Automatic MS Lesion Segmentation by Outlier Detection and Information Theoretic Region Partitioning

Release 0.00

Marcel Prastawa[1] and Guido Gerig[1]

July 17, 2008

[1]Scientific Computing and Imaging Institute
University of Utah, Salt Lake City, UT 84112, USA
{prastawa, gerig}@sci.utah.edu

Abstract

Multiple Sclerosis (MS) is a neurodegenerative disease that is associated with brain tissue damage primarily observed as white matter abnormalities such as lesions. We present a novel, fully automatic segmentation method for MS lesions in brain MRI that combines outlier detection and region partitioning. The method is based on an atlas of healthy subjects and detects lesions as outliers, without requiring the use of training data with segmented lesions. In order to segment lesions as spatially coherent objects and avoid spurious lesion detection, we perform classification on regions (connected groups of voxels) instead of individual voxels. Each voxel location is assigned to a region that would maximize overall relative entropy or Kullback-Leibler divergence between neighboring regions. Our proposed method is fully automatic and does not require manual selection or outlining of specific brain regions. The method can also be adapted to MR images obtained from different scanners and scanning parameters as it requires no training.

Contents

1 Introduction

Multiple Sclerosis (MS) is a neurodegenerative disease where myelin sheathes of the neurons are destroyed by the immune system. This disease is associated with brain tissue damage (e.g., lesions) that can be observed through 3D Magnetic Resonance Imaging (MRI), as shown in Figure 1. The cause and cure for MS is an active area of research in the medical community. Studies involving populations of MS patients require quantitative analysis of lesions observed in vivo through MR imaging. Determining quantities such as the size, shape, location of lesions is crucial for studying the progression of MS lesions and the effect of drug treatments.

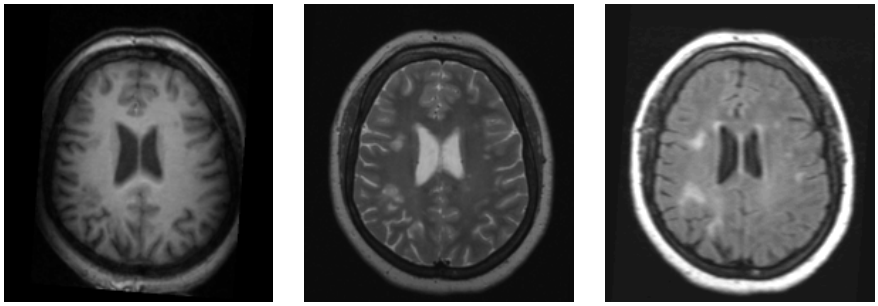

Figure 1: Example MR scan of a subject with MS lesions (UNC test subject 04). From left to right: T1w image, T2w image, FLAIR image. Lesions are shown as the areas with bright intensities in both T2w and FLAIR modalities. The bright areas in FLAIR around the ventricle are artifacts caused by the pulsation of the ventricles.

The segmentation of lesions from 3D MRI is the first step in performing a quantitative analysis of lesion progression. Automatic segmentation methods are of great value in MS population studies due to their high levels of reliability and reproducibility. A number of segmentation methods have been proposed by other researchers. Zijdenbos *et al.* proposed a method based on a neural network classifier [10]. van Leemput *et al.* [7] proposed a method that detects lesions as outliers from the intensity distributions of healthy tissue. Thirion *et al.* [6] and Rey *et al.* [4] segmented lesions based on the local deformation between scans at different time points.

Automatic lesion segmentation methods that have been proposed typically perform classification at the voxel level, and require training images or longitudinal information. We propose a novel, fully automatic iterative segmentation scheme for brain lesions in MRI that requires no training or longitudinal data. Lesions are detected as deviations from normal human brains, as represented by a brain atlas. This approach allows us to segment lesions without explicit training and delineation of lesions in different subjects from a specific scanning sequence. Classification is performed on groups of voxels to directly perform segmentation of lesions as objects formed by spatially coherent voxels, and to significantly reduce false positives inherently linked to conventional voxel-based classification.

2 Method

2.1 Overview

The proposed method uses a brain atlas that functions as a model of the healthy adult human brain. Brain lesions are detected as outliers from the expected healthy anatomy. An example brain atlas from the International Consortium for Brain Mapping (ICBM) is shown in Figure 2. The atlas is registered to the subject

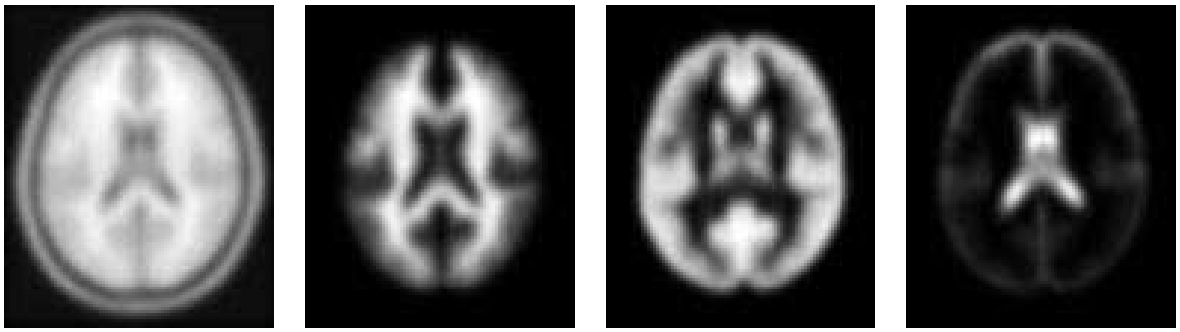

Figure 2: The digital brain atlas provided by the International Consortium for Brain Mapping (ICBM). From left to right: the T1 template image and probability values of white matter, gray matter, and csf.

image using the mutual information image match metric following the approach proposed by Maes *et al.* [2] using affine and B-spline deformable transforms. Our algorithm uses the atlas to determine likely sample locations for healthy tissues and also as spatial priors in a Bayesian segmentation framework.

The algorithm is iterative and it alternates between estimating the intensity probability density functions (pdf), computing voxelwise spatial probabilities, correcting for intensity inhomogeneities, and determining the partitioning of the images into spatially coherent regions. Inhomogeneity correction is performed by using the spatial probabilities, as proposed by van Leemput et al [8]. Figure 3 shows an overview of the iterative segmentation algorithm.

2.2 Outlier Detection

The intensity probability density function for the healthy brain tissue is computed using samples obtained from the atlas. The spatial priors for healthy tissues are thresholded at high values (e.g., 0.9) to determine the likely samples. The Minimum Covariance Determinant (MCD) robust estimation scheme [5] is then applied to the intensity samples for healthy tissues to determine inliers and outliers. The MCD robust estimator computes the hyper-ellipsoid that covers at least half of the input data. Samples with Mahalanobis distance greater than a threshold M are treated as outliers. An example application of the MCD algorithm is shown in Figure 4. The inlier samples are used to form the pdf of the brain tissue intensities using kernel density estimation. This step is identical to the brain tumor segmentation scheme that we have proposed [3]. Outlier samples that follow a user specified rule (e.g., lesions appear bright in FLAIR modality and are brighter than gray matter in T2w modality) are assigned to the lesion class, similar to van Leemput et al's approach [7]. This user specified rule needs to be designed to isolate the relevant outlier samples for MS lesions. For example, the ventricle pulsation artifact in FLAIR generates bright intensities around the ventricle that are detected as outliers, yet are not proper MS lesions.

The intensity pdfs are used to compute spatial posterior probabilities of each tissue in a Bayesian framework. The posterior probability for class c at location x is computed as follows:

$$p_x(c) = p(c|x) = \frac{p(\vec{I}(x)|c)p(c,x)Pr(c)}{\sum_c p(\vec{I}(x)|c)p(c,x)Pr(c)} \tag{1}$$

where $\vec{I}$ denotes the vector of observed image intensities in different modalities, $p(c,x)$ denotes the atlas spatial prior, and $Pr(c)$ denotes the global prior for class c which is a free parameter that is determined empirically.

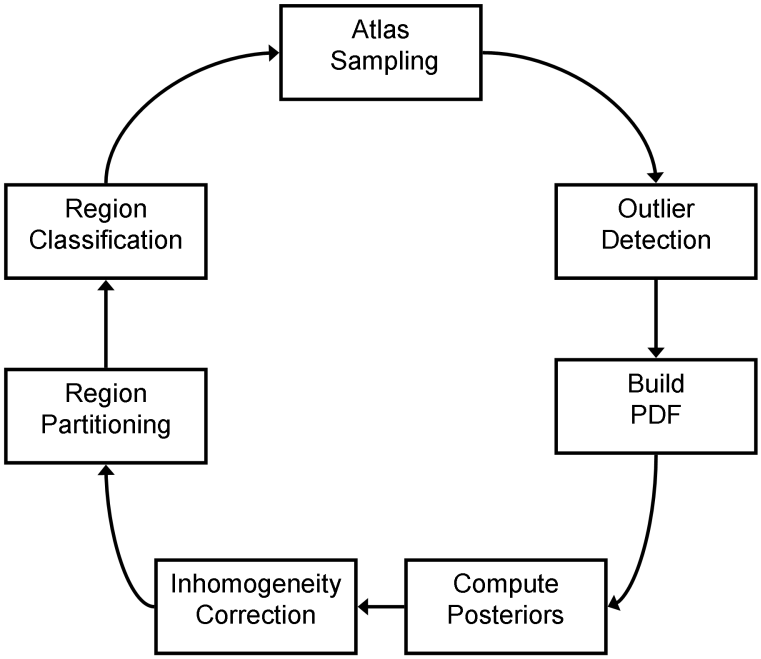

Figure 3: Overview of our proposed automatic segmentation scheme. The algorithm alternately samples the image using the atlas, determines outlier intensity samples, performs Bayesian classification, corrects intensity inhomogeneity, and partitions the MRI into regions.

2.3 Region Partitioning

We simulate a human rater's segmentation approach that groups perceptually similar voxels by partitioning image voxels into relevant regions. The classification for each voxel is based on the region classification, which reduces spurious lesion classification (e.g., in voxels with high noise). Each region is a group of voxels that have similar appearance and anatomy. We assume that both appearance and anatomy are encoded in the set of class posterior probabilities (Equation 1) at each image location.

Region partitioning is initialized by using the watershed image transform on the gradient magnitude of the multimodal images [1]. Our method then reassigns the voxels at the boundary between two regions to maximize the symmetric Kullback-Leibler(KL) divergence or relative entropy between two regions:

$$dist(R_i, R_j) = \frac{1}{2}(KL(\bar{p}, \bar{q}) + KL(\bar{q}, \bar{p})) \tag{2}$$

where c denotes the tissue class, $\bar{p}$ and $\bar{q}$ denote the overall tissue probabilities for region R_i and R_j respectively, and $KL(\bar{p}, \bar{q}) = \sum_c \bar{p}(c) log \frac{\bar{p}(c)}{\bar{q}(c)}$. Maximizing the KL divergence between two neighboring regions ensures that different regions are dissimilar with regards to appearance and anatomy, as encoded by the posterior probabilities. The mean tissue probabilities $\bar{p}$ is chosen to minimize the KL divergence between the mean and the tissue probabilities at each voxel within the region: $\sum_{x \in R_i} KL(p_x, \bar{p})$ which reduces $\bar{p}$ to

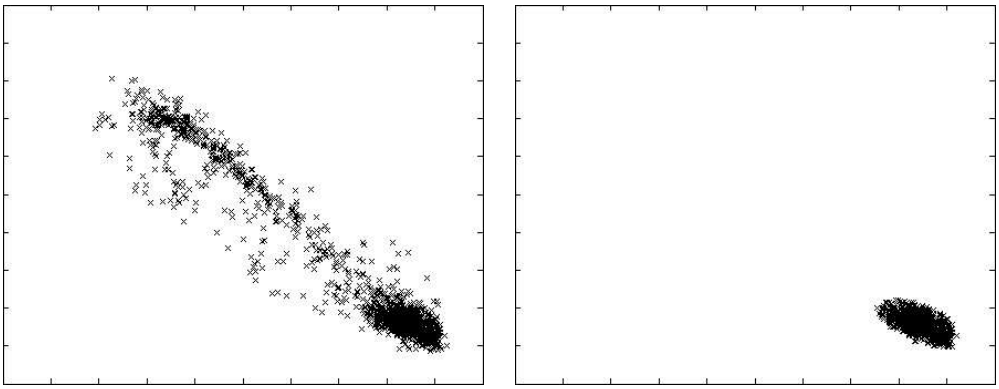

Figure 4: An example application of the MCD robust estimation algorithm. Horizontal axis represents the T1 intensities and the vertical axis represents the T2 intensities. Left: original samples obtained by atlas-guided sampling which is contaminated with samples from other distributions. Right: remaining samples after trimming samples with Mahalanobis distance greater than 3 using the robust MCD estimate.

the normalized geometric mean of the voxel probabilities p_x:

$$\bar{p}(c) = \frac{[\prod_x p_x(c)]^{\frac{1}{|R_i|}}}{\sum_c [\prod_x p_x(c)]^{\frac{1}{|R_i|}}} \tag{3}$$

where $|R_i|$ denotes the number of voxels in region R_i.

3 Results

We have applied our new algorithm to the test datasets from the Children's Hospital of Boston (CHB) and the University of North Carolina (UNC). We use the T1w, T2w, and the FLAIR modalities to generate the results. Example results of our segmentation scheme on a randomly chosen dataset from each institution are shown in Figure 5, These results are generated without using the training data provided by both institutions. The segmentation performance for all the test datasets are shown in the table shown in Figure 6. The performance metrics used are the volume difference, average distance between lesion boundaries, true positive rate, and false positive rate. Performance was also compared against the ground truth computed using the STAPLE algorithm [9].

The results are generated using a high Mahalanobis threshold for intensity outliers ($M = 3.75$) and low global prior for lesion ($Pr(lesion) = 0.1$). This results in a conservative lesion segmentation scheme, which is reflected in the high specificity and low sensitivity scores listed in Figure 6. The lesion rule specified to the algorithm is to select as lesions the intensity samples that are brighter than gray matter in FLAIR, brighter than gray matter in T2w, and darker than white matter in T1w. Application of this lesion rule aids in removing irrelevant outliers that are often caused by ventricle pulsation artifacts in FLAIR that appear bright similar to lesions, yet are typically not present in the T1w and T2w modalities.

4 Discussion

We have presented a fully automatic segmentation scheme for MS lesions that requires minimal user interaction. The segmentation combines a healthy brain atlas and outlier detection, therefore it requires no

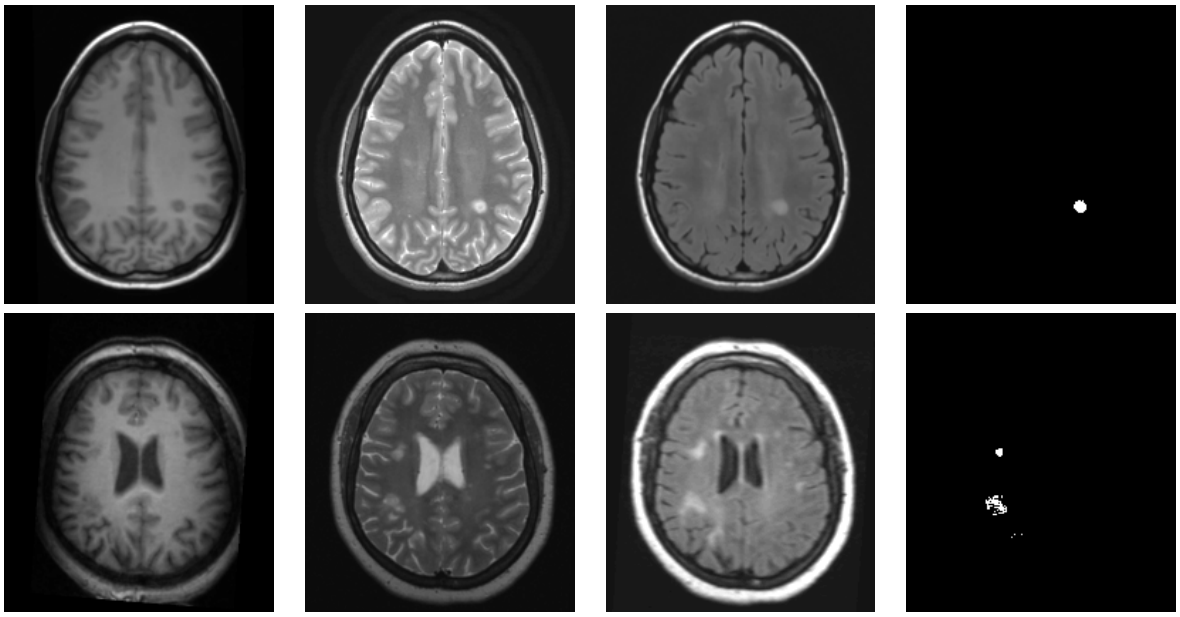

Figure 5: Example results of our segmentation algorithm. Top row: CHB test subject 13. Bottom row: UNC test subject 04. From left to right: T1w image, T2w image, FLAIR image, and the lesion segmentation.

manual delineation or selection of specific structures in MRI; either for training or initialization. As the algorithm requires no training, it can be directly applied to images generated using different MR scanners or different MR scanning parameters. Extension to other lesion characteristics or different MR scanning sequences can be done by specifying a different rule for isolating the intensity samples for lesions. In addition to lesion classification, the algorithm also generates the classication of the whole 3D brain which includes white matter, gray matter, and cerebrospinal fluid.

The method partitions the image into regions that are similar both with regard to anatomy and appearance in MR. Currently, we have only implemented boundary evolution for optimizing the partitioning. In the future, we plan to improve the partitioning scheme by adding region splitting using information theoretic measures. We generated the results using only structural modalities. Our method can be extended to include information from diffusion tensor imaging (DTI). However, there are significant issues to be addressed in fusing DTI and structural information. DT images are typically scanned at lower resolution, and they seem to show different progressions of demyelination.

5 Acknowledgements

This work is part of the National Alliance for Medical Image Computing (NA-MIC), funded by the National Institutes of Health through Grant U54 EB005149. Information on the National Centers for Biomedical Computing can be obtained from http://nihroadmap.nih.gov/bioinformatics.

The segmentation method described in this paper was implemented using the Insight Segmentation Registration toolkit (http://www.itk.org) and the Python interpreter (http://www.python.org).

Ground Truth	UNC Rater								CHB Rater									STAPLE		
All Dataset	Volume Diff.		Avg. Dist.		True Pos.		False Pos.		Volume Diff.		Avg. Dist.		True Pos.		False Pos.		Total	Specificity	Sensitivity	PPV
	[%]	Score	[mm]	Score	[%]	Score	[%]	Score	[%]	Score	[mm]	Score	[%]	Score	[%]	Score				
UNC test1 Case01	89.2	87	14.3	71	2.3	53	99.3	49	84.2	88	15.9	67	6.2	55	98.6	50	65	0.9948	0.0044	0.0382
UNC test1 Case02	95.4	86	14.6	70	13.2	59	48.3	80	99.4	85	14.5	70	4.5	54	10.3	100	76	1.0000	0.0104	0.9843
UNC test1 Case03	87.4	87	5.8	88	9.9	57	37.3	87	83.8	88	5.8	88	11.8	58	34.7	89	80	0.9996	0.1685	0.9621
UNC test1 Case04	85.3	88	7.6	84	21.1	63	64.6	70	77.1	89	4.7	90	37.0	73	59.8	73	79	0.9973	0.1464	0.8186
UNC test1 Case05	100.0	85	62.0	0	0.0	51	100.0	49	100.0	85	63.1	0	0.0	51	100.0	49	46	1.0000	0.0000	0.0000
UNC test1 Case06	99.9	85	51.8	0	0.0	51	100.0	49	99.6	85	48.4	0	0.0	51	100.0	49	46	1.0000	0.0008	0.6344
UNC test1 Case07	88.2	87	8.9	82	9.8	57	16.7	99	72.7	89	7.0	86	23.3	65	30.0	91	82	0.9999	0.1787	0.9850
UNC test1 Case08	98.6	86	21.4	56	6.4	55	0.0	100	97.6	86	12.7	74	16.7	61	0.0	100	77	1.0000	0.0196	1.0000
UNC test1 Case09	92.4	86	52.8	0	0.0	51	100.0	49	89.3	87	53.9	0	0.0	51	100.0	49	47	0.9991	0.0000	0.0000
UNC test1 Case10	93.1	86	24.2	50	5.0	54	86.7	57	75.1	89	14.1	71	16.7	61	80.0	61	66	0.9997	0.0321	0.7875
CHB test1 Case01	98.7	86	19.5	60	8.0	56	46.7	81	98.2	86	16.3	67	19.4	62	20.0	97	74	1.0000	0.0104	0.9903
CHB test1 Case02	85.3	88	15.3	68	9.1	57	75.0	64	93.7	86	25.8	47	10.5	57	50.0	79	68	0.9999	0.0774	0.9747
CHB test1 Case03	75.7	89	12.9	73	7.1	56	99.2	49	88.3	87	14.3	71	13.3	59	98.3	50	67	0.9964	0.0042	0.0176
CHB test1 Case04	99.5	85	40.4	17	0.0	51	100.0	49	99.8	85	38.4	21	0.0	51	100.0	49	51	0.9999	0.0000	0.0000
CHB test1 Case05	78.9	88	11.2	77	22.2	64	91.9	54	96.0	86	11.8	76	34.8	71	89.2	55	71	0.9987	0.0159	0.3935
CHB test1 Case06	96.5	86	6.6	86	8.3	56	39.3	86	96.4	86	6.7	86	9.1	57	40.4	85	78	0.9997	0.0286	0.8790
CHB test1 Case07	99.4	85	30.7	37	1.7	52	85.7	57	99.7	85	21.1	57	2.6	53	85.7	57	61	0.9998	0.0007	0.2098
CHB test1 Case08	91.7	87	9.0	82	25.9	66	88.9	56	94.5	86	10.7	78	17.6	62	86.9	57	72	0.9983	0.0313	0.5016
CHB test1 Case09	100.0	85	81.2	0	0.0	51	100.0	49	100.0	85	74.4	0	0.0	51	100.0	49	46	1.0000	0.0000	0.0000
CHB test1 Case10	80.1	88	7.8	84	26.3	66	92.1	54	90.3	87	9.9	80	24.1	65	90.3	55	72	0.9975	0.0232	0.3105
CHB test1 Case11	98.1	86	26.0	46	4.5	54	60.0	73	99.4	85	34.1	30	6.9	55	60.0	73	63	0.9999	0.0056	0.8036
CHB test1 Case12	100.0	85	128.0	0	0.0	51	0.0	100	100.0	85	128.0	0	0.0	51	0.0	100	59	1.0000	0.0000	nan
CHB test1 Case13	85.3	88	22.6	53	20.0	63	60.0	73	91.0	87	25.0	48	9.5	57	60.0	73	68	0.9999	0.1204	0.9905
CHB test1 Case15	100.0	85	128.0	0	0.0	51	0.0	100	100.0	85	128.0	0	0.0	51	0.0	100	59	1.0000	0.0000	nan
All Average	92.5	86	33.4	49	8.4	56	66.3	68	92.7	86	32.7	50	11.0	58	62.3	70	66	0.9992	0.0366	0.5582
All UNC	93.0	86	26.3	50	6.8	55	65.3	69	87.9	87	24.0	55	11.6	58	61.3	71	66	0.9990	0.0561	0.6210
All CHB	92.1	87	38.5	49	9.5	57	67.1	67	96.2	86	38.9	47	10.6	58	62.9	70	65	0.9993	0.0227	0.5059

Figure 6: Quantitative performance metrics for the proposed segmentation scheme. Our method achieves a high score for specificity which indicates low rates of false lesion identification.

References

[1] H. C. Lee and D. R. Cok. Detecting boundaries in a vector field. *IEEE Trans. Signal Processing*, 39(5):1181–1194, 1991.

[2] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging*, 16(2):187–198, April 1997.

[3] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig. A brain tumor segmentation framework based on outlier detection. *Medical Image Analysis*, 8(3):275–283, 2004.

[4] David Rey, Gérard Subsol, Hervé Delingette, and Nicholas Ayache. Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis. *Medical Image Analysis*, 6(2):163–179, June 2002.

[5] P. J. Rousseeuw and K. Van Driessen. A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41(3):212–223, 1999.

[6] J. P. Thirion and G. Calmon. Deformation analysis to detect and quantify active lesions in three-dimensional medical image sequences. *IEEE TMI*, 18(5):429–441, 1999.

[7] K. van Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens. Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE TMI*, 20(8):677–688, 2001.

[8] K. van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based bias field correction of MR images of the brain. *IEEE Trans. Medical Imaging*, 18:885–896, October 1999.

[9] Simon K. Warfield, Kelly H. Zou, and William M. Wells. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE TMI*, 23(7):903–921, 2004.

[10] A. Zijdenbos, R. Forghani, and A. Evans. Automatic quantification of MS lesions in 3D MRI brain data sets: Validation of INSECT. In *Proc. MICCAI 1998*, LNCS 1496, pages 439–448, 1998.