Centerline Extraction with Principal Curve Tracing to Improve 3D Level Set Esophagus Segmentation in CT Images

Sila Kurugol, Erhan Bas, Deniz Erdogmus, Jennifer G. Dy, Gregory C. Sharp and Dana H. Brooks

Abstract— For radiotherapy planning, contouring of target volume and healthy structures at risk in CT volumes is essential. To automate this process, one of the available segmentation techniques can be used for many thoracic organs except the esophagus, which is very hard to segment due to low contrast. In this work we propose to initialize our previously introduced model based 3D level set esophagus segmentation method with a principal curve tracing (PCT) algorithm, which we adapted to solve the esophagus centerline detection problem. To address challenges due to low intensity contrast, we enhanced the PCT algorithm by learning spatial and intensity priors from a small set of annotated CT volumes. To locate the esophageal wall, the model based 3D level set algorithm including a shape model that represents the variance of esophagus wall around the estimated centerline is utilized. Our results show improvement in esophagus segmentation when initialized by PCT compared to our previous work, where an ad hoc centerline initialization was performed. Unlike previous approaches, this work does not need a very large set of annotated training images and has similar performance.

Index Terms— Curve Tracing, Level Sets, CT, 3D Image Segmentation, Spatial, Shape Model, Radiation Oncology

I. INTRODUCTION

Curvilinear objects are common in biomedical images, e.g., bronchial tree, vessels, neuronal arbors. The methods proposed to segment and trace such structures can be classified into two groups *i*) Global methods [1], [2] that locate the centerline of tubular structures by optimizing a global objective and *ii*) local methods [3], [4] that make local decisions during the trajectory estimation using local evidence. Global methods, are generally more robust to noise and outliers, provided that there is sufficient training data, but are typically less flexible in adaptation to internal variation in the object parameters along its length and also more computationally burdensome. For problems where outliers play a significant role but global methods cannot be applied (for example due to a paucity of training data), a method which fuses the two approaches can be attractive.

Segmentation of the esophagus from 3D CT falls into this category because of absence of both consistent intensity contrast and reliable discriminative features between the esophagus and surrounding mediastinum tissue in thoracic CT scans. Segmentation of the tumor and nearby structures including the esophagus is of critical importance in path planning for radiotherapy to avoid inadvertent damage when

irradiating tumors. However, esophagus is very difficult to locate compared to other thoracic structures. Perhaps due to these difficulties, previous studies are limited. Rousson et al. [5] located the esophagus centerline with a minimal path approach based on locations of left atrium and descending aorta. They segmented the esophageal wall in a limited range of the cranio-caudal axis (around the left atrium) by fitting a 2D ellipse model to each slice using an appearance based cost function with a slice-to-slice smoothness term. Feulner et al. [6] classified candidate 2D ellipses in each slice as being esophagus or not. They combined these decisions and forced smooth slice-to-slice parameter transitions with a Hidden Markov Model. However, their method requires a large amount of training data for correct classification, in part due to variable appearance in the presence of air bubbles, contrast agent, or both.

Recently we presented a model based 3D level set esophagus segmentation algorithm [7] over the entire thoracic range employing a shape model, with a global and a locally deformable component. This model requires initial centerline estimation and we used an ad hoc centerline estimator where the centerline estimation was only performed at the locations of some predefined anatomical landmarks followed by interpolation for the remaining slices. In this work, we extend our previous algorithm by replacing this ad hoc centerline estimation with a more theoretically grounded principal curve tracing (PCT) algorithm adapted from [4]. We extend that work on PCT, addressing problems related to this local tracing method in low contrast regions with the use of prior spatial and appearance models estimated from the training set. We initialize our 3D level set segmentation algorithm from this PCT centerline estimate. We report below on the consequent improvements in the segmentation results compared to our previous work [7].

II. METHODS

We first describe how we register the annotated training CT data sets to learn prior models from them and how we use them to guide the PCT algorithm for centerline detection. We then summarize how the PCT result is used to initialize the level set algorithm for 3D segmentation. Finally we explain the additional prior models learned from the training set and how they are incorporated into the cost function formulated in a level set optimization framework for 3D surface segmentation.

Landmark based registration: Before we build prior models from training sets, for which manual annotations of the esophagi and neighboring structures are available, we first register manually annotated structures in the training set to a common reference set. To do so, we used a simple anatomical landmark based registration algorithm. We chose 7 easily located anatomical landmark points along

the z (cranial-caudal) axis, located at the following positions (superior to inferior): top of the lungs, thoracic vertebrae 2 and 3, bifurcation of trachea, top of heart, thoracic vertebra 8 (and left atrium), and right ventricle. We first matched those landmark slices of the training and the reference data sets. We then interpolated the contours for slices in between landmarks.

A. Learning step for centerline detection

For centerline estimation, we build two prior models: 1) a spatial model of the esophagus center location with respect to the neighboring anatomical structures and 2) an appearance model of the esophagus.

Spatial Model: We learned a model of relative spatial location of anterior-posterior (x) and left-right (y) coordinates of the esophagus center with respect to neighboring structures for each slice from training data. We assumed the segmentations of those neighboring structures are available, either from manual segmentation or a prior use of an automated algorithm. Since these structures are much easier to segment than the esophagus and are already commonly segmented in clinical practice this is not an onerous assumption. These neighboring structures are the vertebra, descending aorta (DA), trachea (or left main bronchi (lmb)) and heart. For the x direction, from landmark slices 1 to 5 the trachea (or lmb) and vertebra are used, from landmark slice 5 to 7 heart and vertebra were used. For the y direction, from landmark slice 1 to 5 only the vertebra was used and from the landmark 5 to 7 the vertebra and DA were used. The spatial model for x and y direction were built for all slices. For each slice, the normalized distance between the x-coordinate of the esophagus center to the first neighboring structure (d_x) and to the second ($1 - d_x$) were calculated over all data sets and similarly for the y coordinates to get d_y's. (See Fig. 5 (a)). We applied kernel density estimation (KDE) to histograms of d_x and d_y over all data sets to estimate their probability density functions (pdf). From these pdfs, for a given voxel in a test set, the x and y components of the spatial centerline probability ($\rho_x(\mathbf{x})$) and ($\rho_y(\mathbf{x})$) were estimated.

Appearance Model: We learn an appearance model from the training set as intensity pdfs, one inside ($p_{in}(I(\mathbf{x}))$) and one outside ($p_o(I(\mathbf{x}))$) the esophagus, estimated using kernel density estimation from the intensity histogram.

The esophagus appearance changes in the presence of air bubbles, which are very dark compared to the esophagus and oral contrast agents, which are very bright. These regions inside the esophagus appear unpredictably in any given scan and so are detected and removed before building the appearance model and before processing in the test case, if they are present. In [6], thresholding was used to locate esophageal air. In our data sets, due to CT artifacts around the boundaries of air bubbles/oral contrast agent which create an artificial intensity range, a single hard threshold did not succeed. Instead, we used a looser threshold resulting in falsely detected air/contrast regions which we then reject in a second step using a classifier trained on simple features such as area and location with respect to neighboring structures. Once detected, we took advantage of the presence of air/contrast regions to locate the centerline and the esophagus, especially in regions lacking contrast. Therefore we incorporated them in the spatial model calculated for a voxel in a test set for centerline estimation by setting the

value of that voxel to the maximum spatial probability value over the entire region ($\max\{\rho_x(\mathbf{x}_j)\rho_y(\mathbf{x}_j)\}$).

B. Probabilistic Centerline Estimation with Principal Curve Tracing Algorithm

PCT is a non-parametric method based on the concept of subspace local maxima. Mathematically, a principal curve of a twice continuously differentiable function, $p(\mathbf{x})$ (obtained from the data samples $\mathbf{x}_i \in \mathbb{R}^3$, $i = 1 \ldots N$) is a set of points which has the property that the local gradient, $\mathbf{g}(\mathbf{x})$, is aligned with the eigenvector corresponding to the smallest absolute eigenvalue of the local Hessian matrix, $\mathbf{H}(\mathbf{x})$, and all the eigenvalues of the corresponding remaining eigenvectors are negative [8]. In order to obtain such points on the curve, first, we defined tangential space as the span of the eigenvector having the smallest absolute eigenvalue. Similarly, remaining eigenvectors are selected as the basis for the normal space. Let $\mathbf{H}_{\|}(\mathbf{x})$ and $\mathbf{H}_{\perp}(\mathbf{x})$ be the tangential and normal components of the local Hessian matrix respectively, such that $\mathbf{H}(\mathbf{x}) = \mathbf{H}_{\|}(\mathbf{x}) + \mathbf{H}_{\perp}(\mathbf{x}) = \lambda_1 \mathbf{q}_1(\mathbf{x})\mathbf{q}_1(\mathbf{x})^T + \sum_{i=2}^{3} \lambda_i \mathbf{q}_i(\mathbf{x})\mathbf{q}_i(\mathbf{x})^T$ where λ_i and $\mathbf{q}_i$ are the ith eigenvalue and eigenvector pairs of $\mathbf{H}(\mathbf{x})$ and $|\lambda_1| < |\lambda_2| < |\lambda_3|$. A measure being on the curve can be given as

$$\zeta(\mathbf{x}) = \text{abs}\left(\frac{\mathbf{g}(\mathbf{x})^T H_{\perp}(\mathbf{x})\mathbf{g}(\mathbf{x})}{\|H(\mathbf{x})\mathbf{g}(\mathbf{x})\|\|\mathbf{g}(\mathbf{x})\|}\right) \qquad (1)$$

such that $\zeta(\mathbf{x})$ vanishes on the principal curve, since the inner product of $\mathbf{g}(\mathbf{x})$ and eigenvectors of $\mathbf{H}_{\perp}(\mathbf{x})$ are zero.

Here, one can use this measure to project all samples to their corresponding principal curves as a dimensionality reduction technique [8] by solving a differential equation at every iteration, however given the size of the data such an approach is not feasible in our task. Moreover, different neighboring tissues have different intrinsic dimensionality and can not be modeled as a curvilinear structure. For that reason we start from a seed location on the center of the esophagus and iteratively trace the ridge of a function estimated from the data having high values at the esophagus center. We used a weighted kernel density estimation technique to obtain a pdf, after replacing the detected air/contrast regions in the original intensities, $I(\mathbf{x}) \in \mathbb{R}^+$, by fitting an intensity surface to the local neighborhood using cubic spline interpolation. We further enhanced the PCT algorithm by incorporating prior appearance and spatial models learned from training data while estimating the feature pdf.

Our iterative tracing algorithm consists of update and correction steps, where in the update step we trace the underlying principal curve of the feature pdf along the tangential direction using fixed length updates. In order to obtain continuous trace we correct the direction of the tracing update such that it will have positive inner product with the previous iteration. Since each tangential update deviates the trajectory of the trace from the underlying curve, we use the correction step to project back to the curve. In the correction step, we use the projection of the gradient on the normal subspace to climb up to the subspace local maxima where the underlying principal curve lies. In this scheme, iterations in the normal plane of the principal curve which carry out correction steps are alternated with update iterations in the direction of the tangential vector.

In our calculations, we restricted the density estimate calculations to a finite ϵ-ball support around a sample point

$(B_\epsilon(\mathbf{x} - \mathbf{x}_i))$ and employed Gaussian kernels for both spatial (G_{Σ_i}) and appearance (G_δ) domain. KDE of feature pdf $\mathbf{p}(\mathbf{x})$ is given as

$$\mathbf{p}(\mathbf{x}) = \sum_{\mathbf{x}_i \in B_\epsilon(\mathbf{x})} \bar{w}_i G_{\Sigma_{\mathbf{x}_i}}(\mathbf{x} - \mathbf{x}_i) \qquad (2)$$

Here, N is the number of samples, $\mathbf{x}_i$ is the position of a voxel in the neighborhood, α is the normalization constant of the kernel, w_i is the weight of the ith kernel, and $\bar{w}_i = \alpha w_i G_{\delta_i}(I(\mathbf{x}) - I(\mathbf{x}_i))$ is the overall effective weight. We estimated Σ_i from the mean shape learned from the training data, whereas δ is selected experimentally. KDE weights were determined as $w(\mathbf{x}_i) = p_{in}(I(\mathbf{x}_i))\rho(\mathbf{x}_i)$ from the appearance model $p_{in}(I(\mathbf{x}_i))$ and the modified spatial model $\rho(\mathbf{x}_i)$ built during learning step. Letting $\beta_i(\mathbf{x}) = \bar{w}_i G_{\Sigma_{\mathbf{x}_i}}(\mathbf{x} - \mathbf{x}_i)$, the gradient and Hessian are

$$\mathbf{g}(\mathbf{x}) = - \sum_{\mathbf{x}_i \in B_\epsilon(\mathbf{x})} \beta_i(\mathbf{x})\Sigma_{\mathbf{x}_i}^{-1}(\mathbf{x} - \mathbf{x}_i) \qquad (3)$$

$$\mathbf{H}(\mathbf{x}) = \sum_{\mathbf{x}_i \in B_\epsilon(\mathbf{x})} \beta_i(\mathbf{x})(\Sigma_{\mathbf{x}_i}^{-1}(\mathbf{x} - \mathbf{x}_i)(\mathbf{x} - \mathbf{x}_i)^T \Sigma_{\mathbf{x}_i}^{-1} - \Sigma_{\mathbf{x}_i}^{-1}) \qquad (4)$$

We used the most likely point according to the spatial model as the initial center seed to start PCT at the first slice. Note that the correction iterations, as well as the tracing updates might result in locations which are not limited to voxel grid centers. This enables us to obtain subpixel accuracy during tracing. However, in order to calculate the intensity differences between the current iteration and its neighborhood at subpixels, we used nearest neighbor interpolation. We recorded the location of the correction step as the trace location $\mathbf{P}_t$ at the tth iteration.

C. Locating The Esophagus Outer Boundary Surface

After centerline estimation with PCT, we initialized of the shape model with the estimated centerline and used our 3D level set algorithm with a locally deformable shape model [7] to find the esophagus wall. For the level set energy function (E) in Eqn. 5, we used standard energy terms [9], [10] including an appearance (E_{app}), a level set regularization (E_{reg}) and a smoothing (curvature) term (E_{sm}) and some problem specific terms including an air/contrast (E_{air}), a neighboring structures term (E_{nb}) and a shape fitting term (E_{shape}) which we will explain next.

$$E = E_{app} + E_{air} + E_{nb} + E_{shape} + E_{reg} + E_{sm} \qquad (5)$$

Shape Model: To model the complex tubular esophagus shape, we used a shape model that has both global and local components. After landmark based registration of the annotated esophagi, centerline is subtracted from each esophagus which allowed modeling of variations around centerline only. To model these variations, the global shape component (ψ) was constructed from the mean shape and the principal component analysis (PCA) modes (eigen-shapes U_i), assuming they are gaussian distributed.

$$\psi = \psi_m + \sum_{i=1}^{k} c_i U_i \qquad (6)$$

where c_i's are the weight of the modes and parameters to be optimized for an esophagus shape in a test data. We also learned a prior on c_i's. To do so, we calculated the histogram of each c_i over the training shapes and constructed the prior assuming a uniform density.

For the local component of the shape model, although we locally estimated the centerline with PCT, we also included the nonlinear local transformation into the shape model to correct for the inaccuracies in center estimation. This model

follows the locally affine transformation model in [11]. To construct this model, N uniformly sampled *action points* z_k through the centerline were chosen. A local transformation, in the form of a translation in x-y plane, to be applied to each z_k was constructed. This translation affects the neighboring slices and this effect smoothly dies off as one moves away from the action points in z-direction. Such a local deformation $\mathbf{A}$ was formally defined in [7].

Next, we define the shape energy (E_{shape}) that drives the level set function ϕ to be similar to our shape model $\psi(\mathbf{A})$. Here δ_ϵ is the dirac delta function.

$$E_{shape}(\phi, c) = \int_\Omega \delta_\epsilon(\phi)(\phi - \psi(\mathbf{A}))^2 d\Omega \qquad (7)$$

The energy term for appearance (E_{app}) [9] uses the appearance model learned from the training data. Here $p_{in}(I)$ and $p_o(I))$ are the pdfs inside and outside esophagus and H_ϵ is the Heaviside function of the level set ϕ.

$$E_{app}(p_i, p_o) = -\int_\Omega [H_\epsilon(\phi) \log(p_i(I))d\Omega \\ - \int_\Omega (1 - H_\epsilon(\phi)) \log(p_o(I))]d\Omega \qquad (8)$$

To make use of the presence of air and contrast regions we incorporate them into the level set framework with an additional air/contrast term (E_{air}). Here p_{air} is a pdf indicating the probability of a voxel being inside esophagus which is close to 1 if the pixel is air/contrast and 0.5 otherwise.

$$E_{air}(p_{air}) = -\int_\Omega [H_\epsilon(\phi) \log(p_{air}(I))d\Omega \\ - \int_\Omega (1 - H_\epsilon(\phi)) \log(p_{air}(I))]d\Omega \qquad (9)$$

We used a similar energy term $(E_{nb}$, same form as $E_{air})$ to exclude neighboring structures from the segmented esophagus. This term includes a probability function that takes low values (~ 0) for the neighboring structure voxels and 0.5 otherwise.

After including level set regularization E_{reg} [10] and smoothness E_{sm} terms [9], the energy functional in Eqn. 5 was obtained. We minimized this function with respect to ϕ and shape parameters to locate the esophagus boundary.

For a test set, after landmark based registration, centerline was estimated by the PCT algorithm and the data was centered around this estimated centerline. Then 3D segmentation algorithm was initialized. We initialized the shape prior as the mean shape. The initial level set function representing the esophagus boundary and the shape prior level set function were updated at each step t by minimizing E. The equation of evolution for ϕ is given by calculus of variations; the optimization of E with respect to mode weights c_i were obtained by solving a linear system [9]. Adding the weight priors results in a constrained least square minimization that was solved by convex optimization [12]. Finally, the minimization of E with respect to local deformation parameters was carried out using calculus of variations. The update equations were explained in [9], [7].

III. RESULTS

We report experiments using thoracic CT scans (resolution $0.98 \times 0.98 \times 3.75\text{mm}^3$) from 8 subjects. We tested our method with a leave-one-out scheme for training and testing. The input to the algorithm is the designation of anatomical landmarks and the segmentation of neighboring structures, which can be obtained using available algorithms [13], [14], [15], [16]. We report the results in comparison to the manually segmented esophagi in all data sets.

We first report the results of the centerline estimation algorithm. The PCT centerline estimation resulted in a mean error of 1.40 ± 0.55 mm in x and 2.44 ± 0.64 mm in y direction over all experiments. Fig. 1 shows a comparison of these results to our previous work [7], which achieved average error of 1.9 mm in the x-direction and 4.1 mm in the y-direction. Fig. 3 illustrates the results of both centerline estimation algorithms for two data sets (numbers 6 and 7). The true centerline is also shown for comparison. The errors in y direction are larger due to the presence of a larger low contrast neighborhood in that direction.

We use the following point-wise distance metric to evaluate the results of the esophagus boundary surface segmentation algorithm. The distance between the points on both contours at the same angle from the x axis were calculated for each slice. Fig. 2 shows point-wise mean distance errors of each data set. Results improved from a point-wise mean error of 2.6 ± 2.1 mm and maximum error of 17.6 mm to 2.1 ± 1.9 mm and maximum error of 15.1 mm. The results for 3 sample axial slices are shown in Fig.5 (b) and a 3D rendering in Fig. 4 for a sample data set.
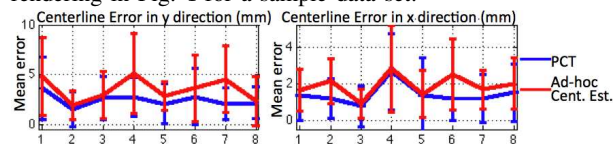

Fig. 1. Mean $\pm$ std error of the centerline estimation algorithms in x (left) and y (right) for each data set. Blue curves represent the PCT results whereas red curves show earlier interpolation results.
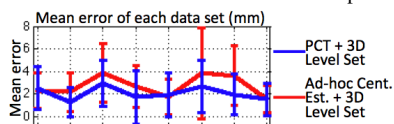

Fig. 2. Mean $\pm$ std error of final segmentations of each data set.
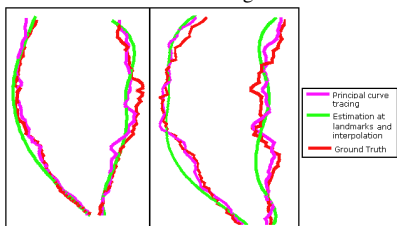

Fig. 3. Results of the centerline estimation algorithm (Red ground truth, magenta PCT, green landmark interpolation) for two different views (left and right) and two data sets (data 6-left, 7-right box).

IV. CONCLUSIONS AND FUTURE WORK

We introduced a PCT algorithm for esophagus centerline estimation that works on the local pdf and incorporates prior models learned from training data to improve performance, especially when intensity contrast is absent. The estimated centerline based on PCT is used to initialize the model based 3D level set segmentation algorithm. This algorithm takes advantage of prior models including appearance and shape models and the presence of air bubbles and contrast agents. Since the algorithm works in 3D, in contrast to existing 2D algorithms with additional slice to slice smoothness constraints, this algorithm directly achieves a smooth segmentation result.

We are currently acquiring a larger data set to fully test the method. We eventually expect to be able to eliminate the required user input by automating the landmark selection

and neighboring structure segmentation processes using the algorithms in [13], [14], [15], [16]. The robustness of our algorithm when manual annotation is replaced by automated segmentation will be evaluated. However, due to smooth appearance of spatial model, the algorithm will compensate for the segmentations errors of a few mm.
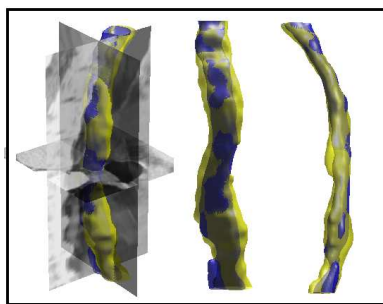

Fig. 4. Segmented esophagi in 3D (yellow-ground truth, blue-algorithm) for three different views, where on the left view the
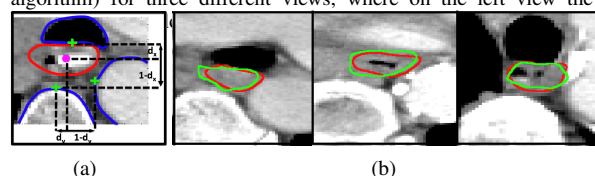

Fig. 5. (a) Spatial Model illustrated on a sample slice. Esophagus center location (pink dot) is learned with respect to neighboring structure locations (green pluses). (b) Axial slices from 3 data sets showing the result of the algorithm (green) and expert (red).

REFERENCES

[1] H. Li and A. Yezzi, "Vessels as 4-D curves: Global minimal 4-D paths to extract 3-D tubular surfaces and centerlines," *IEEE Trans. Med. Imag.*, vol. 26, no. 9, pp. 1213–1223, 2007.

[2] W.C.K. Wong and A.C.S. Chung, "Principal curves to extract vessels in 3D angiograms," in *IEEE Workshop*, 2008.

[3] T. Deschamps and L.D. Cohen, "Fast extraction of minimal paths in 3d images and applications to virtual endoscopy," *Medical Image Analysis*, vol. 5, no. 4, pp. 281–299, 2001.

[4] E. Bas and D. Erdogmus, "Principle curve tracing," in *Europ. Symp. on Artificial Neural Networks*, 2010.

[5] M. Rousson et al., "Probabilistic minimal path for automated esophagus segmentation," in *SPIE*, 2006.

[6] Feulner et. al., "Model-Based Esophagus Segmentation from CT Scans Using a Spatial Probability Map," *MICCAI*, 2010.

[7] S. Kurugol, N. Ozay, J.G. Dy, G.C. Sharp, and D.H. Brooks, "Locally Deformable Shape Model to Improve 3D Level Set based Esophagus Segmentation," *ICPR*, 2010.

[8] U. Ozertem, *Locally Defined Principal Curves and Surfaces*, Oregon Health & Science University, 2008.

[9] M. Rousson et al., "Implicit active shape models for 3D segmentation in MR imaging," in *MICCAI*, 2004.

[10] C. Li et al., "Level set evolution without re-initialization: A new variational formulation," in *CVPR*, 2005.

[11] R. Narayanan et al., "Diffeomorphic nonlinear transformations: A local parametric approach for image registration," in *IPMI*, 2005.

[12] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," June 2009.

[13] H. Lombaert et al., "A multilevel banded graph cuts method for fast image segmentation," in *ICCV*, 2005.

[14] Y. et. al. Kang, "A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data," *IEEE Trans. Med. Imag.*, vol. 22, no. 5, pp. 586–598, 2003.

[15] U. Kurkure et al., "Automated segmentation of thoracic aorta in non-contrast CT images," in *ISBI*, 2008.

[16] O. Ecabert et al., "Automatic model-based seg. of the heart in CT images," *IEEE Trans. Med. Imag.*,, vol. 27, no. 9, 2008.