# Adaptive Segmentation of MRI data

W. M. Wells III[123] , W.E.L. Grimson[24] , R. Kikinis[15] and F. A. Jolesz[16]

[1] Harvard Medical School and Brigham and Women's Hospital, Department of
Radiology, 75 Francis St., Boston, MA 02115
[2] Massachusetts Institute of Technology, Artificial Intelligence Laboratory, 545
Technology Square, Cambridge, MA 02139
[3] sw@ai.mit.edu  [4] welg@ai.mit.edu  [5] kikinis@bwh.harvard.edu
[6] jolesz@bwh.harvard.edu

**Abstract.** Intensity-based classification of MR images has proven prob-
lematic, even when advanced techniques are used. Intra-scan and inter-
scan intensity inhomogeneities are a common source of difficulty. While
reported methods have had some success in correcting intra-scan inhomo-
geneities, such methods require supervision for the individual scan. This
paper describes a new method called *adaptive segmentation* that uses
knowledge of tissue intensity properties and intensity inhomogeneities to
correct and segment MR images. Use of the EM algorithm leads to a
fully automatic method that allows for more accurate segmentation of
tissue types as well as better visualization of MRI data, that has proven
to be effective in a study that includes more than 1000 brain scans.

## 1 Introduction

Medical applications that use the morphologic contents of MRI frequently require
segmentation of the imaged volume into tissue types. Such tissue segmentation
is often achieved by applying statistical classification methods to the signal in-
tensities [1, 2], in conjunction with morphological image processing operations
[3, 4].

Conventional intensity-based classification of MR images has proven prob-
lematic, however, even when advanced techniques such as non-parametric, multi-
channel methods are used. Intra-scan intensity inhomogeneities due to RF coils
or acquisition sequences (e.g. susceptibility artifacts in gradient echo images) are
a common source of difficulty. In addition, the operating conditions and status of
the MR equipment frequently affect the observed intensities, causing significant
inter-scan intensity inhomogeneities that often necessitate manual training on a
per-scan basis. While reported methods [5, 6, 7, 8, 9, 10] have had some success
in correcting intra-scan inhomogeneities, such methods require supervision for
the individual scan.

This paper describes a new method called *adaptive segmentation* that uses
knowledge of tissue properties and intensity inhomogeneities to correct and seg-
ment MR images. Use of the expectation-maximization algorithm leads to a
method that allows for more accurate segmentation of tissue types as well as
better visualization of MRI data. Adaptive segmentation has proven to be an

effective fully-automatic means of segmenting brain tissue in a study including more than 1000 brain scans.

## 2 Description of Method

### 2.1 Bias Field Estimator

We use a Bayesian approach to estimating the bias field that represents the gain artifact in log-transformed MR intensity data. We first compute a logarithmic transformation of the intensity data as follows,

$$Y_i = g(X_i) = (\ln([X_i]_1), \ln([X_i]_2), \dots, \ln([X_i]_m))^T \ , \tag{1}$$

where $X_i$ is the observed MRI signal intensity at the $i$-th voxel, and $m$ is the dimension of the MRI signal.

Similar to other statistical approaches to intensity-based segmentation of MRI [3, 4], the distribution for observed values is modeled as a normal distribution (with the incorporation of an explicit bias field):

$$p(Y_i \mid \Gamma_i, \beta_i) = \mathrm{G}_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i) \ , \tag{2}$$

where

$$\mathrm{G}_{\psi_{\Gamma_i}}(x) \equiv (2\pi)^{-\frac{m}{2}} |\psi_{\Gamma_i}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}x^T \psi_{\Gamma_i}^{-1} x\right)$$

is the $m$-dimensional Gaussian distribution with variance $\psi_{\Gamma_i}$, and where

$Y_i$    is the observed log-transformed intensities at the $i^{th}$ voxel
$\Gamma_i$    is the tissue class at the $i^{th}$ voxel
$\mu(x)$ is the mean intensity for tissue class $x$
$\psi_x$    is the covariance matrix for tissue class $x$
$\beta_i$    is the bias field at the $i^{th}$ voxel.

Here $Y_i$, $\mu(x)$, and $\beta_i$ are represented by $m$-dimensional column vectors, while $\psi_x$ is represented by an $m$-by-$m$ matrix. Note that the bias field has a separate value for each component of the log-intensity signal at each voxel.

A stationary prior (before the image data is seen) probability distribution on tissue class is used, it is denoted $p(\Gamma_i)$. If this probability is uniform over tissue classes, our method devolves to a maximum-likelihood approach to the tissue classification component. A spatially-varying prior probability density on brain tissue class is described in [11]. Such a model might profitably be used within this framework.

The entire bias field is denoted by $\beta = (\beta_0, \beta_1, \dots, \beta_{n-1})^T$, where $n$ is the number of voxels of data. The bias field is modeled by a $n$-dimensional zero mean Gaussian prior probability density. This model allows us to capture the smoothness that is apparent in these inhomogeneities:

$$p(\beta) = \mathrm{G}_{\psi_\beta}(\beta) \ , \text{where} \ \mathrm{G}_{\psi_\beta}(x) \equiv (2\pi)^{-\frac{n}{2}} |\psi_\beta|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}x^T \psi_\beta^{-1} x\right) \tag{3}$$

is the $n$-dimensional Gaussian distribution. The $n \times n$ covariance matrix for the entire bias field is denoted $\psi_\beta$. Although $\psi_\beta$ will be too large to manipulate directly in practice, we will show below that tractable estimators result when $\psi_\beta$ is chosen so that it is banded.

We assume that the bias field and the tissue classes are statistically independent, this follows if the intensity inhomogeneities originate in the equipment. Using the definition of conditional probability and computing a marginal over tissue class leads to the conditional probability of intensity alone:

$$p(Y_i \mid \beta_i) = \sum_{\Gamma_i} p(Y_i, \Gamma_i \mid \beta_i) = \sum_{\Gamma_i} p(Y_i \mid \Gamma_i, \beta_i) p(\Gamma_i) \ . \tag{4}$$

Thus, our modeling has led to a class-independent intensity distribution that is a mixture of Gaussian populations (one population for each tissue class). Since this model is a Gaussian mixture, rather than a purely Gaussian distribution, the estimators that we derive below will be non-linear.

We assume statistical independence of the voxel intensities (in other words, the noise in the MR signal is spatially white).

Bayes' rule may then be used to obtain the posterior probability of the bias field, given observed intensity data as follows, $p(\beta \mid Y) = \frac{p(Y \mid \beta) p(\beta)}{p(Y)}$, where $p(Y)$ is an unimportant normalizing constant.

Having obtained the posterior probability on the bias field, we now use the maximum-a-posteriori (MAP) principle to formulate an estimate of the bias field as the value of $\beta$ having the largest posterior probability,

$$\hat{\beta} = \arg\max_\beta p(\beta \mid Y) \ . \tag{5}$$

A necessary condition for a maximum of the posterior probability of $\beta$ is that its gradient with respect to $\beta$ be zero. We use an equivalent zero-gradient condition on the logarithm of the posterior probability,

$$\left[ \frac{\partial}{\partial [\beta_i]_k} \ln p(\beta \mid Y) \right]_{\beta = \hat{\beta}} = 0 \qquad \forall_{i,k} \ , \tag{6}$$

where $[\beta_i]_k$ is the $k$-th component of the bias field at voxel $i$. Installing the statistical modeling of (2-5) yields the following expression for the zero gradient condition:

$$\left[ \frac{\partial}{\partial [\beta_i]_k} \left( \sum_j \ln p(Y_j \mid \beta_j) + \ln p(\beta) \right) \right]_{\beta = \hat{\beta}} = 0 \qquad \forall_{i,k} \ .$$

Equations 2 and 4, and the fact that only the $i$-th term of the sum depends on $\beta_i$ leads to:

$$\left[ \frac{\sum_{\Gamma_i} p(\Gamma_i) G_{\psi_{\Gamma_i}} (Y_i - \mu(\Gamma_i) - \beta_i) \left[ \psi_{\Gamma_i}^{-1} (Y_i - \mu(\Gamma_i) - \beta_i) \right]_k}{\sum_{\Gamma_i} p(\Gamma_i) G_{\psi_{\Gamma_i}} (Y_i - \mu(\Gamma_i) - \beta_i)} + \frac{\frac{\partial}{\partial [\beta_i]_k} p(\beta)}{p(\beta)} \right]_{\beta = \hat{\beta}} = 0$$

$\forall_{i,k}$.

This expression may be written more compactly as

$$\left[\sum_j W_{ij}\left[\psi_j^{-1}(Y_i - \mu_j - \beta_i)\right]_k + \frac{\frac{\partial}{\partial[\beta_i]_k}p(\beta)}{p(\beta)}\right]_{\beta=\hat{\beta}} = 0 \qquad \forall_{i,k} \quad , \qquad (7)$$

with the following definition of $W_{ij}$, (which are called the *weights*),

$$W_{ij} \equiv \frac{\left[p(\Gamma_i)\mathrm{G}_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i)\right]_{\Gamma_i=\textit{tissue-class-j}}}{\sum_{\Gamma_i} p(\Gamma_i)\mathrm{G}_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i)} \quad , \qquad (8)$$

where subscripts $i$ and $j$ refer to voxel index and tissue class respectively, and defining $\mu_j \equiv \mu(\textit{tissue-class-j})$ as the mean intensity of tissue class $j$. Equation 7 may be re-expressed as follows,

$$\left[\left[\overline{R}_i\right]_k - \left[\overline{\psi^{-1}}_{ii}\beta_i\right]_k + \frac{\frac{\partial}{\partial[\beta_i]_k}p(\beta)}{p(\beta)}\right]_{\beta=\hat{\beta}} = 0 \qquad \forall_{i,k} \qquad (9)$$

with the following definitions for the mean residual, and the mean inverse co-variance,

$$\overline{R}_i \equiv \sum_j W_{ij}\psi_j^{-1}(Y_i - \mu_j) \ , \ \overline{\psi^{-1}}_{ik} \equiv \begin{cases} \sum_j W_{ij}\psi_j^{-1} & \text{if } i = k \\ 0 & \text{otherwise} \end{cases} \ . \qquad (10)$$

The mean residuals and mean inverse covariances defined above are averages taken over the tissue classes, weighted according to $W_{ij}$.

Equation 9 may be re-expressed in matrix notation as

$$\left[\overline{R} - \overline{\psi^{-1}}\beta + \frac{\nabla_\beta p(\beta)}{p(\beta)}\right]_{\beta=\hat{\beta}} = \mathbf{0} \ .$$

After differentiating the last term, the zero-gradient condition for the bias field estimator may be written concisely as

$$\hat{\beta} = H\overline{R} \ , \qquad (11)$$

where the linear operator $H$ is defined by

$$H \equiv \left[\overline{\psi^{-1}} + \psi_\beta^{-1}\right]^{-1} \ , \qquad (12)$$

that is, the bias field estimate is derived by applying the linear operator $H$ to the mean residual field, and $H$ is determined by the mean covariance of the tissue class intensities and the covariance of the bias field.

The bias field estimator of 11 has some resemblance to being a linear estimator in $Y$ of the bias field $\beta$. It is not a linear estimator, however, owing to the fact that the $W_{ij}$ (the "weights") that appear in the expression for $\overline{R}$ and $H$ are themselves non-linear functions of $Y$ (8).

The result of the statistical modeling in this section has been to formulate the problem of estimating the bias field as a non-linear optimization problem embodied in (11).

## 2.2 EM Algorithm

We use the expectation-maximization (EM) algorithm to obtain bias field estimates from the non-linear estimator of (11). The EM algorithm was originally described in its general form by Dempster, Laird and Rubin [12]. It is often used in estimation problems where some of the data are "missing." In this application, the missing data is knowledge of the tissue classes. (If they were known, then estimating the bias field would be straightforward.)

In this application, the EM algorithm iteratively alternates evaluations of the expressions appearing in (11) and (8).

In other words, (8) is used to estimate the weights given an estimated bias field, then (11) is used to estimate the bias, given estimates of the weights.

As frequently occurs in application of the EM algorithm, the two components of the iteration have simple interpretations. Equation (8) (the *E-Step*) is equivalent to calculating the posterior tissue class probabilities (a good indicator of tissue class) when the bias field is known. Equation (11) (the *M-Step*) is equivalent to a MAP estimator of the bias field when the tissue probabilities $W$ are known.

The iteration may be started on either expression. Initial values for the weights will be needed to start with (11), and initial values for the bias field will be needed to start with (8).

It is shown in [12] that in many cases the EM algorithm enjoys pleasant convergence properties – namely that iterations will never worsen the value of the objective function. Provided that the bias estimates are bounded, our model satisfies the necessary conditions for guaranteed convergence (although there is no guarantee of convergence to the global minimum).

In principle, given $\mu(\Gamma_i)$, $\psi_\beta$, and $\psi_j$, we could use the EM algorithm to obtain the needed estimates. In practice, we cannot directly measure $\psi_\beta$, and thus we will seek other estimates of $\psi_\beta$ (in Sect. 2.3).

Although the covariance matrix $\psi_\beta$ that characterizes the prior on bias fields is impractically large in general, tractable estimation algorithms may yet be obtained.

$\psi_\beta$ may be chosen to be a banded matrix. If it is then factored out of $H$, the bias estimation step (11) may be implemented as the solution of a banded linear system (note that $\overline{\psi^{-1}}$ is diagonal).

## 2.3 Determination of the Linear Operator H

We have taken a Bayesian approach to estimating the bias field and tissue classes, and a formal prior model on bias fields has been assumed. This approach has allowed us to derive a version of the EM algorithm for this application. The operator $H$ is related to the prior on the bias field via $\psi_\beta^{-1}$ and to the measurement noise via $\overline{\psi^{-1}}$ (12) . Ideally, $H$ would be determined by estimating the covariance $\psi_\beta$, but given the size of this matrix, such an approach is impractical.

As pointed out above, $H$ is the MAP estimator of the bias field when the tissue probabilities are known, (the "complete-data" case with the EM algorithm).

As such, $H$ is an optimal estimator (with respect to the Gaussian modeling), and is also the optimal *linear* least squares estimator (LLSE) for arbitrary zero-mean models of the bias field whose second-order statistics are characterized by $\psi_\beta$.

A frequent problem that arises in filter design (the present complete-data case included) is that of estimating a slowly-varying signal that has been contaminated with white noise. The optimal filter in such situations will be a low-pass filter [13, Sect. 9.2].

In practice, it is difficult to obtain *the* optimal linear filter. $H$ may be instead chosen as a good engineering approximation of the optimal linear filter (this approach is described in more detail below). In this case, (8) and (11) are still a useful estimator for the missing data case, and the good convergence properties of the EM algorithm still apply. This is the approach we have taken in our implementations, where the filter was selected empirically.

While the low-pass filters $H$ we have used in practice are not the optimal filters for estimating these bias fields, they are reasonable choices, and may correspond to reasonable subjective estimates of the unknown probability law for bias fields, in the sense described by Friden [14, Chapt. 16]. In the end, they are justified empirically by the good results obtained via their use. Because $\psi_\beta$ is required to be positive definite, not all choices of low-pass filter $H$ will correspond to formally valid prior models on the bias field.

**Computationally Efficient Filter.** As argued above, the optimal $H$ will be a linear low-pass filter, when tissue class is constant. We have employed a particularly efficient filter that is characterized as follows

$$\hat{\beta}_i = \frac{[F\overline{R}]_i}{[F\overline{\psi^{-1}}\mathbf{1}]_i} \quad \text{where} \quad \mathbf{1} \equiv (1, 1, 1, \ldots, 1)^T \tag{13}$$

The filter specified above is clearly linear in the mean residual, and it will be a low-pass filter when the tissue class is constant, provided that $F$ is a low-pass filter. It has been designed to have unity DC gain – a spatially constant shift in $Y$ induces the same constant shift in $\hat{\beta}$. If $F$ is chosen to be a computationally efficient low-pass filter, then the filter specified by (13) will also be computationally efficient.

### 2.4   Equal Covariance Case

The formalism simplifies somewhat when the tissue classes have the same covariance. This case, for scalar data, was previously reported in [15], along with scalar formulas for the weights. The bias estimator is then particularly simple when the bias model is stationary. It is a shift-invariant linear low-pass filter applied to the difference between the observed intensities and a prediction of the signal that is based on the weights (which are a good estimator of tissue class).

### 2.5 Non-Parametric Extension

The method that was described in previous sections has two main components: tissue classification and bias field estimation. Our approach in the extended method has been to use the same basic iteration, and to replace the tissue classification component with the technique described in [16]. The classifier described in [16] uses the *Parzen Window* representation for non-parametric probability densities [17] that are derived from training data

The non-parametric tissue class conditional intensity models are derived from training in the "natural" MR intensities. In view of our logarithmic transformation of the intensity data (1), we use the standard formula for transforming probability densities.

In the expressions for the average residual (10) and average covariance (10) we approximate with the empirical tissue class means and covariances from the log transformed training data.

The resulting iterative algorithm is a simple generalization from the Gaussian theory developed in the previous sections. Results obtained using the method are described in Sect. 3.

## 3   Results

This section describes results recently obtained for segmenting MR images from a large, longitudinal study of several dozen patients with multiple sclerosis (MS) [18].

All of the MR images shown in this setion were obtained using a General Electric Signa 1.5 Tesla clinical MR imager [19]. An anisotropic diffusion filter developed by Gerig et al. [20] was used as a pre-processing step to reduce noise.

We used an implementation of the non-parametric extension that is described in Sect. 2.5. This implementation is coded in the C programming language. It accommodates 2 channel data (typically registered proton-density and T2-weighted images), and multiple (more than two) tissue classes having un-equal covariances. Because it can model the important intensities in the imagery (including the background signal) it is able to correct and segment brain images without the need for a previously generated ROI. It uses the computationally-efficient filter described in Sect. 2.3, $F$ is implemented as a moving average filter. Both uniform and non-uniform distributions have been used for the prior on tissue class.

In a typical case, the program was run until the estimates stabilized, typically in $5-10$ iterations, requiring approximately 2 seconds per iteration (per $256^2$ slice pair) on a Sun Microsystems Sparcstation 10 [21].

The method has been found to be substantially insensitive to parameter settings. For a given type of acquisition, intensity variations across patients, scans, and equipment changes have been accommodated in the estimated bias fields without the need for manual intervention. In this sense, the method is fully automatic for segmenting healthy brain tissue.

The data comprised registered proton-density and T2-weighted images for a single multiple-sclerosis patient with multiple white matter lesions. These images represented the same section from 20 scans over time after they were spatially registered using the method described in [22, 23]. The same tissue class conditional intensity models were used to segment all sections.

As expected, the results without intensity correction were dissapointing. These results are equivalent to those which would be obtained using conventional non-parametric intensity-based segmentation (which would more typically be used with per-scan manual training). They showed many gross misclassifications and demonstrated that conventional intensity-based segmentation is unfeasible in this application, at least without per-scan training. Even with per-scan training, significant asymmetries remained in the results due to the spatial intensity inhomogeneities present in the data.

Results using adaptive segmentation are shown in Fig. 1. Tissues are encoded from black to white as follows: background, subcutaneous fat, gray matter, CSF, lesions, white matter. Good stability and symmetry of the cortical gray matter structures are apparent. Similar results have been obtained in processing 23 complete scans for each of 47 patients participating in the study mentioned above, without the need for retraining or manually-generating regions-of-interest. This has facilitated monitoring the evolution of specific white matter lesions over time. Thus, fully automatic segmentation of clinical MRI data has been demonstrated in more than 1000 complete scans, without the need for per-patient or per-scan training or adjustments. The exams occurred over a 2.5 year period that included a major MR equipment upgrade.

The algorithm embodies an interative solver for a non-linear optimization problem, and like other local methods, there is no guarantee of convergence to the global solution. We have found that in practical applications, such as the one described here, the algorithm will reliably converge to a satisfactory solution once the appropriate tissue model has been identified.

## 4   Discussion

The use of multi-channel statistical intensity classifiers was pioneered by Vannier et al.[1]. The classification component of adaptive segmentation is similar to the method described by Gerig et al. and Cline et al. [3, 4]. The classification component of the non-parametric extended method is equivalent to that described [16].

The bias field estimation component of adaptive segmentation method is somewhat similar to homomorphic filtering (HMF) approaches that have been reported. Lufkin et al. [5] and Axel et al. [6] describe approaches for controlling the dynamic range of surface-coil MR images. Lim and Pfferbaum [7] use a similar approach to filtering that handles the boundary in a novel way, and apply intensity-based segmentation to the result.

When started on the "M Step", and run for one cycle, adaptive segmentation is equivalent to HMF followed by conventional intensity-based segmentation. We
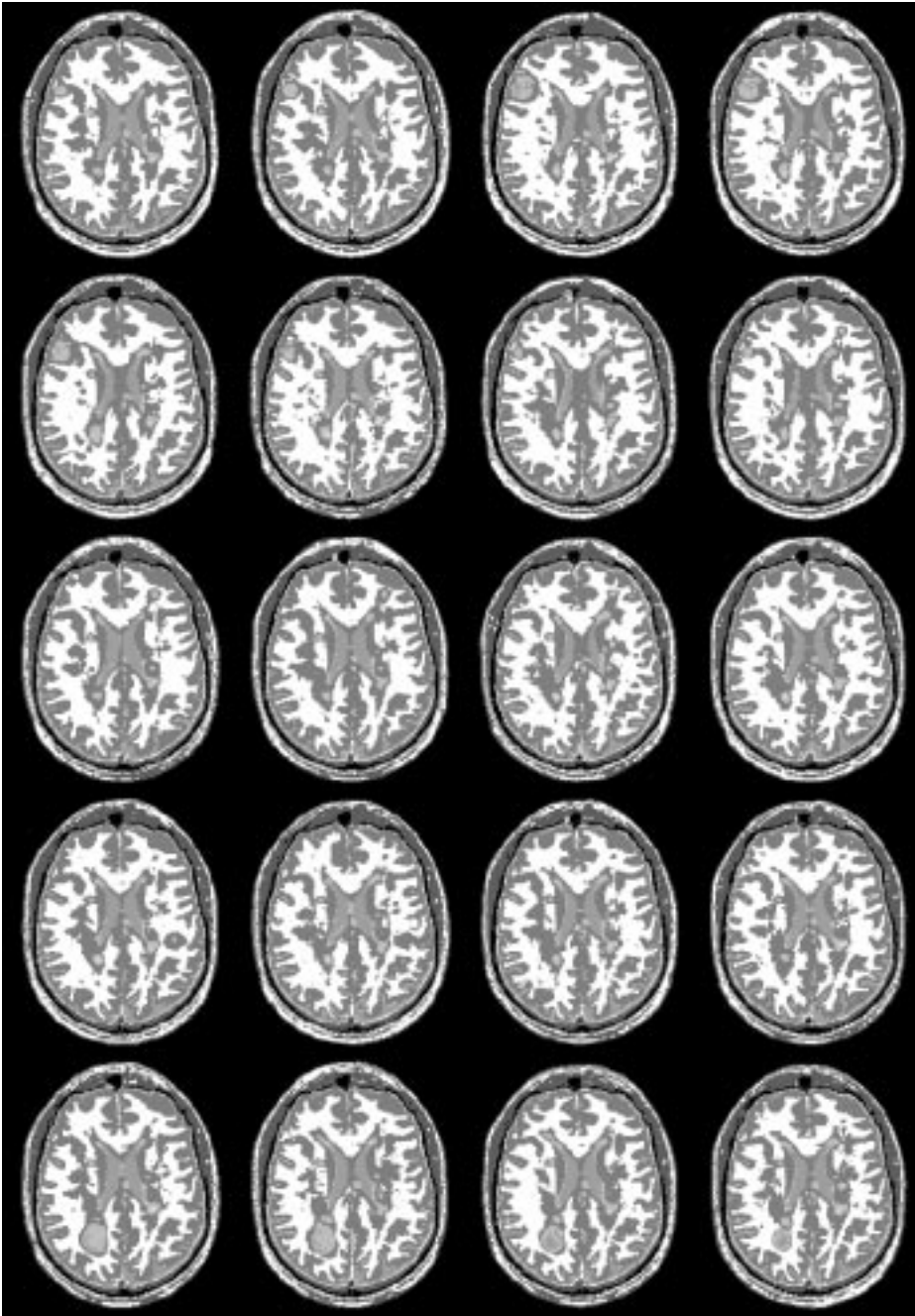
**Fig. 1.** Adaptive segmentation results (without per-scan training)

have discovered, however, that more than one iteration are frequently needed to converge to good results – indicating that adaptive segmentation is more powerful than HMF followed by intensity-based segmentation. The essential difference is that adaptive segmentation utilizes evolving knowledge of the tissue type to make increasingly accurate estimates of the gain field.

Dawant, Zijdenbos and Margolin describe methods for correcting intensities for tissue classification [8]. In one variant, an operator selected points of a tissue class are used to regress an intensity correction. In the other method, a preliminary segmentation is used in determining an intensity correction, which is then used for improved segmentation. This strategy is somewhat analogous to starting adaptive segmentation on the "E step" and running it for one and a half cycles. As in the previous case, our results demonstrate improvement with additional iterations.

Aylward and Coggins describe a two-stage approach that first uses a band-pass intensity corrector. Remaining inhomogeneities are handled by using supervised training to obtain spatially-varying statistics for classifying the corrected MR data [10].

Several authors have reported methods based on the use of phantoms for intensity calibration [6, 9]. This approach has the drawback that the geometric relationship of the coils and the image data is not typically available with the image data (especially with surface coils). Fiducial markers were used to address this problem in [9]. In addition, the calibration approach can become complicated because the response of tissue to varying amounts of RF excitation is significantly non-linear (see [24, Equations 1-3 and 1-16]). In addition, phantom calibration cannot account for possible gain inhomogeneities induced by the interaction of anatomy and the RF coils.

## 5    Acknowledgments

## References

1. M. Vannier, R. Butterfield, D. Jordan, W. Murphy, et al. Multi-Spectral Analysis of Magnetic Resonance Images. *Radiology*, (154):221 – 224, 1985.
2. M. Kohn, N. Tanna, G. Herman, et al. Analysis of Brain and Cerebrospinal Fluid Volumes with MR Imaging. *Radiology*, (178):115 – 122, 1991.
3. G. Gerig, W. Kuoni, R. Kikinis, and O. Kübler. Medical Imaging and Computer Vision: an Integrated Approach for Diagnosis and Planning. In *Proc. 11'th DAGM Symposium*, pages 425–443. Springer, 1989.
4. H.E. Cline, W.E. Lorensen, R. Kikinis, and F. Jolesz. Three-Dimensional segmentation of MR images of the head using probability and connectivity. *JCAT*, 14(6):1037–1045, 1990.

5. R.B. Lufkin, T. Sharpless, B. Flannigan, and W. Hanafee. Dynamic-Range Compression in Surface-Coil MRI. *AJR*, 147(379):379–382, 1986.
6. L. Axel, J. Costantini, and J. Listerud. Intensity Correction in Surface-Coil MR Imaging. *AJR*, 148(4):418–420, 1987.
7. K.O. Lim and A. Pfferbaum. Segmentation of MR Brain Images into Cerebrospinal Fluid Spaces, White and Gray Matter. *JCAT*, 13(4):588–593, 1989.
8. B. Dawant, A. Zijdenbos, and R. Margolin. Correction of Intensity Variations in MR Images for Computer-Aided Tissue Classification. *IEEE Trans. Med. Imaging*, 12(4):770 – 781, 1993.
9. J. Gohagan, E. Spitznagel, W. Murphy, M. Vannier, et al. Multispectral Analysis of MR Images of the Breast. *Radiology*, (163):703 – 707, 1987.
10. S. Aylward and J. Coggins. Spatially Invariant Classification of Tissues in MR Images. In *Proceedings of the Third Conference on Visualization in Biomedical Computing*. SPIE, 1994. to appear.
11. M. Kamber, D. Collins, R. Shinghal, G. Francis, and A. Evans. Model-Based 3D Segmentation of Multiple Sclerosis Lesions in Dual-Echo MRI Data. In *SPIE Vol. 1808, Visualization in Biomedical Computing 1992*, 1992.
12. A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. Roy. Statist. Soc.*, 39:1 – 38, 1977.
13. J.S. Lim. *Two-Dimensional Signal and Image Processing*. Prentice Hall, 1990.
14. B.R. Frieden. *Probability, Statistical Optics, and Data Testing*. Springer-Verlag, 1983.
15. W. Wells, R. Kikinis, W. Grimson, and F. Jolesz. Statistical Intensity Correction and Segmentation of Magnetic Resonance Image Data. In *Proceedings of the Third Conference on Visualization in Biomedical Computing*. SPIE, 1994. to appear.
16. R. Kikinis, M. Shenton, F.A. Jolesz, G. Gerig, J. Martin, M. Anderson, D. Metcalf, C. Guttmann, R.W. McCarley, W. Lorensen, and H. Cline. Quantitative Analysis of Brain and Cerebrospinal Fluid Spaces with MR Imaging. *JMRI*, 2:619–629, 1992.
17. R.O. Duda and P.E. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, 1973.
18. Ron Kikinis et al. in preparation.
19. General Electric Medical Systems, Milwaukee, WI.
20. G. Gerig, O. Kübler, and F. Jolesz. Nonlinear Anisotropic Filtering of MRI data. *IEEE Trans. Med. Imaging*, (11):221–232, 1992.
21. Sun Microsystems Inc., Mountain View, CA.
22. G. Ettinger, W. Grimson, T. Lozano-Pérez, W. Wells, S. White, and R. Kikinis. Automatic Registration for Multiple Sclerosis Change Detection. In *Proceedings of the IEEE Workshop on Biomedical Image Analysis*, Seattle, WA., 1994. IEEE.
23. W.E.L. Grimson, T. Lozano-Pérez, W. Wells, et al. An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced realigy visualization. In *Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, WA., June 1994. IEEE.
24. D. Stark and Jr. W. Bradley, editors. *Magnetic Resonance Imaging*. Mosby Year Book, 1990.