A General Framework for Image Segmentation Using Ordered Spatial Dependency

Mikaël Rousson and Chenyang Xu

Department of Imaging and Visualization
Siemens Corporate Research, Princeton, NJ

Abstract. The segmentation problem appears in most medical imaging applications. Many research groups are pushing toward a whole body segmentation based on atlases. With a similar objective, we propose a general framework to segment several structures. Rather than inventing yet another segmentation algorithm, we introduce inter-structure spatial dependencies to work with existing segmentation algorithms. Ranking the structures according to their dependencies, we end up with a hierarchical approach that improves each individual segmentation and provides automatic initializations. The best ordering of the structures can be learned off-line. We apply this framework to the segmentation of several structures in brain MR images.

1 Introduction

Different anatomical structures often have strong spatial dependency among each other. This spatial dependency is usually present in a hierarchical manner, i.e., the shape and pose variations of one structure is fully or partially bounded by those of other more stable structures. We refer to this type of spatial dependency as *ordered spatial dependency* due to its ordered nature. Radiologists routinely rely on ordered spatial dependency to help them locating and identifying structures that have large variations in shape, pose, and appearance by searching its presence relative to other structures that are much easier to identify. In this paper, we would like to take benefit from this inter-structure ordered spatial dependency in an explicit manner by proposing a novel general image segmentation framework. The proposed framework learns the ordered spatial dependency from pre-segmented training images and applies the learned model to improve both the performance and robustness of individual segmentation algorithms utilized.

A key benefit of the proposed framework is that it is not another new segmentation algorithm but a new general framework that could integrate any existing segmentation algorithms. The motivation of this work arises from the fact that many powerful and effective segmentation algorithms such as seeded region growing[1], watershed[16], active contours[9,3], and graph cuts[2,12] have been proposed and used widely in particular for medical image segmentation applications. The topic of devising a segmentation framework that combines existing segmentation algorithms to achieve better results has recently emerged as a new promising research direction [13]. The work [13] proposed a framework that computes an improved segmentation result based on optimizing parameter

values of multiple segmentation algorithms. Similarly, our proposed framework also improves the segmentation results over the individual algorithms utilized, but it achieves this in a different way by focusing on segmenting objects in a hierarchical manner using the ordered spatial dependency.

Observing that structures often have a strong spatial dependency, we define a new spatial prior based on neighboring structures. This dependency is introduced by registering the structure of interest to a common reference coordinate system based on neighboring structures. This can be achieved by computing the elastic matching of the neighboring structures from one image to a reference one, and then applying it to the structure of interest. This modeling can be implemented for each structure, based on the ones already segmented. This leads us to the definition of a hierarchical segmentation framework.

The proposed work has two important contributions: 1) the explicit modeling and utilization of ordered spatial dependency for segmentation; 2) the estimation of the optimal segmentation sequence for segmenting multiple structures. Our work is closely related to atlas-based segmentation (cf. [5,11,8]), which treats segmentation as a registration problem by elastically matching a pre-segmented atlas to the target image. Atlas-based segmentation approaches are generally better-suited for segmenting structures that are stable over the population of study. Our proposed framework uses elastic matching to enforce the spatial dependency and restricts the plausible segmentation space rather than using it to obtain the final segmentation. The actual segmentation of each structure is still performed using a pre-selected segmentation algorithm.

Other closely related work are the active shape and appearance models [6,7,10], which assume a statistical correlation between the shape or appearance of the organs over population. Our proposed segmentation framework uses a weaker assumption by modeling the relative locations of the structures between one and another. It is also worth noting that recent works on joint segmentation and registration [17,15,14], that also use segmentation and registration in an iterative manner, are primarily used for segmenting two or more images simultaneously and do not use ordered dependency. In fact, one could incorporate active shape and appearance models within our framework as its building blocks like any other segmentation algorithms.

In the following sections, we detail the proposed segmentation framework and demonstrate its utility in segmenting multiple structures from MR brain images.

2   Modeling Inter-structure Spatial Dependency

Let $\{\mathcal{S}_1, \ldots, \mathcal{S}_N\}$ be the set of structures of interest in an image. We assume a dataset of M annotated images to be available. We note $\{s_{ij}; i = 1, \ldots, N, j = 1, \ldots, M\}$ the complete set of structures. Given a manual segmentation s of a structure $\mathcal{S} \in \{\mathcal{S}_1, \ldots, \mathcal{S}_N\}$, we propose a smooth approximation of the conditional probability of an image location x to be inside the structure s:

$$p_{\mathcal{S}}(x|s) \propto \exp(H_\epsilon(\phi(x)) - 1),$$

where ϕ is the distance transform of s and H_ϵ is a regularized Heaviside function with ϵ controlling the level of smoothness [3]. This distribution gives high probability to voxels inside s and low probability to the ones outside, and the smoothness of the transition is related to the distance to the interface. Also, we note the conditional probability of a voxel x to belong to the background of s as $p_{\bar{\mathcal{S}}}(x|s)$:

$$p_{\bar{\mathcal{S}}}(x|s) \propto \exp(-H_\epsilon(\phi(x))).$$

Now, if we want to combine all the annotated instances of a structure $\mathcal{S}_i$ to define the spatial prior probability of the structure $\mathcal{S}_i$, we need to place the manual segmentations in a common reference[1]. It is at this point that we consider the ordered spatial dependency, i.e., $\mathcal{S}_i$'s dependency on known neighboring structures. The principle is to align all the instances s_{ij} of $\mathcal{S}_i$ to a common coordinate system using the known structures as anchors. This is done by estimating a warping between each instance of the anchor structure(s) to a chosen reference. These warpings are obtained using an image based registration algorithm [4] applied on level set representations of each structure instance. If several anchor structures are available, they are merged to form a single shape composed of several components. This allows us to constrain even more the deformation field between the structures.

Then, these warpings are applied to the corresponding structures s_{ij}. Let $\tilde{s}_{ij}$ be the segmentation transformed by the warping ψ_{ij} and $\tilde{\phi}_{ij}$ be its level set representation, the spatial prior probability of $\mathcal{S}_i$ and its background $\bar{\mathcal{S}}_i$ are defined in the reference image as the geometric mean of each individual prior:

$$p_{\mathcal{S}_i}(x) \propto \left(\prod_{j=1}^{M} \exp\left(H_\epsilon(\tilde{\phi}_{ij}(x)) - 1 \right) \right)^{\frac{1}{M}} , p_{\bar{\mathcal{S}}_i}(x) \propto \left(\prod_{j=1}^{M} \exp\left(-H_\epsilon(\tilde{\phi}_{ij}(x)) \right) \right)^{\frac{1}{M}}$$

Up to now, we did not give any detail about which anchor structures were considered to estimate the warping. A priori, we do not know which structure $\mathcal{S}_i$ is spatially dependent on. We propose to learn these dependencies by defining the spatial probability of $\mathcal{S}_i$ with respect to other structures $\{\mathcal{S}_k, k \neq i\}$. We denote $\mathcal{V}_i$ all possible subsets of $\{S_k, k \neq i\}$, $v_i \in \mathcal{V}_i$ a subset of segmented structures, and v_{ij} the corresponding annotated structures in the training image j. With these notations, all structure s_{ij} are registered to a reference image by estimating the warpings that align v_{ij} to a reference set v_{ir}. Therefore, for each choice of subset v_i, we end up with different registrations and hence, a different spatial prior probability for $\mathcal{S}_i$. Since this probability is subject to the selected subset v_i, we will use the notations $p_{\mathcal{S}_i}(x|v_i)$ and $p_{\bar{\mathcal{S}}_i}(x|v_i)$ to denote respectively the prior probabilities of a voxel to be inside and outside the structure $\mathcal{S}_i$, given a set of known segmentations v_i. The choice of the optimal subset of reference for each structure is studied in Section 3.2.

---

[1] The common reference is chosen arbitrary in this work. It would be interesting to study how important this choice is.

In the following section, we incorporate this spatial prior in a general segmentation framework and we detail the complete framework obtained when a level set based approach is considered.

3 Integrating Spatial Prior in the Segmentation Process

Given an image I and a set of segmented structures v, we want to extract another structure s of the class $\mathcal{S}$ (the class $\mathcal{S}$ is one of the classes $\mathcal{S}_i$ considered previously). We consider a statistical formulation of this segmentation problem using a maximum a posteriori estimation. This consists in maximizing the posterior conditional distribution $p(s|I, v)$. Making the assumption that I and v are not correlated, the optimal structure is the one maximizing

$$p(s|I, v) = p(s|I)\, p(s|v).$$

The first term can be expressed with any statistically defined segmentation algorithm whereas the other one can integrate the spatial prior learned in the previous section. To incorporate this prior knowledge, we make the assumption that the prior probabilities of the locations x are independent and identically distributed. This allows us to incorporate the spatial prior probability term introduced in the previous section:

$$p(s|v) = \prod_{x \in s_{\text{in}}} p_{\mathcal{S}}(x|v) \prod_{x \in s_{\text{out}}} p_{\bar{\mathcal{S}}}(x|v),$$

where s_{in} and s_{out} are respectively the parts of the image inside and outside the structure s. This formulation is very general and any efficient segmentation approaches like graph-cuts [2,12] and surface evolutions [9,3] can be considered. In the following, we develop our system using level set based surface evolutions but this should not be seen as the only possibility.

3.1 Level Set Based Segmentation

In the level set framework, the structure of interest is represented as the zero crossing of an embedding function $\phi : \Omega \to \mathbb{R} : s = \{x \in \Omega | \phi(x) = 0\}$. Hence, the problem of finding the surface s becomes the one of finding a real function ϕ that maximizes: $p(s|I, v) \to p(\phi|I, v)$. Equivalently, the optimal solution can be obtained from the minimization of the energy:

$$E(\phi) = -\log p(\phi|I, v) = -\log p(\phi|I) - \log p(\phi|v)$$

We follow [3] to define the first term with a region-based criteria and a regularity constraint. To use the spatial prior, we first need to register the anchor structures from the current image to the reference ones used for modeling. Let ψ be the obtained warping, the whole energy can be written as follows:

$$E(\phi) = -\int_{\Omega} \left(H_{\phi} \log p_{\text{in}}(I(x)) + (1 - H_{\phi}) \log p_{\text{out}}(I(x)) + \nu |\nabla H_{\phi}| \right) \, dx$$
$$- \lambda \int_{\Omega} \left(H_{\phi} \log p_{\mathcal{S}}(\psi(x)|v) + (1 - H_{\phi}) \log p_{\bar{\mathcal{S}}}(\psi(x)|v) \right) \, dx,$$

where p_{in} and p_{out} are the intensity distributions inside and outside the structures. They can be estimated on-line or a priori from the learning set. We minimize this energy using a gradient descent obtained from the Euler-Lagrange equations. This gives the following curve evolution:

$$\phi_t = \delta(\phi)\left(\nu\,\text{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) + \log\frac{p_{\text{in}}(I)}{p_{\text{out}}(I)} + \lambda\,\log\frac{p_{\mathcal{S}}(\psi(x)|v)}{p_{\bar{\mathcal{S}}}(\psi(x)|v)}\right)$$

$$= \delta(\phi)\left(\nu\,\text{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) + \log\frac{p_{\text{in}}(I)}{p_{\text{out}}(I)} + \frac{\lambda}{M}\sum_{j=1}^{M}\left(2H_\epsilon(\tilde{\phi}_{cj}(\psi(x))) - 1\right)\right)$$

where $\tilde{\phi}_{cj}$ stand for the warpings estimated during the modeling phase for the current shape. The segmentation of s is obtained by evolving ϕ according to this equation until convergence (the initialization is discussed in the next paragraph). For an efficient implementation, $\sum_{j=1}^{M}\left(2H_\epsilon(\tilde{\phi}_{cj}(x)) - 1\right)$ can be estimated off-line, and then warped to the current image domain using ψ.

3.2   Hierarchical Segmentation

For a given ordering, we can run the segmentation algorithm on each structure successively. This process can be initialized automatically if we are able to segment the first structure without any spatial prior. In most medical images, this can be done easily by starting with the envelope of the body. Then, to segment each structure, we also need to initialize the associated level set. A straightforward solution is to place seeds inside each structure. Obviously, this would give a good initialization but it requires user interaction. The spatial prior can be used to make these initializations automatic by selecting the voxels with a prior probability greater than a threshold τ. More precisely, the initial level set ϕ_i^0 used to extract the structure $\mathcal{S}_i$ is set as follow[2]:

$$\begin{cases} \phi_i^0(x) = +1, & \text{if} \quad \log\dfrac{p_{\mathcal{S}}(\psi(x)|v)}{p_{\bar{\mathcal{S}}}(\psi(x)|v)} \geq \tau, \\ \phi_i^0(x) = -1, & \text{otherwise.} \end{cases}$$

With this technique, the segmentation of all N structures can be obtained automatically. Only the weights ϵ, ν, λ and τ must be set before starting the process.

4   Estimation of the Optimal Segmentation Sequence

To learn the subset v_i that helps the segmentation of S_i the best, we propose to apply the segmentation on a second set of annotated images. For each v_i, we can measure the quality of the segmentation according to a chosen similarity measure $\mathcal{M}$ between the automatic and "true" segmentation. Assuming that,

---

[2] Once initialized with the spatial prior, the level set is projected to a signed distance functions. This is repeated after each iteration of the level set evolution.
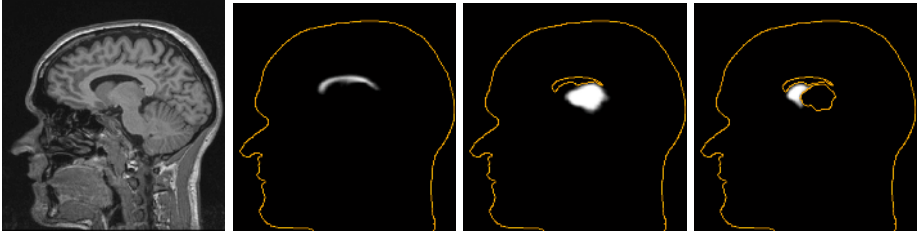
Fig. 1. Location priors corresponding to the optimal ordering - From left to right are the shown the reference image, and the spatial priors of (1) the lateral ventricle given the skull, (2) the thalamus given the skull and the lateral ventricle, and (3) the caudate nucleus given all other structures

if S_j depends on S_k, S_k cannot depend on S_j, the objective is to estimate the best ordering of the structures such that structures classified higher can be used to extract lower-classified ones. Once all the segmentations obtained for a given ordering, we measure the overall quality of the process by comparing the results with the manual segmentations according to a similarity measure $\mathcal{M}$. The optimal ordering is then given by:

$$\hat{O} = \arg\max_{O \in \mathcal{O}} \sum_{i=1}^{N} \sum_{j=1}^{M} \mathcal{M}(s_{ij}, \hat{s}_{ij}(O)),$$

where $\mathcal{O}$ is set of all permissible orderings, and $\hat{s}_{ij}(O)$ is the segmentation obtained automatically in the image j for the structure $\mathcal{S}_i$ using the ordering O. As for the similarity measure, we use the Dice coefficient. In general the number of structures to extract is relatively small ($<$10) and all combinations can be tested. If we fix the first structure, the number of combinations is equal $(N-1)!$. Even though, this number gets high for $N = 10$, this is an off-line process and the user can introduce heuristics to reduce the possible orderings. For example if choosing a given structure at a high level conducts to a bad segmentation of the next one, a whole set of possible orderings can be discarded. We are conducting further investigations to reduce efficiently the possible orderings in a more theoretical way.

5   Hierarchical Segmentation of Brain Structures in MR Images

We validate this segmentation framework on several structures of the brain in MR images: the lateral ventricles, the caudate nucleus, the thalamus and the skull. These structures were annotated manually in 13 different sagittal slices.

The first step is the learning of the optimal ordering. To start this process, we must be able to segment automatically the first anchor structure. For the brain image shown in Figure 1, this is relatively simple if we consider the skull. Initializing the level set with a seed in the background is sufficient to get its
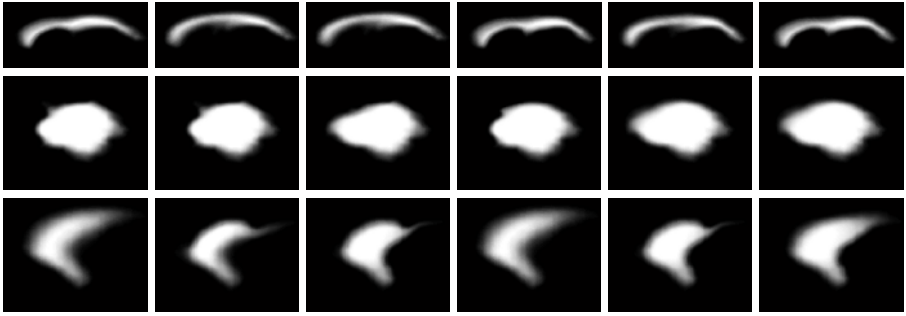
Fig. 2. Location priors for each possible ordering - Each column shows the spatial prior obtained for one particular ordering. Priors corresponding to the optimal ordering are shown in the third column.
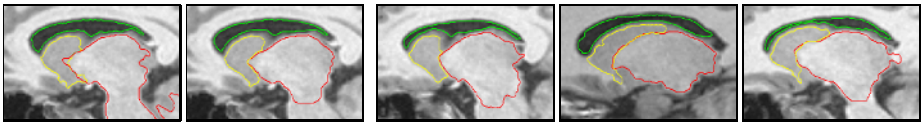

Fig. 3. Examples of automatic segmentations - From left to right: segmentation without prior and with manual initialization, automatic segmentation of the same image, and three different results obtained with the "optimal" ordering

segmentation. Then, each structure can be initialized automatically by following Section 3.2. We set the threshold τ to the maximum value of the location map J minus 0.1. This guarantees to give a seed inside the structure. Having 4 structures, the number of possible orderings is 6. We have tested our algorithm for each possible choice. Figure 1 shows the sequence that maximizes the overall Dice coefficient between the obtained segmentations and the manual ones for the whole training set. Figure 2 shows the location maps estimated for each ordering.

Once the optimal ordering known, we can validate the approach using a leave-one-out strategy on the 13 available images. A few results are presented in Figure 3. As quantitative validation, we computed the Dice coefficient for each of the 52 automatically computed segmentation, giving an average above 0.8.

6   Conclusion

We have presented a novel image segmentation framework that learns the ordered spatial dependency among structures to be segmented and applies it in a hierarchical manner to both provide automatic initializations and improve each individual segmentation algorithm's performance. We demonstrated the efficacy of the proposed framework by applying it to the MR brain image segmentation with level set algorithm as its segmentation algorithm. Future work includes applying this framework to more applications with more types of segmentation algorithms. We believe that the paradigm of "boosting" segmentation performance

by combining existing segmentation algorithms into a systematic framework is a promising research direction and the work presented in this paper is one step along this direction.

References

1. R. Adams and L. Bischof. Seeded region growing. *IEEE T. PAMI*, 16(16):641–647, 1994.
2. Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE T. PAMI*, 23(11):1222–1239, 2001.
3. T.F. Chan and L.A. Vese. Active contours without edges. *IEEE T. IP*, 10(2):266–277, 2001.
4. C. Chefd'hotel. *Geometric Methods in Computer Vision and Image Processing: Contributions and Applications*. PhD thesis, Ecole Normale Suprieure de Cachan, April 2005.
5. D.L. Collins, C.J. Holmes, T.M. Peters, and A.C. Evans. Automatic 3-D model-based neuroanatomical segmentation. *Human Brain Mapping*, 3:190–208, 1995.
6. T.F. Cootes, D. Cooper, C.J. Taylor, and J. Graham. Active shape models - their training and application. *Comput. Vis. Image Und.*, 61(1):38–59, 1995.
7. T.F. Cootes, G.J. Edwards, and C.J. Taylor. Active appearance models. *IEEE T. PAMI*, 23(6):681–685, 2001.
8. S.L. Hartmann, M.H. Parks, P.R. Martin, and B.M. Dawant. Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations: part II, validation on severely atrophied brains. *IEEE T. MI*, 18(10):917–926, 1999.
9. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int'l J. Comp. Vis.*, 1(4):321–331, 1988.
10. A. Kelemen, G. Székely, and G. Gerig. Elastic model-based segmentation of 3-D neuroradiological data sets. *IEEE T. MI*, 18(10):828–839, 1999.
11. R. Kikinis, M.E. Shenton, and et al. A digital brain atlas for surgical planning model-driven segmentation, and teaching. *IEEE T Vis. Comp. Graphics*, 2(3):232–240, 1996.
12. H. Lombaert, Y. Sun, L. Grady, and C. Xu. A multilevel banded graph cuts method for fast image segmentation. In *Proc. ICCV*, pages 259–265, 2005.
13. M. Maddah, K.H. Zou, W.M. Wells, R. Kikinis, and S.K. Warfield. Automatic optimization of segmentation algorithms through simultaneous truth and performance level estimation (STAPLE). In *Proc. MICCAI*, pages 274–282, 2004.
14. G. Unal and G. Slabaugh. Coupled PDEs for non-rigid registration and segmentation. In *Proc. CVPR*, pages 168–175, 2005.
15. B. Vemuri and Y. Chen. Joint image registration and segmentation. In S. Osher and N. Paragios, editors, *Geometric Level Set Methods in Imaging, Vision, and Graphics*, pages 251–269. Springer Verlag, 2003.
16. L. Vincent and P. Soille. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE T. PAMI*, 13(6):583–598, 1991.
17. A. Yezzi, L. Zollei, and T. Kapur. A variational framework for joint segmentation and registration. In *Proc. Workshop on MMBIA*, pages 44–49, 2001.