Segmentation of magnetic resonance images using fuzzy Markov random fields

Su Ruan, Bruno Moretti, Jalal M. Fadili, Daniel Bloyet

Segmentation of Magnetic Resonance Images using Fuzzy Markov Random Fields

Su Ruan*, Bruno Moretti, Jalal Fadili and Daniel Bloyet

GREYC-ISMRA, CNRS UMR 6072, F-14050 Cedex Caen, France
(*) E-mail: su.ruan@greyc.ismra.fr

Abstract - In this paper, we present a fuzzy Markovian method for brain tissue segmentation from magnetic resonance images. Generally, there are three principal brain tissues in a brain dataset: gray matter, white matter and cerebrospinal fluid. However, due to the limited resolution of the acquisition system, many voxels may be composed of multiple tissue types (partial volume effects). The proposed method aims to calculate the fuzzy membership of each voxel to indicate the partial volume degree using a fuzzy Markovian segmentation. Since our method is unsupervised, it first estimates the fuzzy Markovian Random Field model parameters using a stochastic gradient algorithm. The efficiency of the proposed method is quantified on a digital phantom using an absolute average error, and qualitatively tested on real MRI brain data.

Introduction

Image segmentation is a classical problem in computer vision and is of paramount importance to medical imaging. The study of many brain disorders requires accurate tissue segmentation from magnetic resonance (MR) images of the brain. Manual tracing of the three brain tissues white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF) in MR images by a human expert is too time consuming for studies involving large amounts of data.

Automated and reliable tissue classification is complicated by the overlap of MR intensities of different tissue classes (partial volume effects). The problem of partial volume effect is often ignored by many proposed methods, which produce a classification where each voxel is assigned to a single tissue [1][2]. Two main reasons lead to the problem of partial volume effects: the imaging resolution that causes many voxels to be a mixture of two or more than two tissues and the constitution of a brain cannot be restricted to only the three pure tissues (GM, WM and CSF). Therefore, in order to obtain accurate segmentation, it is necessary to determine the degree to which voxels are similar to, or belong to, one or more tissue categories. Recently some approaches have been presented to deal with these problems [3][4]. An extensive review of this issue is given in [5].

It is now agreed by both fuzzy logic and statistics communities that fuzzy and probabilistic approaches are rather complementary than competitive. The distinction between them is based on how uncertainty is captured by each concept. The essence of our model is to exploit the complementarities of fuzzy and probabilistic approaches combining their strengths in what we call fuzzy MRF. In this paper, we consider a brain as 3D fuzzy fields which consist of three main tissues (hard classes), and the mixtures of the pure brain tissues (fuzzy classes or mixclasses). A Fuzzy MRF model is used to take into account the contextual information, statistical information of image signals and the anatomical information of the brain. As our method is unsupervised, the first step is to estimate optimally all parameters used in the proposed models: statistical parameters of distribution of the three hard classes and parameters of the Markovian distribution of the fuzzy field. The fuzzy Markovian segmentation is then carried out using the ICM (Iterated Conditional Modes) algorithm [6] with parameters of the Markovian distribution estimated in the first step.

Fuzzy MRF model

The classical hard Markov fields are widely used in many applications, such as image restoration [7], image segmentation [8] and volumetric object reconstruction [9]. The fuzzy Markov fields, recently introduced by Pieczynski [10], are particularly attractive in the case of presence of fuzzy regions in an image. They were mainly used to deal with the satellite image analysis.

The partial volume effect appears when more than one type of class or material occupies one voxel or pixel of an image. These voxels or pixels are usually called mixels. Thus, the fuzzy segmentation of images consists of allowing each pixel to belong simultaneously to numerous classes. The problem is to associate to each pixel s a vector

$(a_1,...,a_k)\in[0,1]^k$ with $(a_1+...+a_k)=1$. The a_i's represent the partial volume proportions of each "pure" tissue. We consider two random fields $A=(A_s)_{s\in S}$ and $Y=(Y_s)_{s\in S}$. The image data to be segmented is a realisation $Y=y$ of Y and the desired result is a realisation $A=a$ of the field A. The joint distribution of (A,Y) is defined by the prior distribution $P_A(a)$, which is assumed to be stationary and Markovian, and by the posterior distribution $P_{Y/A}(y/a)$:

$$P_{A,Y}(a,y)=P_A(a)P_{Y/A}(y/a) \qquad (1)$$

Firstly, we discuss the observed field Y. Let us denote y_j^p a random variable that is referred to as the pure class j. Both the heterogeneity of pure matter and the imaging noise are taken into account in its probability density $p(y_j^p)$. The intensity value y of a mixel can then be represented by a weighted sum of k pure tissues :

$$y=\sum_{j=1}^{k} a_j y_j^p \quad \text{with} \quad \sum_{k=1}^{N} a_j=1 \ \ and \ a_j>0 \qquad (2)$$

where a_j is the proportion of the pure class j. In our case, the distribution of each pure class $p(y_j^p)$ is assumed to be Gaussian. From (2), it is obvious to see that the observation y is a variable whose fluctuations depend directly on the statistical properties of the random variables y_j^p (i.e. noise variance and tissue heterogeneity).

In the case of two pure tissues $\Omega=\{c_1,c_2\}$, the intensity value of a mixel takes the following form:

$$y=ay_1^p+(1-a)y_2^p \qquad (3)$$

In this linear fuzzy model, we take $a\in[0,1]$, where the values 0 and 1 correspond to the hard classes (0:class 2 and 1: class 1), and $]0,1[$ stands for the fuzzy classes. Assume that the two pure classes are independent of s, normally distributed. Denoting $N(\mu_k,\sigma_k^2)$ the normal distribution of mean μ_k and variance σ_k^2 (k=1,2), the probability density function of y is the convolution of $p(y_1^p)$ and $p(y_2^p)$ with a weight a. It can be written as follows for a given a:

$$p(y/a)=\frac{1}{\sqrt{2\pi}\sqrt{a_2\sigma_1^2+(1-a)_2\sigma_2^2}}\exp\left\{-\frac{[y-(a\mu_1+(1-a)\mu_2)]_2}{2[a_2\sigma_1^2+(1-a)_2\sigma_2^2]}\right\} \qquad (4)$$

$p(y/a)$ is also a Gaussian function. The parameters μ_0 , μ_1, σ_0^2, σ_1^2 of the pure classes define all distributions of Y conditional on A. Let us now consider the a priori model $P_A(a)$. For the hard class c_j, we have $a_j=1$, $a_i=0\ \forall\ i\neq j$. If the field A is considered Markovian with respect to a neighborhood V, in accordance with the Hammersley-Clifford theorem [6], its distribution is defined as a Gibbs distribution:

$$P[A=a]=\frac{1}{Z}\exp(-U_h(a)) \qquad (5)$$

where U_h stands for the energy function, defined on cliques within V, and Z is the normalizing constant. In the fuzzy case, the distribution of A is defined on $\Omega_f=[0,1]$. If we consider that $\frac{1}{Z}\exp(-U_f(a))$ is the density of P_A with respect to the mixels, U_f has the same shape of cliques as the function U_h. Then, it is possible to show exactly, as in the hard case, that A is Markovian with respect to V [10]. Since a_j is a real number in the fuzzy case, it can be infinite, however, in practice, an infinite value is not realistic. Therefore, we have to discretize a.

Segmentation algorithm

As the probability of mixing more than two tissues in a voxel is very low and the mixture of WM and CSF is very poor in a brain data, we can consider that there are three pure classes CSF, GM and WM, and two mix-classes : CG (mixture of CSF and GM), GW (mixture of GM and WM). The fuzzy segmentation consists in finding the three pure classes and the fuzzy membership values of the two mixclasses. Since the intensities of the two mixclasses do not overlap and each mixclass is mixed of only two hard classes, the fuzzy models associated to the two hard classes, described above, can be used directly for each mixclass. We can consider that there are two fuzzy fields to be found, which are independent of each other.

Let us denote a the fuzzy field $a=(a^{CG},a^{GW})$, $a\in[0,1]_2$, which can take the three pure classes and fuzzy values relative to the mixclasses CG and GW. The data energy term concerning Y is expressed as $U_1(y/a)=\ln P(y/a)$. From (4), it is obvious that $U_1(y/a)$ is defined by a set of parameters: $\Phi_y(a)=\{\mu_{CSF},\sigma_{CSF}^2,\mu_{WM},\sigma_{WM}^2,\mu_{GM},\sigma_{GM}^2\}$ which are unknown a priori.

The energy term corresponding to the a priori model is considered in the hard and fuzzy cases. As described in the above section, we have $P(a)=\frac{1}{Z}e^{-U_2(a)}$, where $U_2(a)=U_h(a)+U_f(a)$. As the object is three-dimensional (3D), the system neighborhood V is the spatial 18-connexity. Two kinds of cliques are defined: "horizontal" neighbors, "vertical" neighbors and "inferior-posterior" neighbors belong to the first kind; the other cliques such as "south-west" and "north-west" belong to the

second one. $U_h(a)$ is defined when both pixels are hard:

$$U_h(a)=\left(\sum_{\langle s,t\rangle\in V_1}\beta_1^h(1-\delta(a_s,a_t)+\sum_{\langle s,t\rangle\in V_2}\beta_2^h(1-\delta(a_s,a_t)\right) \quad (6)$$

where V_1 and V_2 denote the two kinds of cliques, and $\delta(.)$ is the Kronecker delta function. $U_f(a)$ is defined when one pixel or both pixels are fuzzy:

$$U_f(a)=\left(\begin{array}{l}\sum_{\langle s,t\rangle\in V_1}\beta_1^{fCG}|(a_s-a_t)|+\sum_{\langle s,t\rangle\in V_2}\beta_2^{fCG}|(a_s-a_t)| \\ +\sum_{\langle s,t\rangle\in V_1}\beta_1^{fGM}|(a_s-a_t)|+\sum_{\langle s,t\rangle\in V_2}\beta_2^{fGM}|(a_s-a_t)|\end{array}\right) \quad (7)$$

The a priori model is then defined by a set of parameters :

$$\Phi_a(a)=\{\beta_1^h,\beta_2^h,\beta_1^{fCG},\beta_2^{fCG},\beta_1^{fGW},\beta_2^{fGW}\} \quad (8)$$

Using the Bayes rule, the fuzzy segmentation objective is in fact to search for a such as : $a=\arg\max_a(P(a)P(y/a))$. This maximization is equivalent to $a=\arg\min_a(U_1(y/a)+U_2(a))$.

In the unsupervised Markovian segmentation case, we have to estimate the parameters used in the models (i.e. β in $P(a)$, means and variances of the Gaussian distributions in $P(y/a)$. The complexity of the estimation problem is due to the absence of the observation a. Since the number of classes is known in our case, the initialization of parameters Φ_y can be obtained by a fit of the image histogram with a mixture of different distributions. The obtained parameters are very close to the optimal ones. Therefore, the method described in [11] is adapted in our parameter estimation.

Let us specify that the prior distribution $P(a/\Phi_a)$ depends on a parameter vector Φ_a and the intensity distribution $P(y/a,\Phi_y)$ depends on the parameter vector Φ_y. We consider estimation procedures for complete data and incomplete data separately. For the complete data, the estimator of Φ_y is chosen as the empirical mean and variance. For the incomplete data, the stochastic gradient method is used to estimate Φ_a. The parameter estimation procedure is outlined below:

(a) Initialize the parameters $\Phi^0=(\Phi_a,\Phi_y)$, iteration n=0.

(b) Segment the image data with the current parameter $\hat{\Phi}_n$ using ICM, denoted $\hat{a}$.

(c) Estimate $\hat{\Phi}_y^{n+1}$ from the segmented image data.

(d) Generate a Markov field using a Gibbs Sampler according to the a priori law $P(a)\propto e^{-U_2(a)}$ with the current parameter value $(\hat{\Phi}_a^n)$, denoted a^{n+1}.

(e) Estimate $\hat{\Phi}_a^{n+1}$ using the stochastic gradient algorithm.

(f)If $\|\hat{\Phi}_{n+1}-\hat{\Phi}_n\|>\varepsilon$, n=n+1, and go back to step (b),

where ε is a threshold depending on the required accuracy. Instead of using two Gibbs Samplers, the segmentation result is used during the iterative procedure. Normally, at the end of the algorithm, the segmentation is also obtained at the same time.

Validation and results

The method has been validated in two ways: (a) a simulation study focused on a quantitative assessment in noise free and stationary conditions; (b) its application to real MRI data was qualitatively validated by experts.

(a)Simulated data

The digital brain phantom, available on the site BrainWeb [12], is used as a gold standard. The simulated MRI image volumes with different noise levels are also available in this site. Each volume data consists of 181×217×181 isotropic voxels of size 1×1×1mm^3.

The segmentation results are evaluated by measuring the absolute average error ξ:

$$\xi=\frac{\sum_{s\in S}|a_s-a_s^f|}{Card(S)} \quad (9)$$

where a_s^f denotes the known tissue proportion at voxel s in the phantom and a_s the estimated proportion in the segmented fuzzy volume. *Card(S)* denotes the number of voxels s in the reference volume S. The average error ξ can be directly related to the false positives and the false negatives errors, in the sense that a difference in proportions for a given tissue is similar to a difference of partial volumes.

According to the fuzzy degree of the phantom, the step of the fuzzy discridization is set to 0.2 for the phantom analysis. As a result, there are finally 11 classes to segment. In Table 1, the error rates of the segmentation were calculated for the entire phantom volume. It can be observed that the performance decays less when the noise level increase. In practical situation, the noise level corresponds generally to 5%. It can be also observed from Table 1 that more errors are found in the gray matter than in the other tissues. This is due to the fact that the gray matter is involved in both GM/WM and GM/CSF mixtures.

(b) Real MRI data

T_1-weighted SPGR MRI data sets with acquisition parameters (TE=7ms, TR=30ms, α=40°) were acquired on a 1.5 Tesla Signa scanner. Each data set consists of 256 ×256×124 voxels with a resolution of 1×1×1.2mm³. The acquisition orientation is axial.

The algorithm was applied to ten MRI datasets. Consistent and stable results of the segmentation algorithm were observed. An axial slice is used here to illustrate the performance of the proposed method. The proportions obtained for the three pure tissues are shown in Figure 1. The voxel intensities correspond to proportion degrees of each pure tissue. Bright intensity represents high proportions. Here, the sampling step of the fuzzy membership a_s is 0.1. Because the intensities of the MRI images change within an interval of less than 200 gray levels, it is not reasonable to take the fuzzy membership value smaller than 0.1. Figure 1 shows that voxels in the center of the brain or along the cortical gray-white matter boundary have, as predicted, significant proportions of both respective tissue types. It should be noted that the sub-cortical locations (in the center of the brain), composed of some internal structures, would be wrongly classified if a hard segmentation were used, while the fuzzy segmentation renders fine details in these regions. As the mixture of WM/CSF is ignored in our brain model, some errors occur along the boundary of the ventricles. However, these misclassified voxels are negligible compared to the entire dataset.

Conclusion

We have described an unsupervised fuzzy segmentation method, based on a mixture model and a MRF model, which seems well adapted and efficient for MRI image segmentation. Quantitative results indicate that the estimated parameters lead to a result as well as the one obtained using the optimal set of parameters. When the real data are fuzzy, such as MRI brain data, the use of the fuzzy segmentation is always more effective than the use of the hard one. Further quantitative validation on the accuracy and stability of the method is necessary, using realistic phantoms and a large number of clinical scans.

Tissu	noise level 3%	noise level 5%	noise level 7%
White matter	4.70	5.97	9.30
Gray Matter	7.23	9.28	12.41
CSF	2.18	2.56	3.89

Table 1 Segmentation errors dealing with the phantom noised with 3%, 5% and 7% levels.

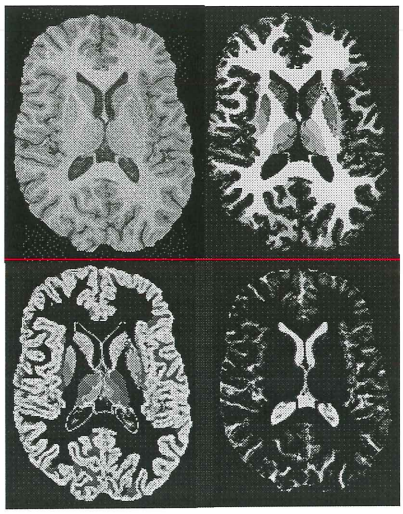

Figure 5 Fuzzy segmentation results. The voxel intensities correspond to proportion degrees of each pure tissue. One original slice (a), and the proportions of the WM (b), GM (c) and CSF (d) are shown respectively.

References

[1] Rajapakse J C and al., "Cerebral magnetic resonance image segmentation using data fusion", J. Comput. assist. Tomogr., Vol.20(2), pp.206-218, 1996.

[2] Held K, Kops E R, Krause B J, Wells W M, Kikinis R, Müller-Gärtner H W, "Markov Random Field segmentation of brain MR images", IEEE. Trans. on Med. Imag., Vol.16(6), pp.878-892, 1997.

[3] Vincken K L, E.Koster A S, Viergever M A, "Probabilistic multiscale image segmentation", *IEEE. Trans. PAMI.* Vol.19(2), 109-120, 1997.

[4] Bedell B J, Narayana P A, "Volumetric analysis of white matter, gray matter and CSF using fractional volume analysis", Magn. Reson Med, Vol.39(6), pp.961-969, 1998.

[5] Laidlaw D.H, Fleischer.K.W, Barr.A.H, "Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms", IEEE. Trans. on Medical. Imaging., Vol.17(1), 878-89,1998.

[6] Beśag J, "Spatial interaction and the statistical analysis of lattice systems", J.Roy. Stat. Soc. Ser. B, Vol.36, pp.192-242, 1974.

[7] Geman S, Geman D, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images", IEEE. Trans. PAMI, Vol.6(6), pp. 721-741, 1984.

[8] Panjwani D.K and Healey G, "Markov random field models for unsupervised segmentation of textured color images", IEEE.Trans.PAMI.,Vol.17(10), pp.939-954, 1995.

[9] Choi S M, Lee.J E, Kim M H, "Volumetric object reconstruction using the 3D-MRF model-based segmentation", IEEE.Trans.Med.Imag.,Vol.16(6), pp.887-892, 1997.

[10] Salzenstein F, Pieczynski W, "Parameter estimation in hidden fuzzy Markov random fields and image segmentation", Graphical and models processing. Vol.59 (4) 205-220, 1997.

[11] Younes L, "Parametric inference for imperfectly observed Gibbsian fields", Prob. Th. Fields. Spring-Verlag, No.82, pp.625-645, 1989.

[12] Cocosco A, Kollokian V, Kwan R K-S and Evans A C, "Brain Web: Online interface to a 3D MRI simulated brain database", Http: //www.bic.mni.mcgill.ca /brainweb.