

# A New Brain Segmentation Framework

Torsten Butz<sup>1</sup>, Patric Hagmann<sup>1</sup>, Eric Tardif<sup>2</sup>, Reto Meuli<sup>3</sup>, and  
Jean-Philippe Thiran<sup>1</sup>

<sup>1</sup> Swiss Federal Institute of Technology (EPFL), Signal Processing Institute (ITS),  
CH-1015 Lausanne, Switzerland

{torsten.butz, patric.hagmann, jp.thiran}@epfl.ch,  
<http://ltswww.epfl.ch/~brain>

<sup>2</sup> University of Lausanne, Institute of Physiology, CH-1005 Lausanne, Switzerland  
eric.tardif@iphsiol.unil.ch

<sup>3</sup> University Hospital, Department of Radiology, CH-1011 Lausanne, Switzerland  
reto.meuli@chuv.hospvd.ch

**Abstract.** We present a new brain segmentation framework which we apply to T1-weighted magnetic resonance image segmentation. The innovation of the algorithm in comparison to the state-of-the-art of non-supervised brain segmentation is twofold. First, the algorithm is entirely non-parametric and non-supervised. We can therefore enhance the classically used gray level information of the images by other features which do not fulfill the parametric Gaussian assumption. This is illustrated by a segmentation algorithm that considers both, voxel intensities and voxel gradients for the segmentation task. The resulting algorithm is called a *non-supervised, non-parametric* hidden Markov random field segmentation. Furthermore we have also to construct an anatomically relevant segmentation model in the resulting two-dimensional feature space. This is the second main contribution of this paper. We construct a morphologically inspired classification model, which is also able to segment the deep structures of the brain into a separate class, resulting in a six class segmentation model. We prove the validity of the introduced mathematical and morphological aspects on simulated T1-weighted magnetic resonance images of the brain.

## 1 Introduction

Based on morphological observations, one can consider that the brain contains 3 main tissue types, gray matter (GM) tissue, white matter (WM) tissue, and cerebrospinal fluid (CSF). Automatic segmentation of MR datasets into these classes is very desirable for several reasons. One of them is that global or local morphological changes of these anatomical tissues are characteristic for some pathologies. That's why several research groups have been addressing the problem of automatic brain segmentation. In particular, statistical approaches with parametric probability density models are widely employed. One example is called the finite Gaussian mixture (FGM) model [1], which uses the voxel intensities and the parametric Gaussian assumption for the intra-class statistics in order to

classify the MR data. The main disadvantage of this histogram-based approach is that it completely ignores any spatial information about the data. Therefore Gaussian hidden Markov model (HMM) segmentation has been proposed [2,3] as HMMs are able to explore spatial dependencies of the data [4,5].

Another limitation of MR brain segmentation lies in the chosen classes, GM, WM, and CSF, into which we want to segment the original MR data. As image voxels extend over some finite physical space, there are voxels in the data which contain several tissue types at once. In particular along the CSF-GM and GM-WM boundaries, voxels will contain both tissue types of the adjacent anatomical classes simultaneously. In order to consider these partial volume effects, a five class segmentation model has been proposed [6]. This model considers the pure classes, CSF, GM, and WM, but also two mixture classes, accounting for the CSF-GM mixture (CSF-GM class), and the GM-WM mixture (GM-WM class).

Common to most unsupervised statistical brain segmentation algorithms is the fact that they explore only voxel intensities as features for the segmentation task. This is due to the fact that Gaussian densities model well the MR intensities, but not necessarily other features. Nevertheless a very promising extension to intensity-based segmentation would explore also other data features. For example for the five class segmentation, the partial volume classes, CSF-GM and GM-WM, are not only characterized by intermediate voxel intensities lying in between the CSF and GM intensities, resp. GM and WM intensities, but also by larger gradients. Therefore it would be very promising to construct a HMM segmentation that explores both features, the voxel intensities and their gradients. Unfortunately the Gaussian HMM will fail for this task, as there is no evidence that the intra-class statistics of the voxels' gradients would fulfill the Gaussian condition. These arguments are true for most features we might want to add to the segmentation algorithm in addition to the voxel intensities. Therefore we developed a segmentation algorithm which does not need any parametric assumptions about the intra-class statistics, resulting in *non-supervised, non-parametric* HMMs (NPHMM). We will introduce the mathematical framework of NPHMMs in the next section. The developments are very general and applicable to any choice of features. In the result section we will enhance the voxel intensity-based segmentation algorithm with the norm of the voxel gradients in order to experimentally validate the theoretical derivations on a particular example.

## 2 Non-supervised, Non-parametric HMMs

Let's consider the MR data,  $S$ , of  $n$  voxels, with associated feature vectors,  $y_i \in \Omega_Y = \mathbb{R}^d$ , where  $d$  is the dimension of the feature vectors.  $Y$  is the continuous random variable associated to the samples,  $y_i$ , over  $\mathbb{R}^d$  and can be estimated by density estimation from the MR data,  $S$ . Our classification task aims to classify the  $n$  voxels into  $n_c$  classes, each class type being labeled by one of the symbols in  $\Omega_c = \{1, \dots, n_c, t\}$  and representing one tissue type. The class label,  $t$ , is associated to the feature subspace,  $\Omega_Y^S \subset \mathbb{R}^d$ , for which no feature vector

of the MR data,  $S$ , exists:  $\Omega_Y^{\bar{S}} = \{y \in \mathbb{R}^d | \#s_i \in S \text{ with feature vector } y\}$ . Let's call  $\mathbf{c} \in \mathcal{C}$  any simultaneous configuration of the class labels,  $c_i$ , where  $\mathcal{C}$  is the set of possible simultaneous configurations.

The random variable over  $\Omega_c$  associated to the class labels is denoted  $C$ . Furthermore we consider a random variable different from  $C$ , called  $C^{est}$ , also over  $\Omega_c$ , which models an estimation of  $C$  from the data  $Y$ . Very naturally we can build the following stochastic process:

$$C \rightarrow Y \rightarrow C^{est} \rightarrow E, \tag{1}$$

where  $E$  is an error random variable being 1 whenever the estimated class label,  $c^{est}$ , is considered a wrong estimate of the initial class label,  $c$ , and 0 otherwise. Using non-parametric Gaussian kernel density estimation, we can get the following probability densities for the transitions of eq. 1:

$$P(y|c) = \sum_{y_1 \in S_c} \frac{1}{|S_c|} G(y - y_1, \sigma_1^2) \tag{2}$$

$$P(c^{est}|y) = \begin{cases} \sum_{y_2 \in S_{c^{est}}} \frac{1}{|y_2|} G(y - y_2, \sigma_2^2) & \text{if } c^{est} \neq t, \\ \int_{\Omega_Y^{\bar{S}}} G(y - y_2, \sigma_2^2) dy_2 & \text{if } c^{est} = t, \end{cases} \tag{3}$$

where  $G(x - a, b)$  denotes a Gaussian kernel with expectation,  $a$ , and variance,  $b$ .  $|y|$  is the number of samples with feature vector  $y$ .  $S_c$  is the subset of  $S$  that contains the samples classified to class  $c$  and  $|S_c|$  is the cardinality of  $S_c$ . Eq. 3 justifies the introduction of the class label  $t$ , which simply ensures that we include the tails of the Gaussian kernels in the probability estimations, i.e. that  $\sum_{c \in \Omega_c} P(c|y) = 1$ . In addition to this, the probability  $P(c)$  is given by  $\frac{|S_c|}{n}$ . We could have chosen to use other transition probabilities than the ones of eqs. 2 and 3. For example we could have imposed that the probability density of the random variable  $C^{est}$  equals the density of  $C$ . In this case,  $P(c^{est}|y)$  would have to be estimated from  $P(y|c)$  using the Bayes theorem.

Let's now use the transition probabilities of eqs. 2 and 3 to estimate a key quantity of eq. 1. It is the probability of error,  $P_{e|c}$ , of the transmission from  $C$  to  $C^{est}$  and equals the expectation of  $E$ ,  $\mu_E$ . This error probability,  $P_{e|c}$ , for a given class label map,  $\mathbf{c}$ , is calculated as follows:

$$\begin{aligned} P_{e|c} = \mu_E &= \sum_{c \in \Omega_c} \int_{\Omega_Y} \sum_{c^{est} \in \Omega_c} P(E = 1 | c^{est}, c) P(c^{est}|y) P(y|c) P(c) dy \\ &= \sum_{c \in \Omega_c \setminus \{t\}} \sum_{c^{est} \in \Omega_c \setminus \{t\}} \frac{P(E = 1 | c^{est}, c)}{n} \sum_{y_1 \in S_c} \sum_{y_2 \in S_{c^{est}}} \frac{G(y_1 - y_2, \sigma_1^2 + \sigma_2^2)}{|y_2|} \\ &+ \sum_{c \in \Omega_c \setminus \{t\}} \frac{P(E = 1 | t, c)}{n} \sum_{y_1 \in S_c} \int_{\Omega_Y^{\bar{S}}} G(y_1 - y_2, \sigma_1^2 + \sigma_2^2) dy_2. \end{aligned} \tag{4}$$

Let's now assume that  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ , and  $P(E = 1|t, c) = 1, \forall c \in \Omega_c \setminus \{t\}$ . Then, by identification of the terms of eq. 4, we can write

$$P_{e|c} = \sum_{c \in \Omega_c \setminus \{t\}} \sum_{c^{est} \in \Omega_c \setminus \{t\}} A_{c,c^{est}} \Gamma_{c,c^{est}} + P(t), \tag{5}$$

$$\text{where } \Gamma_{c,c^{est}} = \frac{1}{n} \sum_{y_1 \in S_c} \sum_{y_2 \in S_{c^{est}}} \frac{1}{|y_2|} G(y_1 - y_2, 2\sigma^2), \forall c, c^{est} \in \Omega_c \setminus \{t\}, \tag{6}$$

is called the transmission matrix and measures the probability that a data sample of  $S$  with initial class label,  $c$ , is transmitted to an output label,  $c^{est}$ ,

$$A_{c,c^{est}} = P(E = 1|c, c^{est}), \forall c, c^{est} \in \Omega_c \setminus \{t\}, \tag{7}$$

is called the distortion matrix which weights the erroneous transmissions through the transmission matrix of eq. 6, and

$$P(t) = \frac{1}{n} \sum_{y_1 \in S} \int_{\Omega_Y^S} G(y_1 - y_2, 2\sigma^2) dy_2 \tag{8}$$

is the probability that any sample gets transmitted into the out-layer,  $\Omega_Y^S$ . It's important to note that  $P(t)$  just depends on the data to be classified but is independent of the specific class labeling,  $\mathbf{c}$ .

One way to apply this framework to MR segmentation would consist of determining the class labels,  $\mathbf{c}$ , that minimize the error probability of eq. 5. This approach would neglect spatial dependencies between the voxels though. But the probabilistic nature of the derived formulae allows us to add a HMM to the error probability minimization, resulting in a NPHMM segmentation:

$$\hat{\mathbf{c}} = \arg \min_{\mathbf{c} \in \mathcal{C}} P(\mathbf{c}) \cdot P_{e|\mathbf{c}}, \tag{9}$$

In complete analogy to parametric HMMs, the prior probabilities,  $P(\mathbf{c})$ , are modeled by a Gibbs distribution [4,5]. The derived non-parametric framework for classification allows us to consider voxel features for which we do not know any parametric model, as it is the case for e.g. voxel gradients. In what follows, we will consider the voxel intensities and the norm of the voxel gradients as a two-dimensional feature space representation of the MR data to perform the segmentation task. We have therefore to check if the five class model (CSF, CSF-GM, GM, GM-WM, and WM) is still adequate to model the anatomy in the resulting feature space. In fact we will show that it is not, and we will derive an anatomically more adapted class model which contains six different anatomical classes.

### 3 A New Anatomical Model for Brain Segmentation

A pathologist who dissects a fixed brain distinguishes darker areas from brighter ones. He denominates them GM and WM respectively. This contrast reflects at a microscopic level the difference in tissue types. Schematically there is GM tissue that is rich in neuronal soma and very poor in myelinated axons, and WM tissue that is exclusively made of axons, that are myelinated or not. One can also speak about a CSF tissue which is roughly said physiologic liquid. Of course in the GM and WM tissues, glial cells and vessels are also to be found but are of no interest in the present discussion. The three above mentioned tissue types have also different MR-relaxation parameters which explains the similar contrast of a T1-weighted MR image and a brain section in pathology. As explained above, most algorithms use either a three or a five class model in order to perform the classification task (CSF, CSF-GM, GM, GM-WM, and WM classes). It happens that pixels falling in the GM and GM-WM classes are made from an anatomical point of view of quite heterogeneous material. This can be summarized for the five class model in the following way:

- The GM class regroups parts of the thalamus and the nucleus caudatus, small parts of the truncus cerebri and most cerebral cortex.
- The GM-WM class regroups the interface between the cerebral cortex and the white matter, the lentiform nucleus, parts of the thalamus, the sub-thalamic region and large parts of the truncus cerebri.

We define from now on the term deep cerebral nuclei (DCN) as an ensemble made of the thalamus, the sub-thalamic region, the lentiform nucleus, the nucleus caudatus and abusively also large parts of the truncus cerebri which is made of a mixture of small nuclei and WM fibers. When reviewing in more detail the morphology of the structures belonging to the GM-WM class, i.e. the cortico-subcortical interface and the DCN, one notices that partial volume at the cortico-sub-cortical interface is mainly due to the overlapping of a pixel over a sharp edge with on the one side GM tissue, i.e. the cortex, and on the other side WM tissue (Fig. 1: Section 3). We could call it partial volume of edge type (GM-WM edge type). The classification of the DCN is usually split over the GM, WM and GM-WM classes. The reason for this heterogeneity can be easily understood when looking at a histological section focused either on the lateral border of the thalamus (Fig. 1: Section 1.a), on the lentiform nucleus (Fig. 1: Section 1.b) or on the pons (Fig. 1: Section 2). In each case there is a micro-metric mixture of myelinated axons and islets of neuronal soma, present in different proportions depending on the location. We propose to speak about partial volume of mosaic type (GM-WM mosaic type). These morphological considerations allow us to naturally introduce a six class model where the GM-WM class is split into two subtypes: edge and mosaic. It should not only enable the distinction of two really different anatomical brain tissues, i.e. the cortico-subcortical interface versus the DCN but also give access to a separate estimation of their respective statistical densities that have a-priori no reason to be identical.

Our approach is also perfectly meaningful from a medical point of view. Depending on the type of disease, atrophy can be predominantly cortical, e.g.

Alzheimer’s disease, age related atrophy, fronto-temporal dementia, or predominantly centered on the DCN, e.g. Parkinson’s and Huntington’s diseases or progressive supra-nuclear palsies. It can also be centered on the WM tissue, e.g. leukodystrophies, some vascular dementia or it can be mixed, involving WM as well as GM tissue, i.e. Lewy body dementia and multi-systemic diseases. It is therefore clear that a six class model has potentially useful clinical advantages.

## 4 Results

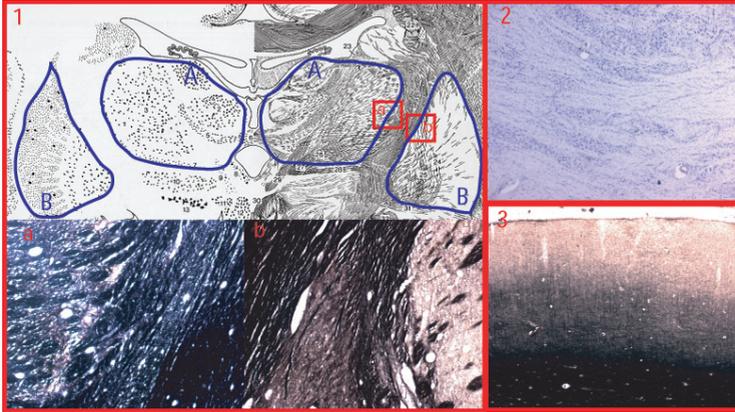
We used the *BrainWeb* MR simulation ([7]) in order to validate the proposed framework for T1-weighted MR brain segmentation. The *BrainWeb* simulation just considers 5 classes. In order to create the distinction between GM-WM edge type and GM-WM mosaic type, we checked for every reference voxel containing both, GM and WM, if it has a 1<sup>st</sup> order neighbor with pure tissue, GM or WM. If yes, this voxel was considered as a voxel of the GM-WM edge type class, while the others of the GM-WM mosaic type class. The resulting six-class reference constituted the ground truth to validate our segmentation results. An axial slice of this reference data set is shown in fig. 2.

Our segmentation algorithms considered voxel intensities and the norms of their gradients as features. We used the independent conditional modes algorithm (ICM) [8] to perform the NPHMM-based segmentation. In table 1, we validate our framework for the MR simulations with noise levels  $n = \{1, 3, 5, 7\}$ . We show the percentages of correct classification and completely misclassified voxels (considering the mixture classes), for both, error probability minimization without hidden Markov random field (Min.  $P_{e|c}$ ) and error probability minimization with hidden Markov random field (NPHMM).

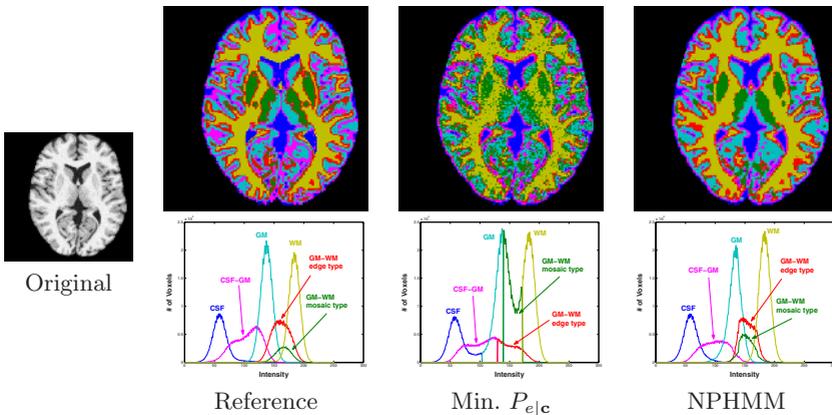
**Table 1.** We show the six class error statistics resulting from non-parametric classification for both, without (Min.  $P_{e|c}$ ) and with the addition of a hidden Markov random field (NPHMM). We applied the algorithm for the *BrainWeb* simulations with noise levels  $n = \{1, 3, 5, 7\}$ . The indicated percentages measure the fraction of correctly classified voxels and the percentage of completely misclassified voxels (considering the mixture classes).

	n=1	n=3	n=5	n=7
Min. $P_{e c}$	71.0% \ 0.2%	64.0% \ 6.1%	59.4% \ 1.4%	54.6% \ 6.9%
NPHMM	71.0% \ 0.2%	71.6% \ 0.7%	71.5% \ 0.3%	64.2% \ 1.4%

We also want to show that the intra-class statistics of the six classes are well approximated by our framework. Therefore we show in fig. 2 the intra-class histograms of the reference dataset and compare them to their analogues resulting from our algorithm. In fig. 2, we also show axial views of our segmentations for the *BrainWeb* dataset with noise level  $n = 5$  and compare them with the reference data. The new class, segmenting the deep structures of the brain (DCN), is shown in green.



**Fig. 1.** Tissue structure in the GM-WM class. 3 types of images: i) Scheme showing the distribution of the neuronal soma on the left and the axonal distribution on the right, ii) Myelin-stain section and iii) Nissl-stain section enhancing respectively myelinated axons and neuronal soma. 1) Demonstration that in the thalamus (A) and in the lentiform nucleus (B) WM and GM tissue, i.e. neuronal soma and axons, mix at the microscopic level in a mosaic like manner. 2) Demonstration that in the pons the cranial nerve nuclei are embedded in a matrix of axons in a mosaic like manner. 3) Section showing the sharp edge between the cortex (GM) and the underlying GM.



**Fig. 2.** On the left, we show an axial slice of the MR simulation with noise level  $n = 5$  (Original). Thereafter we present the same slice of the six-class reference dataset, i.e. the ground truth, and its resulting intra-class histograms (Reference). In the last two columns, we present the non-parametric segmentation results and the resulting intra-class statistics for both, with (NPHMM) and without the addition of a hidden Markov random field (Min.  $P_{e|c}$ ). The new class, segmenting the deep structures of the brain (segmenting the DCN as defined in section 3), is presented in green and distinguishes clearly from the GM-WM edge type class (red), lying predominantly along sharp boundaries between the pure GM class and the pure WM class.

## 5 Discussion and Conclusion

The presented results demonstrate the validity of the mathematical and anatomical framework. Comparing to the six-class reference, the segmentation results of fig. 2 outline nicely the deep structures of the brain (DCN) in green, which, as described in section 3, allows to exploit medically relevant information. The quantitative evaluation of the segmentation results (table 1) further validates the presented framework if we want to explore other features than the voxel intensities of the MR images.

We have presented a new mathematical framework for MR brain segmentation, namely *non-supervised, non-parametric* HMMs. This algorithm doesn't assume that the intra-class statistics belong to any particular parametric family of probability densities. We have applied this novel framework to brain segmentation and were able to consider both, voxel intensities and voxel gradients, for the segmentation task. For this particular case, we have also introduced a new morphological model of the human brain which distinguishes between GM-WM partial volume of mosaic type and GM-WM partial volume of edge type. These anatomical considerations allowed a reliable segmentation of the deep structures of the brain which added valuable information to the segmentation. We have evaluated the algorithm on simulated T1-weighted MR scans.

## References

1. W.M. Wells III, W.E.L. Grimson, R. Kikinis, and F.A. Jolesz, "Adaptive segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 429–442, August 1996.
2. Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, January 2001.
3. David W. Shattuck, Stephanie R. Sandor-Leahy, Kirt A. Schaper, David A. Rottenburg, and Richard M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *Neuro Image*, vol. 13, pp. 856–876, 2001.
4. J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of Royal Statistical Society*, vol. 36, no. 2, pp. 192–236, 1974.
5. S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, June 1984.
6. P. Santago and H.D. Gage, "Quantification of MR brain images by mixture density and partial volume modeling," *IEEE Transactions on Medical Imaging*, vol. 12, no. 3, pp. 566–574, September 1993.
7. R.K.-S. Kwan, A.C. Evans, and G.B. Pike, "MRI simulation-based evaluation of image-processing and classification methods," *IEEE Transactions on Medical Imaging*, vol. 18, no. 11, pp. 1085–1097, November 1999.
8. J. Besag, "On the statistical analysis of dirty pictures," *Journal of Royal Statistical Society*, vol. 48, no. 3, pp. 259–302, 1986.