A Non-parametric Method for Automatic Correction of Intensity Non-uniformity in MRI Data

John G. Sled

Department of Biomedical Engineering

McGill University, Montréal

May, 1997

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Master's of Engineering.

Abstract

A novel approach to correcting for intensity non-uniformity in MR data is described that achieves high performance without requiring supervision. By making relatively few assumptions about the data, the method can be applied at an early stage in an automated data analysis, before a tissue intensity or geometric model is available. Described as Non-parametric Non-uniform intensity Normalization (N3), the method is independent of pulse sequence and insensitive to pathological data that might otherwise violate model assumptions. To eliminate the dependence of the field estimate on anatomy, an iterative approach is employed to estimate both the multiplicative bias field and the distribution of the true tissue intensities. The performance of this method is evaluated using both real and simulated MR data. Preprocessing of MR data using N3 is shown to substantially improve the accuracy of anatomical analysis techniques such as tissue classification and cortical surface extraction.

Résumé

Une nouvelle façon de corriger le problème de non-uniformité d'intensité d'images de résonnance magnétique (RM) est décrite. Cette méthode atteint de hautes performances sans surveillance. Etant donné que la méthode fait peu de suppositions issues des données, elle peut être utilisée dès le premier stade d'analyse automatique, avant même qu'un modèle d'intensité des tissus ou de géometrie ne devienne disponible. Décrite comme une Normalization Non-paramétrique d'intensité Non-uniforme (N3), cette méthode est indépendante de la série de RM et est insensible aux données pathologiques qui auraient violé les suppositions du modèle. Pour que la correction estimée ne dépende pas de l'anatomie, un algorithme itératif est utilisé pour estimer le champ multiplicatif correctif et la distribution d'intensités du tissu reel. La performance de cette méthode est évaluée avec les données de RM réelles et simulées. Le traitement des données RM avec N3 fait preuve d'une grande amélioration sur la précision d'analyse anatomique comme la classification de tissus et la démarcation de la surface du cortex.

Acknowledgments

I am grateful to a number of people for the support and assistance that has made this thesis possible. I would like to thank Vasco Kollokian, Greg Ward, David MacDonald, Peter Neelin, and Louis Collins for providing me with helpful advice and a multitude of software tools and libraries. I would like to thank Colin Holmes for providing me with his average MRI scans, Noor Kabani and others for creating the anatomical model from which my simulated MRI scans are derived, Bruce Pike for insight into MR physics, Chris Cocosco for proofreading, and Veronique Bohbot for helping with the French translation. Most of all, I would like to thank Alex Zijdenbos and Alan Evans who have provided support and guidance at every stage of this work.

In addition, I would like to acknowledge the financial support I have received from the Natural Sciences and Engineering Research Council of Canada (NSERC), for which I am grateful.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

An artifact often seen in MRI is for the signal intensity to vary smoothly across an image. Variously referred to as RF inhomogeneity, shading artifact, or intensity non-uniformity, it is usually attributed to such factors as poor radio frequency (RF) field uniformity, eddy currents driven by the switching of field gradients, and patient anatomy both inside and outside the field of view.

In the past, much of the effort on the correction of intensity non-uniformity has been directed at reducing the extreme variations seen in surface coil images. However, recent interest in automatic segmentation has driven expectations of better uniformity for routine volume acquisitions. While the intensity variations of between 10% and 30% often seen in clinical scanners have little impact on visual diagnosis, the performance of automatic segmentation techniques, which assume homogeneity of intensity within each class, can be significantly degraded by clinically acceptable levels of intensity non-uniformity. As automated methods are increasingly used to process large volumes of data for longitudinal studies and clinical trials [15], a robust, automatic, and inexpensive means of correcting for this artifact is essential for such techniques to be accurate in labelling each voxel with a tissue type.

In considering MR intensity non-uniformity it is important to distinguish between the rapid inter-slice variations sometimes observed with two dimensional multi-slice sequences and the smooth intensity variations present in most acquisitions including those using 3D sequences. The former can be dealt with by methods which normalize the intensities of individual slices [6, 43, 46, 58, 60]. Correction for the latter is the subject of this thesis.

A wide variety of techniques have been proposed to correct for intensity non-uniformity. These can be categorized as those relying on theoretical modelling, those requiring specialized scanning techniques, and those that estimate the non-uniformity directly from the data. Theoretical modelling techniques have not gained widespread acceptance because detailed models of the MR machine and subject are seldom available in a clinical setting. Similarly, specialized scanning is often impractical and precludes retrospective correction of the data. Methods that estimate non-uniformity directly from the data are the most practical and economical; however, the need for expert supervision, in choosing a set of sample voxels expected to have similar intensities for example, has prevented their widespread use.

This thesis describes a novel method of retrospectively correcting for intensity non-uniformity in clinical MR scans of the head. Referred to as non-parametric intensity non-uniformity normalization (N3), the method is shown to be robust, accurate, and fully automatic. These results are validated using both real and simulated MR data. In comparison with other methods, the N3 method is shown to perform favorably, which given that the method is automated and requires few assumptions about the nature of the data, makes it suitable for a broad range of applications. As examples, the tasks of tissue classification, cortical surface extraction, and non-linear co-registration are demonstrated to benefit from using N3.

Chapter 2

Intensity non-uniformity

2.1 Physical causes of non-uniformity

Intensity non-uniformity is the smooth intensity variation often seen in MR images caused by such factors as:

- inhomogeneous radio-frequency (RF) excitation
- non-uniform reception sensitivity
- electrodynamic interactions with the object often described as RF penetration and standing wave effects.

Other less important effects contributing to non-uniformity include:

- eddy currents driven by the switching of field gradients
- mistuning of the RF coil
- bandwidth filtering of the data
- geometric distortion.

Reviewing this list of causes in reverse order, the impact of geometric distortion on intensity non-uniformity can normally be neglected for clinical scanners. Since routine calibrations show at most a few millimetres of in plane geometric distortion over a field of view the size of a head, and these distortion fields are slowly varying, one can expect corresponding changes in intensity on the order of 1%, which is small compared to the non-uniformity caused by other factors. Investigation by Simmons et al. [46] did not find intensity variations caused by geometric distortion in a 1.5 T scanner.

In MR, the imaging process encodes spatial location in terms of resonant frequency; hence, frequency dependent variations in the sensitivity of the RF reception system cause corresponding intensity variations in the resulting image along its frequency encoding direction. A similar problem arises if the filter used to limit the bandwidth of the received signal has a frequency response that is non-uniform. Since the sensitivity of an RF coil is most uniform near its resonant frequency and the coil's resonant frequency depends on the subject within, intensity uniformity is improved by tuning the coil to the MR resonant frequency before each scan. Modern scanners tune themselves to each patient and use filters with relatively flat frequency responses; hence, intensity variations due to mistuning of the RF coil and bandwidth filtering of the data can be considered deficiencies in the scanner.

Simmons et al. [46] have observed that intensity non-uniformity increases as the repetition time of the acquisition sequence decreases. Since the frequency of field gradient switching also increases with decreasing repetition time, it has been postulated that this non-uniformity is caused by eddy currents in the subject induced by the switching of gradients. Although the magnitude of this effect is currently small compared to other effects, the trend in MR technology is towards faster switching gradients for more rapid image acquisition, so the significance of this effect should increase.

MR imaging involves excitation of the subject using a radio-frequency (RF) magnetic field and the subsequent reception of the RF magnetic field emitted by the subject. Spatial variations in both excitation field strength and reception sensitivity can produce significant variations in measured signal intensity. These effects have received considerable attention in the literature and have led to the use of the somewhat misleading term "RF inhomogeneity" being used to describe intensity non-uniformity.

Sensitivity is the ratio of measured signal intensity to emitted signal intensity. As the sensitivity of the reception coil and data acquisition system is generally assumed to be linear, spatial variations in sensitivity can be modelled as a multiplicative field that scales the measured intensities. Given the simple geometry of the receiving coil, this multiplicative field can be assumed to vary smoothly across the field of view.

The relationship between the strength of the excitation field and the measured intensity is non-linear and may also depend on the magnetic relaxation properties, T_1 and T_2, of the subject. Hence, the variations in intensity caused by variations in the excitation field strength cannot be exactly modelled by a smoothly varying multiplicative field.

Assuming linearity of the RF system and coils, one can expect sensitivity variations to have the same form and relative magnitude as variations in the excitation field strength. However, since the impact of sensitivity variations on the measured intensity is more simply modeled than those due to excitation, steps are usually taken to minimize the intensity variations introduced by excitation. For example, standard spin echo sequences are designed to minimize flip angle errors caused by variations in the excitation field. In addition, as signal to noise ratio is not an issue during excitation, a body coil is often used for excitation during head scans to take advantage of the greater uniformity of the larger coil. These factors have led to the widely used model of intensity non-uniformity as a multiplicative field corrupting the measured intensities. The validity of this model is considered further in Section 2.3.2.

2.2 Review of intensity non-uniformity models

A number of authors have conducted theoretical analyses of the excitation field (B_1) and reception sensitivity variations that contribute to intensity non-uniformity. This kind of analysis is greatly simplified by treating the excitation field as static and using the Biot-Savart law to compute magnetic field strength by integrating the contributions of nearby currents [36, 38]. While this approach simplifies the modelling of complicated coil shapes, it neglects penetration and standing wave effects that are present under the true dynamic conditions. Since the MR resonant frequency increases with main magnetic field strength B_0, this approximation of a quasi-static excitation field becomes less valid at higher field strengths.

The results of an analysis assuming static fields will at best match measurements made by placing a small pickup coil within the unloaded RF coil. Head coils have been measured in this manner to have in-plane uniformity with as little as 1% variation when operating at 64 MHz (suitable for a standard 1.5 T magnet). However, since intensity variations of 20% are typically observed at this field strength, a model that predicts 1% variation is clearly unsatisfactory.

Initial attempts at dynamic modelling of the MR imaging process [4, 35] predicted that penetration effects would significantly degrade performance at frequencies above 30MHz. As 64MHz (1.5T) machines have since become standard for clinical imaging, these predictions were overly pessimistic. Some authors [44, 46] account for this discrepancy by suggesting that insulating structures in biological tissue prevent large current loops from forming. However, calculations by Glover et al. [19] show substantially greater field penetration[1] which have been confirmed by experiments with a cylindrical phantom at 64MHz. Zypman [62] has performed a similar calculation taking into account the influence of the RF coil. Analysis of this kind predicts intensity variations of 10% to 20% in a head size structure, which is consistent with what is typically observed. The magnitude of these variations suggest that electrodynamic interactions, rather than other effects such as deficiencies in the scanner, are the primary cause of intensity non-uniformity for 1.5 T machines.

A difficulty with analytic modelling is that closed form solutions are only available for simple geometries. Finite element analysis offers a more flexible approach to dynamic modelling of complicated structures. The problem of modelling the human head at 64MHz and 175MHz has been considered by Šimunić et al. [47] and Vaughan et al [55] respectively. Šimunić has validated a model of the fields generated by a linearly polarized saddle coil *in vivo* using the dual flip-angle technique described in [50]. These results predict variations on the order of 30%. Similar measurements [50] using a bird cage head coil to scan a human head showed less than 10% variation in

[1]A sign error in equation (7) on p. 632 of [4] accounts for this discrepancy. This equation is presented in its correct form in Glover et al. [19].

B_1 field strength. Provided a model of the subject is available, finite element analysis is an accurate method of predicting the field variations in a scan.

2.3 Characterization of intensity non-uniformity

While simulation of intensity non-uniformity is not always a practical method of correcting for it, simulation is a useful tool for investigating its character under a variety of conditions.

2.3.1 Simulated spin echo images

The analytic modelling method described in [19, 48] can be used to derive expressions for the excitation field and reception sensitivity present when scanning a uniform cylinder with elliptic cross section. Using dimensions similar to those of a human head, these expressions are useful to illustrate the form of non-uniformity one can expect to observe. Assuming uniform circularly polarized fields from the unloaded coil and neglecting magnetic relaxation effects yields simulated spin echo images for circular and elliptic geometries as shown in Figures 2.1a and 2.1b. Also shown are profiles through the principal axes for the two cases. Note that the elliptic cylinder shows a diagonal pattern of non-uniformity, while the pattern for the circular cylinder is radially symmetric. The magnitude of the non-uniformity is consistent with the levels of non-uniformity seen in typical brain scans, 15% to 20% variation. Although these simulations only account for non-uniformity caused by electrodynamic interaction with the subject, the evidence suggests that this is the primary cause of intensity non-uniformity. For comparison, three T1 weighted scans of the head are shown in Figure 2.2. All three show a diagonal non-uniformity pattern in the white matter.

2.3.2 Pulse sequence sensitivity

The choice of imaging sequence has an impact on the magnitude of the non-uniformity. The signal measured at location $\mathbf{x}$ for a spin echo pulse sequence, neglecting relaxation

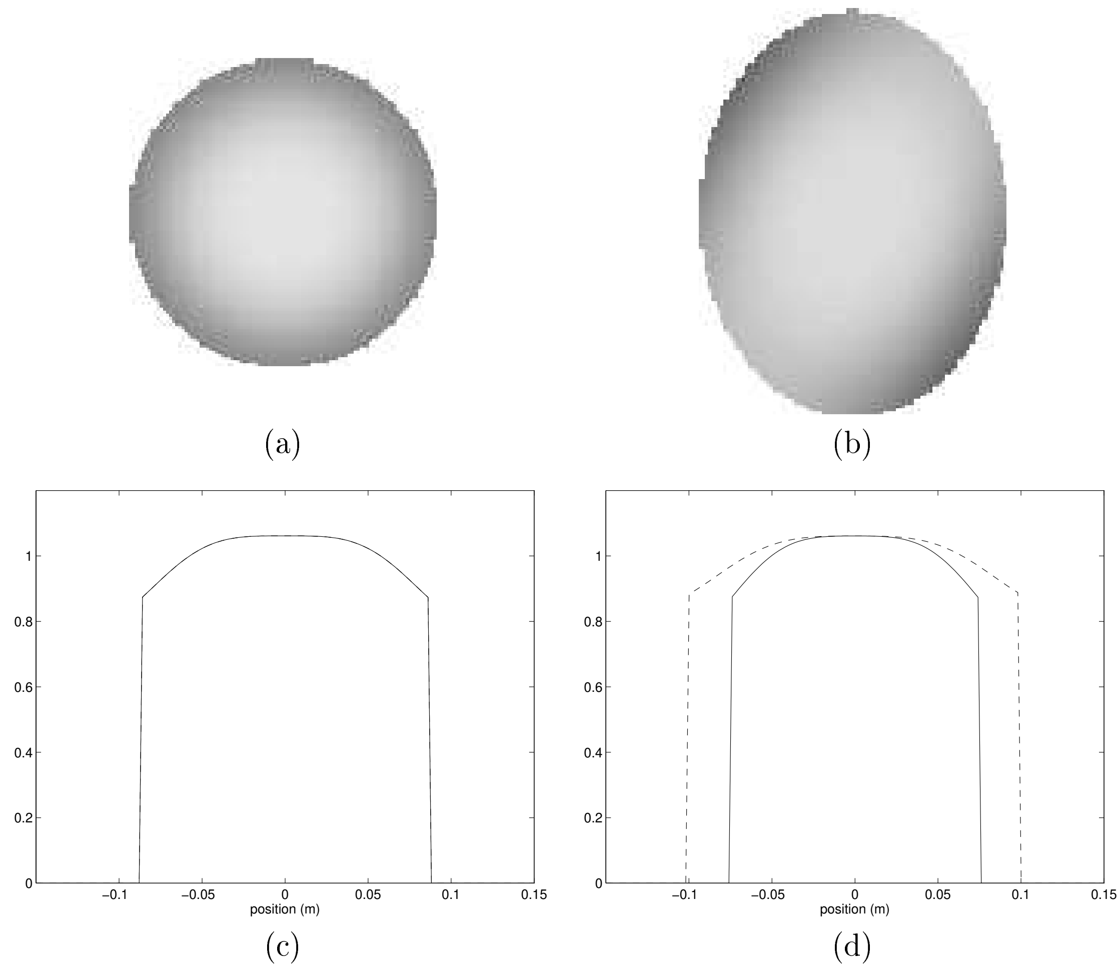


Figure 2.1: Simulated spin echo images through a cylinder with circular or elliptic cross section, using nominal 90^o and 180^o flip-angles. The main magnetic field strength (B_0) is 1.5 T. The cylinders have a relative permittivity $\epsilon_r = 80$ times that of free space and a resistivity $\rho = 2\,\Omega$m. (a) spin echo image of circular section with diameter 15 cm. (b) spin echo image of elliptic section. The major and minor diameters are 20 cm and 15 cm respectively. (c) profile along diameter for circular section. (d) profile along principal axes for elliptic section.

Figure 2.2: T1 weighted brain scans showing diagonal non-uniformity, brighter at the top left and bottom right. The gray scale has been compressed to exaggerate the intensity variations.

effects, is given by the expression [19]

$$\begin{aligned} S(\mathbf{x}) &= R(\mathbf{x})S_{SE}(\mathbf{x}) && (2.1) \\ S_{SE}(\mathbf{x}) &= S_0 \sin^3 \theta(\mathbf{x}) && (2.2) \\ \theta(\mathbf{x}) &= (\pi/2)\,|\mathbf{B}(\mathbf{x})|\,/B_m && (2.3) \\ R(\mathbf{x}) &= R_0 \mathbf{B}^*(\mathbf{x}) && (2.4) \end{aligned}$$

where S_{SE} is the emitted signal, S is the measured signal, θ is the flip angle, $\mathbf{B}$ is the excitation field, B_m is the field strength needed to produce a 90^o flip-angle, and R is the reception sensitivity. Consider a first order perturbation on the measured signal $S = S_m + \delta S$ caused by a perturbation $B = B_m + \delta B$ to the excitation field strength and a corresponding perturbation to the reception sensitivity. Substituting these expressions into equation (2.1) and treating the signal as real leads to an expression for the sensitivity of a spin echo imaging sequence

$$\begin{aligned} \frac{S_m + \delta S}{S_m} &= \frac{B_m + \delta B}{B_m} \sin^3 \left(\frac{\pi}{2} \frac{B_m + \delta B}{B_m} \right) \\ 1 + \frac{\delta S}{S_m} &= \left(1 + \frac{\delta B}{B_m} \right) \cos^3 \left(\frac{\pi}{2} \frac{\delta B}{B_m} \right) \end{aligned}$$

$$1 + \frac{\delta S}{S_m} = \underbrace{\left(1 + \frac{\delta B}{B_m}\right)}_{\text{reception}} \underbrace{\left(1 - \frac{3\pi^2}{8}\left(\frac{\delta B}{B_m}\right)^2 + o\left(\frac{\delta B}{B_m}\right)^3\right)}_{\text{excitation}} \tag{2.5}$$

$$\frac{\delta S}{S_m} = \frac{\delta B}{B_m} - \frac{3\pi^2}{8}\left(\frac{\delta B}{B_m}\right)^2 + o\left(\frac{\delta B}{B_m}\right)^3 \tag{2.6}$$

where the notation $o(x)^n$ denotes terms in x of order n and greater. This expression represents the fractional change in measured signal caused by a given fractional change in field magnitude and reception sensitivity.

It should be noted from equation (2.5) that, while the dependence on reception sensitivity is first order, the dependence on excitation uniformity is second order. Hence the sequence is less sensitive to variations in the excitation field than variations in sensitivity, which supports the argument that intensity variations can be modelled as a multiplicative field.

The comparable expression for a gradient echo sequence which only involves a single 90^o RF excitation is

$$\frac{\delta S}{S_m} = \frac{\delta B}{B_m} - \frac{\pi^2}{8}\left(\frac{\delta B}{B_m}\right)^2 + o\left(\frac{\delta B}{B_m}\right)^3 \tag{2.7}$$

While the sensitivities of 90^o gradient echo and 90^o–180^o spin echo sequences differ only in the second order terms, it should be noted that gradient echo sequences are commonly used with smaller flip-angles. For example, the expression for the sensitivity of a gradient echo sequence with a 45^o degree flip-angle is given by

$$\frac{\delta S}{S_m} = \left(1 + \frac{\pi}{4}\right)\frac{\delta B}{B_m} + \frac{\pi}{4}\left(1 - \frac{\pi}{8}\right)\left(\frac{\delta B}{B_m}\right)^2 + o\left(\frac{\delta B}{B_m}\right)^3 \tag{2.8}$$

which means that this sequence is 78% more sensitive to field variations than the either the spin echo or 90^o gradient echo sequences.

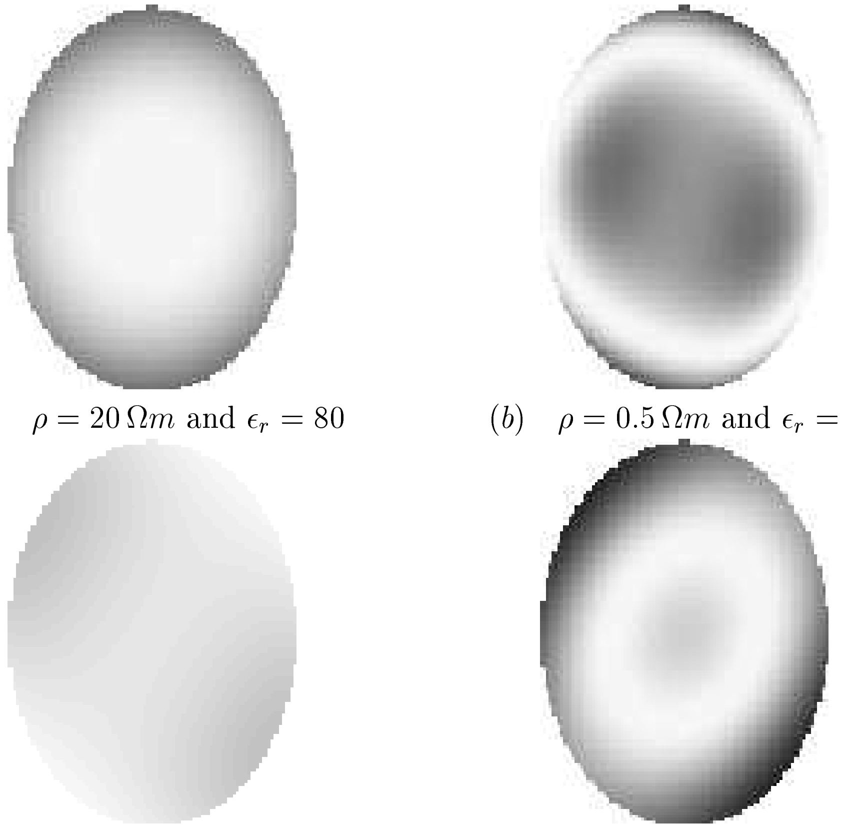
2.3.3 Influence of media properties

In continuing to investigate the character of intensity non-uniformity, consider the effect that the properties of the media have on the shape of the non-uniformity observed. Simulated images for a number of combinations of relative permittivity ϵ_r

and resistivity ρ are shown in Figure 2.3. These values span the typical ranges for biological tissue. Increasing the resistivity of the media from its nominal value of $2\,\Omega$m to $20\,\Omega$m, as shown in Figure 2.3a, decreases the image uniformity by enhancing the signal in the center of the cylinder. This cup shaped intensity variation is referred to as a standing wave effect. Decreasing the resistivity to $0.5\,\Omega$m, as shown in 2.3b, reverses the cup shaped variation and produces the strongest intensity in a rim at the edge. The reduced intensity at the center is often described as an RF penetration effect, caused by induced currents. Reducing the relative permittivity to 20, as in 2.3c, produces the most uniform image. This is consistent with making the properties of the media more like the surrounding space. Increasing the relative permittivity to 160, a value larger than typically found in biological tissue, produces the wave like non-uniformity pattern shown in 2.3d. This behaviour is consistent with the reduced wavelength being smaller than the dimensions of the cylinder. To some extent, RF penetration and standing wave effects are complementary in that one enhances the intensity near the center while the other diminishes it. The properties of biological tissue are such that there is cancelation between the two effects, perhaps accounting for the better than predicted uniformity of 1.5T MR images.

2.4 Practical implications of non-uniformity

While a number of effects contribute to intensity non-uniformity, spatial variation in the reception sensitivity is the most important. In clinical scanners a significant portion of this variation is due to electromagnetic interaction with the subject. Consequently, the artifact is difficult to manage since each scan potentially has a different pattern of variation. In addition, the form of the variation has a complex dependence on pulse sequence, so one cannot assume that the pattern of non-uniformity will be the same for different echos of a multi-echo scan. Given this qualitative understanding of intensity non-uniformity, one is in a position to evaluate proposals for correcting for it.



(a) $\rho = 20\,\Omega m$ and $\epsilon_r = 80$ (b) $\rho = 0.5\,\Omega m$ and $\epsilon_r = 80$

(c) $\rho = 2\,\Omega m$ and $\epsilon_r = 20$ (d) $\rho = 2\,\Omega m$ and $\epsilon_r = 160$

Figure 2.3: Simulated spin echo images of a cylinder with elliptic section for various values of ϵ_r and ρ. Compare with Figure 2.1 where $\epsilon_r = 80$ and $\rho = 2\,\Omega m$, i.e. normal values.

Chapter 3

Review of intensity non-uniformity correction methods

Existing methods for non-uniformity correction can be grouped into three categories: analytic modelling of the non-uniformity [36, 38]; modified acquisition protocols that measure non-uniformity [1, 10, 18, 39, 41, 50, 58]; and data-driven postprocessing [5, 11, 12, 20, 23, 27, 31, 32, 37, 40, 49, 54, 57]. With the exception of [5, 12], all of these methods assume a multiplicative non-uniformity field that is corrected for by dividing it from the acquired image.

Analytic methods as described in the previous chapter are a useful tool for understanding the mechanisms that produce the intensity variations. However, the dependence of the non-uniformity on the subject geometry makes these approaches impractical since a new model is needed for each scan. While a patient independent analytic model may be satisfactory when correcting for the severe non-uniformity caused by surface coils, a patient-independent approach is not sufficiently accurate to correct for the diverse and relatively mild intensity variation seen with standard volume coils.

Among the techniques that involve **modified acquisition protocols**, the fact that the non-uniformity is strongly subject-dependent at 1.5 T is grounds to dismiss those that involve regular scanning of a calibration phantom [1, 10, 17, 58]. Besides, frequent scanning of calibration phantoms is not practical in a clinical setting.

A modified acquisition protocol that can be used *in vivo* is to image with a second reception coil [41] that is more uniform. For example, one might image with both a head coil and a body coil. Differences between the two images that are smoothly varying can be attributed to intensity non-uniformity. However, since the intensity variations caused by electrodynamic interaction with the subject will appear in both images, this method can only detect the variations caused by deficiencies in the head coil.

An alternative approach described by Stollberger and Wach [50] measures the RF excitation field *in vivo* based on the change in non-uniformity with excitation flip angle. While this approach does not yield a map of reception coil sensitivity, one could assume it is related to the excitation field, then use the two maps to correct for intensity non-uniformity using the Bloch equations [3]. However, modified acquisition protocol methods have not been widely used because the extended scan time is impractical in a clinical setting and precludes retrospective analysis of previously acquired data.

The most popular of the **data-driven postprocessing methods** is to use homomorphic filtering to estimate a multiplicative non-uniformity field and divide it from the image. A number of variations on this method have been presented that differ for the most part in their handling of background regions [23, 27, 31, 32, 46]. These approaches assume that the frequency content of the non-uniformity field is lower than that of the anatomy. While this approximation may be satisfactory for surface coil images, the non-uniformity seen in volumetric acquisitions is too subtle for this approximation to apply since one cannot neglect the low frequency intensity variations caused by anatomy. In particular, the contrast seen in T1 weighted scans between regions of grey and white matter is such that when taken together the whole volume appears to have low frequency spatial variations corresponding to the distribution of these regions.

A number of authors have presented schemes to eliminate the low frequency components of the anatomy when estimating non-uniformity. These methods can be distinguished by their modelling assumptions as: (i) tissue intensity model based; (ii) spatial homogeneity based; and (iii) field model based. **Tissue intensity model** based methods such as those described in [12, 40] segment and remove the bright cerebrospinal fluid from T2 and proton density (PD) weighted scans and then apply a homomorphic filter and divide approach to the remaining, approximately isointense, brain tissue. These approaches cannot be applied to T1 weighted scans due to the contrast between grey and white matter intensity.

An alternative tissue model approach described in [11] is to segment the white matter, estimate the non-uniformity in it, and extrapolate that estimate to the rest of the brain. A difficulty with this approach is that there are regions of the brain without white matter in which the quality of the extrapolated non-uniformity estimate may suffer. Tissue intensity methods which address this problem by taking into account the non-uniformity of several classes of tissue have been proposed in [5, 20, 57]. These methods also achieve better accuracy through some form of iteration.

A common difficulty with tissue intensity model based methods is the determination of the model. In general, a different model is needed for each pulse sequence and perhaps for different scanners. While automated methods of developing such a model can be devised, the need for such a model limits the application of these methods.

An alternative model assumption is that of **spatial homogeneity**. The method described in [37] uses LCJ [33] segmentation to identify contiguous regions irrespective of their tissue type. A smooth field is globally fit to the log intensities of all regions allowing for a constant offset in each regions to reflect its tissue type. Given a robust segmentation, this method can be applied to images of any pulse sequence. However, since this method requires contiguous regions of homogeneous tissue, the correction may suffer in areas of the brain such as the cerebellum where such regions are not present. A second difficulty arises if contiguous regions are mistakenly broken into smaller regions creating additional undesired degrees of freedom: the smaller the regions, the less information they contribute to the field estimate. In the limiting case an isolated voxel contributes no information. Since this technique is heavily dependent on a successful segmentation, one can expect it to perform best on images with large scale structure.

A third type of modelling assumption is that of a **field model**: the assumption that non-uniformity blurs the histogram of the data in a way that it can be identified, quantified, and removed. This is the basis of the N3 method, described in the following chapter.

Chapter 4

The N3 method

Non-parametric intensity non-uniformity normalization (N3) is a new data-driven method for correcting for intensity non-uniformity in MRI data. It is non-parametric in the sense that it requires neither a parametric model of tissue intensities nor the decomposition of the intensity distribution into a limited number of tissue classes. While making no assumptions about the kind of anatomy present in a scan, the method is well suited to removing the relatively subtle intensity variations seen in standard clinical scans of the head.

4.1 Non-uniformity model

The problem of correcting for intensity non-uniformity is greatly simplified if it is modelled as a smooth multiplicative field. As described in Chapter 2, this model is consistent with the multiplicative non-uniformity arising from variations in the sensitivity of the reception coil and to a lesser extent the non-uniformity due to induced currents and non-uniform excitation.

Consider the following model of image formation in MR

$$v(\mathbf{x}) = u(\mathbf{x})f(\mathbf{x}) + n(\mathbf{x}) \tag{4.1}$$

where at location $\mathbf{x}$, v is the measured signal, u is the true signal emitted by the tissue, f is an unknown smoothly varying bias field, and n is white Gaussian noise assumed to be independent of u. The problem of compensating for intensity non-uniformity is the task of estimating f. The combination of additive and multiplicative interference makes this task difficult.

Consider a noise free case in which the true intensities u at each voxel location $\mathbf{x}$ are independent identically-distributed random variables. Using the notation $\hat{u}(\mathbf{x}) = \log(u(\mathbf{x}))$ the image formation model becomes additive:

$$\hat{v}(\mathbf{x}) = \hat{u}(\mathbf{x}) + \hat{f}(\mathbf{x}) \tag{4.2}$$

Consider the distribution of values that $\hat{f}$ takes over the region of interest to be the probability distribution of a random variable. For example, if $\hat{f}$ is a linearly increasing field aligned on a rectangular region then $\hat{f}$ will have a uniform distribution.

Let U, V, and F be the probability densities of $\hat{u}$, $\hat{v}$, and $\hat{f}$ respectively. Making the approximation that $\hat{u}$ and $\hat{f}$ are independent or uncorrelated random variables, the distribution of their sum is found by convolution as follows (for details see Appendix B.)

$$V(\hat{v}) = F(\hat{v}) * U(\hat{v}) = \int F(\hat{v} - \hat{u})U(\hat{u})d\hat{u} \tag{4.3}$$

The non-uniformity distribution F can be viewed as blurring the intensity distribution U.

4.2 Correction strategy

From a signal processing perspective, the blurring due to the field reduces the high frequency components of U. The task of correcting for intensity non-uniformity is that of restoring the frequency content of U. Since the shape of the blurring kernel F is not known, it is not clear what frequency components of U need to be restored to get from the observed distribution V to the true distribution U. However, since the non-uniformity field $\hat{f}$ is restricted to be smooth and slowly varying, there are relatively few possible distributions U corresponding to a given distribution V. **In N3, the approach to correcting for non-uniformity is to find the smooth,**

slowly varying, multiplicative field that maximizes the frequency content of U.

As evidence of the simple form of the distribution F, consider the distributions shown in Figure 4.1. These have been derived from putative fields obtained by fitting the intensities of manually identified points of nominally isointense white matter on twelve individuals. Each individual was scanned using the same pulse sequences but on a different MR machine (data collected as part of a multi-center clinical trial). Included in the twelve are machines made by Philips, Siemens, and GE. As shown, the large scale features of F vary little between scans. In particular, F is well approximated by a unimodal distribution. These results suggest that the full width at half maximum (FWHM) of the distribution F lies between 0.1 and 0.4 for typical brain scans.

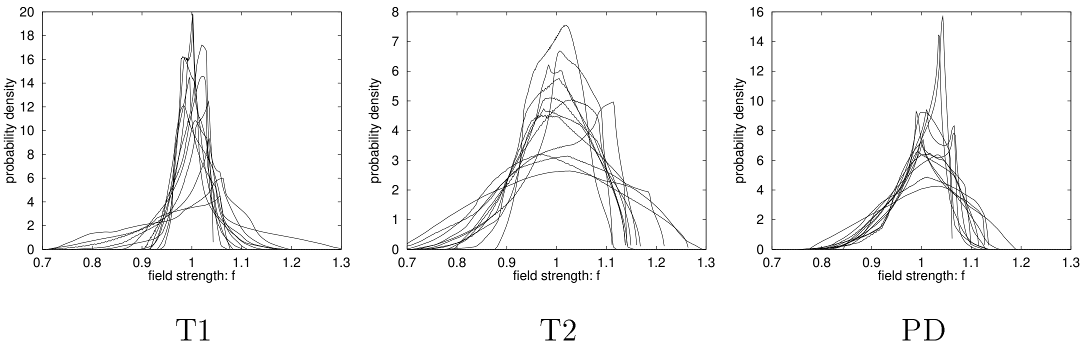


Figure 4.1: Probability densities for non-uniformity fields estimated from white matter on MR scans of twelve individuals taken with twelve different MR scanners. Note that these are the distributions of the dimensionless scale factors f rather than the log of these values $\hat{f}$. The T1 scan is a gradient echo 3D acquisition with TE $= 11$ ms and TR $= 35$ ms, while the proton density (PD) and T2 scans are two echoes TE $= 30$ ms and TE $= 80$ ms of a multi-slice spin echo acquisition with TR $= 3$ s.

Returning to the optimization criterion, one could in principle search through all possible fields $\hat{f}$ to find the one that maximizes the high frequency content of U. However, there are two problems with this approach: the search space of all 3D fields $\hat{f}$ is extremely large; and spectral estimates and related measures such as entropy are difficult to compute with sufficient accuracy to detect subtle changes in U.

The approach adopted is to propose a distribution for U by sharpening the distribution V, and then to estimate the corresponding smooth field $\hat{f}$ which produces a distribution U close to the one proposed. While searching through the space of all distributions U may seem no more tractable than searching through the space of all fields $\hat{f}$, there is an important difference in that one can take advantage of the simple form of the distribution F. Suppose that the distribution of F is Gaussian. Then one need only search the space of all distributions U corresponding to Gaussian distributed F having zero mean and given variance. In this way the space of all distributions U is collapsed down to a single dimension, the width of the F distribution.

In practice, F is only approximately Gaussian and some of our assumptions, such as zero noise, are violated. These difficulties are contended with by taking an iterative approach to estimating the distribution U and corresponding field $\hat{f}$. Since any Gaussian distribution can be decomposed into a convolution of narrower Gaussian distributions, the space of all U distributions corresponding to Gaussian distributed F can be searched incrementally by deconvolving narrow Gaussian distributions from subsequent estimates of U. The benefit of this approach is that between subsequent estimates of U, a corresponding smooth field $\hat{f}$ is estimated. The constraint that the field be smooth changes the shape of the proposed distribution U to one that is consistent with the field. These perturbations of U perturb F from its Gaussian shape and compensate for the distortion of V caused by noise and other factors. The iterative process can be viewed as traveling in the space of all U distributions along a path corresponding to smooth fields $\hat{f}$ with increasingly wider distributions. These iterations proceed until no further changes in $\hat{f}$ or U result from deconvolving narrow Gaussian distributions from V.

4.3 Field estimation

Further theory is presented here to explain the process of proposing distributions for U and estimating corresponding fields. For notational simplicity, assume that the true distribution of intensities U can be arrived at in a single iteration by deconvolving a distribution F, which is Gaussian, from V. The full iterative description of the method is left for a subsequent Section.

Given the distribution U, the method of estimating the corresponding field is as follows. For a measurement $\hat{v}$ at some location $\mathbf{x}$, $\hat{u}$ is estimated using the distributions U and F. Since the choice of the location $\mathbf{x}$ is arbitrary, the measurement $\hat{v}$ can be treated as a random sample from the distribution V. The expected value of $\hat{u}$ given a measurement $\hat{v}$ is as follows

$$E[\hat{u} \mid \hat{v}] = \int_{-\infty}^{\infty} \hat{u}\, p(\hat{u} \mid \hat{v}) d\hat{u} \tag{4.4}$$

$$= \int_{-\infty}^{\infty} \hat{u}\, \frac{p(\hat{u}, \hat{v})}{p(\hat{v})} d\hat{u} \tag{4.5}$$

Writing $p(\hat{v})$ as $V(\hat{v})$ and using equation (4.3) gives

$$E[\hat{u} \mid \hat{v}] = \frac{1}{V(\hat{v})} \int_{-\infty}^{\infty} \hat{u} p(\hat{u}, \hat{v}) d\hat{u} \tag{4.6}$$

$$= \frac{\int_{-\infty}^{\infty} \hat{u} F(\hat{v} - \hat{u}) U(\hat{u}) d\hat{u}}{\int_{-\infty}^{\infty} F(\hat{v} - \hat{u}) U(\hat{u}) d\hat{u}} \tag{4.7}$$

An estimate of $\hat{f}$ can be obtained using the estimate of $\hat{u}$ from equation (4.7) as follows

$$\hat{f}_e(\hat{v}) = E[\hat{f} \mid \hat{v}] = \hat{v} - E[\hat{u} \mid \hat{v}] \tag{4.8}$$

where $\hat{f}_e$ is an estimate of $\hat{f}$ at location $\mathbf{x}$ based on the single measurement of $\hat{v}$ at $\mathbf{x}$. This estimate can be smoothed by the operator S to produce

$$\hat{f}_s(\hat{v}) = S\{\hat{f}_e(\hat{v})\} = S\{\hat{v} - E[\hat{u} \mid \hat{v}]\} \tag{4.9}$$

an estimate of $\hat{f}$ based on all of the measurements in a neighbourhood of $\mathbf{x}$. Smoothing is described in Section 4.5.

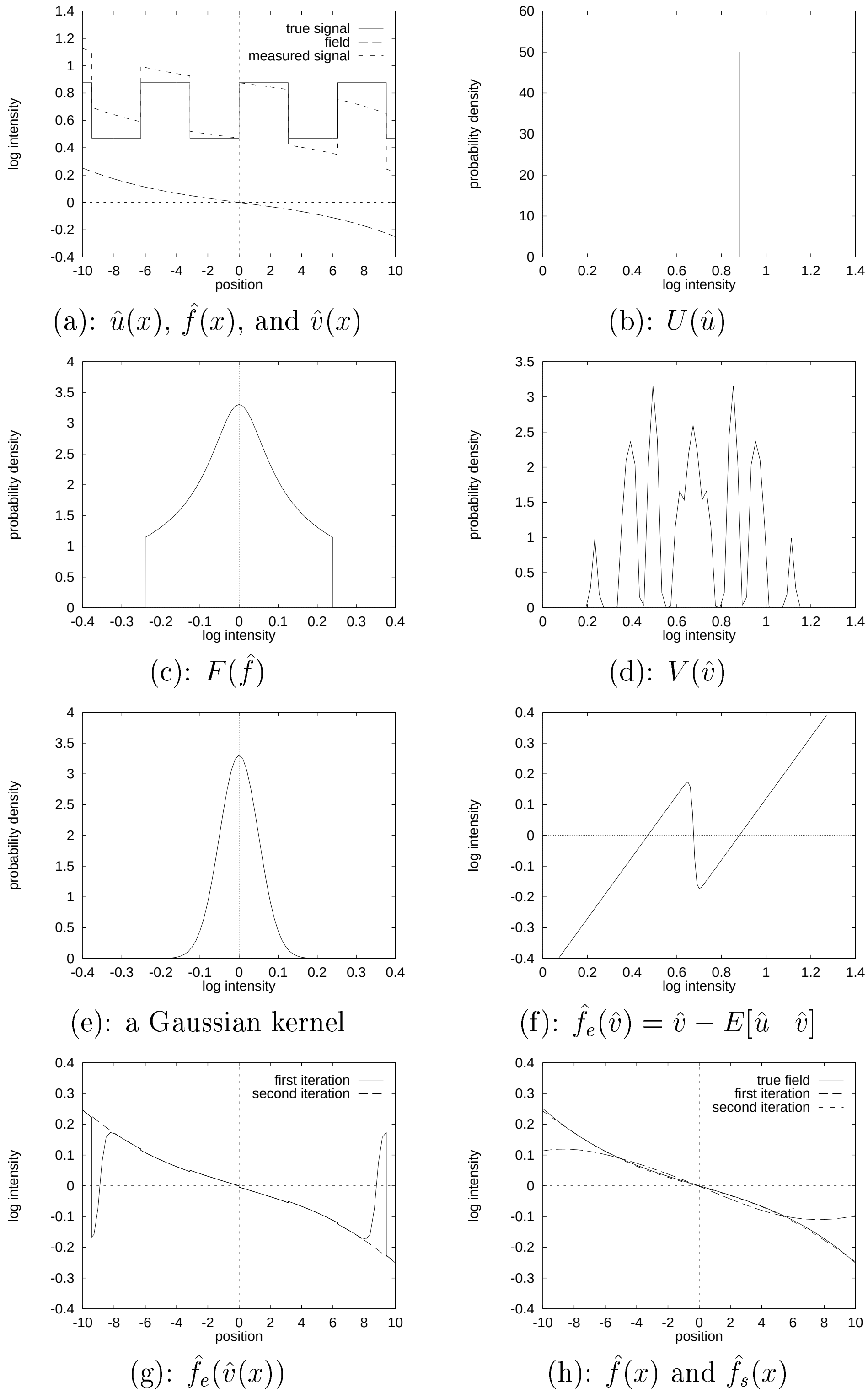
This process of estimating $\hat{f}$ is illustrated in Figure 4.2 for the one dimensional case with $\hat{u}$ as a square wave. The square wave can be viewed as two tissue classes corresponding to the low and high parts of the cycle. The log intensities $\hat{u}$, $\hat{f}$, and

$\hat{v}$ are shown in Figure 4.2a. To make the example non-trivial, the magnitude of the non-uniformity field is large enough that the low intensity tissue on the left overlaps by roughly 10% with the intensity of the high intensity tissue on the right.

One obtains the probability densities U, V, and F by computing a histogram of the signals $\hat{u}$, $\hat{f}$, and $\hat{v}$. These distribution are shown in Figures 4.2b, 4.2c, and 4.2d respectively. These histograms are interpreted as the probability distributions of the given signal. In practice, only V will be known while U and F have to be estimated. Shown in Figure 4.2e is the Gaussian kernel used in place of the actual distribution of F in equation (4.7).

The distributions U and the Gaussian kernel F are used to compute from equation (4.8) the mapping, shown in Figure 4.2f, which maps measured intensities $\hat{v}$ to field estimates $\hat{f}_e$. Estimating the distribution U from V is described in the following section. This mapping when applied to the measured signal produced the field estimate shown in 4.2g. Note the sharp jumps in this estimate in the regions where the intensities of the upper and lower tissue classes overlap. This field estimate is smoothed to produce $\hat{f}_s$ shown in 4.2h. While the degree of smoothing is arbitrary, in filtering terms the smoothing filter should be chosen to have as small a bandwidth as possible yet still pass the non-uniformity field undistorted. In particular, it is not necessary to have a smoothing filter able to completely remove the sharp jumps in $\hat{f}_e$ shown in Figure 4.2g since the residual discontinuities disappear in subsequent iterations.

While the smooth field estimate after one iteration is not sufficient to completely remove the non-uniformity, it is enough of a correction that the tissue distributions no longer overlap. By correcting the measured signal $\hat{v}$ using the the smooth estimate of the field $\hat{f}_s$ and repeating the entire estimation process again, one arrives at a second estimate of the non-uniformity field that is indistinguishable from the true field. The second estimates of $\hat{f}_e$ and $\hat{f}_s$ are shown in Figures 4.2g and 4.2h.



(a): $\hat{u}(x)$, $\hat{f}(x)$, and $\hat{v}(x)$

(b): $U(\hat{u})$

(c): $F(\hat{f})$

(d): $V(\hat{v})$

(e): a Gaussian kernel

(f): $\hat{f}_e(\hat{v}) = \hat{v} - E[\hat{u} \mid \hat{v}]$

(g): $\hat{f}_e(\hat{v}(x))$

(h): $\hat{f}(x)$ and $\hat{f}_s(x)$

Figure 4.2: An example of non-parametric correction of a one dimensional signal. See text for detailed explanation.

4.3.1 Estimating the distribution U

Given a distribution F and the measured distribution of intensities V, the distribution U can be estimated using a deconvolution filter as follows

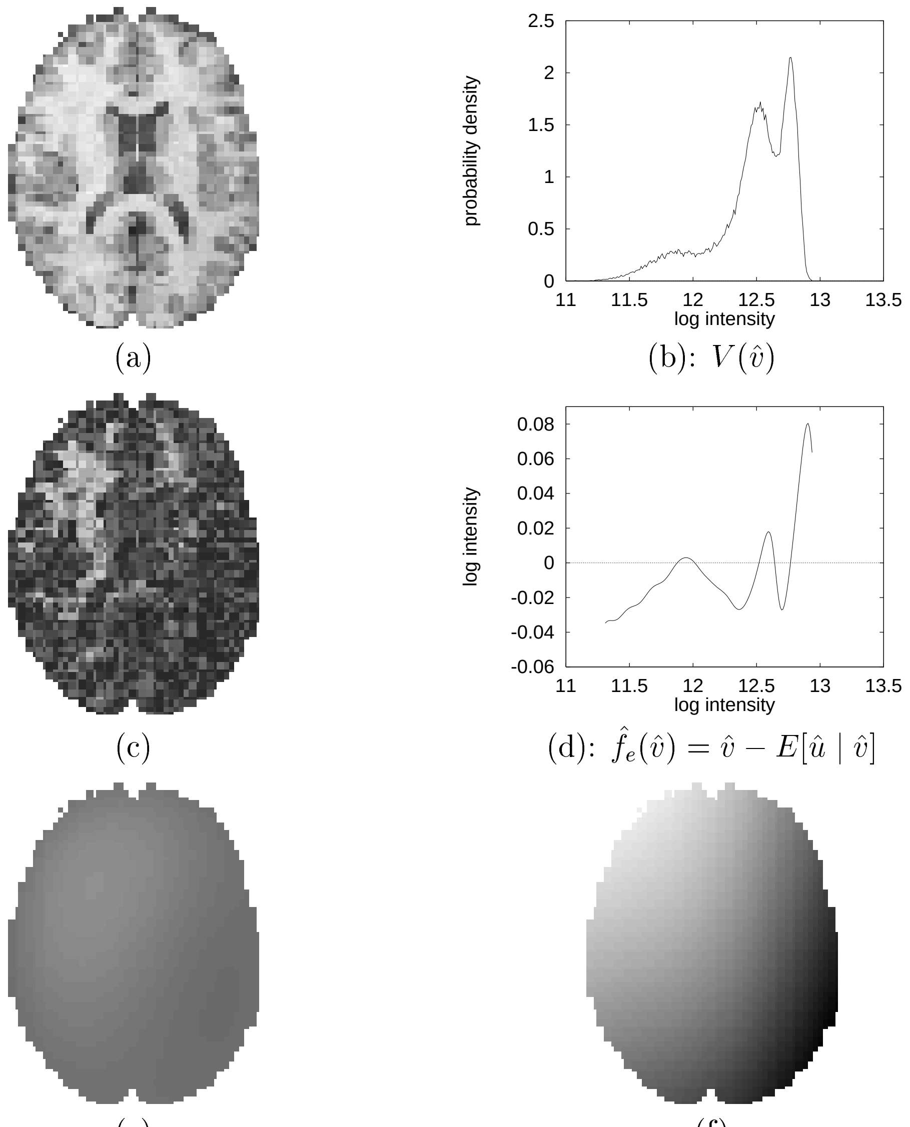
$$\widetilde{G} = \frac{\widetilde{F}^*}{|\widetilde{F}|^2 + Z^2} \tag{4.10}$$

$$\widetilde{U} \approx \widetilde{G}\,\widetilde{V} \tag{4.11}$$

where * denotes complex conjugate, $\widetilde{F}$ is the Fourier transform of F, and Z is a constant term to limit the magnitude of $\widetilde{G}$. This estimate of U is then used to estimate a corresponding field $\hat{f}$.

4.3.2 Field estimation in 3D

As an illustration of the field estimation process in 3D, consider the simulation of a T1 weighted MR scan shown in Figure 4.3a. This simulation is based on a three tissue model of the brain and incorporates intensity non-uniformity, noise, and partial volume effects. The measured intensity distribution V is shown in 4.3b. The mapping between measured intensity and field estimate for this volume based on the distributions of V and a Gaussian distribution F with full width at half maximum (FWHM) of 0.15 is shown in Figure 4.3d. This mapping is applied to produce the volume in 4.3c. Measurement noise causes this estimate to be noisy. Smoothing this estimate produces the field shown in 4.3e. For comparison, the actual field imposed on the data during simulation is shown in 4.3f. Although the estimated field is much smaller in magnitude than the true field, the shape is similar. Experiments, described later, show that with subsequent iterations the field estimates will grow to narrow this difference. A progression of field estimates with iterations is shown for a real MR scan in Figure 4.4. It should be noted that while the initial local field estimates have brain-like structure, this bias diminishes, particularly in the ventricles, with iterations.



(a) (b): $V(\hat{v})$

(c) (d): $\hat{f}_e(\hat{v}) = \hat{v} - E[\hat{u} \mid \hat{v}]$

(e) (f)

Figure 4.3: Field estimates for a simulated MR volume. (a) A slice from a simulated MR volume. (b) The histogram of the volume shown in (a). This is considered an estimate of the distribution V. (c) Bias field estimates f_e created by applying the mapping in (d) to the image in (a). (d) The mapping between image intensity $\hat{v}$ and bias field estimate f_e, derived from the histogram in (b). (e) A smoothed estimate f_s of the bias field created by smoothing the volume shown in (c). (f) The actual bias field present in the volume shown in (a).

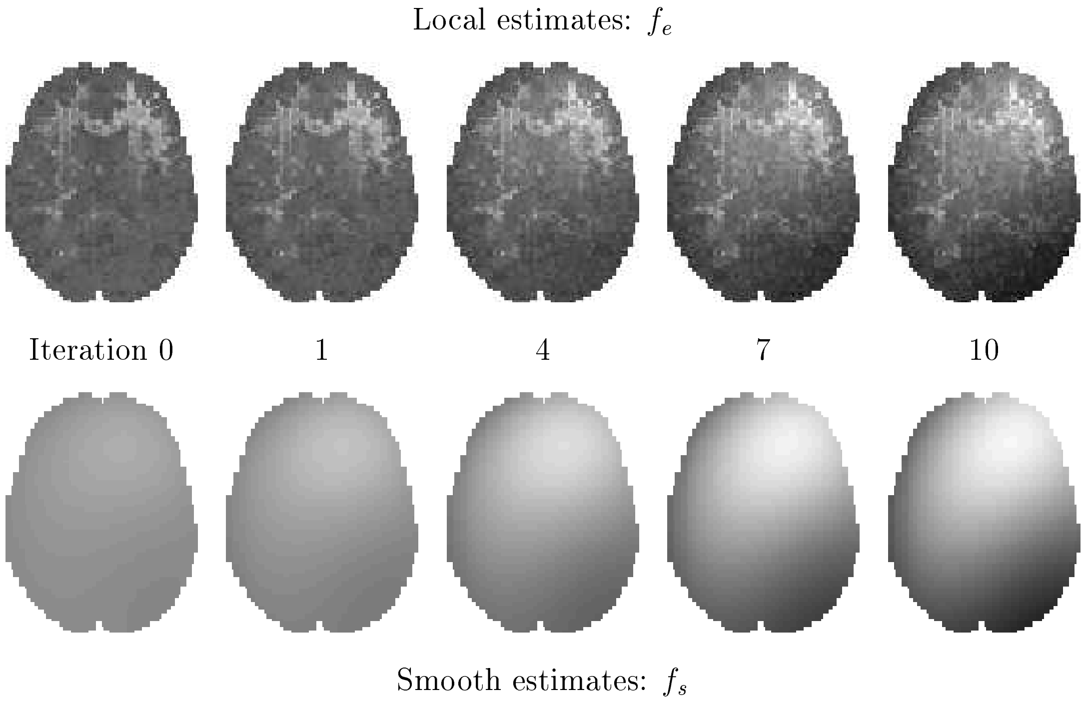


Figure 4.4: Evolution of field estimates with iterations using real data. Local and smooth field estimates are shown during correction of a T1 weighted MR scan.

4.4 Implementation details

A flow chart describing the N3 correction process is shown in Figure 4.5. Besides the processing steps described previously, there are a number of steps needed for practical implementation of the algorithm.

The first step shown in Figure 4.5, "identify foreground," is to segment and remove empty regions from the volume. Besides the numerical problems associated with transforming values near zero to the log domain, these background regions provide no information about the non-uniformity field. The foreground can be determined using a simple threshold chosen automatically by analyzing the histogram of the volume [42].

Another consideration in implementing the N3 approach is measuring the distribution V from the unprocessed MR data. For simplicity, a histogram with equal size bins and a triangular Parzen window [14] is used to estimate V. Given a set of N

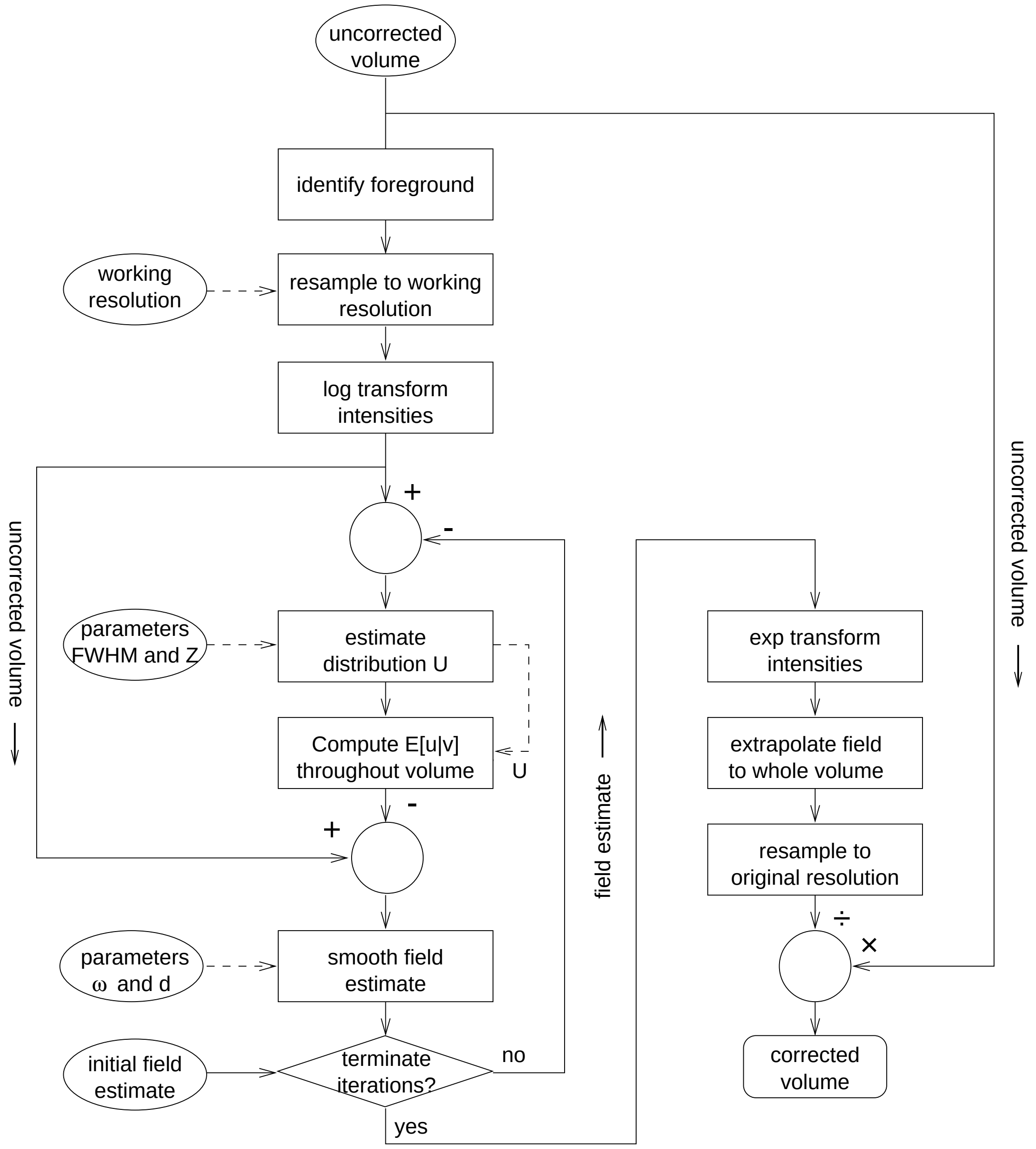


Figure 4.5: A flow chart describing the N3 method. Ellipses represent user-selected parameters and priors. Among these are FWHM and Z which specify the deconvolution kernel, and ω and d which control the smoothness of the field estimate. Rectangles are processing steps. The flow of volumetric data is represented by solid lines and other data by dashed lines. Circles perform arithmetic operations on a voxel by voxel basis. The result of the process is a corrected volume.

measurements $\hat{v}(\mathbf{x}_i)$ and locations $\mathbf{x}_i$, V is estimated as follows

$$V(\hat{v}_j) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{h}\varphi\left(\frac{\hat{v}_j - \hat{v}(\mathbf{x}_i)}{h}\right) \quad (4.12)$$

$$\varphi(s) = \begin{cases} 1-|s| & |s|<1 \\ 0 & \text{elsewhere} \end{cases} \quad (4.13)$$

where $\hat{v}_j$ are the centers of the bins and h is the distance between them. For a typical 20% bias field, the scale factor f ranges from 0.9 to 1.1 which corresponds to $\hat{f}$ between -0.1 and 0.1. MR volumes generally have sufficient data to estimate V at a resolution h better than a tenth of this range.

Smoothing the non-uniformity field at full resolution is computationally expensive, so the MR data is subsampled to a lower resolution. Since the non-uniformity field is slowly varying, reducing a 1 mm isotropically sampled volume to 3 mm has a negligible effect on the field estimate and substantially accelerates computation (see Appendix C). The processing of a volume with ten iterations of the N3 method is reduced from 12 hours to twenty minutes of CPU time on an SGI workstation (floating point performance: 99 SPECfp92) by resampling to the coarser resolution. The final field estimate is resampled to the original resolution and used to correct the original volume.

The measure used to terminate the iterations is the coefficient of variation in the ratio between subsequent field estimates, computed as follows

$$e = \frac{\sigma\{r_n\}}{\mu\{r_n\}}, \quad n = 1 \ldots N \quad (4.14)$$

where r_n is the ratio between subsequent field estimates at the nth location, σ denotes standard deviation, and μ denotes mean. This measure is chosen so as to be insensitive to global scale factors that may accumulate with iterations. Iteration is stopped when e drops below 0.001, typically after ten iterations.

4.5 Smoothing

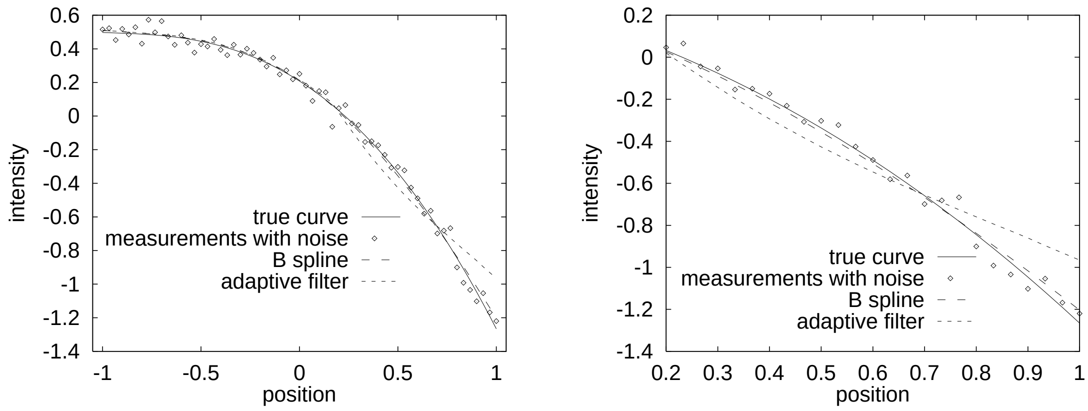
The manner in which the field estimate is smoothed has a significant impact on the performance of the correction method. Smoothing is particularly challenging for this problem because the scale over which the field varies is comparable to the size of the region being smoothed. Conventional filtering techniques proved unsatisfactory for this application since boundary effects significantly degraded overall performance.

A computationally tractable approach to smoothing that performs well on bounded domains is to approximate the data by a linear combination of smooth basis functions. An attractive basis for this is the compactly-supported spline known as a B spline. Spline approximation incorporating smoothness constraints is superior to filtering techniques in dealing with missing data since the behaviour of a spline curve can be constrained even if there is insufficient data to support a basis function.

As an illustration in 1D of the kinds of difficulties that can arise with conventional filtering, consider Figure 4.6 which compares B spline approximation to adaptive filtering. While the smoothing properties of the two are comparable on an infinite domain, the fact that data is only available on the interval $[-1, 1]$ causes the filter to produce an artifact at the boundary. This artifact occurs despite the fact that the filter adapts to the boundary by only making use of data within the domain. Without this adaptive property, the filter produces an even larger artifact. In contrast, the B spline approximation is largely unaffected by the boundary, and in general requires no special treatment to deal with missing data.

Details of computing a B spline approximation can be found in the Appendix A. The smoothness of the approximation is determined by two parameters: ω, referred to as the smoothing parameter, and d, the distance between basis functions. Since splines are being used as a filter for this application, the smoothness of the approximation must be chosen rather than derived from the data [45, 52]. The relationship between the smoothness of the approximation and the smoothing parameter is non-linear. However, since the normalization of the B spline has been chosen to eliminate the dependence of ω on scale and number of data points, ω can be fixed and the distance between basis functions varied instead.

In summary, beyond the task of choosing the smoothing parameter ω, the basis function distance d, and the deconvolution kernel parameters FWHM and Z, N3 is a fully automated method of correcting for intensity non-uniformity that makes few



(a) (b)

Figure 4.6: Comparison of spline smoothing and adaptive filtering. The true curve is given by the function $f(x) = -0.54 + 0.75e^{-x} + x^2$. The curves shown in (b) are an expanded view of the right portion of (a). While both the adaptive filter and B spline have comparable performance at the left boundary where the slope of $f(x)$ is small, the large slope at the right causes the adaptive filter to produce an artifact.

assumptions about the data. As will be shown in the next chapter, these parameters can be fixed at values that are nearly optimal for broad classes of images.

Chapter 5

Validation of the N3 method

A two fold approach was taken to validating the N3 algorithm: (i) the method was established to work on data which satisfies its assumptions; (ii) the method's behaviour was characterized on realistic simulated data for which the assumptions that voxel intensities are independent identically distributed random variables and that measurement noise is negligible are less valid. At the same time, the sensitivity of the method to its parameters was investigated.

5.1 Correcting random fields

5.1.1 Experiments

As a first step in validating the N3 method, consider a case in which the measurements are independent and identically distributed random variables. Suppose the distribution of the true intensities U is as shown in Figure 5.1a. While this distribution was taken from a T1 weighted MR scan, the choice of U for the purpose of this example is arbitrary. The volume is a cube with thirty-two voxels on an edge and the non-uniformity field to be removed is a parabolic function aligned with the center of the volume. The corresponding distribution F of the non-uniformity field is shown in Figure 5.1b along with the distribution V in 5.1c. A slice through the center of this idealized volume is shown in Figures 5.1d and 5.1f before and after the non-uniformity field is applied. Although the volume itself appears noise-like, measurement noise n is zero. Such a volume fully satisfies the assumptions of N3.

The idealized volume shown in Figure 5.1 was used to evaluate the performance of N3. Different widths of the deconvolution kernel were compared for convergence rate and accuracy in removing the non-uniformity field. In addition, N3 was compared with direct B spline filtering of the data. While direct B spline filtering is not acceptable for data that is structured, it is suitable for the random field of this example.

The performance of a correction method is best evaluated by comparing the field estimate to the true field if it is known. For example, a field estimate made for the random volume of Figure 5.1f is shown in Figure 5.1h along with the actual field in 5.1g. Besides differing subtly in shape, the two surfaces also differ by a multiplicative factor. This scale factor has no impact on the quality of the correction since MR intensity is relative. However, in quantifying the performance of the method this factor needs to be removed. Equation (4.14) can also be used for this purpose if r_n is taken as the ratio of estimated to actual field intensity at location $\mathbf{x}_n$. For the example of Figure 5.1, this is the coefficient of variation of the field shown in 5.1i.

5.1.2 Results

Correcting for intensity non-uniformity in a random field illustrates that N3 is able to correct volumes that lack regions of contiguous tissue. The improvement in the field estimate with iterations for four different widths of the estimated field distribution F is shown in Figure 5.2. In all four cases the method was converging in the sense that the distance between subsequent field estimates, computed using equation (4.14), became small (less than 0.02%). Note that the FWHM of the distribution F had an impact on both convergence rate and and, to a secondary degree, on performance. While there was a subtle reduction in final accuracy, there was a substantial increase in convergence rate as the FWHM increased.

For comparison, the same B spline approximation that is used in the N3 method was applied to the uncorrected data directly. The smoothness of the two is comparable

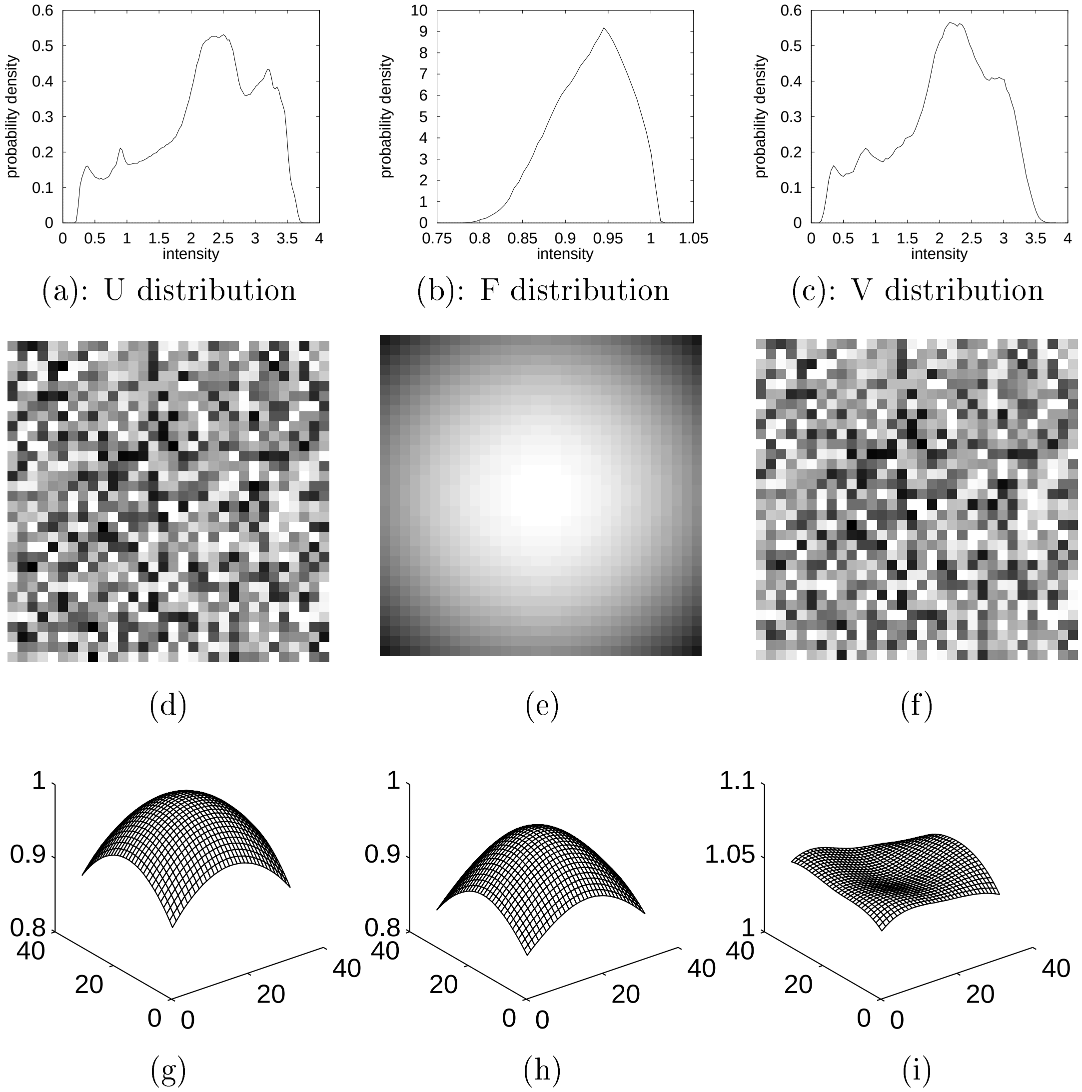


Figure 5.1: (a) Intensity distribution U for an idealized volume. Random samples from this distribution were used to create the volume in (d). (b) Distribution F for a parabolic non-uniformity field. (c) Distribution V corresponding to the volume shown in (f). (d) A slice through a random volume. (e and g) A parabolic non-uniformity field. (f) The volume in (d) multiplied by the parabolic field in (e). (h) Estimated field. (i) Ratio of actual field to estimated field.

since the N3 method, unlike conventional iterative filtering, has been designed such that field smoothness does not accumulate with iterations. As expected, since the N3 method takes advantage of the structure of the distribution U, it outperformed simple smoothing in all four cases.

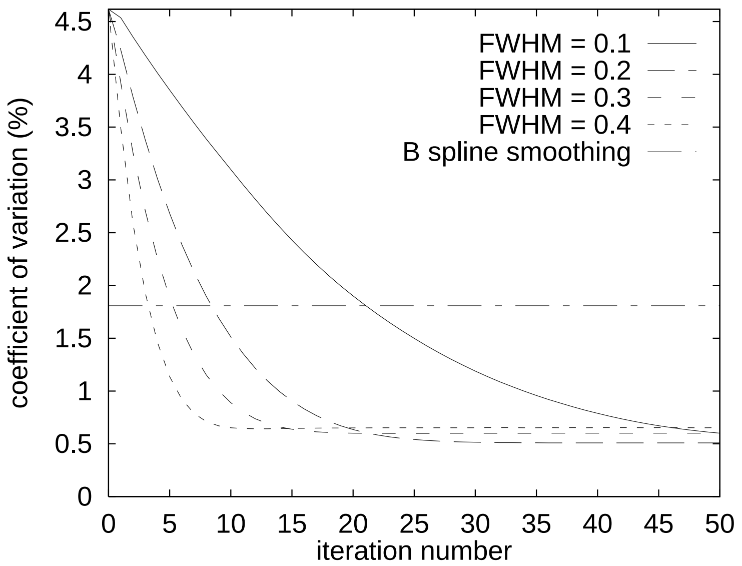


Figure 5.2: Error measure versus iterations for four different widths of the deconvolution kernel F. The FWHMs of the field distributions F used in deconvolution were 0.1, 0.2, 0.3, and 0.4. The maximum coefficient of variation 4.5% corresponds to no correction. The horizontal line is the result of smoothing the uncorrected data directly using B splines.

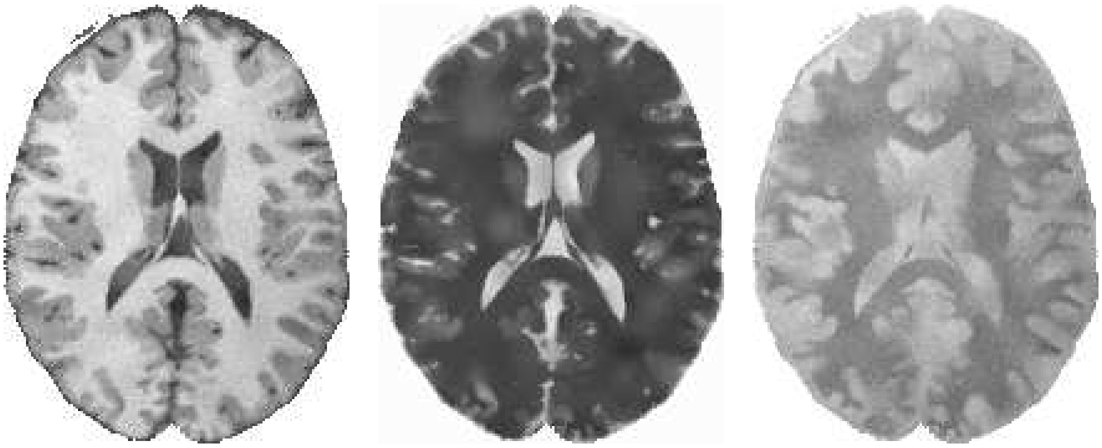
5.2 Simulated MR volumes

Non-uniformity correction methods such as N3 can be validated indirectly with real data using subjective measures of image quality and by assessing the reduction of variability in tasks such as segmentation. However, the large number of uncontrolled factors in such experiments confound attempts to evaluate and optimize performance. In particular, partial volume effects, true anatomical variability, and an unknown non-uniformity field prevent a sensitive analysis of the influence of various methodological parameters. These technical issues were circumvented in this analysis by the use of an MR simulator which incorporates realistic models for noise and partial volume. Slices through simulated T1, T2, and PD volumes are shown in Figure 5.3.

The anatomical model for the simulations was derived from high quality T1, T2, and proton density (PD) weighted scans formed from the average of twenty-seven, twelve, and twelve scans respectively [25, 26] of a normal individual. The three volumes were corrected for intensity non-uniformity using the N3 method and resampled into a standardized three dimensional, so-called stereotaxic, coordinate system[1] [51, 9]. The brain region of the T1 weighted scan was then segmented into the three tissue classes using a manually trained minimum distance [2, 30] classifier. Among the results of several different classifiers, this segmentation was chosen and manually edited by a trained neuroanatomist to improve the classification of deep structures and brain stem. Partial volume regions were formed by eroding each of the tissue regions using a six neighbour structuring element. The proportions of each tissue class in the partial volume regions were taken from a segmentation of the T1 volume using a Bayesian classifier [14, 28]. Non-brain structures such as scalp were not of interest for the present analysis and not included in the anatomical model.

The MR simulator is designed to produce volumes with an intensity histogram similar to that of real data, based on a discrete labelling of each voxel as either cerebrospinal fluid (CSF), grey matter, or white matter. However, representing an MR volume in terms of a few classes of homogeneous tissue is unsatisfactory for simulation since it does not reflect the variability seen in real data. Often the properties within a given tissue class vary for different structures included in that class [22], which leads to a broadening of the tissue peaks in the intensity histogram. The approach taken to produce a simulation with a realistic histogram is to color each region with random noise having medium to low spatial frequency and an intensity distribution

[1]This device allows for any neuroanatomical location to be addressed by a Cartesian coordinate and, by removal of affine differences among individual brains, for population statistic images to be built [29, 16].



T1 T2 PD

Figure 5.3: Simulated T1, T2, and PD weighted volumes. Non-uniform sensitivity of the reception coil has been simulated causing the intensity to drop off at the bottom right. The noise distribution has a standard deviation 3% of the mean intensity of white matter.

appropriate for the given tissue type. This noise can be viewed as "pseudo-anatomy," representing low frequency spatial variations within tissue due to anatomy which gives rise to a 'lumpy' appearance in the image (e.g. see the T_2 image in Figure 5.3). By validating the correction algorithm on a number of realizations of this pseudo-anatomy, the results were not biased towards any preferred shape of the non-uniformity field.

To complete each simulation, a multiplicative intensity non-uniformity field was imposed and Rician distributed noise added throughout. Rician noise, typical of that found in real MR images, is simulated by adding a complex Gaussian distributed random intensity to a voxel, then computing its absolute value.

The intensity distributions used to generate the pseudo-anatomy were created by computing the histogram of intensities within the eroded or pure tissue regions for each of the T1, T2, and PD averages. These distributions are shown in Figure 5.4. It is assumed that these distributions reflect the range of intensities intrinsic to the tissue class uncorrupted by noise, intensity non-uniformity, and partial volume effects. Random fields having these intensity distributions are created by rejection sampling [13]. Each random field is first constructed at 10 mm resolution and subsequently resampled to the desired 1 mm resolution using tricubic interpolation.

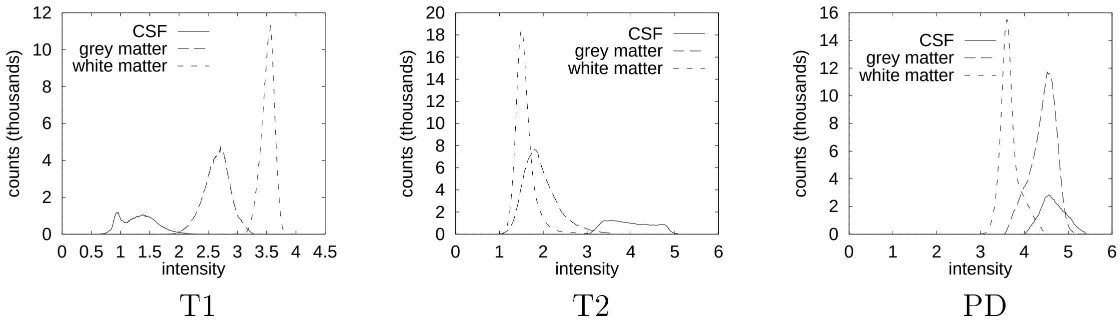


Figure 5.4: Intrinsic intensity distributions for each tissue class in T1, T2, and PD weighted simulations.

5.3 Correcting simulated data

Although N3 does not require a model of the expected tissue distributions, there are still a number of model-independent parameters that need to be selected (see Figure 4.5). These are the smoothing parameter ω, the distance between basis functions d, the noise term in the deconvolution filter Z, and the FWHM of the deconvolution kernel F. Experiments with simulated data show little dependence on Z, which was fixed arbitrarily to be 0.1 for the purpose of this analysis. The parameters ω and d both control smoothness, but the relationship between smoothness and ω is more complex than the one between smoothness and d. As a result, ω was fixed arbitrarily at 1.0. The remaining two parameters, FWHM and d, are considered further.

5.3.1 Characterizing the FWHM parameter

Simulated MR volumes like those shown in Figure 5.3 were used to evaluate the effect of FWHM on correction performance. Recall from Chapter 4 that FWHM refers to the width of the deconvolution kernel used to compute local estimates of the non-uniformity field. For this analysis, two different non-uniformity fields were generated from combinations of linear, quadratic, and Gaussian terms. Both fields vary in magnitude by 20% within the brain volume, which is typical for brain scans (see Figure 4.1), however, the second field has more curvature. Slices from these two fields are shown in Figure 5.5. The first field was used to create the simulations shown in Figure 5.3. The performance of the correction method was evaluated using equation (4.14) to compute the difference between the estimated field and that imposed explicitly during simulation.

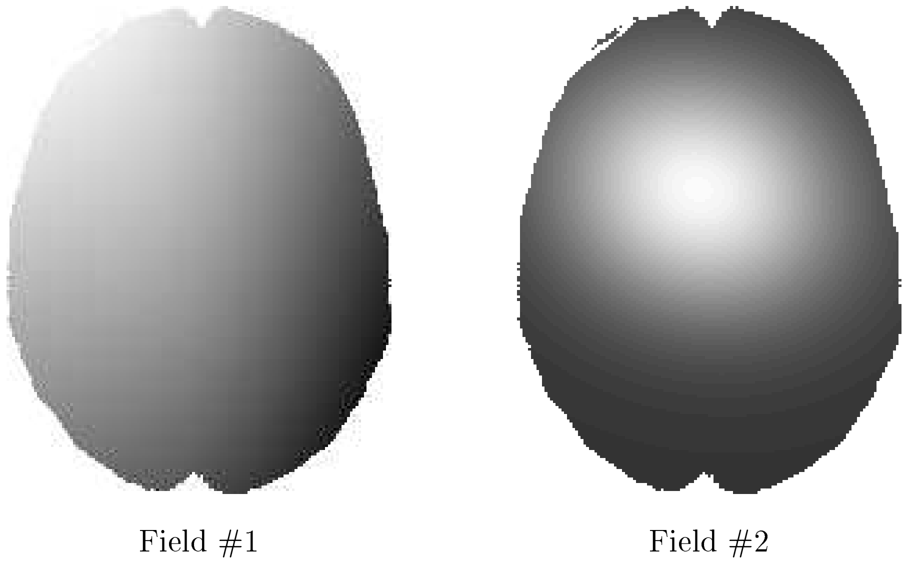


Figure 5.5: Slices through two non-uniformity fields used to construct simulated MR scans.

The performance curves for four simulated brains and a range of the FWHM parameter are plotted in Figure 5.6 for the two different non-uniformity fields shown in Figure 5.5 and fixed basis function distance $d = 200\,\mathrm{mm}$. A coefficient of variation below the lines corresponding to no correction in Figure 5.6 indicates a reduction in intensity non-uniformity. Within most of the range of the FWHM shown, the N3 method substantially improved uniformity. However, the quality of the result depended on the choice of the FWHM parameter. The general trend in Figure 5.6 is for performance to improve with decreasing FWHM. Although much weaker, the same trend is present in the results for random fields shown in Figure 5.2.

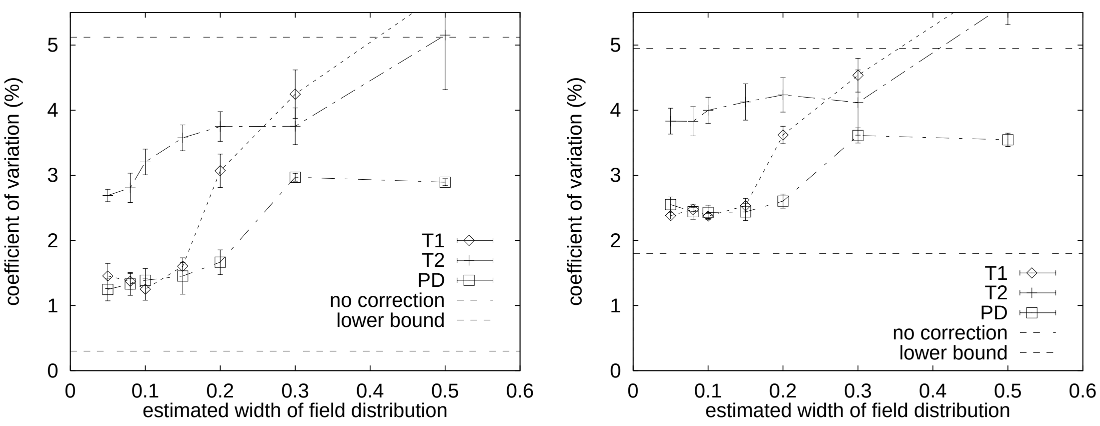
Another trend, not reflected in Figure 5.2, is that the rate of convergence is substantially slower as the FWHM parameter is decreased. For these experiments the stopping criterion was reduced to $e = 0.0002$ from its default value of 0.001 so that iteration would not stop prematurely in cases where the convergence was extremely slow. For example, with a FWHM of 0.05 the method typically required 50 iterations to reach this threshold, significantly longer than the the 20 iterations needed at a FWHM of 0.2. In practical terms, the choice of the FWHM requires a tradeoff between computation time and accuracy. For all subsequent experiments, the FWHM parameter was fixed arbitrarily at 0.15.

It should be noted that zero estimation error is not achievable for these experiments since even if an algorithm were insensitive to the low spatial frequency components of the pseudo-anatomy and partial volume, there is still a low frequency component to the broad-band Gaussian noise that is indistinguishable from intensity non-uniformity. Furthermore, smoothness constraints imposed on the spline fitting operation preclude an exact match to the field. The lower bound imposed by spline approximation, computed on the imposed field directly, is also shown in Figure 5.6. Noise and pseudo-anatomy make the true lower bound somewhat higher.

5.3.2 Characterizing the basis function distance parameter

In addition, an analysis of the effect of basis function distance d on correction performance, for a fixed FWHM, was done using the same simulated data. Recall that the field estimate becomes smoother as d increases. These results were compared against the results of fitting B splines to the imposed field directly.

The results of correcting simulated brains for a range of basis function distances d are shown in Figure 5.7. The greater curvature of the second field produces broad minima in the error curve. As in Figure 5.6, the lower bound on the estimation error given by directly fitting splines to the non-uniformity field is also shown. Basis function distances of 100 mm to 200 mm are nearly optimal for these two cases. It should be noted that due to the need to suppress noise in the simulations, the optimal basis function distance is greater than that which can exactly match the non-uniformity

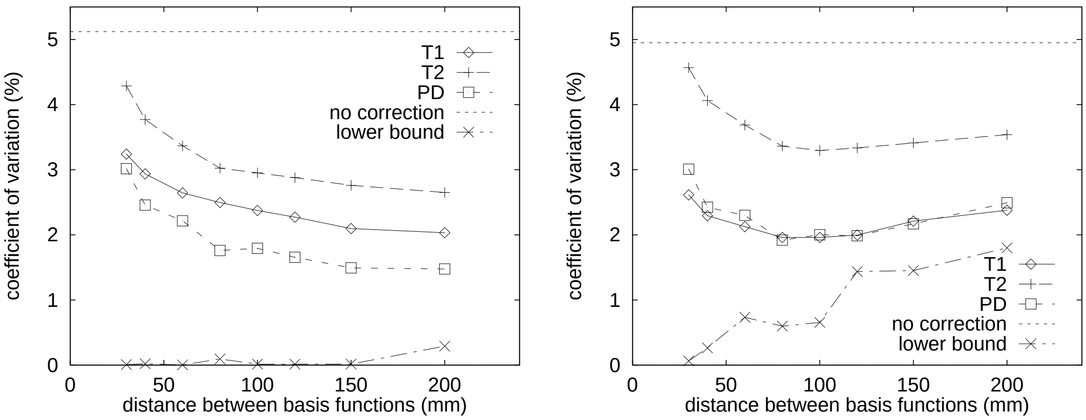


Field #1 Field #2

Figure 5.6: Field estimation error on simulated data for a range of the FWHM parameter and two different non-uniformity fields. Lines are shown for the coefficients of variation corresponding to no correction for the two cases. Also shown is the lower bound on estimation error given by fitting splines to the non-uniformity field directly. The error bars are at plus and minus two standard deviations. These results are based on four realizations of the pseudo-anatomy.

field. In other words, the optimal d for suppressing noise introduces some distortion in the underlying field.

These results show not only that the N3 method is able to substantially reduce non-uniformity in realistic simulated data but also that the parameters FWHM and d can be chosen to be nearly optimal for a broad class of data. This allows the method to be fully automated, freeing the user from determining suitable parameters for each type of scan.



Field #1 Field #2

Figure 5.7: Correction performance as a function of basis function distance on simulated T1, T2, and PD scans. The FWHM parameter for this analysis was 0.15. Also shown is the approximation error for fitting splines to the non-uniformity field directly. The discontinuities in this curve are caused by changes in the number of basis functions, which for efficiency is the minimum needed for full support at the given distance between basis functions.

Chapter 6

Comparison with other methods

Simulated and real data were used to evaluate the performance of the N3 method relative to three other techniques. While these results do not show the N3 method to be superior in every category, the method distinguished itself for uniformly good performance across a range of different of tests.

6.1 Non-uniformity correction methods

Three non-uniformity correction methods were selected to compare their performance to that of N3 in removing typical levels of non-uniformity from volumetric scans of the human head. The methods are referred to respectively as the expectation maximization (EM) [56], revised expectation maximization (REM), and white matter (WM) [11, 60] methods. All four methods are fully automated three dimensional implementations formulated to detect and remove a multiplicative intensity non-uniformity field by iterative approximation.

To facilitate automatic training of the classifiers which the EM, REM, and WM methods rely on, all MR volumes were first resampled into 3D stereotaxic space. This standardized coordinate system was also used to construct 3D tissue probability maps (TPM) for grey matter, white matter, and cerebrospinal fluid (CSF), based on the classification of 53 normal subjects [59].

Classifier training sets were created for each individual brain by randomly selecting voxels in stereotaxic space corresponding to voxels in the TPM thresholded at the 90% level. This approach allows automatic training of the classifier regardless of the volume's pulse sequence [30]. In addition, non-brain tissues were removed automatically using a standard brain mask derived from an average of normal brains previously defined in stereotaxic space [9]. While this mask, which was derived from an average of 305 T1 weighted brain scans, may include non-brain tissues or exclude small portions of cortex in some individuals, this does not limit the use of these correction methods since the derived correction field can be extrapolated to the whole volume using B splines (see Appendix A).

With the exception of the WM method, which terminates after two iterations, all of the methods terminate when the coefficient of variation of the ratio between consecutive field estimates, computed at all intracranial voxels, drops below a threshold of 0.001 . Typically, the methods terminated after ten iterations.

While all of the volumes used in this analysis were 1mm isotropically sampled data, for efficiency these volumes were subsampled without blurring to a 3mm resolution to speed up the correction process. At this resolution all four methods took between twenty and thirty minutes to complete on a workstation with floating point performance of 99 SPECfp92.

6.1.1 Expectation maximization method

The first method considered is the EM method described in [56]. This method iterates between a classification stage, designed to remove anatomical features from the image, and a filtering stage that smoothes the field estimate. This filtering is done with a Gaussian filter whose kernel is truncated to remain within the volume of interest. A Bayesian classifier [14, 29] was used in this implementation; Wells et al. have since described an implementation [57] based on a Parzen window classifier. Both the EM and REM methods use 256 training samples for each of the white matter, grey matter, and CSF tissue classes.

Experiments, not shown here, on simulated data with various widths of the smoothing kernel suggested that a Gaussian kernel with full width at half maximum (FWHM) of 30 mm was nearly optimal. While results produced with a 30 mm kernel appeared visually "lumpy", larger kernels produced a significant edge artifact. This artifact was caused by the increasing phase lag as the filter was truncated when approaching the edge of the volume of interest. For example, if the magnitude of the field was rising as it approached an edge then the truncated filter tended to underestimate the field strength at the edge. This effect is the 3D generalization of that shown in Figure 4.6.

6.1.2 Revised expectation maximization method

The second method evaluated is a revised version, developed by the author, of the EM method. This method takes better advantage of the TPMs by using them as probability priors during classification. These priors stabilize the method by limiting the extent of misclassification in volumes that are extremely non-uniform.

The second difference between the REM method and the EM method is the use of tensor cubic B splines for smoothing (see Appendix A for details). Unlike truncated filtering, B splines do not produce an edge artifact. The distance between basis functions used throughout was 200 mm, significantly larger than the 30 mm kernel used by the EM method.

6.1.3 White matter method

The third method, the WM method, is a revised version of that described in [11, 60]. It relies on an artificial neural network classifier [61] to identify white matter in the brain region. The classifier is trained on 500 points per class selected from TPMs thresholded at the 100% level, that is stereotaxic voxels that were classified as belonging to the same class for all 53 normal subjects. Partial volume voxels are then eliminated from the white matter region using gradient information. Tensor cubic B splines are used to fit a smooth field to the remaining white matter voxels and extrapolate this field to the rest of the volume. For the experiments described here, the smoothing parameters are the same as those used by the REM method.

Once estimated, the non-uniformity field is used to correct the original volume and the process is repeated once. Experiments with real data suggested that beyond two iterations the field estimate is essentially unchanged.

6.1.4 N3 method

On the basis of the results of Chapter 5, the parameters of the N3 method were fixed throughout this analysis, irrespective of the type of volume. In particular, the FWHM of the histogram blurring kernel was fixed at 0.15, the deconvolution parameter $Z = 0.1$, and the smoothing parameters $d = 200$mm and $\omega = 1$. These smoothing parameters are the same as those used by the REM and WM methods.

6.2 Correcting simulated data

Simulated MR volumes like those shown in Figure 5.3 were used to evaluate the performance of the four methods at various levels of noise and non-uniformity. The non-uniformity field used in these experiments was Field #1 shown in Figure 5.5. This field was created using a real MR scan as a guide and has typical curvature. Correction performance was computed in terms of the coefficient of variation in the ratio of the estimated non-uniformity field to the known field imposed during simulation using equation (4.14).

6.2.1 Sensitivity to noise level

A Rician noise distribution with standard deviation 3% of the mean intensity of white matter is typical of T1 weighted volumes. A comparable level of noise was chosen for T2 and PD volumes and termed 3% noise.

The results of experiments on simulated data with varying noise level are shown in Figures 6.1a through 6.1c. Also shown is the level of non-uniformity present before correction (baseline). Results above this line indicate that the method has made the non-uniformity worse. To estimate the intersubject variability in these results the experiments were repeated for four realizations of the pseudo-anatomy. This variability is reflected in the error bars shown on Figure 6.1.

To address concerns about the sensitivity of the EM method to the training set, its performance was also measured using an ideal training set. This training set consists of samples chosen within regions of pure tissue taken directly from the model used to generate the simulated data. These results are also shown on Figure 6.1.

6.2.2 Sensitivity to non-uniformity magnitude

A similar set of experiments was conducted for increasing levels of non-uniformity using the same shape of the non-uniformity field, but different magnitudes. Fifty percent non-uniformity denotes a non-uniformity field that varies between 0.75 and 1.25, that is, the range of field variations within the brain region is $\pm 25\%$. The results of these experiments are shown in Figures 6.1d through 6.1f. Also shown is the level of non-uniformity corresponding to no correction. As described in Section 5.3.1, this level does not include the implicit field, which is present even at zero non-uniformity, due to the noise and random anatomy.

6.3 Correcting real data

The second type of experiment performed was a comparison of the reduction in variability in the intensity of tissue manually labelled on real data. For this experiment, regions of pure grey and white matter were manually labelled on twelve sets of T1, T2, and PD weighted scans acquired on twelve different MR scanners. Included in the twelve are machines made by Philips, Siemens, and GE. The volumes were transformed into stereotaxic space so that the same labelling could be applied to each of the three modalities. All twelve individuals are MS patients having a moderate number of white matter lesions.

For each volume, the coefficient of variation in white and grey matter intensity was computed before and after correction. Since for real data, this measure cannot distinguish between intensity non-uniformity, noise, and anatomical intensity vari-

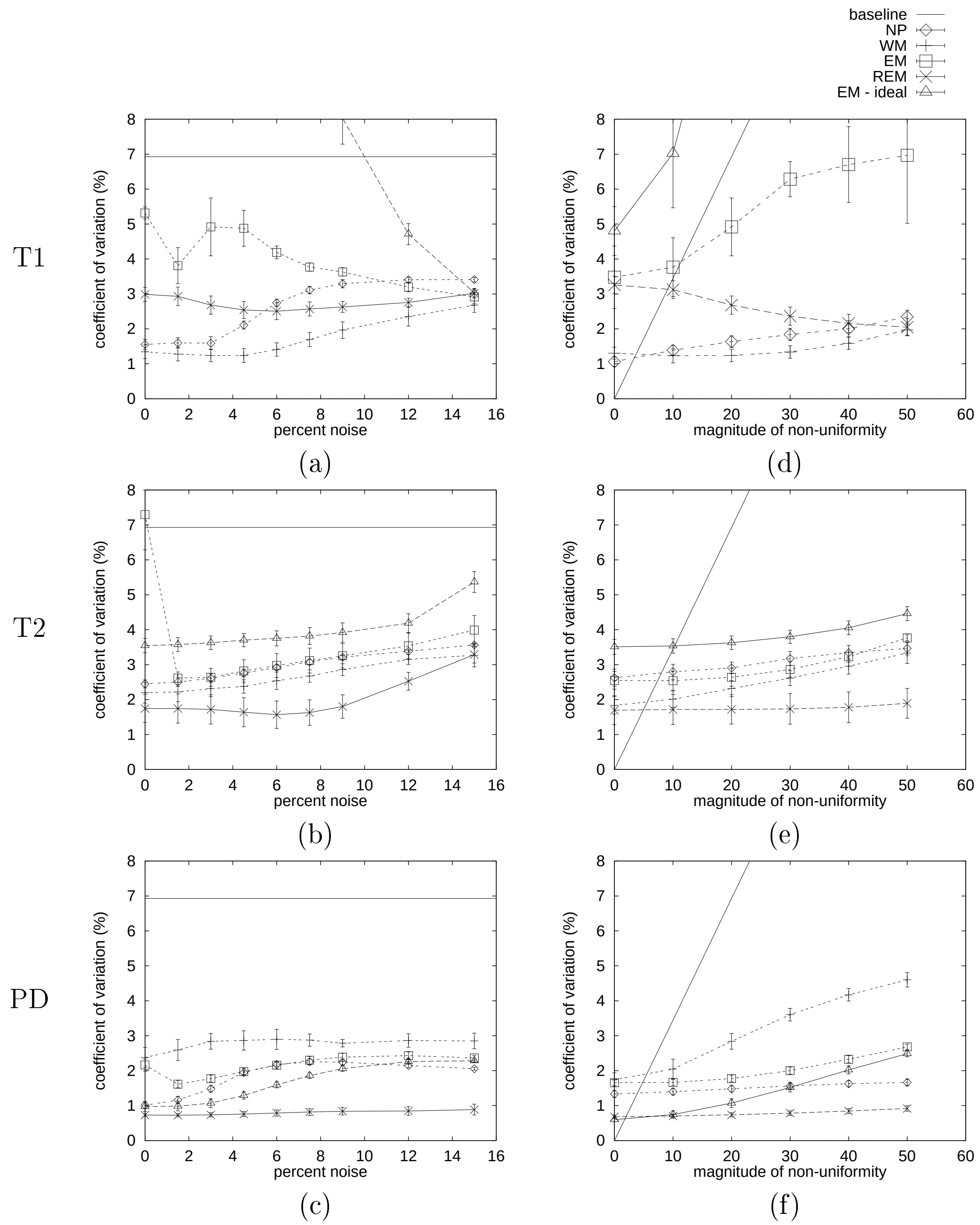


Figure 6.1: A comparison using simulated data of the four non-uniformity correction methods. (a) through (c) are for increasing noise level and (d) through (e) are for increasing levels of intensity non-uniformity. The trace “EM - ideal” is the performance of the EM algorithm with an ideal training set (see text). The error bars are at ±1 standard deviation based on four trials.

Table 6.1: Mean coefficient of variation in white matter tissue intensity.

	T1 weighted	T2 weighted	PD weighted
no correction	5.8%	10.0%	6.4%
EM	7.2% (+)	8.4% (-)	4.4% (-)
REM	6.6% (+)	8.4% (-)	4.7% (-)
WM	4.4% (-)	8.6% (-)	4.8% (-)
N3	5.1% (-)	9.0% (-)	4.9% (-)

ations, it is only suitable for showing that intensity non-uniformity is qualitatively reduced.

The coefficient of variation, the ratio of standard deviation to mean, for white and grey matter intensity was computed for each patient in each modality. The mean across patients is shown in Tables 6.1 and 6.2 for white and grey matter respectively, before and after correction for intensity non-uniformity. The statistical significance of these results was computed using a one sided Wilcoxon signed rank test for a paired difference experiment. The notation $(-)$ and $(+)$ denotes significant decreases and increases in non-uniformity tested at the 95% confidence level.

It should be noted that when these experiments were first conducted, the EM and REM methods performed extremely poorly, tending to make non-uniformity worse. Hypothesizing that this was caused by dark regions in the sinuses that occasionally lie within the average brain mask [9], the experiments were repeated using an intensity threshold to removed low intensity and background voxels. This threshold was chosen automatically using the technique described in [42]. These results are the ones presented in Tables 6.1 and 6.2.

As a final result, the T1 weighted twenty-seven average scan is shown before and after correction by each method in Figure 6.2. This low-noise volume is useful for illustration since intensity non-uniformity that would normally be obscured by noise is clearly visible.

The three types of comparisons made between the methods have their respective

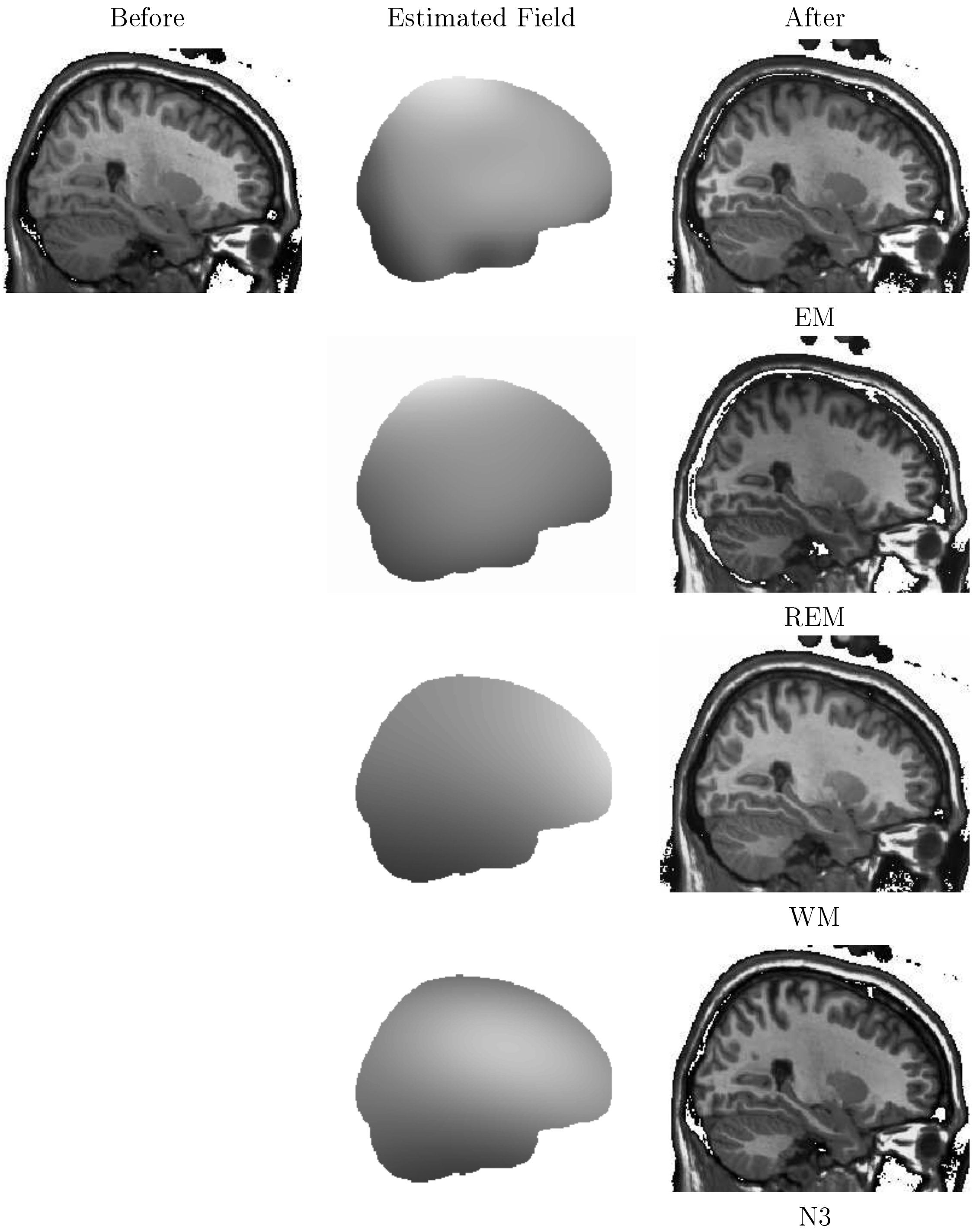


Figure 6.2: Intensity non-uniformity correction of a T1 weighted gradient-echo MR scan by each of the four methods.

Table 6.2: Mean coefficient of variation in grey matter tissue intensity.

	T1 weighted	T2 weighted	PD weighted
no correction	10.1%	15.4%	8.8%
EM	12.6% (+)	14.3% (-)	6.8% (-)
REM	9.6%	13.8% (-)	6.4% (-)
WM	9.7%	15.7%	7.3% (-)
N3	9.8%	14.3% (-)	6.8% (-)

merits. Comparison using simulated data is well controlled and quantitative, yielding an absolute measure of the reduction in non-uniformity. As shown by Figure 6.1, the individual algorithms have their strengths. With the possible exception of T1 data, the REM method performed best in these trials. However, the validity of simulation experiments is limited by the fact that neither the MR imaging process nor the anatomy have been completely characterized.

Real data, while unable to provide quantitative results, is important for uncovering limitations of a method that are caused by factors not reflected in the simulations. For example, the EM and REM methods proved unexpectedly sensitive to the choice of brain mask. Even when this masking problem was fixed, these methods performed poorly on T1 data, tending to make non-uniformity worse. While it is not clear what aspect of real T1 data causes the REM method to perform poorly, it should be noted that the REM method is highly model dependent, relying on a three class tissue model and TPMs as probability priors. It is possible that the mild atrophy present in MS patients or perhaps merely natural anatomical variability is a sufficient violation of these model assumptions to cause the method to fail.

Consulting the T1 volume shown in Figure 6.2 sheds some light on the way the methods behave. The EM and REM methods have excessively reduced the intensity at the top of the head, a typical failing of these methods. In contrast, both the WM and N3 methods have reduced the intensity of the anterior portion of the head. The N3 method has also significantly raised the intensity of the cerebellum, the large structure at the bottom left. Visual inspection of the results suggests these corrections are appropriate. Although, it is useful to show that non-uniformity can be substantially reduced in an absolute sense, these results need to be put in perspective by determining what impact the reduction has on real applications. This is the subject of the next chapter.

Chapter 7

Applications

Growing interest in measuring anatomical differences in brain structure among normal and disease populations has led to the development of a variety of techniques for automatically identifying these structures on standard MRI scans. However, such analysis may be compromised by MRI intensity non-uniformity artifacts. In this investigation, simulated MRI data was used to assess the impact of non-uniformity on three different anatomical brain mapping techniques: (i) tissue classification, (ii) cortical surface extraction, and (iii) non-linear co-registration. The advantage of using simulated data is that anatomical structures identified by the three techniques can be compared to the anatomical model upon which the simulations are based.

7.1 Tissue classification

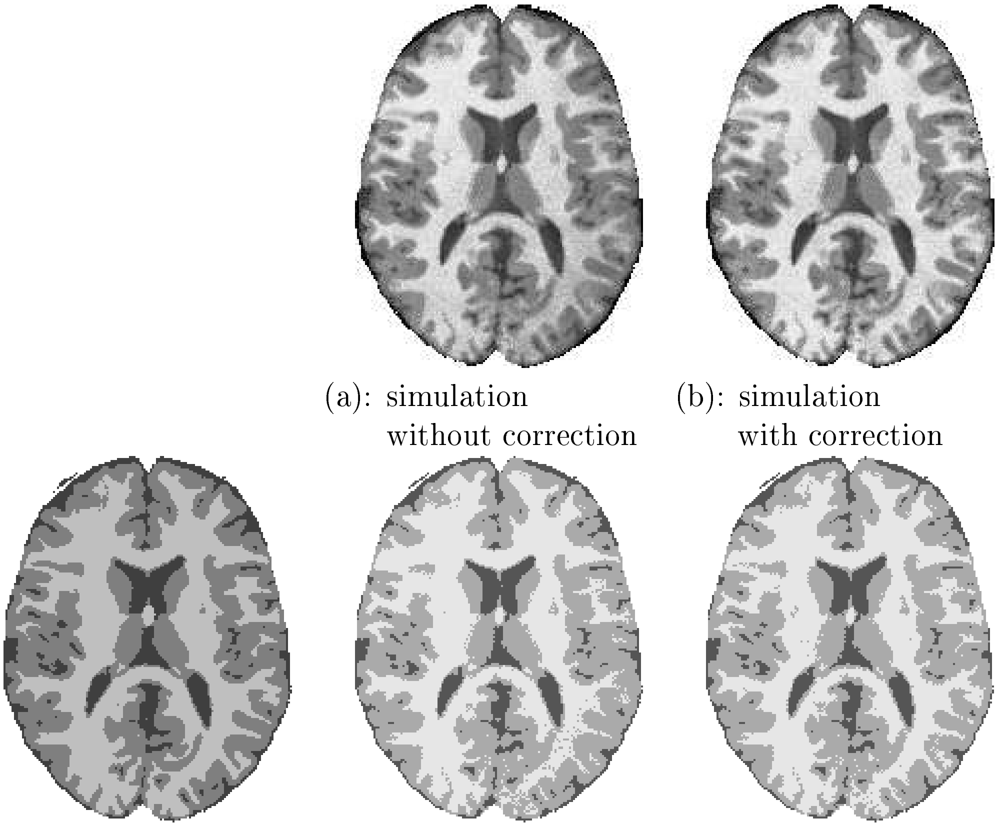
Tissue classification experiments were conducted using an artificial neural network tissue classifier (ANN) trained using probability priors defined in stereotaxic space [30]. This is a fully automated method of labelling each voxel as belonging to a particular tissue class, which for these experiments means one of grey matter, white matter, or cerebrospinal fluid (CSF). The training data was created using 1000 samples[1] per tissue class chosen with prior probability of at least 90% using tissue probability maps

[1] ANN classifiers require more training data than Bayesian or minimum distance classifiers since they have more degrees of freedom.

(TPM).

The quality of the classification was evaluated by comparing the labelled volume to the original discrete tissue model from which the simulations were derived. For example, a slice from the simulation model is shown in Figure 7.1c. This model was used to create the simulation with 20% non-uniformity shown in 7.1a. A slice from the labelled volume created by classifying this simulation is shown in Figures 7.1d. Non-uniformity causes white matter at the bottom right of the volume shown in 7.1d to be misclassified as grey matter. The volume in 7.1e is the same except that it has been corrected using the N3 algorithm prior to classification, resulting in improved classification of the white matter at the bottom right. The measure *kappa* [7, 30], a similarity measure corrected for chance, is used to quantify the improvement between the two cases. A kappa of unity indicates perfect agreement, while a kappa of zero indicates that any agreement is purely coincidental. Kappa values computed within the brain region are shown for each labelling in Figure 7.1. Note that the classifier does not achieve a kappa of unity even in the absence of non-uniformity because noise and partial volume degrades its performance.

To assess the impact that non-uniformity has on correction and the benefits of using N3, T1 weighted simulated volumes like those described in Chapter 5 were classified before and after correcting for non-uniformity. These volumes have field #1 from Figure 5.5 at various levels of severity. For the experiments, four realization of the pseudo-anatomy at six levels of non-uniformity were used. The results, plotted in Figure 7.2, show a clear reduction in classification performance for uncorrected volumes as non-uniformity increases. However, this degradation is avoided by first correcting the volume using the N3 method. The low variability in these results suggest that all of the differences are significant. Given that classifiers such as ANN work by choosing intensity thresholds that differentiate the various tissue classes, it is not surprising that their performance depends on having uniform data.



(a): simulation without correction

(b): simulation with correction

(c): model
kappa = 1.0

(d): classification without correction
kappa = 0.81

(e): classification with correction
kappa = 0.83

Figure 7.1: Volumes labelled by tissue type using an artificial neural network classifier. (a) A simulated T1 weighted volume with 20% non-uniformity. (b) The volume in (a) after correction with N3. (c) The discrete model used to generate the simulation. (d) A labelling of the volume in (a). (e) A labelling of the volume in (b).

7.2 Cortical surface extraction

Similar experiments were performed using the cortical surface extraction method described in [34]. This method starts with an inflated surface model of the cortex that is iteratively shrunk to fit tightly around the cortex defined by the change in image intensity at the brain-CSF interface.

The standard to which the extracted cortical surfaces were compared was created by fitting a surface to the grey matter portion of the simulation model. A rendering

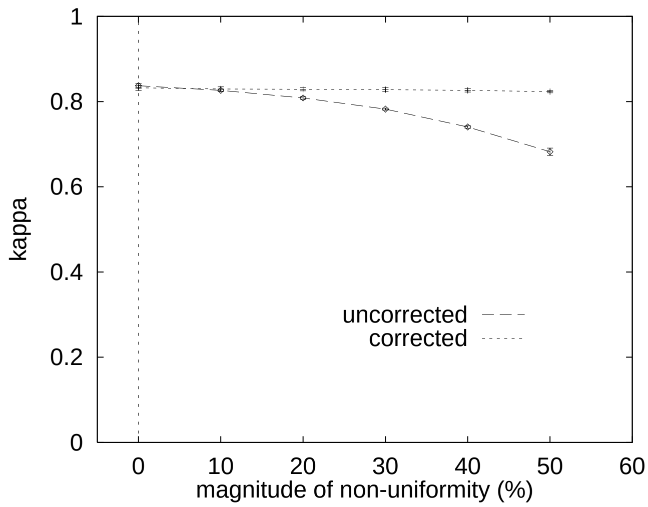
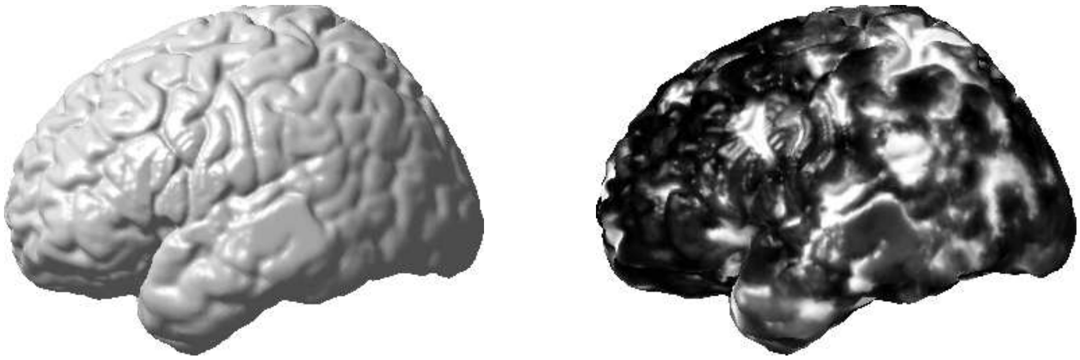


Figure 7.2: Classification performance versus severity of intensity non-uniformity for corrected and uncorrected T1 weighted volumes. The error bars are at plus and minus one standard deviation computed for four individuals.

of this surface is shown in Figure 7.3 along with a map indicating the displacement of a surface extracted from a simulation with 20% non-uniformity. Surfaces are represented internally by the software as being tessellated by triangles. The method is designed to match anatomical features in the surface model to those found in the data. For example, in a correct fit to the data, major sulci on the extracted surface should line up with those marked on the model. Hence, if two surfaces are the same in this sense, then vertices of the corresponding tessellations should match. The error measure used in these experiments is the root mean squared (RMS) distance between corresponding vertices of the tessellated surfaces.

The results of fitting cortical surfaces to the same simulated data as used to evaluate tissue classification performance are shown in Figure 7.4. Like tissue classification, the quality of the surface fitting deteriorates as the severity of the non-uniformity increases. Once again, these effects are largely eliminated by first correcting for intensity non-uniformity. These results are consistent with the dependence of the surface extraction method on intensity thresholds.



(a) (b)

Figure 7.3: Cortical surfaces fit to simulated T1 weighted MR data. (a) a rendering of the cortical surface derived from the anatomical model. (b) the cortical surface in (a) texture-mapped with the distance between vertices of the surface fit to a simulated MRI and those of the anatomical model. The appearance of white around the sulci indicates that the method has the greatest difficultly finding the cortical surface in these regions. The density range is black for zero displacement through to white for displacements of 1.2 mm or more.

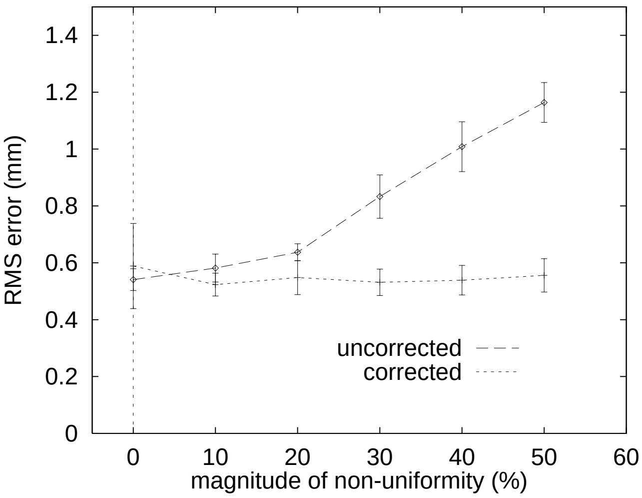


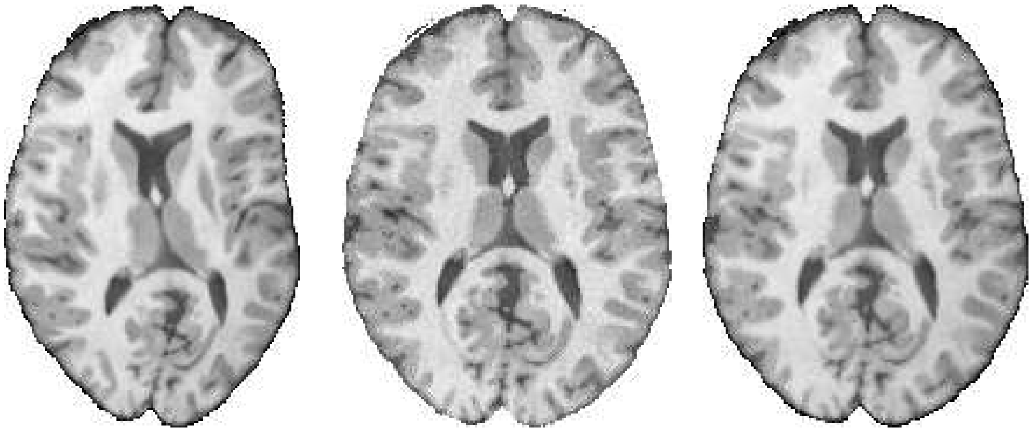
Figure 7.4: Cortical surface extraction error versus severity of intensity non-uniformity. These results are for four realizations of T1 weighted simulated MR volumes. Error bars are at one standard deviation.

7.3 Non-linear co-registration

The third brain mapping application considered is non-linear co-registration. This technique determines the geometric deformation needed to match the anatomy of one brain to another. One of its applications is in automatically labelling brain structures by warping a labelled model to match a given brain. For these experiments, the ANIMAL method [8] has been used to recover a random but known deformation applied to a simulated volume. While this test does not mimic the typical application of the method which is to match the anatomies of different individuals, it is suitable for assessing the impact that intensity non-uniformity has on the ANIMAL method.

In conducting the experiments, a random deformation was created which was then used to distort a simulated volume containing some level of non-uniformity. ANIMAL was then used to match the deformed simulation to an undeformed uniform (i.e. 0% intensity non-uniformity) simulation of the same individual. For example, the deformed volume shown in Figure 7.5a which suffers from mild intensity non-uniformity was co-registered to the uniform simulation shown in 7.5b to produce the volume shown in 7.5c. For these experiments the recovered deformation is computed at a resolution of 8 mm, i.e. the deformation field is represented as displacement vectors on an 8 mm lattice with intermediate displacement vectors determined by interpolation. Even under ideal conditions the recovered deformation will not exactly cancel features of the applied deformation that are beyond this resolution. To measure the quality of this cancellation the RMS of the residual deformation field is computed within the brain region.

The results of co-registering simulated volumes with and without correction for increasingly severe levels of non-uniformity are shown in Figure 7.6. While performance does decrease as the magnitude of the non-uniformity increases, the difference is only significant at extreme magnitudes, where correcting for non-uniformity does improve the results. However, it should be noted that 100% non-uniformity is extreme and rarely found in volumetric scans. The robust behavior of the ANIMAL algorithm is due to its use of image gradients rather than image intensities to define the deforma-



(a): deformed (b): uniform (c): co-registered

Figure 7.5: Simulated T1 weighted MRI volumes deformed and co-registered. (a) a deformed simulated volume having intensity non-uniformity. (b) a uniform and undeformed volume. (c) the volume in (a) after co-registration with the volume in (b).

tion fields. Image gradients are largely unaffected by the smooth variations associated with intensity non-uniformity.

The experiments described in this chapter show that intensity non-uniformity significantly degrades the performance of the tissue classification and cortical surface extraction methods tested. This is attributable to their reliance on intensity thresholds. In contrast, the co-registration method, which relies on image gradients rather than absolute intensities, was unaffected except by severe levels of non-uniformity. In combination with the N3 method all three mapping techniques provide good performance independent of the level of non-uniformity present. As anatomical analysis using these techniques can be employed to answer a broad range of questions of scientific and clinical interest, these results emphasize the need for intensity non-uniformity correction.

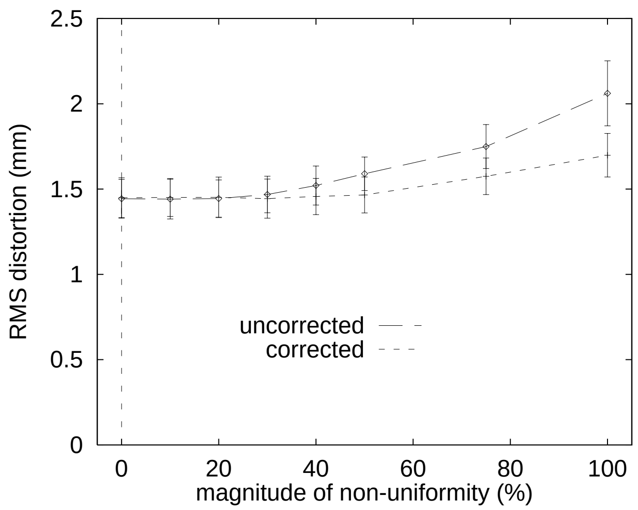


Figure 7.6: Error in deformation estimate versus severity of non-uniformity. Error bars are at one standard deviation, based on four different deformations.

Chapter 8

Discussion and conclusions

8.1 Characteristics of the N3 method

A new method, called N3, for correcting for intensity non-uniformity in 3D MR images has been described. N3 does not rely on a parametric model of tissue intensities, nor on segmenting a volume into contiguous regions. Instead, non-uniformity is assumed to blur the histogram of the data in a way that can be identified, quantified, and removed. The assumption that all of the information needed to derive a smooth correction field for the volume is present in the histogram is a major advantage of the N3 method, as it does not need to know anything about what an MR scan contains or how it was acquired. However, relying heavily on the histogram also has drawbacks, as it involves restrictive assumptions about the statistical properties of the data that are rarely satisfied in practice. To understand why the method is robust and accurate in correcting MRI data despite this apparent weakness, we must consider the assumptions of the method in more detail.

The first of these is that non-uniformity can be modelled as a smooth multiplicative field. While this model is shared by most other correction methods, it is particularly important to N3. As described in Chapter 2, the multiplicative field model accurately describes non-uniform sensitivity of a reception coil, but it is only an approximation to the tissue dependent character of excitation field non-uniformity. One can make two arguments that the discontinuities in the smooth field, caused by dependence of excitation non-uniformity on tissues properties, are small. First, standard spin echo sequences are designed to be insensitive to low levels of excitation field inhomogeneity. Second, excitation is often done with a scanner's body coil, which is larger and more uniform than the head coil used for reception. However, neither argument applies in all situations since other imaging sequences, such as gradient echo, can be sensitive to excitation field variations and, as was described in Chapter 2, much of the non-uniformity observed with clinical scanners is due to electrodynamic interaction with the subject rather than uniformity of the coils. Instead, one can draw upon the observation that, while some tissues may be more sensitive to excitation field variations than others for a particular pulse sequence, the trend in the non-uniformity field is still the same. Given that excitation non-uniformity accounts for at most half of the combined excitation-reception non-uniformity and that differences in sensitivity between tissues are a fraction of that half, one should accept that the multiplicative field estimated is an average of the similar but not identical multiplicative fields present in individual classes of tissue.

The second assumption of the N3 method is that tissue intensities are independent identically distributed random variables. This stationarity assumption essentially says that the various tissues present should be well mixed, much like a random field. Clearly, in objects of interest for MRI this is not the case. The brain, for example, has regions of homogeneous tissue corresponding to the various anatomical structures. Yet despite this difficulty, the method is able to correct for non-uniformity in brain scans.

Two factors allow N3 to tolerate violations of the stationarity assumption. The first is the simple form of the non-uniformity field. Since the field is required to be smooth, field variations introduced by localized structures in the image are suppressed by smoothing the estimated field. This is the same argument that was given for the filter and divide approach [23, 27, 31, 32, 46] to non-uniformity correction. However, smoothing alone is not sufficient to separate intensity non-uniformity from large anatomical structures. It is the second factor that distinguishes the N3 method from these other methods.

The N3 method is able to tolerate violations of the stationarity assumption because it is the local field estimates that are smoothed rather than the images intensities. For example, two different brain structures having widely different image intensities can still yield field estimates that are consistent with the true field since this information is derived from the sharpening of the histogram rather than the image intensity alone. However, not all tissues are equal in their power for estimating non-uniformity.

Consider correcting the non-uniformity in a brain scan using only grey and white matter. Since white matter is found in large contiguous regions, the peak in the histogram corresponding to white matter is narrower and better resolved than that for grey matter which is more prone to partial volume[1]. As a result of this well resolved peak, white matter is more powerful for estimating non-uniformity. Since there is a relatively narrow range of intensities corresponding to white matter, intensities outside of this range are more likely to be caused by non-uniformity. This estimation power is reflected in the N3 method's more aggressive correction of white matter.

At first glance, this would seem to ensure that the field estimate will be biased by the anatomy. However, recall that N3 is an iterative optimization, not a single estimate. By iterating towards a field estimate that is both smooth and maximizes the high frequencies in the histogram, the fact that certain types of tissue are more powerful estimators of non-uniformity becomes less important. While the field estimate near structures containing these more powerful estimators will converge more rapidly, the field estimate in other regions can continue to improve after these regions have converged. This trend is apparent in the first ten iterations of the correction of the T1 volume shown in Figure 4.4 where the field estimate grows more rapidly in white matter regions than in grey matter or CSF.

A potential difficulty with iterative optimization methods such as N3 is that one is never certain whether the solution found is the global or a local minimum in the objective function. While it cannot be proven exhaustively that the method converges to a global minimum, the fact that the method has always been observed to converge

[1]Note that resolved peaks corresponding to isointense classes of tissue are not necessary for the proper functioning of N3. This is merely an illustrative example.

to the same solution regardless of the initial field suggests that the optimization is not prone to local minima. This observation has been confirmed with random fields, other kinds of simulated data, and real data.

To summarize, the N3 method makes the assumptions that non-uniformity can be modelled as a smooth multiplicative field and that the image intensities are independent identically distributed random variables. While neither assumption is strictly valid for MR data, the N3 method is tolerant of data that violates these assumptions, suffering a degradation in performance rather than failure. The strongest argument in support of this is the experimental evidence.

8.2 Experimental results

Correcting for non-uniformity in a random field is useful to understand the behaviour of N3. Aside from the trivial illustration that the method is able to correct volumes lacking contiguous regions of homogeneous tissue, the experiment described in Section 5.1 is important as an illustration of the benefits of using a statistical approach. A random field is fair data upon which to compare N3 to a filter and divide approach. Since there is no structure in the volume, the usual criticism of filter and divide methods, that the spatial frequencies of the anatomy overlap with those of the non-uniformity field, does not apply. Despite this fact, the performance of the N3 method was significantly better than direct filtering since it takes advantage of structure in the histogram of the data.

Simulated data was used to validate the N3 method under a variety of conditions. These experiments show that it is able to correct T1, T2, and PD weighted scans for a range of noise levels and a number of non-uniformity fields. Furthermore, all of these results were obtained without adjustment of the algorithm. In typical trials non-uniformity was reduced to 25%–50% of its previous level. It should be noted that this is a conservative estimate since it does not take into account the non-uniformity implicit in the noise and natural variability of tissue. Since in practice the goal of correction is to produce a uniform volume, the performance of N3 is even better since smooth variations due to noise, natural tissue variability, and true non-uniformity are all reduced. Experiments measuring this combined variability using real data, while not quantitative, showed significant reductions in non-uniformity using N3.

To put these results in perspective the benefits of correcting for non-uniformity were investigated for three different anatomical mapping applications: tissue classification, cortical surface extraction, and non-linear co-registration. N3 created substantial improvements in accuracy on volumes with typical levels of non-uniformity. Furthermore, the performance of these techniques became largely independent of the severity of the non-uniformity when used in conjunction with the N3 method.

8.3 Implementation issues

For the practical use of the N3 method one needs to consider how to choose the various parameters. In particular, one needs to choose the smoothing parameter ω, the basis function distance d, the initial field estimate, the stopping threshold e, the working resolution, and the two parameters controlling the deconvolution kernel, FWHM and Z. Fortunately, there are criteria for choosing these parameters.

As described previously, since both ω and d control the smoothness of the field estimate, one of these parameters can be fixed arbitrarily. Throughout the analysis, ω has been fixed at 1.0. The basis function distance d is then used to control the smoothness of the estimated field, with larger d corresponding to smoother fields. Experiments on simulated data showed performance to have little dependence on this parameter (see Figure 5.7) and it was fixed at 200 mm.

Experiments with simulated data have also shown performance to have little dependence on the choice of the deconvolution parameter Z. It has been fixed at 0.1. The choice of the second deconvolution parameter FWHM involves a compromise. The results shown in Figure 5.6 suggests that performance generally improves as the FWHM parameter is reduced; however, the rate of convergence drops rapidly. Choosing this parameter involves a tradeoff between accuracy and computation time and should take into account the working resolution and stopping condition, both of which affect computation time. Throughout this analysis the FWHM was fixed at 0.15.

As mentioned earlier, the N3 method has been found to converge to the same solution irrespective of the initial field estimate. For simplicity, the initial estimate is taken to be zero field. Another consideration is the choice of working resolution. The results of Appendix C suggest that substantial reductions in computation time can be gained with little or no reduction in accuracy by subsampling a volume to a coarser resolution. In light of this data, it appears that the decision to use 3mm isotropic sampling as the working resolution throughout this analysis was conservative.

Another consideration is the choice of the threshold e used to stop iteration. The idea behind measuring the difference between subsequent field estimates is that this measure will become small when the method has converged. As shown by Figure 5.2, the convergence of the algorithm is regular and predictable. Hence, the parameter e should be chosen based on the convergence rate for a given FWHM. In summary, criteria are available for choosing all of the parameters needed by the N3 method. The choice of these various parameters control the smoothness of the field estimate and the tradeoff between accuracy and computation time; they do not determine the success or failure of the algorithm.

8.4 Comparison with other methods

To put the strengths and weaknesses of the N3 method in perspective it is worthwhile to compare it to other methods. While the results of the experiments with simulated data described in Chapter 6 suggest that no one method is best in all cases, the relative performance of the methods in each case can be rationalized from their design. Consider the WM method which relies on segmenting the white matter from the rest of the brain. On T1 weighted volumes, where the contrast between white matter and other tissues is greatest, the white matter method gave the best performance. On PD weighted volumes where contrast is low, the performance of the white matter method is relatively poor.

Unlike the white matter method, the contrast seen in T1 scans works against the

EM method. In principle the EM method has an advantage over the WM method in that, by using the whole volume, it is less affected by noise. However, a number of authors have described this method as overly sensitive to the training of its classifier [5, 20, 31]. The author's experience has been that the method makes excessively large corrections to voxels that fall outside the classifier's tissue model. These outliers may dominate the correction process and cause poor results. If one trains the classifier on samples of pure tissue and the intensity distributions of the different classes are well resolved, as is the case in T1 scans, then regions of partial volume will produce intensities outside the tissue model and receive extreme corrections. This is reflected in the data by the fact that the method performed worst given an ideal training set consisting entirely of pure tissue (see Figure 6.1). In addition, in some cases as the level of noise increased, the EM method performed better (see Figure 6.1a). This is due to the greater variability in the training data and reduced contrast in these images. Similarly, on PD weighted scans, with less contrast, the method performed better. Overall, the performance of this method was erratic.

Compared to the EM method, the REM method performed substantially better, suggesting that the use of probability priors and spline smoothing significantly improves overall performance. Like the EM method, REM performed better on PD than T1 weighted scans.

The non-parametric method also gave consistently good performance on all three modalities for the tests shown in Figure 6.1. Although REM performed better than N3 on simulated T2 and PD data, what distinguished the N3 method from the REM method is the results for real data. While real data is unable to provide quantitative results, it uncovered limitations of the methods caused by factors not reflected in the simulations. As discussed in Section 6.3, the EM and REM methods proved unexpectedly sensitive to the choice of brain mask. Even when this masking problem was fixed, the REM method performed poorly on real T1 data, tending to make non-uniformity worse. While it is not clear what caused this failure, it is likely that some aspect of the data violates the sophisticated model employed by the REM method. In contrast, the N3 method showed no such sensitivity and required no special precautions in moving from simulated to real data.

The comparison of the four methods on a sagittal slice of a T1 weighted scan in Figure 6.2 illustrates the typical behaviour of the four methods. Both the EM and REM methods make their greatest correction at the top of the head. Diminished intensity at the top of the head is often seen in gradient echo T1 weighted scans. However, both of these methods have made the problem worse. This is likely due to a poor initial classification of this region. It should be noted that the field estimated by the EM method has higher frequency content than that of the other methods due the use of the truncated filter.

While the fields estimated by the WM and N3 methods are similar, the WM method makes its greatest correction at the front edge of the brain while the N3 method corrects most strongly the region just forward of the middle of the brain. This latter behaviour is typical of the N3 method and may be due to a breakdown of the stationarity assumption discussed in Section 8.1.

Another difference between the N3 method and the other methods is that it tends to raise the intensity of the cerebellum. This difference is likely due to the mixing of grey and white matter in this region. While N3 is designed to correct randomly mixed fields of tissue, the other methods may have difficultly finding pure tissue on which to base their correction.

8.5 Conclusions

An iterative method for correction of intensity non-uniformity in MR volumes has been described that avoids some of the restrictive model assumptions that plague other methods. In particular, this Non-parametric intensity Non-uniformity Normalization (N3) method does not require a model of the tissue intensities in terms of discrete tissue classes, nor does it rely on a segmentation of the volume into homogeneous regions. Instead, a non-parametric model of the tissue intensities is derived directly from the data. The behaviour of N3 is controlled by two parameters: one controlling the smoothness of the estimated non-uniformity, the other controlling the tradeoff between convergence rate and accuracy. Experiments with simulated data have shown that both of these parameters can be chosen to provide uniform performance independent of pulse sequence, subject, and field shape. This is a considerable advantage in automated data analysis as the method can be applied at an early stage, without prior knowledge of the data. For instance, no special precautions need to be taken for pathological data that might otherwise violate tissue model assumptions.

A wide variety of experiments with simulated MR volumes show that by conservative estimates N3 reduces non-uniformity from a typical level of 20% to a level of 5%–10%. Furthermore, real data corrected by the method is visually uniform and shows a statistically significant reduction in tissue intensity variation. These improvements in uniformity were shown to translate into substantial improvements in the accuracy of three different anatomical mapping techniques. Robust, fully automatic, and requiring little domain specific knowledge, N3 is attractive as a preprocessing step for a variety of MRI analysis applications.

8.6 Future Work

While this work establishes the core ideas of the N3 method, there are a number avenues that remain unexplored. The first of these is the optimization of accuracy and convergence rate. While one could further explore the relationship between the stopping condition, deconvolution kernel FWHM, and working resolution in terms of accuracy and computation time, a more interesting investigation is to vary the FWHM with iterations. If the final accuracy is determined by the FWHM of the final deconvolution kernel, then one can envision a scheme in which the width of the deconvolution kernel starts large to increase the initial convergence rate but ends small to improve final accuracy.

Another aspect of N3 that remains to be explored is the choice of the kernel function for deconvolution. In fact, the whole process of proposing new intensity distributions could be replaced by a different heuristic such as one that proposes intensity distributions with higher entropy rather than enhanced high frequency content. Any scheme that could propose incrementally sharper histograms based on the measured histogram would be suitable.

A third avenue of investigation is to apply the method to other types of data. While all of the analysis described here has focused on correction of volumetric brain scans, there are no anatomy specific assumptions in N3. Any MR image used for quantitative analysis would likely benefit from reduced intensity non-uniformity. Furthermore, any imaging modality that suffers from a smooth multiplicative or, with some modification, additive bias field could benefit from correction using N3.

Appendix A

Spline smoothing

The tensor cubic B spline approximation of a function is given by

$$p(\mathbf{x}) = \sum_{i=1}^{M_x} \sum_{j=1}^{M_y} \sum_{k=1}^{M_z} \theta_{ijk} B_i(x) B_j(y) B_k(z) \tag{A.1}$$

where $\mathbf{x} = [x\, y\, z]$ and B is a one dimensional cubic B spline in the variable x, y, or z. For example, the B spline for x is given by

$$B_i(x) = \sum_{s=0}^{4} \frac{(-1)^s}{d^3} \binom{4}{s} \left(x - \lambda_{i-s}^{(x)}\right)^3 \mu\left(x - \lambda_{i-s}^{(x)}\right) \tag{A.2}$$

$$\mu(x) = \begin{cases} x & : \quad x > 0 \\ 0 & : \quad \text{elsewhere} \end{cases} \tag{A.3}$$

where $\lambda_i^{(x)}$ is referred to as a knot location and d is the distance between knots. A spline $B_i(x)$ only takes non-zero values on the interval $[\lambda_{i-4}^{(x)}, \lambda_i^{(x)}]$.

The B spline coefficients θ_{ijk} for least squares approximation of a set of data are found by minimizing as follows

$$\min_{\theta} E(\theta) + \omega R(\theta) \tag{A.4}$$

where E reflects the closeness of fit to the data and R reflects the roughness of the approximate function. Choosing the parameter ω determines the tradeoff between the two.

Suppose there are N measurements P_n at locations $\mathbf{x}_n$. Then E and R are as follows

$$E(\theta) = \frac{1}{N}\sum_{n=1}^{N}(P_n - p(\mathbf{x}_n))^2 \tag{A.5}$$

$$R(\theta) = \frac{1}{V}\int_D \sum_{i=1}^{3}\sum_{j=1}^{3}\left[\frac{\partial^2 p(\mathbf{x})}{\partial x_i \partial x_j}\right]^2 d\mathbf{x} \tag{A.6}$$

where D contains the region of interest and V is the volume of D.

In matrix form, the solution of equation (A.4) for θ is given by

$$\theta = (A^T A + \omega J)^{-1} A^T Z \tag{A.7}$$

where

$$
\begin{aligned}
A_{n,ijk} &= \frac{\partial z(\mathbf{x}_n)}{\partial \theta_{ijk}} = B_i(x_n)B_j(y_n)B_k(z_n) \;\; \text{is an element of an } N \text{ by } Q \text{ matrix.}\\
Q &= M_x M_y M_z \\
J &= \sum_{\substack{\alpha_x,\alpha_y,\alpha_z \geq 0 \\ \alpha_x+\alpha_y+\alpha_z=2}} \frac{2}{\alpha_x!\,\alpha_y!\alpha_z!} J_x^{(\alpha_x)} \otimes J_y^{(\alpha_y)} \otimes J_z^{(\alpha_z)} \;\; \text{is a } Q \text{ by } Q \text{ matrix.}\\
J_{x\ i,j}^{(l)} &= \frac{1}{V}\int_D B_i^{(l)}(x)B_j^{(l)}(x)dx \;\; \text{is an element of an } M_x \text{ by } M_x \text{ matrix.}
\end{aligned}
$$

The symbol $\otimes$ denotes Kronecker product. θ and Z are column vectors with elements in their natural order. D is most conveniently taken as

$$D = [\lambda_0^{(x)}, \lambda_{M_x-3}^{(x)}] \times [\lambda_0^{(y)}, \lambda_{M_y-3}^{(y)}] \times [\lambda_0^{(z)}, \lambda_{M_z-3}^{(z)}] \tag{A.8}$$

although this may be larger than the region of interest.

Equation (A.7) may not be solvable directly if the smoothing parameter is very small or zero since the matrix $A^T A$ is often ill conditioned. Hayes and Halliday [24] have proposed a technique using Householder transformations to stabilize the method. However, in practice ω can be chosen large enough to avoid numerical difficulties.

Appendix B

Sum of Random Variables

The idea that adding a field $\hat{f}$ to the log intensity image $\hat{u}$ has the effect of blurring the intensity distribution U is central to the N3 method. To clarify this relationship, the mathematics of adding independent random variables are described here. For further details, consult [13].

Suppose $\hat{f}$ and $\hat{u}$ are independent random variable with probability distributions given by $F(\hat{f})$ and $U(\hat{u})$. $\quad \hat{f} \in \Re \quad \hat{u} \in \Re$

The distribution $V(\hat{v})$ where $\hat{v} = \hat{f} + \hat{u}$ can be computed as follows. The joint probability distribution of $\hat{v}$ and $\hat{u}$ is given by

$$P(\hat{v}, \hat{u}) = F(\hat{v} - \hat{u})U(\hat{u}) \left| \frac{\partial(\hat{v}, \hat{u})}{\partial(\hat{f}, \hat{u})} \right| \tag{B.1}$$

where the term $\frac{\partial(\hat{v},\hat{u})}{\partial(\hat{f},\hat{u})}$ is the Jacobian factor for the change of variables from $(\hat{f}, \hat{u})$ to $(\hat{v}, \hat{u})$.

$$\left| \frac{\partial(\hat{v}, \hat{u})}{\partial(\hat{f}, \hat{u})} \right| = \begin{vmatrix} \frac{\partial \hat{v}}{\partial \hat{f}} & \frac{\partial \hat{v}}{\partial \hat{u}} \\ \frac{\partial \hat{u}}{\partial \hat{f}} & \frac{\partial \hat{u}}{\partial \hat{u}} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1 \tag{B.2}$$

Hence

$$P(\hat{v}, \hat{u}) = F(\hat{v} - \hat{u})U(\hat{u}) \tag{B.3}$$

The distribution $V(\hat{v})$ is given by the marginal distribution of $P(\hat{v}, \hat{u})$ as follows

$$V(\hat{v}) = \int_{-\infty}^{\infty} P(\hat{v}, \hat{u}) d\hat{u} = \int_{-\infty}^{\infty} F(\hat{v} - \hat{u})U(\hat{u}) d\hat{u} \tag{B.4}$$

which is the same as equation (4.3) on page 18 and corresponds to the convolution of F and U.

Appendix C

Working resolution

Simulated T1 weighted MR volumes, described in Section 5.2, were used to investigate the connection between working resolution, correction accuracy, and computation time. For these experiments, four realizations of the pseudo-anatomy were tested for both of the non-uniformity fields shown in Figure 5.5. The estimation error and computation time for these experiments are plotted in Figure C.1. These results show no significant difference in estimation error for the range of working resolutions shown. Although the data point in Figure C.1a for 1 mm working resolution is not available, other experiments not shown here suggest no improvement in accuracy at 1 mm resolution.

As shown by Figure C.1b, there is a substantial decrease in computation time as the voxel size of the working resolution increases. This trend reflects the fact that at larger working resolutions there are fewer voxels to process. It should noted that data is converted to the coarser resolution by subsampling without filtering, effectively throwing away some fraction of the data. The fact that so much of the data can be removed with little impact on correction accuracy suggests that the problem of estimating non-uniformity is highly overdetermined.

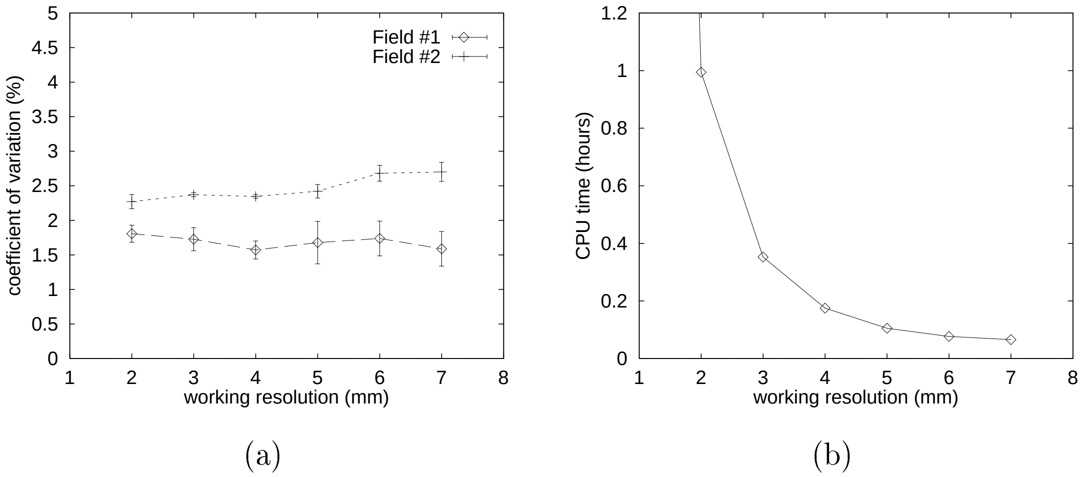


Figure C.1: Accuracy and computation time versus working resolution. (a) Estimation error versus working resolution. (b) Computation time versus working resolution. CPU times are for an SGI workstation with floating point performance of 99 SPECfp92.

Appendix D

Glossary

artificial neural network classifier An artificial neural network classifier (ANN) employs a non-linear function formed from simple terms, corresponding to nodes in a network, to label measurements as belonging to a particular class [61]. The coefficients of this function are derived using a set of correctly labelled examples. See classifier.

$\mathbf{B_0}$ The notation B_0 denotes the static magnetic field produced by an MR scanner.

$\mathbf{B_1}$ In MR, the notation B_1 denotes the magnetic field, oscillating in the radio frequency range, used to induce magnetic resonance in the subject.

B splines B splines are smooth piecewise polynomial functions of limited extent that can be used as a basis for constructing smooth functions. See Appendix A.

bandwidth filter In MR, a filter, referred to as a bandwidth filter, is used to select the range of frequencies present in the measured signal. Since spatial location is encoded as frequency, frequency dependent variations in the sensitivity of this filter lead to intensity variations in the frequency encoding direction of the resulting image.

basis function A set of functions that are linearly independent, in the sense that no member can be constructed from a linear combination of the others, is termed a basis and the individual functions are called basis functions. See B spline and Appendix A.

Bayesian classifier The restricted definition of a Bayesian classifier used for this work is that each tissue class is modelled as having normally distributed intensities of given mean and covariance. The classifier labels a measurement as belonging to the class having maximum aposterior probability (MAP) [14, 29].

bias field For the purposes of this work, the smooth multiplicative field associated with intensity non-uniformity is termed the bias field.

Biot-Savart law The Biot-Savart law is an expression for the magnetic field at a point expressed as an integral of current density over all space. This expression is valid when the distribution of current density is static.

bird cage coil A kind of antenna for transmitting and receiving RF pulses, the bird cage coil is often used for volumetric MR scans of the head. It is formed from two loops connected by a number of spokes and produces a circularly polarized field oriented perpendicular to the coil's axis of symmetry. See [53].

Bloch equations The Bloch equations [3] are a set of coupled differential equations that empirically describe the time evolution of the net magnetization of a subject in an MR scanner given knowledge the magnetic fields present.

body coil The large antenna fixed permanently in an MR scanner and suitable for whole body scans is called the body coil.

broad-band noise Noise having a frequency spectrum with a wide range of frequencies is said to be broad-band.

cerebellum The cerebellum is a large dorsally projecting brain structure with thin layers of gray and white matter barely resolvable at the resolution of a typical MR scan.

cerebrospinal fluid Cerebrospinal fluid is a substance similar to blood plasma found surrounding the brain and within the ventricles.

circularly polarized coil A circularly polarized coil produces a magnetic field of constant magnitude but rotating orientation. Such a coil can be formed from two linearly polarized coils oriented at right angles and driven in quadrature.

classifier A tool for labelling each voxel as belonging to a particular class on the basis on one or more features associated with that voxel is said to be a classifier. This definition can be relaxed to allow fuzzy or partial membership in a class.

co-registration The task of determining the transformation that relates measurements in one coordinate to measurements of the same object in another is that of co-registration. This definition can be relaxed to include objects that are not the same, but merely similar.

compactly-supported spline A compactly supported spline is only non-zero in a finite region. The matrices associated with solving the linear problem of approximating a function in the least squares sense by splines having this property are sparse.

cortical surface The cortical surface is the outermost surface of the brain.

cortical surface extraction Cortical surface extraction is a brain mapping technique for identifying the cortical surface in an MR scan.

echo The reappearance of the MR signal after the initial signal caused by excitation has died away is referred to as an echo. See [21].

echo time Denoted TE, the echo time is the time after an initial excitation at which the echo signal occurs. The echo time of a pulse sequence is one of the factors determining the weighting or contrast in the resulting MR image.

eddy currents Eddy currents are electrical currents induced in an object by transient magnetic fields.

entropy Entropy is a measure of disorder. For random variables, lower entropy corresponds to a more tightly clustered probability distribution.

excitation field The excitation field in an MR scanner is the oscillating radio-frequency magnetic field used to induce magnetic resonance in the subject.

field gradient In MR, spatial variations in the magnitude of the main magnetic field B_0 are used to encode spatial location. These linear variations are referred to as field gradients.

finite element analysis Finite element analysis is a numerical method suitable for solving partial differential equations by approximating the solution in terms of a number of discrete elements.

flip angle The angle by which the net magnetization rotates in response to an RF excitation pulse is the pulse's flip angle. More precisely, this angle is a nutation of the net magnetization as it precesses about the direction of the main magnetic field B_0.

frequency encoding direction In a spin warp imaging sequence, the direction in which spatial location is encoded by resonant frequency is the the frequency encoding direction.

geometric distortion Geometric distortion is displacement of voxels in the image that causes the geometry of the subject to be misrepresented.

gradient echo pulse sequence An imaging sequence which produces an echo by the switching of field gradients alone without the aid of a second RF excitation is referred to as a gradient echo pulse sequence.

grey matter Grey matter is brain tissue characterized by unmyelinated neurons and a greyish appearance upon dissection.

head coil An antenna specifically for imaging the human head is a head coil. Most head coils are of bird cage design.

imaging sequence See pulse sequence.

independent random variables Random variables are said to be independent if the values they take on are unrelated and if any correspondence between them is purely coincidental.

intensity non-uniformity Intensity non-uniformity is the variation in image intensity caused by the imaging process rather than true variation in the subject. Distinct from noise, intensity non-uniformity is usually smoothly varying and not random.

kappa Kappa is a measure of similarity, corrected for chance [7, 30]. A kappa of unity indicates perfect agreement, while a kappa of zero indicates that any agreement is purely coincidental.

kernel In signal processing terminology, a kernel is a function describing the response of a filter to an impulse at its input.

LCJ segmentation An acronym for Liou, Chiu, and Jain, LCJ segmentation [33] is an unsupervised technique for identifying homogeneous regions within an image.

linearly polarized magnetic coil A linearly polarized coil produces an oscillating magnetic field oriented in a particular direction. See [19].

magnetic relaxation In MR, magnetic relaxation is the process by which the signal emitted following excitation decays away and the net magnetization of the subject returns to its equilibrium state.

minimum distance classifier A minimum distance classifier labels a given measurement as belonging to a particular class on the basis of which class has the prototypical example to which the measurement is closest [2, 30].

multiple sclerosis Multiple sclerosis is a degenerative brain disease. The associated brain lesions can be observed with MRI.

multi-slice sequence A multi-slice sequence is a pulse sequence in which the slices that together form the image volume are acquired one at a time.

net magnetization For MR imaging, the net magnetization refers to the magnetic field produced by the hydrogen atoms within the subject. The phenomenon of magnetic resonance occurs when the net magnetization is rotated away from its equilibrium orientation, parallel to the main magnetic field (B_0).

non-parametric method Parametric methods use a limited number of parameters to specify a model whose form is chosen to fit the problem at hand. In contrast, non-parametric methods use a large number of parameters to specify a model whose form is general to a broad class of problems.

partial volume effect The process of image formation in MR is such that voxels that are partially filled with tissue have a proportionally smaller signal strength. When this process leads to voxels of intermediate intensity at the boundary between regions of homogeneous material, it is referred to as a partial volume effect.

Parzen window A method of estimating the distribution of a random variable from a set of measurements is to interpret each measurement as a distribution of measurements clustered around the actual measurement. This latter distribution is termed the Parzen window [14]. Distributions estimated using a Parzen window tend to be more accurate than traditional approaches when the number of measurements is small.

penetration effect When a time varying electromagnetic field is attenuated or prevented from entering a material by the currents it induces in that material, the phenomenon is termed a penetration effect.

permittivity Permittivity is a measure of a material's ability to store electrical potential energy.

phantom A phantom is an object used as a subject of an MR scan for the purposes of calibration or testing.

pickup coil In analyzing an RF coil for use in MR, an antenna small enough to have little effect on the coil being analyzed, yet suitable to detect the magnetic field fluctuations produced by that coil is referred to as a pickup coil.

proton density weighted An MR image is said to be proton density weighted if the image contrast is largely determined by the density of hydrogen atoms present.

pulse sequence A particular orchestration of RF pulses, switched field gradients, and data acquisition periods used to produce an image is referred to as a pulse sequence.

quasi-static A system in which electromagnetic fields are changing slowly enough that magnetostatic equations (i.e. Ampere's law and the Biot-Savart law) apply is said to be quasi-static.

reception coil The antenna used in MR to measure the radio waves emitted by the subject is a reception coil.

reception sensitivity The ratio of the measured signal to that which is emitted by the subject is the reception sensitivity. Reception sensitivity that is spatially varying contributes to intensity non-uniformity in the resulting image.

rejection sampling Rejection sampling is a technique for generating random samples with an arbitrary probability distribution. The method uses the probability density of the desired distribution to randomly reject samples from a uniform distribution. See [13].

relative permittivity The relative permittivity of a material is the ratio of its permittivity to that of free space. Permittivity is a measure of a material's ability to store electrical potential energy.

relaxation See magnetic relaxation.

repetition time The period of the repeating cycle of excitation followed by data acquisition during a pulse sequence is the repetition time (TR). The repetition time of a pulse sequence determines the T1 weighting of the resulting image.

resample Resampling is an image processing technique for transforming a sampled representation of an image to another representation sampled at different locations but corresponding to the same image.

resonant frequency The frequency, proportional to the strength of the static magnetic field B_0, at which net magnetization precesses is the resonant frequency.

Rician distribution A Rician distribution is formed by computing the magnitude of the sum of a complex Gaussian distributed random variable and a constant.

saddle coil A saddle coil is a pair of saddle shaped loop antennas used to produce or receive a linearly polarized RF field.

scan A scan is a volume of data acquired in a single session with an MRI machine.

scanner In this context, a scanner refers to an MRI machine, a device capable of volumetric imaging of the anatomy.

segmentation The set of labels identifying the class that each voxel in a volume belongs to is the segmentation of that volume.

sensitivity See reception sensitivity.

sharpening Sharpening is a signal processing technique that enhances the high frequency components of a signal.

spatial frequency An image can be decomposed in terms of sinusoidal basis functions. The number periods a basis function has within a given distance is its spatial frequency. Low spatial frequencies are associated with smooth image feature, while high frequencies are associated with abrupt changes and edges.

SPECfp92 SPECfp92 is a measure of a computer's floating point performance developed by the Standard Performance Evaluation Corporation (SPEC).

spectral estimation The task of estimating a signal's power within a given frequency range is that of spectral estimation.

spin echo pulse sequence A spin echo sequence is a pulse sequence that employs a 90^o and a 180^o RF pulse to produce an signal echo, during which data is acquired.

spin warp acquisition The process of MR data acquisition can be viewed as sampling the Fourier transform of the image. A spin warp acquisition sequence samples this image on a grid, one row at a time. The rows and columns of this grid correspond to the frequency encoding and phase encoding directions respectively.

standing wave effect A standing wave oscillates but does not travel. In MR, the radio waves reflected within a subject may interfere constructively and appear not to travel. This enhancement of B_1 field strength is referred to as a standing wave effect.

stationarity An image is said to be stationary if the statistics describing each element and relating its neighbours do not change with position in the image.

stereotaxic space Stereotaxic space is a coordinate system in which affine differences among individual brains have been removed so that any neuroanatomical location can be addressed by a Cartesian coordinate [51, 9].

structuring element In binary morphology, the structuring element is a binary mask that defines which voxels are considered neighbours of a given voxel.

sulcus A sulcus is an inward fold of the cortical surface.

surface coil A flat antenna placed on the surface of the subject is a surface coil. Surface coils have high sensitivity that drops off rapidly with distance from the coil.

$\mathbf{T_1}$ weighted An image is said to be T_1 weighted if its contrast is largely due to differences in the intrinsic tissue property T_1. T_1 is the time constant describing the rate at which the net magnetization returns to equilibrium after excitation. This process is referred to as spin-lattice relaxation.

$\mathbf{T_2}$ weighted T_2 weighting indicates that the MR image contrast is largely dependent on the intrinsic tissue property T_2, a property describing the rate of so called spin-spin relaxation.

tessellate A surface is tessellated by approximating it in terms of simple geometric surface elements.

tissue classification Tissue classification is the task of labelling each voxel by tissue type. See **classifier**.

tissue probability map Maps showing the probability that each location in stereotaxic space has a given tissue type are call tissue probability maps (TPM).

training set A set of correctly labelled example measurements from which a classifier learns its behaviour is a training set.

tricubic interpolation Tricubic interpolation is a method employing cubic splines to estimate the value of a function on the basis of nearby measurements taken in three dimensions.

unloaded coil An MR scanner's RF coil is said to be unloaded if there is no subject within it. Viewed as an electrical circuit the RF coil is magnetically coupled to the subject, which presents an impedance during excitation. Since, this impedance absorbs energy it is said to be a load on the electrical circuit, hence the terminology of loaded and unloaded coils.

volumetric acquisition While any imaging sequence in which a 3D volume of data is acquired is technically a volumetric acquisition, the term is used here to refer to scans in which the sensitivity is relatively uniform throughout the volume, such as with head or body coils.

voxel Analogous to a pixel, a voxel is an element of a 3D or volumetric data set.

warping Resampling an image to produce a possibly non-linear spatial distortion is called warping.

white matter White matter is brain tissue characterized by myelinated nerve fibers and a whitish color upon dissection.

Wilcoxon signed rank test The Wilcoxon signed rank test is a non-parametric statistical test to determine if two sets of observation come from the same or different populations.

Bibliography

[1] L. Axel, J. Costantini, and J. Listerud. Intensity correction in surface-coil MR imaging. *American Journal of Roentgenology*, 148:418–420, Feb. 1987.

[2] J. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms.* Plenum, New York, 1981.

[3] F. Bloch, W. Hanson, and M. Packard. Nuclear induction. *Physical Review*, 69:127, Feb. 1946.

[4] P. A. Bottomley and E. R. Andrew. RF magnetic field penetration, phase shift and power dissipation in biological tissue: implications for NMR imaging. *Physics in Medicine and Biology*, 23(4):630–43, Jul 1978.

[5] C. Brechbuhler, G. Gerig, and G. Szekely. Compensation of spatial inhomogeneity in MRI based on a parametric bias estimate. In *Forth International Conference on Visualization in Biomedical Computing*, pages 141–146, 1996.

[6] H. S. Choi, D. R. Haynor, and Y. Kim. Partial volume tissue classification of multichannel magnetic resonance images - a mixel model. *IEEE Transactions on Medical Imaging*, 10(3):395–407, Sept. 1991.

[7] J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological measurements*, 20:37–46, 1960.

[8] D. Collins, C. Holmes, T. Peters, and A. Evans. Automatic 3D model-based neuroanatomical segmentation. *Human Brain Mapping*, 3(3):190–208, 1996.

[9] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, 18(2):192–205, 1994.

[10] B. R. Condon, J. Patterson, D. Wyper, et al. Image non-uniformity in magnetic resonance imaging: its magnitude and methods for its correction. *Br. J. Radiology*, 60:83–87, 1987.

[11] B. M. Dawant, A. P. Zijdenbos, and R. A. Margolin. Correction of intensity variations in MR images for computer-aided tissue classification. *IEEE Transactions on Medical Imaging*, 12(4):770–781, Dec. 1993.

[12] C. DeCarli, D. G. M. Murphy, D. Teichberg, G. Campbell, and G. S. Sobering. Local histogram correction of MRI spatially dependent image pixel intensity nonuniformity. *Journal of Magnetic Resonance Imaging*, 6(3):519–528, 1996.

[13] L. Devroye. *Non-Uniform Random Variate Generation.* Prentice Hall, 1986.

[14] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis.* John Wiley, 1973.

[15] A. C. Evans, J. A. Frank, J. Antel, and D. H. Miller. The role of MRI in clinical trials of multiple sclerosis: Comparison of image processing techniques. *Annals of Neurology*, 41(1):125–132, Jan. 1997.

[16] A. C. Evans, S. Marrett, P. Neelin, et al. Anatomical mapping of functional activation in stereotactic coordinate space. *NeuroImage*, 1:43–53, 1992.

[17] D. T. Glendon, L. P. Clarke, R. P. Velthuizen, and M. L. Silbiger. MR image non-uniformity correction techniques. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 13, pages 0089–0090, 1991.

[18] D. T. Glennon, L. P. Clarke, R. P. Velthuizen, and M. L. Silbiger. MR image non-uniformity correction techniques. In *Proceedings of the 13th Annual Meeting of the IEEE Engineering in Medicine and Biology Society (EMBS)*, pages 89–90, Orlando, Florida, Oct./Nov. 1991.

[19] G. Glover, C. Hayes, N.J.Pelc, W.A.Edelstein, O.M.Mueller, H. Hart, C. Hardy, M. D'Donnel, and W. Barber. Comparison of linear and circular polarization for magnetic resonance imaging. *Journal of Magnetic Resonance*, 64:255–270, 1985.

[20] R. Guillemaud and M. Brady. Enhancement of MR images. In *Forth International Conference on Visualization in Biomedical Computing*, pages 107–116, 1996.

[21] E. Hahn. Spin echoes. *Physical Review*, 80(4):580–594, Nov. 1950.

[22] I. Harvey, P. S. Tofts, J. K. Morris, et al. Sources of T_1 variance in normal human white matter. *Magnetic Resonance Imaging*, 9(1):53–59, 1991.

[23] J. Haselgrove and M. Prammer. An algorithm for compensation of surface-coil images for sensitivity of the surface coil. *Magnetic Resonance Imaging*, 4(6):469–472, 1986.

[24] J. G. Hayes and J. Halliday. The least-squares fitting of cubic spline surfaces to general data sets. *Inst. Maths Applics*, 14:89–103, 1974.

[25] C. J. Holmes, R. Hoge, L. Collins, and A. C. Evans. Enhancement of T1 MR images using registration for signal averaging. In *Second International Conference on Functional Mapping of the Human Brain*, page S28, 1996.

[26] C. J. Holmes, R. Hoge, R. P. Woods, A. C. Evans, and A. W. Toga. Enhancement of T2 and proton density MR images using registration for signal averaging. In *Second International Conference on Functional Mapping of the Human Brain*, page S28, 1996.

[27] B. Johnston, M. S. Atkins, B. Mackiewich, and M. Anderson. Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI. *IEEE Transactions on Medical Imaging*, 15(2):154–169, Apr. 1996.

[28] M. Kamber, R. Shinghal, D. L. Collins, et al. Model-based 3D segmentation of multiple sclerosis lesions in magnetic resonance brain images. *Pattern Recognition*, 1993. (submitted).

[29] M. Kamber, R. Shinghal, D. L. Collins, G. S. Francis, and A. C. Evans. Model-based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images. *IEEE Transactions in Medical Imaging*, 14(3):442–453, Sept. 1995.

[30] V. Kollokian. Performance analysis of automatic techniques for tissue classification in magnetic resonance images of the human brain. Master's thesis, Concordia University, Montreal, QC, Nov. 1996.

[31] S. K. Lee and M. W. Vannier. Post-acquisition correction of MR inhomogeneities. *Magnetic Resonance in Medicine*, 36:275–286, 1996.

[32] K. O. Lim and A. Pfefferbaum. Segmentation of MR brain images into cerebrospinal fluid spaces, white and gray matter. *Journal of Computer Assisted Tomography*, 13(4):588–593, July/Aug. 1989.

[33] S.-P. Liou and R. C. Jain. An approach to three-dimensional image segmentation. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 53(3):237–252, 1991.

[34] D. MacDonald, D. Avis, and A. Evans. Multiple surface identification and matching in magnetic resonance images. In *Visualization in Biomedical Computing 1994, Proceedings of the SPIE*, volume 2359, pages 160–169, 1994.

[35] P. Mansfield and P. G. Morris. *NMR Imaging in Biomedicine*, pages 181–191. Academic Press, New York, 1982. Advances in Magnetic Resonance, Suppl. 2.

[36] E. R. McVeigh, M. J. Bronskill, and R. M. Henkelman. Phase and sensitivity of receiver coils in magnetic resonance imaging. *Medical Physics*, 13(6):806–814, Nov./Dec. 1986.

[37] C. R. Meyer, P. H. Bland, and J. Pipe. Retrospective correction of intensity inhomogeneities in MRI. *IEEE Transactions on Medical Imaging*, 14(1):36–41, Mar. 1995.

[38] S. E. Moyher, D. B. Vigneron, and S. J. Nelson. Surface coil MR imaging of the human brain with an analytic reception profile correction. *Journal of Magnetic Resonance Imaging*, 5(2):139–144, Mar./Apr. 1995.

[39] J. W. Murakami, C. E. Hayes, and E. Weinberger. Intensity correction of phased-array surface coil images. *Magnetic Resonance in Medicine*, 35:585–590, 1996.

[40] P. A. Narayana and A. Borthakur. Effect of radio frequency inhomogeneity correction on the reproducibility of intra-cranial volumes using MR image data. *Magnetic Resonance in Medicine*, 33(3):396–400, Mar. 1995.

[41] P. A. Narayana, W. W. Brey, M. V. Kulkarni, and C. L. Sievenpiper. Compensation for surface coil sensitivity variation in magnetic resonance imaging. *Magnetic Resonance Imaging*, 6(3):271–274, 1988.

[42] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Biomedical Engineering*, 9:63–66, 1979.

[43] M. Özkan, B. M. Dawant, and R. J. Maciunas. Neural-network-based segmentation of multimodal medical images: A comparative and prospective study. *IEEE Transactions on Medical Imaging*, 12(3):534–544, Sept. 1993.

[44] P. Roschmann. Radiofrequency penetration and absorption in the human body: limitations to high-field whole-body nuclear magnetic resonance imaging. *Medical Physics*, 14(6):922–31, Nov-Dec 1987.

[45] L. L. Schumaker and F. I. Utreras. On generalized cross validation for tensor smoothing splines. *SIAM Journal of Scientific and Statistical Computing*, 11(4):713–731, July 1990.

[46] A. Simmons, P. S. Tofts, G. J. Barker, and S. R. Arridge. Sources of intensity nonuniformity in spin echo images. *Magnetic Resonance in Medicine*, 32:121–128, 1994.

[47] D. Simunic, P. Wach, W. Renhart, and R. Stollberger. Spatial distribution of high-frequency electromagnetic energy in human head during mri: numerical results and measurements. *IEEE Transactions on Biomedical Engineering*, 43(1):88–94, Jan 1996.

[48] J. G. Sled, A. C. Evans, and G. B. Pike. Standing-wave and RF penetration artifacts caused by elliptic geometry: an electrodynamic analysis of MRI. In *International Society for Magnetic Resonance in Medicine*, 1997. in press.

[49] J. G. Sled, A. P. Zijdenbos, and A. C. Evans. A non-parametric method for automatic correction of intensity non-uniformity in MRI data. *IEEE Transactions on Medical Imaging*, 1996. (submitted).

[50] R. Stollberger and P. Wach. Imaging of the active B1 field in vivo. *Magnetic Resonance in Medicine*, 35:246–251, 1996.

[51] J. Talairach and P. Tournoux. *Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System - an Approach to Cerebral Imaging.* Thieme Medical Publishers, New York, NY, 1988.

[52] A. Thompson, J. Brown, and T. Titterington. A study of methods of choosing the smoothing parameters in image restoration by regularization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):326–339, 1991.

[53] J. Tropp. The theory of the bird-cage resonator. *Journal of Magnetic Resonance*, 82:51–62, 1989.

[54] M. W. Vannier, C. M. Speidel, and D. L. Rickman. Magnetic resonance imaging multispectral tissue classification. *News in Physiol Sci (NIPS)*, 3:148–154, Aug. 1988.

[55] J. T. Vaughan, H. P. Hetherington, J. O. Otu, J. W. Pan, and G. M. Pohost. High frequency volume coils for clinical NMR imaging and spectroscopy. *Magnetic Resonance in Medicine*, 32(2):206–18, Aug 1994.

[56] W. M. Wells III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz. Statistical intensity correction and segmentation of MRI data. In *Visualization in Biomedical Computing 1994, Proceedings of the SPIE*, volume 2359, pages 13–24, 1994.

[57] W. M. Wells III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz. Adaptive segmentation of MRI data. *IEEE Transactions on Medical Imaging*, 15(4):429–442, 1996.

[58] D. A. G. Wicks, G. J. Barker, and P. S. Tofts. Correction of intensity nonuniformity in MR images of any orientation. *Magnetic Resonance Imaging*, 11(2):183–196, 1993.

[59] A. Zijdenbos, A. Evans, F. Riahi, J. Sled, H.-C. Chui, and V. Kollokian. Automatic quantification of multiple sclerosis lesion volume using stereotaxic space. In *Forth International Conference on Visualization in Biomedical Computing (VBC)*, Hamburg, Germany, 1996.

[60] A. P. Zijdenbos, B. M. Dawant, and R. A. Margolin. Intensity correction and its effect on measurement variability in the computer-aided analysis of MRI. In *Proceedings of the 9th International Symposium and Exhibition on Computer Assisted Radiology (CAR)*, pages 216–221, Berlin, Germany, June 1995.

[61] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer. Morphometric analysis of white matter lesions in MR images: Method and validation. *IEEE Transactions on Medical Imaging*, 13(4):716–724, 1994.

[62] F. R. Zypman. MRI electromagnetic field penetration in cylindrical objects. *Computers in Biology & Medicine*, 26(2):161–175, Mar 1996.