Joint Segmentation of Image Ensembles via Latent Atlases

T. Riklin Raviv[1], K. Van-Leemput[1,2,3], W. M Wells III[1,4] and P. Golland[1]

[1] Computer Science and Artificial Intelligence Laboratory, MIT, USA
[2] Helsinki University of Technology, Finland
[3] Department of Neurology, MGH, Harvard Medical School, USA
[4] Brigham and Womens Hospital, Harvard Medical School, USA

Abstract. Spatial priors, such as probabilistic atlases, play an important role in MRI segmentation. However, the availability of comprehensive, reliable and suitable manual segmentations for atlas construction is limited. We therefore propose a joint segmentation of corresponding, aligned structures in the entire population that does not require a probability atlas. Instead, a latent atlas, initialized by a single manual segmentation, is inferred from the evolving segmentations of the ensemble. The proposed method is based on probabilistic principles but is solved using partial differential equations (PDEs) and energy minimization criteria. We evaluate the method by segmenting 50 brain MR volumes. Segmentation accuracy for cortical and subcortical structures approaches the quality of state-of-the-art atlas-based segmentation results, suggesting that the *latent atlas* method is a reasonable alternative when existing atlases are not compatible with the data to be processed.

1 Introduction

Probabilistic atlases are crucial for most MR segmentation methods due to the absence of well defined boundaries. Derived from comprehensive sets of manually labeled examples, atlases provide statistical priors for tissue classification and structure segmentation [1, 8, 13, 14, 18]. Although atlas-based segmentation methods often achieve accurate results, the need for spatial priors can be problematic. First, the availability of suitable atlases is limited since manual segmentation of a significant number of volumes requires expensive effort of an experienced physician. Second, the suitability of existing atlases for images from different populations is questionable. Examples include using normal adult brain atlas for brain parcellation of young children or patients with severe brain pathologies.

Recently, a few methods have been proposed to reduce or avoid the dependency on possibly incompatible atlases. In the atlas-based segmentation method in [3], topological constraints are used to avoid possible bias introduced by the atlas. In [19], manually labeled structures are used to support the automatic segmentation of neighboring structures within the same image. Tu *et al.* [17] propose a discriminative approach for the segmentation of adjacent brain structures using a set of discriminative features learned from training examples. Lord *et al.* [11] suggest a group-wise smoothing, segmentation and registration method for cross sectional MR scans. Bhatia *et al.* [5] update an initial atlas constructed from a different population using the evolving segmentations of multiple images.

Here we propose and demonstrate a method that does not use a set of training images or probabilistic atlases as priors. Instead we extract an ensemble of corresponding structures simultaneously. The evolving segmentation of the entire image set supports each of the individual segmentations. In practice, a subset of the model parameters, called the spatial parameters, is inferred as part of the joint segmentation processes. These latent spatial parameters, which can be viewed as a 'dynamic atlas', are estimated exclusively from the data at hand and a single manual segmentation. The latent atlas is used as a Markov Random Field (MRF) prior on the tissue labels. The main novelty of the method with respect to other group-wise segmentation methods such as [11, 5] is the consistent statistically-driven variational framework for MR ensemble segmentation.

Our contribution is two-fold. We introduce a level set framework, that is based on probabilistic principles, in which segmentation uncertainty is expressed by the logistic function of the associated level set values, similar to [15]. We then use it for group-wise segmentation. We evaluate our method by segmenting the amygdala, temporal gyrus and hippocampus in each hemisphere in 50 MR brain scans. The dice scores achieved by our method approach the atlas-based segmentation results of [14].

2   Problem definition and probabilistic model

Our objective is to segment a particular structure or region of interest in N aligned MR images. Specifically, we consider the 2-partition problem where each voxel in image I_n ($n = 1 \ldots N$) is assigned to either the foreground (structure of interest) or the background.

Let each image $I_n \colon \Omega \to \mathbb{R}^+$, be a gray level image with V voxels, defined on $\Omega \subset \mathbb{R}^3$ and $\Gamma_n \colon \Omega \to \{0, 1\}$ be the unknown segmentation of the image I_n. We assume that each Γ_n is generated iid from a probability distribution $p(\Gamma \mid \theta_\Gamma)$ where θ_Γ is a set of unknown parameters. We also assume that Γ_n generates the observed image I_n, independently of all other image-segmentation pairs, with probability $p(I_n | \Gamma_n, \theta_{I,n})$ where $\theta_{I,n}$ are the parameters corresponding to image I_n. We assign a specific set of intensity parameters to each image since the acquisition conditions might vary across subjects.

Let $\{I_1 \ldots I_N\}$ be the given set of aligned images that form the observed variable in our problem and let $\Gamma = \{\Gamma_1 \ldots \Gamma_N\}$ be the corresponding segmentations. The joint distribution $p(I_1 \ldots I_N, \Gamma_1 \ldots \Gamma_N | \Theta)$ is governed by the composite set of parameters $\Theta = \{\theta_\Gamma, \theta_{I,1} \ldots \theta_{I,N}\}$. Our goal is to estimate the segmentations Γ.

We jointly optimize for the segmentations Γ and the parameters Θ, assuming that $I_1 \ldots I_N$ are independent:

$$\{\hat{\Theta}, \hat{\Gamma}\} = \arg \max_{\{\Theta, \Gamma\}} \log p(I_1 \ldots I_N, \Gamma_1 \ldots \Gamma_N; \Theta) \tag{1}$$

$$= \arg \max_{\{\Theta, \Gamma\}} \sum_{n=1}^{N} \log p(I_n, \Gamma_n; \Theta) \tag{2}$$

$$= \arg \max_{\{\Theta, \Gamma\}} \sum_{n=1}^{N} \left[\log p(I_n | \ \Gamma_n; \theta_{I,n}) + \log p(\Gamma_n; \theta_\Gamma) \right]. \tag{3}$$

We propose to alternate between estimating the maximum a posteriori (MAP) segmentations and updating the model parameters. For a given value of the model parameters $\hat{\Theta}$, Equation (3) implies that the segmentations can be estimated by solving N separate MAP problems:

$$\hat{\Gamma}_n = \arg \max_{\Gamma_n} \left[\log p(I_n | \ \Gamma_n; \theta_{I,n}) + \log p(\Gamma_n; \ \theta_\Gamma) \right]. \tag{4}$$

We then fix $\hat{\Gamma}$ and estimate the model parameters $\Theta = \{\theta_\Gamma, \theta_{I,1}, \dots \theta_{I,N}\}$ by solving two ML problems:

$$\hat{\theta}_{I,n} = \arg \max_{\theta_{I,n}} \log p(I_n | \ \Gamma_n, \theta_{I,n}) \tag{5}$$

$$\hat{\theta}_\Gamma = \arg \max_{\theta_\Gamma} \sum_{n=1}^{N} \log p(\Gamma_n | \ \theta_\Gamma). \tag{6}$$

In the following sections we present a level-set framework that is motivated by this probabilistic model. We reformulate the estimation problem stated in Eq. (4) such that the soft segmentations $p(\Gamma_n)$ rather then the Γ_n are estimated.

3   Probabilistic view of the level set framework

Now we draw the connection between the probabilistic model presented above and the level set framework for segmentation. Let $\phi_n \colon \Omega \to \mathbb{R}$ denote a level set function associated with image I_n. The zero level $C_n = \{\mathbf{x} \in \Omega | \ \phi_n(\mathbf{x}) = 0\}$ defines the interface between the partitions of I_n. Vese and Chan [6] represent the image partitions by a regularized variant of the Heaviside function of ϕ_n, e.g., $\tilde{H}_\epsilon(\phi_n) = \frac{1}{2}(1 + \frac{2}{\pi} \arctan(\frac{\phi_n}{\epsilon}))$. Alternatively, we can use the hyperbolic tangent to achieve the same goal:

$$\tilde{H}_\epsilon(\phi_n) = \frac{1}{2} \left(1 + \tanh \left(\frac{\phi_n}{2\epsilon} \right) \right) = \frac{1}{1 + e^{-\phi_n/\epsilon}}. \tag{7}$$

For $\epsilon = 1$, the function $\tilde{H}_\epsilon(\cdot)$ is the logistic function. Similar to [15], we define the level set function ϕ_n using the log-odds formulation instead of the conventional signed distance function:

$$\phi_n(\mathbf{x}) \triangleq \epsilon \ \mathrm{logit}(p) = \epsilon \ \log \frac{p(\mathbf{x} \in w)}{1 - p(\mathbf{x} \in \omega)} = \epsilon \ \log \frac{p(\mathbf{x} \in \omega)}{p(\mathbf{x} \in \Omega \setminus \omega)}. \tag{8}$$

The scalar ϵ determines the scaling of the level set function ϕ_n with respect to the ratio of the probabilities. Substituting this definition into Eq. (7) we obtain

$$\tilde{H}_\epsilon(\phi_n(\mathbf{x})) = \frac{1}{1 + p(\mathbf{x} \in \Omega \setminus \omega)/p(\mathbf{x} \in \omega)} = p(\mathbf{x} \in \omega), \tag{9}$$

which implies that the function $\tilde{H}_\epsilon(\phi_n(\mathbf{x}))$ can be viewed as the probability that the voxel in location $\mathbf{x}$ belongs to the foreground region. The functions $\tilde{H}_\epsilon(\phi_n(\mathbf{x}))$ and $H(\phi_n(\mathbf{x}))$ therefore represent soft and hard segmentations, respectively. To simplify the notation we omit the subscript ϵ in the rest of the paper.

In the following subsections we relate the terms in Eq. (3) to the energy terms in the classical level set functional.

3.1   Image likelihood term

Let us first consider the image likelihood term in Eq. (3):

$$\log p(I_n|\ \Gamma_n, \hat{\theta}_{I,n}) = \sum_{\{v|\Gamma_n^v=1\}} \log p_{\text{in}}(I_n^v; \theta_{I,n}) + \sum_{\{v|\Gamma_n^v=0\}} \log p_{\text{out}}(I_n^v; \theta_{I,n}), \quad (10)$$

where p_{in} and p_{out} are the probability distributions of the foreground and background image intensities, respectively.

Let $E_I \cong -\log p(I_n|\ \Gamma_n, \hat{\theta}_{I,n})$ define the energy term associated with the image likelihood term. Using the level-set formulation and replacing the binary labels Γ_n in Eq. (10) with a soft segmentation represented by $\tilde{H}(\phi_n)$, we get:

$$E_I(\phi_n, \Theta) = -\int_\Omega \left[\log p_{\text{in}}(I_n; \theta_{I,n})\tilde{H}(\phi_n(\mathbf{x})) + \log p_{\text{out}}(I_n; \theta_{I,n})\tilde{H}(-\phi_n(\mathbf{x}))\right] d\mathbf{x}. \quad (11)$$

If we use, for example, Gaussian densities for p_{in} and p_{out} we get the familiar minimal variance term [6, 12]. Here, we use a Gaussian mixture to model the background, as described later in the paper.

3.2   Spatial prior term

We define the prior probability $p(\Gamma_n|\ \theta_\Gamma)$ to be a Markov Random Field (MRF):

$$p(\Gamma_n|\theta_\Gamma) = \frac{1}{Z(\theta_\Gamma)} \prod_{v=1}^V (\theta_\Gamma^v)^{\Gamma_n^v} (1 - \theta_\Gamma^v)^{(1-\Gamma_n^v)} e^{-f(\Gamma_n^v, \Gamma_n^{\mathcal{N}(v)})}, \quad (12)$$

where $Z(\theta_\Gamma)$ is the partition function and $\mathcal{N}(v)$ are the closest neighbors of voxel v. The function $f(\cdot)$ accounts for the interactions between neighboring voxels. If we omit the pairwise term in Eq. (12), the prior on segmentations $p(\Gamma_n|\theta_\Gamma)$ reduces to a Bernoulli distribution, where the parameters θ_Γ represent the probability map for the structure of interest. The introduction of the pairwise clique potentials complicates the model but encourages smoother labeling configurations.

Using the continuous level set formulation with soft segmentation, we define the spatial energy term. as follows:

$$E_S(\phi_n, \Theta) = -\int_\Omega \left[\log \theta_\Gamma(\mathbf{x})\tilde{H}(\phi_n(\mathbf{x})) + \log(1 - \theta_\Gamma(\mathbf{x}))\tilde{H}(-\phi_n(\mathbf{x}))\right] d\mathbf{x} \quad (13)$$

As in [4] we ignore the partition function, approximating the MRF model above. The logarithm of the pairwise clique potential term $f(\cdot)$ can be configured to act

as a finite difference operator approximating the gradient of Γ_n at the voxel v [10]. It can be therefore viewed as an approximation of the continuous term

$$E_{\text{LEN}}(\phi_n) = \int_{\Omega} |\nabla \tilde{H}(\phi_n(\mathbf{x}))| d\mathbf{x}, \tag{14}$$

which is the commonly used smoothness constraint as reformulated in [6].

3.3 The unified energy functional

We now construct the cost functional for $\phi_1 \ldots \phi_N$ and the parameters Θ by combing Eqs. (11),(13) and (14):

$$E(\phi_1 \ldots \phi_N, \Theta) = \gamma E_{\text{LEN}} + \beta E_I + \alpha E_S, \tag{15}$$

where $\alpha = 1 - \beta - \gamma$. As in [16] we tune the weights such that the contributions of the energy terms E_{LEN}, E_I and E_S to the overall cost are balanced.

4 Alternating minimization algorithm

We optimize the unified functional (15) in a set of alternating steps. For fixed model parameters Θ, the evolution of each of the level set functions ϕ_n follows the gradient descent equation:

$$\frac{\partial \phi_n}{\partial t} = \tilde{\delta}(\phi_n) \left\{ \gamma \text{ div } \left(\frac{\nabla \phi_n}{|\nabla \phi_n|} \right) + \beta \left[\log p_{\text{in}}(I_n(\mathbf{x}); \hat{\theta}_{I,n}) - \log p_{\text{out}}(I_n(\mathbf{x}); \hat{\theta}_{I,n}) \right] \right.$$
$$\left. + \alpha \left[\log \hat{\theta}_{\Gamma} - \log(1 - \hat{\theta}_{\Gamma}) \right] \right\}, \tag{16}$$

where $\tilde{\delta}(\phi_n) \triangleq \tilde{\delta}_\epsilon(\phi_n) = \frac{d\tilde{H}_\epsilon(\phi_n)}{d\phi_n}$ is the derivative of the Heaviside function, i.e.,

$$\tilde{\delta}_\epsilon(\phi_n) = \frac{1}{2\epsilon} \text{sech}(\frac{\phi_n}{2\epsilon}) = \frac{1}{\epsilon \cosh(\frac{\phi_n}{\epsilon})}.$$

For fixed segmentations ϕ_n the model parameters are recovered by differentiating the cost functional (15) with respect to each parameter.

4.1 Intensity parameters

We assume that the intensities of the structure of interest are drawn from a normal distribution, i.e., $p_{\text{in}}(I_n; \theta_{I,n}) = \mathcal{N}(I_n; \mu_n, \sigma_n^2)$. The distribution mean and variance of the foreground region of I_n are updated at every iteration. The intensities of the background tissues are modeled as a K-Gaussian mixture:

$$p_{\text{out}}(I_n; \theta_{I,n}) = \text{GMM}(\mu_n^1 \cdots \mu_n^K, \sigma_n^1 \cdots \sigma_n^K, w_n^1 \cdots w_n^K),$$

where w_n^k is the mixing proportion component k in the mixture. The Gaussian mixture model parameters are estimated using expectation maximization (EM) [7].

4.2 Spatial parameters

We estimate the spatial function $\theta_\Gamma(\mathbf{x})$, constructing a dynamically evolving latent atlas, by optimizing the sum of the energy terms the depend on θ_Γ :

$$\hat{\theta}_\Gamma = \arg\max_{\theta_\Gamma} \sum_{n=1}^N \int_\Omega [\tilde{H}(\phi_n(\mathbf{x})) \log(\theta_\Gamma(\mathbf{x})) + (1 - \tilde{H}(\phi_n(\mathbf{x}))) \log(1 - \theta_\Gamma(\mathbf{x}))]d\mathbf{x},$$

yielding $\hat{\theta}_\Gamma(\mathbf{x}) = \frac{1}{N} \sum_{n=1}^N \tilde{H}(\phi_n(\mathbf{x}))$.

5 Experimental Results

We test the proposed approach on 50 MR brain scans. Some of the subjects in this set are diagnosed with the first episode schizophrenia or affective disorder. The MR images (T1, $256 \times 256 \times 128$ volume, $0.9375 \times 0.9375 \times 1.5\text{mm}^3$ voxel size) were acquired by a 1.5-T General Electric Scanner. The data was originally acquired for brain morphometry study [9]. In addition to the MR volumes, manual segmentations of three structures (superior temporal gyrus, amygdala, and hippocampus) in each hemisphere were provided for each of the 50 individuals and used to evaluate the quality of the automatic segmentation results. MR images are preprocessed by skull stripping. The volumes were aligned using B-spline registration according to [2].

Assuming that the manual segmentation of a single instance is given, we initialize the latent atlas θ_Γ by the Heaviside function of a single manual segmentation smoothed with a Gaussian kernel of width σ ($\sigma = .35$ in our experiments). We ran the experiments twice, initializing the level-set functions of the signed distance function of the given manual segmentation or a sphere, with center and radius corresponding to the manually segmented structure. The results obtained using the first mode of initialization were slightly better.

We excluded the image associated with the given manual segmentations from the ensemble.The algorithm was implemented in Matlab and ran on the average 7.5 minutes for 50 cropped volumes, excluding the time of the initial estimate of the background intensity of the Gaussian mixture. About 7 iterations were needed until convergence which was obtained when the update of the level-set functions did not induce changes in the corresponding boundaries.

We used the Dice score to evaluate segmentations. The results are shown in Figure 1. In contrast to [14] we do not use spatial priors, neither do we use hierarchical multi-stage segmentation model. Nevertheless, the Dice scores obtained by our method approache the state-of-the-art atlas-based segmentation results reported there. To exemplify the significance of a latent atlas, generated concurrently with the segmentation, we also show a comparison to the segmentation results obtained by using an atlas constructed from a single manual segmentation smoothed by a Gaussian kernel, without the update procedure. Figure 2 shows segmentation examples of the three pairs of structures in representative individual brains. Coronal views of the resulting 3D atlases for each pair of structures are shown in the fourth column of Figure 2.
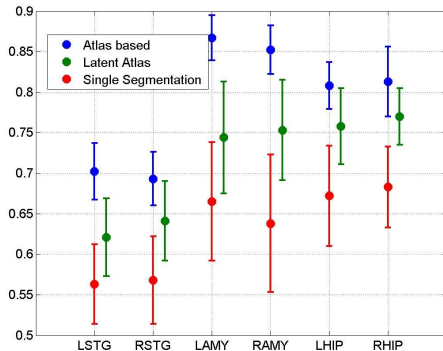
Fig. 1. The mean and standard deviation of the Dice scores calculated for six structures in the ensemble. The latent atlas segmentation (green) is compared to the atlas-based segmentation (blue) reported in [14] and to the segmentation obtained by using a single manual segmentation as an atlas (red).

6   Discussion and Conclusions

We presented a level set framework for segmentation of MR image ensembles, that is motivated by a generative probabilistic model. Unlike most previous methods, we do not use spatial priors in the form of a probabilistic atlas. Instead, spatial latent parameters, which form a 'dynamic atlas', are inferred from the data set through an alternating minimization procedure.

The quality of the segmentation results obtained for ensembles of brain structures shows that the proposed method presents a reasonable alternative to standard segmentation techniques when a compatible atlas is not available. An ongoing research is now conducted to demonstrate the ability of the proposed algorithm to handle pathological cases (e.g., in a longitudinal or a multimodal study) where the atlas-based approach still fails.

References

1. J. Ashburner and K. Friston. Unified segmentation. *NeuroImage*, 26:839–851, 2005.
2. Balci *et al*. Free-form B-spline deformation model for groupwise registration. In *Proc. of MICCAI 2007 Statistical Registration Workshop: Pair-wise and Groupwise Alignment and Atlas Formation*, pages 23–30, 2007.
3. P.L. Bazin and D.L. Pham. Statistical and topological atlas based brain image segmentation. In *MICCAI*, 94–101, 2007.
4. J. Besag. Statistical analysis of non-lattice data. *The Statistician*, 24(3):179–195, 1975.
5. Bhatia *et.al*. Groupwise combined segmentation and registration for atlas construction. In *MICCAI*, volume 4791, 532–540, 2007.
6. T.F. Chan and L.A. Vese. Active contours without edges. *TIP*, 10(2):266–277, 2001.
7. A. Dempster, N. Laird, and D. Rubin. Maximal likelihood form incomplete data via the EM algorithm. *Proceedings of the Royal Statistical Society*, 39:1 – 38, 1977.
8. Fischl *et al*. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002.
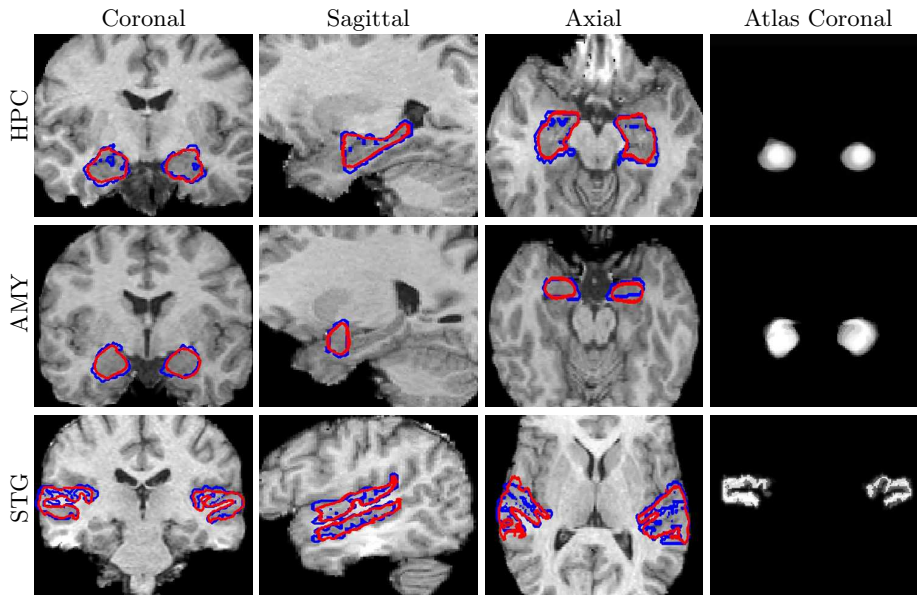
Fig. 2. Three cross-sections of 3D segmentations of Hippocampus, Amygdala and Superior Temporal Gyrus in the left and right hemispheres. Automatic segmentation is shown in red. Manual segmentation is shown in blue. Fourth column: Coronal views of the resulting atlases for each pair of structures.

9. Hirayasu *et al*. Lower left temporal lobe MRI volumes in patients with first-episode schizophrenia compared with psychotic patients with first-episode affective disorder and normal subjects. *Amer. J. Psychiatry*, 155(10):1384–1391, 1998.
10. S.Z. Li. *Markov Random Field Modeling in Computer Vision*. Springer-Verlag, 1995.
11. N.A. Lord, J. Ho, and B.C. Vemuri. Ussr: A unified framework for simultaneous smoothing, segmentation, and registration of multiple images. In *ICCV*, 1–6, 2007.
12. N. Paragios and R. Deriche. Geodesic active regions: A new paradigm to deal with frame partition problems in computer vision. *JVCIR*, 13:249–268, 2002.
13. Pohl *et al*. A bayesian model for joint segmentation and registration. *NeuroImage*, 31(1):228 – 239, 2006.
14. Pohl *et al*. A hierarchical algorithm for MR brain image parcellation. *IEEE TMI*, 26(9):1201–1212, 2007.
15. Pohl *et al*. Using the logarithm of odds to define a vector space on probabilistic atlases. *Medical Image Analysis*, 11(6):465–477, 2007.
16. T. Riklin-Raviv, N. Sochen, and N. Kiryati. Shape-based mutual segmentation. *IJCV*, 79:231–245, 2008.
17. Tu *et al*.   Brain anatomical structure segmentation by hybrid discriminative/generative models. *IEEE TMI*, 27(4):495 – 508, 2008.
18. Van Leemput *et al*. Automated model-based tissue classification of MR images of the brain. *IEEE TMI*, 18(10):897–908, 1999.
19. J. Yang and J.S. Duncan. Joint prior models of neighboring object for 3D image segmentation. In *CVPR*, 314–319, 2004.