Effects of Anatomical Asymmetry in Spatial Priors on Model-Based Segmentation of the Brain MRI: A Validation Study

Siddarth Srivastava, Frederik Maes, Dirk Vandermeulen, Wim Van Paesschen, Patrick Dupont, and Paul Suetens

Katholieke Universiteit Leuven, Faculties of Medicine and Engineering, Medical Image Computing (Radiology - ESAT/PSI), University Hospital Gasthuisberg, Herestraat 49, B-3000 Leuven, Belgium.
`Siddharth.Srivastava@uz.kuleuven.ac.be`

Abstract. This paper examines the effect of bilateral anatomical asymmetry of spatial priors on the final tissue classification based on maximum-likelihood (ML) estimates of model parameters, in a model-based intensity driven brain tissue segmentation algorithm from (possibly multispectral) MR images. The asymmetry inherent in the spatial priors is enforced on the segmentation routine by laterally flipping the priors during the initialization stage. The influence of asymmetry on the final classification is examined by making the priors subject-specific using non-rigid warping, by reducing the strength of the prior information, and by a combination of both. Our results, both qualitative and quantitative, indicate that reducing the prior strength alone does not have any significant impact on the segmentation performance, but when used in conjunction with the subject-specific priors, helps to remove the misclassifications due to the influence of the asymmetric priors.

1 Introduction

A normal human brain has systematic bilateral asymmetry, exhibited, most notably, in the frontal and the occipital regions. This effect is referred to as the *brain torque*. For a normal population, the right frontal lobe is larger than the left, the left occipital lobe is larger than the right ([8] and references therein). This normal inter-hemispheric morphological variability is reflected in all databases that are built upon a large representative population study. An example of such a database is the digital brain atlas provided with SPM99 [6] (figure 1) which contains spatially varying probability maps for gray matter (GM), white matter (WM) and cerebro-spinal fluid (CSF), obtained by averaging deterministic classifications of respective tissue classes from a large normal population, aligned to a standard reference space by affine transformations [2]. The asymmetry inherent in the normal brains, hence, is reflected in the inter-hemispheric asymmetry of the priors. This atlas has wide applications in the field of automated brain tissue segmentation, wherein it is used as a first estimate of the classification, and to spatially constrain the classification [3]. The use of the priors
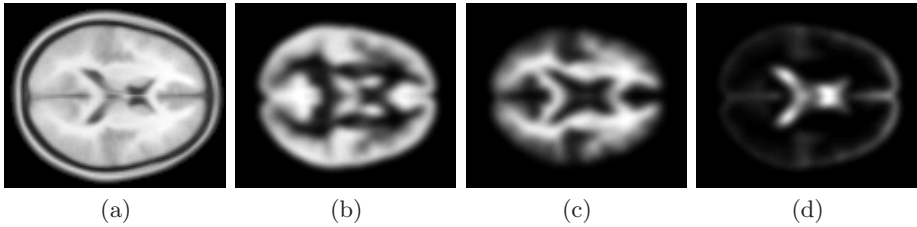
Fig. 1. Digital Brain Atlas as provided by SPM99 [6]. (a) Normal T1 brain template, (b)-(d): apriori spatial probability maps for gray matter, white matter and csf respectively

in their usual form, however tends to introduce a spatial bias in the derived classifications, of a nature similar to the asymmetry content in the priors. Furthermore, asymmetry studies and anatomical lateralization over and above the normal morphogenic incongruity of the human brain are strong indicators of widespread and local neurodegenerative disorders. Typically, in such computational neuroscience methodologies, it is the segmentation map which is under examination. This fact, combined with the ubiquity of the automated segmentation methodologies using a brain atlas as spatial *a-priori* information merits a study examining the influence that the asymmetric priors have on the final results, and devise ways to attenuate or ideally, obliterate their influence. Such a study will ensure that the morphological findings are intrinsic to the data under examination, and not an artifact of the intermediate processing steps.

2   Materials and Methods

We use high-resolution simulated 3D T1 weighted brain MR images provided by Brainweb [1]. All the data sets have 1 mm isotropic voxel dimensions. The intensity homogeneity was chosen to be 20% to engage the bias correction component of the segmentation scheme. The experiments were performed for noise levels $\sigma = \{3, 5, 7, 9\}\%$. Corresponding to these data sets were anatomical deterministic ground truths G per tissue class, which were used as a reference to which we compare the segmentation maps produced by each segmentation methodology. We use the spatial priors P^o in their original orientation, as provided in the standard distribution of SPM99 [6]. We also derive a new set of priors P^f by flipping P^o bilaterally in the native space. This is done to transfer the intrinsic bilateral asymmetry in the spatial priors to the opposite side. Because of the flipping in the native space of the atlas, no re-alignment of P^f to P^o was performed.

2.1   Standard Segmentation

The standard segmentation strategy under examination is the method presented in [3], which parameterizes the tissue-specific intensity distribution as a gaussian, resulting in a mixture model for the entire data set. We denote this strategy by

$\mathcal{U}$. At the core is the classical Expectation-Maximization (EM) loop which is initialized by model parameter estimates from P^o or P^f. The standard processing methodology involves reformating the spatial priors for each tissue class, to the space of the study subject, which has already been aligned to a standard reference space by a 12 parameter affine transformation using maximization of mutual information [4] as an image similarity measure. The realigned spatial priors are then used to initialize the EM loop, and also to spatially constrain the tissue classification during the search for ML estimates of model parameters. The results of the segmentation, however, differ, depending on whether the classification has been initialized by P^o or P^f. The effect is clearly visible in figure 2, which shows an axial slice of the GM classification map obtained by initialization and constraining the classification by P^o, and by P^f respectively, and the absolute difference between the two classifications.

(a)   (b)   (c)
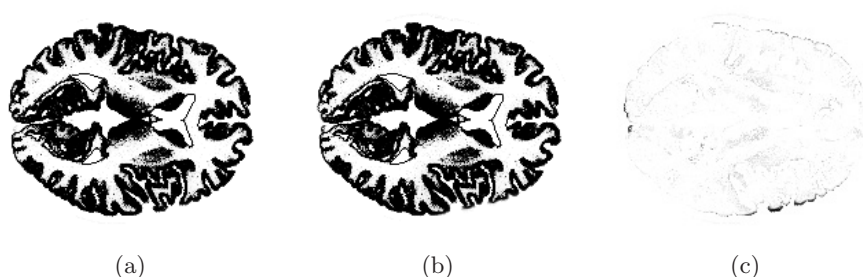
Fig. 2. (a) Probabilistic GM classification obtained by model-based tissue classification initialized with P^o, (b) probabilistic GM classification obtained on initialization with P^f, (c) absolute difference of (a) and (b), showing a systematic misclassification on the outer periphery of the cortex, predominantly at the frontal and contralateral occipital lobe. The colormap is inverted (high(dark) to low(bright)).

2.2   Segmentation with Subject Specific Priors

An affine mapping of the priors P^o or P^f to the study subject ensures global anatomical correspondence between the subject and the prior. Local morphological variability in the brain of the subject to be segmented is not taken into account by this methodology. A more optimal way of providing the EM segmentation routine with initial estimates would be to normalize the priors to the study subject by using a high-dimensional non-rigid registration technique [7]. Warping the priors to the subject brings the priors in better anatomical correspondence with the subject, resulting in a more accurate initialization, which is likely to improve the overall segmentation. Starting from the initial affine alignment of the spatial priors to the study subject, a set of deformation fields were estimated from the T1 weighted template provided with SPM99 [6], using the non-rigid registration tool of SPM99 [6], with default parameter settings. These

deformation fields were then applied to the priors, enhancing the local anatomical correspondence. The deformed priors were then provided to the segmentation routine as initial estimates and for successive spatial constraints. We denote this strategy by $\mathcal{W}$.

2.3   Segmentation with Reduced Priors

In [5] it is argued that it is not convenient to initialize the segmentation with actual *a-priori* probabilities because the empirical frequencies in the atlas tissue maps encode a wide variation in population, and are different from the anatomical variation in the study subject. As proposed in [5] we moderate the priors by a small factor before initialization and subsequent processing steps. If $p_t(r)$ denotes the a-priori probability of voxel r being of the tissue type t, then the prior is modified into

$$q_t(r) = \alpha p_t(r) + (1 - \alpha)\frac{1}{K} \tag{1}$$

where K is the number of classes under consideration, and $\alpha \in [0, 1]$ is the weighing parameter that serves to reduce the influence of the spatial priors on classification. We have applied this prior reduction strategy to each of the segmentation strategies mentioned in sections 2.1 and 2.2, for values of $\alpha = \{0.3, 0.5, 0.7, 0.9, 1.0\}$, where $\alpha = 1.0$ refers to unmodified spatial priors. In cases where reduced priors were used in combination with $\mathcal{W}$ (section 2.2), the reduction was applied after the non-rigid registration.

2.4   Segmentation Performance Index

Each segmentation methodology yields a classification map for each of the tissue classes, which we denote by $\mathcal{C}$. The free parameters for our experiments are the segmentation methodologies $\mathcal{M} \in \{\mathcal{U}, \mathcal{W}\}$ representing the affine and warped priors respectively, orientations $k \in \{o, f\}$ of the priors P^o and P^f respectively and the reduction fraction $\alpha \in \{0.3, 0.5, 0.7, 0.9, 1.0\}$ for the priors, and the noise content of the data $\sigma \in \{3, 5, 7, 9\}\%$. For ease of reference, a segmentation for a particular tissue class performed on data with a particular noise content σ, with methodology $\mathcal{M}$, and initialization performed with priors in orientation k will be denoted by $\mathcal{C}_\alpha^{k,\mathcal{M}}$. Based on this notation, we have, for a particular orientation k of the prior, and for a particular noise level, 10 possible segmentations for each tissue class. Since the basic objective of the work is to check for the effects of asymmetry of the priors on the segmentation, our hypothesis is that a segmentation $\mathcal{C}_\alpha^{k,\mathcal{M}}$ is immune to the asymmetry effects if it agrees well with its corresponding ground truth G, and also if $\mathcal{C}_\alpha^{o,\mathcal{M}}$ and $\mathcal{C}_\alpha^{f,\mathcal{M}}$ agree well with each other. We use the Dice Similarity Coefficient (DSC) described in [9] as a measure of agreement between the segmentations for a particular tissue class using different methodologies. This similarity index is defined as $DSC(\mathcal{C}_1, \mathcal{C}_2) = \frac{2V(\mathcal{C}_1 \cap \mathcal{C}_2)}{V(\mathcal{C}_1) + V(\mathcal{C}_2)}$, where $V(\mathcal{C})$ is the volume of the classification map $\mathcal{C}$, defined, in our case, as the

total number of voxels classified in a classification map (all voxels with non-zero probability). This measure is used to evaluate a ratio of twice the number of commonly classified voxels, to the number of all the voxels classified in C_1 and C_2, without taking into account the probability of the classification. This makes the measure more stringent as compared to discretizing both the volumes by deterministic assignment (hard segmentation) [3], and more responsive to spatial misclassifications even with a small membership probability. The value of this DSC lies between 0 and 1, and serves as a performance index increasing from 0 to 1, indicating an excellent agreement for $DSC > 0.7$ [7]. For ease of reference, we define $s_\alpha^{k,\mathcal{M}} = \mathrm{DSC}(C_\alpha^{k,\mathcal{M}}, G)$ and $s_\alpha^{\mathcal{M}} = \mathrm{DSC}(C_\alpha^{o,\mathcal{M}}, C_\alpha^{f,\mathcal{M}})$, where G refers to the ground truth corresponding to the classification map C.

3   Results

We have studied the segmentation performance based on DSC, on simulated Brainweb data [1]. Since most of the systematic errors occur in the GM (figure 2(c)), we present results only for the GM, for noise content of 5% and 7%. This is the noise content from moderate to extreme that is typically present in most of the high resolution acquisitions that are used. Table 1 shows the DSC based performance index for GM, and its variations across methodologies and noise level. Figure 3 shows the GM segmentations for all methodologies (for $\alpha = 1.0$ and $\alpha = 0.7$), for $\sigma = 5\%$, allowing visual inspection of the improvement in segmentation for each methodology.

Table 1. Similarity index (DSC) for GM, for all methodologies. The results are shown for noise levels $\sigma = 5\%$ and $\sigma = 7\%$. The left-most column shows the segmentation methodology, standard ($\mathcal{U}$) or warped ($\mathcal{W}$), along with the reduction fraction α.

$\sigma \Longrightarrow$	$\sigma = 5\%$			$\sigma = 7\%$		
$(\mathcal{M}, \alpha) \Downarrow$	$s_\alpha^{o,\mathcal{M}}$	$s_\alpha^{f,\mathcal{M}}$	$s_\alpha^{\mathcal{M}}$	$s_\alpha^{o,\mathcal{M}}$	$s_\alpha^{f,\mathcal{M}}$	$s_\alpha^{\mathcal{M}}$
$(\mathcal{U}, 1.0)$	0.725612	0.722171	0.986609	0.709364	0.708018	0.992829
$(\mathcal{W}, 1.0)$	0.730665	0.730551	0.992660	0.713761	0.712719	0.994128
$(\mathcal{U}, 0.3)$	0.722913	0.714043	0.980580	0.709440	0.708352	0.993268
$(\mathcal{U}, 0.5)$	0.721188	0.714287	0.981220	0.708900	0.708214	0.993152
$(\mathcal{U}, 0.7)$	0.721942	0.713215	0.980316	0.708662	0.708607	0.993026
$(\mathcal{U}, 0.9)$	0.722741	0.715180	0.981435	0.708951	0.709029	0.992880
$(\mathcal{W}, 0.3)$	0.730177	0.729552	0.993097	0.713373	0.712819	0.994267
$(\mathcal{W}, 0.5)$	0.730270	0.729546	0.992827	0.713594	0.712928	0.994203
$(\mathcal{W}, 0.7)$	0.731153	0.729602	0.992718	0.713463	0.712810	0.994210
$(\mathcal{W}, 0.9)$	0.730947	0.729415	0.992680	0.713562	0.712726	0.994411
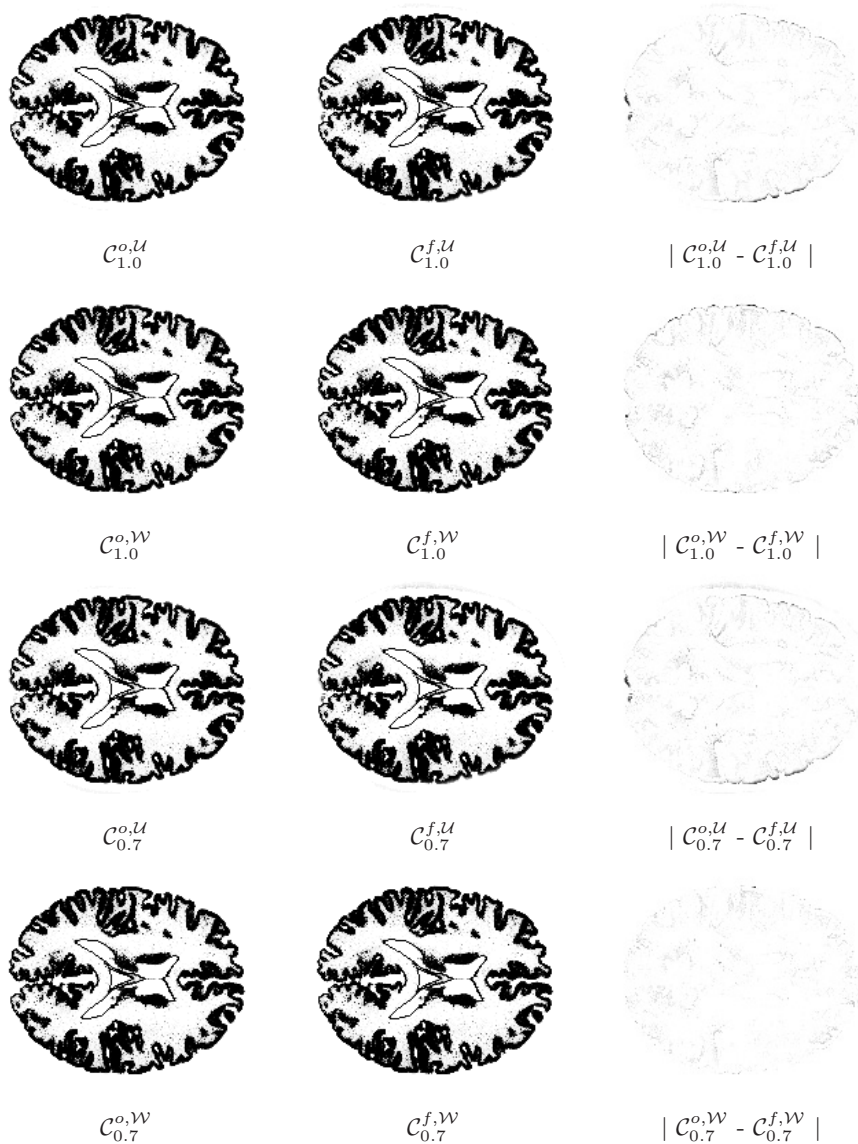
Fig. 3. Column-wise: segmentations initialized with P^o, P^f, and their absolute differences respectively. Methods vary row-wise, as designated by the corresponding labels. The noise level for the data is $\sigma = 5\%$

4   Discussion

Table 1 gives a global overview of segmentation performance of various methodologies, and their ability to make the segmentation maps indifferent to the spatial

asymmetry content in the prior probability maps. We observe a small improvement, with respect to the similarity index, between the affine-transformed, and non-rigidly deformed prior based initialization strategy, with un-modified ($\alpha = 1.0$) priors. Upon inclusion of the reduction strategy (section 2.3) in the standard segmentation procedure $\mathcal{U}$ there is a drop in the DSC measure, and it remains less than the performance of $\mathcal{U}$ for $\alpha = 1.0$. The use of $\mathcal{W}$, however, performs better than $\mathcal{U}$ for all values of α, and performs better than ($\mathcal{W}$,1.0) (table 1) for $0.7 \leq \alpha \leq 0.9$. This holds true for all noise ranges $\sigma \in \{3, 5, 7, 9\}\%$. Within a particular methodology, there seems to be a very minimal dependence of α upon the DSC based comparison with the ground truth. This does not agree well with the findings reported in [5]. The differences could be because of the fact that we calculate volumes in DSC metric as voxel counts in the posterior probability maps, and not as voxel counts in hard segmented classification maps. Since the basic objective of this work is to make the segmentations immune to the anatomical asymmetry in the priors, a successful strategy will be one that gives high values of performance index s_α^k, indicating that the segmentations of the same study subject, obtained by initializations with P^o and P^f respectively agree well with each other, while at the same time are in good agreement with their respective ground truths. Table 1 demonstrates that some improvement is gained in s_α^k by changing the normalization strategy from affine-only to non-rigid. Further, there is a slight improvement, for the non-rigid normalization, for values of $0.3 \leq \alpha < 1.0$, in comparison to the values for ($\mathcal{U}$,1.0) and ($\mathcal{W}$, 1.0). It should be mentioned here that the similarity across segmentations, as demonstrated under column labeled s_α^k in table 1 is high even without any modification to the segmentation methodology. Since the misclassifications due to asymmetry represent a small fraction of connected voxels at anatomically specific locations, a removal of this effect lends a very small fraction to the performance index as an improvement, making small changes in performance very significant. Residual errors ($1.0 - DSC$) in the most optimal segmentation (figure 3, bottom row) can be attributed to the remaining misclassified voxels spread over the entire segmentation map, while segmentation errors due to the misclassified voxels at the outer edge of the cortex are completely removed.

5   Conclusion

This paper investigates the effect of anatomical bilateral asymmetry in the spatial prior probability maps used to initialize, and spatially guide intensity based segmentation, on the resulting classifications. The premise for the experiments performed was that the same study-subject brain, segmented under laterally flipped conditions (P^o or P^f) should result in the same segmentation, and should also agree well with the corresponding anatomical ground truth. To this end, we have used the volume similarity as the performance index, and compared gray matter segmentations obtained after initialization of the classification method with P^o and P^f respectively. Since the effects of asymmetry are more visible in the gray matter segmentation maps, we have presented results only for the gray

matter. Our results indicate that anatomical asymmetry in the prior probability maps influence the segmentation by severely misclassifying small but significant regions of the brain, especially at the outer surface of the cortex. Such an effect might get swamped in volumetric studies, but is of significance in studies where topological and geometrical properties of the cortex are under examination. Further, the use of non-rigid registration improves the segmentation performance, as compared to the usual affine-only spatial correspondence, and also helps to remove the effects of asymmetry from the final GM classification map. Reducing the prior strength alone does not result in any significant improvement of the segmentation. Further, visual analysis of figure 3 (bottom row) shows that the reduced prior strategy, when used in conjunction with the non-rigid registration, serves to remove the effect of the spatial prior asymmetry from the classification.

References

1. C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, and A.C. Evans. Brainweb: Online interface to a 3d mri simulated brain database. *NeuroImage*, 5(4):S425, 1997. part 2/4.
2. A.C. Evans, D.L. Collins, S.R. Mills, E.D. Brown, R.L. Kelly, and T.M. Peters. 3d statistical neuroanatomical models from 305 mri volumes. *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference*, pages pp 1813–1817, 1993.
3. K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model based tissue classification of MR images of the brain. *IEEE Transactions on Medical Imaging*, 18(10):pp 897–908, 1999.
4. F. Maes, A. Collignon, D. Vandermeulen, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):pp 187–198, 1997.
5. J.L. Marroquin, B.C. Vemuri, S. Botello, F. Calderon, and A. Fernandez-Bouzas. An accurate and efficient bayesian method for automatic segmentation of brain MRI. *IEEE Transactions on Medical Imaging*, 21(8):pp 934–945, 2002.
6. Wellcome Department of Cognitive Neurology. Statistical Parametric Mapping (SPM). http://www.fil.ion.ucl.ac.uk/spm/.
7. Killian M. Pohl, William M. Wells, Alexandre Guimond, Kiyoto Kasai, Martha E. Shenton, Ron Kikinis, W. Eric L. Grimson, and Simon K. Warfield. Incorporating non-rigid registration into expectation maximization algorithm to segment mr images. *MICCAI*, LNCS 2488:pp 564–571, 2002.
8. S. Prima, S. Ourselin, and N. Ayache. Computation of the mid-saggital plane in 3d brain images. *IEEE Transactions on Medical Imaging*, 21(2):pp 122–138, 2002.
9. A.P. Zijdenbos, B.M. Dawant, R.A. Margolin, and A.C. Palmer. Morphometric analysis of white matter lesions in MR images: Methods and validation. *IEEE Transactions on Medical Imaging*, 13(4):pp 716–724, 1994.