Research Article

Esophagus Segmentation from 3D CT Data Using Skeleton Prior-Based Graph Cut

Damien Grosgeorge,¹ Caroline Petitjean,¹ Bernard Dubray,² and Su Ruan¹

¹ *Université de Rouen, LITIS EA 4108, 22 Boulevard Gambetta, 76183 Rouen Cedex, France*

² *Centre Henri Becquerel, rue d'Amiens, 76038 Rouen Cedex 1, France*

Correspondence should be addressed to Caroline Petitjean; caroline.petitjean@univ-rouen.fr

Received 14 May 2013; Accepted 30 July 2013

Academic Editor: Liang Li

Copyright © 2013 Damien Grosgeorge et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The segmentation of organs at risk in CT volumes is a prerequisite for radiotherapy treatment planning. In this paper, we focus on esophagus segmentation, a challenging application since the wall of the esophagus, made of muscle tissue, has very low contrast in CT images. We propose in this paper an original method to segment in thoracic CT scans the 3D esophagus using a skeleton-shape model to guide the segmentation. Our method is composed of two steps: a 3D segmentation by graph cut with skeleton prior, followed by a 2D propagation. Our method yields encouraging results over 6 patients.

1. Introduction

Lymphoma is a common tumor in the mediastinum, which is often located close to the trachea and the esophagus. Before radiotherapy, organs at risk such as heart, lung, and esophagus, must be outlined, in order to minimize the quantity of irradiation. In particular, the esophagus is very challenging to delineate because of the really low contrast of its boundaries, its complex shape and its inhomogeneous appearance (Figure 1). Today, the segmentation of the esophagus is performed manually by clinicians and is a tedious task, prone to intra- and interobserver variabilities.

Due to its difficulty, the literature on (semi)automatic esophagus segmentation is quite light. In [1], Rousson et al. proposed a two-step segmentation method in cardiac CT: first, the esophagus centerline is extracted using probabilistic spatial and appearance modeling, and then the outer surface of the esophagus is extracted using multiple and coupled ellipse fittings. However, the method requires as input two points on the centerline of the esophagus and manual segmentation of the aorta and the left atrium. In [2], Ragan et al. segment several organs using deformable models but fail to accurately contour the esophagus. In [3], Huang et al. propose a semiautomated method consisting in manually drawing

one contour in an axial slice, which is propagated to other slices of based registration, but no quantitative evaluation is proposed in this paper. Fieselmann et al. in [4] propose, from several contours in axial slices provided by user input, to interpolate all missing contours in the frequency domain. A fully automatic method is presented in [5] by Feulner et al. that consists in first finding the approximate shape using a “detect and connect” approach, and then a classifier is trained to find short segments of the esophagus which are approximated by an elliptical model.

Most of these methods require a significant amount of user input or are based on an elliptical model. As shown in Figure 1, the esophagus is a deformable organ and has a more complex shape which can hardly be approximated by an elliptic shape. In this work we propose a method to segment the esophagus on thoracic CT scans in two steps: (i) a 3D segmentation by graph cut using a skeleton prior and (ii) a detection of the first slice of inaccurate segmentation, an oversegmentation due to the same gray level of neighboring organs such as aorta, and a 2D segmentation by graph cut for the next slices. Our method requires a very simple user input, has a low computational cost, and gives encouraging preliminary results on 6 patients.

FIGURE 1: CT scan ((a), (b)) and cropped ROI from CT scan ((c), (d)) with esophagus manually segmented in red. Note the variable shape of the esophagus and how its grey levels are similar to surrounding tissues.

2. Method: 3D Graph Cut Segmentation with Shape Prior

In this section, we first outline the graph cut segmentation framework as described in [6]. Then, we introduce our method, which includes the construction of the skeleton model, the 3D segmentation using graph cuts, and the 2D graph cut propagation.

2.1. 3D Graph Cut Segmentation. Let us consider the volume V as a graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, where $\mathcal{V}$ is the set of nodes (voxels) and $\mathcal{E}$ the set of edges. Each pair of nodes $(p, q) \in \mathcal{E}$ in a neighborhood $\mathcal{N}$ is connected by a segment called *n-link* and weighted by $B_{p,q}$, a regularization or boundary term, designed to provide spatial coherence in a neighborhood of voxels. $B_{p,q}$ is typically defined as

$$B_{p,q} \propto \exp\left(-\frac{(V_p - V_q)^2}{2\sigma^2}\right) \cdot \frac{1}{\text{dist}(p, q)}, \quad (1)$$

where V_p and V_q are the gray levels of voxels p and q , $\text{dist}(p, q)$ the Euclidean distance between p and q , and σ a constant usually related to acquisition noise. Consider two additional nodes, called terminal nodes: the source $\mathcal{S}$ representing the object $\mathcal{O}$ (in our case, the RV cavity) and the sink $\mathcal{T}$ representing the background $\mathcal{B}$. Each node $p \in \mathcal{V}$ is

connected to the terminal nodes $\mathcal{S}$ and $\mathcal{T}$ by two respective segments called *t-links* and weighted by the so-called region term denoted by R_p and defined by

$$R_p(\omega) = -\ln \Pr(V_p | \omega), \quad (2)$$

where $\Pr(V_p | \omega)$ is the likelihood of observing V_p given that voxel p belongs to class ω that is intensity distribution of class ω .

A cut $\mathcal{C}$ in the graph consists in cutting *t-links* and *n-links* to attribute a label $\mathcal{O}$ or $\mathcal{B}$ to each voxel p of the image, which boils down to segment the volume. The energy of a cut $\mathcal{C}$ is defined by

$$E(\mathcal{C}) = \sum_{p \in \mathcal{V}} R_p(\omega_p) + \lambda \sum_{p, q \in \mathcal{N}} B_{p,q} \cdot \delta(\omega_p \neq \omega_q), \quad (3)$$

where $\delta(\omega_p \neq \omega_q)$ is 0 if p and q have the same label, 1 otherwise. The 3D optimal segmentation is obtained by searching for the cut of minimal energy. This global search can be very efficiently performed due to mincut-maxflow algorithms, in polynomial time [7].

2.2. Proposed Graph Cut Segmentation Framework Using Shape Prior. We propose a two-step method to segment the esophagus on thoracic CT scans based on graph cuts with a skeleton prior. In Section 2.2.1, we present the construction

of a skeleton model based on a principal component analysis (PCA). Using this model, the first step of the method consists in segmenting the esophagus based on 3D graph cut (Section 2.2.2). We show 3D segmentation overestimate from a certain slice, taking aorta into esophagus segmentation. In this case, the variability of the skeletons is so important that the 3D shape model cannot well guide the segmentation. The second step consists then in detecting this slice, called breaking slice bs in the following, and realizes from it 2D segmentation by graph cut with a skeleton prior for the rest slices (Section 2.2.3).

2.2.1. Construction of the Skeleton Model. Let us consider N 3D esophagus obtained by expert's manual delineation on CT scans. On each slice s , a gravity center c_s^n can be computed $1 \leq n \leq N$ and a skeleton is made of all gravity centers $C^n = \{C_1^n, C_2^n, \dots, C_S^n\}$ to where S is the number of slices. Gravity centers are first decimated to allow skeletons to have the same number of points. Following the definition of a point distribution model (PDM) as first introduced in [8], skeletons are then rigidly aligned using a Procrustes analysis (see Figure 2). A mean skeleton $\bar{\Phi}$ is then computed:

$$\bar{\Phi} = \frac{1}{N} \sum_{n=1}^N C^n. \quad (4)$$

A PCA is then performed on the set of centered skeletons and yields eigenmodes denoted by Φ_i , with $i = 1, \dots, N$, and their associated eigenvalues, denoted by λ_i . The number $k \leq N$ of eigenmodes is retained, with k chosen large enough to account for the most important skeleton variations present in the training set.

Let us now describe how a single skeleton prior is computed from the PCA. Our aim is to isolate areas of variation of the mean skeleton, for each principal axis. We thus generate deformed skeleton instances for each axis:

$$\gamma_i^\alpha = \bar{\Phi} \pm \alpha \sqrt{\lambda_i} \Phi_i, \quad \forall i = 1, \dots, k, \alpha \in [-3 : 3]. \quad (5)$$

α has to be bounded in order to represent variations in the training set of 3D esophagus. Classical boundary values of α in PCA framework are $[-3; +3]$ [9]. A lower value would not capture all variations given by the training set. A higher value would include too large variations and would not ensure smooth deformations. All possible positions of skeletons are summed up in γ_i^α . A prior volume is then obtained by computing the minimum distance between a point p and its nearest skeleton point:

$$P_{3D}(p) = \min_{\alpha, i} \text{dist}(p, \gamma_i^\alpha), \quad (6)$$

where $\text{dist}(p, \gamma_i^\alpha)$ is the euclidean distance between a point p and its nearest gravity center from all possible skeletons. Thus, the lower the distance, the higher the probability to be inside the esophagus.

2.2.2. 3D Segmentation Using Graph Cuts. Now how can this prior map be integrated into the graph cut framework? In

the literature, additional energy terms on the t -links [10] or the n -links [11] are added to the graph cost function. In any case, the prior must be rigidly registered onto the image to be segmented. The user is thus required to point out two landmarks: inside the esophagus, in the first and last slices of the volume. Esophagus intensity values are in the range -100 – 200 in CT scans. The prior map $P_{3D}(p)$ gives the distance of pixel p to a possible skeleton of the esophagus. We thus suggest that the shape prior contributes to weighting t -links. The region term with skeleton prior R_p^S can straightforwardly be defined with

$$R_p^S(\mathcal{O}) = \begin{cases} +\infty & \text{if } V_p < -100 \text{ or } V_p > 200 \\ \frac{(P_{3D}(p))^2}{2\sigma_r^2} & \text{otherwise,} \end{cases} \quad (7)$$

$$R_p^S(\mathcal{B}) = \begin{cases} 0 & \text{if } V_p < -100 \text{ or } V_p > 200 \\ \frac{(P_{3D}(p))^2}{2\sigma_r^2} & \text{otherwise} \end{cases}$$

with V_p being the gray level of voxel p , $P_{3D}(p)$ the distance prior map, and σ_r a parameter defining what value of distance can be considered as being near to the possible skeleton.

We use the classical definition of $B_{p,q}$ to weight the n -links as defined in (1). The final energy of a cut $\mathcal{C}$ for our 3D graph integrating a shape prior is then

$$E(\mathcal{C}) = \lambda \sum_{p,q \in \mathcal{N}} B_{p,q} \cdot \delta(\omega_p \neq \omega_q) + \sum_{p \in V} R_p^S(\omega_p), \quad (8)$$

where λ weights the relative contributions of the n -link and t -link terms.

However, as esophagus is a moving cavity, compressed by the other organs, possible skeletons are various. Moreover, esophagus shares the same gray level with neighboring organs (in particular the aorta). As a result, from a certain slice, an oversegmentation is observed with the 3D graph (see Figure 3(c)). We propose to detect slices with oversegmentation and use 2D propagation by graph cut to improve segmentations.

2.2.3. 2D Propagation by Graph Cut. To detect the slice position where the oversegmentation begins, we use a simple heuristic: esophagus area does not change consequently from one slice to the other. Our aim is to detect the slice where oversegmentation starts by browsing through slices from top to bottom. This slice level is called breaking slice bs and is determined by

$$bs = \min_{s \in S} : \frac{\mathcal{A}_s}{\mathcal{A}_{1, \dots, s-1}} < 1 - \ell \quad \text{or} \quad 1 + \ell < \frac{\mathcal{A}_s}{\mathcal{A}_{1, \dots, s-1}}, \quad (9)$$

where $\mathcal{A}_s$ is the esophagus area of slice s , $\mathcal{A}_{1, \dots, s-1}$ the mean esophagus area of previous slices, and ℓ the threshold of area variation. From bs , we use a 2D segmentation by graph cut with prior for each slice. The final energy of the 2D graph is the same as 3D graph, as defined in (8). The prior is defined by

FIGURE 2: Superimposition of skeletons after alignment.

FIGURE 3: Example of the breaking slice detection (at slice $s = 22$). s is the slice number. (a)–(d) Segmentation with 3D graph and (e)–(g) results of the 2D segmentation. Red: manual contour; green: automatic contour.

the distance to the gravity center of the previously segmented slice $s - 1$:

$$P_{2D_s}(p) = \text{dist}(p, c_{s-1}), \quad (10)$$

where c_{s-1} is the gravity center of the previous segmented slice $s - 1$.

3. Experimental Results and Discussion

The proposed segmentation method has been applied on thoracic CT scans of 6 patients (which have been acquired with different scanners). Each CT volume includes between 73 and 108 slices. Voxel size is $0.98 \times 0.98 \times 2.0$ mm or $0.98 \times 0.98 \times 2.5$ mm.

3.1. Skeleton Model Construction and Method Parameterization. Following a leave-one-out cross-validation strategy

(LOOCV), 6 skeleton models are constructed using a training set of 5 esophagus skeletons. Preliminary registration is performed by superposing the first and the last gravity centers of each skeleton on an arbitrary reference.

Parameters are derived empirically: $\sigma_r = 7$, $\sigma = 40$, $\ell = 0.5$, and $\lambda = 2$ for 3D and 2D graph segmentations. The implementation of Boykov and Kolmogorov of the mincut-maxflow algorithm (available online at <http://pub.ist.ac.at/~vnk/software.html>) is used to compute the cut of minimal cost in the graph [7].

3.2. Segmentation Results. Our segmentation algorithm is run on 6 patients following an LOOCV strategy. For each volume, the user is required to point out two landmarks in order to register the skeleton prior: a point inside the esophagus in the first slice and a point inside the esophagus in the last slice. Our segmentation results are compared to manual ground truth through the Dice metric (DM),

FIGURE 4: Different views of 3D reconstruction of the esophagus. Red: manual contour; green: automatic contour.

TABLE 1: Mean ($\pm$ standard deviation) Dice metric (DM) between automatic and manual delineations of the esophagus and proportion P of segmented slice considering the total number of slices.

	DM	P (%)
3D Seg	0.78 ± 0.03	33.5 ± 5.1
2D Seg	0.48 ± 0.09	55.4 ± 13.8
3D + 2D Seg (total)	0.61 ± 0.06	88.9 ± 11.9

a standard overlap measure for comparing two surfaces. The Dice coefficient is given for 3D results and 2D results after slice bs and combined 3D and 2D results. The Dice score is stopped from being computed when a zero value is found and considered next slices as not segmented. Proportions P of segmented slices considering bs over the total number of slices of esophagus are also computed. Results are provided in Table 1.

Not surprisingly, as shown in Table 1, segmentation results are better for the first slices of esophagus with 3D segmentation, as the esophagus contours are better defined. Figure 3 shows the breaking slice bs detection and the segmentation differences between the 3D and 2D methods. An example of 3D reconstruction of an esophagus is also shown in Figure 4.

The use of 2D segmentation allows avoiding overestimation given by 3D graph method in slices much more difficult to segment and to obtain better results. Considering this challenging application, our method yields encouraging results. However, as shown in Table 1, an average of 88.9% ($\pm 11.9\%$) of esophagus slices are segmented with our method (from top to bottom). Room for improvement is left in the last slices of esophagus. Note that, thanks to the well-known high computational efficiency of graph cut, the computation time of our method is about 15 seconds by patient on a regular PC hardware, a time compatible with clinical practice.

4. Conclusion and Perspectives

In this paper, we have presented a method to segment the esophagus in CT scanner, based on a graph cut approach with incorporation of a shape prior. The shape model is a 1D PDM, constructed via a PCA from a set of skeletons of the esophagus obtained by manual segmentation. This shape model is then integrated into the graph cut cost function as prior term, in order to guide the segmentation. A combined 3D and 2D segmentation method is proposed. Results have been presented over 6 3D CT scans. If results are satisfying for the first slices, room for improvement is left in the remaining ones.

Apart from a validation on a larger database of patients, future works will focus on improving the 3D model of skeleton. We are currently investigating how to use air hole to guide the 2D segmentation process. Other works on 3D curve shape models such as [12] and medial tubular models [13] could also be fruitfully investigated.

References

- [1] M. Rousson, Y. Bai, C. Xu, and F. Sauer, "Probabilistic minimal path for automated esophagus segmentation," in *Medical Imaging: Image Processing*, vol. 6144 of *Proceedings of the SPIE*, pp. 49–51, February 2006.
- [2] D. Ragan, G. Starkschall, T. McNutt, M. Kaus, T. Guerrero, and C. W. Stevens, "Semiautomated four-dimensional computed tomography segmentation using deformable models," *Medical Physics*, vol. 32, no. 7, pp. 2254–2261, 2005.
- [3] T.-C. Huang, G. Zhang, T. Guerrero, G. Starkschall, K.-P. Lin, and K. Forster, "Semi-automated CT segmentation using optic flow and Fourier interpolation techniques," *Computer Methods and Programs in Biomedicine*, vol. 84, no. 2-3, pp. 124–134, 2006.
- [4] A. Fieselmann, S. Lautenschläger, F. Deinzer, M. John, and B. Poppe, "Esophagus segmentation by spatially-constrained shape interpolation," in *Bildverarbeitung für die Medizin*, pp. 247–251, Springer, Berlin, Germany, 2008.

- [5] J. Feulner, S. Zhou, A. Cavallaro, S. Seifert, J. Hornegger, and D. Comaniciu, "Fast automatic segmentation of the esophagus from 3d ct data using a probabilistic model," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '09)*, pp. 255–262, 2009.
- [6] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images," in *Proceedings of the 8th International Conference on Computer Vision*, pp. 105–112, July 2001.
- [7] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.
- [8] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models—their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.
- [9] A. Tsai, A. Yezzi Jr., W. Wells et al., "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 137–154, 2003.
- [10] Z. Song, N. Tustison, B. Avants, and J. Gee, "Adaptive graph cuts with tissue priors for brain MRI segmentation," in *Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 762–765, April 2006.
- [11] D. Freedman and T. Zhang, "Interactive graph cut based segmentation with shape priors," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 755–762, June 2005.
- [12] I. Corouge, S. Gouttard, and G. Gerig, "Towards a shape model of white matter fiber bundles using diffusion tensor MRI," in *Proceedings of the 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano*, pp. 344–347, April 2004.
- [13] R. Saboo, J. Levy, E. Chaney, and S. Pizer, "Medial models of populations of nearly tubular objects," in *Proceedings of the Probabilistic Models for Medical Image Analysis (MICCAI '09)*, 2009.