MRI Bias Field Correction Based on Tissue Labeling

Lianli Liu, Jiyang Chu*, Jie Li, Zhen Zeng
Electrical Engineering and Computer Science, University of Michigan.
Nuclear Engineering and Radiological Sciences, University of Michigan*.
Ann Arbor, MI 48109
{llliu,chujy,ljlijie,zengzhen}@umich.edu
EECS 556 Apr 26th 2014

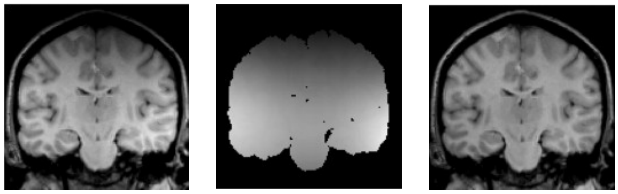
***Abstract*—Intensity inhomogeneity, also known as "bias field" in MR images, mainly arises from imperfection in RF profiles, and it harms the performance of many image analysis algorithms. Traditional Segmentation based bias field correction methods is constrained by its low intensity based segmentation accuracy. To improve segmentation accuracy, we propose to integrate rich-feature segmentation in bias field estimation. A Laplacian regularization scheme is also designed to encourage the smoothness of estimated bias field. We use synthetic data from BrainWeb [1] and augmented them with artificial bias field. Experiment on the dataset shows that our method provides more robust and accurate segmentation result, which results in better bias field estimation.**

I. Introduction

Intensity nonuniformities, also known as the bias field in magnetic resonance imaging (MRI), arise from various factors such as imperfections in the radiofrequency (RF) pulse profile; nonuniform flip angles caused by an inhomogeneous transmit field; nonuniform reception sensitivity; RF penetration effects dependent upon the electromagnetic parameters of the object; wave behavior when the object size is equal to or more than one-half the wavelength of electromagnetic radiation and finally, gradient eddy currents related to the coupling between the object and gradient coils [2]. Such bias field manifests as a slowly varying signal intensity variation across tissue regions that should be uniform, as shown in Fig. 1. Human eyes are robust to such intensity inhomogeneity, thus medical experts can perform analysis tasks correctly. However, many intensity based image analysis algorithms are very sensitive to such intensity variation; thus correction of bias field is of great importance for accurate image analysis results.

Various methods have been proposed to address this issue, including prospective methods [3][4][5], where bias field is corrected during imaging, and retrospective ones, where bias field is estimated from images. Retrospective methods are more general in application as they do not rely on specific MR scanners and are able to remove patient induced inhomogeneity. Segmentation based method is one of the important categories in retrospective methods.

However, most existing segmentation based methods only rely on image intensities or simple statistics referred from intensities [6][7][8][9]. Other potentially useful information, such as geometric shape, texture are ignored. The state of the art segmentation techniques developed in computer vision are deployed to obtain better segmentation results. In this project, we investigated different features for supervised based segmentations, and integrated the segmentation algorithm in bias field correction.



(a) Image corrupted by bias field (b) Estimated bias field (c) Corrected image

Figure 1. Bias field example

Specifically, our project includes the following parts

1) **Replication**: We will first reproduce existing work on segmentaion based bias field correction, as proposed by Chen et al. [9], as well as rich-feature segmentation method developed by Hoiem et al. [10].
2) **Extension**: After replication, we will integrate rich-feature segmentation technique into our bias filed correction framework to see if better segmentation techniques can lead to improvement in bias field correction.

The hypothesis we base our plan on are:

- Existing segmentation based bias field correction, can lead to approximate 15% improvement in segmentation accuracy comparing to direct segmentation on bias corrupted images, measured by the percentage of correctly classified voxels, as reported by Ahmed et al. [11]
- The rich feature tissue segmentation provides a more accurate segmentation than fuzzy c-means based segmentation methods. The quality of segmentation will be evaluated as the pixel-wise label accuracy, which can be measured against ground truth (human labeled or provided by synthetic data). We expect the segmentation accuracy to be 30% higher than fuzzy c-means [12].
- Integrating improved segmentation method with bias field correction will lead to higher segmentation accuracy compared with existing work[9] we replicate. We expect our performance to be improved by 30%.

II. Related Work

A. Segmentation Based Bias Field Correction

Among various retrospective methods, segmentation based methods have received much attention as they incorporate anatomy information, such as tissue type, into correction process. For example, maximum-likelihood (ML) or maximum a posteriori probability (MAP) criterion are applied to estimate image intensity probability distributions by parametric models, where bias field is modeled as a Gaussian mixture model as in [13][14]. Fuzzy c-means based segmentation is also insensitively studied as it assigns partial membership to voxels, which is consistent with partial volume effect in MR images. Standard fuzzy c-means objective function is modified in various ways to incorporate bias field effect [15][11][16]. Other methods based on non-parametric models are also proposed as in [17][18], using maxshift or meanshift clustering.

All these methods rely on image intensity only for segmentation, while over the past few decades, image segmentation has been an active area in computer vision, where numerous methods have been proposed that rely on more complicated features extracted from image instead of simple pixel intensity.

B. Image Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. It has long been investigated in the field of image processing and computer vision. Depending on whether prior information is provided, image segmentation can be categorized into unsupervised segmentation and supervised segmentation. Unsupervised segmentation clusters pixels based on their similarities in specific aspects (features). Examples of the specific aspects being used in computer vision community are color of a pixel, or richer features such as texture and shapes. Felzenszwalb [19] et al. proposed a graph-based algorithm to segment an image into regions. They start with a graph with each node corresponding to each pixel, and each pixel is connected to its neighbors by edges. Then they iteratively prune the edges if two connected components are dissimilar. The final graph with separated connected components correspond to separate regions in the segmentation result. Their work presented good segmentation results that align with the visual boundaries, but an image can be oversegmented in a sense that each segment does not correspond to a semantic entity, such as an object.

If we want to get semantic level segmentation, such as object segmentation, then prior knowledge on the object could be utilized for better segmentation. Supervised segmentation groups pixels into regions such that each region gets assigned to a specific pre-trained semantic label. These methods typically start with oversegmentation results of an image, and then learn a prior on the appearance for each semantic label, so that they can estimate the posterior probability of assigning a specific label to one segment based on the appearance of the segment. Hoiem et al. [10] learned the appearance model for some 3D geometric labels, such as "horizontal" and "vertical". Their method segments a given image into several geometric classes. The features that they used are simple and several 3D geometric features designed specifically for their task. In the case that hand crafted features are powerful enough, we can deploy deep learning [20] algorithm to automatically generate a hierarchy of useful features. Farabet et al. [20] used deep learning for generating a set of features in the task of scene labeling. The drawback of this line of work is the requirement of a huge amount of training data, which is hard to get due to the human efforts needed on labeling.

In this project we will investigate more robust segmentation methods for better bias field correction, as well as the corresponding regularization schemes to validate the above hypothesis.

III. Technical Approach

In this project, we plan to integrate rich feature tissue segmentation into the workflow of bias field correction, thus validating the hypothesis presented previously. Therefore, existing works that we plan to replicate are two-fold: 1) rich feature tissue segmentation; 2) segmentation based bias field correction.

A. Rich Feature Tissue Segmentation

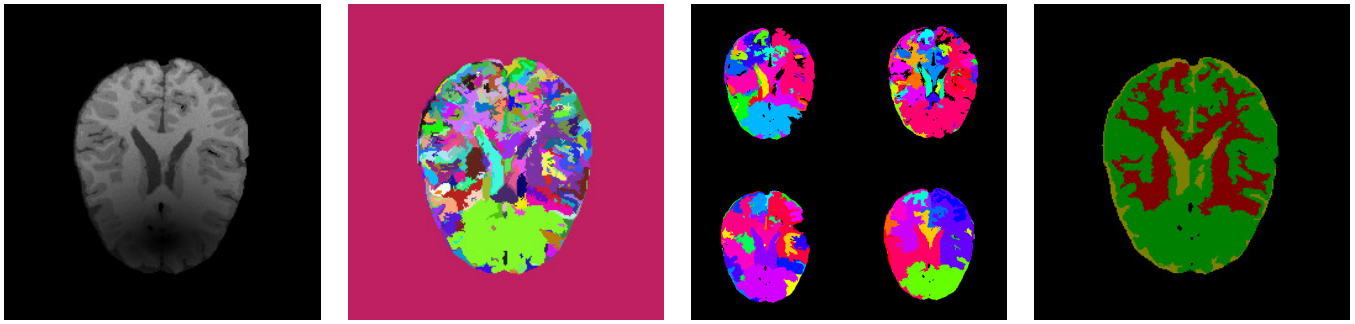
We aim to segment a given MRI image into regions that correspond to 3 tissue classes, e.g. white matter, gray matter, and fluid. Appearance based features can be extracted to model each class. Integrating different features into a segmentation process is not a trivial task. Different features ask for different spatial support. Some are small-scale features like color or texture. Some are large-scale features like shape. How to provide the supporting area greatly affects the robustness and accuracy of a segmentation algorithm. Inspired by the work of Hoiem et al. [10] scene labeling, we proposed to follow Hoiem's multi-level segmentation pipeline to segment tissues in MRI images.

As shown in Fig. 2, our proposed tissue labeling has 4 parts: 1) superpixel generation; 2) multiple hypothesis generation; 3) labeling. And then we will introduce how we trained the appearance models and the features being used.

1) Superpixel Generation: We directly apply the segmentation methods by Felzenszwalb [19] to get a set of over segmented superpixels. As shown in Fig. 2, the oversegmentation method also works with images with bias field. The goal of the tissue segmentation is to label each superpixel as one of the pre-defined tissue classes, with the hope that all the pixels within one superpixel should belong to the same tissue class.

2) Multiple Hypothesis Generation: By changing the thresholds in the segmentation methods [19], we can have different oversegmentation results. It is intractable to evaluate every possible oversegmentation result and find the one that best segments the image into the pre-learned classes. Thus we first over segmented the image into very small

s (typically around 400 superpixels for an image in our dataset), and then grouped the superpixels of the oversegmentation into bigger regions based on the grouping likelihood $P(y_i = y_j||x_i - x_j|)$, where y_i, y_j are the labels for superpixel i and j, and x_i, x_j are their features. The likelihood function is



(a) Input Image (b) Over Segmentation (c) Multiple level Segments Hypothesis (d) Final Segmentation Result

Figure 2. Illustration figures for different steps in our proposed segmentation framework.

learned in our training stage, as explained later. And the target number of the final regions can vary, thus we can get different groupings of superpixels, and each grouping is referred as one of our hypothesis.

3) Labeling: For each hypotheses, we assume that each region belongs to the same tissue class, meaning all the superpixels within that region should be assigned to the same tissue label. Because of multiple hypothesis, each superpixel will be a member of several different regions from different hypothesis. The superpixel label confidence Eq. (1) is measured by a weighted sum of the label likelihoods of the regions that contain it, and the weights depend on the homogeneity likelihood of the regions [10]:

$$C(y_i = v|x) = \sum_{j}^{n_h} P(y_i = v|x, h_{ji})P(h_{ji}|x) \qquad (1)$$

where y_i is the superpixel label, and v is one of the possible labels, and x is the image, n_h is the total number of hypothesis, and h_{ji} is the regions that contains the ith superpixel for the j^{th} hypothesis, and y_j is the tissue label for the region h_{ji}. The label likelihood function of the regions, and the homogeneity likelihood function of the regions are learned in our training stage, as explained below.

Training In the training stage, for each training image, we generate multiple hypothesis. And for each hypothesis, we label the region as one of the pre-defined tissue class if the region corresponds well to one tissue class, or label the region as "mixed" if the region contains multiple tissue classes. Features are extracted within each region, and the features being used are as shown in Fig. 3. Details about features extractions are explained later. The grouping likelihood of two superpixels are learned based on the parent-child relations between the regions and superpixels. The label likelihood function of a region given the image data within that region is learned in a one versus rest fashion. And the homogeneity likelihood function of a region is learned by classifying homogeneously labeled (i.e. label as just one of the tissue classes) vs "mixed" labeled. All the likelihood functions can be estimated using logistic regression version of Adaboost [21], or using other classifiers such as SVM [22]. In our case, we used the online available matlab codes developed based on Adaboost [21].

Features As shown in Fig. 3, we have several types of features that we can extract within each region. We chose this set of features since they can help recognize the tissue type as explained below.

Feature Description
Intensity I1: T1 intensity
Texture X1: LM filters: mean absolute response (15 filters) X2: LM filters: histogram of maximum responses (15 bins)
Location L1: normalized x and y, mean L2: relative location to center point. L3: relative distance to center point
Shape S1: number of pixels S2: normalized area in image S3: ratio of width by height

Figure 3. Features to extract

Intensity: Different MRI image types (T1 weighted, T2 weighted, etc.) have different tissue contrast, and one single MRI volume may not be sufficient to separate all tissue types. For example, it has been reported that and T2-weighted images was more sensitive to bone density variation compared to other MRI image types; calculated fat image using Dixon method has strong contrast of fat with other tissue yet is not capable of separating other tissue types [23]. Most existing bias field uses single phase image (mainly T1 only) [24], thus it may be useful to study bias field correction using multi-phase images, utilizing characteristics of different MRI volume types. Here we represent the intensity using two intensity channels: T1, and T2, and calculate their mean (I1) and histograms (I2).

Texture: To explore whether texture gives us any cues for the tissue class, we extracted texture features and include them in the training stage. We apply a subset of the filter banks by Leung and Malik [25]. The texture is represented by the absolute filter response of each filter (X1) and the histogram of maximum responses over pixels within a segment (X2).

Location: Since MRI images using the same coil are mostly aligned together, the location is a very strong cue of what the tissue type is. The locations are represented by the normalized pixel locations of the tissue type: the mean of the locations

(L1), and also the 10th and 90th percentile of the pixel position of a region in the image (L2).

Shape: The shape and extension of the region is another cue that we can utilize. We represent the extension of the region by the number of superpixels (S1), and the shape is roughly represented by the normalized area in the image (S2). Depending on the results, we might incorporate more discriminative features for shape later.

B. Segmentation Based Bias Field Correction

The joint segmentation and bias field method proposed by Chen et al. [9] formulate the optimization problem the as

$$\begin{aligned} J(b,u) &= \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m}\|x_k - b_k - v_i\|^2 \\ \text{s.t. } &\sum_{i=1}^{c} u_{ik} = 1, \forall k \\ &0 \le u_{ik} \le 1, \forall k, i \end{aligned} \tag{2}$$

where x_k, b_k are log transformed image data and bias field at voxel k, u_{ik} is the fuzzy membership of voxel k belongs to tissue type i, c_i is the prototype vector of tissue class i and m controls the fuzzy degree. An iterative low pass filtering is applied to the estimated b after each iteration to guarantee smoothness.

This algorithm is computational expensive as an iterative 3D filtering is required at each algorithm iteration. Also, low pass filtering does not have a clear meaning in optimality nor do the authors specify which low pass filter is used in the paper.

To overcome this problem, instead of applying post processing step, we incorporate the constraint of smoothness into our optimization problem directly. This results in a Laplacian regularized problem

$$\begin{aligned} J(b,u) &= \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m}\|x_k - b_k - v_i\|^2 + \lambda\nabla\mathbf{b} \\ \text{s.t. } &\sum_{i=1}^{c} u_{ik} = 1, \forall k \\ &0 \le u_{ik} \le 1, \forall k, i \end{aligned} \tag{3}$$

In our attempt to solve this problem, we find that the segmentation error term and Laplacian regularization term are of different scales, as bias field usually have a unit mean while MRI voxel values can be very large. Experimental results show that though the value of objective function continue decreasing during iterations, regularization term is actually increasing. Even when we increase the weight of regularization term, the value of regularization term drops down yet the increasing trend is not reversed.

Therefore, we resort to post processing techniques inspired by the Laplacian regularization above — Laplacian regularized least square fitting as in Eq. (4). That is we first solve the unregularized problem and get a residual image. Then we perform Laplacian regularized least square fitting (which is equivalent to cubic B-spline smoothing) to the residual image and get the estimated smooth bias field.

$$\hat{f} = \arg\min \frac{1}{n}\sum_{i=1}^{n}|y_i - f(t_i)|^2 + \beta\int |f^{(m)}(t)|^2 dt \tag{4}$$

Furthermore, we notice that although log transform is a necessary step to decouple multiplicative bias field and true MRI signal, it is not a linear operation and will distort the contrast of original image. Such contrast distortion will effect the segmentation algorithm as it relies heavily on contrast between different tissue types. Therefore, we replace Euclidean distance with a Gaussian distance on log transformed data, i.e $D(x,y) = 1 - \exp(\frac{-\|x-y\|^2}{\sigma^2})$ to recover original image contrast.

To conclude, our algorithm operates as follows

1) Evaluate membership:
$$u_{ik} = \sum_{j=1}^{c}\left(\frac{D(x_k - b_k, v_i)}{D(x_k - b_k, v_j)}\right)^{\frac{-1}{m-1}}$$
2) Update centroid:
$$v_i^* = \frac{\sum_{k=1}^{n} u_{ik}^m(1 - D(x_k - b_k, v_i)(x_k - b_k)}{\sum_{k=1}^{n} u_{ik}^m(1 - D(x_k - b_k, v_i))}$$
3) Calculate bias field:
$$b_k^* = \left[x_k - \frac{\sum_{i=1}^{c} u_{ik}^m(1 - D(x_k - b_k, v_i))(x_k - v_i)}{\sum_{i=1}^{c} u_{ik}^m(1 - D(x_k - b_k, v_i))}\right]$$
4) Regularize bias field using B-spline smoothing
5) If the change of the updated centroid is smaller than a given threshold, terminate the algorithm and return results, else go back to step 1)

C. Integrate Segmentation into Bias Field Correction

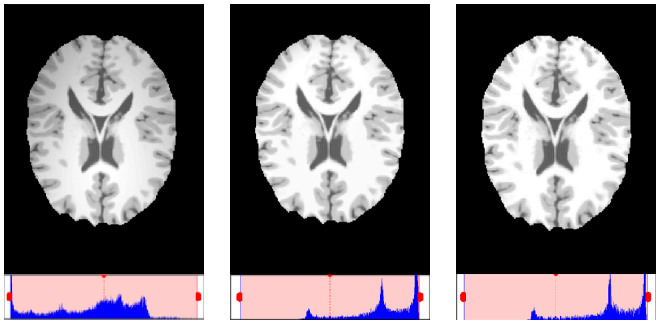
Given the segmentation results in part A, bias field can be estimated based on the assumption that same tissue type have uniform intensity. We propose two strategies to incorporate the rich features based segmentation into bias field correction, corresponding to the non-iterative and iterative cases.

Non-iterative Case When segmentation is invariant to intensity change, bias field will have little impact on segmentation results. Thus we purely rely on the segmentation results to generate a bias free image, i.e. the fuzzy membership of pixel for different tissue types are fixed. Given the label of each voxel u_{ik} and the prototype of each voxel c_i (which can be obtained from expert knowledge or supervised based tissue labeling), the voxel intensity z_k bias free image will be

$$z_k = \sum_{i=1}^{c} u_{ik}c_i \tag{5}$$

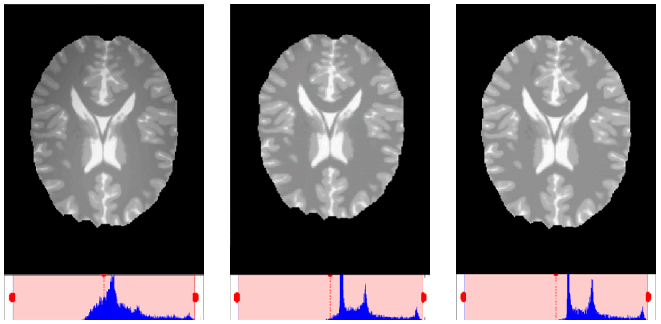
with this bias free image, we can easily estimate bias field as

$$\mathbf{b} = \mathbf{X}/\mathbf{Z} \tag{6}$$



(a) T1 image corrupted by 80% bias field and and histogram (b) Corrected T1 results and histogram (c) Ground truth T1 and histogram

Figure 4. IV-A1 T1 example.



(a) T2 image corrupted by 80% bias field and and histogram (b) Corrected T2 results and histogram (c) Ground truth T2 and histogram

Figure 5. IV-A1 T2 example.

where X is original image and Z is bias free image constructed by Eq. (6). However, the bias field estimation is an inverse problem and ill-posed. Therefore we apply the following regularized version, as proposed in [26]:

$$\hat{\mathbf{b}} = \arg\min \frac{1}{2}\|\mathbf{X} - \mathbf{Z}\mathbf{b}\|^2 + \frac{\lambda}{2}\|\mathbf{R}\mathbf{b}\|_2^2 \tag{7}$$

where $\mathbf{R}$ is the finite differencing matrix to ensure the smoothness of estimated bias field. Following the algorithm proposed in [26], this problem can be solved efficiently using variable splitting and Augmented Lagrangian(AL) methods.

Iterative Case In this case, segmentation and bias correction are performed jointly for an optimal solution. To do this, we defined the following objective function

$$\begin{aligned} J(b,u) &= \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^m \|x_k - b_k - v_i\|^2 + R(\mathbf{b}) \\ \text{s.t. } &\sum_{i=1}^{c} u_{ik} = 1 \forall k \\ &0 \le u_{ik} \le 1, \forall k, i \end{aligned} \tag{8}$$

where b, u, c has the same meaning as in Eq. (2) Similar with Eq. (3), a regularization item $R(\mathbf{b})$ is introduced to enforce the smooth constraints on bias field. Implementation details on $R(\mathbf{b})$ will be discussed in the experiment section IV.

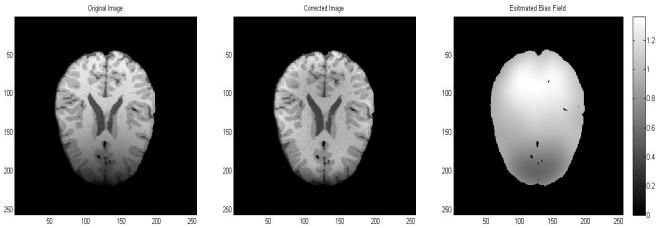
We started with rich feature based segmentation, a bias field is estimated using Eq. (8) and the bias corrected image is fed back into the segmentation algorithm to update u_{ik}. Thus we update the estimated bias field and segmentation iteratively util certain terminating condition is met. We used thresholds on the absolute differences of the variables between each iteration for the terminating condition.

IV. Experiment

Our experiment consists of four parts: datasets, the evaluation of our replicated bias correction methods, the evaluation of our replicated rich-feature segmentation methods, and the evaluation of our segmentation methods integrated with our bias correction methods.

A. Datasets

We used synthetic data downloaded from BrainWeb [1]. Two types of synthetic data are involved for tests:



(a) Corrupted image (b) Corrected image (c) Estimated Bias Field

Figure 6. Dataset IV-A2 Artificial bias field applied and result.

1) T1-T2 Dataset: $217 \times 181 \times 181$ T1-weighted and T2-weighted images with bias field are included. The true bias field generated from real MRI scanners is provided on the website, thus gives the ground truth for bias field estimation. Specifically, the bias filed is at a strength of 80%. No ground truth labels of tissue classes are provided, thus this dataset is only used to provide some qualitative result.

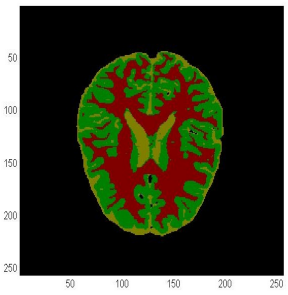
2) T1 Dataset with tissue label: T1-weighted images simulated from 20 brains anatomical model are included. The simulation images are bias free. Both hard tissue labels and soft classification results are provided, thus we used this dataset for segmentation evaluations. In addition, we augmented artificial bias field on this dataset for evaluations on our bias field correction method.

In our experiment, those clean images from the dataset is corrupted by artificial bias field which is generate using Prof. Fessler's tool box [27]. We use four different pattern of artificial bias field, as shown in Fig. 9

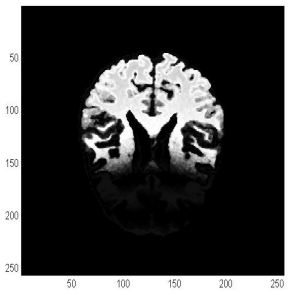
B. Fuzzy c-means based bias field correction

Three tissue types within brain region are considered: white matter, gray matter and fluid. We tested our replications on dataset IV-A1, and the NRMSE (normalized root of mean square error) of the estimated bias field is 0.38. As shown in Fig. 4, the intensity histogram of the corrected T1 image is very similar to the intensity histogram of the ground truth image. We also show similar performance on T2 images in Fig. 5. These results validate that our replication is successful.

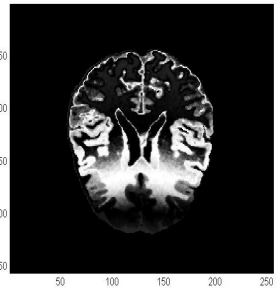
We also tested our replications on IV-A2, with images augmented by artificial bias field from Prof. Fessler's image reconstruction toolbox [27]. Parameter settings are the same with those in the tests on IV-A1. An example result is shown



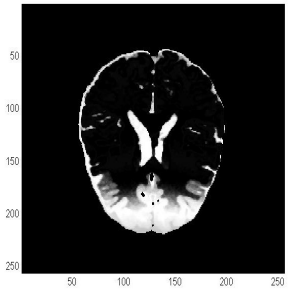
(a) Ground truth labels



(b) White matter segmentation

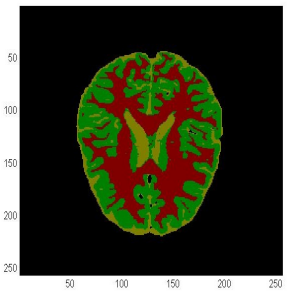


(c) Gray matter segmentation

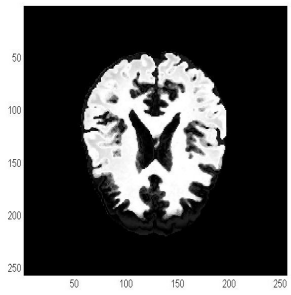


(d) Fluid segmentation

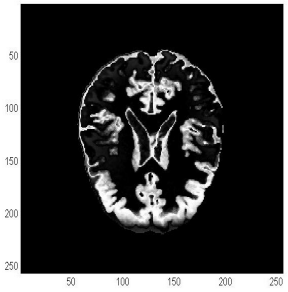
Figure 7. Before bias correction. Red: white matter; Greed: gray metter; Yellow: fluid.



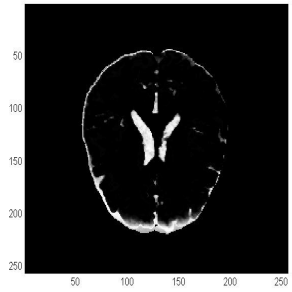
(a) Ground truth labels



(b) White matter segmentation



(c) Gray matter segmentation



(d) Fluid segmentation

Figure 8. After bias correction. Red: white matter; Greed: gray metter; Yellow: fluid.

in Fig. 6. When the image is corrupted by a strong bias field, fuzzy c-means based segmentation without bias field correction suffers a lot as shown in Fig. 7. After we applied our bias field correction methods, the segmentation results of the corrected image is much better as shown in Fig. 8. Before the bias field correction, the average accuracy of fuzzy c-means segmentation is only 22%; after the bias field correction based on our replications, the accuracy is 31%.

C. Rich-feature segmentation

To evaluate the robustness and accuracy of our rich feature tissue segmentation framework, we tested our algorithm on T1-weighted synthetic data in IV-A2. Artificial bias field is included using Prof. Fessler's tool box [27]. We used 50 clean image and corrupted them with 4 patterns of bias field, which composed to our training set. We then used 200 different clean images and corrupted them in the same way, which composed to our testing set. Given the ground truth labels of synthetic images, we evaluated the performance by averaging the label accuracy of each pixel. We compared our algorithm with fuzzy c-means method (FCM) and fuzzy c-means with bias field estimation (FCM-B).

The experiment results are shown in Fig. 12. An instance in our test set is also shown in Fig. 10. The results demonstrate that our proposed rich-feature segmentation dramatically outperforms fuzzy c-means segmentation (by 51.1%), and also fuzzy c-means with bias field estimation (by 41.8%). The results fit our second hypothesis. It can be seen from the result that Fuzzy c-means method is quite sensitive to bias field, while our approach provide similar performance on both clean images and biased images.

To better evaluate our segmentation method. We carried out some intensive analysis on feature selection and classifier robustness. We firstly did feature ablation study to analyze the importance of each type of features in our framework. As shown in Fig. 13, if we delete only one type of features, the performance won't drop to much. This result indicates that our method is robust to noise or bias in only one type of features. We also test the performance by using only one type of features. It turns out the texture is the most important feature in our case, while shape can be a weak feature when used alone.

In our experiment, we use 4 kinds of bias patterns as shown in Fig. 9. A general concern about supervised classification, which is used in our approach, is the robustness of classifier on different datasets. To evaluate this, we perform a cross testing on different bias patterns, as shown in Fig. 11. We train one classifier on images with one specific bias pattern, then test its performance on the other bias pattern. The classifiers will always perform the best on the pattern they are trained on, which is the case for most supervised learning algorithm. However, the performance is still tolerable and much better than the baseline method when they are tested on unfamiliar patterns. We also found that if the classifier is trained on clean images, they will be sensitive to biased images, because they tend to believe that the same tissue share the same intensity.

D. Integrate rich-feature segmentation with bias field correction

Our main approach is to completely trust the rich feature segmentation result. Thus in the iteration of bias field correction, tissue labels are fixed on rich feature segmentation

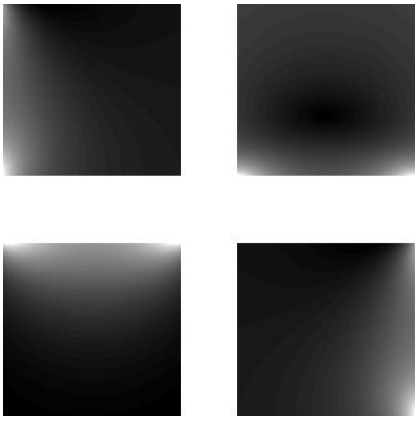


Figure 9. Four kinds of patterns.

labels. Another approach is to do rich feature segmentation and bias field correction recursively. In each iteration, the output of bias field correction will be labeled with rich feature segmentation, then the labels are used in next bias field correction iteration. However, in our experiments, we noticed that bias field correction is not gauranteed to be helpful for our rich feature segmentation, which is actually reasonable. Our segmentation is not purely intensitive based. When the classifier is trained on biasd images, there is no reason that they will perform better on less biased image after the bias field correction compared to the biased ones before that.

To quantively evaluate our bias field estimation and correction, we use normalized root of mean square error of the estimated bias field with respect to the ground truth. The result is shown in Fig. 14. Our proposed method was able to acheive more accurate bias field correction in all testing patterns than the baseline. However, both FCM-B and our approache performed badly in pattern 2, where bias field was extremly strong in the center of the image, not reflecting the realistic cases.

V. Conclusion

In this project, we replicated fuzzy c-means based bias field correction method, and showed that the bias field can be mostly corrected when it is not strong, and improves in average 9% on segmentation accuracy compared to fuzzy c-means segmentation without bias field correction method. Thus our results validate our first hypothesis in that the one with bias field correction method will outperform the one without bias field correction, but not as high as 15% improvement in segmentation accuracy as we previously expected. To improve the performance of bias field correction in cases of strong bias, we extended the previous fuzzy c-means based bias field correction with a rich feature based segmentation method, which provides the prior knowledge on the fuzzy memberships of each pixel to different tissue types. Our new segmentation method is based on supervised learning of the appearance model for different tissue types, and we are able to acheive above 60% tissue label accuracy even when the bias field is very strong, which is 51.1% higher than fuzzy c-means segmentation thus validating our second hypothesis. The bias field correction benefits from our segmentation method, and the estimation of the bias field is more accurate than the fuzzy c-means based method, specifically, the performance is boosted by 41.7%, thus validating our third hypothesis. For further exploration of this problem, a more regularization method adaptive to bias field with higher order variation.

References

[1] B. Aubert-Broche, A. C. Evans, and L. Collins, "A new improved version of the realistic digital brain phantom," *NeuroImage*, vol. 32, no. 1, pp. 138–145, 2006. http://brainweb.bic.mni.mcgill.ca/brainweb/.

[2] J. Wang, M. Qiu, Q. X. Yang, M. B. Smith, and R. T. Constable, "Measurement and correction of transmitter and receiver induced nonuniformities in vivo," *Magnetic resonance in medicine*, vol. 53, no. 2, pp. 408–417, 2005.

[3] H. Watanabe, N. Takaya, and F. Mitsumori, "Non-uniformity correction of human brain imaging at high field by RF field mapping of B+ and B-," *Journal of Magnetic Resonance*, vol. 212, no. 2, pp. 426–430, 2011. http://www.sciencedirect.com/science/article/pii/S1090780711002588.

[4] H. Mihara, N. Iriguchi, and S. Ueno, "A method of rf inhomogeneity correction in mr imaging," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 7, no. 2, pp. 115–120, 1998. http://www.sciencedirect.com/science/article/pii/S1352866198000672.

[5] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger, *et al.*, "Sense: sensitivity encoding for fast mri," *Magnetic resonance in medicine*, vol. 42, no. 5, pp. 952–962, 1999. http://macduff.usc.edu/ee591/library/Pruessmann-SENSE.pdf.

[6] M. Styner, C. Brechbuhler, G. Szckely, and G. Gerig, "Parametric estimate of intensity inhomogeneities applied to MRI," *Medical Imaging, IEEE Transactions on*, vol. 19, no. 3, pp. 153–165, 2000. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=845174.

[7] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *Medical Imaging, IEEE Transactions on*, vol. 17, no. 1, pp. 87–97, 1998. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=668698.

[8] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4itk: improved n3 bias correction," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 6, pp. 1310–1320, 2010. http://www.ncbi.nlm.nih.gov/pubmed/20378467.

[9] W. Chen and M. L. Giger, "A fuzzy c-means (fcm) based algorithm for intensity inhomogeneity correction and segmentation of mr images," in *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pp. 1307–1310, IEEE, 2004. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1398786.

[10] D. Hoiem, A. A. Efros, and M. Hebert, "Geometric context from a single image," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 1, pp. 654–661, IEEE, 2005. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1541316&tag=1.

[11] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, "A modified fuzzy c-means algorithm for bias field estimation and segmentation of mri data," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 3, pp. 193–199, 2002. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=996338.

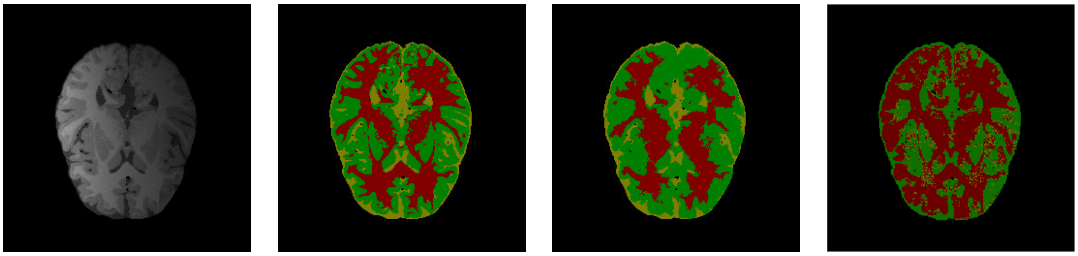
[12] S. Chen and D. Zhang, "Robust image segmentation using fcm with spatial constraints based on new kernel-induced distance measure," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 34, no. 4, pp. 1907–1916, 2004. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1315771.

[13] W. M. Wells III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of mri data," *Medical Imaging, IEEE Transactions on*, vol. 15, no. 4, pp. 429–442, 1996.

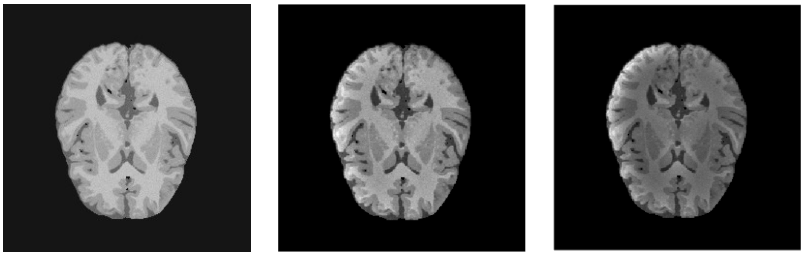
[14] A. H. Andersen, Z. Zhang, M. J. Avison, and D. M. Gash, "Automated segmentation of multispectral brain mr images," *Journal of neuroscience methods*, vol. 122, no. 1, pp. 13–23, 2002. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=511747.

[15] M. N. Ahmed, S. M. Yamany, A. A. Farag, and T. Moriarty, "Bias field estimation and adaptive segmentation of mri data using a modified fuzzy c-means algorithm," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, vol. 1, IEEE, 1999. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=786947.

[16] D. L. Pham and J. L. Prince, "An adaptive fuzzy c-means algorithm for image segmentation in the presence of intensity inhomogeneities," *Pattern Recognition Letters*, vol. 20, no. 1, pp. 57–68, 1999. http://www.sciencedirect.com/science/article/pii/S0167865598001214.



(a) Input: input image with bias field (b) Input: ground truth tissue label (c) Labeling of proposed method (d) Labeling of fuzzy c-means



(e) Ground truth image (f) Corrected image by proposed method (g) Corrected Image by fuzzy c-means

Figure 10. Sample results: tissue labeling and bias field correction.

classifier	No Bias	Pattern 1	Pattern 2	Pattern 3	Pattern 4	Mix
No Bias	**81.5**%	56.7%	68.9%	69.5%	56.6%	**79.2**%
Pattern 1	37.2%	**79.2**%	68.2%	71.4%	66.3%	**76.6**%
Pattern 2	47.8%	52.0%	**76.1**%	65.5%	51.2%	**69.8**%
Pattern 3	43.1%	64.5%	74.9%	**77.8**%	64.1%	**75.8**%
Pattern 4	39.9%	64.9%	67.1%	71.2%	**77.3**%	**75.2**%

Figure 11. Segmentation accuracy of different types of classifiers tested on different types of data. Each colomn shares the same classifier trained on images with certain bias pattern; each row shares the same testing images with certain bias pattern. The best and second accurate classifier for each row is highlighted in bold.

Methods	Proposed	FCM	FCM-B
Images with bias field	**73.4**%	22.3%	31.6%
Images without bias field	79.2%	**82.1**%	76.7%

Figure 12. Accuracy of segmentation. Proposed: our proposed rich-feature segmentation; FCM: cuzzy c-means segmentation without bias field estimation; FCM-B: fuzzy c-means segmatation with bias field estimation and B-spline regularization.

Features	All	Intensity	Texture	Location	shape
w/o	–	68.7%	72.7%	71.3%	73.0%
Only use	73.4%	63.3%	71.2%	63.9%	46.2%

Figure 13. Segmentation accuracy based on different feature selections. First row: we delete certain type of features in our feature list; Second row: We use only one type of feature in the feature list.

Bias pattern	Pattern 1	Pattern 2	Pattern 3	Pattern 4
Proposed	**0.1789**	**0.4300**	**0.1470**	**0.1568**
FCM-B	0.1944	0.4995	0.1643	0.1860

Figure 14. Normalized Root of Mean Square Error (NRMSE) of estimated bias field on different types of bias pattern.

[17] B. Likar, J. Derganc, and F. Pernus, "Segmentation-based retrospective correction of intensity nonuniformity in multispectral mr images," in *Medical Imaging 2002*, pp. 1531–1540, International Society for Optics and Photonics, 2002. http://www.sciencedirect.com/science/article/pii/S0167865598001214.

[18] J. Derganc, B. Likar, and F. Pernus, "Nonparametric segmentation of multispectral mr images incorporating spatial and intensity information," in *Medical Imaging 2002*, pp. 391–400, International Society for Optics and Photonics, 2002. http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=879732.

[19] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004. http://www.cs.cornell.edu/~dph/papers/seg-ijcv.pdf.

[20] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 8, pp. 1915–1929, 2013. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6338939.

[21] M. Collins, R. E. Schapire, and Y. Singer, "Logistic regression, adaboost and bregman distances," *Machine Learning*, vol. 48, no. 1-3, pp. 253–285, 2002. http://link.springer.com/article/10.1023/A:1013912006537.

[22] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995. http://link.springer.com/article/10.1007/BF00994018.

[23] S.-H. Hsu, Y. Cao, K. Huang, M. Feng, and J. M. Balter, "Investigation of a method for generating synthetic ct models from mri scans of the head and neck for radiation therapy," *Physics in medicine and biology*, vol. 58, no. 23, p. 8419, 2013. http://www.ncbi.nlm.nih.gov/pubmed/24217183.

[24] U. Vovk, F. Pernus, and B. Likar, "A review of methods for correction of intensity inhomogeneity in mri," *Medical Imaging, IEEE Transactions on*, vol. 26, no. 3, pp. 405–421, 2007. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4114560.

[25] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *International Journal of Computer Vision*, vol. 43, no. 1, pp. 29–44, 2001. http://link.springer.com/article/10.1023/A:1011126920638.

[26] M. J. Allison, S. Ramani, and J. A. Fessler, "Accelerated regularized estimation of mr coil sensitivities using augmented lagrangian methods," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 3, pp. 556–564, 2013. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6361300.

[27] J. A. Fessler, "Image reconstruction toolbox," http://web.eecs.umich.edu/~fessler/code/index.html.