

A Robust Parametric Method for Bias Field Estimation and Segmentation of MR Images

Chunming Li¹, Chris Gatenby¹, Li Wang², John C. Gore¹

¹ Vanderbilt University of Imaging Science, Nashville, TN 37232, USA

² Nanjing University of Science and Technology, Nanjing, China

lchunming@gmail.com

Abstract

This paper proposes a new energy minimization framework for simultaneous estimation of the bias field and segmentation of tissues for magnetic resonance images. The bias field is modeled as a linear combination of a set of basis functions, and thereby parameterized by the coefficients of the basis functions. We define an energy that depends on the coefficients of the basis functions, the membership functions of the tissues in the image, and the constants approximating the true signal from the corresponding tissues. This energy is convex in each of its variables. Bias field estimation and image segmentation are simultaneously achieved as the result of minimizing this energy. We provide an efficient iterative algorithm for energy minimization, which converges to the optimal solution at a fast rate. A salient advantage of our method is that its result is independent of initialization, which allows robust and fully automated application. The proposed method has been successfully applied to 3-Tesla MR images with desirable results. Comparisons with other approaches demonstrate the superior performance of this algorithm.

1. Introduction

In quantitative processing and analysis of magnetic resonance (MR) images, major difficulties arise from variations in intensity due to B1 and B0 field inhomogeneities. Such intensity non uniformities cause intensity variations even for a single tissue, which may mislead many image analysis algorithms, such as segmentation and registration. Therefore, correction for such intensity inhomogeneities is often a mandatory step before quantitative analysis of the image data.

Bias correction methods can be broadly categorized into two classes: prospective methods [19, 16] and retrospective methods [18, 6, 14, 11, 8]. Prospective methods try to avoid intensity inhomogeneity in the acquisition process by

using special hardware or specific sequences. These methods are able to correct some of the intensity inhomogeneity caused by the MR scanner. However, they cannot correct for sources of inhomogeneity that are patient dependant, which makes them of limited value for clinical use [10]. In contrast to the prospective methods, retrospective methods rely exclusively on the information within the acquired image and thus can be applied to remove patient dependant effects.

One of the most popular types of methods for bias field correction are segmentation based approaches [18, 7, 4, 11, 8, 9]. In these methods, the tasks of bias field correction and segmentation are interleaved in an iterative process such that they benefit from each other to yield better results. In [18], Wells *et al.* developed an approach based on an expectation-maximization (EM) algorithm for interleaved bias field correction and segmentation. This method was later improved by Guillemaud and Brady in [4]. However, methods based on the EM algorithm require a good initialization for either the bias field or for the classification estimate [15]. They typically require manual selections of representative points for each tissue class to perform initialization. Such initializations are subjective and often irreproducible [7]. Moreover, the final correction and segmentation are sensitive to the specific choices of initial conditions [11, 15, 17].

Based on the EM framework in [18], Leemput *et al.* [7] proposed an explicit parametric model of the bias field. Instead of manual intervention, their method used a digital brain atlas that provides *a priori* probability maps for white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). Although this method is claimed to be more robust than the method of Wells *et al.*, the initialization of the parameters remains critical [15]. In [11], Pham and Prince proposed an energy minimization approach for segmentation and bias field estimation in which a fuzzy C-means (FCM) algorithm was used for segmentation. In their proposed energy function, a term was introduced to ensure the smoothness of the computed bias field. The coefficient of the smoothing term is, however, sometimes difficult to ad-

just for desirable results [17]. In addition, this method is also computationally intensive [2], due to the introduction of the bias field smoothing term.

In this paper, we propose a new energy minimization approach for joint bias field estimation and tissue segmentation. A bias field is modeled as a linear combination of smooth basis functions, and hence parameterized as the coefficients of the basis functions. We define an energy that depends on the coefficients of the basis functions, the membership functions of the tissues in the image, and the constants approximating the true signal from the corresponding tissues. This energy is convex in each of its variables. Bias field estimation and image segmentation are simultaneously achieved as the result of minimizing this energy. A salient advantage of our method is that its result is independent of initialization, which allows robust and fully automated application.

2. Problem Formulation and an Energy Minimization Method

The intensity inhomogeneity in an MR image can be modeled as a multiplicative component of an observed image described by

$$I(x) = \tilde{b}(x)\tilde{J}(x) + n(x) \quad (1)$$

where $I(x)$ is the measured image intensity at location (voxel) x , \tilde{J} is the true signal to be restored, \tilde{b} is an unknown bias field, n is additive noise. Ideally, the true signal \tilde{J} from each tissue should take a specific value of the physical property being measured (e.g. the proton density for MR images). Therefore, it can be assumed that the true signal \tilde{J} is piecewise approximately constant. More specifically, we assume that there are N tissues in the region of interest, denoted by Ω , and that the true signals \tilde{J} originated from the i -th tissue are approximately a constant \tilde{c}_i . The bias field \tilde{b} is commonly assumed to be slowly varying.

It is our goal to estimate the unknown bias field \tilde{b} and true signal \tilde{J} from the measured image I . In this work, the bias field is estimated by a linear combination of a set of basis functions. Let g_1, \dots, g_M be a set of basis functions defined on Ω . We estimate the bias field by a linear combination of the basis functions

$$b(x) = \sum_{k=1}^M w_k g_k(x) \quad (2)$$

where $w_k \in \mathfrak{R}$, $k = 1, \dots, M$, are the combination coefficients. Theoretically, any function can be approximated by a linear combination of a set of basis functions up to arbitrary accuracy [12], given a sufficiently large number of basis functions. In our current implementation, we use orthogonal polynomials as the basis functions, i.e. the basis

functions g_1, \dots, g_M satisfy

$$\int_{\Omega} g_i(x)g_j(x)dx = \delta_{ij} \quad (3)$$

where $\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ij} = 1$ for $i = j$.

The segmentation is given by an optimal estimation of the true signal \tilde{J} by a piecewise constant map J , which takes a constant value c_i in the region Ω_i of the i -th tissue. These regions $\{\Omega_i\}_{i=1}^N$ of N tissues form a *partition* of the image domain Ω in the sense that $\cup_{i=1}^N \Omega_i = \Omega$ and $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$. Thus, the piecewise constant map J can be written as

$$J(x) = \sum_{i=1}^N c_i u_i(x). \quad (4)$$

where $u_i(x)$ is the membership function of the region Ω_i such that

$$u_i(x) = \begin{cases} 1, & x \in \Omega_i \\ 0, & \text{else.} \end{cases} \quad (5)$$

and

$$\sum_{i=1}^N u_i(x) = 1. \quad (6)$$

2.1. Energy Formulation

We formulate the problem of segmentation and bias field estimation as a task of seeking the best piecewise constant map J and bias field b such that their product $b \cdot J$ best fits the measured intensity image I . The piecewise constant map J can be expressed as $J = \sum_{i=1}^N c_i u_i$ as in Eq. (4), while the bias field b is modeled as a linear combination $b = \sum_{k=1}^M w_k g_k$ as in Eq. (2). Therefore, we define the fitting error

$$F = \int_{\Omega} |I(x) - [\sum_{k=1}^M w_k g_k(x)][\sum_{i=1}^N c_i u_i(x)]|^2 dx \quad (7)$$

This fitting error is the proposed energy in terms of the constants c_1, \dots, c_N , the membership functions u_1, \dots, u_N , and the coefficients w_1, \dots, w_M . The minimization of this energy gives the optimal membership functions u_1, \dots, u_N as a segmentation result, while the optimal coefficients w_1, \dots, w_M of the basis functions define the estimated bias field.

The scalar constants c_1, \dots, c_N and w_1, \dots, w_M , and the functions u_1, \dots, u_N and g_1, \dots, g_M can be represented in the form of column vectors, i.e. $\mathbf{c} = (c_1, \dots, c_N)^T$, $\mathbf{w} = (w_1, \dots, w_M)^T$, $U(x) = (u_1(x), \dots, u_N(x))^T$, and $G(x) = (g_1(x), \dots, g_M(x))^T$. Thus, the above energy F can be rewritten as

$$F(U, \mathbf{c}, \mathbf{w}) = \int_{\Omega} |I(x) - (\mathbf{w}^T G(x))(\mathbf{c}^T U(x))|^2 dx \quad (8)$$

From Eq. (5), we have $\mathbf{c}^T U(x) = c_i$ for $x \in \Omega_i$. Thus, the energy F can be rewritten as

$$\begin{aligned} F(U, \mathbf{c}, \mathbf{w}) &= \sum_{i=1}^N \int_{\Omega_i} |I(x) - (\mathbf{w}^T G(x))(\mathbf{c}^T U(x))|^2 dx \\ &= \sum_{i=1}^N \int_{\Omega_i} |I(x) - (\mathbf{w}^T G(x))c_i|^2 dx \quad (9) \\ &= \sum_{i=1}^N \int_{\Omega} |I(x) - (\mathbf{w}^T G(x))c_i|^2 u_i(x) dx \end{aligned}$$

From the expression of F in the last line, it is easy to minimize with respect to the membership functions u_1, \dots, u_N .

2.2. Energy Minimization

It is worth noting that the energy $F(U, \mathbf{c}, \mathbf{w})$ is convex in each of its variables. The energy $F(U, \mathbf{c}, \mathbf{w})$ can be minimized by an iterative process of interleaved minimization with respect to each variable. The minimizer of $F(U, \mathbf{c}, \mathbf{w})$ in each variable, U , \mathbf{c} , or \mathbf{w} , is given below.

Minimization with respect to U . For fixed \mathbf{c} and \mathbf{w} , we minimize $F(U, \mathbf{c}, \mathbf{w})$ with respect to U under a constraint that $U = (u_1, \dots, u_N)^T$ satisfies Eq. (6). It can be shown that the minimizer $\hat{U} = (\hat{u}_1, \dots, \hat{u}_N)^T$ is given by

$$\hat{u}_i(x) = \begin{cases} 1, & i = i_{\min}(x); \\ 0, & i \neq i_{\min}(x). \end{cases}, \quad i = 1, \dots, N. \quad (10)$$

where

$$i_{\min}(x) = \arg \min_i \{|I(x) - (\mathbf{w}^T G(x))c_i|^2\}.$$

Minimization with respect to \mathbf{c} . For fixed U and \mathbf{w} , there is a unique minimizer of the function $F(U, \mathbf{c}, \mathbf{w})$ in the variable \mathbf{c} . This unique minimizer, denoted by $\hat{\mathbf{c}} = (\hat{c}_1, \dots, \hat{c}_N)^T$, is given by

$$\hat{c}_i = \frac{\int_{\Omega} I(x) b(x) u_i(x) dx}{\int_{\Omega} b^2(x) u_i(x) dx}, \quad i = 1, \dots, N. \quad (11)$$

Minimization with respect to \mathbf{w} . For fixed U and \mathbf{c} , to minimize $F(U, \mathbf{c}, \mathbf{w})$ with respect to \mathbf{w} , we take derivative of F with respect to \mathbf{w} , we get

$$\frac{\partial F}{\partial \mathbf{w}} = -2\mathbf{v} + 2A\mathbf{w}$$

where

$$\mathbf{v} = \int_{\Omega} I(x) G(x) J(x) dx,$$

and

$$A = \int_{\Omega} G(x) G(x)^T J^2(x) dx, \quad (12)$$

with $J(x) = \mathbf{c}^T U$. Note that A is an $M \times M$ matrix, with M being the number of the basis functions. It can be shown that the above matrix A is nonsingular (see below). Therefore, the linear equation $\frac{\partial F}{\partial \mathbf{w}} = -2\mathbf{v} + 2A\mathbf{w} = 0$ has a unique solution

$$\hat{\mathbf{w}} = A^{-1}\mathbf{v} \quad (13)$$

The entire procedure of minimization of the energy $F(U, \mathbf{c}, \mathbf{w})$ is described as below:

Step 1. Initialization of \mathbf{c} , \mathbf{w} and U ;

Step 2. Update \mathbf{c} to be $\hat{\mathbf{c}}$ given by Eq. (11);

Step 3. Update \mathbf{w} to be $\hat{\mathbf{w}}$ given by Eq. (13);

Step 4. Update U to be \hat{U} given by Eq. (10);

Step 5. Check convergence criterion. If convergence has been reached, stop the iteration, otherwise, go to Step 2.

2.3. Matrix Analysis for Numerical Issues

The non-singularity of matrix A is verified as the following. We first define $h_m(x) \triangleq g_m(x) \sqrt{\sum_{i=1}^N c_i^2 u_i(x)}$. Thus, the (m, k) entry of A can be expressed as the inner product of h_m and h_k given by

$$\langle h_m, h_k \rangle = \int_{\Omega} h_m(x) h_k(x) dx.$$

Therefore, the matrix A is the *Gramian matrix* of h_1, \dots, h_M . By linear algebra [5], the Gramian matrix of h_1, \dots, h_M is non-singular if and only if they are linearly independent. It is easy to see that the above defined functions h_1, \dots, h_M are linearly independent, which implies the non-singularity of A .

Numerical stability is an important issue in computing the inverse matrix A^{-1} in Eq. (13), and may have a significant impact on the accuracy of the final result of our method. In general, the numerical stability of computing an inverse matrix A^{-1} can be characterized by the *condition number* [3] of the matrix A . The condition number of a positive-definite matrix A is given by

$$\kappa(A) = \lambda_{\max}(A) / \lambda_{\min}(A),$$

where $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ are the minimal and maximal eigenvalues of matrix A , respectively. The stability is ensured when the condition number is bounded by a reasonably small number.

For the above defined matrix A in Eq. (12) with the basis functions g_1, \dots, g_M satisfying the orthogonality condition in Eq. (3), we are able to prove that

$$0 < \min_i \{c_i^2\} \leq \lambda_{\min}(A) \leq \lambda_{\max}(A) \leq \max_i \{c_i^2\}$$

Therefore, an upper bound of the condition number of A is provided by

$$\kappa(A) \leq \frac{\max_i \{c_i^2\}}{\min_i \{c_i^2\}}. \quad (14)$$

For example, if $\max_i \{c_i\} = 250$ and $\min_i \{c_i\} = 50$, by the inequality (14), we have $\kappa(A) \leq \frac{250^2}{50^2} = 25$. We have observed that the condition numbers are less than 26.0 for all the experiments in this paper, which is sufficient to ensure the numerical stability of the inversion operation in our algorithm.

3. Results

In this section, we demonstrate the effectiveness of the proposed method, especially its robustness to initialization. To verify the robustness of our method, we randomly generate different initializations of all the variables U , c , and w in using our method. Experimental results have confirmed that our algorithm converges to the same result from different initializations that are generated randomly. Moreover, the convergence is reached after a small number of iterations. The typical number of iterations needed for convergence is between 10 and 20 for most of images.

Our method has been applied to 3-tesla MR images, which have significant intensity inhomogeneities. We first show the results of our method for the 3-tesla MR images in the first column of Fig. 1. The estimated bias fields, the segmentation results, and bias corrected images are shown in the second, third, and fourth columns, respectively. As can be seen from third column, the tissue segmentation results are quite consistent with the brain anatomy. In the bias corrected images in the fourth column, the intensities within each tissue become quite homogeneous.

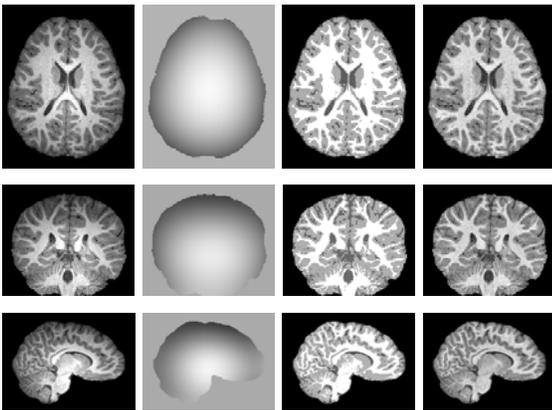


Figure 1. Results of our method on 3T brain MR images. Column 1: Original images; Column 2: Estimated bias fields; Column 3: Segmentation results. Column 4: Bias corrected images.

We compared our method with the methods of Wells *et*

al. [18], Leemput *et al.* [7]. We tested these three methods on a synthetic image of circles with intensity inhomogeneity (top left in Fig. 2) and an MR brain image with noise 1%, intensity non-uniformity (INU) 80% (top right Fig. 2), which was obtained from McGill Brain Web [1]. Since the algorithms of Wells *et al.* and Leemput *et al.* are sensitive to the initialization of the parameters (mean, variances and *a priori* probability for each tissue), proper initialization of these parameters is necessary in using their methods. Unless otherwise specified, the initialization of the parameters in their methods are obtained from a preliminary segmentation and estimation of the parameters using the K-means algorithm. When such initializations are used, the segmentation results of Wells *et al.*'s method (Figs. 2(c) and 2(e)) and Leemput *et al.*'s method (Figs. 2(g) and 2(i)) are acceptable. However, when the initialization of these parameters are changed, the segmentation results of these two methods vary significantly, as shown in the two results in Figs. 2(d), 2(f), 2(h), 2(j). The results of our method for two randomly generated initializations are shown in the bottom row of Fig. 2. No visible difference can be seen from the results of our method for the two randomly generated initializations.

Figs. 3 and 4 show the comparison results for two synthetic images with severe intensity inhomogeneities. The results of the methods by Wells *et al.*, Leemput *et al.*, and the proposed one are shown in the first, second and third rows, respectively. The estimated bias field, segmentation results, and bias corrected image are shown in every row. While it is difficult to visually compare the bias corrected images, the segmentation results of our method are more accurate than the other two methods, especially in the lower part of the image.

To quantitatively evaluate the performance of the algorithms, we use Jaccard similarity (JS) [13] as an indicator of the segmentation accuracy. The JS between two regions S_1 and S_2 is defined as the ratio between the areas of the intersection and the union of them, namely, $J(S_1, S_2) = \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|}$. To evaluate the accuracy of segmentation, we compute the JS between the segmented region S_1 by the algorithm and the corresponding region S_2 given by the ground truth. The closer the JS values to 1, the better the segmentation and bias correction.

We tested the three methods on 30 images from McGill brain data. For the methods of Wells *et al.* and Leemput *et al.*, we used 20 different initializations of the means, while the variances and *a priori* probability are initialized properly by preliminary estimation of them. The JS values for WM and GM of the results obtained by the three methods are shown in Fig. 5. The JS values of the Wells *et al.*, the Leemput *et al.* and our method are plotted with red squares, green circles and blue diamonds respectively. The JS values of our method for the 20 different initializations show

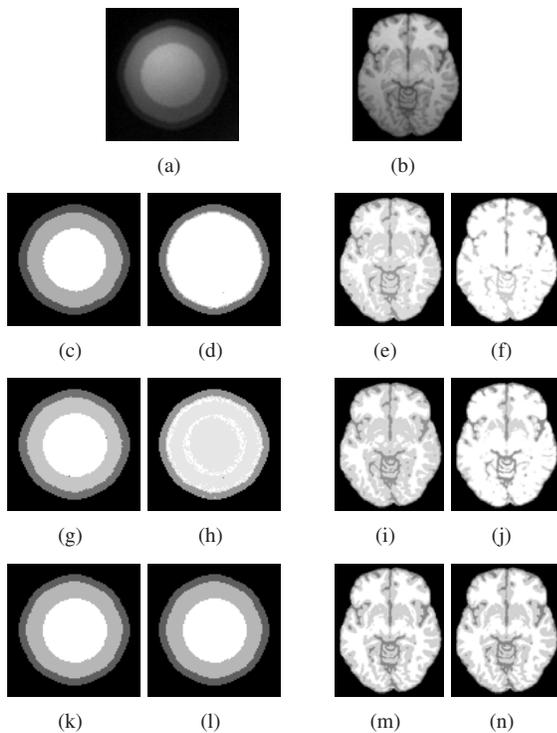


Figure 2. Results of the method of Wells *et al.*, the method of Leemput *et al.* and our method with two different initializations. The segmentation results with good guess are shown in columns 1 and 3; the segmentation results without good guess are shown in columns 2 and 4. The results of the methods of Wells *et al.*, Leemput *et al.*, and the proposed method are shown in rows 2,3,4, respectively.

no visible difference in Fig. 5, which demonstrate the independence of initialization of our method. By contrast, there is large variability in the JS values for the results obtained by the methods of Wells *et al.* and Leemput *et al.* for 20 different initializations, as can be seen in Fig. 5.

We have applied our method to 3D MR images. Fig. 6 shows the result of our method for a 3-tesla 3D MR image. To visualize the results, we select four sagittal slices as shown in the first row of Fig. 6. The corresponding estimated bias fields, final segmentation results and corrected images are shown in the second, third, and fourth rows respectively. The intensities with each tissue become quite homogeneous in the bias corrected images. Meanwhile, the segmentation results show high agreement with the brain anatomy.

4. Conclusion

In this paper, we have presented a new energy minimization framework for simultaneous estimation of the bias field and segmentation of tissues for magnetic resonance images.

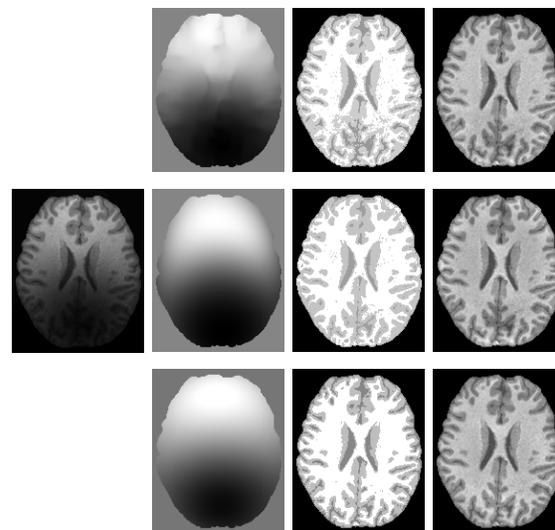


Figure 3. Comparison results for a synthetic image shown in the first column. The estimated bias fields, segmentation results, and bias corrected images are shown in the second, third, and fourth columns, respectively. The results of the methods of Wells *et al.*, Leemput *et al.*, and the proposed method are shown in rows 1,2,3, respectively.

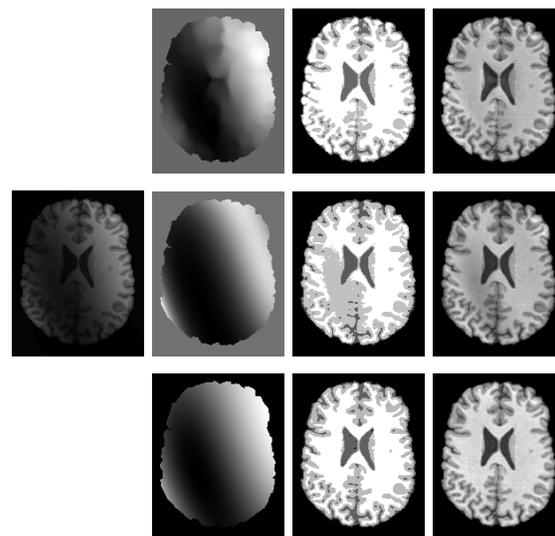


Figure 4. Comparison results for a synthetic image shown in the first column. The estimated bias fields, segmentation results, and bias corrected images are shown in the second, third, and fourth columns, respectively. The results of the methods of Wells *et al.*, Leemput *et al.*, and the proposed method are shown in rows 1,2,3, respectively.

We provide an efficient iterative algorithm for energy minimization, which converges to the optimal solution at a fast rate. A salient advantage of our method is that its result is independent of initialization, which allows robust and fully

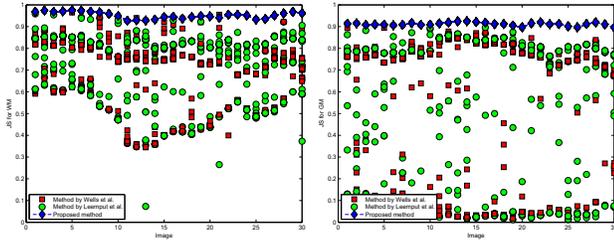


Figure 5. Test of sensitivity to initialization for our method and the methods of Wells *et al.* and Leemput *et al.* The x-axis represents 30 images, and the y-axis represents the JS values for WM (left) and GM (right) of the three methods for 20 different initializations.

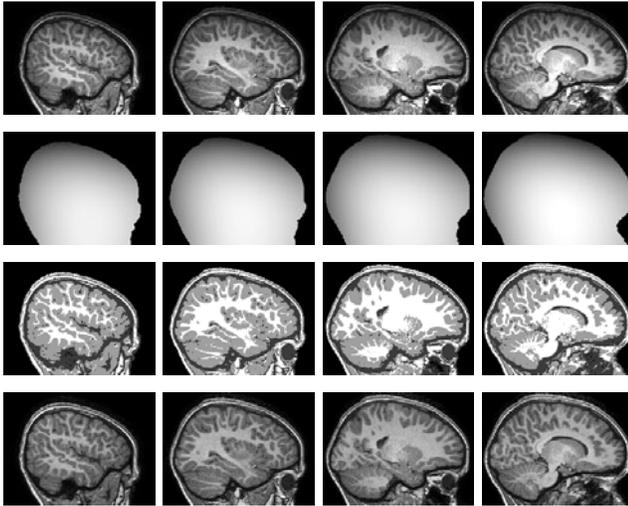


Figure 6. Sagittal view of the 3D segmentation and bias correction results. Row 1: Original images. Row 2: Estimated bias fields. Row 3: Segmentation results. Row 4: Bias corrected images.

automated application. The proposed method has been successfully applied to 3-Tesla MR images with desirable results. Comparisons with other approaches demonstrate the superior performance of this algorithm.

References

- [1] <http://www.bic.mni.mcgill.ca/brainweb/>. 4
- [2] M. Ahmed, S. Yamany, N. Mohamed, A. Farag, and T. Moriarty. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Trans. Med. Imaging*, 21(3):193–199, March 2002. 2
- [3] G. Golub and C. V. Loan. *Matrix Computations*. The Johns Hopkins University Press, 3rd edition, 1996. 3
- [4] R. Guillemaud and J. Brady. Estimating the bias field of MR images. *IEEE Trans. Med. Imag.*, 16(3):238–251, June 1997. 1
- [5] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985. 3
- [6] B. Johnston, M. S. Atkins, B. Mackiewicz, and M. Anderson. Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI. *IEEE Trans. Med. Imag.*, 15(2):154–169, April 1996. 1
- [7] V. Leemput, K. Maes, D. Vandermeulen, and P. Suetens. Automated model-based bias field correction of MR images of the brain. *IEEE Trans. Med. Imag.*, 18(10):885–896, October 1999. 1, 4
- [8] C. Li, R. Huang, Z. Ding, C. Gatenby, D. Metaxas, and J. Gore. A variational level set approach to segmentation and bias correction of medical images with intensity inhomogeneity. In *Proceedings of Medical Image Computing and Computer Aided Intervention (MICCAI)*, volume LNCS 5242, Part II, pages 1083–1091, 2008. 1
- [9] C. Li, C. Xu, A. Anderson, and J. Gore. MRI tissue classification and bias field estimation based on coherent local intensity clustering: A unified energy minimization framework. In *Proceedings of Information Processing in Medical Imaging (IPMI)*, Williamsburg, VA on July 5-10, 2009. 1
- [10] B. Likar, M. Viergever, and F. Pernus. Retrospective correction of mr intensity inhomogeneity by information minimization. *IEEE Trans. Med. Imag.*, 20(12):1398–1410, December 2001. 1
- [11] D. Pham and J. Prince. Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Trans. Med. Imag.*, 18(9):737–752, September 1999. 1
- [12] M. J. D. Powell. *Approximation Theory and Methods*. Cambridge University Press, Cambridge, 1981. 2
- [13] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy. Magnetic resonance image tissue classification using a partial volume model. *Neuroimage*, 13:856–876, May 2001. 4
- [14] J. Sled, A. Zijdenbos, and A. Evans. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging*, 17(1):87–97, January 1998. 1
- [15] M. Styner, C. Brechbuhler, G. Szekely, and G. Gerig. Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Trans. Med. Imag.*, 19(3):153–165, March 2000. 1
- [16] M. Tincher, C. R. Meyer, R. Gupta, and D. M. Williams. Polynomial modeling and reduction of rf body coil spatial inhomogeneity in MRI. *IEEE Trans. Med. Imag.*, 12(2):361–365, 1993. 1
- [17] U. Vovk, F. Pernus, and B. Likar. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Trans. Med. Imag.*, 26(3):405–421, March 2007. 1, 2
- [18] W. Wells, E. Grimson, R. Kikinis, and F. Jolesz. Adaptive segmentation of MRI data. *IEEE Trans. Med. Imag.*, 15(4):429–442, 1996. 1, 4
- [19] D. A. G. Wicks, G. J. Barker, and P. S. Tofts. Correction of intensity nonuniformity in MR images of any orientation. *Magn. Reson. Imag.*, 11(2):183–196, 1993. 1