

Estimation of the partial volume effect in MRI

Miguel Ángel González Ballester*, Andrew P. Zisserman, Michael Brady

Robotics Research Group, Department of Engineering Science, University of Oxford, Parks Rd., Oxford OX1 3PJ, UK

Received 12 February 2001; received in revised form 3 August 2001; accepted 30 November 2001

Abstract

The partial volume effect (PVE) arises in volumetric images when more than one tissue type occurs in a voxel. In such cases, the voxel intensity depends not only on the imaging sequence and tissue properties, but also on the proportions of each tissue type present in the voxel. We have demonstrated in previous work that ignoring this effect by establishing binary voxel-based segmentations introduces significant errors in quantitative measurements, such as estimations of the volumes of brain structures. In this paper, we provide a statistical estimation framework to quantify PVE and to propagate voxel-based estimates in order to compute global magnitudes, such as volume, with associated estimates of uncertainty. Validation is performed on ground truth synthetic images and MRI phantoms, and a clinical study is reported. Results show that the method allows for robust morphometric studies and provides resolution unattainable to date.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Partial volume effect; Morphometry; Confidence bounds; Segmentation

1. Introduction

Medical image analysis has developed rapidly in recent years. Initially, most work aimed at generating *qualitative* information in the form of images that could engage the clinician's perceptual abilities. Increasingly, however, there is a need for medical image analysis to deliver *quantitative* information, for example about the size of a tumour, or the extent to which an anatomical structure has responded to a drug therapy. Since images are intrinsically noisy and sample continuous anatomical structures, results can never be given with complete precision. Error bounds on measurements are inevitably necessary, as they are in all engineering applications. However, most published medical image analysis methods develop segmentations and then derive measurements of certain structures or lesions, but confidence bounds on such measurements are rarely provided. The need for precise quantitative measure-

ments of structures, with associated error bounds, is particularly important in magnetic resonance image (MRI) analysis, because the voxel size of most MRI image acquisitions is significant at the scale of the object to be measured. The MRI intensity in a particular voxel depends on the entire contents of the corresponding anatomical volume and the sequence that is used. If only a single tissue type is present in the voxel, the signal intensity will be characteristic of that tissue type. However, if more than one tissue type is present, the signal will be a combination of the contributions of the different tissues. This is known as the Partial Volume Effect (PVE), and this paper presents a method for estimating it, and correcting segmentations on the basis of that estimate.

In previous research (Gonzalez Ballester et al., 2000), we described a methodology for morphometric studies of brain structures from MRI data sets, and this paper is a continuation of that research. The method in (Gonzalez Ballester et al., 2000) was based on a combination of statistical and geometrical information to perform segmentations and derive shape measurements. PVE was handled by detecting voxels that do not conform to the expected statistical distributions of the tissues present in

*Corresponding author. Present address: Projet Epidaure, INRIA, 2004 Route des Lucioles BP 93, 06902 Sophia Antipolis Cedex, France.

E-mail address: Miguel.Gonzalez@sophia.inria.fr (M.Á. González Ballester).

the image, then building surface bounds that isolate such voxels and which enforce geometric continuity via an active shape model. In that way, PVE voxels are considered uncertain, and upper and lower estimates on the shape measurements (e.g. volume) are derived from the inner and outer surface bounds. The particular choice of the shape model that we use (a mesh of triangular Gregory–Bezier patches) enforces geometric continuity and reduces the width of the confidence interval with respect to voxel-based measurements.

The results reported for the method demonstrate clearly that the effect of discrete sampling in boundary locations is crucial, and that it can lead to volume measurements with errors in the range 20–60% of the volume of the object being measured, particularly for complex brain structures for which the ratio of surface area to volume is relatively large¹.

In this paper, we propose decoding the information contained in PVE voxels in order to segment objects to sub-voxel accuracy and thereby to provide narrower confidence bounds on the measurements. More precisely, we develop a statistical framework within which to model the partial volume effect and then build a statistical distribution for it. Once that is in place, the uncertainties inherent in each PVE voxel are propagated in order to construct a distribution on the measurements derived from the data (e.g. volume, area, etc.). Since the approach is statistical, the end product is not a single value for the shape descriptor of choice—volume will be used for the examples reported in this paper, though the approach is more generally applicable—rather, it is a statistical distribution from which we can obtain its probability density function, mode, mean, etc. In addition, confidence bounds may be computed to a certain probability value.

Section 2 provides a review of the existing literature regarding PVE estimation. Section 3 introduces the proposed PVE model, and then Section 4 provides a more detailed description for the special case of Gaussian tissue models. Section 5 shows how confidence bounds on the estimated proportion of each tissue in a PVE voxel can be established, while Section 6 presents a methodology to propagate such local confidence bounds in order to develop bounds for global shape descriptors (e.g. volume). Validation on synthetic images is described in Section 7. Section 8 discusses possible complications when applying the method to real MRI data sets, and Sections 9, 10 and 11 address these issues. Section 12 shows the results of experiments performed on MRI phantoms, and Section 13 describes a clinical study performed following our method. Finally, Section 14 provides discussion and conclusions.

¹See also (Niessen, 1997) for a study of the effect of PVE using synthetic brain phantoms. Percentages of volume encompassed by PVE voxels are consistent with the estimates provided by our method.

2. Literature review

The quantitative limitations imposed by the partial volume effect have received growing attention in the last decade. Some methods consider PVE as a corrupting factor and try to *correct* for it. On the other hand, the work described in this and other recent works is based on using PVE as an important image feature, and the focus is on *estimating* it in order to obtain sub-voxel accuracy.

Roll et al. (1994) use an intensity thresholding segmentation approach to assign a tissue type tag to each voxel and establish an optimal segmentation threshold and a correction factor to take into account PVE. The particular application is MS lesion quantification. Although complicated by the fact that they take into account all voxels in the image, instead of only PVE voxels, their approach is rather simple, and concludes that the optimal threshold is halfway between lesion and background intensities.

Several authors have proposed approaches for estimating PVE by solving a linear system using the information provided by several imaging MR sequences. Soltanian Zadeh et al. (1993) develop a method for creating images with intensities proportional to the quantity of a certain tissue in each voxel, while maximising signal-to-noise ratio. However, it requires having at least the same number of (perfectly registered) image sequences as the number of tissues that we are interested in segmenting. Similarly, Thacker et al. (1998) set up a linear system with the intensities from two image sequences in order to segment grey matter, white matter, and CSF from neurological data sets, and provide a method for estimating the expected accuracy of the results. Choi et al. (1991) argue that methods that combine different sequences linearly are very sensitive to noise, and they propose a Markov random field (MRF) prior as a way to smooth out the results. This introduces the typical problems of MRFs, i.e. determining the correct parameters and the weighting between the prior and the data, as well as huge computational cost. It must be noted that several imaging sequences are not always readily available, and in such cases a single-channel PVE estimation method is necessary.

Another approach is to model the statistical distribution of pure tissue and mixture voxels and to fit these distributions to the image. Santiago and Gage (1993) use a Gaussian model for the tissue classes based on a mean intensity and a common noise variance for all the tissues, and employ a uniform distribution to describe the behaviour of the proportion of each tissue in PVE voxels. Laidlaw et al. (1998) use the same tissue model assuming a single common variance for all tissues, and provide a sophisticated means for fitting the distributions to the histograms of the whole image and of single voxels, while constraining neighbour continuity. No quantitative comparison with other methods is provided, and visual comparison is performed with methods that do not give appropriate results for the example object. Therefore, it is

difficult to draw conclusions about the performance of the method. Unfortunately, this technique is extremely time-consuming, even though it was implemented on parallel hardware. Our method is highly related to these techniques, especially (Santago and Gage, 1993). The main differences are generality in tissue distributions (not limited to Gaussian distributions with common variance), prior segmentation of PVE voxels (we argue that a uniform distribution for the proportion of each tissue is only correct if non-PVE voxels are not included), and an in-depth consideration of the issue of error propagation in global measurements (volume, in this paper).

Marais (1999) proposes a different PVE estimation strategy by using a model of the width of each tissue in profiles traversing the brain surface. His work was motivated by the need to deal with sparse, low resolution data sets, and the assumptions taken to model anatomical structure are too simplistic for its use on high resolution MRI.

3. Mixture model

To simplify the presentation, the model will be presented only for PVE voxels containing two different types of tissue. The extension to multiple tissues is relatively straightforward though mathematically tedious. Note that in any case, this restriction only applies locally and does not limit the number of tissue types modelled for the whole image. This is because the two tissues present in PVE voxels in separate locations of the image may be entirely different. A second reason for imposing this restriction is that, in practice, cases in which more than two tissue types are present in a voxel are rare in brain studies. In short, although our model for PVE voxels can be extended to include such cases quite straightforwardly, the practical improvement expected from such an extension is questionable, and it could introduce inaccuracy (if a mixture model of three or more distributions is fitted to a PVE voxel which actually contains only two tissues) and computational overhead. This restriction to two tissue classes was also adopted in (Laidlaw et al., 1998), amongst others.

The distribution of intensities generated from samples of a tissue type j is assumed to follow a certain distribution $p_j(I)$. No assumptions are made regarding this distribution; it can be modelled either in terms of a known analytical statistical model or it can be constructed empirically from histograms derived from a suitable training set. In the next section, the model is elaborated for the special case of a Gaussian distribution, since it arises so frequently in the literature.

PVE arises from the presence of more than one tissue in a voxel, which, in accordance with the physics of MRI, generates an intensity value that depends linearly on the

proportion of each tissue in the voxel² (Rinck, 1993). Based on the assumption of just two tissue types present in PVE voxels, the intensity of a voxel is determined by the following linearly-weighted sum:

$$I^i = \alpha^i I_1^i + (1 - \alpha^i) I_2^i, \quad (1)$$

where i indexes all PVE voxels, and I_j^i is drawn from p_j , the distribution for tissue j , $j = \{1, 2\}$. In this context, α^i denotes the proportion of tissue 1 in the voxel i , $\alpha^i \in [0, 1]$.

The distributions for individual tissue types can be combined in order to build a mixture model, $p(I^i|\alpha^i)$. This distribution expresses the likelihood of a particular intensity value in a PVE voxel, given the proportion of each tissue in it (Fig. 1). The expression for this mixture model can be obtained either analytically (cf. next section) or by means of Monte Carlo simulation. See (Everitt and Hand, 1981) for further detail on mixture modelling.

In order to estimate PVE, we are interested in obtaining $p(\alpha^i|I^i)$, i.e. the statistical distribution of the proportion α of each tissue given the intensity of the PVE voxel i . Having this distribution, it is possible to compute its mode, which corresponds to the most likely value of α^i , as well as confidence intervals (cf. Section 5). Using Bayes' theorem, the expression for $p(\alpha^i|I^i)$ is

$$p(\alpha^i|I^i) = \frac{p(I^i|\alpha^i)p(\alpha^i)}{p(I^i)}. \quad (2)$$

The term $p(I^i)$ is a normalising constant, and is computed as

$$p(I^i) = \int p(I^i|\alpha^i)p(\alpha^i) d\alpha. \quad (3)$$

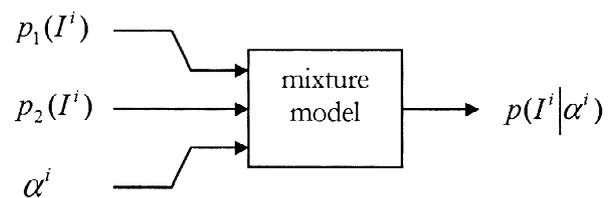


Fig. 1. A distribution for the intensities of PVE voxels, given the proportion of the two tissues present in it, is built based on the two tissue distributions p_1 and p_2 .

²This holds for most common imaging modalities. It should be noted, however, that inversion recovery sequences may present problems in cases where one of the tissues present in the PVE voxel has negative magnetisation, while the other tissue has positive magnetisation. In such a case, PVE voxels can have lower intensities than the tissues present in them. This is an uncommon case, and although theoretically possible, we did not find it in any of the images in our data base.

$p(\alpha^i)$ is the prior on α , and expresses the probability of a voxel i having a proportion α of tissue 1 (and therefore a proportion $1 - \alpha$ of tissue 2). In this work we propose that, if we consider only PVE voxels, this prior can be modelled as a uniform distribution in the range $[0 \dots 1]$.

$$p(\alpha^i) = 1, \quad \forall i. \quad (4)$$

The prior is thus reduced to a constant, transforming the ‘maximum a posteriori’ expression in (2) into a ‘maximum likelihood’ estimator. Other works assuming uniform priors include (Santago and Gage, 1993; Laidlaw et al., 1998).

Note that this assumption only holds for the set of PVE voxels. If all voxels in the image were to be considered, a prior giving more probability to the values 0 and 1 would be more appropriate, since *pure* (non-PVE, i.e. $\alpha=0$ or $\alpha=1$) voxels are in general far more common than PVE voxels (see Fig. 2). Examples of such distributions and further discussion about this issue can be found in (Roll et al., 1994). The problem of using a uniform distribution in the presence of many pure voxels is that small variations from 0 or 1 in the α 's of such voxels produce considerable errors, due to the high number of pure voxels in an image. It should be noted, however, that choosing the correct ‘U’-shaped prior is not trivial, and a bad choice can influence the results for the estimated α 's. A prior with wide tails, which tends to ‘round off’ too many values close to 0 and 1 will underestimate PVE, while a prior with narrow tails will not cancel the error due to small variations from 0 or 1 in pure voxels.

We now consider the independence assumptions that can be made within the above framework. First, we suppose that $p(I^i|\alpha^i)$ is independent across voxels. A reminder of the intuitive meaning of the distribution should convince that this is a reasonable assumption. $p(I^i|\alpha^i)$ represents the probability of an intensity value in a voxel, given the proportion of each tissue. This only depends on the distributions of the tissues, p_1 and p_2 . The value of α is given, and there is no reason why neighbouring voxels should influence this distribution. Let us now consider $p(I^i)$, the probability of a given intensity in a voxel. This

distribution is a direct function of $p(I^i|\alpha^i)$ and $p(\alpha^i)$ as stated in (3). In addition, the effect of $p(I^i)$ in the computation of $p(\alpha^i|I^i)$ is irrelevant, since it acts as a normalising constant for the distribution.

The assumption of independence for $p(\alpha^i)$ is more problematical. If pure voxels are taken into account, a neighbourhood relation should be incorporated, since voxels whose neighbours are pure tissue have a higher likelihood of being pure, due to the piecewise constant nature of the tissues. Markov random fields could be employed to this end, and we are currently investigating that possibility (Zhang et al., 2000). In our case, however, only PVE voxels are considered, which opens the possibility for assuming independence. The variability of shapes passing through voxels is infinite (unless some assumptions are introduced to limit their local geometry). Hence, no reliable prediction can be made to infer α^i from the values of α at neighbouring voxels. Independence is thus a reasonable choice. Further consideration to this argument is given in (Gonzalez Ballester, 1999).

4. Example: Gaussian distributions

We elaborate the framework described in the previous section for the case of a Gaussian probability distribution, which is the model most commonly used in MRI studies³. The intensity distribution for tissue j is

$$p_j(I) = G(\mu_j, \sigma_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(I - \mu_j)^2}{2\sigma_j^2}\right). \quad (5)$$

The expression for the mixture model can be computed analytically. Given a value of α^i , and assuming the mixture model in (1), the distribution for the combined intensities follows a linear combination of two Gaussians, which is itself a Gaussian (Fig. 3):

³Note that Wells et al. (1996) and Guillemaud and Brady (1997) use a Gaussian distribution on the logarithm of the intensities. However, this choice is not well suited for dark intensities (Gonzalez Ballester, 1999).

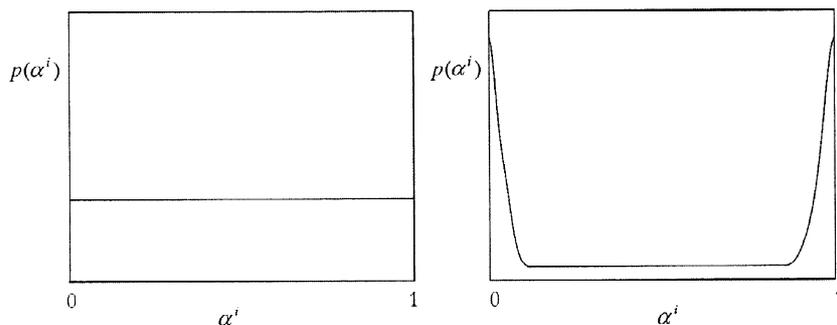


Fig. 2. Distributions for $p(\alpha^i)$. Left: Uniform distribution, corresponding to a set of PVE voxels; Right: ‘U’-shaped distribution, corresponding to an image containing pure tissue voxels.

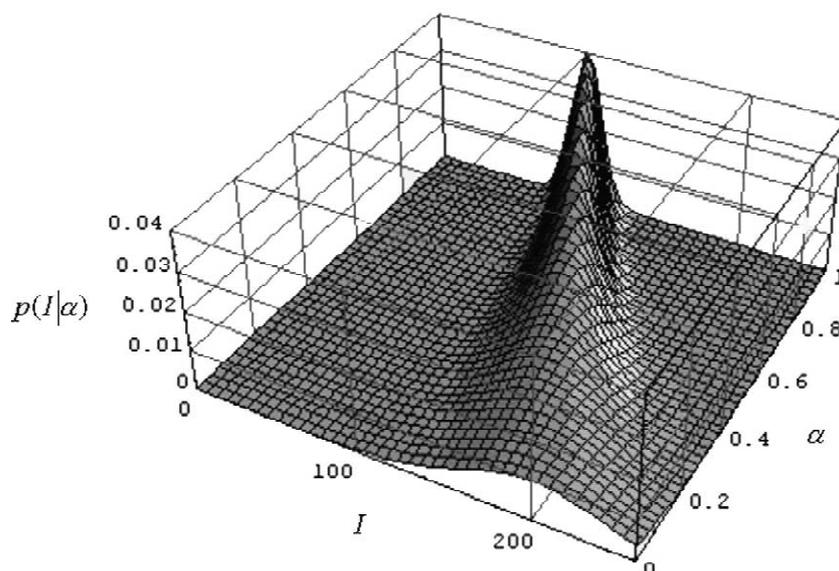


Fig. 3. $p(I|\alpha^i)$ plotted for all values of $\alpha^i \in [0 \dots 1]$. Each cross-section for a given value of α is a Gaussian following (6). In this example, $\mu_1 = 100$, $\sigma_1 = 10$, $\mu_2 = 200$ and $\sigma_2 = 40$.

$$p(I^i|\alpha^i) = G(\alpha^i \mu_1 + (1 - \alpha^i) \mu_2, \sqrt{\alpha^i \sigma_1^2 + (1 - \alpha^i) \sigma_2^2}). \tag{6}$$

Bayes' theorem can now be used to compute $p(\alpha^i|I^i)$, applying (2). As noted above, we assume that $p(\alpha^i)$ is uniform in the interval $[0 \dots 1]$, so $p(\alpha^i) = 1$. The resulting formula has the same form as (6), except for the normalising constant $p(I^i)$. However, this time it is not a Gaussian, since it is no longer a function of the intensity I ,

but of the proportion α (Fig. 4). A few cross-sections at chosen values of I^i are shown in Fig. 5. Note that the function is not symmetric.

5. Confidence bounds

The information contained in the individual distributions $p(\alpha^i|I^i)$ can be summarised by computing the most likely value of α^i , as well as upper and lower bounds at

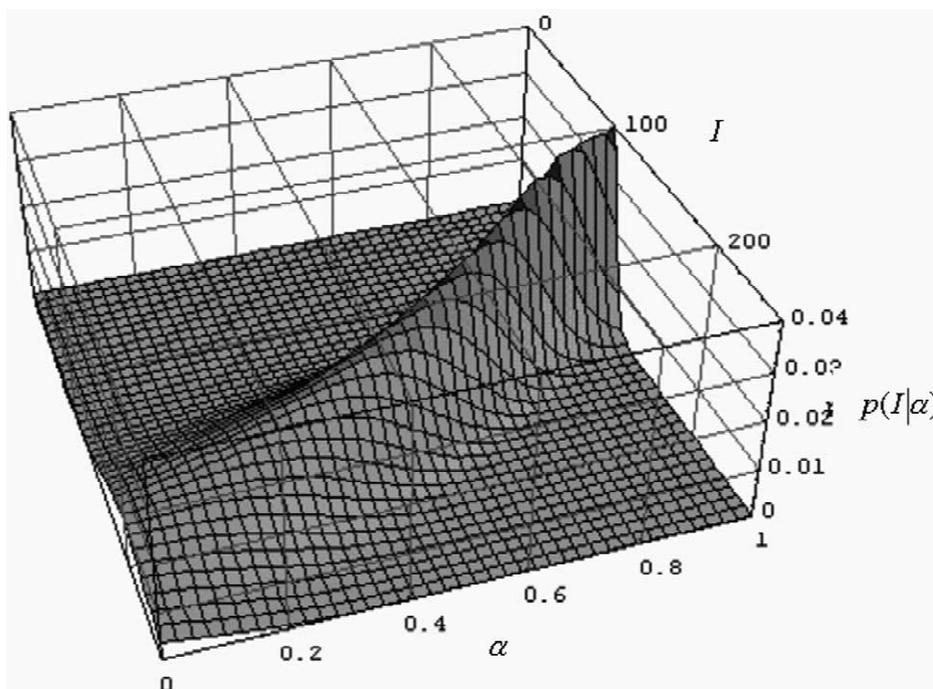


Fig. 4. $p(\alpha^i|I^i)$ plotted for all values of $I^i \in [0 \dots 255]$.

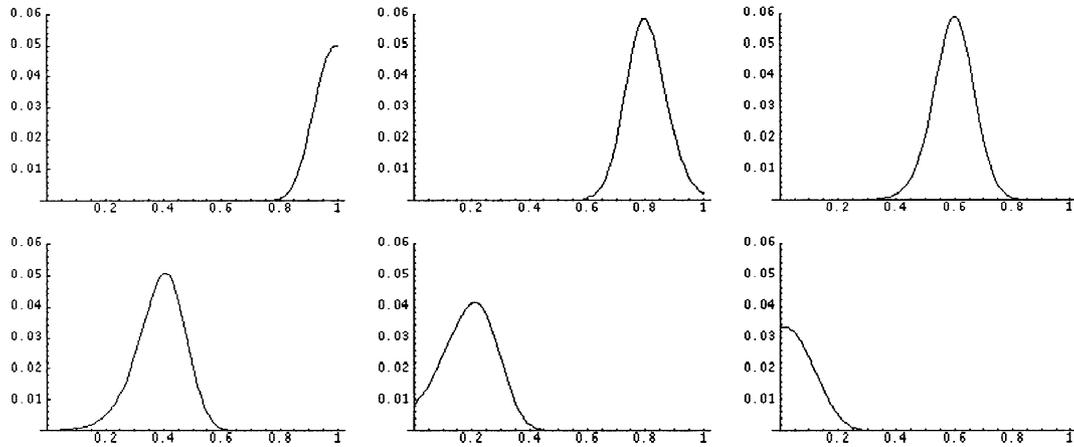


Fig. 5. Cross-sections of $p(\alpha^i | I^i)$ at values $I = 100 (\mu_1)$, 120, 140, 160, 180 and 200 (μ_2).

particular confidence levels. It was shown above that $p(\alpha^i | I^i)$ is not necessarily symmetric, so the mean value α_{mean} is not, in general, the value of highest probability. Assuming that the distribution is unimodal, i.e. has only one maximum (the mode, α_{mode}), a simple search, for example starting from the mean value, suffices to compute this value. α_{mode} represents the most likely value of α given the intensity of the voxel and the distributions for the tissues present in it.

Similarly, confidence intervals on α^i can be set by constraining the area below the probability density function (pdf) of $p(\alpha^i | I^i)$. Upper and lower bounds are found by locating the values α_{upper} and α_{lower} , respectively, which encompass a certain area surrounding the mode α_{mode} . This area is the lateral confidence interval (c_{lat}) for α .

$$\int_{\alpha_{\text{lower}}^i}^{\alpha_{\text{mode}}^i} p(\alpha^i | I^i)(\alpha) d\alpha = \int_{\alpha_{\text{mode}}^i}^{\alpha_{\text{upper}}^i} p(\alpha^i | I^i)(\alpha) d\alpha = c_{\text{lat}}. \quad (7)$$

For example, an area of $c_{\text{lat}} = 0.45$ to each side of the

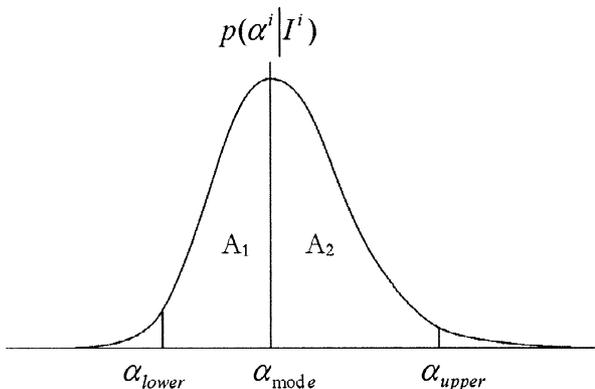


Fig. 6. Confidence bounds on $p(\alpha^i | I^i)$. α_{upper} and α_{lower} are set so as to make the areas A_1 and A_2 equal to c_{lat} .

mode enforces a 90% confidence interval, i.e. there is a 90% chance that the value of α^i given I^i is between the bounds computed in such a way. This point is illustrated in Fig. 6.

6. Propagation of uncertainty

The framework developed up to this point applies to individual PVE voxels. In order to compute quantitative estimates of the volume of a region bounded by PVE voxels, a method must be devised to combine the statistical information contained in the individual distributions in order to create a distribution for the volume

$$p(\alpha^i | I^i), \quad \forall i \rightarrow p(V). \quad (8)$$

6.1. Conservative bounds

The first approach towards propagating local information makes use of the confidence intervals developed in Section 5. In particular, upper and lower bounds on the volume V can be set by employing the upper and lower estimates of α for every PVE voxel. This obviously produces an overestimation of the confidence intervals (Pennec, 1996). The bounds computed in such a way will be referred to as *conservative bounds*.

In the case of volume, the framework is straightforward, since v can be expressed as

$$V = v_{\text{voxel}} \left(n_{\text{pure}} + \sum_{i \in \text{PVE}} \alpha^i \right), \quad (9)$$

where v_{voxel} is the volume of a voxel; n_{pure} is the number of pure voxels; and i indexes all PVE voxels containing tissue of the object being measured.

Upper and lower estimates of the volume can then be computed as

$$V_{\text{upper}} = v_{\text{voxel}} \left(n_{\text{pure}} + \sum_{i \in \text{PVE}} \alpha_{\text{upper}}^i \right), \quad (10)$$

$$V_{\text{lower}} = v_{\text{voxel}} \left(n_{\text{pure}} + \sum_{i \in \text{PVE}} \alpha_{\text{lower}}^i \right). \quad (11)$$

However, this method has two drawbacks. First, it produces a considerable overestimation of the confidence bounds. This means that using local bounds for each voxel at a confidence level c_{lat} will, in general, produce bounds for the global magnitude V with much greater confidence. This implies an unnecessary loss of resolution in the results reported by this method. Second, establishing confidence bounds is a step back in the process of establishing a solid statistical model for the PVE. In fact, what we are after is a statistical distribution for V , not just two confidence bounds.

6.2. Monte Carlo

One way to construct the statistical distribution for V is to use Monte Carlo methods (Leon-Garcia, 1994). The idea is to sample values from the individual distributions for each PVE voxel. Taking a sample from each PVE voxel, it is possible to compute a sample of V . By repeating this process a large number of times, the distribution for V is simulated and can thus be reconstructed. The process is spelled out in detail in Fig. 7. In order to generate samples from the distributions, a rejection method was used (Leon-Garcia, 1994).

Monte Carlo methods have been extensively used in Bayesian estimation frameworks (Chen et al., 2000), and they are often computationally expensive, to the point they quickly become ineffective in practical applications. In our case, however, experiments carried out for computing volumes (cf. following section) were performed in reasonable times (typically in the range of 2–5 min). This is partly due to the fact that the process is applied only to

PVE voxels previously segmented, not to the whole 3D image.

7. Validation with synthetic images

In order to provide ground truth for our measurements, we first employ a recursive subdivision procedure to simulate MR acquisitions of objects of known geometry, for which an analytic formula is available (e.g. an ellipsoid). At each voxel location, it is determined whether the 8 diagonal neighbours of the voxel are inside the simulated object by evaluating its formula. If this is the case, the voxel is assigned intensity value I_{in} , whereas value I_{out} is given to voxels whose 8 corners are all outside the object. The remaining voxels correspond to boundaries and their intensity values should simulate the partial volume effect. In order to determine the proportion α of the voxel that is inside the object, a recursive subdivision procedure is initiated by dividing the voxel by its centre into 8 smaller cubes. Testing for inclusion continues in the manner described above until all sub-voxels are assigned a value or a recursion limit is reached. Then, the intensity value

$$I = \alpha I_{\text{in}} + (1 - \alpha) I_{\text{out}} \quad (12)$$

is assigned to the voxel. The recursion limit is chosen so that the contribution to the final value of α in the limit is smaller than a certain small value ϵ (typically 1×10^{-10}). Tissue-dependent Gaussian noise can also be applied by modelling I_{in} and I_{out} as Gaussian distributions.

A synthetic data set was created using the program described above. The phantom data set contains a sphere of volume $V_{\text{ground}} = 1145.7$ in a voxel grid of $20 \times 20 \times 20$ voxels, each with dimensions $1 \times 1 \times 1$ (Fig. 8). Only two tissues (inside and outside the sphere) are represented. They are modelled as Gaussian distributions with the following parameters: $\mu_{\text{in}} = 200$, $\sigma_{\text{in}} = 2.5$, $\mu_{\text{out}} = 100$, $\sigma_{\text{out}} = 2$.

```

Algorithm MONTE-CARLO
  For  $n=1$  To number of Monte Carlo iterations Do
    volume[n]=0
    For  $i=1$  To number of PVE voxels Do
       $\hat{\alpha}^i = \text{sample from } p(\alpha^i | I^i)$ 
      volume[n] = volume[n] +  $\hat{\alpha}^i$ 
    End For
    volume[n] = volume[n] * voxel volume
  End For
  pdf = normalised histogram of volume[]
End Algorithm

```

Fig. 7. Algorithm for Monte Carlo propagation of uncertainty.

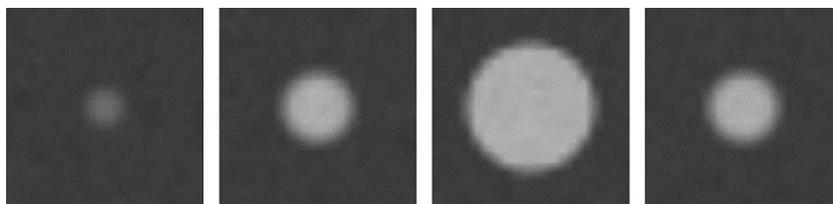


Fig. 8. Slices number 4, 5, 10 and 16 of simulated data set.

PVE voxels are isolated by building a mask of values different from μ_{in} and μ_{out} in an image generated simulating the same shape but without noise. The volume encompassed by these PVE voxels is 656, which establishes an area of uncertainty based on voxels of 57.3% of V_{ground} . We then proceed to fit an inner and outer surface in order to establish bounds as described in (Gonzalez Ballester et al., 2000). The width of these confidence bounds is computed as

$$\frac{(V_{outer} - V_{inner}) \times 100}{V_{ground}}. \quad (15)$$

The additional smoothness of surface representations is bound to reduce the uncertainty area, when compared to the one based on voxels. In particular, the width of the confidence interval on volume using 3D graph representations (simplex meshes) as surface model is 36.70% of V_{ground} , and this width reduces to 23.30% when using G^1 -continuous triangular Gregory–Bezier patches (Gonzalez Ballester et al., 2000).

We applied the PVE estimation framework to the noisy image, as described in the previous sections. The mixture model is constructed based on the ground truth parameters for the tissue distributions. Using the mode values α_{mode} of $p(\alpha^i | I^i)$ for every PVE voxel, the estimated volume is $V_{mode} = 1146.3$. The error with respect to V_{ground} is 0.05%.

Conservative bounds are established as described in Section 6.1. Several confidence thresholds were used, and the results are shown in Table 1. Note that these confidence thresholds do not reflect the confidence in the value of the volume, but on local voxel-based estimates. This means that these confidence bounds are very conservative. In fact, the real volume is comfortably included between the bounds even in cases when a low threshold is used. Experiments performed on real MRI data suggest that a

Table 1
Conservative confidence bounds on volume for the synthetic phantom^a

Confidence threshold	Lower bound (%)	Upper bound (%)	Width %
80%	1131.6 (−1.23%)	1159.3 (+1.19%)	2.42%
90%	1127.6 (−1.58%)	1162.3 (+1.45%)	3.03%
95%	1121.3 (−2.13%)	1164.3 (+1.62%)	3.75%
99%	1012.4 (−11.63%)	1166.8 (+1.84%)	13.47%

^a Upper and lower bounds are computed using the values of α_{upper} and α_{lower} , respectively, as described in Section 6.1. Percentages are computed with respect to $V_{ground} = 1145.7$.

Table 2
Results of the Monte Carlo simulation^a

Confidence	Lower bound (%)	Upper bound (%)	Width %
$2\sigma_{MC}$	1144.6 (−0.10%)	1146.6 (+0.08%)	0.18%
$3\sigma_{MC}$	1144.1 (−0.14%)	1146.9 (+0.10%)	0.24%

^a Confidence intervals were computed as a function of the standard deviation computed from 10 000 samples.

value of 80% confidence ($c_{lat} = 0.4$) is reasonable, and a more conservative estimate is obtained when using a threshold of 90% confidence. Increasing this threshold to 99% is exaggerated, and the width of the confidence bounds in this case is unreasonably high. The values in Table 1 should be compared to the width of the confidence bounds using the simplex mesh (36.70%) and tGB patches (23.30%). There is a reduction by an order in magnitude.

Finally, a Monte Carlo process is used to generate 10 000 samples of volume estimates computed from random samples of the distributions of $p(\alpha^i | I^i)$ for every PVE voxel i , as described in Section 6.2. The sample mean was $\mu_{MC} = 1145.5$ (−0.02% error with respect to V_{ground}), and the standard deviation was $\sigma_{MC} = 0.4559$. Confidence bounds based on these values are reported in Table 2. A further reduction by an order of magnitude in the width of the confidence intervals was obtained, the ground truth volume still being contained between the bounds.

It should be emphasised that the result of this Monte Carlo simulation is not a set of bounds, but a statistical distribution for the volume, from which bounds and other useful information can be derived.

8. Discussion and practical considerations

The results reported in the previous section show an impressive improvement by two orders of magnitude in the

resolution of the measurements obtained from MRI data. It should be noted, however, that in this synthetic example we made use of three pieces of information that are not likely to be available for real data:

- The exact parameters for the distribution of the tissues, $p_j(I)$.
- The composition of PVE voxels (trivial in the synthetic example, which contains only two tissues).
- The location of PVE voxels of interest.

These issues are discussed in more detail in the following three sections, and solutions are offered for each of them. In the following section, the sensitivity of the method to slight deviations from the correct tissue parameters is analysed, and a solution based on explicitly modelling the uncertainty in the estimation of such parameters is proposed. Next, a method to enhance the PVE maps obtained in (Gonzalez Ballester et al., 2000) and which includes information about the composition of PVE voxels, is described in Section 10. That section also describes how contextual knowledge may be used to obtain good voxel-based segmentations including PVE information. Finally, a method for generating PVE masks from the inner and outer surfaces developed in (Gonzalez Ballester et al., 2000) is described in Section 11.

9. Sensitivity to tissue parameters: modelling uncertainty in parameter estimation

The method, as it has been described so far, is quite sensitive to the tissue parameters. In fact, the method assumes that the tissue distributions $p_j(I)$ are a perfect model. Thus, in the synthetic case presented in the previous section we obtained very narrow bounds around the ground truth volume. However, such extremely narrow bounds also mean that small errors in the means of the tissue distributions tend to shift the distribution for the volume, yielding bounds that do not contain the ground truth⁴. For example, if in the synthetic phantom we change μ_{in} from 200 to 201 and run the Monte Carlo simulation, the $3\sigma_{MC}$ bounds become [1141.3, 1144.1], which do not contain $V_{ground} = 1145.7$. Conservative bounds are not so sensitive to the values of the parameters, and although they slightly shift when incorrect parameters are used, V_{ground} is consistently between the bounds.

For our experiments with MRI data, tissue distributions are built from samples obtained using manual selection of voxels. Tissue parameters estimated in this way may vary in their accuracy and precision. One possible solution is to

create more accurate tissue models based on the physics of MRI. We contend that a careful analysis of acquisition parameters and chemical properties of particular tissue types could yield a satisfactory model. More details are given in (Gonzalez Ballester, 1999).

Another solution is to sacrifice some of the resolution of the method, i.e. to ‘widen’ the pdf of the volume $p(V)$ in favour of accuracy. One possible approach is to model tissue parameters as statistical variables, in order to reflect the uncertainty in the estimation of the parameters. Thus, means and variances can have a most likely value and some standard deviation around it. This is the approach we adopt. Tissue parameters are estimated from the data set as described above, i.e. obtaining samples of the intensities of tissues by manual selection. However, this process is repeated several times, each yielding an estimate of the tissue parameters (e.g. the mean and standard deviation, for Gaussian models) of the distribution on intensities of the tissue. Next, the mean and standard deviation of such parameters are computed. Therefore, this model assumes that the distribution of the noise in the estimation of the parameters is Gaussian.

Monte Carlo simulation is performed in order to generate the final tissue model. At each iteration of the Monte Carlo process, samples from the distributions of the tissue parameters are drawn. This produces an instance of the intensity distribution for the tissue. A sample from such a distribution is obtained and stored, and the process is repeated. After a number of iterations, the histogram of the samples is analysed to build the final intensity model for the tissue (Fig. 9). Effectively, the process sketched above is a non-parametric generation of distributions.

Since the framework described in this paper allows for any type of tissue distribution to be used, a mixture model is then constructed from the tissue distributions computed above and the remaining steps in the PVE estimation framework are as described in the previous sections.

Alternatively, a scheme fitting a pre-determined number of tissue distributions to the histogram of the image could be employed to compute the tissue distributions automatically. See for example (Santago and Gage, 1993) for an application using Gaussian tissue models. However, this technique assumes that all intensities in the image can be explained using the modelled distributions. In other words, either distributions are created for all tissues present in the image or the fitting will be incorrect, as it will try to explain alien intensities using the available distributions.

10. Use of contextual information to determine the contents of PVE voxels

Generally, MRI scans are of anatomy that contains more than two tissue types that are of interest. It is therefore necessary to determine which two tissues are contained in particular PVE voxels. This is a vital requirement for the construction of the PVE mixture model. As one possible

⁴Note that the values of the variances of the distributions in the simulated data set of Section 7 are very narrow. This was forced as an illustrative example to show the sensitivity of the method to incorrect parameters here. Thus, the simulated data set should be regarded as an illustrative example of the workings of the methodology, not a data set representative of clinical settings (refer to the following sections for such examples).

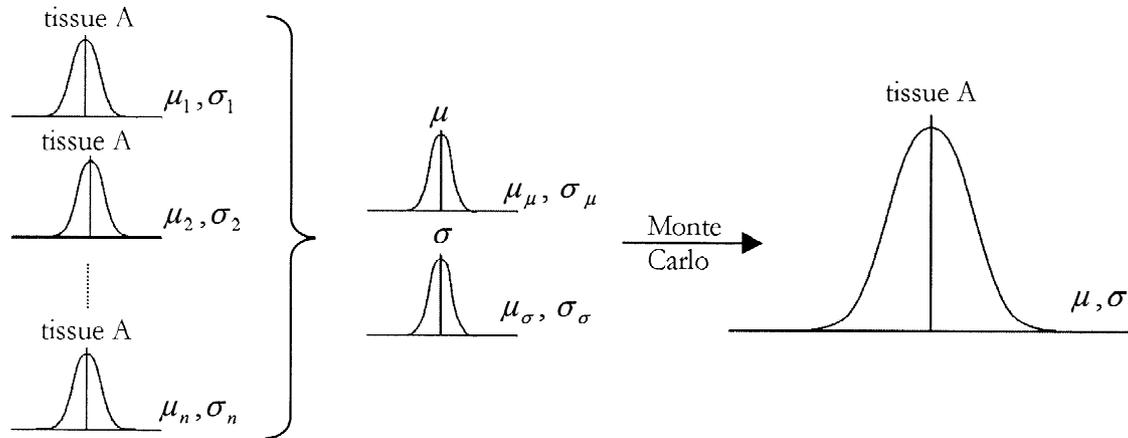


Fig. 9. Uncertainty on the estimation of tissue parameters is explicitly modelled by representing tissue parameters as statistical distributions. A Monte Carlo simulation is then used to build the final tissue model.

approach, we can make use of the voxel classification used as part of the morphometric framework described in (Gonzalez Ballester et al., 2000). In particular, a method based on the EM algorithm (Guillemaud and Brady, 1997) was used to interleave the classification of voxels into tissue classes and also correct for intensity inhomogeneities (bias fields). PVE voxels were detected by setting a threshold on the minimum value of the tissue likelihood for each voxel. Therefore, if a particular voxel has a probability value greater than the threshold for a particular tissue, the voxel is identified as containing only that tissue type, and otherwise it is marked as PVE. An example is shown in Fig. 10.

In the following, it will be shown that the use of contextual information significantly improves the results of such voxel-based classifications. Two intuitively plausible ideas can be used to identify the contents of voxels marked as PVE, by looking at their 26 neighbours:

- (1) If a voxel is marked as PVE but all non-PVE neighbours have been given the same tissue tag t , the voxel has a large probability of having been misclassified as PVE. Tissue tag t is then assigned to it. In

order to increase the robustness of this criterion, a minimum number of non-PVE neighbours is established.

- (2) Otherwise, the two tissues present in the PVE voxel are determined as the two most frequent tissue tags in its neighbourhood.

A further consideration may be taken into account in order to improve further the results. Let us focus on the T_1 -weighted MRI slice shown in Fig. 10. Cerebrospinal fluid (CSF) appears as dark intensities, white matter (WM) as bright intensities, while grey matter (GM) occupies the range of intensities in between. It is obvious that the intensities of PVE voxels containing CSF and WM overlap, and so could be wrongly assigned to the distribution of GM. This is illustrated in Fig. 11.

The following criterion is used to avoid such misclassifications:

- (3) If a voxel is classified as GM but some of its neighbours are CSF, then it must be a PVE voxel. If the most prevalent tissue occurring in its neighbours is WM, it is classified as PVE(CSF, WM); if it is GM, it is identified as PVE(CSF, GM). This rule works for

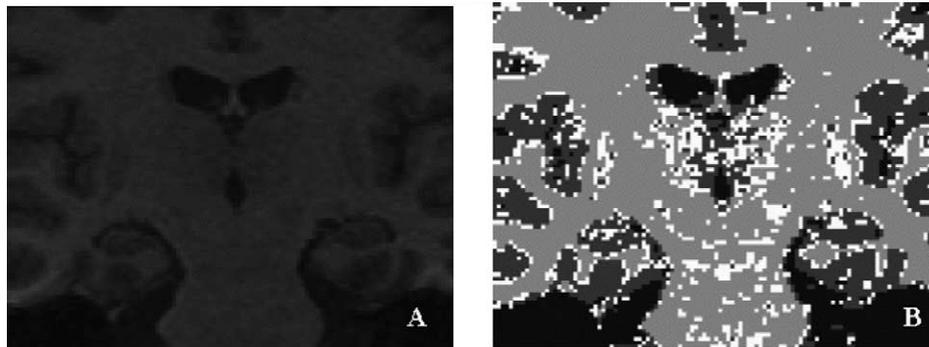


Fig. 10. Voxel-based segmentation obtained as described in (Gonzalez Ballester et al., 2000). (A) Original image; (B) segmentation (cerebrospinal fluid: black, white matter: light grey; grey matter: dark grey; PVE: white).

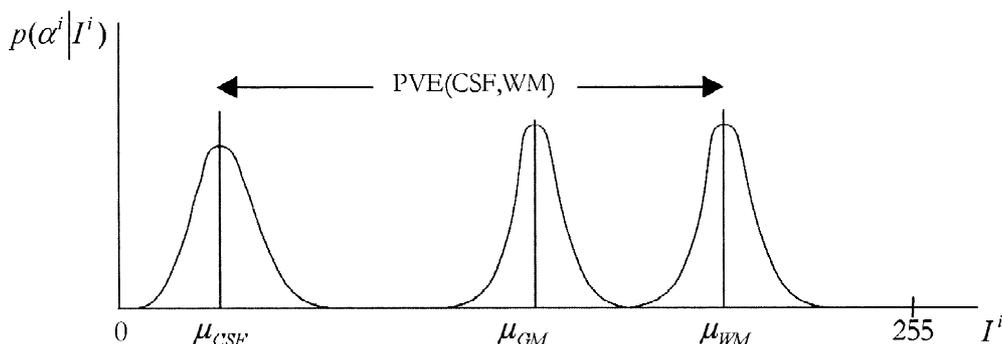


Fig. 11. The distribution of GM is in the range of intensities of PVE voxels containing CSF and WM. Thus, such PVE voxels may be wrongly classified as GM.

three tissue types, which proves sufficient for most practical applications. If more tissues need to be modelled, additional criteria can be established.

Fig. 12 shows the improvement on the voxel-based classification achieved using these three criteria. Note that noise in central areas of white matter has been considerably reduced, and the contents of PVE voxels determined. Additionally, voxels around the ventricles, which are PVE(CSF, WM) and were wrongly classified as GM in Fig. 10 are now correctly identified.

11. Construction of PVE masks from inner and outer surface bounds

In order to isolate the PVE area around the structure of interest, we can make use of the surface bounds created as described in (Gonzalez Ballester et al., 2000). These surfaces are constructed to encompass the region of uncertainty in the segmentation due to PVE. Once the inner and outer surfaces are constructed, the problem is reduced to creating a mask with the same voxel structure as the original image, highlighting the region encompassed by the two bounds. The two masks are then combined and



Fig. 12. Improvement on the voxel-based classification by using contextual information. Compare with Fig. 10. CSF: black, WM: light grey, GM: dark grey, PVE(WM, GM): white, PVE(CSF, GM): red, PVE(CSF, WM): green.

simple region growing techniques are used to include voxels contained between the bounds (combining the inside of the inner bound and the outside of the outer bound). We now describe how to construct the masks. The technique described here is closely related to the algorithm used in Section 7 to create synthetic images for phantom validation. In that case, an analytical formula for the object was available, and an oct-tree recursive subdivision process was used by testing the corners of the (sub-)voxels for inclusion in the object. The main difference with respect to the present case rests on the fact that the definition of the object was a volumetric one, whereas the bounds to be used here are surfaces.

The two-dimensional nature of the surface bounds, defined as a set of triangular Gregory–Bezier (tGB) patches (Gonzalez Ballester et al., 2000), makes it quite difficult to test for inclusion of a 3D point inside the object enclosed by the surface. Fortunately, for the application that lies behind the work presented in this section, we only need to locate and mark the voxels the surface passes through. To this effect, we can take advantage of the parametric definition of the surface. In general, any surface can be defined by a parametric mapping of the form

$$S(u, v) = (x(u, v), y(u, v), z(u, v)) \quad u \in D_u, v \in D_v, \quad (16)$$

where D_u and D_v are the domains for the parameters u and v , respectively. Without loss of generality, in the following we will assume that $D_u = D_v = [0, 1]$.

For the particular case of tGB patches, a prior step must be taken in order to establish a map \mathbf{b} from surface coordinates in the domain $[0, 1] \times [0, 1]$ to barycentric coordinates, used in the definition of tGB patches⁵ (see Fig. 13),

$$\begin{aligned} [0, 1] \times [0, 1] &\rightarrow \{[0, 1] \times [0, 1] \times [0, 1] \mid u' + v' + w' = 1\} \rightarrow \mathbb{R}^3 \\ (u, v) &\rightarrow \mathbf{b}(u, v) = (u', v', w') \rightarrow \mathbf{tGB}(u', v', w') = (x, y, z) \end{aligned} \quad (17)$$

⁵Note that a simpler approach, consisting of tessellating the triangle formed by the barycentric domain, could be employed, avoiding the need for the map \mathbf{b} . The description above is of a more general nature, valid not only for barycentric definitions of the surface.

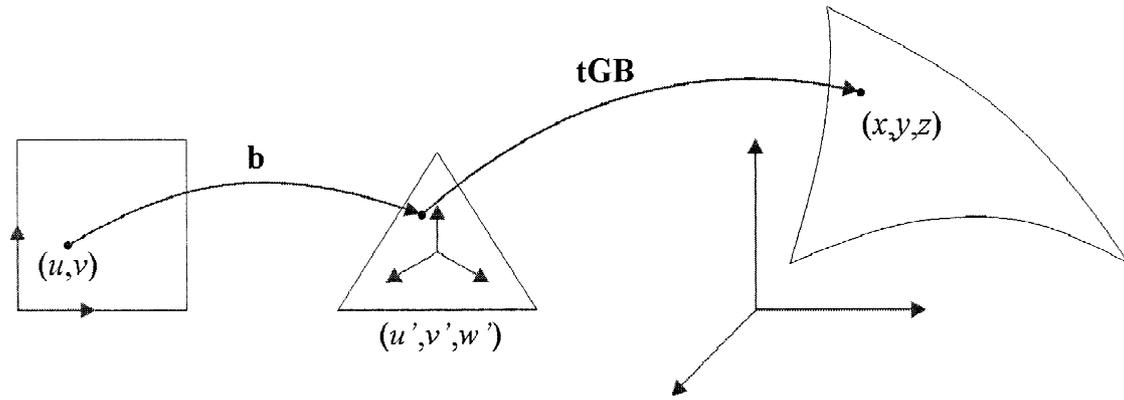


Fig. 13. Mapping from surface coordinates (u, v) to three-dimensional points in a tGB patch. First, surface coordinates must be mapped to barycentric coordinates, in order to index the formula for the patch.

This prior mapping is given by the formula

$$\mathbf{b}(u, v) = (u, v \cdot (1 - u), 1 - u - v). \quad (18)$$

We now consider the more general problem of creating a mask for a surface defined parametrically in the form of (16), for any voxel grid. A possible solution would be to densely sample the surface after guaranteeing that there is sufficient resolution to prevent missing any voxels. However, it is difficult to establish a sampling step in the parametric space of the surface with these characteristics, unless an extremely conservative—and computationally inefficient—approach is taken.

A recursive subdivision technique is used. Instead of acting on the three-dimensional space of the surface and the voxel grid, this time the subdivision is performed on the two-dimensional parameter space. Therefore, the algorithm used is a *quad-tree* recursive subdivision. The approach consists of dividing the parametric space into squares and computing the location of the 3D points corresponding to the corners of each parametric square. The stopping criterion for the recursive subdivision of these squares is a minimum distance between the four 3D

corner points. This minimum distance must, of course, be related to the voxel size. Typically, conservative values around 10^{-8} times the smallest side of a voxel are used, in order to not miss any voxel. The algorithm is illustrated in Fig. 14.

This method is exhaustive in the coverage of intersected PVE voxels. Alternative faster methods based on transforming the surface model into a triangulation and then computing the intersection of the triangles with the planes defining the imaging matrix could also be employed. However, the size of such triangles should be kept minimal in order not to miss intersected voxels.

It should be noted that by exploiting information about the approximate (or candidate) locations of the boundary pixels in order to isolate PVE voxels we also reduce the effect of voxels incorrectly classified as PVE. For example, voxels which are pure but whose intensities fall in the tails of the tissue distribution may be classified as PVE, thus biasing the results. The number of such voxels is very small, and their incidence is spread over the whole image. Therefore, the number of such voxels occurring at the boundary (i.e. the area isolated by the mask) will be extremely small.

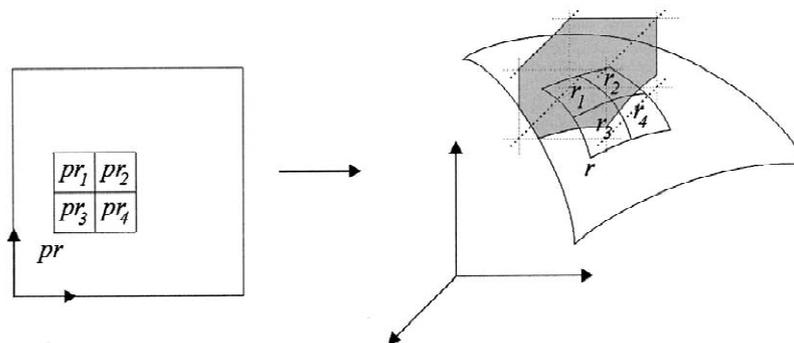


Fig. 14. Illustration of the quad-tree subdivision process. The rectangle pr in parametric space maps, through $\mathbf{tGB} \circ \mathbf{b}$, to the surface sub-patch r in 3D space. A voxel is shown, together with its intersection with the tGB patch. Since r does not fall completely inside the voxel, it is subdivided into 4 smaller rectangles, by subdividing pr . The rectangle r_1 falls completely inside the voxel, so the voxel will be marked. The remaining patches will continue the subdivision process to guarantee that no voxels are missed. A threshold area for the sub-patch is used as a recursion limit.

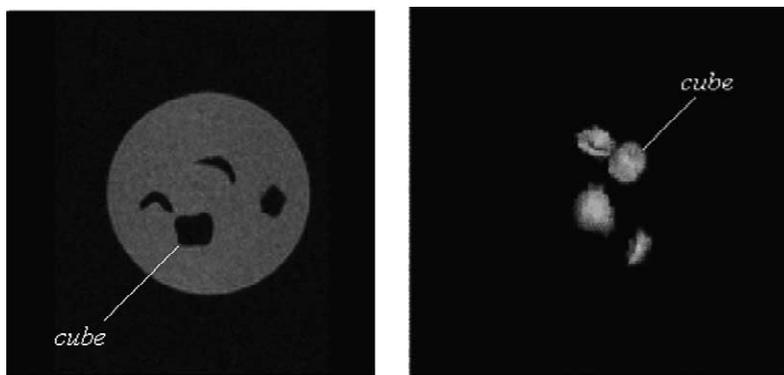


Fig. 15. One slice through and a rendering of the MRI phantom. Object ‘cube’, for which results are reported below, is indicated.

12. Experiments on MRI phantoms

A phantom consisting of a group of shapes made from paraffin wax and embedded in an agarose gel was used in a second experiment. By measuring the density of the wax, the true volume can be derived from their weight to within a confidence interval of 2% (Roll et al., 1994). The phantoms were developed to simulate the size and shape of MS lesions, which are usually very small relative to the resolution of the MR acquisition. Fig. 15 shows a slice through the MRI phantom (1 mm^3 voxels, no inter-slice gap) and a rendering of the shapes embedded in it.

Following the morphometric framework of (Gonzalez Ballester et al., 2000), we first applied a bias field correction, and then inner and outer surface bounds were fitted to the paraffin wax shapes. Volume confidence bounds were computed using voxel-based measurements, simplex mesh, and tGB patches. The results are summarised in Table 3. The ground truth interval is provided for reference in the last row.

PVE voxels are identified by rendering the inner and outer surfaces into the voxel grid, as described in the previous section. Voxels contained between the bounds are included into the PVE mask (Fig. 16). The PVE mixture model is constructed based on tissue parameters computed from sampling a slice of the data set several times. As we noted in Section 9, this enables us to model explicitly the uncertainty in the estimates of the tissue parameters inherent in the sampling technique. Conservative bounds to

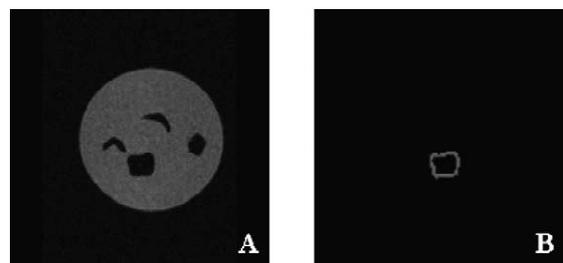


Fig. 16. (A) A slice of the phantom data set; (B) corresponding slice from the PVE mask generated from the outer and inner surfaces computed as described in (Gonzalez Ballester et al., 2000).

a 90% confidence level ($c_{\text{lat}} = 0.45$) were established, a Monte Carlo simulation was used to generate 10 000 samples of the volume, and statistics were derived from such samples. The mean volume was 1029.36, and the standard deviation was 1.47. $\pm 3\sigma$ confidence bounds are shown in Table 3.

All the estimated bounds are correct, in the sense that they contain the range provided as ground truth. Nevertheless, the confidence interval computed by means of the Monte Carlo simulation is actually smaller than the ground truth. This interval is contained inside the ground truth bounds, and the mean is almost identical to the ground truth mean, which suggests that the result is sensible. Additionally, $\pm 3\sigma$ bounds comfortably contain the ground truth mean. However, it is not possible to determine

Table 3
Summary of the results obtained for the MRI phantom ‘cube’^a

Method	Confidence interval	Width %
Simplex mesh	[899.6, 1264.7]	35.5%
Voxel	[821.0, 1152.0]	32.2%
tGB mesh	[954.8, 1166.9]	20.6%
PVE conservative (90% bounds)	[980.9, 1112.0]	12.7%
PVE Monte Carlo	$1029.4 \pm 3\sigma = [1024.6, 1033.4]$	0.9%
Ground truth	$1028 \pm 2\% = [1007, 1049]$	2.0%

^a Percentages are computed with respect to the mean value of the ground truth confidence interval, 1028.

whether or not the Monte Carlo bounds are correct, since the resolution of the ground truth interval provided by the manufacturer of the phantom is insufficient.

13. Clinical study

We applied our method to a neurological study concerning early onset schizophrenia. In particular, the aim of this study was to determine the possible relevance (or otherwise) of the asymmetry of the temporal horns to the aetiology of schizophrenia. The temporal, or lateral, horns of the ventricles are extremely thin structures traversing the brain in the anterior–posterior direction. Their minute diameter means that these structures are barely visible in MRI scans, and most of the voxels intersected by them exhibit PVE (Figs. 17 and 18). As a consequence, no accurate quantitative imaging study relating these organs has been possible to date.

High resolution coronal T_1 -weighted MRI data sets of 8 schizophrenic patients (SC) and 8 normal controls (NC) were analysed. The data sets each comprise 124 slices (slice thickness 1.5 mm), each slice consisting of 256×256 voxels of dimensions $0.9375 \times 0.9375 \text{ mm}^2$. Therefore, the field of view (FOV) is $240 \times 240 \times 186 \text{ mm}^3$. Typical slices through one of the data sets are shown in Fig. 17. These scans were acquired at the State University of New York (SUNY) in Stonybrook, NY, USA, and formed a part of the data pool gathered for the European Project BIOMORPH (Colchester et al., 1996), of which the present study forms a part.

Example images of intermediate processing steps were shown above (Figs. 10 and 12). It should be noted that, in some regions where the thickness of the temporal horns is minimal, the automated classification of voxels into PVE classes failed. Manual editing was used in such cases, which were infrequent.

Voxel-based ‘expert’ segmentations setting upper and lower thresholds on an advanced region-growing technique (Colchester et al., 1996) report a volume of uncertainty due to PVE of 228% with respect to ‘pure’ voxels. After

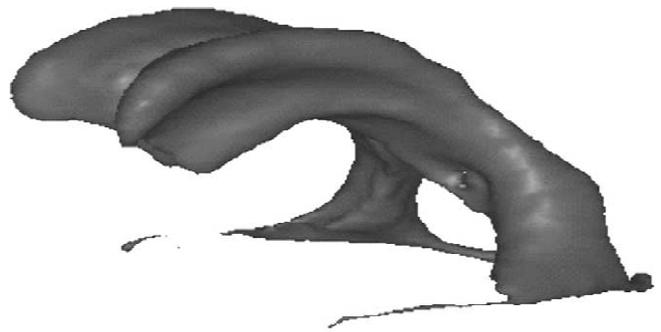


Fig. 18. Smoothed rendering of a voxel-based segmentation of the ventricles. Note that the temporal horns are unconnected. It is very difficult to perform a binary classification, due to the extremely small size of the temporal horns with respect to the voxel size.

applying our methodology, confidence bounds reduce to 35%, for the conservative case, and 6% for the Monte Carlo method (Figs. 19 and 20). These results show an improvement in quantitative resolution of two orders of magnitude, sufficient to convert a problem that has been considered intractable to date into one that may feasibly be tackled.

The clinical implications of the results are analysed next. Symmetry (rather a loss of normal asymmetry) is the main clinical focus of the present study, so it is first necessary to define a suitable measure. To this end, we define the following normalised symmetry coefficient:

$$S = \frac{L - R}{L + R},$$

where L denotes the volume of the left temporal horn, and R is the volume of the right temporal horn. Perfect symmetry gives a value $S = 0$, whereas the more asymmetric the horns, the larger the value of S . Standard error propagation (Chatfield, 1983) is used to compute the standard deviation of the symmetry coefficient as a function of the standard deviations for L and R . Values for all data sets are provided in Table 4.

A t -test (Chatfield, 1983) on this data may be applied to test for significant differences between schizophrenics and

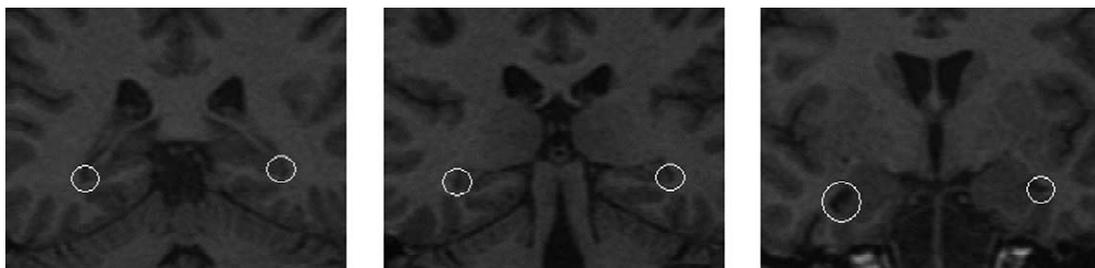


Fig. 17. Three MRI slices (numbers 25, 29 and 47, from back to front) showing the temporal horns. Notice that they traverse very few voxels. The middle slice, in particular, presents a very unclear delineation of them, highlighting the need for sub-voxel resolution.

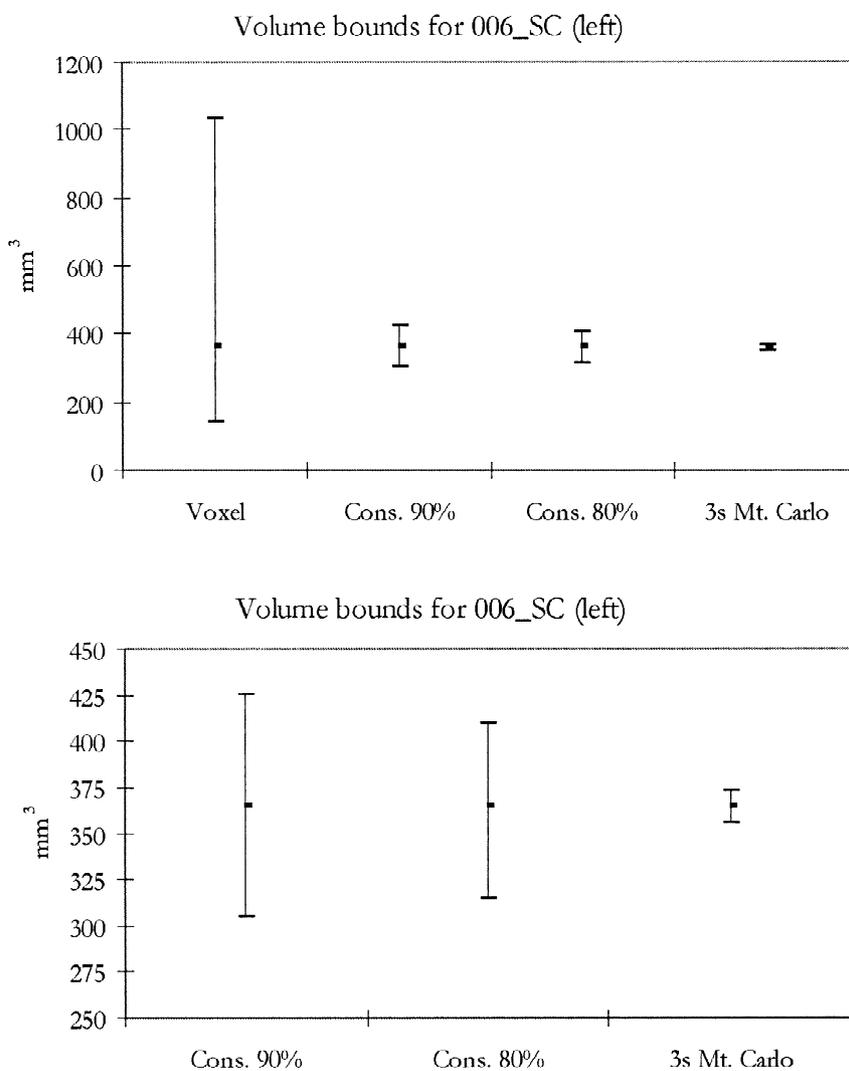


Fig. 19. Volume bounds, in mm^3 , for the left temporal horn of patient 006_SC. The top graph shows results for voxel-bounds, conservative bounds to 90% and 80% local confidence levels, and $\pm 3\sigma$ bounds computed from the results of the Monte Carlo simulation. For clarity, the bottom graph shows a close-up of the results obtained using the last three methods. The mean value of the Monte Carlo volume distribution is marked as a dark dot.

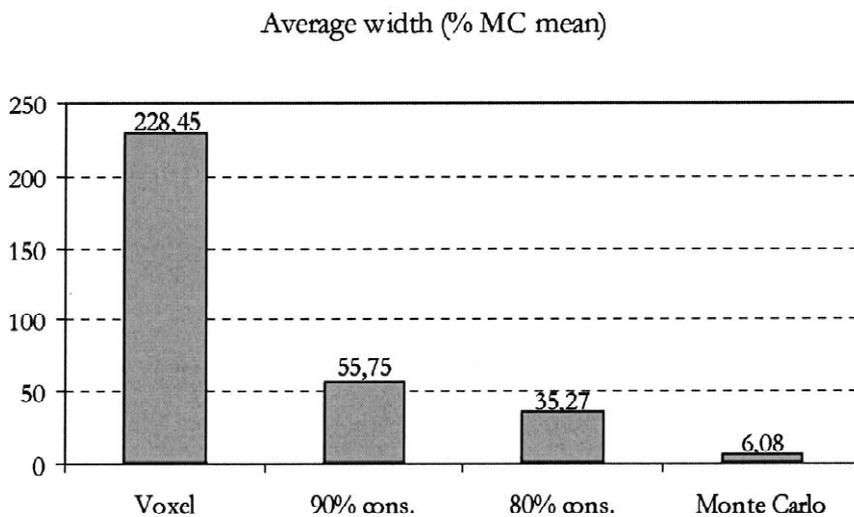


Fig. 20. Average width of confidence intervals on temporal horn volumes. From left to right, voxel bounds, 80% and 90% conservative bounds, and $\pm 3\sigma$ bounds on Monte Carlo results. Values are expressed as percentages of the mean value of the Monte Carlo simulation.

Table 4

Mean and standard deviation of the symmetry coefficients computed on all data sets (SC=schizophrenic, NC=normal control)

Data set	Mean sym.	Std. sym.
006_SC	-0.20	0.005
007_SC	-0.03	0.009
008_SC	0.27	0.006
010_SC	0.14	0.008
024_SC	-0.10	0.006
025_SC	-0.12	0.008
033_SC	-0.48	0.006
039_SC	0.39	0.009
104_NC	0.14	0.008
105_NC	-0.51	0.008
106_NC	-0.15	0.009
107_NC	-0.47	0.007
112_NC	0.19	0.006
114_NC	-0.08	0.006
117_NC	0.15	0.005
118_NC	0.00	0.009

normal controls. The t -test determines the probability of two samples being drawn from distributions with the same mean, and is especially suited to a small number of sample values. Using a two-tailed t -test, the probability of both sets of symmetry coefficients (SC and NC) being drawn from a distribution with the same mean is 0.6. This means that *there are no significant group differences in left–right symmetry of the temporal horns between the schizophrenics and normal controls analysed in this study.*

14. Future work

Future work will concentrate on the explicit modelling and propagation of uncertainty in the early stages of the framework. In particular, the technique as it stands makes use of a set of tools to isolate PVE voxels prior to the estimation of their contents. This allows to use a simple uniform prior for the proportion of each tissue, removes noise coming from pure voxels, and speeds up computation to the point that Monte Carlo methods run in efficient times. Validation studies for these tools were presented in previous work (Gonzalez Ballester et al., 2000). However, quantification of the uncertainty resulting from this prior segmentation is still an interesting problem to be addressed. To this end, error propagation techniques (Pennek, 1996) could be applied.

Validation with more realistic MR phantoms will also be performed. To this end, phantoms such as MNI's Brain-Web (Collins et al., 1998) or Yale's MRI phantom (Zubal et al., 1994) could be employed.

Analytical computation of the pdf on global measurements will also be investigated. As suggested by one of the anonymous reviewers, computation of the pdf for the volume from the individual voxel PVE pdfs may be performed in closed form. Assuming independence across

PVE voxels, iteratively convolving all their pdfs would result in a pdf for the volume. This would also allow to compute the mode of the volume distribution, instead of the mean, which would further improve the accuracy of the method. This technique will be investigated, and particular limitations induced by the independence assumption will be revised. In particular, a technique based on the use of Markov random fields is under study (Zhang et al., 2000). Alternative approaches include the use of anatomical models (Kapur et al., 1998; Shattuck and Leahy, 2000).

15. Conclusions

We have presented a framework for modelling and estimating the partial volume effect. The mixture model allows for sub-voxel resolution in measurements obtained from MRI data. Such a mixture model is propagated to build a distribution on global shape descriptors (e.g. volume), or estimate confidence bounds. Three key issues for the correct use of the mixture model in real MRI data sets were tackled, namely: sensitivity to the tissue prior distributions on the estimated distribution for the global shape descriptor, and the need to isolate PVE voxels and determine their contents. Results show an improvement in quantitative resolution of two orders of magnitude with respect to previous methods. This opens the possibility to perform clinical studies on small structures where PVE voxels contain a high percentage of the volume of the structure, as illustrated by the application of the technique to the study of the temporal horns.

Acknowledgements

This work was supported by the EC-funded BIOMORPH project 95-0845, a collaboration between the Universities of Kent and Oxford (UK), ETH Zurich (Switzerland), INRIA Sophia Antipolis (France) and KU Leuven (Belgium). The authors would like to thank Prof. Tim Crow from the Department of Psychiatry of the University of Oxford, for driving our research towards clinically useful areas. The MRI phantom used in Section 12 was generously provided by Dr. Fernando Bello and Prof. Alan Colchester of the NeuroMedIA group of the University of Kent at Canterbury. We would also like to thank the anonymous reviewers for very useful suggestions.

References

- Chatfield, C., 1983. Statistics for Technology. Chapman & Hall.
- Chen, M.-H., Shao, Q.-M., Ibrahim, J.G., 2000. Monte Carlo Methods in Bayesian Computation. Springer-Verlag.

- Choi, H.S., Haynor, D.R., Kim, Y., 1991. Partial volume tissue classification of multichannel magnetic resonance images—a mixel model. *IEEE Trans. Med. Imag.* 10 (3), 395–407.
- Colchester, A.C.F., Zhao, J., Holton-Tainter, K.S., Henri, C.J., Maitland, N., Roberts, P.T.E., Harris, C.G., Evans, R.J., 1996. Development and preliminary evaluation of VISLAN, a surgical planning and guidance system using intra-operative video imaging. *Medical Image Analysis* 1 (1), 73–90.
- Collins, D.L., Zijdenbos, A.P., Kollokian, V., Sled, J.G., Kabani, N.J., Holmes, C.J., Evans, A.C., 1998. Design and construction of a realistic digital brain phantom. *IEEE Trans. Med. Imag.* 17 (3), 463–468.
- Everitt, B.S., Hand, D.J., 1981. *Finite Mixture Distributions*. Chapman and Hall, London.
- Gonzalez Ballester, M.A., 1999. Morphometric analysis of brain structures in MRI. DPhil Thesis. Dept. of Engineering Science, University of Oxford.
- Gonzalez Ballester, M.A., Zisserman, A., Brady, M., 2000. Segmentation and measurement of brain structures in MRI including confidence bounds. *Medical Image Analysis* 4 (3), 189–200.
- Guillemaud, R., Brady, M., 1997. Estimating the bias field of MR images. *IEEE Trans. Med. Imag.* 16 (3), 238–251.
- Kapur, T., Grimson, W.E.L., Kikinis, R., Wells, W.M., 1998. Enhanced spatial priors for segmentation of magnetic resonance imagery. In: *Proc. of Medical Image Computing and Computer Assisted Intervention MICCAI'98*. First International Conference. Springer Verlag, Berlin, Germany, pp. 457–468.
- Laidlaw, D.H., Fleischer, K.W., Barr, A.H., 1998. Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms. *IEEE Trans. Med. Imag.* 17 (1), 74–86.
- Leon-Garcia, A., 1994. *Probability and Random Processes for Electrical Engineering*. Addison-Wesley Publishing.
- Marais, P.C., 1999. The segmentation of sparse MRI images. DPhil Thesis. Department of Engineering Science, University of Oxford.
- Niessen, W.J., 1997. Multiscale medical image analysis. PhD Thesis. University of Utrecht.
- Pennec, X., 1996. L'incertitude dans les problèmes de reconnaissance et de recalage. Application en imagerie médicale et biologie moléculaire. PhD Thesis. École Polytechnique.
- Rinck, P.A., 1993. *Magnetic Resonance Imaging*. Blackwell Scientific Publications.
- Roll, S.R., Colchester, A.C.F., Summers, P.E., Griffin, L.D., 1994. Intensity-based object extraction from 3D medical images including a correction for partial volume errors. In: *Proc. of BMVC94*. Proceedings of the 5th British Machine Vision Conference. BMVA Press, Guildford, UK, p. 1.
- Santago, P., Gage, H.D., 1993. Quantification of MR brain images by mixture density and partial volume modeling. *IEEE Trans. Med. Imag.* 12 (3), 566–574.
- Shattuck, D.W., Leahy, R.M., 2000. Brainsuite: an automated cortical surface identification tool. In: *Proc. of MICCAI'2000*, Lecture Notes in Computer Science. Springer-Verlag, pp. 51–61.
- Soltanian Zadeh, H., Windham, J.P., Yagle, A.E., 1993. Optimal transformation for correcting partial volume averaging effects in magnetic resonance imaging. *IEEE Trans. Nucl. Sci.* 40 (4), 1204–1212.
- Thacker, N., Jackson, A., Zhu, X.P., Li, K.L., 1998. Accuracy of tissue volume estimation in NMR images. In: *Proc. of MIUA'98*, Leeds, UK.
- Wells, III W.M., Grimson, W.E.L., Kikinis, R., Jolesz, F.A., 1996. Adaptive segmentation of MRI data. *IEEE Trans. Med. Imag.* 15 (4), 429–442.
- Zhang, Y., Brady, M., Smith, S., 2000. A hidden Markov random field model for segmentation of brain MR images. In: *Proc. of Medical Imaging 2000*, SPIE Proceedings, p. 3979.
- Zubal, I.G., Harrell, C.R., Smith, E.O., Rattner, Z., Gindi, G., Hoffer, P.B., 1994. Computerized 3-dimensional segmented human anatomy. *Med. Phys.* 21 (2), 299–302.