ITERATIVE IMAGE RECONSTRUCTION IN MRI
WITH SEPARATE MAGNITUDE AND PHASE REGULARIZATION

Jeffrey A. Fessler[*]

EECS Dept., The University of Michigan
fessler@umich.edu

Douglas C. Noll

BME Dept., The University of Michigan
dnoll@umich.edu

ABSTRACT

Iterative methods for image reconstruction in MRI are useful in several applications, including reconstruction from non-Cartesian k-space samples, compensation for magnetic field inhomogeneities, and imaging with multiple receive coils. Existing iterative MR image reconstruction methods are either unregularized, and therefore sensitive to noise, or have used regularization methods that smooth the complex valued image. These existing methods regularize the real and imaginary components of the image equally. In many MRI applications, including T_2^*-weighted imaging as used in fMRI BOLD imaging, one expects most of the signal information of interest to be contained in the magnitude of the voxel value, whereas the phase values are expected to vary smoothly spatially. This paper proposes separate regularization of the magnitude and phase components, preserving the spatial resolution of the magnitude component while strongly regularizing the phase component. This leads to a non-convex regularized least-squares cost function. We describe a new iterative algorithm that monotonically decreases this cost function. The resulting images have reduced noise relative to conventional regularization methods.

1. INTRODUCTION

The most prevalent method for image reconstruction in MRI is to compute the inverse fast Fourier transform (FFT) of the measurements, under the assumption that the measurements are samples of the Fourier transform (k-space data) of the object being imaged. In the case of non-Cartesian k-space samples, usually the FFT is preceded by interpolation (gridding), *e.g.*, [1].

Iterative methods for image reconstruction in MRI are useful in several applications, including reconstruction from non-Cartesian k-space samples [2], compensation for the effects of magnetic field inhomogeneities [3], and reconstructing from multiple receiver coil signals [4]. Existing iterative

image reconstruction methods for MRI are either unregularized, and therefore sensitive to noise, or have used regularization methods that encourage the complex-valued image to be smooth or piece-wise smooth by penalizing some function of the absolute differences between the complex values of neighboring voxels. In essence, existing methods regularize equally the real and imaginary components of the image. In many MRI applications, including T_2^*-weighted imaging as used in fMRI BOLD imaging, one expects most of the signal information of interest to be contained in the magnitude of the voxel value, whereas the phase values are expected to vary smoothly spatially. Indeed, some methods have even constrained the phase to be zero, but this is often unrealistic in practice since there can be residual phase components that are non-negligible but spatially smooth.

In this paper, we propose to regularize the magnitude and phase components separately, aiming to preserve the spatial resolution of the magnitude while strongly regularizing the phase component. We note that in the radar literature, the "opposite" strategy has been employed. The phase in radar imagery is often highly fluctuating, so Cetin *et al.* used a regularization method that smoothed only the magnitude component, leaving the phase component unregularized [5]. The problem that we address is more challenging since the phase term enters the signal model "more nonlinearly" than the magnitude term. In fact, the resulting regularized least-squares cost function is non-convex. Nevertheless, we are able to derive an iterative algorithm that monotonically decreases this cost function, by applying the principle of optimization transfer [6]. That principle has been used frequently in the PET literature for deriving image reconstruction methods with desirable monotone convergence properties *e.g.*, [7–9], but to our knowledge this paper is the first such development in MRI.

In the "homodyning" approach of Noll *et al.* [10], the central portion of k-space was used to estimate a spatially smooth phase map, and then the real image component was reconstructed from (a bit more than) half of the usual k-space data. Depending on the type of k-space data available, there were three primary motivations for that work: reducing noise, reconstruction from partial k-space data, and pre-

serving the sign information in inversion recovery MR images. These motivations also underly the present work, with the additional challenges that we want to compensate for the effects of magnetic field inhomogeneities during the reconstruction, and we want to accommodate non-Cartesian k-space trajectories. The non-iterative approach of [10] would be difficult to extend to these more complicated cases; in contrast, the regularized least-squares (aka MAP) approach can accommodate these effects naturally.

2. THEORY

For completeness, we first describe the "conventional" approach to iterative MR image reconstruction with compensation for the effects of field inhomogeneity and allowing non-Cartesian k-space samples. For simplicity we focus on the case of a single receive coil. The extension to the case of multiple receive coils (*i.e.*, sensitivity encoded imaging [4, 11]) is straightforward for a model-based iterative approach.

We assume that the fieldmap $\omega(\vec{r})$ describing the local off-resonance frequency is known, *i.e.*, is determined from a shifted-echo sequence at the start of the study, as is common practice in many fMRI centers. (Alternatively the fieldmap can be ignored if off-resonance effects are negligible.) The (baseband) signal equation for MRI, including off-resonance effects, is the following:

$$s(t) = \int f(\vec{r}) \, \mathrm{e}^{-\imath \omega(\vec{r})} \, \mathrm{e}^{-\imath 2\pi \vec{k}(t) \cdot \vec{r}} \, \mathrm{d}\vec{r}, \qquad (1)$$

where $\vec{r}$ denotes object-space coordinates (2D or 3D), $f(\vec{r})$ denotes the unknown object magnetization that is to be reconstructed, and $\vec{k}(t)$ denotes the k-space trajectory. (One could also include receive coil sensitivity in the model (1).) We measure N noisy samples of this signal:

$$y_i = s(t_i) + \varepsilon_i, \qquad i = 1, \ldots, N. \qquad (2)$$

The goal is to estimate f from the "k-space data" vector $\boldsymbol{y} = (y_1, \ldots, y_N)$.

This problem is inherently ill-posed since the unknown object f is a continuous-space function yet only a finite number of measurements are available. Therefore we approximate f using a finite-series expansion:

$$f(\vec{r}) \approx \sum_{j=1}^{M} f_j b(\vec{r} - \vec{r}_j),$$

where b denotes a basis function such as a rect function with voxel width, and $\vec{r}_j$ denotes the center of the jth voxel. Substituting this object model into the signal equation (1) yields

$$s(t_i) \approx \sum_{j=1}^{M} a_{ij} f_j,$$

where the following values are known:

$$a_{ij} \triangleq \int b(\vec{r}) \, \mathrm{e}^{-\imath \omega(\vec{r})} \, \mathrm{e}^{-\imath 2\pi \vec{k}(t_i) \cdot \vec{r}} \, \mathrm{d}\vec{r}.$$

In matrix form, we rewrite (2) as follows:

$$\boldsymbol{y} = \boldsymbol{A} \boldsymbol{f} + \boldsymbol{\varepsilon}, \qquad (3)$$

where the $N \times M$ system matrix $\boldsymbol{A}$ has elements $\{a_{ij}\}$ and where $\boldsymbol{f} = (f_1, \ldots, f_M)$ denotes the vector of unknown image voxel values that we wish to reconstruct from $\boldsymbol{y}$.

2.1. Conventional regularized least-squares

Usually $M \approx N$, *i.e.*, the number of unknown voxels is on the order of the number of data points, so typically the system of equations (3) is ill-conditioned if not outright under determined. Therefore, our previous reconstruction methods have used a regularized least-squares cost function of the following form:

$$\hat{\boldsymbol{f}} = \arg \min_{\boldsymbol{f} \in \mathbb{C}^M} \Psi_0(\boldsymbol{f})$$

$$\Psi_0(\boldsymbol{f}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{f}\|^2 + \alpha R(\boldsymbol{f}), \qquad (4)$$

where the penalty function $R(\boldsymbol{f})$ is chosen to control noise by discouraging excessive image roughness, *i.e.*,

$$R(\boldsymbol{f}) \approx \int \|\nabla f(\vec{r})\|^2 \, \mathrm{d}\vec{r}.$$

The regularization parameter α controls the tradeoff between spatial resolution and noise. We have developed a fast and efficient iterative algorithm for minimizing the cost function (4) by combining the conjugate gradient iteration with nonuniform FFT (NUFFT) methods [12] and min-max temporal interpolation [3]. However, the roughness penalty function $R(\boldsymbol{f})$ acts on complex image values $\boldsymbol{f} \in \mathbb{C}^M$, so the real and imaginary components are penalized equally. This yields good images in our experience [3], but they may still have less than the best possible SNR since the *a priori* knowledge that the phase of f is smooth has not been exploited. Alternatively, in the context of sensitivity encoded imaging (SENSE) with multiple receive coils [4, 11], the conventional approach will not yield the maximum possible speedup factor.

2.2. Proposed regularization

To facilitate separate regularization of the magnitude and phase components, we write the image $\boldsymbol{f}$ explicitly in terms of its "magnitude" and phase as follows:

$$f_j = m_j \, \mathrm{e}^{\imath x_j}, \qquad j = 1, \ldots, M. \qquad (5)$$

We allow the "magnitudes" $\boldsymbol{m} = (m_1, \ldots, m_M)$ to include negative values; this is desirable for inversion recovery sequences, and, more generally, we want the phase image $\boldsymbol{x} = (x_1, \ldots, x_M)$ to be smooth spatially, so it should not have to absorb any π jumps that would occur if m_j could not have negative values. Our goal then is to reconstruct both the magnitude map $\boldsymbol{m} \in \mathbb{R}^M$ and the phase map $\boldsymbol{x} \in \mathbb{R}^M$ from the measured data $\boldsymbol{y}$, with minimal or modest regularization of the magnitude map $\boldsymbol{m}$ but substantial regularization of the phase map $\boldsymbol{x}$. To this end we define the following regularized least-squares cost function:

$$\Psi(\boldsymbol{m}, \boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{f}(\boldsymbol{m}, \boldsymbol{x})\|^2 + \beta R(\boldsymbol{x}) + \gamma R(\boldsymbol{m}),$$

where $f_j(\boldsymbol{m}, \boldsymbol{x}) = m_j\, e^{\imath x_j}$ and we intend to choose $\beta \gg \gamma$. We wish to jointly estimate $\boldsymbol{m}$ and $\boldsymbol{x}$ by minimizing this cost function:

$$(\hat{\boldsymbol{m}}, \hat{\boldsymbol{x}}) = \underset{\boldsymbol{m}\in\mathbb{R}^M,\ \boldsymbol{x}\in\mathbb{R}^M}{\arg\min} \Psi(\boldsymbol{m}, \boldsymbol{x}).$$

We follow the typical alternating minimization approach for such optimization problems; after determining initial estimates of $\boldsymbol{m}$ and $\boldsymbol{x}$ using some type of "conventional" (iterative or non-iterative) estimation method for $\boldsymbol{f}$, we then alternate between updating $\boldsymbol{m}$ and $\boldsymbol{x}$ as follows:

$$\boldsymbol{x}^{\mathrm{new}} = \underset{\boldsymbol{x}\in\mathbb{R}^M}{\arg\min} \Psi(\boldsymbol{m}^{\mathrm{old}}, \boldsymbol{x}) \qquad (6)$$

$$\boldsymbol{m}^{\mathrm{new}} = \underset{\boldsymbol{m}\in\mathbb{R}^M}{\arg\min} \Psi(\boldsymbol{m}, \boldsymbol{x}^{\mathrm{new}}), \qquad (7)$$

where we always use the most recent estimates of $\boldsymbol{m}$ and $\boldsymbol{x}$. This approach monotonically decreases the cost function Ψ. Since the cost function Ψ is non-convex, stronger convergence claims than montonicity are difficult to obtain, and typically the algorithm will converge to a local minimizer near the initial point. If the initial point is sufficiently good, then the algorithm should converge to a global minimizer or at least a local minimizer with desirable properties. We use the iterative algorithm described below (4) to initialize the iterations proposed here, so we are confident of having a good initial estimate.

Since $\boldsymbol{f}(\boldsymbol{m}, \boldsymbol{x})$ is linear in $\boldsymbol{m}$, the magnitude update (6) is very easy to implement using our previous conjugate gradient algorithm [3] with a minor modification to account for the fact that the magnitude $\boldsymbol{m}$ is real.

In contrast, the phase minimization problem (6) is more challenging since $\boldsymbol{f}(\boldsymbol{m}, \boldsymbol{x})$ in (5) is a highly nonlinear function of the phase $\boldsymbol{x}$. One could apply a classical gradient-based iteration such as steepest descent to try to update $\boldsymbol{x}$, but such methods require either a 1D line-search, which would be prohibitively expensive computationally for the MRI application, or require a cumbersome quest for an appropriate step size parameter, e.g., [13]. Instead, here we use the optimization transfer principle [6, 14] to derive a

simple update for the phase $\boldsymbol{x}$ that is guaranteed to monotonically decrease Ψ each sub-iteration.

To simplify notation, since we are focusing here on the minimization problem (6) we define $\Phi(\boldsymbol{x}) = \Psi(\boldsymbol{m}^{\mathrm{old}}, \boldsymbol{x})$. We would like to minimize this cost function. Since Φ is difficult to minimize directly, the optimization transfer approach is to find a sequence of surrogate functions $\phi^{(n)}(\boldsymbol{x})$ that must be chosen by the algorithm desinger to satisfy the following *majorization* conditions:

$$\phi^{(n)}\Big(\boldsymbol{x}^{(n)}\Big) = \Phi(\boldsymbol{x}^{(n)})$$

$$\phi^{(n)}(\boldsymbol{x}) \geq \Phi(\boldsymbol{x}), \qquad \forall \boldsymbol{x} \in \mathbb{R}^M,$$

and then to iteratively minimize the surrogates as follows:

$$\boldsymbol{x}^{(n+1)} = \underset{\boldsymbol{x}^{(n)}\in\mathbb{R}^M}{\arg\min} \phi^{(n)}(\boldsymbol{x}). \qquad (8)$$

It follows from the majorization conditions that this approach will decrease Φ (and hence Ψ) monotonically. The "art" lies in choosing suitable surrogate functions $\phi^{(n)}$. The expectation maximization (EM) approach [15] provides one possible recipe for surrogate construction using statistical principles, but those principles provide little insight here. Instead, we adapt and extend techniques that have been used for other tomographic reconstruction problems [7–9].

Space constraints preclude providing a complete derivation. Like many other optimization transfer methods, the final iteration is a type of diagonally preconditioned gradient descent algorithm with a diagonal preconditioner that ensures monotone decreases in the cost function.

3. SIMULATIONS

To date we have implemented a simplified version of the algorithm without field inhomogeneity correction. We simulated data using the analytical Fourier transform of the object shown in Fig. 1 consisting of rectangles. The k-space trajectory was 400 samples in an Archimedian spiral with 10 rotations, i.e., appropriate for a 20×20 image. We reconstructed a 32×28 image, so this is a "partial k-space" problem.

We first ran 40 iterations of the CG algorithm of [3] to form a preliminary estimate of $\hat{f}$. We then applied the proposed algorithm alternating 10 times between the magnitude update (10 subiterations each) and phase update (40 subiterations each). These are all probably overkill.

Fig. 1 shows the results of iterative reconstruction using the conventional regularization approach and that from the proposed method. Both used quadratic regularization with first-order finite differences between each pixel and its 8 neighbors. The phase estimates with the new method are improved substantially thanks to the stronger regularization. The NRMS errors for the magnitude component were 22%
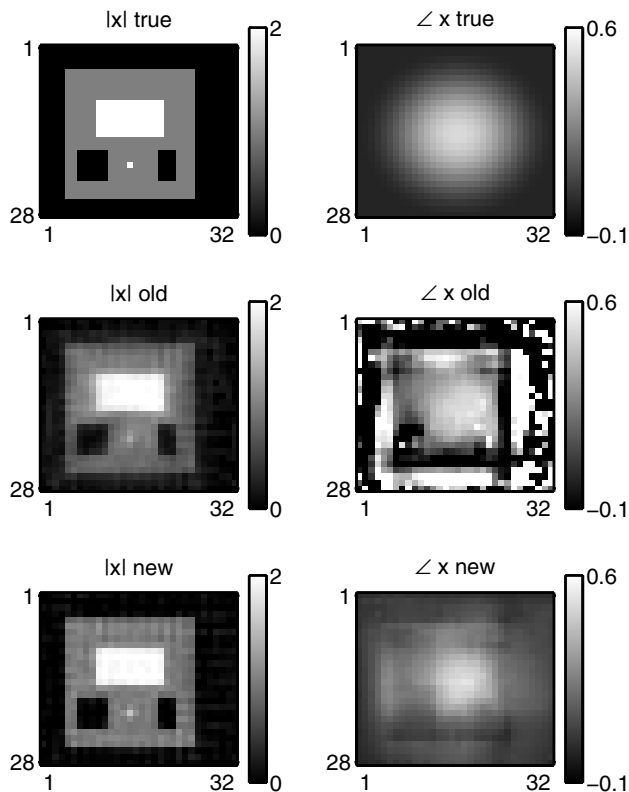
Fig. 1. True magnitude and phase image and reconstructed estimates from under-sampled k-space data.

for the conventional approach and 12% for the proposed approach. Cramer-Rao bound analysis for the case of a non-singular problem with exactly known phase predicts a factor of 2 reduction in variance relative to the case of completely unknown phase.

4. DISCUSSION

There are several interesting potential extensions of the algorithm described here. In this derivation, we used implicitly the same basis functions for the magnitude and phase terms. It could be useful to use wider basis functions for the phase terms since the phase varies slowly spatially. The general principles are also applicable to other problems involving phase estimation, including joint estimation of the image $f(\vec{r})$ and the field map $\omega(\vec{r})$, and other phase retrieval problems.

We plan further investigations that include real data, field inhomogeneity effects, and sensitivity encoded imaging.

5. REFERENCES

[1] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski, "Selection of a convolution function for Fourier inversion using gridding," *IEEE Tr. Med. Im.*, vol. 10, no. 3, pp. 473–8, Sept. 1991.

[2] F. T. A. W. Wajer, R de Beer, M. Fuderer, A. F. Mehlkopf, and D van Ormondt, "Bayesian image reconstruction from an arbitrarily and sparsely sampled K-space without gridding and density correction," in *Proc. Intl. Soc. Mag. Res. Med.*, 1998, p. 667.

[3] B. P. Sutton, D. C. Noll, and J. A. Fessler, "Fast, iterative image reconstruction for MRI with in the presence of field inhomogeneities," *IEEE Tr. Med. Im.*, vol. 22, no. 2, pp. 178–88, Feb. 2003.

[4] K. P. Pruessmann, M. Weiger, P Börnert, and P. Boesiger, "Advances in sensitivity encoding with arbitrary k-space trajectories," *Mag. Res. Med.*, vol. 46, no. 4, pp. 638–51, Oct. 2001.

[5] M. Cetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," *IEEE Tr. Im. Proc.*, vol. 10, no. 4, pp. 623–31, Apr. 2001.

[6] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," *J. Computational and Graphical Stat.*, vol. 9, no. 1, pp. 1–20, Mar. 2000.

[7] A R De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Tr. Med. Im.*, vol. 14, no. 1, pp. 132–137, Mar. 1995.

[8] H Erdoğan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Tr. Med. Im.*, vol. 18, no. 9, pp. 801–14, Sept. 1999.

[9] J. A. Fessler, "Statistical image reconstruction methods for transmission tomography," in *Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis*, M. Sonka and J. Michael Fitzpatrick, Eds., pp. 1–70. SPIE, Bellingham, 2000.

[10] D. C. Noll, D. G. Nishimura, and A. Macovski, "Homodyne detection in magnetic resonance imaging," *IEEE Tr. Med. Im.*, vol. 10, no. 2, pp. 154–63, June 1991.

[11] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Mag. Res. Med.*, vol. 42, no. 5, pp. 952–62, Nov. 1999.

[12] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Tr. Sig. Proc.*, vol. 51, no. 2, pp. 560–74, Feb. 2003.

[13] B. P. Sutton, D. C. Noll, and J. A. Fessler, "Dynamic field map estimation using a spiral-in / spiral-out acquisition," *Mag. Res. Med.*, 2003, Submitted.

[14] K. Lange, *Numerical analysis for statisticians*, Springer-Verlag, New York, 1999.

[15] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.