

# A conditional random field approach for coupling local registration with robust tissue and structure segmentation

B. Scherrer<sup>1,3</sup>, F. Forbes<sup>2,3</sup>, M. Dojat<sup>1,3</sup>

<sup>1</sup> INSERM, U836, Grenoble, F-38042, France

<sup>2</sup> INRIA, MISTIS, Grenoble, France

<sup>3</sup> Université Joseph Fourier, Grenoble, France

**Abstract.** We consider a general modelling strategy to handle in a unified way a number of tasks essential to MR brain scan analysis. Our approach is based on the explicit definition of a Conditional Random Field (CRF) model decomposed into components to be specified according to the targeted tasks. For a specific illustration, we define a CRF model that combines robust-to-noise and to nonuniformity Markovian tissue and structure segmentations with local affine atlas registration. The evaluation performed on both phantoms and real 3T images shows good results and, in particular, points out the gain in introducing registration as a model component. Besides, our modeling and estimation scheme provide general guidelines to deal with complex joint processes for medical image analysis.

## 1 Introduction

The analysis of MR brain scans is a complex task that requires several sources of information to be taken into account and combined. The analysis is frequently based on segmentations of tissues and of subcortical structures performed by human experts. For automatic segmentation, difficulties arise from the presence of various artifacts such as noise or intensity nonuniformities. For structures, the segmentation requires in addition the use of prior information usually encoded via a pre-registered atlas. Recently growing interest has been on tackling this complexity by combining different approaches. As an illustration, Yang *et al.* [1] propose to use a region based tissue classification approach followed by a watershed algorithm to label brain sulci while Yu *et al.* [2] combine a region-based bias field estimation and a level set method to segment the cortex. A step further the combinaison of methods is *coupling*, giving the possibility to introduce mutual interactions between components of a model. Such a coupling can be naturally expressed in a statistical framework via the definition of joint distributions. In this vein, Ashburner and Friston [3] couple a global statistical tissue segmentation approach with the estimation of a bias field and a global registration of an atlas of tissue probability maps. Another growing feature in the literature is to locally estimate model parameters on the image to better fit local image properties. For instance, Scherrer *et al.* [4] couple a *local* tissue

segmentation approach with a structure segmentation approach; Pohl *et al.* [5] couple structure segmentation with the *local affine* registration of an atlas.

In this paper, we propose to go further towards coupling methods by constructing a *Conditional Random Field* (CRF) model that performs a number of essential tasks. We will focus on developing a statistical framework that allows 1) tissue segmentation using *local* Markov Random Field (MRF) models, 2) MRF segmentation of structures and 3) *local affine* registration of an atlas. All tasks are linked and completing each one of them can help in refining the others. The idea is to capture in a single model all the relationships that could be formalized between these tasks. Our basis toward a solution is similar to that in [4] with the major difference that therein a joint model was not explicitly given but defined through the specification of a number of compatible conditional MRF models. In this work, we specify directly a joint model from which the conditional models are derived. As a result, cooperation between tissues and structures is treated in a more symmetric way which results in new more consistent conditional models. In addition, interaction between the segmentation and registration steps is easily introduced. An explicit joint formulation has the advantage to provide a strategy to construct more consistent or complete models that are open to incorporation of new tasks. For estimation, we provide an appropriate variational EM framework allowing a Bayesian treatment of the parameters. The evaluation performed on both phantoms and real 3T brain scans shows good results and demonstrates the clear improvement provided by coupling the registration step to tissue and structure segmentation.

## 2 A CRF approach to segmentation and registration

We consider a finite set  $V$  of  $N$  voxels on a regular 3D grid. Our tissue and structure segmentation task is recast into a missing data framework in which the observed data  $\mathbf{y} = \{y_1, \dots, y_N\}$  are the intensity values observed respectively at each voxel and the missing data  $\mathbf{z} = (\mathbf{t}, \mathbf{s})$  is made of two sets: the tissue classes  $\mathbf{t} = \{t_1, \dots, t_N\}$  and the subcortical structure classes  $\mathbf{s} = \{s_1, \dots, s_N\}$ . The  $t_i$ s take their values in  $\{e_1, e_2, e_3\}$  that represents the three tissues cephalo-spinal-fluid, grey matter and white matter. Each  $e_k$  is a 3-dimensional binary vector whose  $k$ th component is 1, all other components being 0. For the subcortical structure segmentation we consider  $L$  structures, the  $s_i$ s taking their values in  $\{e'_1, \dots, e'_L, e'_{L+1}\}$  where  $e'_{L+1}$  corresponds to an additional background class. Tissues and structures are linked and we denote by  $T^{s_i}$  the tissue of structure  $s_i$  at voxel  $i$ . The model parameters  $\theta = (\psi, \mathcal{R})$  include both the intensity distributions parameters  $\psi$  and the registration parameters  $\mathcal{R}$ . We consider them in a Bayesian framework as realizations of random variables that take their values in  $\Theta = \Psi \times \mathcal{R}$ .

To capture interactions between the various fields  $\mathbf{y}$ ,  $\mathbf{t}$ ,  $\mathbf{s}$  and  $\theta$  we adopt a *conditional random field* approach which consists in specifying a conditional model  $p(\mathbf{t}, \mathbf{s}, \theta | \mathbf{y})$ . We define  $p(\mathbf{t}, \mathbf{s}, \theta | \mathbf{y})$  as a Gibbs measure with energy function  $H(\mathbf{t}, \mathbf{s}, \theta | \mathbf{y})$  *ie.*  $p(\mathbf{t}, \mathbf{s}, \theta | \mathbf{y}) \propto \exp(-H(\mathbf{t}, \mathbf{s}, \theta | \mathbf{y}))$  where the energy is decomposed in the following terms. We denote by  $g(y_i | t_i, s_i, \psi_i)$  positive functions of  $y_i$  and

consider the decomposition:

$$\begin{aligned}
 H(\mathbf{t}, \mathbf{s}, \theta | \mathbf{y}) = & H_T(\mathbf{t}) + H_S(\mathbf{s}) + H_{T,S}(\mathbf{t}, \mathbf{s}) + H_{T,\mathcal{R}}(\mathbf{t}, \mathcal{R}) + H_{S,\mathcal{R}}(\mathbf{s}, \mathcal{R}) \\
 & + H_\Psi(\psi) + H_{\mathcal{R}}(\mathcal{R}) + \sum_{i \in V} \log g(y_i | t_i, s_i, \psi_i) .
 \end{aligned} \tag{1}$$

**A number of essential tasks.** In what follows, we show how the terms in (1) can be specified so that the model performs the tasks listed below.

*Robust-to-noise segmentation.* Robust-to-noise segmentation is generally addressed via MRF modelling. It introduces local spatial dependencies between voxels, providing a labelling regularization. For tissue and structure segmentations, we use standard Potts models setting  $H_T(\mathbf{t}) = \sum_{i \in V} \sum_{j \in \mathcal{N}(i)} \eta_T \langle t_i, t_j \rangle$  and  $H_S(\mathbf{s}) = \sum_{i \in V} \sum_{j \in \mathcal{N}(i)} \eta_S \langle s_i, s_j \rangle$ , where  $\langle \cdot, \cdot \rangle$  denotes the scalar product,  $\mathcal{N}(i)$  represents the voxels neighboring  $i$  and  $\eta_T$  and  $\eta_S$  are additional interaction strength parameters.

*Local approach to deal with nonuniformity.* Generally, tissue intensity models are estimated globally through the entire volume and then suffer from imperfections at a local level. We adopt as in [4] a local segmentation alternative. The principle is to locally compute the tissue models in various subvolumes of the initial volume. These models better reflect local intensity distributions and are likely to handle different sources of intensity nonuniformity. We consider intensity models that depend on the tissue class  $k$  but also on the voxel localization so that  $\psi$  decomposes into  $\psi = \{\psi_i, i \in V\}$  where  $\psi_i = {}^t(\psi_i^k, k = 1, 2, 3)$ . Although possible in our Bayesian framework, this general setting results in too many parameters which could not be estimated accurately. The local approach [4] provides an intermediate efficient solution where the  $\psi_i$ 's are first considered as constant over subvolumes. Let  $\mathcal{C}$  be a regular cubic partitioning of the volume  $V$  in a number of nonoverlapping subvolumes  $\{V_c, c \in \mathcal{C}\}$ . We write  $\psi = \{\psi_c, c \in \mathcal{C}\}$  where  $\psi_c = {}^t(\psi_c^k, k = 1, 2, 3)$  is the common value of all  $\psi_i$  for  $i \in V_c$ . In addition to ensure consistency and spatial regularity between the local estimations of the  $\psi_c$ 's we consider a MRF prior  $p(\psi) \propto \exp(H_\Psi(\psi))$ . The specific form of  $H_\Psi(\psi)$  is the same as in [4]. When Gaussian intensity distributions are considered, it corresponds to assign *auto-normal* Markov priors to the mean parameters. Outside the issue of estimating  $\psi$ , having voxel dependent  $\psi_i$ 's is not a problem. We easily go back to this case, from estimated  $\psi_c$ 's, by using a cubic splines interpolation step.

*Incorporating a priori knowledge via local affine atlas registration.* The *a priori* knowledge required for structure segmentation is classically provided via a global non-rigid atlas registration. Most methods first register the prior information to the medical image and then segment the image based on that aligned information. Although reliable registration methods are available, it is still important, in the subsequent segmentation task, to overcome biases caused by commitment to the initial registration. Also segmentation results provide information that can be used for feedback on registration. Global registration approaches generally lead to a high dimensional minimization problem which is computationally greedy and subject to a high number of local optima. We rather choose a hierar-

chical local affine registration model as in [5]. We consider 1) a global affine transformation given by parameters  $\mathcal{R}^G$ , which describes the non structure-dependent deformations, and 2) one local affine structure-dependent deformation for each structure, defined in relation to  $\mathcal{R}^G$  and capturing the residual structure-specific deformations. It follows  $L+2$  affine transformation parameters  $\mathcal{R} = (\mathcal{R}^G, \mathcal{R}_1^S, \dots, \mathcal{R}_{L+1}^S)$  to be estimated. Interactions between labels and registration parameters are introduced through  $H_{T,\mathcal{R}}(\mathbf{t}, \mathcal{R})$  and  $H_{S,\mathcal{R}}(\mathbf{s}, \mathcal{R})$ . Similarly to [5], the interaction between  $\mathbf{S}$  and  $\mathcal{R}$  is chosen so as to favor configurations for which the segmentation of a structure  $l$  is aligned on its prior atlas. We denote by  $\zeta_S = \{\zeta_S^l, l = 1, \dots, L+1\}$  the statistical atlas of the brain subcortical structures under consideration and by  $\rho(\mathcal{R}^G, \mathcal{R}_l^S, i)$  the interpolation function assigning a position in the atlas space to the image space. We compute the spatial *a priori* distribution  $f_S^l(\mathcal{R}, \cdot)$  of one structure  $l$  by  $f_S^l(\mathcal{R}, i) = \frac{\zeta_S^l(\rho(\mathcal{R}^G, \mathcal{R}_l^S, i))}{\sum_{l'=1..L+1} \zeta_S^{l'}(\rho(\mathcal{R}^G, \mathcal{R}_{l'}^S, i))}$ . The normalization across all structures is necessary as  $\mathcal{R}_l^S$  are structure-dependent parameters and multiple voxels in the atlas space could be mapped to one location in the image space. Although some atlas are potentially available for tissues, in our setting we build  $f_T$ , the spatial *a priori* distribution of the  $K = 3$  tissues, from the  $f_S^l$ 's:  $f_T^k(\mathcal{R}, i) = \sum_{l \text{ st. } T^l = k} f_S^l(\mathcal{R}, i) + \frac{1}{K} f_S^{L+1}(\mathcal{R}, i)$ . Agreement between structure segmentation and atlas is then favored by setting  $H_{S,\mathcal{R}}(\mathbf{s}, \mathcal{R}) = \sum_{i \in V} \langle s_i, \log(f_S(\mathcal{R}, i) + \epsilon) \rangle$ , with the vectorial notation  $f_S = {}^t(f_S^1, \dots, f_S^{L+1})$ . The logarithm and a positive scalar  $\epsilon$  are introduced respectively for homogeneity between probabilities and energies, and to ensure the existence of the logarithm. We choose  $\epsilon = 1$ , making in addition  $H_{S,\mathcal{R}}(\mathbf{s}, \mathcal{R})$  positive, but the overall method does not seem sensitive to its exact value. Similarly, we define the interaction between  $\mathbf{t}$  and  $\mathcal{R}$  by  $H_{T,\mathcal{R}}(\mathbf{t}, \mathcal{R}) = \sum_{i \in V} \langle t_i, \log(f_T(\mathcal{R}, i) + \epsilon) \rangle$ . Then, the term  $H_{\mathcal{R}}(\mathcal{R})$  can be used to introduce *a priori* knowledge to favor estimation of  $\mathcal{R}$  close to some average registration parameters computed from a training data set if available. In our case, no such data set were available and we set  $H_{\mathcal{R}}(\mathcal{R}) = 0$ .

Cooperative tissue and structure segmentations. Tissues and structures are linked: a structure is made of a specific tissue and knowledge on structures locations provides information for tissue segmentation. Inducing cooperation between tissue and structure segmentations can be done through the term  $H_{T,S}(\mathbf{t}, \mathbf{s})$ . We set  $H_{T,S}(\mathbf{t}, \mathbf{s}) = \sum_{i \in V} \langle t_i, e_{T^{s_i}} \rangle$ , so as to favor situations for which the tissue  $T^{s_i}$  of structure  $s_i$  is the same as the tissue given by  $t_i$ . Cooperation between tissue and structure labels also appear via the energy data term  $\sum_{i \in V} g(y_i | t_i, s_i, \psi_i)$ .

Considering Gaussian intensity distributions, we denote by  $\mathcal{G}(\cdot | \mu, \lambda)$  the Gaussian distribution with mean  $\mu$  and precision  $\lambda$  (*i.e.* the inverse of the variance). Denoting  $\psi_i^k = \{\mu_i^k, \lambda_i^k\}$ , we see  $\psi_i$  as a 3-dimensional vector, so that when  $t_i = e_k$ , then  $\mathcal{G}(y_i | \langle t_i, \psi_i \rangle)$  denotes the Gaussian distribution with mean  $\mu_i^k$  and precision  $\lambda_i^k$ . To account for both tissue and structure information, we set:  $g(y_i | t_i, s_i, \psi_i) = \mathcal{G}(y_i | \langle t_i, \psi_i \rangle)^{\frac{1+(s_i, e'_{L+1})}{2}} \mathcal{G}(y_i | \langle e_{T^{s_i}}, \psi_i \rangle)^{\frac{1-(s_i, e'_{L+1})}{2}}$ . When tissue and structure segmentations contain the same information at voxel  $i$ , *ie.* either  $t_i = e_{T^{s_i}}$  or  $s_i = e'_{L+1}$ , then the expression of  $g$  above reduces to the usual

$\mathcal{G}(y_i|t_i, \psi_i)$ . When this is not the case, the expression of  $g$  above leads to  $\mathcal{G}(y_i|t_i, \psi_i)^{1/2}\mathcal{G}(y_i|e_{T^{s_i}}, \psi_i)^{1/2}$  which is a more appropriate compromise.

### 3 A Bayesian EM estimation framework

We consider the EM algorithm and more specifically its *Maximization-Maximization* interpretation as a general estimation technique in the presence of missing data. Let  $\mathcal{T}$  and  $\mathcal{S}$  be respectively the spaces in which  $\mathbf{t}$  and  $\mathbf{s}$  take their values. We denote by  $\mathcal{D}$  the set of probability distributions on  $\mathcal{Z} = \mathcal{T} \times \mathcal{S}$ . In a Bayesian framework, EM can be used to find Maximum A Posteriori (MAP) estimations (see eg [6]) and leads to the alternating maximization over  $q \in \mathcal{D}$  and  $\theta \in \Theta$  of the function defined by  $F_{\text{MAP}}(q, \theta) = \sum_{\mathbf{z} \in \mathcal{Z}} \log p(\mathbf{y}, \mathbf{z} | \theta) q(\mathbf{z}) + \log p(\theta) + I[q]$ , where  $I[q] = -\mathbb{E}_q[\log q(\mathbf{Z})]$  is the entropy of  $q$  ( $\mathbb{E}_q$  denotes the expectation with regard to  $q$  and capital letters indicate random variables while small letters denote their realizations). However, the dependencies between the missing data usually make the optimization over  $\mathcal{D}$  intractable. We then propose to use a Variational EM approach [7] in which the E-step is not performed exactly. The optimization is solved over a restricted class of probability distributions which factorize as  $q(\mathbf{t}, \mathbf{s}) = q_T(\mathbf{t}) q_S(\mathbf{s})$  where  $q_T$  (resp.  $q_S$ ) belongs to the set  $\mathcal{D}_T$  (resp.  $\mathcal{D}_S$ ) of probability distributions on  $\mathcal{T}$  (resp. on  $\mathcal{S}$ ). Further generalizing by dividing the approximate E-step into two stages, it follows a variant that falls in the modified Generalized Alternating Minimization (GAM) procedures family [8]. From the definition of  $F_{\text{MAP}}$ , we therefore derive a 3-steps algorithm. At iteration  $r + 1$ , with current estimates denoted by  $q_T^{(r)}$ ,  $q_S^{(r)}$  and  $\theta^{(r)}$ , it consists of:

$$\mathbf{E-T-step:} \quad q_T^{(r+1)} = \arg \max_{q_T \in \mathcal{D}_T} \mathbb{E}_{q_T} [\mathbb{E}_{q_S^{(r)}} [\log p(\mathbf{T}|\mathbf{S}, \mathbf{y}, \theta^{(r)})]] + I[q_T] \quad (2)$$

$$\mathbf{E-S-step:} \quad q_S^{(r+1)} = \arg \max_{q_S \in \mathcal{D}_S} \mathbb{E}_{q_S} [\mathbb{E}_{q_T^{(r+1)}} [\log p(\mathbf{S}|\mathbf{T}, \mathbf{y}, \theta^{(r)})]] + I[q_S] \quad (3)$$

$$\mathbf{M-step:} \quad \theta^{(r+1)} = \arg \max_{\theta \in \Theta} \mathbb{E}_{q_T^{(r+1)} q_S^{(r+1)}} [\log p(\theta|\mathbf{T}, \mathbf{S}, \mathbf{y})] . \quad (4)$$

Equations (2-4) show that for inference the specification of the three conditional distributions  $p(\mathbf{t}|\mathbf{s}, \mathbf{y}, \theta)$ ,  $p(\mathbf{s}|\mathbf{t}, \mathbf{y}, \theta)$  and  $p(\theta|\mathbf{t}, \mathbf{s}, \mathbf{y})$  is sufficient. These models can be easily deduced from the conditional distribution  $p(\mathbf{t}, \mathbf{s}, \theta|y)$  confirming that there is no need to define the complete joint model  $p(\mathbf{t}, \mathbf{s}, \mathbf{y}, \theta)$  and emphasizing the rational of using a CRF approach for segmentation purpose. Moreover, the model definition in (1) induces that the conditional models  $p(\mathbf{t}|\mathbf{s}, \mathbf{y}, \theta)$  and  $p(\mathbf{s}|\mathbf{t}, \mathbf{y}, \theta)$  are MRF with energy functions denoted by  $H(\mathbf{t}|\mathbf{s}, \mathbf{y}, \theta)$  and  $H(\mathbf{s}|\mathbf{t}, \mathbf{y}, \theta)$  obtained by omitting in expression (1) the terms that do not depend on  $\mathbf{t}$ , resp. on  $\mathbf{s}$ . The two-stage E-step (2) and (3) requires then to compute  $H_T^{(r+1)}(\mathbf{t}) = \mathbb{E}_{q_S^{(r)}} [H(\mathbf{t}|\mathbf{S}, \mathbf{y}, \theta^{(r)})]$  and  $H_S^{(r+1)}(\mathbf{s}) = \mathbb{E}_{q_T^{(r+1)}} [H(\mathbf{s}|\mathbf{T}, \mathbf{y}, \theta^{(r)})]$ . Neglecting terms not depending on  $\mathbf{t}$ , it comes:

$$H_T^{(r+1)}(\mathbf{t}) = \sum_{i \in \mathcal{V}} \left[ \langle t_i, \log(\tilde{f}_T^{(r)}(\mathcal{R}^{(r)}, i)) \rangle + \sum_{j \in \mathcal{N}(i)} \eta_T \langle t_i, t_j \rangle + \log(g_{T_i}(y_i|t_i)) \right], \quad (5)$$

where  $g_{T_i}(y_i|t_i) = \mathcal{G}(y_i|t_i, \psi_i^{(r)}) \frac{1+q_{S_i}^{(r)}(e'_{L+1})}{2}$  and  $\tilde{f}_T^{(r)} = t(\tilde{f}_T^{k(r)}, k = 1, 2, 3)$  with  $\tilde{f}_T^{k(r)}$  defined by  $\log(\tilde{f}_T^{k(r)}(\mathcal{R}, i)) = \log(f_T^k(\mathcal{R}, i) + \epsilon) + \sum_{l \text{ st. } T^l=k} q_{S_i}^{(r)}(e'_l)$ . In the latter expression, the term  $\sum_{l \text{ st. } T^l=k} q_{S_i}^{(r)}(e'_l)$  is the probability, given the current

distribution  $q_{S_i}^{(r)}$ , that voxel  $i$  belongs to a structure whose tissue is  $k$ . Intuitively, the higher this probability the more favored is tissue  $k$ . Similarly,

$$H_S^{(r+1)}(\mathbf{s}) = \sum_{i \in V} \left[ \langle s_i, \log(\tilde{f}_S^{(r)}(\mathcal{R}^{(r)}, i)) \rangle + \sum_{j \in \mathcal{N}(i)} \eta_S \langle s_i, s_j \rangle + \log(g_{S_i}(y_i | s_i)) \right], \quad (6)$$

where  $g_{S_i}(y_i | s_i) = (\prod_{k=1}^3 \mathcal{G}(y_i | \psi_i^{k(r)})^{q_{T_i}^{(r+1)}(e_k)})^{\frac{1 + \langle s_i, e'_{L+1} \rangle}{2}} \mathcal{G}(y_i | \langle e_{T^{s_i}}, \psi_i^{(r)} \rangle)^{\frac{1 - \langle s_i, e'_{L+1} \rangle}{2}}$  and  $\log \tilde{f}_S^{(r)}(\mathcal{R}, i) = \log(f_S^l(\mathcal{R}, i) + \epsilon) + q_{T_i}^{(r+1)}(e_{T^l})(1 - \langle e'_l, e'_{L+1} \rangle)$  where the term  $q_{T_i}^{(r+1)}(e_{T^l})(1 - \langle e'_l, e'_{L+1} \rangle)$  favors a structure whose tissue is  $T^l$  if  $l$  is a proper structure. We recognize in (5) and (6) the standard decomposition of a MRF model into three terms: an external field, a regularizing spatial term and a data term. Then, solving the current E-T and E-S steps is equivalent to solve the segmentation task for standard MRFs whose definition depends on the previous iteration. In this work we consider a Mean field like algorithm to actually compute  $q_T^{(r+1)}$  and  $q_S^{(r+1)}$  but any other MRF estimation strategy could be possible.

The independence of  $\psi$  and  $\mathcal{R}$  then leads to a two-stage M-step **M- $\psi$**  and **M- $\mathcal{R}$** . For the **M- $\psi$**  step, the choice of a Markovian prior energy  $H_\Psi(\psi)$  as in [4] requires the use of a Mean Field like approximation for the maximization. Similarly to [4], we update the  $\psi_c^k = \{\mu_c^k, \lambda_c^k\}$ 's with the values obtained at convergence of the following scheme ( $(\nu)$  denotes the iteration number):

$$\begin{aligned} \mathbf{M-}\psi : \quad \mu_c^{k(\nu+1)} &= \frac{\lambda_c^{k(\nu)} \sum_{i \in V_c} a_{ik} y_i + \lambda_c^{0k} |\mathcal{N}(c)|^{-1} \sum_{c' \in \mathcal{N}(c)} \mu_{c'}^{k(\nu)}}{\lambda_c^{k(\nu)} \sum_{i \in V_c} a_{ik} + \lambda_c^{0k}} \\ \lambda_c^{k(\nu+1)} &= \frac{\alpha_c^k + \sum_{i \in V_c} a_{ik} / 2 - 1}{b_c^k + 1 / 2 [\sum_{i \in V_c} a_{ik} (y_i - \mu_c^{k(\nu+1)})^2]}, \end{aligned}$$

where  $\{\lambda_c^{0k}, \alpha_c^k, b_c^k, c \in \mathcal{C}\}$  are hyperparameters to be specified,  $\mathcal{N}(c)$  denotes the indices of the subvolumes neighboring subvolume  $c$ ,  $|\mathcal{N}(c)|$  the number of them and  $a_{ik} = \frac{1}{2}(q_{T_i}(e_k) + q_{T_i}(e_k)q_{S_i}(e'_{L+1}) + \sum_{l \text{ st. } T^l = k} q_{S_i}(e'_l))$ . The first term in  $a_{ik}$  is the probability for voxel  $i$  to belong to tissue  $k$  without any structure knowledge. The sum over  $k$  of the two other terms is one and they can be interpreted as the probability for  $i$  to belong to the tissue class  $k$  when information on structure segmentation is available. Parameter values per voxel are then computed by cubic splines interpolation between  $\psi_c$  and  $\psi_{c'}$  for all  $c' \in \mathcal{N}(c)$  so that smooth variations between neighboring subvolumes are ensured and the intensity nonuniformity is handled inside each subvolume. For the **M- $\mathcal{R}$**  step, we get from (4),  $\mathcal{R}^{(r+1)} = \arg \max_{\mathcal{R} \in \mathcal{R}} (H(\mathcal{R}) + \mathbb{E}_{q_T^{(r+1)}}[H_{T, \mathcal{R}}(\mathbf{T}, \mathcal{R})] + \mathbb{E}_{q_S^{(r+1)}}[H_{S, \mathcal{R}}(\mathbf{S}, \mathcal{R})])$ . In practice, the global parameters  $\mathcal{R}^G$  are determined in a pre-processing step using some standard intensity based method such as FLIRT<sup>4</sup>. For the other transformations, we adopt a relaxation approach and update the 12 parameters defining each local affine transformation  $\mathcal{R}_i^S$  by maximizing in turn:

$$\begin{aligned} \mathbf{M-}\mathcal{R} : \quad \mathcal{R}_i^{S(r+1)} &= \arg \max_{\mathcal{R}_i^S} \left( H(\mathcal{R}) + \sum_{i \in V} \sum_{k=1}^3 q_{T_i}^{(r+1)}(e_k) \log(f_T^k(\mathcal{R}, i) + \epsilon) \right. \\ &\quad \left. + \sum_{i \in V} \sum_{l=1}^{L+1} q_{S_i}^{(r+1)}(e'_l) \log(f_S^l(\mathcal{R}, i) + \epsilon) \right). \end{aligned}$$

There exists no simple expression and the optimization is performed numerically using a variant of the Powell algorithm.

<sup>4</sup> <http://www.fmrib.ox.ac.uk/fsl/flirt/>

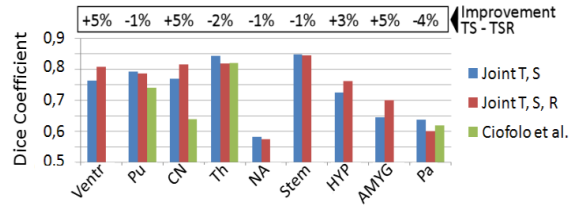


Fig. 1. Evaluation on IBSR v2 (9 right structures) and comparison with [9].

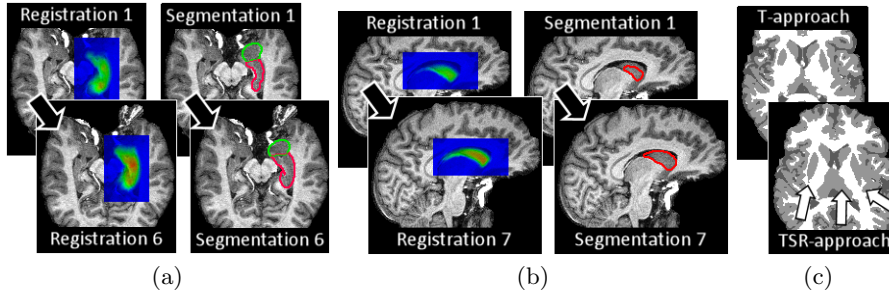


Fig. 2. (a) Evolution of hippocampus local affine registration and segmentation; (b) Evolution of the caudate atlas registration and segmentation after an artificial perturbation of the initial registration; (c) Tissue segmentations with T and TSR approaches.

## 4 Evaluation

We chose to set parameters  $\eta_T$  and  $\eta_S$  to the inverse of a decreasing temperature as generally done. The precision parameters  $\lambda_c^{0k}$  were set to  $N_c \lambda_g^k$  where  $N_c$  is the number of voxels in  $c$  and  $\lambda_g^k$  is a rough precision estimation for class  $k$  obtained by a standard global EM algorithm. The  $\alpha_c^k$ 's were set to  $|\mathcal{N}(c)|$ ,  $b_c^k$  to  $|\mathcal{N}(c)|/\lambda_g^k$  and the tissue subvolumes size to  $20^3$  voxels. The atlas used was the Harvard-Oxford subcortical probabilistic atlas. For a fair comparison, we first carried out tissue segmentation only. The results were equivalent to that in [4] and quantitatively comparable to the results from FAST and SPM5 for lower computational times. They showed robustness both to noise and to nonuniformities. We then rather focused on evaluating the coupling performance. We computed via STAPLE a 3-structure BrainWeb gold standard from three manual expert segmentations of the left caudate, left putamen and left thalamus. When combining tissue and structure segmentations (TS approach), the mean Dice metric over 8 experiments (phantoms with 3%, 5%, 7%, 9% of noise, and 20% or 40% of nonuniformity) was respectively 74%, 90% and 90% for the three structures (computational time: 25 min on a Pentium 2Ghz, 2Go RAM). When adding registration in the combination (TSR approach), it reached respectively 91%, 95% and 94% (computational time: 50 min), showing great improvement for the caudate whose atlas was initially badly registered. Comparatively, [4] reported respectively 74%, 85% and 91%. We then considered 18 images from the IBSR v2 database. The mean Dice metric for the 9 right structures (17 were segmented) is reported in Fig. 1). Most structure segmentations were improved

by the introduction of registration in the coupling (mean improvement: +4.5% ; mean degradation: -1.2%). Then Fig. 2 shows the results for a real 3T brain scan. Fig. 2(a) illustrates how registration and structure segmentation improve with iterations while in Fig. 2(b), the initial caudate registration was perturbed artificially to point out the ability of our approach to correct the mis-alignment and recover a correct segmentation. Eventually, Fig. 2(c) shows that the final tissue segmentation is much better with the TSR approach.

## 5 Discussion

Our approach provides general guidelines to deal with complex joint processes. It is based on the initial specification of a joint probabilistic model decomposed into parts to account for various type of interactions. We used this strategy to integrate an atlas registration with a tissue and structure segmentation process. We proposed a model that captures several level of interactions 1) spatial dependencies between voxels for robustness to noise, 2) spatial dependencies between local intensity models to ensure their consistency, 3) relationships between tissue and structure labels and 4) relationships between labels and local affine atlas registration parameters. In addition to the inclusion of registration, we built on the approach in [4] by introducing new tissue and structure interaction terms. As a result of the joint approach, these new terms correspond to a more symmetric cooperation between tissues and structures. Besides, we obtained very good results that confirmed the benefits of allowing symmetric interactions and including registration as part of the model components rather than as a separate step. Further refinements include the introduction of an *a priori*  $H(\mathcal{R})$  for the registration and the addition of a sulci lines segmentation process. We believe the use of training data as in [5] will facilitate registration parameters estimation and further improve the results. Also, interactions between sulci lines and tissue segmentation could reduce the over-regularization effect of MRF around sulci.

## References

1. Yang, F., Kruggel, F.: Automatic segmentation of human brain sulci. *Medical Image Analysis* **12**(4) (2008) 442–451
2. Yu, Z.Q., Zhu, Y., Yang, J., Zhu, Y.M.: A hybrid region-boundary model for cerebral cortical segmentation in MRI. *Comp. Med. Imag. and Graph.* **30**(3) (2006) 197–208
3. Ashburner, J., Friston, K.J.: Unified Segmentation. *NeuroImage* **26** (2005) 839–851
4. Scherrer, B., Forbes, F., Garbay, C., Dojat, M.: Fully Bayesian Joint Model for MR Brain Scan Tissue and Structure Segmentation. In: *MICCAI*. (2008) 1066–1074
5. Pohl, K.M., Fisher, J., Grimson, W.E.L., Kikinis, R., Wells, W.M.: A Bayesian model for joint segmentation and registration. *NeuroImage* **31**(1) (2006) 228–239
6. Gelman, A., et al: *Bayesian Data Analysis*. Chapman & Hall, 2nd edition (2004)
7. Jordan, M., Ghahramani, Z., Jaakkola, T., Saul, L.: An introduction to variational methods for graphical models. In: *Learning in Graphical Models*. (1998) 105–162
8. Byrne, W., Gunawardana, A.: Convergence theorems of Generalized Alternating Minimization Procedures. *J. Machine Learning Research* **6** (2005) 2049–2073
9. Ciofolo, C., Barillot, C.: Shape analysis and fuzzy control for 3D competitive segmentation of brain structures with level sets. In: *ECCV*. (2006) 458–470