast implementations of the bilateral filter have recently been proposed including a separable approximation [18] which practically extends to higher-dimensional data.

As MRI data is inherently complex, gradient sparsity can be addressed in several different manners. We simply follow the approach described in [5] where sparsity is assessed in the real and imaginary channels independently such as is common for many denoising applications. Additionally, note that the constraints in (1) and (2) can both be enforced via projection of the Fourier measurements on Φ to lie inside the

L_2 error ball, namely

$$\Phi \hat{v}^t = \begin{cases} \Phi \hat{v}^t, & E(\hat{v}^t) \leq \epsilon \\ \Phi \cdot \left[\hat{f} + \sqrt{\epsilon} \frac{\hat{v}^t - \hat{f}}{\|\Phi \hat{v}^t - \Phi \hat{f}\|_2} \right], & E(\hat{v}^t) > \epsilon \end{cases}, \quad (13)$$

where $E(\hat{v}^t) = \|\Phi \hat{v}^t - \Phi \hat{f}\|_2^2$; for (1), $\epsilon = 0$. Given (12) and (13), complex multiscale L_0 -minimization can be achieved by the following iterative algorithm:

Let $\hat{u}^0 = \Phi \hat{f}; \sigma \gg 0$

(1) $\Re v^{t+1} = \text{BilateralFilter}[\Re u^t; \sigma]$

(2) $\Im v^{t+1} = \text{BilateralFilter}[\Im u^t; \sigma]$

(3) Project v^{t+1} **into the** L_2 -**error ball**

(4) if $\|v^{t+1} - u^t\|_2 < tol, \sigma = \sigma \times \beta$; **else** $u^{t+1} = v^{t+1}$; **go to Step 1**

In the above algorithm, \Re and \Im denotes the real and imaginary operators, respectively, tol is an error tolerance threshold and $\beta \in (0, 1)$ controls the reduction rate of σ in the continuation procedure. The number of iterations can be specified

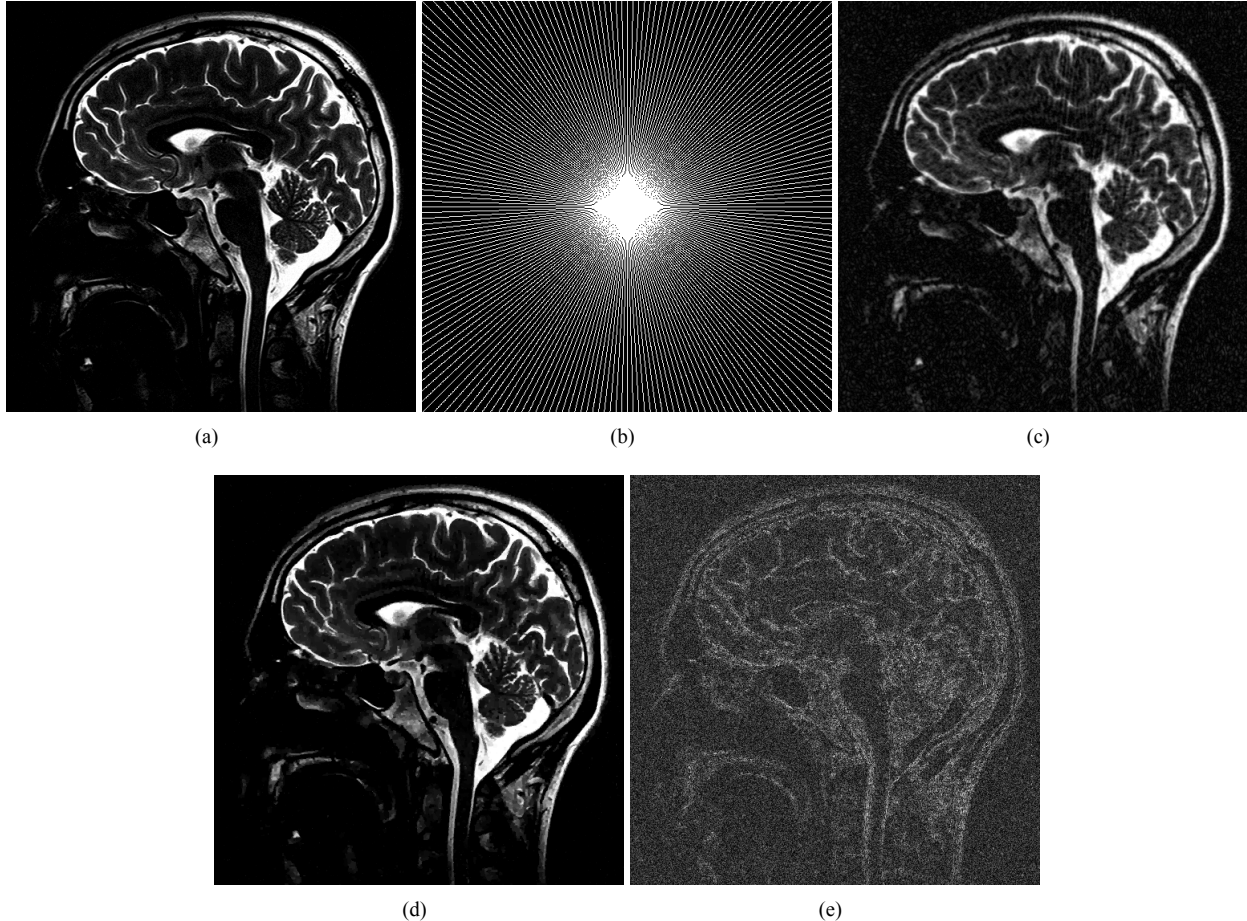


Fig. 4. Example: Sagittal view of T2-weighted MRI of the brain (a), simulated K-space trajectory (Cartesian radial, 77% undersampling), minimum-energy reconstruction (c), proposed reconstruction (d), reconstruction error (amplified 10x, RMS = $9.91e-4$).

a priori or an intelligent termination scheme may be used; at this preliminary stage, we have simply fixed the number of iterations in advance. We also note that, during the numerical continuation process, $\sigma = 0$ is never truly reached; however, the ability to get arbitrarily close to this state is conjectured to have no discernable difference in practice.

3. EXAMPLES

For both presented examples, $tol = 1e-4$, $\beta = 0.5$, $\sigma|_{t=0} = 2 \cdot \max(|u^0|)$, $\epsilon = 1e-4$, $iter = 80$. A C++ implementation of the algorithm in Section 2.3 with the FFTW library and Intel compiler (v9.1) was used and separable bilateral filters were employed as defined in [18]. Additionally, all images were normalized to unity prior to computation. On a single-processor 3.4 GHz Pentium IV machine, the reconstruction runs at roughly 80ms/iteration, yielding a total execution time of 7s; additionally, when a 16-node parallel cluster dedicated for real-time MRI processing is used, sub-second reconstruc-

tions were obtained.

Figure 3 shows the reconstruction of a 256×256 standard resolution phantom after 78% undersampling using a simulated random phase encoding K-space trajectory - encoding lines were chosen from a Gaussian random distribution. Using a simple region of interest (ROI) placed in the image background, ϵ was estimated to be roughly 0.005. While the minimum energy reconstruction obtained by zero-filling clearly has aliasing artifacts, note that our reconstruction has completely removed these components and at no loss of morphological detail. In particular, note the crisp recovery of the comb object and General Electric logo in Figure 3d. Additionally, the root mean square (RMS) error between the fully sampled and reconstructed image was only $1.14e-4$ per pixel, indicating a very accurate recovery.

Figure 4 shows the reconstruction of a 256×256 sagittal view of a T2-weighted MRI of the brain after 77% undersampling using a simulated Cartesian radial K-space trajectory. For this example, $\epsilon = 0.02$. While the reconstructed image

in Figure 4d exhibits a loss of texture from the original fully sampled image in Figure 4a, note the suppression of aliasing artifacts in areas such as the spinal canal and occipital lobe. For the example in Figure 4, the RMS error of the reconstruction was only $9.91e-4$ per pixel.

4. SUMMARY

In this work, we have presented a novel approach to the L_0 -minimization problem and shown its applicability to sparse MRI reconstruction from highly-undersampled Fourier measurements. As the proposed technique not only directly attacks the L_0 semi-norm problem but also incorporates multiscale information, we are able to achieve rapid image reconstructions with minimal implementation overhead, especially compared to many of the contemporary L_1 -norm based method found in practice. Consequently, we hope the presented method will bring sparse MRI reconstruction one step closer to clinical practicality, especially when considering the drive towards 3D and dynamic imaging.

5. REFERENCES

- [1] E. Candés, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Info. Theory*, 52(2):489–509, 2006.
- [2] D. Donoho. Compressed sensing. *IEEE Trans. Info. Theory*, 52(4):1289–1306, 2006.
- [3] E. Candés, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure. Appl. Math.*, 59:1207–1223, 2006.
- [4] M. Lustig, D. Donoho, and J. Pauly. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Manuscript*, 2007.
- [5] L. He, T. 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 Reports*, 06-35, 2006.
- [6] R. Boubertakh, J. Giovanelli, A. De Cesare, and A. Herment. Non-quadratic convex regularized reconstruction of MR images from spiral acquisitions. *Sig. Proc.*, 86:2479–2494, 2006.
- [7] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. A method for large-scale L_1 -regularized least squares problems with applications in signal processing and statistics. *Manuscript*, 2007.
- [8] S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM J.Sci. Comp.*, 20(1):33–61, 1998.
- [9] R. Chartrand. Exact reconstruction of sparse signal via nonconvex minimization. *Manuscript*, 2007.
- [10] T. Chan, H. Zhou, and R. Chan. Continuation method for total variation denoising. *UCLA CAM Reports*, 95-28, 1995.
- [11] P. Huber. *Robust Statistics*. Wiley, New York, 1981.
- [12] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Patt. Anal. Mach. Intel.*, 12(7):629–639, 1990.
- [13] M. Black, G. Sapiro, D. Marimon, and D. Heeger. Robust anisotropic diffusion. *IEEE Trans. Imag. Proc.*, 7(3):421–432, 1998.
- [14] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Trans. Patt. Anal. Mach. Intel.*, 14(3):367–383, 1990.
- [15] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. IEEE ICIP*, 1998.
- [16] H. Takeda, S. Farsiu, and P. Milanfar. Kernel regression for image processing and reconstruction. *IEEE Trans. Imag. Proc.*, 16(2):349–366, 2007.
- [17] M. Elad. On the origin of the bilateral filter and ways to improve it. *IEEE Trans. Imag. Proc.*, 11(10):1141–1151, 2002.
- [18] T. Pham and L. van Vliet. Separable bilateral filtering for fast video preprocessing. In *Proc. IEEE ICME*, 2005.