Joint Segmentation-Registration of Organs Using Geometric Models

Alper Ayvaci and Daniel Freedman

Abstract— **In this paper, we present a novel method for the segmentation of the organs found in CT and MR images. The proposed algorithm utilizes the shape model of the target organ to gain robustness in the case where the objective organ is surrounded by other organs or tissue with the similar intensity profile. The algorithm labels the image based on the graph-cuts technique and incorporates the shape prior using a technique based on level-sets. The method requires proper registration of the shape template for an accurate segmentation, and we propose a unified registration-segmentation framework to solve this problem. Furthermore, to reduce the computational cost, the algorithm is designed to run on watershed regions instead of voxels. The accuracy of the algorithm is shown on the medical examples.**

Keywords: segmentation; registration; graph cuts; shape priors; watershed transformation.

I. Introduction

Accurate and robust segmentation of the organs found in medical images is often a challenging task, as illustrated in Figure 1. Specifically, the following problems cause difficulties.

- The presence of the significant noise in CT and MR images often forms strong edges inside organs.
- The target organ is often surrounded by other structures which causes shallow boundaries.
- The shape of the organs varies with many factors such as age, sex, weight and height.

In many locations, the low level image features are weak or deceptive, as a result of the first two problems. Any segmentation algorithm which relies heavily on these features would simply fail. Therefore, shape models are necessary for an accurate segmentation.

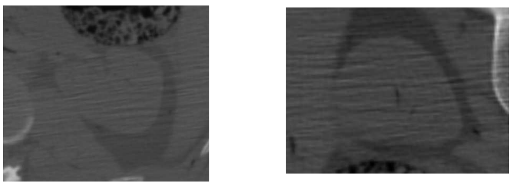

Fig. 1. Segmentation of organs is often a difficult task due to: (left) surrounding structures with similar intensity profile and (right) the presence of significant noise in images.

The algorithm proposed in this paper relies on graph cuts based optimization. To incorporate shape models into this framework, it modifies and extends the technique introduced by Freedman and Zhang [4]. Specifically, it uses the distance map of the shape template which has a zero level set on its boundary. If this map is registered on the image properly, it would form a minimal cut around the organ even though the low level features are weak. However, proper registration of the template is the key obstacle. The work by Freedman and Zhang [4] relied on a combination of user input and heuristics. In this paper, we both eliminate the need for user input, as well as pose the problem of template registration in a principled way. In particular, we propose to combine the registration and the segmentation steps in such a way that the registration step tries to minimize the segmentation energy as the cost in each iteration. If it were to run on raw voxel data, this algorithm would be expected to have high computational cost. Hence we designed the algorithm to run on watershed regions instead of voxels, and this reduces the cost dramatically.

The paper is organized as follows. In Section II, the related work is examined. Section III presents the segmentation algorithm in detail. In Section IV, implementation issues are discussed, and the value of the algorithm is demonstrated on two examples: segmentation of the bladder from a scan of the mid-section, and segmentation of the corpus callosum from a scan of the brain. Finally, Section V concludes the paper.

II. Related Work

Two sub-categories of the object segmentation bear on the work presented in this paper: segmentation using graph-cuts minimization and shape model based segmentation. We briefly address both of them.

The idea of performing interactive segmentation using graph-cuts was introduced by Boykov and Jolly [1], [2]. Based on a similar concept, Li et al. [6] presented a method combining watershed segmentation and graph cut minimization to build an interactive image cut-out system. Lombaert et al. [7] improved the idea of Boykov and Jolly [1], [2] with a multilevel banded heuristic for computation of graph cuts that is motivated by the well-known narrow band algorithm in level set computation.

To utilize shape models, some researchers offer to transform shape templates to have the best fit for segmentation. Freedman and Zhang [4] incorporated the shape template into the graph-cuts formulation as a distance function. Their method transforms the template using the Procrustes Method [5], and through the calculation a Gaussian pyramid of the image. This paper relies crucially on having user input. Sclaroff and Liu [10] present a two stage algorithm: oversegmentation using a traditional region segmentation algorithm, followed by a simultaneous region merging and

This work was supported by NIH Grant 1R01CA116743-01

A. Ayvaci and D. Freedman are with Department of Computer Science, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, USA 12180 ayvaca@cs.rpi.edu and freedman@cs.rpi.edu

object identification. A somewhat different strategy is to unify registration and segmentation. Pohl et al. [9] uses an Expectation Maximization-based algorithm which simultaneously estimates image inhomogeneities, anatomical label map, and a mapping from the atlas to the image space

III. The Algorithm

In this section, the segmentation algorithm is described in detail. Section III-A explains how to over-segment the input image. Section III-B gives a brief description of segmentation of the objects with graph cuts and describes how this algorithm is modified to operate on watershed regions. Section III-C presents the incorporation of shape priors into the energy function and its final form. In Section III-D, the energy minimization mechanism is clarified, and in Section III-E, we propose an algorithm which segments the object while registering the template at the same time.

A. Over-Segmenting the Input Image

In order to reduce the computational cost of the minimization step, the proposed segmentation algorithm operates on homogeneous regions instead of voxels. Graph-cuts based minimization runs much faster on regions than on voxels, because the number of operands is reduced dramatically. In this step, the set of homogeneous regions R is constructed via the watershed transformation, which is implemented based on a top-down, gradient descent strategy.

B. Basic Graph Cuts Segmentation

Boykov and Jolly [1], [2] introduced a method for segmentation based on computing the min-cut / max-flow of a graph. A segmentation consists of classifying the voxels into two sets: object and background, and it is scored based on the appearance and the boundary models. Unlike their method, the operands of our algorithm are watershed regions instead of voxels. Therefore, their energy function should be modified to be compatible with regions. Let $r \in R$ and $A_r = 0$ or 1 if r is in the background or the object, respectively. Let $R_r(A_r)$ be the individual region matching cost for region r; let B_{rs} be a measurement of the similarity between the regions r and s. Then the modified segmentation energy is

$$E_R = \sum_{r \in \mathcal{R}} D_r(A_r) + \sum_{(r,s) \in \mathcal{N}_{\mathcal{R}} : A_r \neq A_s} B_{rs}$$

where $\mathcal{N}_{\mathcal{R}}$ is the set of neighboring regions. The precise forms of the term B_{rs} is given in Section IV.

A polynomial time combinatorial algorithm exists for minimizing E_R. However, a good segmentation generally results only with the aid of some user input. We will remove this requirement of supplying user input by adding in a shape prior, which can be automatically registered to the image.

C. Addition of Shape Priors

Our shape model is a single binary object template. Because of the restrictions on the energy function which can be minimized using graph cuts, incorporating a shape model into the energy function is a tough problem. Freedman and Zhang [4] proposed a method to incorporate a template-based shape prior into this type of energy function. Basically, this method specifies the template as a distance function whose zero level-set corresponds to the object boundary. Let $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ be such that

$$\mathbf{S} = \{x \in \mathbb{R}^3 : \phi(x) = 0\}$$

where $\mathbf{S}$ is the point-set identifying the surface which is the boundary of the object.

Our method uses the same idea to append the shape model to the energy function, but in a different way. First, for each region, its distance function value is added as a data term to the energy function instead of as a pair-wise term. Second, we use a signed distance function. Thus, the regions, overlapping with the template have negative distance values, and those not overlapping have positive values. Third, the new shape term is plausible, given our energy function which is defined on the regions. Consider a region which is inside the template, but close to its boundary. Although the shape prior has a preference for this region to be classified as part of the object, it is easy to assign it to the background if that is desirable, as the cost is only the distance of the region from the boundary, which is small. In this way, we encourage competition between the object and the background on the boundary of the template. By contrast, the regions which are closer to the center of the template will be more difficult to reclassify, as they have higher distance values.

In this context, the new shape term for the regions can be written as

$$E_{R_S} = \sum_{r \in \mathcal{R}} \varphi_r(A_r)$$

when the function $\varphi_r(A_r)$ is defined by

$$\varphi_r(A_r = 0) = \begin{cases} 0 & \phi(\mathcal{C}_r) < 0 \\ |\phi(\mathcal{C}_r)| & \phi(\mathcal{C}_r) \geq 0 \end{cases}$$

$$\varphi_r(A_r = 1) = \begin{cases} |\phi(\mathcal{C}_r)| & \phi(\mathcal{C}_r) < 0 \\ 0 & \phi(\mathcal{C}_r) \geq 0 \end{cases}$$

where $\mathcal{C}_r$ is the centroid of the region r.

Note that the minimization of E_{R_S} needs accurate registration of the shape template to the target object. This problem will be discussed in the section III-E.

With all the terms defined, the final energy function is

$$E_R = \nu \sum_{r \in \mathcal{R}} \left((1-\lambda) D_r(A_r) + \lambda \varphi_r(A_r)\right) + \sum_{(r,s) \in \mathcal{N}_{\mathcal{R}} : A_r \neq A_s} (1-\nu) B_{rs} \tag{1}$$

D. Minimizing the Energy

Graph cut techniques may be used to find the global optimum in certain combinatorial optimization problems [3] and they are widely used in computer vision. In the case of minimization of the energy function (1), the graph cuts method are applicable. The key step in using these techniques

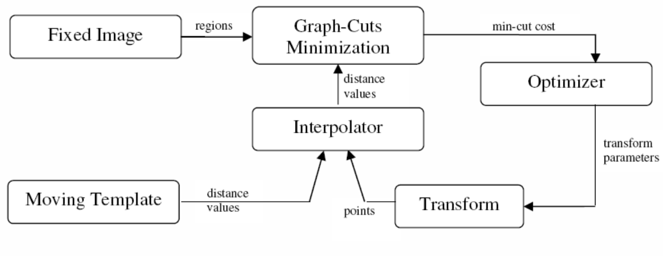

Fig. 2. The steps of the joint segmentation-registration algorithm.

on (1) is the construction of the relevant graph. Let $G = (V, E)$ be the graph corresponding to the our energy function E_R. The minimum-cut on the graph G gives the minimum energy of the function E_R. [3] presents an optimal algorithm that finds the minimum cut on the graph G.

E. Registration

The algorithm, operating on a properly registered template, captures the boundary of the object, even if it is diffuse or weak. However, registration of the template on the input image is the key problem. We propose a unified solution that combines registration and segmentation steps to overcome this problem. Specifically, the registration algorithm uses the segmentation energy as a measure of how well the template is fit to the input image, and minimizes that energy over the space of transformations. In this way, the proposed algorithm registers the template and segments the image simultaneously.

Figure 2 clarifies the steps of our algorithm. The *Fixed Image* comprises the watershed regions extracted from the input image, and the *Moving Template* is the distance map of the shape model. Our algorithm has an iterative structure. For each iteration, the *Moving Template* is first transformed under the parameters determined by the *Optimizer*. Subsequently, the transformed distance map and the watershed regions are used to construct the energy function E_R; this function is then minimized by graph-cut minimization to label (segment) the image. The *Optimizer* uses the capacity of minimal cut to estimate the transformation parameters for the next iteration. In this way, each subsequent iteration minimizes the corresponding capacity of the relevant graph, in other words, the segmentation energy. This process continues until the capacity of the minimal cut converges. The moving template approaches its proper position on the objective in each iteration. The final segmentation based on the optimum parameters is also the desired one.

Given this model, let θ be the transformation, and ϕ_θ be the distance map of the shape template under the transformation θ. Then, the segmentation energy is

$$E_{seg}(\theta) = min_{\{A_r\}} E_R\left(\{A_r\}, \phi_\theta\right)$$

where as before A_r is the labeling for region r. The optimum transformation θ^* and the minimum segmentation energy $E_{seg}(\theta^*)$ is then the solution of the following optimization problem;

$$min_\theta \; E_{seg}\left(\theta\right)$$

In this case, our objective segmentation is the labeling $\{A_r^*\}$ that minimizes the energy function $E_R\left(\mathcal{R}, \phi_{\theta^*}\right)$.

Fig. 3. This figure illustrates the VIP-Man model. Each organ is colored by experts.

We mainly focus on the group of Euclidean similarity transformations in the registration step. However, this approach could easily incorporate affine or even generic non-rigid transformations; there is no restriction on the transformation space.

On the other side, to minimize the function E_{seg} over the transformation space, we have chosen to use the Nelder-Mead method [8], as it does not require explicit computation of the gradient of the objective function.

IV. Experiments

Before presenting the results, some points about implementation should be clarified. First, in our experiments, the pair-wise term B_{rs} is defined in the following form;

$$B_{rs} = |\partial r \cap \partial s| \, \frac{1}{1 + \left(\mu_r - \mu_s\right)^2}$$

where $r, s \in \mathcal{R}$, μ_r and μ_s are the intensity means of the voxels in the regions r and s, and $\partial r \cap \partial s$ is their common boundary. Second, the appearance model D_r is ignored during our experiments. We observed that the due to surrounding objects all possessing similar apperarance models, the appearance term does not aid in segmentation. Third, we incorporated the shape model into the energy function as it is described in the section III-C. It is voxel-based one and is extracted from the VIP-Man model (see Figure 3), which consists of a full-body scan which has been pre-segmented by experts. Finally, our algorithm needs an initial location for the template for the registration step. This initial position is currently determined by the user picking a single point on the objective organ, though other heuristics could easily be used (for example, in the case of lower abdominal area, the pelvic bones could be used to initialize the position).

We tested our algorithm on a whole body image; the running time for the unified registration/segmentation algorithm was under 30 seconds in our experiments. Figure 4 presents our results on the bladder. The first and the second rows of the figure show the position of the template at the iterations 1, 15, and 195 for slices 1 and 10. Observe that the shape template approaches its proper position on the organ. At the final step, they do not overlap completely, even though it is registered properly. The reason is the difference between the shape of the objective organ and the shape template. The last row of the Figure 4 illustrates the segmentation results for the corresponding iterations. As expected, the initial segmentation is far from the desired one and its segmentation energy is 90.10. The second column depicts the segmentation result of the 15^{th} iteration. Even though it is an early step, the result captures most of the bladder without leaking to the other organs. The segmentation energy decreases from

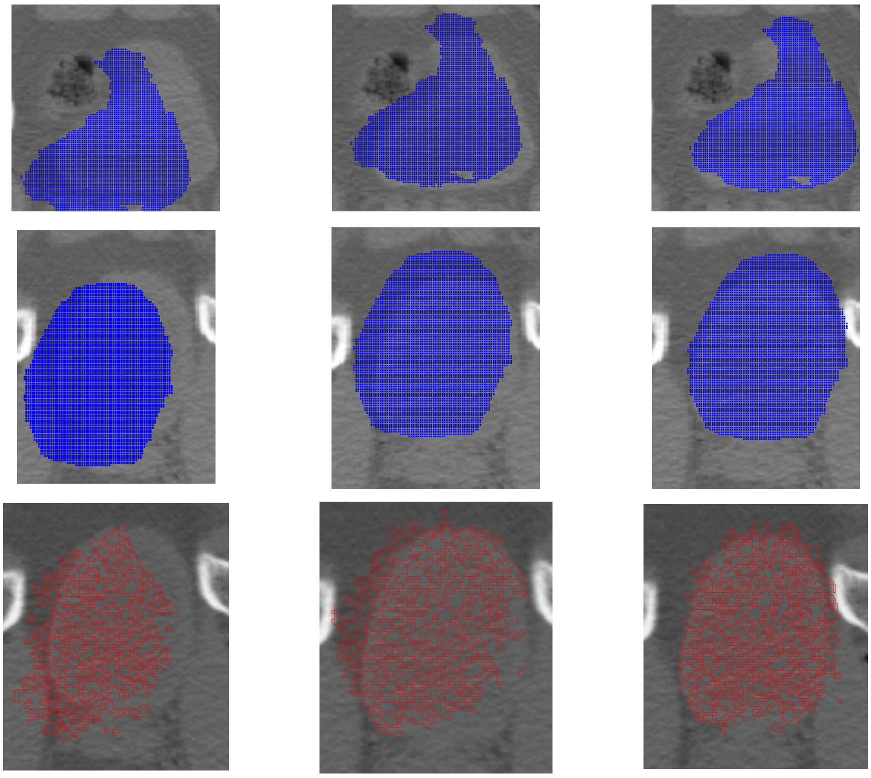

Fig. 4. This figure illustrates the unified segmentation-registration on bladder. The first and the second rows show the position of the template at the iterations 1, 15, and 195 for slices 1 and 10, and the last row depicts the segmentation results for those iterations.

90.10, observed at iteration 1, to 65.81 in iteration 15. Finally, the algorithm converges at the segmentation energy 55.92 in iteration 195 and outputs the segmentation which is illustrated in the third column.

We also experimented our algorithm on the corpus collosum. Figure 5 illustrates the accurate segmentation of the objective structure.

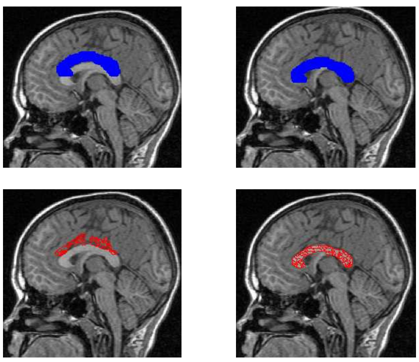

Fig. 5. This figure illustrates the segmentation of the corpus collosum: (left) The initial and (right) the final positions of the template and segmentation results.

Note that the boundary of the final segmentation of the bladder is somewhat irregular. This is due to the fact that in contrast to voxels, the addition of one watershed region to the object changes the boundary dramatically, because of their size and geometrical irregularity. In order to correct the boundary, a similar solution to the one proposed in [6] can be used. Specifically, instead of watershed regions, a graph cuts based optimization can be applied on the voxels in a small band around the final boundary.

V. Conclusion

In this paper, we present a joint registration-segmentation method that extends the work by Freedman and Zhang [4] in two ways. First, it removes the user interaction on segmentation. Second, it resolves the problem of template registration in a principled way.

Incorporation of the single shape template solves the first two problems mentioned in Section I. However, to overcome the third problem, the transformation space should be extended to a non-rigid one as a future improvement.

References

[1] Y. Boykov and M.-P. Jolly. Interactive organ segmentation using graph cuts. In *Proceedings of MICCAI*, pages 276–286, 2000.

[2] Y. Boykov and M.-P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. In *Proceedings of ICCV*, volume 1, pages 105–112, 2001.

[3] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision. *IEEE Transaction on PAMI*, 26(9):1124–1137, 2004.

[4] D. Freedman and T. Zhang. Interactive graph cut based segmentation with shape priors. In *In Proc. of CVPR*, volume 1, pages 755–762, 2005.

[5] I.L.Dryden and K.V.Mardia. *Statistical Shape Analysis.* Chichester New York, 1998.

[6] Y. Li, J. Sun, C.K. Tang, and H.Y. Shum. Lazy snapping. In *In Proc. of SIGGRAPH 2004*, volume 23, pages 303–308, 2004.

[7] H. Lombaert, Y. Sun, L. Grady, and C. Xu. A multilevel banded graph cuts method for fast image segmentation. In *In Proc. of ICCV*, volume 1, pages 259–265, 2005.

[8] J.A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7:308–313, 1964.

[9] K. Pohl, J. Fisher, J. Levitt, M. Shenton, R. Kikinis, W. Grimson, and W. Wells. A unifying approach to registration, segmentation, and intensity correction. pages 310–318, 2005.

[10] S. Sclaroff and L. Liu. Deformable shape detection and description via model-based region grouping. *IEEE Transactions on PAMI*, 23(5):475–489, May 2001.