Validation of Tissue Modelization and Classification Techniques in T1-Weighted MR Brain Images

M. Bach Cuadra[1], B. Platel[2], E. Solanas[1], T. Butz[1], and J.-Ph. Thiran[1]

[1] Signal Processing Institute (ITS),
Swiss Federal Institute of Technology (EPFL), Switzerland
{Meritxell.Bach,Eduardo.Solanas,Torsten.Butz,JP.Thiran}@epfl.ch
http://ltswww.epfl.ch/~brain
[2] Department of Biomedical Imaging
Eindhoven University of Technology (TUE), The Nehterlands
{b.platel}@stud.tue.nl

Abstract. We propose a deep study on tissue modelization and classification Techniques on T1-weighted MR images. Three approaches have been taken into account to perform this validation study. Two of them are based on Finite Gaussian Mixture (FGM) model. The first one consists only in pure Gaussian distributions (FGM-EM). The second one uses a different model for partial volume (PV) (FGM-GA). The third one is based on a Hidden Markov Random Field (HMRF) model. All methods have been tested on a Digital Brain Phantom image considered as the ground truth. Noise and intensity non-uniformities have been added to simulate real image conditions. Also the effect of an anisotropic filter is considered. Results demonstrate that methods relying in both intensity and spatial information are in general more robust to noise and inhomogeneities. However, in some cases there is no significant differences between all presented methods.

1 Introduction

The study of many brain disorders requires an accurate tissue segmentation from magnetic resonance (MR) images. Manual tracing of the three brain tissue types, white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF) in MR images by an expert is too time consuming because most studies involve large amounts of data. Automated and reliable tissue classification is complicated due to different tissue intensities overlapping (partial volume effect, PVE), presence of noise and intensity non-uniformities caused by the inhomogeneities in the magnetic field of the MR scanner. Different approaches have been presented in recently to deal with this key topic. There is a need to explicitly take into account the pve but most used methods, as [1] only use a FGM representing three main tissue types. In other cases [2] pve is added as normal Gaussian distribution. In more evolved methods such as [3] mixing proportions are equally alike in PV voxels resulting in a density function. Recently, other methods that consider the

T. Dohi and R. Kikinis (Eds.): MICCAI 2002, LNCS 2488, pp. 290–297, 2002.

mixing proportions changing according to a MRF received an increasing interest [4]. Finally, non-parametric techniques can be considered when no assumption on the intensity distributions can be done [5]. Here we propose a detailed comparative study and validation of three different classification techniques, a fourth technique using a statistical non-parametric method will be considered in a near future. With this comparative study and taking into account the conditions of our problem we will be able to propose the most suitable technique to solve the classification problem.

2 Image Model

Theory behind voxel intensities is the same as the one used by Santago et al. in [3]. In an ideal case, only three main tissues have to be considered: CSF, GM and WM. But because of the finite resolution of an MR image the well-known partial volume effect (PVE) appears. Voxel intensity (ν) is modeled as the weighted sum of intensities of the tissues present in the voxel volume, plus an error term. It is assumed that no particular combination of underlying tissues is more probable than any other. If PVE is considered the probability density function $\rho_\nu(\nu)$ is defined by:

$$\rho_\nu(\nu) = \sum_t Pr\,[t]\,\rho_{\nu_t}(\nu_t), \quad \text{where } t \in \{csf, gm, wm, cg, cw, gw, cgw\}, \tag{1}$$

where $Pr\,[t]$ is the probability of tissue t and $\rho_{\nu_t}(\nu_t)$ is the probability density function of ν given tissue t and cg, cw, gw, cgw, respresnts the mixtures for csf/gm, csf/wm, gm/wm and csf/gm/wm respectivetly. Because $Pr\,[cgw]\,\rho_{\nu_{cgw}}$ (ν_{cgw}) and $Pr\,[cw]\,\rho_{\nu_{cw}}(\nu_{cw})$ are insignificant in $\rho_\nu(\nu)$, they will not be considered. $\rho_\nu(\nu)$ for single a tissue voxel and a mixed two-tissue (ab) voxel:

$$\rho_{\nu_i}(\nu_i) = \frac{1}{\sigma_\eta\sqrt{2\pi}}\; Exp\left[\frac{-(\nu - I_i)^2}{2\sigma_\eta^2}\right], \tag{2}$$

and

$$\rho_{\nu_{ab}}(\nu_{ab}) = \int_0^1 \frac{1}{\sigma_\eta\sqrt{2\pi}}\; Exp\left[\frac{-(\nu - (\alpha I_a + (1-\alpha)I_b))^2}{2\sigma_\eta^2}\right] d\alpha, \tag{3}$$

with mean and variance given by:

$$\mu_{\nu_{ab}} = \frac{1}{2}(I_a + I_b) \quad \text{and} \quad \sigma^2_{\nu_{ab}} = \frac{1}{12}I_a^2 - \frac{1}{6}I_aI_b + \frac{1}{12}I_b^2 + \sigma_\eta^2. \tag{4}$$

We can observe that PV function varies between a block function and a Gaussian function depending on I_a (mean intensity of main tissue type a), I_b (mean intensity of main tissue type b) and σ_η^2 (noise variance of main tissues), see Fig. 1. This means that for certain cases (for larger σ_η's, and when I_a and I_b are very different) equation 3 can be replaced by a simple Gaussian. When validating (see Fig. 2in section 6), this is a reasonable assumption, if the variance of the

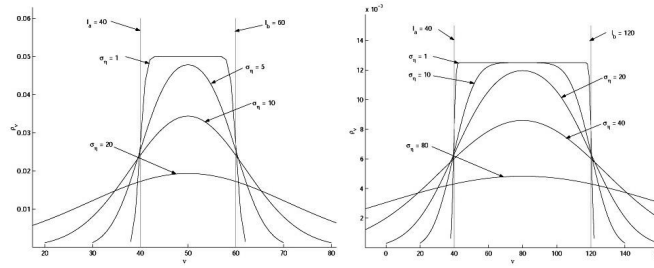


Fig. 1. Plot of Equation 3 with (a) varying σ_η (where $I_a = 40$, $I_b = 60$) and (b) varying σ_η (where $I_a = 40$, $I_b = 120$)

Gaussian functions representing the tissue mixtures is considered independent of σ_η. Then, MR image histogram can be modeled solely by Gaussian distributions even for PV representation. As mentioned in [3], the continuous physical reality $\rho_\nu(\nu)$ is approximated by the normalized image histogram (h_i in Eq. 6). Finally, we have to find the set of parameters that solves the minimization problem:

$$\mathbf{p} = \{Pr\left[i\right], \mu_j, \sigma_i\} \quad \text{for } i \in \{csf, gm, wm, cg, cw, gw\} \quad \text{and } j \in \{csf, gm, wm\} \tag{5}$$

$$\sum_{i=1}^{I} (h_i - \sum_{t} Pr\left[t\right] \rho_{\nu_j}(i))^2 \quad \text{for } t \in \{csf, gm, wm, cg, cw, gw\} \tag{6}$$

The goal is to optimally classify an instance of ν, $\tilde{\nu}$, in one of the possible categories. It can be shown that the optimal classification of $\tilde{\nu}$, in terms of the probability error is the Bayes classification.

3 Method A: Pure Finite Gaussian Mixture Model

Here, the intensity histogram is modeled by 3, 4 or 5 Gaussian distributions. We deal with the mixture-density parameter estimation problem (see Eq. 6). A possible approach is to find the maximum of a mixture likelihood function. One of the most used methods is the Expectation Maximization (EM) algorithm. Here we follow the steps as explained by [6]. The following probabilistic model as defined in Eq. 2 is assumed:

$$p(\mathbf{x}|\Theta) = \sum_{i=1}^{M} w_i p_i(\mathbf{x}|\theta_i), \tag{7}$$

where the parameters are $\Theta = (w_1, \ldots, w_M, \theta_1, \ldots, \theta_M)$ such that the sum of all the weights (w_i) is one and each p_i is a density function parameterized by θ_i. So it is assumed that there are M component densities mixed together with M mixing coefficients $P(w_l)$. In the case of the univariate normal mixture, the

maximum likelihood estimates for the next iteration step $P(w_l)^{(t+1)}$ of the mixture coefficients, $\mu_l^{(t+1)}$ of the mean and $\sigma_l^{2(t+1)}$ of the variance are expressed as follows:

$$P(w_l)^{(t+1)} = \frac{1}{N}\sum_{i=1}^{N} p(w_l|x_i,\Theta)^{(t)} \quad \text{and} \quad \mu_l^{(t+1)} = \frac{\sum_{i=1}^{N} x_i p(w_l|x_i,\Theta)^{(t)}}{\sum_{i=1}^{N} p(w_l|x_i,\Theta)^{(t)}} \tag{8}$$

$$\sigma_l^{(t+1)} = \frac{\sum_{i=1}^{N} p(w_l|x_i,\Theta)^{(t)}(x_i-\mu_l^{(t)})(x_i-\mu_l^{(t)})^T}{\sum_{i=1}^{N} p(w_l|x_i,\Theta)^{(t)}} \tag{9}$$

where

$$p(w_l|x_i,\Theta)^{(t)} = \frac{p(x_i|w_l,\Theta_l)P(w_l)^{(t)}}{\sum_{j=1}^{M} p(x_i|w_j,\Theta_j)P(w_j)^{(t)}} \tag{10}$$

These four equations are used for the numerical approximation of the parameters of the mixture. Initial estimations of the parameters $P(w_l)$, μ_l and σ_l have to be specified and only pure Gaussian distributions are considered.

4 Method B: Finite Gaussian Mixture Model Considering a Partial Volume Model

In this case the main three tissue types are modeled by normal Gaussians (Eq. 2) and both CSF/GM and GM/WM partial volumes are modeled by the PVE presented in Eq. 3. PVE is numerically solve it with Newton's method and a *genetic algorithm* is used to find the mixture parameters as defined in 5. The genetic algorithm used here is the same presented in [1]. Genes are composed by the means of the three Gaussians (Eq. 2) μ_1, μ_2, μ_3, by the noise variance σ_η (considered the same for all tissues) and by a weight for either Gaussian or PV distributions w_1, w_2, w_3, w_4, w_5, where $\sum_{i=1}^{5} w_i = 1$.

5 Method C: Hidden Markov Random Field Model

A hidden Markov random field (HMRF) model [4] is used to improve the classification by using spatial information. The importance of the HMRF model derives from MRF theory, in which the spatial information in an image is encoded through contextual constraints of neighboring pixels. By imposing such constraints it is expected that neighboring pixels will have the same class labels. This is achieved by characterizing mutual influences among pixels using conditional MRF distributions. A Pure Gaussian Mixture model will be used for the HMRF method:

$$p(y_i|x_{\mathcal{N}_i},\theta) = \sum_{l\in\mathcal{L}} g(y_i;\theta_l)p(l|X_{\mathcal{N}_i}), \tag{11}$$

where $\theta_l = (\mu_l, \sigma_l)$ and $g(y;\theta_l)$ is the guassian density function. This type of HMRF is often referred to as the *Gaussian hidden Markov random field*

(GHMRF) model. Note that finite mixture (FM) models is a particular case of HMRF where $p(l|X_{\mathcal{N}_i}) = w_l$ (note Eq. 7). The HMRF-EM algorithm will be used:

$$u_l^{(t+1)} = \frac{\sum_{i \in S} P^{(t)}(l|y_i) y_i}{\sum_{i \in S} P^{(t)}(l|y_i)} \quad \text{and} \quad \left(\sigma_l^{(t+1)}\right)^2 = \frac{\sum_{i \in S} P^{(t)}(l|y_i)(y_i - \mu_l)^2}{\sum_{i \in S} P^{(t)}(l|y_i)}, \quad (12)$$

which are the same update equations as for the finite Gaussian mixture model (Eq. 9), except for

$$P^{(t)}(l|y_i) = \frac{g^{(t)}(y_i; \theta_l) \cdot P^{(t)}(l|x_{\mathcal{N}_i})}{p(y_i)}. \quad (13)$$

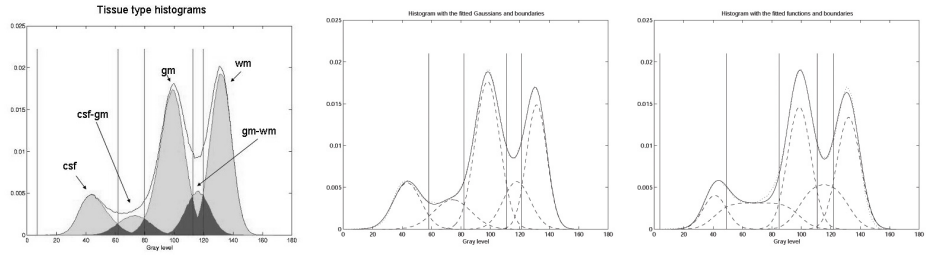
$P^{(t)}(l|x_{\mathcal{N}_i})$ calculation involves the estimation of class labels, which are obtained through MRF-MAP estimation. Means, variances and label map estimation by EM algorithm is used to initialize the HMRF-EM algorithm.

6 Validation

Validation has been done using the *digital brain phantom* images [7]. Tissue classification (CSF, GM, WM and mixtures of these) of these phantoms are known *a priori*. This makes them suitable for segmentation algorithm assessment. The Brainweb web-site [7] provides several simulated MRI acquisitions of this phantom including RF non-uniformities and noise levels. It is possible to split phantom histogram into specific ones for each pure tissue type and its mixtures (see Fig. 2(a)). Note that three main tissues are pure Gaussians while mixture densities look a bit like a Gaussian, but with a larger variance. Quantitative results of Brainweb segmentation using Bayes classificator are shown in first row of table 4. The percentage of correctly classified voxles is computed with respect to this reference data. Pixels of the background are not considered. The overall percentage of the classification is calculated over all the voxels of the brain. The percentage of different tissue type volumes is also calculated. Classification has been done on MR images containing 5, 7 and 9% noise and of 0 or 20% RF non-uniformity. We also considered the effect of an anisotropic diffusion filter on the classification results. Test have been done in both filtered and not filered images.

7 Results

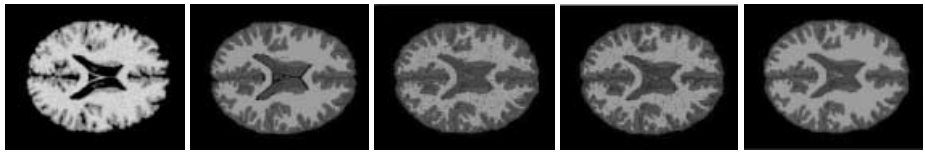
All tests have been done on Brainweb images (image size is 181x217x181 and isotropic voxel size of 1mm). Method A has been tested for 3, 4 and 5 Gaussians (see A3, A4, A5 in table 4). Results show that using 5 Gaussian distributions does not always yield better results and that the amount of PV is overestimated. This is due to the fact that the assumption of using a normal distribution for a PV is false in this cases (see Fig. 2(b)). The segmentation results of method B



(a) (b) (c)

Fig. 2. (a) Segmentation from a Brainweb phantom with 5% noise and 0% RF, not filtered. Dark area represents errors made when using only histogram data. (b) FGM-EM technique, method A. (c) FGM-GA technique, method B

are not as good as expected since it models the PV more in line with physics. It actually highly overestimates PV (Fig. 2(c)). Restrictions on PV wheights when using genetic algorithm could probably improve final classifiaction. Method C yields in the best results even if it doesn't reach 100% correct classification. It performs the less noisy classification (see Fig. 3). Most of the errors performed by this method are locate in PV voxels. This can be easily explained: in a HMRF model the neighborhood of a voxel is considered. If we look at the neighborhood of a PV voxel its neighbours are mostly pure tissue voxels, in which case HMRF model will assign a pure tissue label to the PV voxel. This behaviour tends to eliminate PV voxels in label map (that's reflected on table 1 where PV volumes are the smallest). A possibility to solve this problem could be not to consider neighborhood pixels in the partial volume regions. Another possibility could be to adapt the way the *clique potentials* [4] are calculated, so that PV voxels are no longer assigned to a low probability.

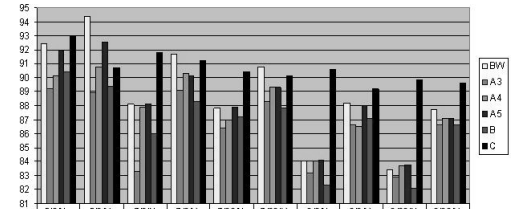


(a) (b) (c) (d) (e)

Fig. 3. Label map of image (a) with 7% noise found by different classification methods. (b) BW segmentation. (c) Method A. (d) Method B. (e) Method C.

8 Discussion

As we have seen in section 6, if the validation is done solely on intensity image spectra, the tissue classification can never be a 100% due to the noise (see



	5/0/No	5/0/Y	7/0/N	7/0/Y	7/20/N	7/20/Y	9/0/N	9/0/Y	9/20/N	9/20/Y
BW	92.4	94.4	88.1	91.7	87.8	90.8	84	88.2	83.4	87.7
A3	89.2	88.9	83.3	89.1	86.4	88.3	83.2	86.6	82.9	86.6
A4	90.1	90.8	87.9	90.3	87	89.3	84	86.5	83.7	87.1
A5	92	92.5	88.1	90.1	87.9	89.3	84.1	88	83.8	87.1
B	90.4	89.4	86	88.3	87.2	87.8	82.3	87.1	82.1	86.6
C	93	90.7	91.8	91.2	90.4	90.1	90.6	89.2	89.9	89.6

Fig. 4. Percentages of correctly classified pixels for different images: %noise/%RF/ filtered. Brainweb histogram segmentation (BW), FGM-EM segmentation using 3, 4 and 5 Gaussian (A3, A4 and A5, respectively), FGM-GA segmentation (B), and HMRF segmentation (C).

Table 1. The percentages of the volumes of the different tissue types.

	Image	csf	wm	gm	gm-wm	csf-gm
BW		11.8	34.7	36.8	10.5	6.17
A	5% noise, not filtered	13.0	26.5	36.1	12.8	11.6
	7% noise, not filtered	16.6	26.0	33.3	13.9	10.1
	7% noise, filtered	15.7	30.0	27.0	15.8	11.3
	9% noise, not filtered	15.0	30.3	28.0	14.6	12.1
	9% noise, filtered	19.2	15.9	31.7	27.2	6.00
B	5% noise, not filtered	7.62	24.3	26.8	23.2	18.3
	7% noise, not filtered	6.88	25.2	27.3	19.0	21.9
	7% noise, filtered	7.76	23.6	25.3	23.8	18.9
	9% noise, not filtered	8.20	26.6	29.1	15.8	20.5
	9% noise, filtered	5.96	25.1	26.1	18.8	24.2
C	5% noise, not filtered	12.1	28.7	47.2	7.07	9.441
	7% noise, not filtered	15.7	31.0	44.1	2.91	6.21
	7% noise, filtered	16.7	24.7	35.9	14.1	8.67
	9% noise, not filtered	17.2	31.2	43.4	2.60	5.64
	9% noise, filtered	19.6	24.3	36.9	13.8	5.39

Fig. 3). Actually, the less noise there is in the image the better the segmentation. When the noise variance is small, the overlap between the histograms of every tissue type and its mixtures is then also smaller. Results show (see Fig. 4) that the application of an anisotropic diffusion filter improves the classification techniques (2% mean error reduction for 5% of noise) and this improvement is more significative when dealing with more noisy images (until 4% mean error reduction for 9% noise). However, this filtering have no so much impact on intensity non-uniformities. We have seen also that FGM models (method A and B) are solely histogram-based methods. They don't consider any spatial information. Such a limitation causes these methods to work efficiently only on images with low level of noise. Unfortunately this is not often the case with MR images. Actually, method C shown to be more suited for MR brain image modelling in the sense that it has the ability to encode both statistical and spatial image properties. When the three methods are compared, method C is by far the best, not only because it scores the best in the percentage of correctly classified voxels and it gives the best volume estimation of the main tissue types. But mostly because the label map resulting from the classification hardly contains any noise

(Fig. 2). However, it presents also some weak points. It tends to eliminate PV and it approximates PV having a Gaussian distribution and that is not always realistic. Finally, it takes prohibitive computation time (it could take up to 8 hours with 181x217x181 image dimension) since it has been implemented in 3D. However, it would be very interesting to compare these methods with a statistical non-parametrical method such as the one we presented in[5] which is suposed to be better since it is a supervised method and uses spatial information.

9 Conclusion

We have presented here a validation study of MR brain images classification techniques. To perform this validation three different classification methods widely used for this application have been presented. All tests have been done in a ground truth image considering different noise and intensity non-uniformity levels. Also the effect of an anisotropic filter has been also considered when comparing between filtered and not filtered images. Results have shown that the techniques considering spatial information lead in better results when high noisy images are considered and that the application of an anisotropic filter lead in better a classification. However, it has been also demonstrated that in other cases histogram-based techniques lead to comparable results.

References

1. Schroeter, P., et al.: Robust parameter estimation of intensity distributions for brain magnetic resonance images. IEEE Transactions on Medical Imageing **17** (1998) 172–186
2. Ruan, S., Jaggi, C., Xue, J., Bloyet, J.: Brain tissue classification of magnetic resonance images using partial volume modeling. IEEE Transactions on Medical Imaging **19(12)** (2000) 172–186
3. Santago, P., Gage, H.D.: Quantification of mr brain images by mixture density and partial volume modeling. IEEE Trans. Medical Imaging **12** (1993) 566–574
4. Zhang, Y., et al.: Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. IEEE Trans. Medical Imaging **20** (2001) 45–57
5. E. Solanas, V. Duay, O.C., Thiran, J.P.: Relative anatomical location for statistical non-parametric brain tissue classification in mr images. Proc. of the 7th ICIP (2001)
6. Bilmes, J.A.: A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. Technical report, International Computer Science Institute, Berkeley California (1998)
7. McConell. World Wide Web, http://www.bic.mni.mcgill.ca/brainweb/ (1998)