Topology-corrected segmentation and local intensity estimates for improved partial volume classification of brain cortex in MRI.

Andrea Rueda, Oscar Acosta, Michel Couprie, Pierrick Bourgeat, Jurgen Fripp, Nicholas Dowson, Eduardo Romero, Olivier Salvado

Topology-corrected segmentation and local intensity estimates for improved partial volume classification of brain cortex in MRI

Andrea Rueda[a,b], Oscar Acosta[a,d], Michel Couprie[c], Pierrick Bourgeat[a], Jurgen Fripp[a], Nicholas Dowson[a], Eduardo Romero[b], Olivier Salvado[a]

[a] *CSIRO Preventative Health National Research Flagship, ICTC, The Australian e-Health Research Centre - BioMedIA, Herston, Australia.*
[b] *BioIngenium Research Group, Universidad Nacional de Colombia, Bogotá, Colombia*
[c] *Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, ESIEE Paris, France*
[d] *Laboratoire de Traitement du Signal et de l'Image (LTSI), INSERM U642, Université de Rennes 1, France.*

Abstract

In magnetic resonance imaging (MRI), accuracy and precision with which brain structures may be quantified are frequently affected by the partial volume (PV) effect. PV is due to the limited spatial resolution of MRI compared to the size of anatomical structures. Accurate classification of mixed voxels and correct estimation of the proportion of each pure tissue (fractional content) may help to increase the precision of cortical thickness estimation in regions where this measure is particularly difficult, such as deep sulci. The contribution of this work is twofold: on the one hand, we propose a new method to label voxels and compute tissue fractional content, integrating a mechanism for detecting sulci with topology preserving operators. On the other hand, we improve the computation of the fractional content of mixed voxels using local estimation of pure tissue intensity means. Accuracy and precision were assessed using simulated and real MR data and comparison with other existing approaches demonstrated the benefits of our method. Significant improvements in gray matter (GM) classification and

cortical thickness estimation were brought by the topology correction. The fractional content root mean squared error diminished by 6.3% ($p < 0.01$) on simulated data. The reproducibility error decreased by 8.8% ($p < 0.001$) and the Jaccard similarity measure increased by 3.5% on real data. Furthermore, compared with manually-guided expert segmentations, the similarity measure was improved by 12.0% ($p < 0.001$). Thickness estimation with the proposed method showed a higher reproducibility compared with the measure performed after partial volume classification using other methods.
Keywords: Brain tissue segmentation, Partial volume classification, Magnetic resonance imaging, Topology correction, Sulci detection, Cortical thickness estimation

1. Introduction

Accurate segmentation of Magnetic Resonance (MR) images into different brain tissues, namely gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), can allow *in-vivo* quantification of structural modifications appearing during neurodegenerative diseases. However, MR-related artifacts, such as intensity inhomogeneity, noise and partial volume (PV) effects, can hamper the precision of this task. Inhomogeneities can be characterized by a low frequency multiplicative bias field and are mostly due to the sensitivity of the receiver coils and, in some cases, to non-homogeneous tissue MR properties. The noise is Rician distributed and it has be shown to strongly affect the tissue classification (Van Leemput et al., 2003). Finally, PV effects appear when the size of anatomical features being imaged is comparable to the voxel size, causing blurring at the interfaces between tissues. In some cases, e.g. with opposed banks of GM in deep sulci, misclassification problems appears, affecting further processings such as cortical thickness estimation.

Topological operators and constraints have been widely used to correct and achieve accurate cortical tissue segmentations (Ségonne, 2008; Bazin and Pham, 2005; Han et al., 2002; Kriegeskorte and Goebel, 2001). It has been assumed that the cerebral cortex is a folded sheet of GM built upon the WM, which would have the topology of a hollow sphere if the midline hemispheric connections were artificially removed. Due to MR artifacts, the segmentation process cannot guarantee this assumption, generating deviations from the true anatomy of the structures of interest. Proposed approaches that address this issue can be classified in two categories: methods that include topological constraints directly into the segmentation process, based on active contours (Ségonne, 2008), topology adaptive snakes (McInerney and Terzopoulos, 1999), digital topology models (Bazin and Pham, 2005, 2007) or segmentation by registration to an atlas (Kriegeskorte and Goebel, 2001); and retrospective techniques that correct the topology after the segmentation process (Han et al., 2002). Those approaches are focused on ameliorating the topology of the segmented tissues, working directly on a voxel or on a mesh (surface) space. Voxel-based methods operate directly on the volumetric tissue segmentations, by removing or adding voxels according to topological constraints. However, remotion or addition of a whole voxel in thin structures such as the GM may considerably modify the measure of thickness (ranging between $\pm$ 1 voxel) if any mechanism such as partial volume is not used to compensate for the structural modifications. In contrast, mesh-based techniques requires an initial 3D reconstruction (triangular mesh) of the volumetric segmentations. The approaches for segmentation and cortical thickness estimation operating directly with the surfaces, such as CLASP (Kim et al., 2005), BrainVISA (Mangin et al., 1995) or Freesurfer (Dale et al., 1999; Fischl et al., 1999; Fischl and Dale, 2000), incorporate mechanisms to pre-

vent self-intersection of surfaces or topology correction, imposing also some smoothness constraints. Mesh-based approaches are however computationally more expensive, because of the needed additional reconstruction step. Overall, after or during the mesh generation, most of the methods tackle the elimination of tunnels and handles (Fischl et al., 1999; Florent Ségonne and Fischl, 2007; Jaume et al., 2005; Zhou et al., 2007).

On the other hand, PV estimation has received considerable attention in the last few years and different approaches have been proposed for classification and computation of fractional content (Santago and Gage, 1993; Laidlaw et al., 1998; Shattuck et al., 2001; Noe and Gee, 2001; Van Leemput et al., 2003; Tohka et al., 2004; Chiverton and Wells, 2008). Most techniques model voxel intensity as a linear combination of the intensity distributions of the possible tissue types within each voxel (Choi et al., 1991; Noe and Gee, 2001). Computing the fractional content of voxels therefore requires both pure and mixed voxels to have been previously classified. Shattuck et al. (2001) implemented a maximum *a posteriori* (MAP) classifier, which combined a tissue measurement model with a prior model of the local spatial interactions to obtain six tissue types: three pure and three mixed. The fractional content for the mixed voxels was calculated based on the global intensity mean of pure tissue types. Tohka et al. (2004) proposed an algorithm which used statistical estimators, based on the MAP estimation (Shattuck et al., 2001). Recently, Chiverton and Wells (2008) presented a local adaptive Gradient-controlled spatial regularizer (GSR) using a Markov Random Field to model the class membership and a Markov chain Monte Carlo (MCMC) simulation to adapt the model to the observed data. The labelling error may remain high because the intensity inhomogeneities (not explicitly modelled) and the noise may lead to misdetection of mixed voxels mainly in tight sulci,

representing a portion of GM/CSF/GM within the same voxel.

The approaches previously presented have been focused on solving either the PV estimation or the topology correction. Our contribution consists in demonstrating that better results and performance are obtained if both strategies are combined together with a spatial intensity variation modeling. In this paper, we propose a new method aimed at improving both PV classification and fractional content computation, working at a voxel level in order to be accurate and computationally efficient. The improved classification is achieved by imposing topological constraints to the binary segmentation and thus detecting hidden mixed voxels in zones of tight sulci. The accurate fractional content estimation is attained by computing the fractional content as a linear relation between robust local intensity averages of pure tissue voxels. The spatially dependent averaging helps to overcome the problems of intensity inhomogeneity for a given tissue across the image.

In the next section we describe our methods, followed by experiments using simulated and real data. We also compare the results with other previously proposed methods. We demonstrated the utility of our approach by integrating the whole process to our voxel-based cortical thickness estimation pipeline.

2. Methods

The proposed strategy follows the steps depicted in Figure .1: Firstly, an initial classification of voxels into pure tissues WM, GM and CSF and mixed tissues WM/GM and GM/CSF is performed. Secondly, topology-constraints are introduced in the classification assuming that the GM is a continuous layer covering the WM. A topology preserving dilation of the WM over GM adds robustness to the delineation of mixed voxels GM/CSF in deep sulci.

Finally, the estimation of fractional content for mixed voxels is adaptively performed based on a local averaging of the pure tissue voxels.

INSERT FIGURE .1 HERE

2.1. Pure tissue segmentation

A first segmentation of pure brain tissues into GM, WM and CSF is performed based on an implementation of the expectation-maximisation (EM) segmentation method as in (Van Leemput et al., 1999). Here, the Colin atlas and associated priors are first affinely registered to the data using a robust block matching approach (Ourselin et al., 2001), followed by a diffeomorphic Demons non-rigid registration (Vercauteren et al., 2007). Probabilistic tissue maps associated with the atlas were used to initialize the EM segmentation and enforce spatial consistency throughout the segmentation. The probability density functions of the tissues are modelled with 6 Gaussians (WM, GM, CSF and 3 for non brain tissues, skull and background). Finally, hard segmentations are obtained after the EM segmentation by labelling each voxel with the most probable tissue.

2.2. Initial partial volume labelling

Using the hard segmentations, a first labelling of partial volume voxels are identified within the hard segmentations and along the interfaces of pure tissues. Three pure tissue classes and two mixture classes are considered $\Gamma = \{\mathrm{GM}, \mathrm{CSF}, \mathrm{WM}, \mathrm{CSF/GM}, \mathrm{GM/WM}\}$. A maximum *a posteriori* classification (MAP) is made and labels the voxels as belonging to the set Γ. This procedure, relying on both intensity and spatial information, extends the method proposed by (Shattuck et al., 2001), but we assume that each voxel contains at most two tissues (Santago and Gage, 1993), and PV classification is restricted to the region formed by a dilated GM region (radius 2) because

only the cortical thickness is sought. To take into account dependency on the neighbouring tissue types, a Markov prior that models local spatial interactions was implemented using a Potts model in order to perform the labelling. As in (Shattuck et al., 2001; Tohka et al., 2004; Kim et al., 2005), we use the Iterated Conditional Modes (ICM) algorithm as explained in (Besag, 1986) to search for the optimal labelled image. According to this, every voxel is updated once per iteration until no label changes occur between iterations. This model favors classification of contiguous regions of GM, WM and CSF and encourages configurations of voxels that make physical sense such as GM/CSF or GM/WM voxels adjacent to GM.

2.3. A topology preserving segmentation

After the MAP labelling, some of the sulci may be misdetected, as the intensity of buried PV GM/CSF voxels is close to that of the GM. In order to refine the segmentation and identify such buried GM/CSF voxels, we used a homotopic dilation of the consolidated $\boldsymbol{WM} = \{\mathrm{WM}, \mathrm{WM/GM}\}$ constrained by the GM, leading to a better delineation of deep sulci. To preserve this folds during dilation, the set $\boldsymbol{WM}$ is corrected first to assure that shares the topology of a filled sphere.

The homotopic transformations that we used are topology-preserving procedures that consist of sequentially deleting or adding single points (voxels) as described in (Bertrand and Malandain, 1994). The algorithms used are detailed in Appendix A. Our topology preserving segmentation of the WM consists in performing a homotopic dilation of a seed set of voxels, called **S**, constrained to only add voxels from the set **WM**, knowing that **S** is topologically equivalent to a filled sphere. The result of this operation is denoted by **SWM**. For example, **S** could be made of single voxels chosen in the white matter, but we describe below a way to obtain a seed that is closer to the

expected result, and thus leads to a more robust segmentation.

To obtain the seed **S**, we first compute a surface skeleton **SK** of **WM**, by dilating using Algorithm 2 as described in Appendix A. Then, we perform an homotopic erosion, constrained by **SK**, of a full cuboid that includes **SK**. Finally, we perform an homotopic dilation of the same seed set **S**, constrained by the set $\mathbf{SWM} \cup \mathbf{GM}$ to only add GM and WM voxels, and we substract **SWM** from the result to obtain the corrected GM.

This method is performed on 3D sets, but for clarity we illustrate it on a 2D reduced example in Figure .2. Notice that small black components in Figure 2(b) can correspond to tunnels in the 3D image, thus simple connected component filtering would not give the correct region. Figures .3 and .4 show further examples in 3D.

INSERT FIGURE .2 HERE

INSERT FIGURE .3 HERE

INSERT FIGURE .4 HERE

2.4. Partial volume relabelling and fractional content

The main contribution of the topology is the relabelling of missegmented GM voxels in hidden sulci as mixed GM/CSF. Once the topologically corrected WM, GM, CSF, WM/GM and GM/CSF segmentations are obtained, the portion of pure tissue, called here fractional content F, is computed for each mixed voxel by estimating the local contribution of each pure tissue. We assume that each voxel contains at most two tissues and the new labelling corresponds only to the mixed voxels WM/GM and GM/CSF. For each mixed voxel, the fractional content F ranges between $[0, 1]$ depending on the amount of pure tissue. Thus, for pure tissue voxels the fractional

content F_j are set to 1 for the class j and 0 otherwise. For mixed voxels ($x \in WM/GM, GM/CSF$), the fractional content $F_{j/k}$ between both pure tissues j and k is computed using the intensity $I(x)$ of the image and the robust local averages of the closest pure tissue voxels $\mu_j(x)$ and $\mu_k(x)$, such that:

$$F_{j/k}(x) = U\left(\frac{\mu_k(x) - I(x)}{\mu_k(x) - \mu_j(x)}\right) \quad (1)$$

where $U(\cdot)$ is a limiter restricting the range of the fractional content to $[0, 1]$. Unlike (Shattuck et al., 2001), which uses the same linear relation between global means of tissues to compute fractional content, we compute μ_k and μ_j as robust local averages rather than global means. This is done by computing the mean of the median 50% of pure tissue intensities (interquartile mean) within a $5mm$ radius sphere, thus rejecting local outliers, over a denoised version of the original MR image. The noise is removed by applying the optimized non-local means method proposed in (Coupe et al., 2008).

Pure tissue voxels are selected by eroding pure tissue segmentations using a $2mm$ radius, therefore reducing the influence of any mixed voxel. Finally, the computed averages are propagated back towards the location of the mixed voxels x, resulting in values of $\mu_j(x)$ and $\mu_k(x)$ that represent the average of the closest pure tissue voxels (Figure .5). The GM fractional content map is eventually defined as $F_{\mathrm{GM/WM}} \cup F_{\mathrm{GM}} \cup F_{\mathrm{GM/CSF}}$. Using a robust local mean overcomes issues related to intensity inhomogeneities and variations of pure tissue signal across the image, weighting accordingly the signal when computing the fractional content.

INSERT FIGURE .5 HERE

INSERT FIGURE .6 HERE

Regional differences in the cell structure and the distribution of different layers of the cortex result in variation of regional intensity differences for the same tissue across the brain. These differences produces local variation of contrast between the tissues that might be pronounced with ageing Salat et al. (2009). Whereas global homogeneity assumptions will bias the voxel fractional content estimation, a local computation of intensity averages for pure tissue yields a more accurate value, which accounts for the changes in cytoarchitecture visibles in MR. A local estimate allows also to overcome the issues of intensity inhomogeneities due to the artifacts during the acquisition.

To illustrate the spatial differences in signal, GM intensity was measured over the population of 20 young adults scans, acquired as described in Subsection 3.3. Figure .6 shows the local average intensity of GM across the brain for an individual. In this example, precentral gyrus presented a higher average value than the temporal or occipital lobe. The same pattern appear in average in all the healthy individuals. The contrast between the tissues has been also measured using the Equation 2 as

$$F_{Contrast} = \frac{\mu_{WM} - \mu_{GM}}{\mu_{GM} - \mu_{CSF}} \tag{2}$$

where μ_{WM}, μ_{GM} and μ_{CSF} are the regional averages of WM, GM and CSF respectively, which can be considered as a measure of the contrast between WM and GM normalized by the CSF. Figure .7 shows the regional differences for the population of 20 young controls.

INSERT FIGURE .7 HERE

3. Experiments

To evaluate our method, named hereafter as Topologically-corrected Partial Volume (TPV), we used different brain MR data sets including simulated and real images. The purpose was twofold, firstly to illustrate the effect of the topology correction in the estimation of fractional content for mixed voxels, and secondly to compare the obtained results with those publicly available in the area. After that, the method was integrated to our voxel-based cortical thickness estimation pipeline. Experiments demonstrated that the overall method showed a better estimate of thickness and a high reproducibility on real data.

3.1. Simulated MR data

A set of 15 simulated MR brain images was obtained from the BrainWeb Simulated Brain Database, maintained by the McConnell Brain Imaging Centre at the Montreal Neurological Institute (Cocosco et al., 1997) and available at `www.bic.mni.mcgill.ca/brainweb`. Each simulation was a $1mm^3$ isotropic T1-weighted MRI volume with dimensions $181 \times 217 \times 181$, generated with varying noise level and intensity inhomogeneity. We tested our method on each combination of 1%, 3%, 5%, 7% or 9% noise levels together with 0%, 20% or 40% intensity nonuniformities. BrainWeb also provides the fuzzy tissue membership volumes, one for each tissue class, together with a discrete anatomical model of the simulated normal brain.

3.2. Manually segmented real MR data

20 normal MR brain data sets and their manual segmentations were obtained from the Internet Brain Segmentation Repository (IBSR), provided by the Center for Morphometric Analysis at Massachusetts General Hospital and available at `www.cma.mgh.harvard.edu/ibsr`. The data sets were acquired along the coronal axis with slice dimension of 256×256 and $1mm^2$

resolution. Interslice distance is $3mm$ and the number of slices for each volume varies between 60 and 65. The data sets have various levels of artifacts, as low contrast and relatively large intensity gradients, that further affects performance of the algorithm. CMA also provides expert tissue labellings of each brain into WM, GM, and CSF, together with reference similarity values for some classification techniques.

3.3. Cross sectional series of real MR scans

20 young healthy subjects (12 female, 8 male; age between 19 - 34 years), who underwent 4 scans at baseline and 4 more scans during a subsequent session after a short delay (less than 90 days), were randomly selected from the Open Access Series of Imaging Studies (OASIS) database (Marcus et al., 2007), available at `www.oasis-brains.org`. For each session, an average motion-corrected image (co-registered average of all available data) was used for our tests. The scans were T1-weighted Magnetization Prepared RApid Gradient Echo (MP-RAGE) in sagittal orientation with isotropic $1mm^3$ resolution ($256 \times 256 \times 128$ pixels). This data was used to assess the precision of the method when classifying partial volume voxels. We also tested the robustness when the method was integrated in our voxel-based cortical thickness estimation pipeline (Acosta et al., 2009), particularly when the detection of deep sulci was improved.

3.4. Error and similarity measures

To quantitatively evaluate performance of the method over both simulated and real MR data sets and compare these results with other well-known results, we used two different metrics: the root mean square (RMS) error for comparison of PV classification maps, and the Jaccard similarity measure for comparison of the corresponding crisp tissue segmentations. The RMS error

was used to quantify the differences between the fractional content calculated for each tissue and the corresponding values in the ground truth fuzzy membership images. As in (Shattuck et al., 2001), the RMS error between two images X and Y is calculated as

$$e_{RMS}(X,Y) = \sqrt{\frac{1}{|\Omega|}\sum_{k\in\Omega}|y_k - x_k|^2}$$

were Ω is the brain region, x_k and y_k are the image intensities at position k.

The Jaccard similarity metric, also known as the Tanimoto coefficient, measures the amount of overlap (agreement) between two images X and Y by taking the ratio between the size of their intersection and the size of their union:

$$J(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}$$

This metric yields values between 0 and 1, where 0 means complete dissimilarity and 1 stands for identical images.

4. Results and discussion

4.1. BrainWeb

Performance of our TPV method was firstly assessed on the simulated brain images from BrainWeb. One example of the resulting PV maps for WM, GM and CSF, compared with the available ground truth, on the synthetic brain volume, 3% noise level and 20% bias field, is depicted in Figure .8. Comparisons between our method and a classical MAP approach are shown in Figure .9 for the computed GMPVC fractional content map. It must be noted that compared to a classical MAP approach as in (Shattuck et al., 2001), the sulci were better delineated by introducing the topological constraints (Figure 9(g)). In this example, a deep sulci voxel with similar intensity to the average GM, will be classified as GM and not as a mixed GM/CSF voxel

unless anatomical constraints are introduced. The mean RMS error of fractional content over the entire BrainWeb data set significantly decreased to 6.1% ($p < 0.01$) for the obtained GMPVC map, as compared with the results reported in (Shattuck et al., 2001). Overall, a good agreement was shown between the computed PV maps and the ground truth, available as fuzzy tissue membership volumes. RMS errors for different noise and intensity nonuniformity levels are shown in Table .1. As expected, the computed error was robust to the bias field, which additionally validates the local averaging approach rather than the global one.

INSERT FIGURE .8 HERE

INSERT FIGURE .9 HERE

INSERT TABLE .1 HERE

The variability between different regions in the brain may affect the performance of PV classifiers (Chiverton and Wells, 2008). To illustrate this effect, we used the automated anatomical labeling (AAL) template (Tzourio-Mazoyer et al., 2002) to calculate the RMS error within each region as in (Chiverton and Wells, 2008). Averaged results for different levels of noise are shown in Figure .10. As a low variability with respect to the bias field was observed, the depicted value corresponds to the average over all the bias field levels (0%, 20% and 40%). The smallest errors appeared in the amygdala (42xx), the insula (30xx), the supplementary motor area (24xx) and the olfactory (25xx); while lower agreement was found in the basal ganglia (70xx), the middle occipital (52xx) and the parietal superior (61xx).

INSERT FIGURE .10 HERE

INSERT FIGURE .11 HERE

We also compared our TPV method with the results reported by Chiverton and Wells (2008) (GSR) and Shattuck et al. (2001) (SMAP). The results are depicted in Figure .11. Evidence suggests that the local average intensity strategy makes the classification more robust to bias field variations, and on average performs better than other methods for low levels of noise (1% to 7%) and bias field of 20%. We point out the fact that GSR does not explicitly take into account the bias field, hence its effect appears in the reported results.

4.2. Real MR Data

4.2.1. OASIS

The reproducibility was measured by applying the method to two of the MR scans from the same individual from the OASIS database. We compared the results with the MAP classifier as in (Shattuck et al., 2001). Significant improvements in GM PV estimation were brought by the topology correction. The reproducibility error decreased by 8.8% in GM and 8.5% in WM ($p < 0.001$), measured as the RMS between the PV maps obtained on the rigidly registered baseline and repeat scans. Likewise, when comparing the crisp segmentations obtained by thresholding by 0.5 the baseline and repeat GM PV maps, the Jaccard similarity measure increased by 3.5% in GM. To compute crisp segmentations, each mixed voxel was assigned to the tissue class with the highest fractional content and the obtained segmentation were subsequently compared.

4.2.2. IBSR

Our method was also compared with both TMCD (trimmed minimum covariance determinant) (Tohka et al., 2004) and MMC (mixture model clus-

tering) (Noe and Gee, 2001) on the IBSR data sets. Since the ground truth is available as manual segmentations performed by clinical experts, we compared the segmentations obtained from the crisped PV maps. Figure .12 shows an example of the ground truth provided by IBSR and a hard segmentation calculated after applying our method. Figure 13(b) depicts the results of the comparison for the GM in the 20 normal subjects. As in (Chiverton and Wells, 2008), results of manual expert segmentation and pure tissue classification presented by Ibrahim et al. (2006) (HMM, hidden Markov model) were included for reference. Significant improvements in GM classification were demonstrated using the TPV, compared to a MAP classifier. The similarity measure (Jaccard) was improved by 8.7% in GM and 2.6% in WM ($p < 0.001$).

INSERT FIGURE .12 HERE

Poor similarity results were obtained in 5 cases, which exhibited strong shading artifacts that impeded a reliable GM and WM classification. Similar findings were presented in (Noe and Gee, 2001), who excluded them from the analysis. We also observed that the anisotropy in the images biased the computation of the local averages. Table .2 summarizes the mean ($\pm$ standard deviation) of the Jaccard similarity values for each method, excluding the volumes with too severe intensity inhomogeneity. In average, our TPV method performed better for WM and GM compared to the others, excepting averaged GM segmentation against (Noe and Gee, 2001). It must be noted that when the PV maps were used to generate the crisp segmentations, the mixed GM/CSF voxels in deep sulci with fractional content above 0.5 might be wrongly reclassified as GM. Under those conditions, the contribution of topology correction in the segmentation can not be fully and

accurately validated with this experiment. Nonetheless, we report these results for completeness.

INSERT TABLE .2 HERE

INSERT FIGURE .13 HERE

4.3. Computational performance

On each image of the BrainWeb data set, after the initial MAP segmentation, the topology correction and PV fractional content estimation takes less than 10 minutes. For the OASIS data sets, the procedure takes about 9 minutes, while for the IBSR images the topology correction and PV fractional content estimation takes less than 4 minutes. Operations were encoded in a single-thread application and then executed in a standard Intel Core 2 Duo (3.00GHz, 2 GB RAM) machine running Linux.

4.4. Deep sulci cutting and cortical thickness estimation on real data

We integrated the proposed sulci detection and improved partial volume classification methods to our cortical thickness estimation pipeline (Acosta et al., 2009), as depicted in Figure .14. Then, we computed the thickness, at two different acquisition times, for the same 20 young healthy subjects from the OASIS database (Marcus et al., 2007) used in the experiment described in Section 3.3. The reproducibility was assessed by using the Pearson correlation coefficient for each Region Of Interest (ROI) of the AAL template (Tzourio-Mazoyer et al., 2002), excluding the cerebellum and subcortical *nuclei* from the analysis.

INSERT FIGURE .14 HERE

Thickness estimation with the proposed method (TPV) showed a higher reproducibility compared with the measure performed after partial volume classification using (Shattuck et al., 2001). As can be seen in Figure .15, the differences in cortical thickness between scans were reduced after applying the TPV. The Pearson correlation coefficient was 0.915 in average and a paired t-test did not reveal any significant differences between the two measurements ($p < 0.1$). Also, the difference between scans was decreased by 13.7% in average, as shown in Table .3.

INSERT FIGURE .15 HERE

INSERT TABLE .3 HERE

By using the proposed method, we found a mean ($\pm$ std. dev.) cortical thickness over the whole brain of $2.08mm$ ($\pm$ 0.11) for all the subjects, which is within the accepted range of cortical thickness for healthy young adults. In previous studies, when the PV is not taken into account as in (Yezzi and Prince, 2003), the computed mean thickness for the same population was $4.69mm$ ($\pm$ 0.11). And when the PV classification method proposed by (Shattuck et al., 2001) is used, without any topology correction, the computed mean thickness was $3.06mm$ ($\pm$ 0.25); using those same PV maps, but correcting the topology problems, decreases the mean thickness to $2.75mm$ ($\pm$ 0.17).

INSERT FIGURE .16 HERE

Fig. .16 depicts in histograms the impact of the topology correction and the accurate PV estimation on the cortical thickness calculation task. The higher thickness values produced after the first PV classification dissapeared when the topology of GM is corrected and the accurate PV value is computed

with the TPV. Fig. 16(a) shows the histogram of the average thickness for the 20 MR subjects before any topological modifications, after the topology correction and with TPV. Fig. 16(b) depicts the differences for each of the cortical thickness histograms between Step 1 and Step 2, illustrating the improvement after the TPV. The number of voxels above $4mm$ in average has beed dramatically reduced. Fig. 16(c) shows the differences between topology corrections and TPV, in average in this further step the number of voxels above $2.5mm$ has been reduced.

5. Conclusion

We have described a simple and fast technique to improve PV estimation of brain tissues from T1W MRI. It improves the detection of hidden mixed voxels in deep sulci by correcting for the topology errors in the segmentation and uses local averages to better estimate the fractional content. We show that fractional tissue content estimation can be improved for low levels of noise and regardless the intensity inhomogeneity, resulting in superior brain tissue segmentations.

Topology correction improved the classification of mixed voxels in opposed banks of buried sulci by assuming GM as a continuous layer following the WM, with the topology of a filled sphere. Local modelling of tissue intensities helps to overcome the issues related with local intensity inhomogeneity and tissue MR properties across the image. Even with a preprocessing stage to correct the intensity inhomogeneities, pure cortical tissues show different intensity levels in the MRI. This suggests that the tissue properties are different depending on the region of the brain. Accuracy and precision were demonstrated and comparisons with other methods showed comparative performance with simulated and real MR data.

We demonstrated the usefulness of the method to improve the accuracy of the cortical thickness estimation. By labelling mixed GM/CSF voxels in deep sulci and by recomputing a spatially compensated PV map, the measure of thickness in difficult regions is improved. Our method showed a high reproducibility on real data, with an extremely good agreement between the baseline and repeat scans. The computed values of thickness for young adults are similar to the ones reported previously in the literature. In the future, we plan to use our technique on clinical data to study cortical atrophy in Alzheimer's disease and other neurodegenerative diseases. We intend also to develop voxel-based techniques for inter-subject comparisons, a challenging issue given the large anatomical variability between patients.

AppendixA

Topology preservation and homotopic transformations

Homotopic transformations are topology-preserving procedures that consist of sequentially deleting or adding *simple points*. This operation works only on binary images, such as the pure tissue segmentations, where each voxel is considered as a point. Informally, a simple point of an object X is a point that can be added or removed from X without changing the topological characteristics of X. It is possible to locally characterize simple points in 3D using two *topological numbers* T and Tb (Bertrand and Malandain, 1994).

Thus, skipping some technical details, let $A(x)$ be the set of points of $X \setminus \{x\}$ lying in a neighborhood of x, and let $Ab(x)$ be the set of points of the complement of X (background) lying in a neighborhood of x. Then, $T(x)$ (resp. $Tb(x)$) is the number of connected components of $A(x)$ (resp. $Ab(x)$). A point x is simple if and only if $T(x) = Tb(x) = 1$. Topological numbers are useful for classifying points of an object X based on local topological characteristics: for example, a point x such that $Tb(x) > 1$ characterizes a region of the object which separates (locally) its background into several parts.

Based on these notions, given an object X, a subset I of X and a priority function P, Algorithm 1 computes an *homotopic erosion of X constrained by I*, that is, an object that is topologically equivalent to X, that contains I and that has no simple point outside I. In this algorithm, the priority function P is usually chosen as the inverse of the distance to I, in order to select in the first place the points that are farthest to the set I. This choice will be assumed in the remaining operations.

Applying Algorithm 1 to the complementary sets of X and I, then in-

Algorithm 1 Homotopic erosion of X constrained by I

repeat

Select $x \in X \setminus I$ such that $P(x)$ is minimal

if x is simple for X **then**

$X = X \setminus \{x\}$

end if

until stability

verting the result, yields an *homotopic dilation of X constrained by I*. In a similar way, Algorithm 2 (Bertrand and Couprie, 2007) computes a *surface skeleton of X* which contains medial surfaces of the original object (provided that the priority function P is a distance map of X).

Algorithm 2 Surface skeleton of X

Let C be a null image

repeat

Select $x \in X$ such that x is simple for X, $C(x) == \emptyset$ and $P(x)$ is minimal

$X = X \setminus \{x\}$

for all y in the neighborhood of x **do**

if $Tb(y) > 1$ **then**

$C(y) = 1$

end if

end for

until stability

References

Acosta, O., Bourgeat, P., Zuluaga, M. A., Fripp, J., Salvado, O., Ourselin, S., 2009. Automated voxel-based 3D cortical thickness measurement in a combined lagrangian-eulerian PDE approach using partial volume maps. Medical Image Analysis 13, 730–743.

Bazin, P.-L., Pham, D. L., 2005. Topology correction using fast marching methods and its application to brain segmentation. In: Medical Image Computing and Computer-Assisted Intervention 2005. Lecture Notes in Computer Science. Springer Verlag, pp. 484 – 491.

Bazin, P.-L., Pham, D. L., 2007. Topology-preserving tissue classification of magnetic resonance brain images. IEEE Transactions on Medical Imaging 26 (4), 487–496.

Bertrand, G., Couprie, M., 2007. Transformations topologiques discrètes. In: Coeurjolly, D., Montanvert, A., Chassery, J.-M. (Eds.), Géométrie discrète et images numériques. Hermès, pp. 187–209.

Bertrand, G., Malandain, G., 1994. A new characterization of three-dimensional simple points. Pattern Recognition Letters 15 (2), 169–175.

Besag, J., 1986. On the statistical analysis of dirty pictures. Journal of the Royal Statistical Society 48, 259–302.

Chiverton, J., Wells, K., 2008. Adaptive partial volume classification of mri data. Physics in Medicine and Biology 53 (20), 5577–5594.

Choi, H., Haynor, D., Kim, Y., 1991. Partial volume tissue classification of multichannel magnetic resonance images - a mixel model. IEEE Transactions on Medical Imaging 10 (3), 395–407.

Cocosco, C., Kollokian, V., Kwan, R.-S., Evans, A., 1997. Brainweb: Online interface to a 3D MRI simulated brain database. NeuroImage 5 (4), S425.

Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C., 2008. An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images. IEEE Transactions on Medical Imaging 27 (4), 425–441.

Dale, A., Fischl, B., Sereno, M., 1999. Cortical surface-based analysis I: Segmentation and surface reconstruction. NeuroImage 9 (2), 179–194.

Fischl, B., Dale, A., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences of the United States of America 97 (20), 11050–11055.

Fischl, B., Sereno, M., Dale, A., 1999. Cortical Surface-Based Analysis II: Inflation, Flattening, and a Surface-Based Coordinate System. NeuroImage, 195–207.

Florent Ségonne, J. P., Fischl, B., 2007. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. IEEE Transactions on Medical ImaginG 26 (4), 518–529.

Han, X., Xu, C., Braga-Neto, U., Prince, J., 2002. Topology correction in brain cortex segmentation using a multiscale, graph-based algorithm. IEEE Transactions on Medical Imaging 21 (2), 109–121.

Ibrahim, M., John, N., Kabuka, M., Younis, A., 2006. Hidden markov models-based 3d mri brain segmentation. Image and Vision Computing 24 (10), 1065–1079.

Jaume, S., Rondao, P., Macq, B., 2005. Open Topology: A Toolkit for Brain Isosurface Correction. In: MICCAI Open Source Workshop.

Kim, J., Singh, V., Lee, J., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J., Kim, S., Evans, A., 2005. Automated 3D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. NeuroImage 27 (1), 210–221.

Kriegeskorte, N., Goebel, R., 2001. An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical mr volumes. Neuroimage 14, 329–346.

Laidlaw, D., Fleischer, K., Barr, A., 1998. Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms. IEEE Transactions on Medical Imaging 17 (1), 74–86.

Mangin, J.-F., Frouin, V., Bloch, I., Régis, J., López-Krahe, J., 1995. From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. Journal of Mathematical Imaging and Vision 5 (4), 297–318.

Marcus, D. S., Wang, T. H., Parker, J., Csernansky1, J. G., C., M. J., L., B. R., 2007. Open access series of imaging studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. Journal of Cognitive Neuroscience 19, 1498–1507.

McInerney, T., Terzopoulos, D., 1999. Topology Adaptive Deformable Surfaces for Medical Image Volume Segmentation. IEEE Transactions on Medical Imaging 18 (10), 840–850.

Noe, A., Gee, J., 2001. Partial volume segmentation of cerebral MRI scans with mixture model clustering. In: International Conference on Information Processing in Medical Imaging. Vol. 2082 of LNCS. Springer, pp. 423–430.

Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N., 2001. Reconstructing a 3D structure from serial histological sections. Image and Vision Computing 19 (1), 25–31.

Salat, D.H., Lee, S.Y., van der Kouwe, A.J., Greve, D.N., Fischl, B., Rosas, H.D., 2009. Age-associated alterations in cortical gray and white matter signal intensity and gray to white matter contrast. NeuroImage 48 (1), 21–28.

Santago, P., Gage, H., 1993. Quantification of MR brain images by mixture density and partial volume modeling. IEEE Transactions on Medical Imaging 12 (3), 566–574.

Ségonne, F., 2008. Active contours under topology control - genus preserving level sets. International Journal of Computer Vision 79 (2), 107 – 117.

Shattuck, D., Sandor-Leahy, S., Schaper, K., Rottenberg, D., Leahy, R., 2001. Magnetic resonance image tissue classification using a partial volume model. Neuroimage 13 (5), 856–876.

Tohka, J., Zijdenbos, A., Evans, A., 2004. Fast and robust parameter estimation for statistical partial volume models in brain MRI. NeuroImage 23 (1), 84–97.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. NeuroImage 15 (1), 273–289.

Van Leemput, K., Maes, F., Vandermeulan, D., Suetens, P., 1999. Auto-

570 mated model-based bias field correction of MR images of the brain. IEEE
571 Transactions on Medical Imaging 18 (10), 885–896.

572 Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P., 2003. A unifying
573 framework for partial volume segmentation of brain MR images. IEEE
574 Transactions on Medical Imaging 22 (1), 105–119.

575 Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2007. Non-parametric
576 diffeomorphic image resgistration with the demons algorithm. In: Medi-
577 cal Image Computing and Computer-Assisted Intervention. Vol. 4792 of
578 LNCS. Springer Berlin / Heidelberg, Brisbane, Australia, pp. 319–326.

579 Yezzi, A.J., Prince, J.L., 2003. An Eulerian PDE approach for computing
580 tissue thickness. IEEE Transactions on Medical Imaging 22 (10), 1332–
581 1339.

582 Zhou, Q.-Y., Ju, T., Hu, S.-M., 2007. Topology repair of solid models using
583 skeletons. IEEE Transactions on Visualization and Computer Graphics
584 13 (4), 675–685.

Tables

	Intensity inhomogeneity					
	0%		20%		40%	
Noise	WM	GM	WM	GM	WM	GM
1%	0.129	0.130	0.129	0.131	0.125	0.132
3%	0.139	0.142	0.140	0.141	0.140	0.142
5%	0.174	0.174	0.172	0.171	0.170	0.171
7%	0.214	0.216	0.210	0.213	0.208	0.212
9%	0.251	0.261	0.245	0.258	0.242	0.256

Table .1: Fractional content RMS error on BrainWeb.

	MMC (Noe and Gee, 2001)	TMCD (Tohka et al., 2004)	TPV
WM	0.648 (± 0.198)	0.696 (± 0.050)	0.701 (± 0.042)
GM	0.753 (± 0.120)	0.697 (± 0.064)	0.708 (± 0.045)

Table .2: Mean (± standard deviation) of Jaccard similarity index for each method.

	Correlation coefficient		Differences between scans	
Brain lobule	SMAP	TPV	SMAP	TPV
Frontal	0.922	0.930	0.090	0.090
Limbic	0.901	0.883	0.158	0.121
Occipital	0.902	0.904	0.101	0.063
Parietal	0.906	0.920	0.058	0.060
Temporal	0.932	0.938	0.105	0.106
Average	0.912	0.915	0.102	0.088

Table .3: Pearson correlation coefficient and differences between scans for the OASIS dataset, grouped by brain lobules.

Figures and Legends

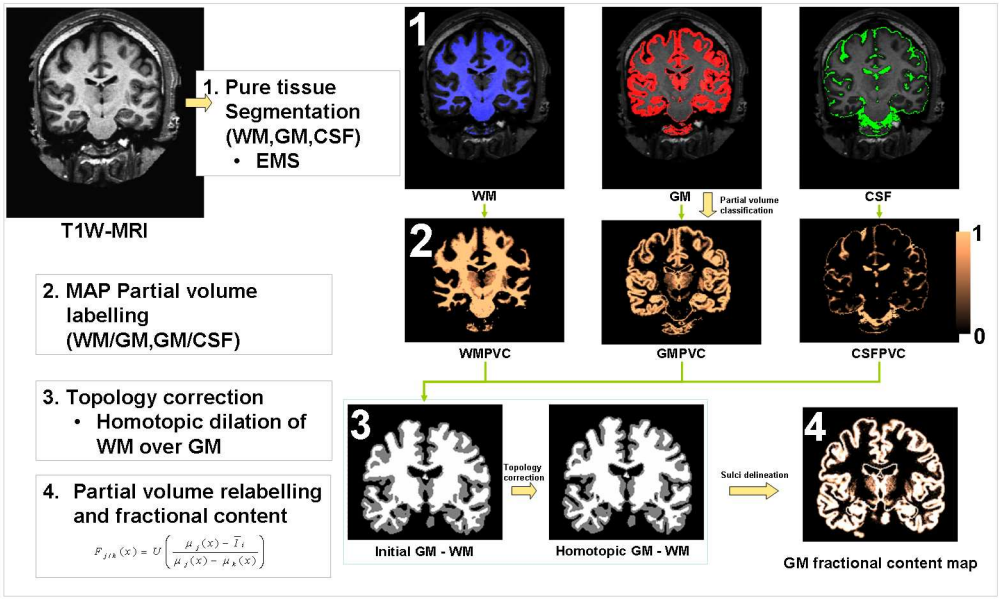


Figure .1: Overall process for topology-corrected PV estimation in MR images

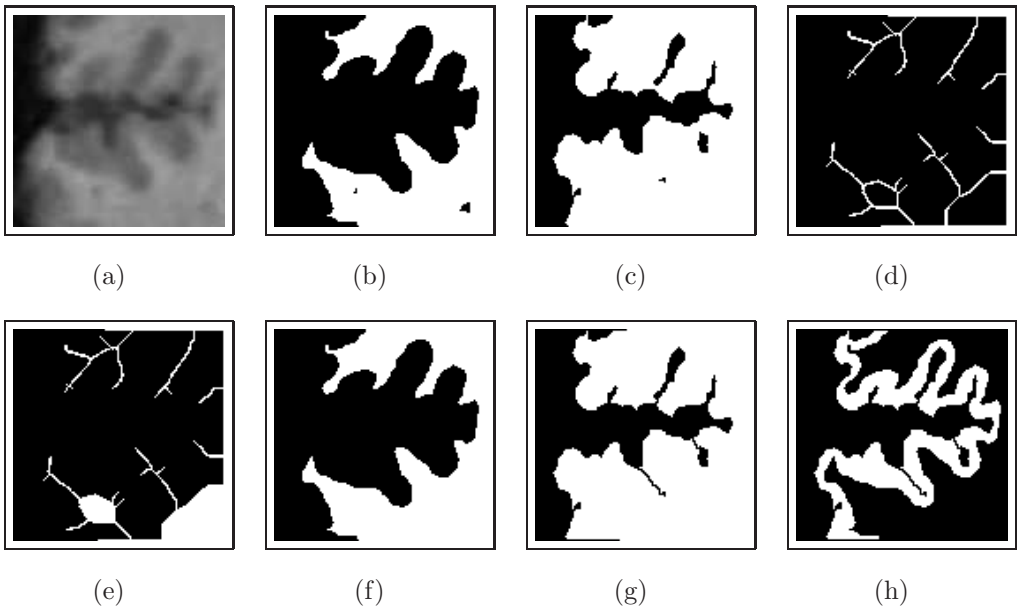
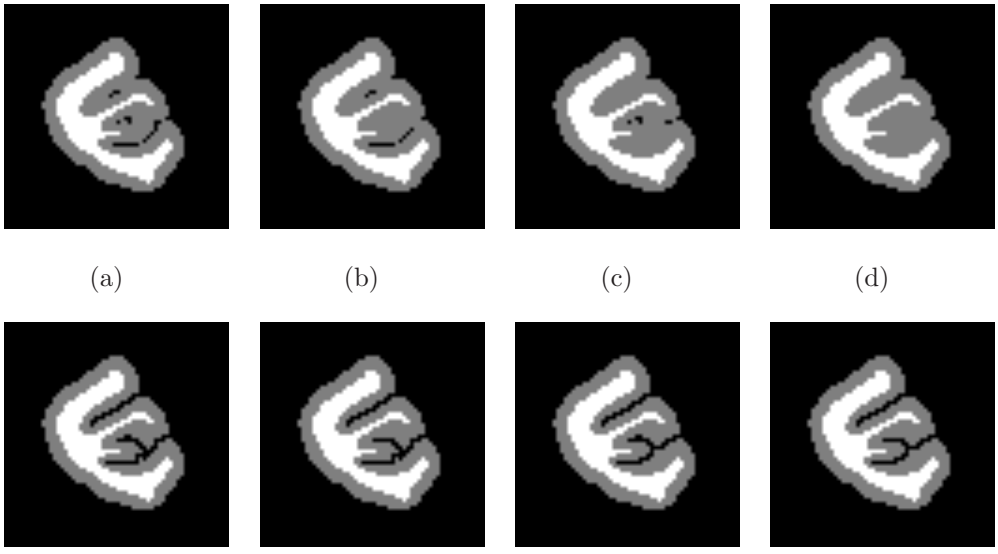


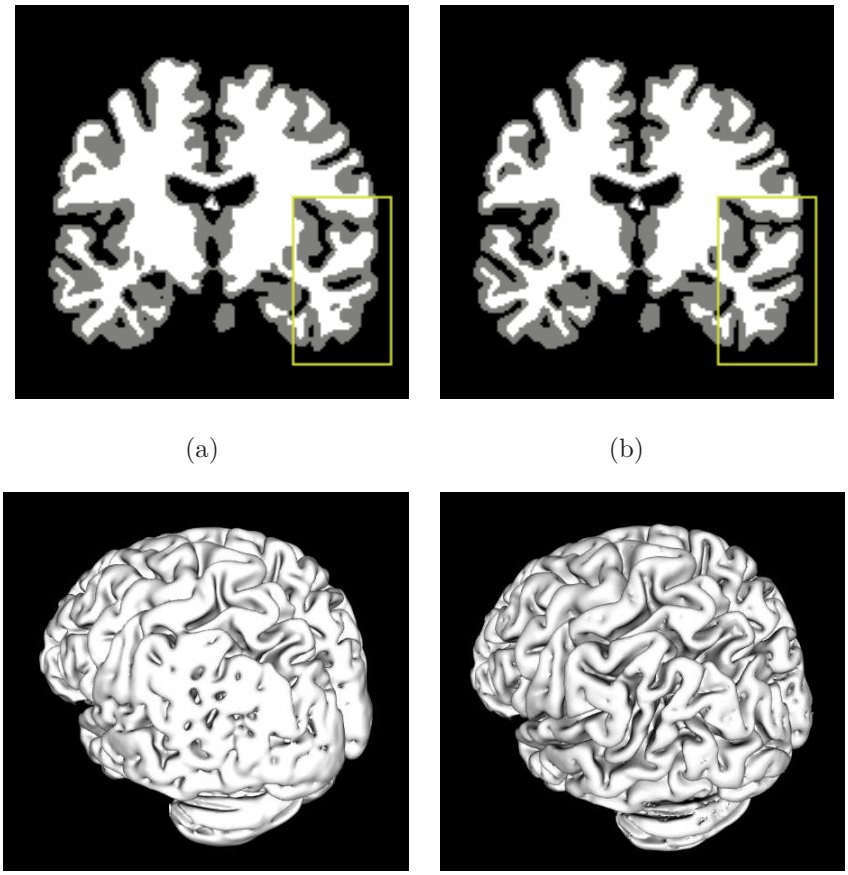
Figure .2: (a): original grayscale image. (b): segmented white matter (set **WM**). (c): segmented white and gray matter (set **WM** $\cup$ **GM**). (d): surface skeleton of **WM** (set **SK**). (e): seed set (set **S**). (f): corrected white matter (set **SWM**). (g): corrected white and gray matter formed by further homotopic dilation. (h): corrected gray matter (final result) formed by substracted images (g) and (f).



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure .3: First row (a) - (d): Different initial configurations of a synthetic phantom. Second row (e) - (h): Corresponding topologically corrected WM-GM segmentations.



(a) (b)

(c) (d)

Figure .4: (a) Initial and (b) topologically corrected WM-GM segmentations, highlighted within the rectangle; (c) marching cubes reconstruction of GM before and (d) after the topology correction procedure.

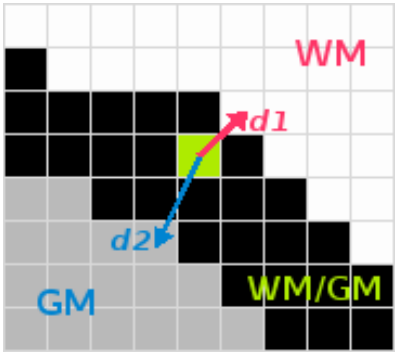


Figure .5: Schematic view of the local tissue averages for a given mixed voxel, where $d1$ and $d2$ relates to the closest voxels in the pure tissues.

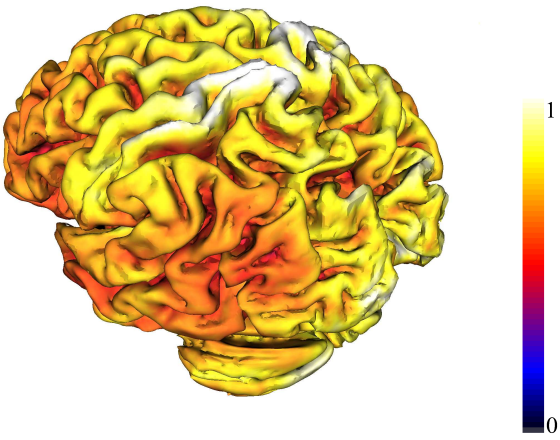


Figure .6: Averaged intensity within the connected components of the pure GM, computed as the interquartile mean (IQM) within a $5mm$ radius sphere on an OASIS example data, normalized by the Maximum of intensity. The differences between the regions clearly appear. Thus, GM tissue intensity will be different between the regions and global homogeneity assumptions will slightly bias the computation of partial volume.

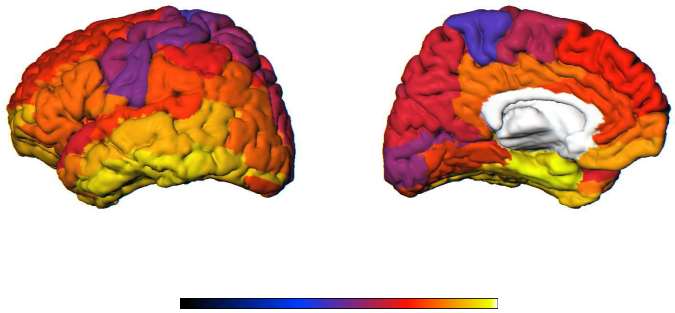
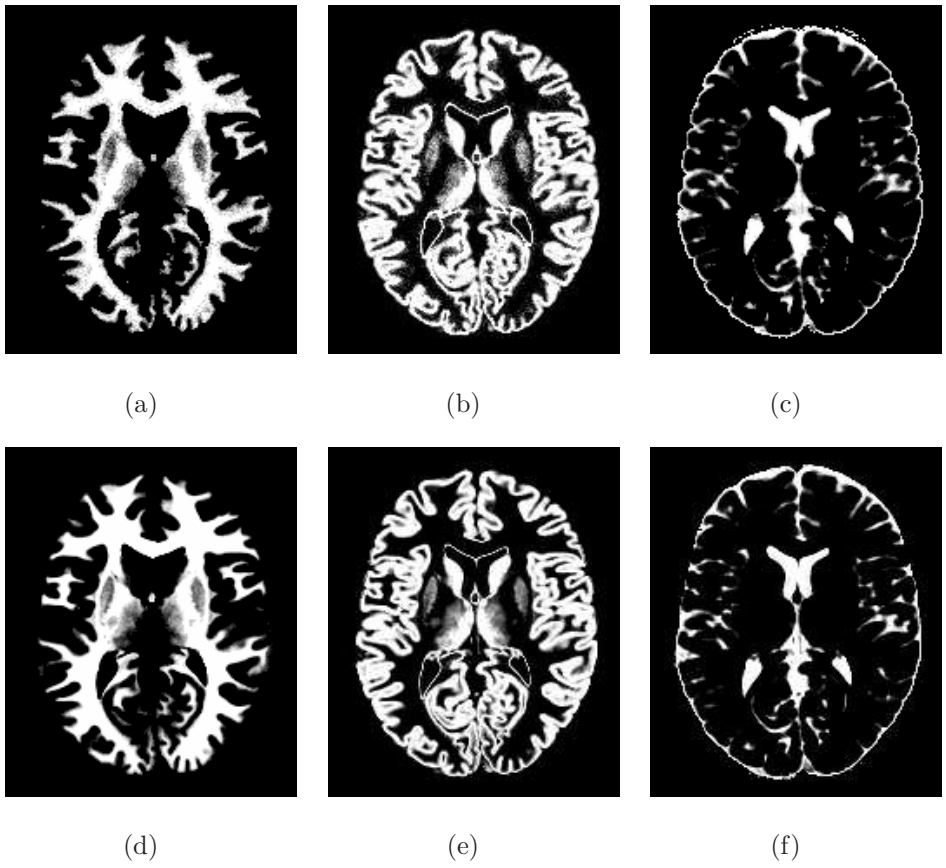


Figure .7: AAL template showing the regional differences incontrast between WM and GM over the surface, by calculating the ratio $\frac{\mu_{\mathrm{WM}}-\mu_{\mathrm{GM}}}{\mu_{\mathrm{GM}}-\mu_{\mathrm{CSF}}}$. Darkest colours indicate bigger ratios, light colours indicate small values. Left: lateral and Right: medial views.



(a) (b) (c)

(d) (e) (f)

Figure .8: Partial volume segmentation of a simulated BrainWeb volume (3% noise, 20% bias field). PV maps for (a) WM, (b), GM (c) and CSF. Ground truth: (d) WM, (e), GM and (f) CSF.

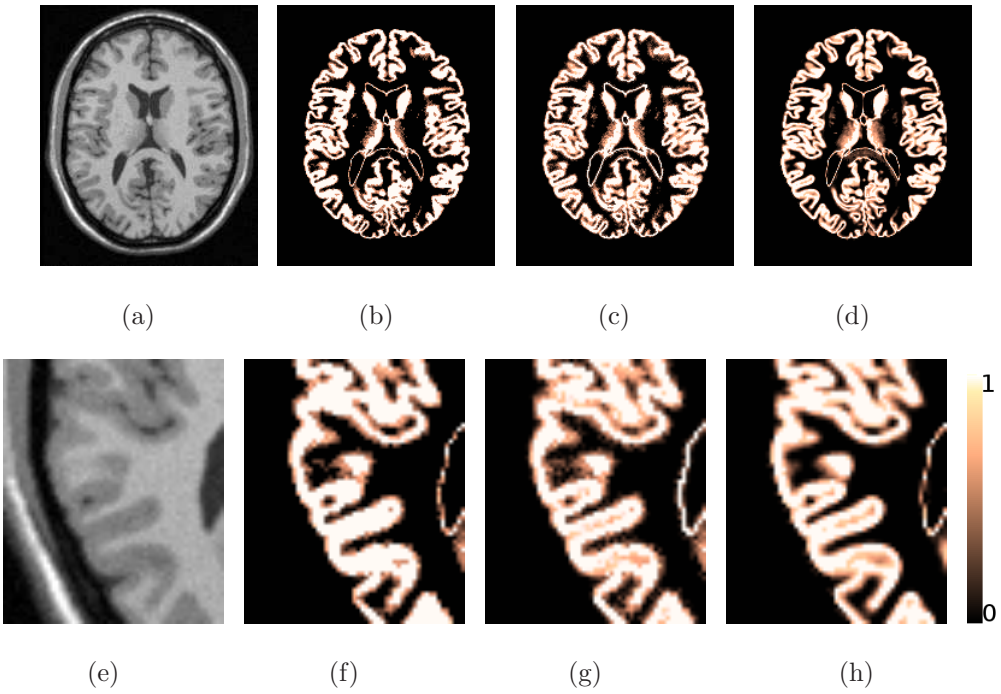
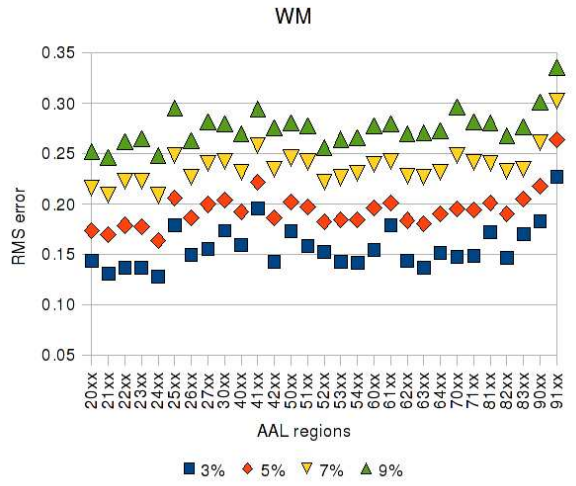
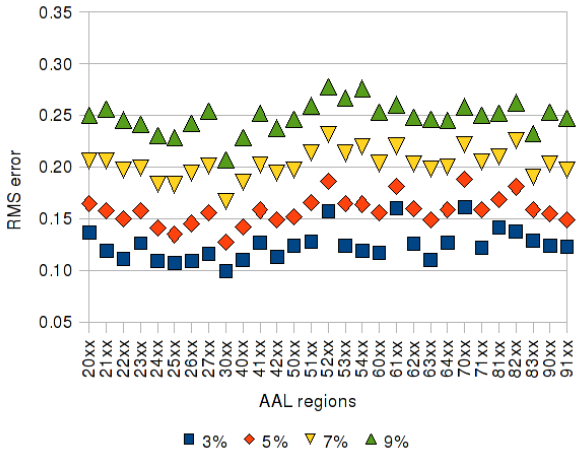


Figure .9: Example of PV estimation of a simulated BrainWeb volume (3% noise, 20% bias field). (a),(e) Original image, (b),(f) MAP PV estimation, (c),(g) Topologically-corrected PV, (a),(h) ground truth. In the detailed views we can observe the improvement in deep sulci, (g) relative to (f), brought by the topology correction.

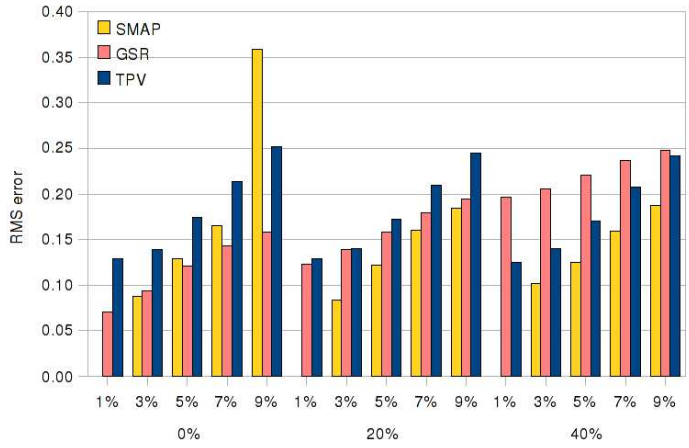


(a) WM

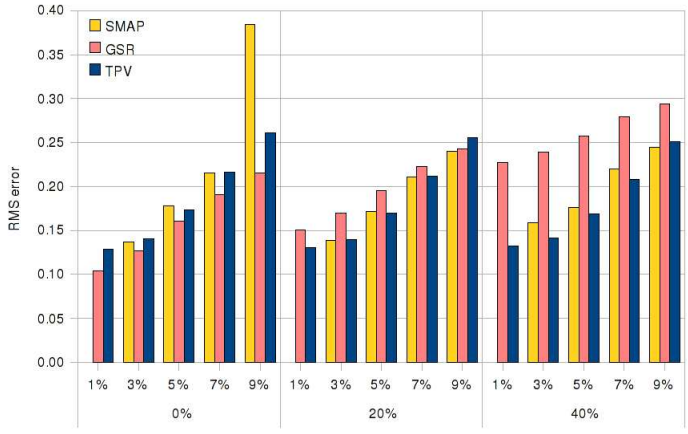


(b) GM

Figure .10: RMS error per AAL region (a) WM and (b) GM regions, for different noise levels using the same labels as (Chiverton and Wells, 2008).

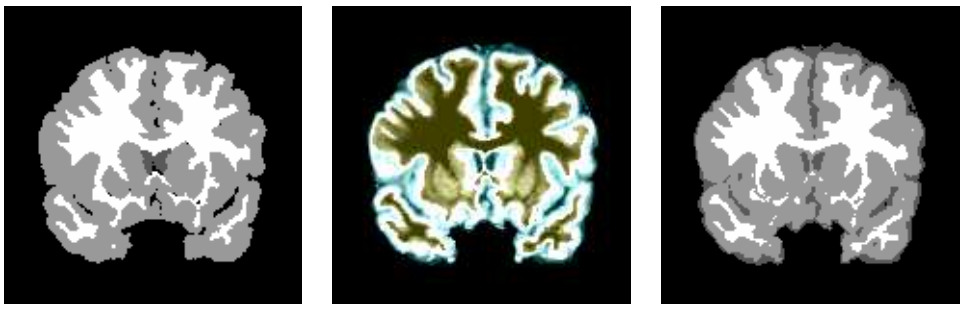


(a) WM



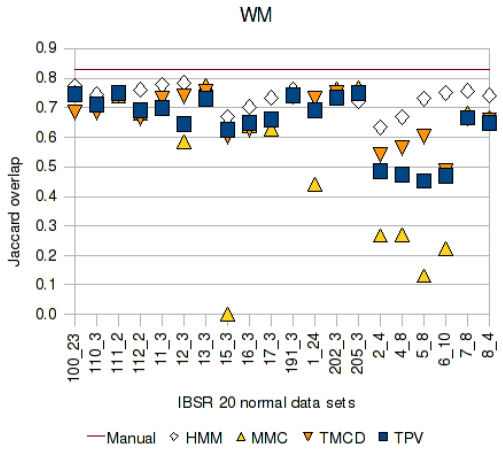
(b) GM

Figure .11: PV estimation errors for (a) WM and (b) GM on BrainWeb, for different noise and bias field levels. (SMAP results for 1% noise not publicly available)

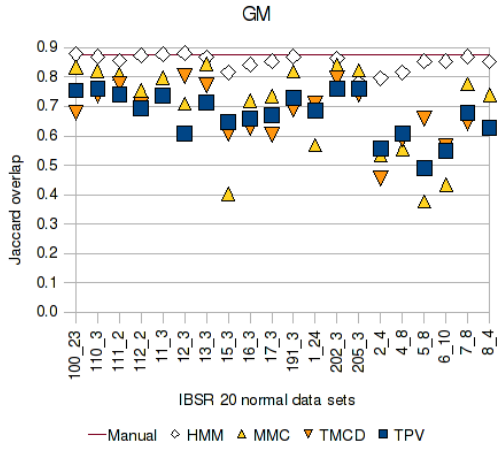


(a) (b) (c)

Figure .12: (a) IBSR Ground truth pure tissue classification. (b) Estimated PV maps (blue: GM/CSF, white: GM, yellow: GM/WM) and (c) computed crisp segmentation.

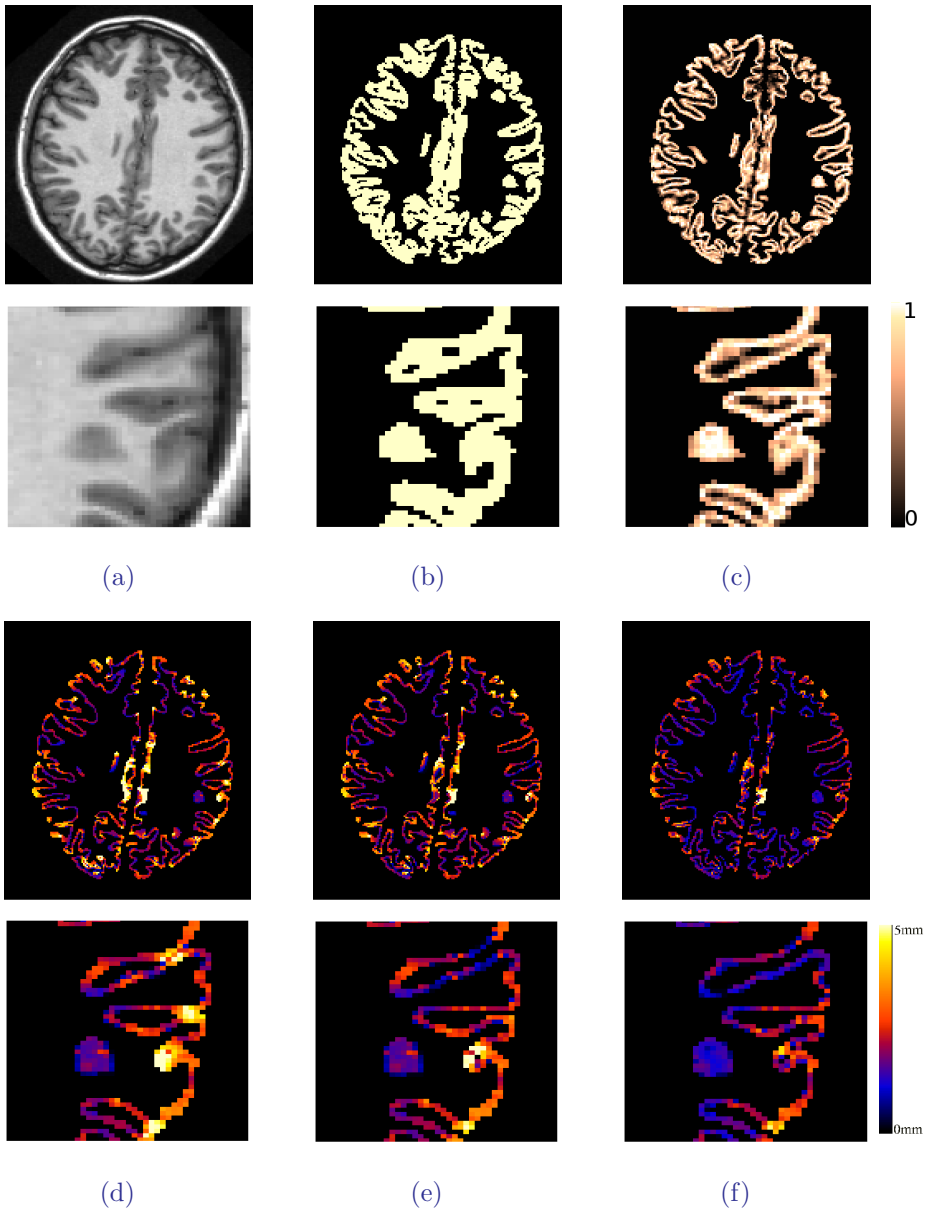


(a) WM



(b) GM

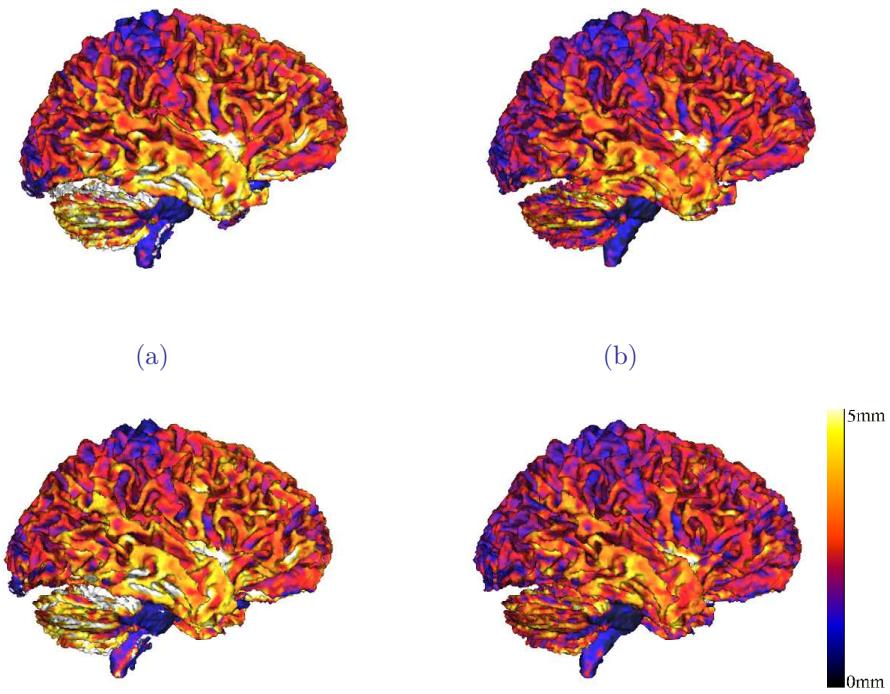
Figure .13: Jaccard similarity results for WM (a) and GM (b).



(a) (b) (c)

(d) (e) (f)

Figure .14: Example of cortical thickness estimation from MR. (a) Original T1-W MRI, (b) GM segmentation, (c) Topologically-corrected GM PV map. Cortical thickness maps (d) without any topology modifications, (e) after topology correction only, (f) after TPV. In the detailed views we can observe the improvement brought by the topology to delineate deep sulci zones, which allows an accurate measurement of the cortical thickness.



(a) (b)

(c) (d)

Figure .15: Surface representation of cortical thickness, computed at different steps for two scans of a single subject (OASIS). *Top row:* Scan 1, *Bottom row:* Scan 2. (a),(c) Without topology modifications, and (b),(d) with topologically-corrected GM PV map (TPV). Overall, we can observe the high values of thickness corrected with the TPV method.

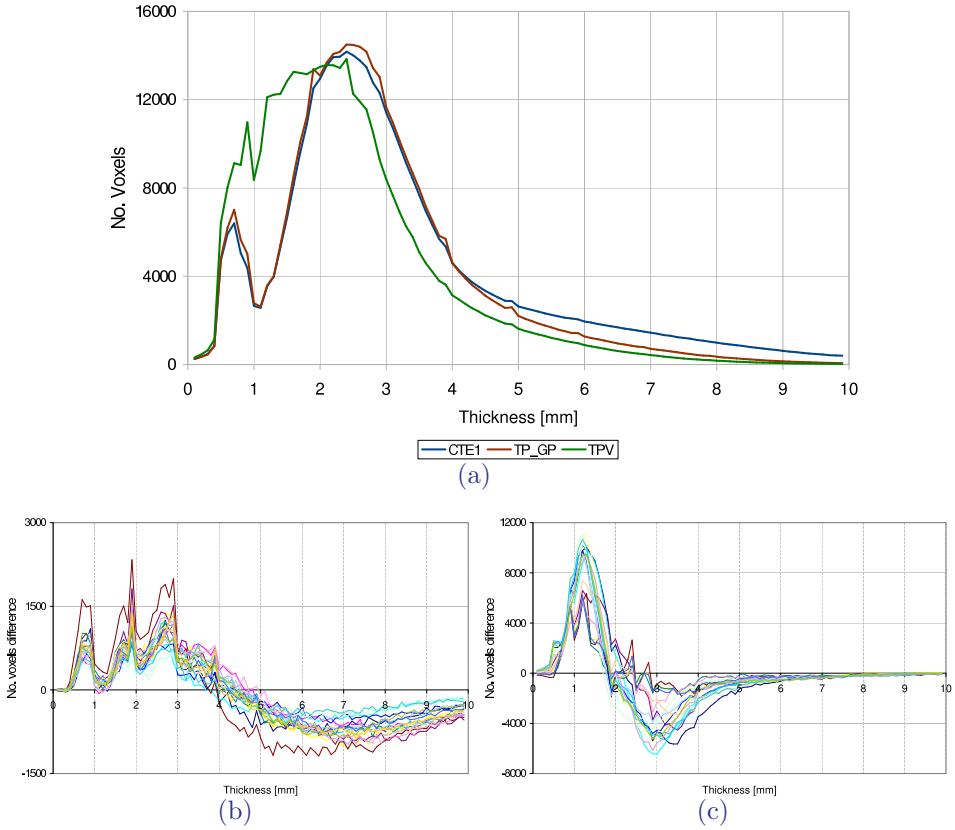


Figure .16: (a) Histogram of the average thickness for the 20 MR before topology correction (step 1), after topology correction (step 2) and with TPV. It is shown how the number of higher thickness voxels was reduced. (b) Differences in cortical thickness histograms between steps 1 and 2 for the 20 MR. This figure depicts the improvement after the topology. The number of voxels above $4mm$ in average has been dramatically reduced. (c) Differences between topology and TPV, in average the number of voxels above $2.5mm$ has been reduced consolidating the average thickness around 2.5 mm (typical value for young adults).