Maximum Likelihood Estimation of the Bias Field in MR Brain Images: Investigating Different Modelings of the Imaging Process

Sylvain Prima[1], Nicholas Ayache[1], Tom Barrick[2], and Neil Roberts[2]

[1] INRIA Sophia Antipolis, EPIDAURE Project, France
[2] MARIARC, University of Liverpool, United Kingdom

Abstract. This article is about bias field correction in MR brain images. In the literature, most of the methods consist in modeling the imaging process before identifying its unknown parameters. After identifying two of the most widely used such models, we propose a third one and show that for these three models, it is possible to use a common estimation framework, based on the Maximum Likelihood principle. This scheme partly rests on a functional modeling of the bias field. The optimization is performed by an ECM algorithm, in which we have included a procedure of outliers rejection. In this way, we derive three algorithms and compare them on a set of simulated images. We also provide results on real MR images exhibiting a bias field with a typical "diagonal" pattern.

1 Introduction

Imperfections of the RF coil and patient-dependent electrodynamic interactions (often referred to as RF penetration and standing-wave effects [20]) systematically cause smooth, biologically meaningless, variations of the tissue intensities across MR images, which can amount to as much as 30% of the signal amplitudes. Generally, this "bias field" has little effect on visual interpretation, but the artificial intra-tissue variability it causes can significantly affect the outputs of image processing tools (segmentation, rigid or non-rigid registration, *etc.*) and subsequent quantitative analyses. Classical methods are routinely used to improve the intensity uniformity during the acquisition process [13]; they are generally able to correct most of the gross nonuniformities due to the coil defects, but unable to eliminate those due the patient anatomy. As an alternative, numerous retrospective methods have been devised to estimate and correct intensity variations after the image has been acquired. The common general approach consists in modeling the imaging process that connects the true emitted signal (uncorrupted by the bias field effects and the random noise due to the measuring device) with the observed data (*i.e.*, the MR image), and then estimating the parameters of the model best fitting the data. More precisely, given a voxel i with coordinates v_i in a MR image, its intensity y_i is widely considered to be related to the true emitted signal x_i according to:

$$y_i = b_i x_i + \varepsilon_i^{mea}$$

W. Niessen and M. Viergever (Eds.): MICCAI 2001, LNCS 2208, pp. 811–819, 2001.

Throughout the paper, given there are n voxels in the 3D volume, we note the bias field $b = (b_i)_{i=1...n}$, the MR image $y = (y_i)_{i=1...n}$, and the "ideal", uncorrupted, brain image $x = (x_i)_{i=1...n}$. The bias field is generally considered to be a slowly spatially varying function of the coordinates (*i.e.*, $b_i = b(v_i)$), and assumed to be multiplicative, consistently with the intrinsic nature of the corrupting physical processes [21]. In MR magnitude images, the random noise ε_i^{mea} due to the measuring device is known to have a Rician p.d.f. [19], which is shown to be quasi-Gaussian at high signal-to-noise ratio (SNR ¿ 3). Thus, the assumption of an additive, stationary, spatially white and Gaussian noise with a low standard deviation σ in the intracranial cavity is generally made: $\varepsilon_i^{mea} = \varepsilon^{mea} \sim N(0, \sigma^2)$. A commonly used and convenient modeling is to consider that every head voxel reflects the biological properties of one single underlying structure. In the following, the brain segmentation is noted $c = (c_i)_{i=1...n}$, c_i being the label of voxel i. A simple assumption is to model the intracranial cavity as composed of a reduced set of m tissues of interest. Considering cerebro-spinal fluid, grey matter and white matter, which generally exhibit distinct average grey levels in magnitude MR images, is a usual choice. Within a given neuroanatomical structure ω_k, natural intensity variations are always observed, due to changes in its composition across the head volume. Thus, the signal emitted by a particular structure is often assumed to be distributed around a mean value μ_k, which can be simply written $x_i = \mu_k + \varepsilon_i^{bio}$ if the voxel i belongs to ω_k. This "biological noise" ε_i^{bio}, due to a natural variability, is observed to be spatially correlated, but in many works related to brain segmentation, it is widely modeled as spatially white, with a stationary tissue-conditional Gaussian p.d.f. of low variance σ_k^2: $\varepsilon_i^{bio} = \varepsilon^{bio} \sim N(0, \sigma_k^2)$ [1,5,10,18]. This leads to rewrite the general model:

$$y_i = b_i(\mu_k + \varepsilon_i^{bio}) + \varepsilon^{mea},\ \varepsilon^{mea} \sim N(0, \sigma^2) \text{ and } \varepsilon_i^{bio} \sim N(0, \sigma_k^2) \text{ if } c_i = \omega_k \quad (1)$$

Following this very general model, many authors have proposed additional hypotheses and subsequent algorithms for estimation of the bias field and correction of the intensity nonuniformities in brain MR images. Most of these algorithms are based on two simplified versions of (1), that we call Models 1 and 2. What motivates this article is the observation that the variety of the algorithms proposed so far makes it difficult to compare the models they rest on.

In this paper, we suggest a third simplification of (1), which we call Model 3 (Section 2). Then, we propose to compare these three models, apart from the conceptually very different methods of the literature that has been developed so far to identify their underlying parameters (in particular, the bias field) ; our aim is to know which model is best suited to real data. For this purpose, we adopt a common estimation strategy, already proposed in [23,24] for Model 2. It consists in a probabilistic interpretation of these models, a functional modeling of the bias field, and a Maximum Likelihood estimation of the unknown parameters (tissue characteristics and bias field coefficients) by way of an Expectation/Conditional Maximization algorithm (Section 3). This estimation approach leads to simple iterative schemes, extensively described in a research report available on the

web (http://www.inria.fr/rrrt/index.en.html). Moreover, we propose to include in the three such derived algorithms a procedure of outliers rejection, inspired by the LTS estimation [16]. In particular, this technique allows to eliminate voxels affected by partial volume effects. We present results on simulated and real MR images (Section 4).

2 Three Different Models

In some works [2,12,15,14,22], x is supposed to be (at least locally) piecewise constant, depending on the underlying tissue ω_k: $x_i = \mu_k$ if $c_i = \omega_k$. This amounts to neglect the intensity variations due to the biological intra-tissue variability with respect to those due to the measurement noise ε^{mea} ($\varepsilon_i^{bio} \ll \varepsilon^{mes}$). Then the general model (1) becomes:

$$y_i = b_i \mu_k + \varepsilon^{mea} \text{ if } c_i = \omega_k, \text{ with } \varepsilon^{mea} \sim N(0, \sigma^2) \qquad (\textbf{Model 1})$$

In other works [23,24,25], it is supposed that the measurement noise can be at least partially removed by a low-pass prefiltering of the data (for instance, anisotropic diffusion [5]). A risk is to suppress also the relevant biological information conveyed by the intra-tissue natural variability [7]. However, as this latter is spatially correlated, it is less likely to be removed by such a filtering than the spatially white ε^{mea}. Then, (1) simply reduces to $y_i = b_i x_i$. In [25,7,23,24], mainly for reasons of computational simplicity, a logarithmic transform is applied to this model, which turns the multiplicative bias field b into an additive one. Then, it is particularily convenient to suppose that (within a given anatomical structure ω_k) the intensities x_i fluctuate around a mean value μ'_k, these fluctuations conveying a "biological noise" $\varepsilon_i^{'bio}$ following a Gaussian noise $N(0, \sigma_k^{'2})$. This hypothesis is contradictory to the traditional assumption done in algorithms of MR images segmentation, which suppose that ε_i^{bio} is Gaussian [1,5,10,18]; in this case, $\varepsilon_i^{'bio}$ is not. This point is not evoked in the concerned papers [25,7,23,24]. However, following this hypothesis, (1) becomes:

$$\log y_i = \log b_i + \mu'_k + \varepsilon_i^{'bio}, \text{ with } \varepsilon_i^{'bio} \sim N(0, \sigma_k^{'2}) \text{ if } c_i = \omega_k \qquad (\textbf{Model 2})$$

If, contrary to Model 2, we choose to keep the hypothesis of a Gaussian law for ε_i^{bio}, which seems more natural, we get a third model, defined as follows:

$$y_i = b_i(\mu_k + \varepsilon_i^{bio}), \text{ with } \varepsilon_i^{bio} \sim N(0, \sigma_k^2) \text{ if } c_i = \omega_k \qquad (\textbf{Model 3})$$

3 Three Algorithms for Bias Field Correction

3.1 Maximum Likelihood Formulation

In the following, we consider that there are $m = 3$ tissues of interest (cerebrospinal fluid, grey matter, white matter) and n voxels in the intracranial volume. We propose to make a functional parameterization of the bias field (detailed in Section 3.3), represented by a low number of coefficients $\alpha_1, \ldots, \alpha_d$. The objective is to estimate the optimal parameters of the three models according to the Maximum Likelihood principle. These parameters are gathered in the vector $\Theta = (\mu_1, \ldots, \mu_m, \sigma_1, \ldots, \sigma_m, \alpha_1, \ldots, \alpha_d)$ (or $\Theta = (\mu_1, \ldots, \mu_m, \sigma, \alpha_1, \ldots, \alpha_d)$ for Model 1). The optimal vector Θ maximizes the likelihood L of the image y, which can be written, using the theorem of total probabilities:

$$L(y;\Theta) = L(y_1, ..., y_n;\Theta) = \prod_{i=1}^{n} p(y_i;\Theta) = \prod_{i=1}^{n}\sum_{k=1}^{m} p(y_i|c_i = \omega_k;\Theta)P(c_i = \omega_k;\Theta)$$

Practically, we consider that the *a priori* probabilities $P(c_i = \omega_k;\Theta)$ are independent of the model parameters: $P(c_i = \omega_k;\Theta) = P(c_i = \omega_k)$. As suggested in [8,23,24], these probabilities $P(c_i = \omega_j)$ are obtained by affine registration of a probabilistic brain atlas for the three tissues of interest and give a rough *a priori*, fixed and spatially varying, knowledge of the tissue parameters and locations in the MR volume. We use the atlas from the Montral Neurological Institute [4].

3.2 Maximization of the Criterion

Analytical maximization of L is impossible. Given an initial estimate $\Theta^{(0)}$ of the parameters, the algorithm Expectation/Maximization (EM) [3] consists in building a series of vectors $(\Theta^{(p)})$ such that L converges towards a (at least local) maximum. This iterative process is composed of one "M-step" and one "E-step" defined as follows:

$$\nabla_\Theta Q(\Theta, \Theta^{(p)}) = \sum_{i=1}^{n}\sum_{k=1}^{m} P(c_i = \omega_k|y;\Theta^{(p)})\nabla_\Theta[\log p(y_i|c_i = \omega_k;\Theta)] = 0 \text{ (\textbf{M-step})},$$

$$\text{where } P(c_i = \omega_k|y;\Theta^{(p)}) \propto p(y_i|c_i = \omega_k;\Theta^{(p)})P(c_i = \omega_k) \text{ (\textbf{E-step})}$$

After choosing a functional parametrization of the bias field as described in Section 3.3, the M-step yields a system of $(2m + d)$ non-linear equations with $(2m + d)$ unknown parameters (or $(m + d + 1)$ for Model 1). This non-linearity makes it impossible to maximize the likelihood by a classical EM algorithm; a generalization is proposed in Section 3.4 to tackle this problem.

3.3 Modeling the Bias Field

As suggested in [23,24] for Model 2, an adapted functional parameterization of the bias field allows to characterize it with a limited number of parameters, while ensuring its spatial smoothness. We propose three functional modelings of the bias field adapted to each of the three models of the imaging process presented in Section 1. The objective is to obtain simple formulae for the p.d.f. of y_i (or $\log y_i$ for Model 2), such that the ML estimation of the whole set of parameters is made possible. Given the monomials ϕ_j, $j = 1, \ldots, d$, where $\phi_1 : (x, y, z) \mapsto (x - tx/2)$, $\phi_2 : (x, y, z) \mapsto (y - ty/2)$, $\ldots$,$\phi_7 : (x, y, z) \mapsto (x - tx/2)(y - ty/2)$, *etc.*, where tx, ty, tz are the sizes of the 3D image in the x, y and z directions, we make the following choices:

- **Bias Field Model 1**: $b_i = b(v_i) = 1 + \sum_{j=1}^{d} \alpha_j \phi_j(v_i)$
- **Bias Field Model 2**: $b_i = b(v_i) = \exp(\sum_{j=1}^{d} \alpha_j \phi_j(v_i))$
- **Bias Field Model 3**: $b_i = b(v_i) = 1/(1 + \sum_{j=1}^{d} \alpha_j \phi_j(v_i))$

3.4 The ECM Algorithm

Practically, the M-step of the classical EM algorithm yields intractable equations (see the research report): the optimal tissue or noise parameters (means and variances) explicitly depend on the optimal bias field coefficients in a non-linear fashion. To tackle this problem, a variant of the EM algorithm has been proposed in [11], that consists in partitioning the parameters vector as $\Theta = (\Theta_1, ..., \Theta_N)$, and replacing the original M-step by N successive Conditional Maximization steps (CM-steps) as described below. As the original EM approach, this Expectation/Conditional Maximization (ECM) algorithm provides a series of vectors $(\Theta^{(p)})$ such that L converges towards a (at least local) minimum [11].

- **CM-step 1:** $\Theta_1^{(p+1)} = \max_{\Theta_1} Q((\Theta_1, ..., \Theta_N), \Theta^{(p)})$
  $(\Theta_2, ..., \Theta_N) = (\Theta_2^{(p)}, ..., \Theta_N^{(p)})$ are held fixed.
- **CM-step 2:** $\Theta_2^{(p+1)} = \max_{\Theta_2} Q((\Theta_1^{(p+1)}, \Theta_2, ..., \Theta_N), \Theta^{(p)})$
  $(\Theta_1, \Theta_3, ..., \Theta_N) = (\Theta_1^{(p+1)}, \Theta_3^{(p)}, ..., \Theta_N^{(p)})$ are held fixed.
- And so on...
- After the last **CM-step** N, the estimated vector $\Theta^{(p+1)} = (\Theta_1^{(p+1)}, ..., \Theta_N^{(p+1)})$ is the input of the following E-step.

We consider the partition $\Theta = (\Theta_1, \Theta_2)$ with $\Theta_1 = (\mu_1, \ldots, \mu_m, \sigma_1, \ldots, \sigma_m)$ (or $\Theta_1 = (\mu_1, \ldots, \mu_m, \sigma)$ for Model 1), and $\Theta_2 = (\alpha_1, \ldots, \alpha_d)$. The equations for the mean and variance parameters constitute the first CM-step, which yields a straightforward explicit solution $\Theta_1^{(p+1)}$, the bias field coefficients being held fixed at the value $\Theta_2^{(p)}$. Then, given the estimated $\Theta_1^{(p+1)}$, the second CM-step amounts to solve a linear system of d equations, for Models 1 and 2.

In case of Model 3, the second CM-step is still a non-linear system of d equations, and is analytically intractable. Thus, we use an approximation, called

One-Step-Late (OSL), and proposed in [6] for reconstruction from SPECT data. The idea is to replace the non-linear terms of the CM-Step 2 of iteration $(p+1)$ of the ECM algorithm by their values estimated at the previous iteration (p), *i.e.*, to replace $1/(1+\sum_{j=1}^{d} \alpha_j^{(p+1)} \phi_j(v_i))$ by $1/(1+\sum_{j=1}^{d} \alpha_j^{(p)} \phi_j(v_i))$. Then the system becomes linear as in Models 1 and 2. An heuristic justification is to say that the ECM algorithm, as the original EM, is known to converge slowly, and this non linear term will not be much different between iterations (p) and $(p+1)$. We call the three such derived schemes Algorithms 1, 2 and 3.

3.5 Outliers Rejection Scheme

So far, we have proposed to minimize $\sum_{i=1}^{n} s_i$ with respect to Θ, where $s_i = -\log p(y_i;\Theta)$, and $\Theta = (\mu_1, \ldots, \mu_m, \sigma_1, \ldots, \sigma_m, \alpha_1, \ldots, \alpha_d)$. Typically, s_i is large when the intensity y_i fits the presumed underlying mixture model badly. In particular, this is the case for voxels affected by partial volume effects, and thus far from the m classes of interest. These voxels can severely offset the ML estimation. By an analogy with the LTS estimation [16], to eliminate these meaningless voxels and thus achieve a better robustness, we propose to minimize $\sum_{i=1}^{h} (s)_{i:n}$, where $(s)_{1:n} \leq \ldots \leq (s)_{n:n}$ are the ordered "residuals" and h is an integer superior to $\lfloor n/2 \rfloor$. In [17], an iterative scheme is proposed to compute a (at least) local minimum of the LTS criterion, which amounts to successive simple LS computations; following the same idea, we propose the following scheme (we do not give proof of convergence of this heuristic procedure, which practically gives good results, see Fig. 1) to minimize $\sum_{i=1}^{h} (s)_{i:n}$:

- Step 1: compute the ML estimate $\tilde{\Theta}$ of Θ on the whole dataset $(y_1, \ldots, y_n)$ by an ECM algorithm
- Step 2: compute the residuals $s_i = -\log p(y_i;\tilde{\Theta})$ on the whole dataset
- Step 3: sort out the residuals s_i
- Step 4: recompute the ML estimate $\tilde{\Theta}'$ on the data that exhibit the h lowest residuals, likely to be best suited to the model, by an ECM algorithm
- Step 5: go back to Step 2, set $\tilde{\Theta} = \tilde{\Theta}'$, and iterate until convergence

4 Validation

4.1 Experiments on Synthetic MR Images

Which of these three models is best suited to real MR data? As there is no ground truth for this latter, we propose a validation and comparison of the three models based on the MR simulator of the Montreal Neurological Institute [9], incorporating realistic models for bias field, measurement noise, uncorrupted intra-tissue intensity distributions and partial volume effects. Practically, we simulated 6 isotropic T1-weighted MR volumes (of voxel size 1mm^3) with different levels of noise (0%, 3%, 7%) and bias (20%, 40%). We use the RMS difference

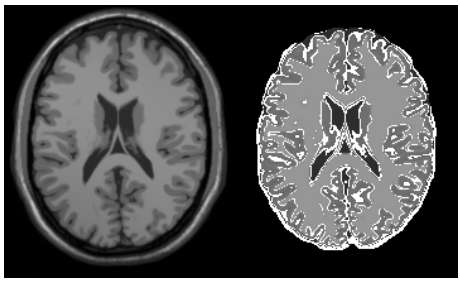

Fig. 1. Robust estimation of the parameters. Left: synthetic T1-weighted MR image, generated by the MNI simulator [9], with noise level 0% and bias field level 40%. Right: intensity-corrected MR image (by Algorithm 3); the voxels in white have been rejected from the estimation. Most of these voxels are close to the tissue boundaries, and are affected by partial volume effects.

between the applied and the computed bias field as a measure of error to evaluate the precision of the three algorithms. Figure 2 shows the results of these 6 experiments. When the noise is weak (0 or 3%), Algorithm 3 performs slightly better than Algorithm 2, and both of them are largely better than Algorithm 1. When the noise level is higher (7%), Algorithm 2 performs largely better than Algorithm 1 and 3. On average, Algorithm 2, and thus the underlying Model 2, seems to be the best.

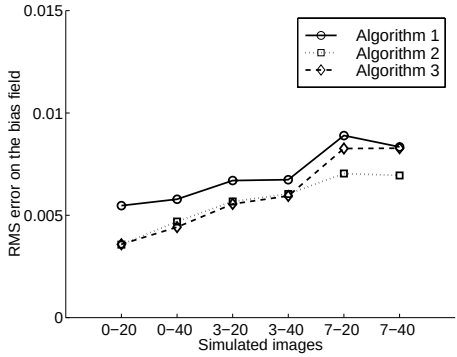

Fig. 2. RMS errors on the bias field for the three algorithms applied to the six simulated images. The degree of the three polynomials of the bias field models is $D = 2$. In average, Algorithm 2 performs better than the two others.

4.2 Experiments on Real MR Images

We applied Algorithm 3 on 10 images of healthy subjects, provided by the MARIARC, University of Liverpool, UK. Acquired by a MR scanner GE SIGNA 1,5 T using a circularly polarized coil, they are of size $256 \times 256 \times 124$ (voxel size $0.78125 \times 0.78125 \times 1.6$). In Figure 3, we display 10 axial slices of the original MR images and the estimated bias fields, which have a characteristic "diagonal" structure. For each subject, the voxels in the right temporal and frontal lobes have higher intensities than their counterparts in the other hemisphere; this pattern is inverted in the occipital lobes. This result confirms works of other authors [20], who link this bias field asymmetry with the elliptical shape of the head.

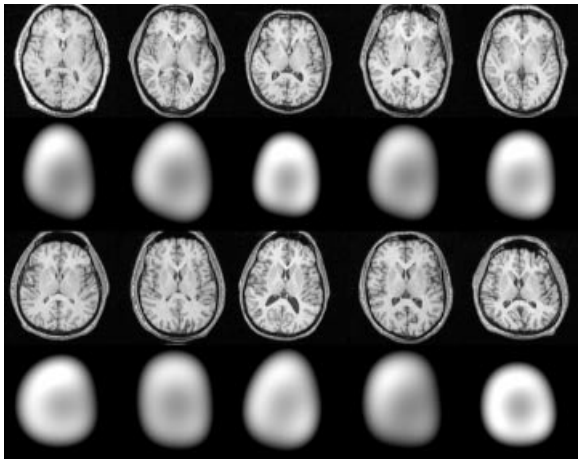

Fig. 3. Bias field estimation on 10 MR brain images. These images are in neurological conventions. $D = 4$. The voxels in the temporal and frontal lobes have higher intensities than their counterparts in the other hemisphere; this pattern is inverted in the occipital lobes.

5 Conclusion

In this article, we have identified two widely used models of the imaging process for bias field estimation. We have proposed a third model, and shown that it is possible to use a common estimation framework for these three models to identify the set of their unknown parameters. This iterative scheme rests on the principle of Maximum Likelihood. The optimization is performed by an ECM algorithm, in which we have included a procedure of outliers rejection. In this way, we have derived three algorithms and compared them on a set of simulated images. We have given a set of results on real MR images for which the bias field has a typical diagonal pattern.

References

1. H.E. Cline *et al.* Three-Dimensional Segmentation of MR Images of the Head Using Probability and Connectivity. *Journal of Computer Assisted Tomography*, 14(6):1037–1045, 1990.
2. B.M. Dawant *et al.* Correction of Intensity Variations in MR Images for Computer-Aided Tissue Classification. *IEEE TMI*, 12(4):770–781, December 1993.
3. A. P. Dempster *et al.* Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39:1–38, 1977.
4. A.C. Evans *et al.* 3D statistical neuroanatomical models from 305 MRI volumes. In *IEEE Nuclear Science Symposium and Medical Imaging Conference*, pages 1813–1817, San Francisco, USA, October 1993.
5. G. Gerig *et al.* Nonlinear Anisotropic Filtering of MRI Data. *IEEE TMI*, 11(2):221–232, June 1992.
6. P.J. Green. Bayesian Reconstructions From Emission Tomography Data Using a Modified EM Algorithm. *IEEE TMI*, 9(1):84–93, March 1990.
7. R. Guillemaud and M. Brady. Estimating the Bias Field of MR Images. *IEEE TMI*, 16(3):238–251, June 1997.

8. M. Kamber *et al.* Model-based 3D segmentation of multiple sclerosis lesions in dual-echo MRI data. In *VBC'92*, volume 1808 of *SPIE*, pages 590–600, Chapel Hill, USA, October 1992.
9. R.K.-S. Kwan *et al.* An Extensible MRI Simulator for Post-Processing Evaluation. In *VBC'96*, volume 1131 of *LNCS*, pages 135–140, Hamburg, Germany, September 1996. Springer-Verlag. MRI simulator: http://www.bic.mni.mcgill.ca/brainweb/.
10. Z. Liang *et al.* Parameter estimation and tissue segmentation from multispectral MR images. *IEEE TMI*, 13:441–449, September 1994.
11. X.L. Meng and D.B. Rubin. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, 80(2):267–278, 1993.
12. C.R. Meyer *et al.* Retrospective Correction of Intensity Inhomogeneities in MRI. *IEEE TMI*, 14(1):36–41, March 1995.
13. P.A. Narayana *et al.* Compensation for surface coil sensitivity variation in magnetic resonance imaging. *Magnetic Resonance Imaging*, 6(3):271–274, 1988.
14. D.L. Pham and J.L. Prince. A Generalized EM Algorithm for Robust Segmentation of Magnetic Resonance Images. In *33rd Annual Conference on Information Sciences and Systems , CISS'99*, pages 558–563, Baltimore, USA, March 1999.
15. J.C. Rajapakse and F. Kruggel. Segmentation of MR images with intensity inhomogeneities. *Image and Vision Computing*, 16(3):165–180, 1998.
16. P.J. Rousseeuw and A.M. Leroy. *Robust Regression and Outlier Detection.* Wiley Series in Probability and Mathematical Statistics, 1987.
17. P.J. Rousseeuw and K. Van Driessen. Computing LTS Regression for Large Data Sets. Technical report, Statistics Group, University of Antwerp, 1999. submitted.
18. P. Schroeter *et al.* Robust Parameter Estimation of Intensity Distributions for Brain Magnetic Resonance Images. *IEEE TMI*, 17(2):172–186, April 1998.
19. J. Sijbers *et al.* Maximum Likelihood estimation of Rician distribution parameters. *IEEE TMI*, 17(3):357–361, 1998.
20. J.G. Sled and G.B. Pike. Magnetic Resonance Imaging - Standing-Wave and RF Penetration Artifacts Caused by Elliptic Geometry: An Electrodynamic Analysis of MRI. *IEEE TMI*, 17(4):653–662, August 1998.
21. J.G. Sled and A.P. Zijdenbos. A Nonparametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data. *IEEE TMI*, 17(1):87–97, February 1998.
22. M. Styner *et al.* Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE TMI*, 19(3):153–165, March 2000.
23. K. Van Leemput *et al.* Automated Model-Based Bias Field Correction of MR Images of the Brain. *IEEE TMI*, 18(10):885–896, October 1999.
24. K. Van Leemput *et al.* Automated Model-Based Tissue Classification of MR Images of the Brain. *IEEE TMI*, 18(10):897–908, October 1999.
25. W. M. Wells III *et al.* Adaptive Segmentation of MRI Data. *IEEE TMI*, 15(4):429–442, August 1996.