

A Unified Variational Approach to Denoising and Bias Correction in MR

Ayres Fan¹, William M. Wells III^{2,3}, John W. Fisher III^{1,3}, Mijdat Çetin¹, Steven Haker², Robert Mulkern², Clare Tempany², Alan S. Willsky¹

¹ Laboratory for Information and Decision Systems,
Massachusetts Institute of Technology, Cambridge, MA USA

² Brigham and Women’s Hospital, Harvard Medical School, Boston, MA USA

³ Artificial Intelligence Laboratory, MIT, Cambridge, MA USA

fan@mit.edu

Abstract. We propose a novel bias correction method for magnetic resonance (MR) imaging that uses complementary body coil and surface coil images. The former are spatially homogeneous but have low signal intensity; the latter provide excellent signal response but have large bias fields. We present a variational framework where we optimize an energy functional to estimate the bias field and the underlying image using both observed images. The energy functional contains smoothness-enforcing regularization for both the image and the bias field. We present extensions of our basic framework to a variety of imaging protocols. We solve the optimization problem using a computationally efficient numerical algorithm based on coordinate descent, preconditioned conjugate gradient, half-quadratic regularization, and multigrid techniques. We show qualitative and quantitative results demonstrating the effectiveness of the proposed method in producing debiased and denoised MR images.

1 Introduction

In magnetic resonance (MR) image acquisition, there is a fundamental trade-off between noise and spatially-homogeneous signal response. An uncorrupted image (which we refer to as the true image or the intrinsic image) would depend solely on the underlying tissue and the imaging parameters. Receiving with a body coil (BC) results in low signal-to-noise ratio (SNR) but good spatial homogeneity. Surface coils (SCs) have strong signal response near the coil, but the intensity rapidly diminishes with distance [1]. This variable response allows better visualization of the region of interest (ROI) but results in a systematic intensity inhomogeneity known as the *bias field*. The intensity distortions caused by the bias field can significantly impair both visual inspection and image processing tasks, and separating the bias field from the true underlying image is an under-constrained and ill-posed problem—there are half the number of observations as there are free variables.

The earliest bias correction techniques relied on phantoms [2] or homomorphic unsharp filtering [11], but both methods have severe limitations. Dawant *et*

al. [8] fit thin-plate splines to the bias field using a least-squares penalty. Likar *et al.* [15] compute a parameterized bias field estimate that minimizes the entropy of the reconstructed image. Wells *et al.* [21] exploit the duality behind the segmentation and bias correction problems by using the expectation-maximization (EM) algorithm [9] to alternately segment and debias brain images. Many have improved on this framework including Zhang *et al.* [22] who use a Markov random field to model the bias field. Sled *et al.* [19] sharpen the histogram of the observed image using deconvolution and use the resulting *a priori* density to do Bayes least-squares estimation of the true image.

A few techniques capture a BC image to help correct the SC image. Brey and Narayana [5] estimate the bias field as the ratio of the two low-pass filtered observation images. Lai and Fang [14] estimate the bias field by fitting a membrane model to the ratio of the SC and BC images. Pruessmann *et al.* [17] fit local polynomials at every point in the image.

Our method is related to the imaging framework proposed by Brey and Narayana. We exploit the homogeneity of the BC and the high SNR of the SC to create a composite image that has higher SNR than either observation image and a minimal bias field. We construct a general variational framework which can be adapted to a number of different imaging setups. We introduce a computationally efficient approach to solve the variational problem, and we demonstrate our algorithm on a variety of MR imaging applications.

2 Problem Formulation

2.1 Observation Model

We formulate our observation model in a discrete manner. We place the BC and SC observation image pixels into column vectors \mathbf{y}_B and \mathbf{y}_S respectively. We assume the SC has a bias field \mathbf{b}^* , and the BC has a constant gain field⁴. We stipulate that both observations have the same intrinsic image \mathbf{f}^* :

$$\mathbf{y}_B = \mathbf{f}^* + \mathbf{n}_B \tag{1}$$

$$\mathbf{y}_S = \mathbf{b}^* \circ \mathbf{f}^* + \mathbf{n}_S \tag{2}$$

In the above equation, \circ represents the Hadamard product [12] (or Schür product or entrywise product). Each element of the noise vectors \mathbf{n}_B and \mathbf{n}_S is assumed to be independent and identically distributed (IID). This is justified by the thermal nature of the noise. In the ROI, \mathbf{b}^* tends to be significantly larger than 1 which results in higher SNR for \mathbf{y}_S than \mathbf{y}_B .

We introduce two diagonal matrices \mathbf{B}^* and \mathbf{F}^* which have \mathbf{b}^* and \mathbf{f}^* respectively as their diagonal entries. We can then rewrite (2) as

$$\mathbf{y}_S = \mathbf{B}^* \mathbf{f}^* + \mathbf{n}_S = \mathbf{F}^* \mathbf{b}^* + \mathbf{n}_S \tag{3}$$

⁴ We can only specify \mathbf{f}^* and \mathbf{b}^* up to a multiplicative constant. Generally, \mathbf{f}^* and $2\mathbf{f}^*$ are equivalent. Without loss of generality, we set the gain of the BC to be 1.

The noise in magnitude MR images is accurately modeled by a Rician distribution [16]. Rician random variables are generated by taking the norm of a Gaussian random vector with arbitrary mean. As the SNR increases, the Rician probability density function (PDF) approaches the Gaussian PDF. The Rician PDF is unwieldy to work with, so we treat the noise as Gaussian and zero-mean in our algorithm. Rician noise has a positive mean, so this assumption results in a biased estimator. In most applications, the SNR in tissue regions is high enough so that our Gaussian noise assumption is reasonable, and only a moderate upward bias is imparted.

2.2 Variational Formulation

We formulate a variational problem with a statistical interpretation. This results in an energy functional which we seek to minimize. We do not take the log transform of our observations, but instead pose our energy functional directly in the original multiplicative form. This leads to a cleaner formalism but imposes the need to do nonlinear estimation. We define an energy functional:

$$E(\mathbf{f}, \mathbf{b}) = \lambda_B \|\mathbf{y}_B - \mathbf{f}\|^2 + \lambda_S \|\mathbf{y}_S - \mathbf{b} \circ \mathbf{f}\|^2 + \alpha \|\mathbf{L}\mathbf{b}\|^2 + \gamma \|\mathbf{D}\mathbf{f}\|_p^p \quad (4)$$

and choose our optimal estimates $\hat{\mathbf{f}}$ and $\hat{\mathbf{b}}$ as the vectors that minimize $E(\mathbf{f}, \mathbf{b})$:

$$\hat{\mathbf{f}}, \hat{\mathbf{b}} = \arg \min_{\mathbf{f}, \mathbf{b}} E(\mathbf{f}, \mathbf{b}) . \quad (5)$$

λ_B , λ_S , γ , and α are positive weights. $\|\cdot\|_p$ represents the ℓ_p norm, and $\|\cdot\|$ represents the ℓ_2 norm. We design \mathbf{L} and \mathbf{D} to approximate derivative operators (generally either gradient or Laplacian operators) as finite differences.

The ℓ_2 norms for our data fidelity terms (the first two terms) in (4) imply a Gaussian noise assumption if the problem is formulated as a maximum *a posteriori* (MAP) estimation problem. From this perspective, we see that the scalar weights λ_B and λ_S should be proportional to the inverse noise variances for each observation image. We use Tikhonov-type regularization to make our intrinsic image and bias field estimates conform to our prior knowledge of the signals [10]. Specifically, we ensure that our bias field estimate is smooth and our intrinsic image estimate is piecewise constant. The regularization on $\hat{\mathbf{f}}$ is similar to putting an anisotropic edge-preserving filter into our method. It is well known that ℓ_2 norms tend to overpenalize large derivative values associated with edges. Hence, using ℓ_2 regularization in image reconstruction tends to oversmooth edges, and ℓ_p norms with $p < 2$ are said to be edge preserving.

2.3 Extension to Multiple SCs

Multiple SC images can be simultaneously captured using carefully crafted coil arrays without requiring additional image acquisition time [18]. Multiple coils are used due to the typically sharp drop-off in sensitivity far away from SCs. By distributing the coils spatially, we achieve better signal coverage. One way to process multiple SC images is to combine them into one composite SC image

using a method such as Roemer’s sum-of-squares technique [18] and then use our formulation in (4). However, there are advantages to processing the SC measurements individually. We introduce a new measurement model where we receive one BC image and K SC images:

$$\mathbf{y}_B = \mathbf{f}^* + \mathbf{n}_B \quad (6)$$

$$\mathbf{y}_{S,k} = \mathbf{b}_k^* \circ \mathbf{f}^* + \mathbf{n}_{S,k} \quad (1 \leq k \leq K) . \quad (7)$$

We can extend (4) to handle this more general case:

$$E = \lambda_B \|\mathbf{y}_B - \mathbf{f}\|^2 + \sum_{k=1}^K \lambda_{S,k} \|\mathbf{y}_{S,k} - \mathbf{b}_k \circ \mathbf{f}\|^2 + \sum_{k=1}^K \alpha_k \|\mathbf{L}_k \mathbf{b}_k\|^2 + \gamma \|\mathbf{D}\mathbf{f}\|_p^p . \quad (8)$$

We obtain superior results minimizing (8) because we can optimally combine the SC observations by waiting until we have each $\hat{\mathbf{b}}_k$. Additionally, with the composite SC image, α and \mathbf{L} are determined by the least homogeneous coil response. Processing the SC images individually allows us to choose α_k and \mathbf{L}_k to individually tune the regularization for each coil.

2.4 Extension to Multiple Pulse Sequences

Multiple scans of the same location using different pulse sequences (*e.g.*, T_1 -weighted and T_2 -weighted) are commonly acquired. The bias fields in all of the SC images are nearly identical, so we can achieve satisfactory results using only one BC image. Our measurement model for this case again involves one BC image and K SC images, but this time each SC image has the same bias field but different intrinsic images:

$$\mathbf{y}_B = \mathbf{f}_1^* + \mathbf{n}_B \quad (9)$$

$$\mathbf{y}_{S,k} = \mathbf{b}^* \circ \mathbf{f}_k^* + \mathbf{n}_{S,k} \quad (1 \leq k \leq K) . \quad (10)$$

Without loss of generality, we have assigned \mathbf{f}_1^* to correspond to the intrinsic image in the BC image. We can again generalize (4) to handle this case:

$$E = \lambda_B \|\mathbf{y}_B - \mathbf{f}_1\|^2 + \sum_{k=1}^K \lambda_{S,k} \|\mathbf{y}_{S,k} - \mathbf{b} \circ \mathbf{f}_k\|^2 + \alpha \|\mathbf{L}\mathbf{b}\|^2 + \sum_{k=1}^K \gamma_k \|\mathbf{D}_k \mathbf{f}_k\|_p^p . \quad (11)$$

Additionally, more complex permutations beyond the two extensions we have presented can also be handled (*e.g.*, M pulse sequences captured with N SCs).

3 Solution of the Optimization Problem

This section details the solution to the optimization problem defined in Sec. 2. We will only describe the solution to (4). Extensions for (8) and (11) as well as 3D volumes are straightforward. A closed-form solution for (4) does not exist, and gradient descent on the full energy functional is slow and cumbersome. Therefore, we minimize (4) using coordinate descent. This is an iterative technique that

alternately minimizes the energy for \mathbf{f} and \mathbf{b} . This results in estimates $\hat{\mathbf{f}}^{(i)}$ and $\hat{\mathbf{b}}^{(i)}$ at each iteration i . Coordinate descent is useful in problems where computing solutions over all of the variables is difficult, but computing solutions over a subset is relatively easy. At each iteration, we refer to the computation of $\hat{\mathbf{f}}^{(i)}$ and $\hat{\mathbf{b}}^{(i)}$ as a f-step and a b-step respectively. A stationary point obtained through coordinate descent is also a stationary point of the overall minimization problem. In order for coordinate descent to terminate, the derivative for each coordinate must be zero. Thus the gradient of the complete energy functional is zero.

3.1 Bias Field Solution

For a given \mathbf{f} , (4) is quadratic in terms of \mathbf{b} . Thus setting the gradient of E with respect to \mathbf{b} equal to zero results in a simple linear equation:

$$(\lambda_S \mathbf{F}^2 + \alpha \mathbf{L}_b^T \mathbf{L}_b) \hat{\mathbf{b}}^{(i)} = \lambda_S \mathbf{F} \mathbf{y}_S \quad (12)$$

Although we could solve (12) by direct matrix inversion, we note that $(\lambda_S \mathbf{F}^2 + \alpha \mathbf{L}_b^T \mathbf{L}_b) \geq 0$, so the subproblem is convex. Hence we can use an iterative algorithm such as preconditioned conjugate gradient [3] to efficiently compute solutions. We use as a preconditioner the tridiagonal matrix composed of the main diagonal and the adjacent subdiagonals of $(\lambda_S \mathbf{F}^2 + \alpha \mathbf{L}_b^T \mathbf{L}_b)$ in order to make our preconditioners easy to construct and apply.

3.2 Intrinsic Image Solution

No Regularization on $\hat{\mathbf{f}}$ To provide some insight, we examine the minimization of (4) for a given \mathbf{b} and $\gamma = 0$. We take the gradient of E with respect to \mathbf{f} and set it equal to zero to obtain a pointwise solution at each pixel index n :

$$\hat{\mathbf{f}}^{(i)}[n] = \frac{\lambda_B \mathbf{y}_B[n] + \lambda_S \mathbf{b}[n] \mathbf{y}_S[n]}{\lambda_B + \lambda_S \mathbf{b}^2[n]} \quad (13)$$

Because λ_B and λ_S are related to the inverse noise variances, $\hat{\mathbf{f}}^{(i)}[n]$ is the noise-weighted convex combination of $\mathbf{y}_B[n]$ and $\mathbf{y}_S[n]/\mathbf{b}[n]$ with a spatially varying weighting factor. In contrast, Brey and Narayana [5] only use the data from \mathbf{y}_S to construct $\hat{\mathbf{f}}$. This works well when $\mathbf{b}[n] \gg 1$, but in regions where the SC response is weak, using both observation images can be advantageous.

Half-Quadratic Solution We now describe the f-step for a general γ . When $p \neq 2$, the optimization problem for \mathbf{f} with a given \mathbf{b} is non-quadratic, and we obtain a nonlinear condition for the minimum. The ℓ_p norm for $p \leq 1$ is non-differentiable at zero, so we use a smoothed approximation:

$$\|\mathbf{x}\|_p^p \approx \sum_n (\mathbf{x}^2[n] + \xi)^{p/2} . \quad (14)$$

As $\xi \rightarrow 0$, the approximation approaches the unsmoothed norm.

Half-quadratic optimization is a fixed-point iterative scheme pioneered by Geman and Reynolds [10] that constructs a weighted- ℓ_2 approximation at each sub-iteration j . It has been demonstrated [20] that half-quadratic optimization provides superior convergence rates compared with gradient descent. Using half-quadratic optimization results in a linear condition on $\hat{\mathbf{f}}^{(i,j)}$:

$$\left(\lambda_B \mathbf{I} + \lambda_S \mathbf{B}^2 + \gamma \mathbf{D}^T \mathbf{W}^{(i,j)} \mathbf{D}\right) \hat{\mathbf{f}}^{(i,j)} = \lambda_B \mathbf{y}_B + \lambda_S \mathbf{B} \mathbf{y}_S \quad (15)$$

with the weighting matrix $\mathbf{W}^{(i,j)}$ being diagonal with the following entries:

$$\mathbf{W}^{(i,j)}[n, n] = \frac{p}{2} \left(((\mathbf{D} \hat{\mathbf{f}}^{(i,j-1)})[n])^2 + \xi \right)^{p/2-1}. \quad (16)$$

This preserves edges by weighting the ℓ_2 norm less in regions with large derivatives. Equation (15) is a positive definite linear system which we can again solve using preconditioned conjugate gradient.

One of the key features of (15) is that the effective amount of regularization is spatially varying—less smoothing is performed in regions where \mathbf{B} is large. This is superior to applying an anisotropic post-processing filter to the output of our algorithm. Depending on the regularization strength, post-processing will either oversmooth in high SNR regions or undersmooth in low SNR regions.

3.3 Convergence and Speed

The energy functional E in (4) is non-convex due to the cross-multiplication between \mathbf{b} and \mathbf{f} . Our algorithm possesses convergence qualities similar to the EM algorithm [9]. Each f- and b-step decreases the energy, so our algorithm will at least find a local minimum of E . In practice, we have found excellent convergence properties with the algorithm converging to identical reasonable solutions for random initializations.

Multigrid techniques [6] can help avoid local minima and improve computation speed for large problems. We use a basic form of multigrid with a single coarse-to-fine sweep. We downsample our data to the coarsest level we wish to process. We then run our coordinate descent solver at this level and upsample the results to the next finest level. This cycle repeats until we have a solution at the original scale. The key advantage of multigrid is that the low-frequency components of the solution can be more efficiently computed at the coarser scales.

3.4 Parameter Selection and Initialization

There are a number of parameters that need to be set in our energy functional: λ_B , λ_S , α , γ , and p . We generally use $p = 1$ because it is the smallest value of p that allows the f-step to remain convex. Ideally, we would specify α and γ based on training data (e.g., phantom scans of the SC profiles, long acquisition-time BC images). In practice, we choose the parameters based on subjective visual assessment of the results. Because we use an iterative solver, we must specify initial values for both $\hat{\mathbf{f}}^{(i)}$ and $\hat{\mathbf{b}}^{(i)}$. The convergence speed of our solver can be

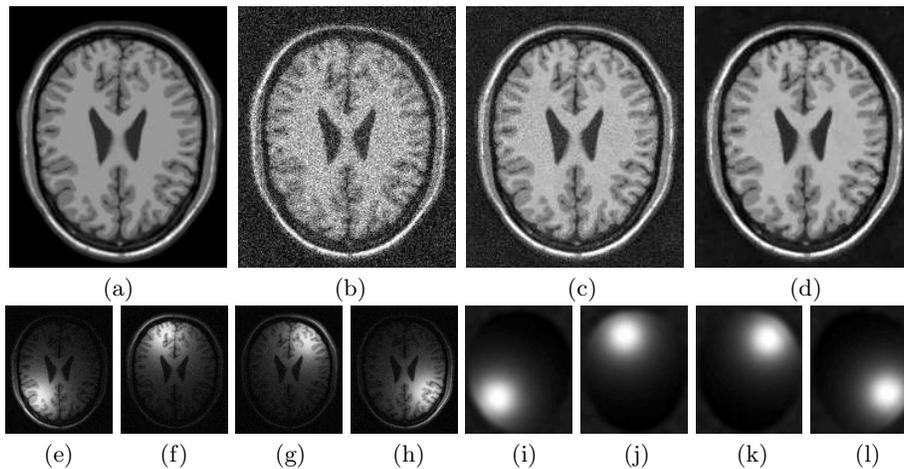


Fig. 1. Synthetic axial T_1 -weighted brain images. (a) True image (f^*). (b) BC image (y_B). Estimated intrinsic image (\hat{f}) computed with (c) Brey-Narayana and (d) proposed method using $\gamma = 0.014$. (e)–(h) SC images ($y_{S,1}$ – $y_{S,4}$). (i)–(l) Estimated bias fields (\hat{b}_1 – \hat{b}_4). $\alpha_k = 2000$. Convergence in 63 sec.

greatly impacted by these choices. We use the bias correction method of Brey and Narayana to produce simple and effective initializations.

We stated that λ_B and λ_S should be related to the inverse noise variances of y_B and y_S respectively. We can estimate the noise variances directly from the images using the method from Nowak [16]. The true signal should be uniformly zero in air-filled regions, so the second moment of y_B in these regions should then equal $2\sigma_B^2$. We can approximate the expected value by taking the sample average over a large air-filled region to obtain σ_B^2 . Note that the bias field has no effect in air-filled regions, so we can perform this same technique for y_S .

When using a multigrid solver, we fix λ_B , λ_S , α , and γ at the original scale. We must also choose parameters at each scale s so that the solutions at the coarser and original scales are similar. The λ 's should scale by 4^s (or 8^s in 3D) due to noise reduction from spatial averaging. For wavelet-based reconstruction, others have found that multiplicative scaling of the regularization parameters is effective [4]. Hence we multiply α and γ at each scale s by experimentally determined scalars k_1^s and k_2^s respectively.

4 Results

In this section, we demonstrate results on real and synthetic data. All real data in this section were captured on General Electric Signa 1.5-T machines. We computed results on a Pentium 4 1.8 GHz workstation using our multigrid solver and stopped the algorithm when the energy changed by less than 0.01%. Convergence times are indicated in the figure captions. For all results, we use Laplacian

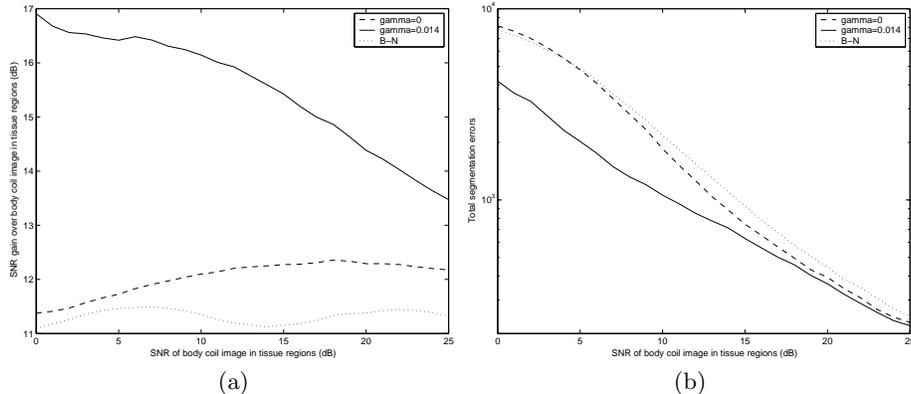


Fig. 2. Performance provided by Brey-Narayana and proposed correction method with varying SNR levels. (a) SNR gain over the BC image (\mathbf{y}_B). (b) Total GM and WM segmentation errors. Averaged over 10 Monte Carlo trials.

	\mathbf{y}_B	B-N	$\hat{\mathbf{f}}, \gamma = 0$	$\hat{\mathbf{f}}, \gamma = 0.014$
MSE (tissue)	196,542	20,820	20,428	10,901
MAE (tissue)	353.91	113.59	112.64	81.99
GM errors	64.4%	14.6%	14.3%	9.9%
WM errors	24.7%	3.6%	3.3%	2.4%

Table 1. Quantitative comparisons using the MNI brain phantom. Corrected images are generated using Brey-Narayana and the proposed method with $\gamma = 0$ and $\gamma = 0.014$. The first two lines are the mean squared error and mean absolute error (based on the true image \mathbf{f}^*) computed only in tissue regions. The last two lines are the percentage of misclassified points in GM and WM regions. Results averaged over 20 random trials.

regularization on $\hat{\mathbf{b}}$ and gradient regularization on $\hat{\mathbf{f}}$. The numerical values of γ presented in this section are not very informative because of scaling variations in the examples. The α values are a measure of relative smoothness because the bias field is unchanged if \mathbf{y}_B and \mathbf{y}_S are both scaled equally.

We begin with synthetic results using the Montreal Neurological Institute (MNI) [7, 13] BrainWeb simulator. We used the T_1 -weighted images with 1 mm slice thickness and constructed synthetic bias fields that simulate a four-coil phased array. We then added Rician noise to obtain our BC and SC images. For tissue regions of \mathbf{y}_B , the noise resulted in SNR of 13 dB and a bias of 2-3%. Estimates were computed within our multiple SC framework by minimizing (8).

We present the observation and corrected images in Fig. 1. The bias field estimates are largely independent of the tissue, and our method produces a $\hat{\mathbf{f}}$ with noticeably superior noise properties than Brey-Narayana. These visual impressions are confirmed with our quantitative results in Table 1 with mean squared error 48% lower than Brey-Narayana. Segmentation accuracy is another way to quantify the quality of the bias correction. We generated gray matter (GM) and white matter (WM) segmentation results using a thresholding scheme

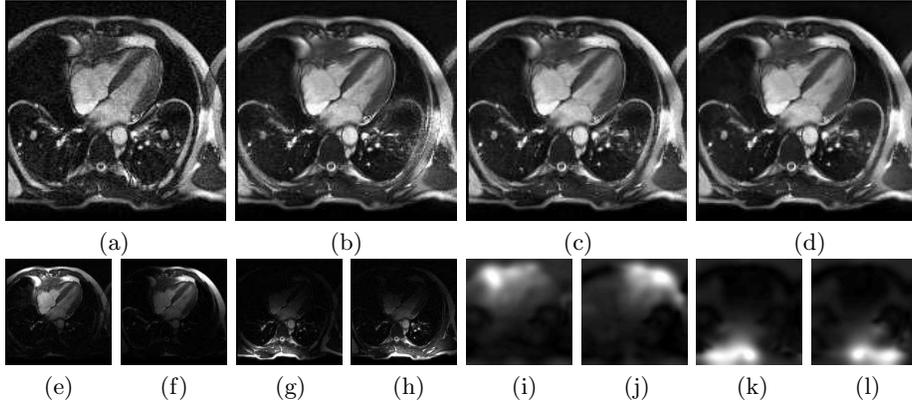


Fig. 3. Gated cardiac MR images. (a) BC image (\mathbf{y}_B). Estimates of the intrinsic image ($\hat{\mathbf{f}}$) using (b) Brey-Narayana and proposed method with (c) $\gamma = 0$ and (d) $\gamma = 1800$. (e)–(h) SC images ($\mathbf{y}_{S,1}$ – $\mathbf{y}_{S,4}$). (i)–(l) Estimated bias fields ($\hat{\mathbf{b}}_1$ – $\hat{\mathbf{b}}_4$). $\alpha_k = 3000$. Convergence in 71 sec.

with manual skull peeling on \mathbf{f}^* and the corrected images. Compared with Brey-Narayana, we reduce overall segmentation error by 33%.

In Fig. 2, we show how the different bias correction schemes function as the SNR is varied. Our method with $\gamma = 0$ consistently outperforms Brey-Narayana due to better bias field estimates. In high SNR regions, all methods provide similar results. As the SNR is decreased, our method with regularization on $\hat{\mathbf{f}}$ builds up a significant advantage over the other methods. At 0 dB, Brey-Narayana and our method with $\gamma = 0$ produce segmentation error rates of 51% (which is approximately equivalent to random guessing), while using regularization on $\hat{\mathbf{f}}$ reduces the error to 27%.

Next, we apply our algorithm to one time step from a gated cardiac MR sequence in Fig. 3. For the SC images, a four-element phased array was used. The images have a field of view (FOV) of $32 \text{ cm} \times 32 \text{ cm}$, resolution of 160×192 , and slice thickness of 8 mm. To obtain our results, we applied our multiple SC correction framework and minimized (8). The main differences between the Brey-Narayana estimate in Fig. 3(b) and our result in Fig. 3(c) are in regions where none of the SCs have good response such as the middle and the right-hand side of the image. This is because our method uses the BC information while Brey-Narayana does not. Fig. 3(d) (using $\hat{\mathbf{f}}$ regularization) is moderately better than Fig. 3(c), but the high SNR in \mathbf{y}_B and $\mathbf{y}_{S,k}$ limit the benefits of filtering.

In Fig. 4, we display the results of our algorithm on a real prostate image. The SCs used were an endorectal coil along with a four element pelvic phased-array coil. We captured T_2 -weighted images using the BC and SCs and T_1 -weighted images using just the SCs. The T_1 -weighted images do not show the internal structure of the prostate but are useful in finding the borders of the gland; the T_2 -weighted images are useful for differentiating regions of the prostate as well

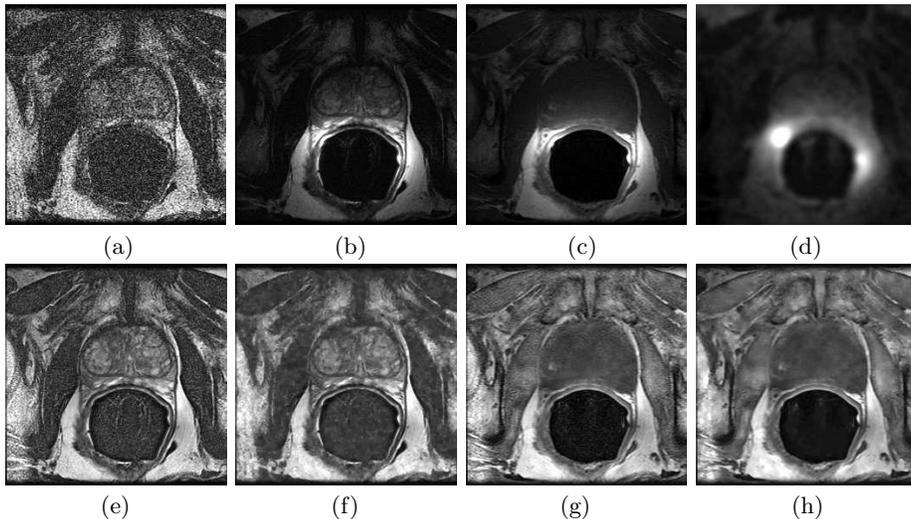


Fig. 4. Prostate images. T_2 -weighted (a) BC (\mathbf{y}_B) and (b) composite SC ($\mathbf{y}_{S,1}$) images. (c) Composite SC T_1 -weighted image ($\mathbf{y}_{S,2}$). (d) Estimated bias field ($\hat{\mathbf{b}}$). T_2 -weighted intrinsic image estimates ($\hat{\mathbf{f}}_1$) using (e) Brey-Narayana and (f) proposed method with $\gamma_1 = 0.018$. T_1 -weighted intrinsic image estimates ($\hat{\mathbf{f}}_2$) using (g) Brey-Narayana and (h) proposed method with $\gamma_2 = 0.010$. $\alpha = 125$. Convergence in 24 sec.

as for tumor detection. The FOV is $12 \text{ cm} \times 12 \text{ cm}$, resolution is 256×256 , and slice thickness is 3 mm. Estimates were computed by minimizing (11) using composite SC images because individual SC data were not available to us.

The prostate is the most challenging example we consider here. The FOV is small so \mathbf{y}_B has very low SNR (about 7 dB in the prostate). To compensate, the endorectal coil produces a strong local response profile which results in a severe intensity artifact. Because the reception profile of the endorectal coil is much less homogeneous than that of the pelvic phased-array coil, the prostate would probably benefit significantly from processing each coil separately. Fig. 4(d) shows that using a composite surface coil image causes $\hat{\mathbf{b}}$ to be under-regularized in regions where the endorectal coil does not dominate. Figs. 4(e)–(h) (when viewed under sufficiently high resolution) demonstrate that our method preserves edges while resulting in lower noise than Brey-Narayana. Fig. 4(h) shows that even without a BC image, we can obtain reasonable intrinsic image estimates for the T_1 -weighted sequence.

We show axial brain images in Fig. 5. The SCs were a four-element phased array. We captured gradient-recalled echo (GRE) images using both the BC and SCs and fluid attenuated FLAIR images using the SCs. FOV is $24 \text{ cm} \times 24 \text{ cm}$, resolution is 192×256 , and slice thickness is 3 mm. We minimize a hybrid of (8) and (11) to obtain our results. All of the SCs are weak in the middle of the brain, so our final estimate of the FLAIR image in Fig. 5(h) is still noisy in the

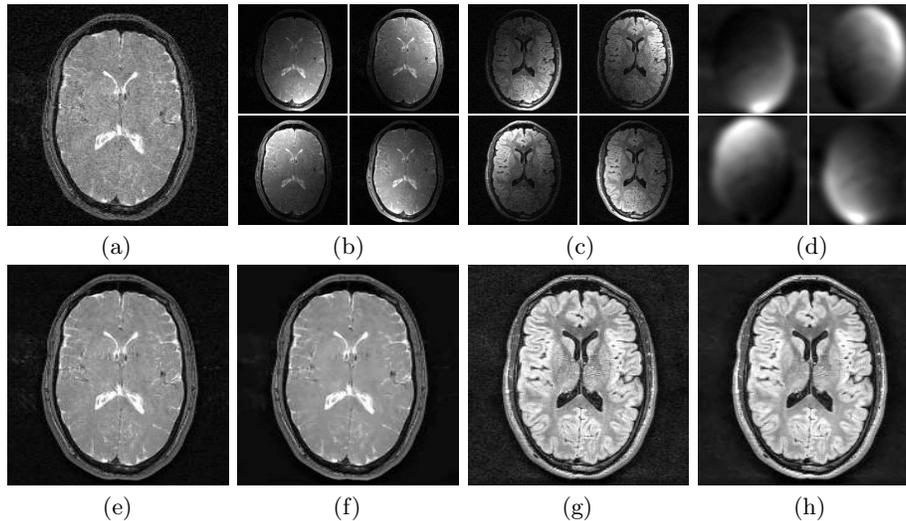


Fig. 5. Axial brain images. (a) GRE BC image (\mathbf{y}_B). (b) GRE SC images ($\mathbf{y}_{S,1}-\mathbf{y}_{S,4}$). (c) FLAIR SC images ($\mathbf{y}_{S,5}-\mathbf{y}_{S,8}$). (d) Estimated bias fields ($\hat{\mathbf{b}}_1-\hat{\mathbf{b}}_4$). Estimated GRE intrinsic images ($\hat{\mathbf{f}}_1$) using (e) Brey-Narayana and (f) proposed method with $\gamma_1 = 1000$. Estimated FLAIR intrinsic images ($\hat{\mathbf{f}}_2$) using (g) Brey-Narayana and (h) proposed method with $\gamma_2 = 1200$. $\alpha_k = 1000$. Convergence in 103 sec.

middle, even with the ℓ_p reconstruction. This artifact is not present in our GRE estimate in Fig. 5(f) because the BC image ensures a minimum SNR level.

Acknowledgements

The authors would like to thank W. Kyriakos and S. Thibodeau at BWH for their help in acquiring data. Research was supported in part by a NDSEG graduate fellowship, ONR grant N00014-00-10089, NSF Johns Hopkins ERC 8810274, and NIH grants P41 RR13218, P01 CA67165, R33 CA99015, and R01 AG19513.

References

1. L. Axel. Surface coil magnetic resonance imaging. *J. of Comp. Asst. Tomography*, 8:381–384, 1984.
2. L. Axel, J. Costantini, and J. Listerud. Intensity correction in surface-coil MR imaging. *Am. J. Roentgenology*, 148:418–420, 1987.
3. D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 2nd edition, 1999.
4. M. Bhatia, W. C. Karl, and A. S. Willsky. A wavelet-based method for multiscale tomographic reconstruction. *IEEE Trans. Med. Imag.*, 15(1):92–101, 1996.
5. W. W. Brey and P. A. Narayana. Correction for intensity falloff in surface coil magnetic resonance imaging. *Med. Phys.*, 15(2):241–245, 1988.

6. W. L. Briggs. *A Multigrid Tutorial*. Society for Industrial and Applied Mathematics, Philadelphia, 1987.
7. D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Trans. Med. Imag.*, 17(3):463–468, June 1998.
8. B. M. Dawant, A. P. Zijdenbos, and R. A. Margolin. Correction of intensity variations in MR images for computer-aided tissue classification. *IEEE Trans. Med. Imag.*, 12:770–781, 1993.
9. A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc.*, 39:1–38, 1977.
10. D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Trans. Patt. Anal. Mach. Intell.*, 14(3):367–383, 1992.
11. J. Haselgrove and M. Prammer. An algorithm for compensation of surface-coil images for sensitivity of the surface coil. *Mag. Res. Imag.*, 4:469–472, 1986.
12. R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
13. R. K.-S. Kwan, A. C. Evans, and G. B. Pike. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Trans. Med. Imag.*, 18(11):1085–1097, Nov. 1999.
14. S.-H. Lai and M. Fang. Intensity inhomogeneity correction for surface-coil MR images by image fusion. In *Proceedings of the International Conference on Multisource-Multisensor Information Fusion*, pages 880–887. CSREA Press, 1998.
15. B. Likar, M. A. Viergever, and F. Pernus. Retrospective correction of MR intensity inhomogeneity by information minimization. In *MICCAI 2000*, pages 375–384, 2000.
16. R. D. Nowak. Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Trans. Imag. Proc.*, 8(10):1408–1419, 1999.
17. K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. SENSE: Sensitivity encoding for fast MRI. *Mag. Res. Med.*, 42:952–962, 1999.
18. P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, and O. M. Mueller. The NMR phased array. *Mag. Res. Med.*, 16:192–225, 1990.
19. J. G. Sled, A. P. Zijdenbos, and A. C. Evans. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imag.*, 17:87–97, 1998.
20. C. R. Vogel and M. E. Oman. Iterative methods for total variation denoising. *SIAM J. Sci. Comp.*, 17(1):227–238, 1996.
21. W. M. Wells, W. E. L. Grimson, R. Kikinis, and F. Jolesz. Adaptive segmentation of MRI data. *IEEE Trans. Med. Imag.*, 15(4):429–442, 1996.
22. Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden Markov random field model and the Expectation-Maximization algorithm. *IEEE Trans. Med. Imag.*, 20(1):45–57, 2001.