Computer-aided lymph node segmentation in volumetric CT data

Reinhard R. Beiche^{a)}

*Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, Iowa 52242;
Department of Internal Medicine, The University of Iowa, Iowa City, Iowa 52242; and The Iowa Institute for
Biomedical Imaging, The University of Iowa, Iowa City, Iowa 52242*

Yao Wang

Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, Iowa 52242

(Received 3 March 2012; revised 5 July 2012; accepted for publication 18 July 2012; published 16 August 2012)

Purpose: The purpose of this work was to develop and validate a computer-aided method for the 3D segmentation of lymph nodes in CT images. The proposed method can be utilized to facilitate applications like biopsy planning, image guided radiation treatment, or assessment of response to therapy.

Methods: An optimal surface finding based lymph node segmentation method was developed. Based on the approximate center point of a lymph node of interest, a graph is generated, which represents the local neighborhood around the lymph node at discrete locations (graph nodes). A cost function is calculated based on a weighted edge and region homogeneity term. By means of optimization, a surface-based segmentation of the lymph node is derived. In addition, an interactive segmentation refinement algorithm was developed, which allows the user to quickly correct segmentation errors, if needed. For assessment of segmentation accuracy, 111 lymph nodes of mediastinum, abdomen, head/neck, and axillary regions from 35 volumetric CT scans were utilized. For accuracy analysis, lymph nodes were divided into three test sets based on lymph node size and spatial resolution of the CT scan. The average lymph node size for test set I, II, and III was 1056, 1621, and 501 mm³, respectively. Spatial resolution of test set II was lower than for test sets I and III. To generate an independent reference standard for comparison, all 111 lymph nodes were segmented by an expert with a live wire approach.

Results: All test sets were segmented with the proposed approach. Out of the 111 lymph nodes, 40 cases (36%) required computer-aided refinement of initial segmentation results. The refinement typically required 10 s per lymph node. The mean and standard deviation of the Dice coefficient for final segmentations was 0.847 ± 0.061 , 0.836 ± 0.058 , and 0.809 ± 0.070 for test sets I, II, and II, respectively. The average signed surface distance error was 0.023 ± 0.171 , 0.394 ± 0.189 , and 0.001 ± 0.146 mm for test sets I, II, and II, respectively. The time required for locating the approximate center point of a target lymph node in a scan, generating an initial OSF segmentation, and refining the segmentation, if needed, is typically less than one minute.

Conclusions: Segmentation of lymph nodes in volumetric CT images is a challenging task due to partial volume effects, nearby strong edges, neighboring structures with similar intensity profiles and potentially inhomogeneous density of lymph nodes. The presented approach addresses many of these obstacles. In the majority of cases investigated, the initial segmentation method delivered results that did not require further processing. In addition, the computer-aided segmentation refinement framework was found to be effective in dealing with potentially occurring segmentation errors.

© 2012 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4742845>]

Key words: lymph node segmentation, optimal surface finding

I. INTRODUCTION

The lymphatic system of the human body is a component of the immune system that plays an important role in dealing with viruses, bacteria, and other illnesses. In medicine, the term lymphadenomegaly is used to describe an abnormal lymph node that is not palpable like in the mediastinum or abdomen. It is well known that critical causes of lymphadenomegaly include cancer, sarcoidosis, or fungal infections such as histoplasmosis.

Multidetector computed tomography (MDCT) has become the primary lymph node imaging modality in clinical rou-

tine and offers good spatial resolution for measuring lymph nodes.¹ Assessment of lymph nodes is required for monitoring and treatment of diseases that directly or indirectly affect the lymphatic system (e.g., cancer). In current clinical practice, lymph nodes imaged with MDCT are analyzed manually based on measures of long and/or short axis length to assess response to therapy in follow-up examinations. Also, the revised response evaluation criteria in solid tumors (RECIST) 1.1 now explicitly include the assessment of lymph nodes.² The accuracy and reproducibility of size measurements is critical for determining response in clinical practice and informed research studies.¹ However, uni- or

bidimensional (length) measurement were found to be subject to inter- and intraobserver variability.^{3,4} A study by Fabel *et al.*⁵ showed that volumetric analysis of lymph nodes by means of semiautomatic segmentation promises to significantly reduce variability and interobserver bias compared to RECIST. In addition, other applications like biopsy planning or image guided radiation treatment (IGRT) require accurate geometric models (i.e., segmentations) of lymph nodes.

Lymph node segmentation in volumetric CT data is a challenging task due to low contrast to adjacent structures, inhomogeneous density values (e.g., calcification), and high variation in size and shape. Cordes *et al.*⁶ utilized a manual segmentation approach for neck lymph node segmentation to aid planning of neck dissections. A fast marching approach for semiautomatic segmentation of lymph nodes in 2D CT data has been reported by Yan *et al.*⁷ The authors applied their algorithm on a slice-by-slice basis for the 3D segmentation of lymph nodes in volumetric image data. No quantitative evaluation results were provided. Dornheim *et al.*⁸ presented a 3D mass-spring model for the segmentation of neck lymph nodes in CT data. The volumetric segmentation error ranged between 39% and 52%.⁸ The approach presented by Maleike *et al.*⁹ utilized a combination of a deformable 3D surface and statistical shape model for lymph node segmentation. In addition, a tool for manual intervention was provided to help the algorithm in converging to the desired object contours. In 6.9% out of 29 cases, the authors reported problems in getting a “usable segmentation”.⁹ Lu *et al.*¹⁰ utilized a semiautomatic 2D/3D live-wire-based approach for the segmentation of central chest lymph nodes in MDCT images. Unal *et al.*¹¹ proposed a semiautomatic method for lymph node segmentation in nanoparticle-enhanced lymphotropic magnetic resonance (MR) images in the context of prostate cancer screening studies. The algorithm is based on coupled ellipse flows and implements a joined segmentation and registration approach. It first propagates in 2D on corresponding MR sequences and extends then to other slices in 3D. In conclusion, the range of existing lymph node segmentation methods for CT images is quite diverse. It is generally accepted that a manual segmentation is too time consuming for routine application. 2D slice-by-slice approaches (e.g., Yan *et al.*⁷) tend to introduce inconsistencies between slices. Some of the 3D approaches show a rather large segmentation error (Dornheim *et al.*⁸), mainly depend on user interaction (Lu *et al.*¹⁰), or have a manual refinement option that does not allow the user to produce desired results in some cases (Maleike *et al.*⁹).

In this paper, we propose a new computer-aided method for 3D lymph node segmentation that consists of a two-stage approach. First, the user identifies the lymph node to be analyzed by specifying the approximate center of the lymph node, and an automated segmentation method is applied to generate a segmentation. This segmentation approach is based on preliminary work published by Wang *et al.*¹² Second, the user inspects the segmentation result, and if needed, can use a method for user-steered and computer-aided refinement of incorrect segmentation results. Both the initial segmentation and refinement method are based on an optimal surface finding (OSF) approach. This has the advantage that segmenta-

FIG. 1. Comparison of lymph node segmentation methods. (a) Result of an edge-based segmentation approach. Arrows indicate major segmentation errors. (b) Result generated with the proposed method. Major segmentation errors due to nearby strong edges are avoided.

tion refinement is efficient, because the user directly interacts with a segmentation algorithm during refinement, compared to manually editing the result. In the validation study presented in this paper, the proposed method was applied to 111 lymph nodes of mediastinum, abdomen, head/neck, and axillary regions, and the segmentation error was assessed in addition to the required refinement effort.

II. METHODS

For lymph node segmentation in volumetric CT data, an OSF approach is utilized. OSF was introduced by Li *et al.*^{13,14} The basic idea behind OSF is to convert a segmentation problem into a graph optimization problem. For this purpose, a graph structure and a cost function need to be defined, which represent the segmentation problem. Based on this information, the OSF algorithm generates a surface that represents the globally optimal solution according to the defined cost function. For more information on OSF-based segmentation, the reader is referred to Li *et al.*¹⁴

The proposed approach avoids common problems in lymph node segmentation by utilizing a cost function that consists of a weighted edge and a region homogeneity term in combination with a surface smoothness constraint (Fig. 1). In addition, the OSF method will also be utilized for refining the initial segmentation result, if needed.

Our method consists of four main processing steps. First, the user is required to provide the approximate center point (voxel) $\mathbf{c}_k$ of lymph node k to be segmented. Second, a directed spherical graph, whose center is located at $\mathbf{c}_k$, is constructed, so that the local region around the lymph node is transformed into a graph-based representation. Then the segmentation problem is formulated as an optimization (graph search) problem.¹⁴ We define a cost function that enables us to segment lymph nodes and to avoid common shortcomings of existing approaches. Third, the optimization problem is solved to produce a segmentation. Fourth, the user inspects the segmentation result and can correct the segmentation, if needed. In Secs. II.A–II.E, we describe our segmentation approach in detail.

II.A. Preprocessing

Lymph nodes have approximately CT densities in a range between -100 and 150 Hounsfield units (HU). Thus, the CT density-values were truncated to this HU range. In addition,

FIG. 2. 2D illustration of the graph generation process required for OSF-based segmentation. For this case, $\Delta = 1$ was used for graph construction.

all CT scans were resampled to isotropic voxel size before processing, and a $3 \times 3 \times 3$ voxel median filter was applied to reduce image noise.

II.B. Graph construction

A node-weighted directed graph $G = (V, E)$ is generated, with node set V and edge set E , representing a spherical volume of interest $VOI(\mathbf{c}_k)$ around the approximate center point $\mathbf{c}_k$ of lymph node k . In order to accomplish this task, a sphere-shaped triangular mesh (Fig. 2) is built around $\mathbf{c}_k$ with radius r . The radius r is a constant and chosen to be larger than the largest expected radius of lymph nodes. Let n_v be the number of vertices of the spherical mesh. For each mesh vertex $\mathbf{p}_i$ with $i \in \{1, \dots, n_v\}$, the image volume is sampled along the line between center point $\mathbf{c}_k$ and vertex $\mathbf{p}_i$ in an equidistant fashion using linear interpolation. The line between center point $\mathbf{c}_k$ and vertex $\mathbf{p}_i$ is denoted as column i .

The discrete sample points define nodes in the graph structure that will be utilized for segmentation. In this context, note that $\mathbf{p}_i$ represents already a sample point/node, whereas $\mathbf{c}_k$ is not utilized. The gray-value samples on the line between $\mathbf{c}_k$ and $\mathbf{p}_i$ form the elements $g_i(j)$ of column i with $j \in [0, 1, \dots, (n_e - 1)]$. The number of elements per column is a constant and denoted as n_e , and $g_i(n_e - 1)$ represents the gray-value density sample at the location of vertex $\mathbf{p}_i$. The node $v_i(j) \in V$ corresponds to gray-value sample point $g_i(j)$.

Because the goal of the algorithm is to find the minimum cost surface in the graph, elements that belong to the lymph node boundary should be assigned low cost values. The calculation of $c_i(j)$ is described in detail in Sec. II.C. Once the node-weighted directed graph is constructed, it is converted to an edge-weighted directed graph, as described by Li *et al.*¹⁴

Overall, there are $n_e \times n_v$ nodes in the graph. Considerations regarding the selection of sampling parameters n_v and n_e are as follows. Essentially, both numbers need to be adapted to the resolution of the CT image data. If n_v is selected too small relative to image voxel size, the mesh will be too sparse and unable to represent the lymph node's surface accurately. Conversely, if n_v is chosen too large, the computation time will increase with no major benefits regarding segmentation accuracy. Similar considerations apply for the selection of n_e , the number of sample points along columns.

The neighborhood relation between columns is defined by the mesh structure. If (p, q) is an edge of the triangular mesh, then column p and column q are adjacent. A surface smoothness constraint Δ between any two adjacent columns is utilized to specify the maximum allowable change/difference in surface node locations.¹⁴ Figure 2 depicts a 2D example of the graph generation. By constructing the graph G as described above, spherical, elliptical, or slightly kidney-shaped lymph nodes can be segmented. In this context, the selection of an adequate smoothness constraint Δ is of importance for our method. If the smoothness constraint is too small, the surface will not be able to follow the lymph node surface in some cases (e.g., elongated lymph nodes). On the other hand, if the smoothness constraint is too large, the resulting surface can be noisy.

II.C. Cost function

For segmentation, the following cost function

$$c_i(j) = [1 - w_{\text{shape}}(j)c_{\text{edge}_i}(j)] + \alpha c_{\text{gh}_i}(j) \quad (1)$$

(a)

FIG. 3. Gray-value information utilized for calculating the cost function. (a) Cross-sectional CT image showing a lymph node. The user-selected center $\mathbf{c}_k$ is indicated by a cross. The line indicates one of the columns utilized to build the graph structure shown in Fig. 2. (b) Gray-value profile corresponding to the column marked in (a).

FIG. 4. Gray-value homogeneity function for the gray-value profile shown in Fig. 3(b).

is utilized, which defines costs $c_i(j)$ for each node $v_i(j) \in V$, where $c_{\text{edge}_i}(j)$ represents an edge term and $c_{\text{gh}_i}(j)$ a gray-value-based homogeneity cost term. $w_{\text{shape}}(j)$ is a global shape weight, which is derived from all the n_v columns of the graph structure. It captures the overall edge distribution around the center $\mathbf{c}_k$. The constant α is used to adjust the influence of the homogeneity term relative to the edge term. Note that the cost function is formulated such that low costs are assigned to nodes which are assumed to belong to the lymph node surface. In the following paragraphs, we will give a detailed explanation of all the components in the cost function.

A cross-sectional CT image of a lymph node is shown in Fig. 3(a). The line in Fig. 3(a) represents a column of the graph. The corresponding gray-value sample points $g_i(j)$ are depicted in Fig. 3(b). The gray-value homogeneity term $c_{\text{gh}_i}(j) = \max_{a=0,1,\dots,j} \{g_i(a)\} - \min_{a=0,1,\dots,j} \{g_i(a)\}$ measures the variation of gray-values along the path from column element $g_i(0)$ to $g_i(j)$ (Fig. 4).

Since the variation of lymph node gray-values is typically in a limited range, the larger the value for $c_{\text{gh}_i}(j)$, the more unlikely it is that the element j of column i belongs to the lymph node.

As mentioned before, lymph nodes and the environment around them can vary considerably, thus an edge term that solely relies on edge magnitude is problematic, as demonstrated by the example shown in Fig. 1(a). To avoid this problem, an edge cost function $c_{\text{edge}_i}(j)$ that identifies potential edge locations, but does not directly utilize edge magnitude information, is generated. First, the derivative of $c_{\text{gh}_i}(j)$ (Fig. 5) is calculated by using a central difference function: $c'_{\text{gh}_i}(j) = \frac{1}{3} \sum_{a=1}^3 [c_{\text{gh}_i}(j+a) - c_{\text{gh}_i}(j-a)]$. Note that values for $c'_{\text{gh}_i}(j)$ are only generated for $j = 3, 4, \dots, (n_e - 4)$ to avoid dealing with undefined border values, and a lymph node boundary is unlikely to be located at the beginning or end of a column. Second, all local maxima of $c'_{\text{gh}_i}(j)$ are detected and the corresponding locations are stored in the set Λ_i . Third, the edge term is calculated by using $c_{\text{edge}_i}(j) = \max_{a \in \Lambda_i} \{p(j, a)\}$ with $p(j, a) = e^{-\frac{(j-a)^2}{2\sigma^2}}$. The function $p(j, a)$ is used to model uncertainty regarding the exact edge location (Fig. 6).

FIG. 5. Derivative of the gray-value homogeneity function in Fig. 4.

FIG. 6. Edge-based cost function for the gray-value profile in Fig. 3.

The relative importance of possible edge locations is globally estimated by the function $w_{\text{shape}}(j) = [\sum_{i=0}^{n_v-1} c_{\text{edge}_i}(j)] / [\max_{j=0,1,\dots,(n_e-1)} \{\sum_{i=0}^{n_v-1} c_{\text{edge}_i}(j)\}]$, and the corresponding plot is shown in Fig. 7.

The idea behind this approach is as follows. Since the user specifies the approximate center $\mathbf{c}_k$ of a lymph node, its edges approximately appear in concentric patterns around $\mathbf{c}_k$. In contrast, other nearby structures (e.g., vessels) within radius r do not lead to such a consistent pattern. Therefore, it is likely that the weight elements of $w_{\text{shape}}(j)$ have larger values in proximity of the real lymph node edge. Thus, $w_{\text{shape}}(j)$ is utilized to weight the edge cost function $c_{\text{edge}_i}(j)$, which helps avoiding problems as shown in Fig. 1(a). Finally, Fig. 8 depicts the node costs for the profile/column shown in Fig. 3(b).

II.D. Lymph node segmentation

Once the graph G is generated and all costs $c_i(j)$ are calculated, a maximum flow algorithm is used to solve the graph optimization problem,¹⁴ which runs in low degree polynomial time. The utilized OSF approach guarantees to produce a globally optimal surface captured by our graph G according to the utilized cost function $c_i(j)$. For the representation of the segmentation result, the initial spherical triangle mesh is utilized. The position of vertices along the radial direction are adjusted to the surface position found by graph search. Since no topology changes of the mesh structure are required, a mesh of the segmentation result can be quickly generated. Note that mesh-based representation M of the segmentation results can be converted into a volume-based representation S by using a voxelization method,¹⁵ if needed.

II.E. Segmentation refinement

As mentioned in Sec. I, lymph node segmentation is a non-trivial problem and local segmentation errors can occur despite the effort in cost function design. Thus, our approach allows the user to correct errors while keeping the time required for segmentation refinement low. This implies that the user must not be required to delineate all the missed portions of the lymph node boundary in a slice-by-slice fashion to correct a segmentation. Instead, the same OSF segmentation framework introduced above will be utilized to make the

FIG. 7. Global shape weight.

FIG. 8. Cost function for the column depicted in Fig. 3.

error correction process fast and intuitive. Figure 9 illustrates the main idea behind the developed refinement approach. In areas that are incorrectly segmented, the user specifies a point located at the correct lymph node boundary and the algorithm will update/manipulate the segmentation in 3D by identifying the affected surface patch, locally changing the OSF node costs, and rerunning the OSF segmentation. The developed refinement algorithm consists of four major steps, which are described in the following.

Step I. Extract gray-value information around the user-selected boundary point.

1. The user locates a local segmentation error and specifies a point $\mathbf{p}_u$ on the correct boundary.
2. The algorithm searches for the closest node on all columns. The closest node is denoted as $v_{i^*}(j^*)$, where i^* and j^* are the column and node of the closest node, respectively. Column i^* will be denoted as user selected column.
3. A gray-value profile around node j^* from column i^* is extracted, which is represented by a set of gray values:

$$S(i^*, j^*) = \{g_{i^*}(j) | j \in [j^* - n_{re}, j^* - n_{re} + 1, \dots, j^* + n_{re}]\}, \quad (2)$$

where n_{re} is an integer which controls the length of the gray-value profile.

Step II. By utilizing a breadth-first-search on the mesh structure, similar neighboring columns are identified and

FIG. 9. Illustration of the segmentation refinement process. (a) and (c) Initial segmentation results, which are partially incorrect. Surface points provided by the user to refine the segmentation in 3D are indicated by dots. (b) and (d) Segmentation refinement results derived from initial results shown in (a) and (c), respectively.

added to the region set Θ . As we have mentioned before in Sec. II.B, the neighborhood relation between columns is defined by the mesh topology. The search starts at column i^* and ends when no more similar columns can be added to Θ or the number of edges on the shortest path from the considered column to the user selected column i^* is larger than a constant n_{max} . The similarity criteria for generating the region set Θ is defined as follows. Let the column i be the column that we want to compare with the user selected column i^* . First, a set of gray-value profiles $\Omega(i, j^*) = \{S(i, j^* - \gamma), S(i, j^* - \gamma + 1), \dots, S(i, j^* + \gamma)\}$ for column i is produced, where γ is a variable position offset to take vertical changes of the surface between two columns into account. If γ is selected too large, profiles far away from the user selected lymph node surface point may be found. If selected too small, nearby similar profiles might not be detected due to the local shape changes. Thus, we utilize $\gamma = \min\{d(v_{i^*}(j^*) - v_i(j^*))/2, \eta\}$, which adapts γ based on the dissimilarity between $v_{i^*}(j^*)$ and $v_i(j^*)$. In this context, η is a constant which is utilized to limit the search range, and $d()$ represents the Euclidean distance. Second, each gray-value profiles $S(i, j) \in \Omega(i, j^*)$ of column i is compared to $S(i^*, j^*)$ by utilizing the following boolean function:

$$\phi(S(i^*, j^*), S(i, j)) = \left(\max_{a=-n_{re}, -n_{re}+1, \dots, n_{re}} \{abs(g_{i^*}(j^* + a) - g_i(j + a))\} \right) \leq \nu, \quad (3)$$

where ν is a threshold. If $\phi(S(i^*, j^*), S(i, j))$ is true for one of the gray-value profiles in $\Omega(i, j^*)$, the column i is considered to be similar to $S(i^*, j^*)$.

Step III. All the node costs of the user selected column i^* and columns in the region set Θ are updated. First, the user selected column i^* is updated. The cost of node $v_{i^*}(j^*)$ is set to 0 and all the other nodes of column i^* are assigned a large constant cost c_{max} . Thus, the surface is forced to pass through the user defined point. Second, all the other columns which passed the similarity criteria are updated. For column i passing the similarity criteria, the cost function is calculated by

$$c_i(j) = (1 - w_i(j)c_{edge_i}(j)) + \alpha c_{gh_i}(j), \quad (4)$$

where the original global shape weight $w_{shape}(j)$ is replaced by a Gaussian weighting function: $w_i(j) = e^{-\frac{(j-\tilde{j})^2}{2\sigma^2}}$, where $\tilde{j}$ is determined as follows. If there are local maxima located on column i in the range of $[j^* - \gamma, j^* + \gamma]$, then $\tilde{j}$ equals to the corresponding location of the local maximum $c'_{gh_i}(j)$ closest to j^* , or else $\tilde{j}$ is set to j^* .

Step IV. Once the user has specified the point $\mathbf{p}_u$, a new segmentation is calculated based on the modified node costs. The new result is displayed immediately after the maximum flow calculation is finished. The refinement process can be repeated, if needed. Note, if the user selected column i^{*1} in the first refinement process, it is also considered as a similar column in the second refinement process, but the node cost of column i^{*1} will remain unchanged in Step III of the second refinement iteration. In addition, the developed user interface allows the user to “undo” a refinement step.

III. EXPERIMENTAL METHODS

In Subsections III.A–III.D, the experimental setup as well as the utilized image data, independent reference standard, quantitative indices, and algorithm parameters are described. A physician oversaw image data collection, trained and supervised image analysts, and inspected/approved the independent reference standard. This study was approved by the Institutional Review Board (IRB) at the University of Iowa.

III.A. Image data and experimental setup

For validation, two experiments were performed, which are described in Secs. III.A.1 and III.A.2.

III.A.1. Segmentation accuracy analysis

For assessment of segmentation accuracy, 111 lymph nodes of the mediastinum, abdomen, head/neck, and axillary regions were selected from 35 volumetric CT data sets. The majority of the lymph nodes was enlarged. Fourteen of the CT data sets were contrast enhanced. The CT data were divided into three test sets:

- *Test set 1* consisted of 22 lymph nodes from three CT scans. Scans in test set 1 were from patients with large lung cancer masses. Pixel spacing ranged between 0.62 and 0.67 mm, and the slice thickness was 0.50, 0.60, and 3.00 mm, respectively. The majority of lymph nodes in test set 1 was imaged with approximately isotropic voxel size.
- *Test set 2* contained 45 lymph nodes from nine CT scans. The scans were taken from cancer patients and were acquired for radiation treatment planning. The spatial resolution for test set 2 is lower compared to test set 1. Pixel spacing was between 0.98 and 1.27 mm, and all scans had a slice thickness of 2.00 mm.
- *Test set 3* consists of 44 lymph nodes from 23 CT scans from patients with lung nodules. Consequently, the lymph nodes in this set are smaller compared to the other two test sets. The pixel spacing was between 0.54 and 0.81 mm. The slice thickness ranged between 0.63 and 1.00 mm.

Based on the independent reference standard (Sec. III.B), the average lymph node volume in test sets 1–3 was 1056, 1621, and 501 mm³, respectively.

For validation, the user (trained image analyst) was asked to select the center of the lymph node, apply the algorithm, and inspect the segmentation result. If the result had errors, the user was required to refine the segmentation.

III.A.2. Center sensitivity analysis

For the initial segmentation, the center position $\mathbf{c}_k$ of a lymph node to be analyzed is required. Placing $\mathbf{c}_k$ at different positions may lead to different segmentation results. Also, the center position selected by different users might vary. Hence, the center sensitivity of our method was analyzed in two dif-

FIG. 10. Axial images of the image data utilized for center sensitivity analysis. The calculated center point $\mathbf{c}_k$ is marked by the red cross.

ferent cases (Fig. 10). For the first case (volume 1), an elliptical, slightly kidney shaped lymph node was selected, whose long axis length was 14 mm and short axis length was 10 mm. The voxel size of the data set was $0.6719 \times 0.6719 \times 0.5$ mm. For the second case (volume 2), an approximately spherical lymph node was chosen whose diameter was approximately 10 mm, and the voxel size was $0.6152 \times 0.6152 \times 0.6$ mm.

First, the center position $\mathbf{c}_k$ was calculated from the corresponding independent reference standard (Sec. III.B) for both volumes. Second, the center position was shifted along the x -, y -, and z -axis from the original center $\mathbf{c}_k$, respectively, and the proposed segmentation method (without refinement) was applied. Third, the mean unsigned distance error (Sec. III.C) was calculated for each center location.

III.B. Independent reference standard

The independent reference was generated by an expert image analyst in a slice-by-slice fashion using a semiautomatic live wire¹⁶ segmentation tool. The live-wire method was chosen to reduce the time required by the user for generating the reference segmentations. A potential issue of this approach might be that object boundaries were identified more consistently compared to completely manual segmentation. On the other hand, the user had the option to influence the boundary location by using path-cooling techniques.¹⁶ The independent reference was utilized for evaluation of the developed lymph node segmentation method.

III.C. Quantitative indices

For segmentation error assessment, surface-based and volume-based measures were used. Let V_S and V_R be binary volume data sets describing the segmentation result and the reference segmentation, respectively. Corresponding surfaces are denoted as S and R . The distance $d(x, A)$ between a point x and a surface A was defined as $d(x, A) = \min_{x' \in A} \|x - x'\|$. The mean unsigned distance error was determined by evaluating $d_U(S, R) = [I(S, R) + I(R, S)]/[|S| + |R|]$ with $I(A, B) = \int_{x \in A} d(x, B) dx$. Similarly, the mean signed distance error d_S was calculated. Distances are counted positive, if the point of the segmentation result surface is located outside of the boundary of the independent reference standard, negative otherwise. To measure the volume error of segmentations, the Dice coefficient¹⁷ $D(V_S, V_R) = 2|V_S \cap V_R|/[|V_S| + |V_R|]$ was determined.

III.D. Parameter selection

For all experiments, the following parameters were utilized for segmentation. For graph construction, r was set to 20 mm to be able to segment larger lymph nodes. The spherical mesh consisted of $n_v = 642$ vertices. The number of elements per column was $n_e = 60$. The selection of n_v and n_e is based on the expected voxel size of CT images to be segmented as well as considerations presented in Sec. II.B. Based on the selected values for n_v and n_e , Δ was set to four, which is sufficient to segment ellipsoidal-shaped lymph nodes with a long to short axis ratio of approximately 1.5. Another aspect in selecting Δ is that abnormal/diseased lymph nodes are frequently spherically shaped, whereas normal lymph nodes are usually elongated (i.e., bean-shaped). For cost calculation, $\alpha = 0.001$ and $\sigma = 1.5$ were used. The following parameters were used for the refinement method: $n_{re} = 5$, $n_{max} = 6$, $\eta = 3$, $v = 30HU$, $c_{max} = 1000$, and $\tilde{\sigma} = 5$. All parameters were determined on a separate CT image set.

IV. RESULTS

IV.A. Segmentation accuracy analysis

Results on the three test data sets for the segmentation without refinement are summarized in Table I(a). For test set 1 and test set 3, the signed mean border positioning error is close to 0, indicating a low border positioning bias. For test set 2, the signed mean border positioning error is 0.421 mm. The spatial resolution for images in test set 2 is low, compared to test set 1. For volumetric CT data with lower resolution, lymph node boundaries are getting blurred [Fig. 15(e)] and the segmentation performance is impacted by partial volume effects. In such cases, even the manual segmentation of lymph nodes becomes a difficult task. The unsigned mean border positioning error for all three test sets is less than a voxel. Once the approximate center of a lymph node of interest is known, the initial OSF segmentation can be calculated in less than 1 s on a 3 GHz workstation.

The number of lymph nodes that required refinement and the average number of refinement iterations ("clicks") for each test set is summarized in Table II. The average user interaction time required for refinement was 10 s per lymph node (range: 4–23 s). Table I(b) summarizes the validation results after refinement for all 111 lymph node segmentations inves-

TABLE I. Average and standard deviation of validation results on all 111 test sets.

	Test set 1	Test set 2	Test set 3
(a) Before refinement			
d_S (mm)	0.027 ± 0.178	0.421 ± 0.216	-0.088 ± 0.217
d_U (mm)	0.524 ± 0.149	0.853 ± 0.176	0.572 ± 0.235
D (-)	0.845 ± 0.059	0.828 ± 0.064	0.775 ± 0.082
(b) After refinement			
d_S (mm)	0.023 ± 0.171	0.394 ± 0.189	0.001 ± 0.146
d_U (mm)	0.518 ± 0.146	0.824 ± 0.146	0.501 ± 0.157
D (-)	0.847 ± 0.061	0.836 ± 0.058	0.809 ± 0.070

TABLE II. Segmentation refinement statistics per test set.

	Number of lymph nodes refined	Average number of points ("clicks") required	Minimum/maximum of points required
Test set 1	4	1.25	1/2
Test set 2	16	1.62	1/2
Test set 3	20	1.55	1/2

tigated. For lymph nodes that required refinement, Table III shows error indices before and after refinement for all three test sets. Examples of segmentation results are depicted in Figs. 11 and 12. Figure 13 shows examples of segmentations before and after refinement.

For the subset of lymph node segmentations that were refined, a Student's t -test at a significance level of 0.05 was performed to determine whether the mean values of d_S , d_U , and D after refinement were significantly different (improved) than before refinement. For test set 1 ($n = 4$), the error indices were not found to be significantly different [Table III(a)]. In the case of test set 2 ($n = 16$), only the mean signed surface distance error d_S was not found to be significantly different, all other indices were significantly different [Table III(b)]. All indices of test set 3 ($n = 20$) were found to be significantly different [Table III(c)].

IV.B. Center sensitivity analysis

Figure 14 depicts the segmentation error for volume 1 and volume 2 when the center position is shifted along the x -, y -, and z -axis between the lymph node boundaries given by the reference segmentation. The location 0 corresponds to the center of a lymph node derived from the independent standard. Figure 14 shows that the error stays approximately in the same low range close to 0.5 mm around the lymph node center. The segmentation error increased when c_k was approaching the lymph node boundary.

TABLE III. Comparison of validation results before and after refinement for cases in test sets 1 to 3 which required refinement.

	Before refinement	After refinement	Statistical significance of improvement after refinement
(a) Test set 1			
d_S (mm)	0.071 ± 0.209	0.044 ± 0.170	$p = 0.7025$
d_U (mm)	0.607 ± 0.050	0.571 ± 0.060	$p = 0.2786$
D (-)	0.883 ± 0.022	0.894 ± 0.020	$p = 0.1017$
(b) Test set 2			
d_S (mm)	0.560 ± 0.233	0.485 ± 0.201	$p = 0.0833$
d_U (mm)	1.001 ± 0.166	0.920 ± 0.141	$p = 0.0038$
D (-)	0.796 ± 0.079	0.820 ± 0.074	$p = 0.0134$
(c) Test set 3			
d_S (mm)	-0.184 ± 0.257	0.011 ± 0.161	$p = 0.0051$
d_U (mm)	0.723 ± 0.245	0.566 ± 0.164	$p = 0.0002$
D (-)	0.736 ± 0.080	0.812 ± 0.072	$p = 0.0001$

FIG. 11. Comparison between segmentation results generated with the proposed method (without refinement) and independent reference standard.

V. DISCUSSION

The comparison between segmentation results (without refinement) and independent standard in Fig. 11 shows mostly local differences despite the variation in shape, size, and imaging conditions. The method successfully deals with the variations in lymph node shape depicted in Fig. 12—none of these examples required segmentation refinement. Traditional edge based segmentation methods fail when there are strong

FIG. 12. Examples of lymph node segmentations generated with the proposed method depicting the variation in size, shape, and texture of lymph nodes. None of the results needed refinement by the user.

FIG. 13. Examples of initial segmentation ((a) and (c)) and refinement results ((b) and (d)).

FIG. 14. Unsigned mean distance error for volume 1 (a), (c), and (e) and volume 2 (b), (d), and (f) in dependence of center location along the x -, y -, and z -axis, respectively. All distances along the horizontal axes are in mm. The minimum and maximum values on the horizontal axis correspond to boundary locations in the reference segmentation.

FIG. 15. Comparison of segmentation results produced by three experts. (a) Axial CT cross-section showing an enlarged lymph node. (b)–(d) Manual segmentation results generated by experts one to three, respectively. (e) Explanation of image features; arrow A points to a region with edge blurring due to partial volume effects and arrow B points to a vessel in close proximity to the lymph node, which is difficult to discern in this particular image slice. (f)–(h) Segmentations generated by experts one to three, respectively, produced with the proposed segmentation approach. In all cases, only the approximate center was specified and no refinement was performed/required.

edges around a lymph node [Fig. 1(a)]. In our method, the influence of neighboring strong edges is reduced [Fig. 1(b)].

A comparison of results on test sets 1 and 2 (Table I) shows that for the test set with low resolution CT scans (test set 2), segmentation errors, independent of the used indices, are larger before and after refinement. Differences in segmentation performance between test sets 1 and 3 can mainly be observed for the Dice coefficient, which is impacted by the size of the analyzed lymph nodes (Sec. III.A.1).

In our cost function design, we assumed that lymph nodes have a spherical or slightly ellipsoid shape, as is the case for most abnormal lymph nodes. However, as can be seen in Figs. 13(a) and 13(c), lymph nodes can deviate from this assumption, which can cause segmentation errors.

If the initial segmentation contains errors, the user can utilize the developed segmentation refinement approach to correct the segmentation result, which takes only a few seconds of user interaction. Examples for initially incorrect lymph node segmentations and corresponding results after refinement are presented in Figs. 9 and 13. In our evaluation on 111 lymph nodes, the expert decided to refine 40 cases (36%) of initial results. The segmentation refinement method is efficient; all segmentation performance indices improved (Table III), and a maximum of two refinement iterations (i.e., mouse clicks) were required (average: 1.55). In addition, the improvement was statistically significant for the majority of calculated indices. The time required for locating a target lymph node in a scan, specifying the center, generating an initial OSF segmentation, and refining the segmentation, if needed, was typically less than one minute. In comparison, the process of generating an independent reference segmentation with the live wire tool in a slice-by-slice fashion required approximately 10 min per lymph node, which is one order of magnitude more compared to our approach.

The developed segmentation refinement method is intuitive to use and allows the user to produce segmentations that can be utilized for assessment of lymph nodes. In addition, the approach has the potential to decrease interobserver variability. For example, Fig. 15 depicts the outcome of an experiment where three expert image analysts were asked to segment a lymph node manually slice-by-slice [Figs. 15(b)–15(d)] and with the proposed method [Figs. 15(f)–15(h)]. As can be seen in Fig. 15, manual segmentation results show more variation when compared across experts and to results of the proposed semiautomatic approach. This can be explained by factors like subjective interpretation of image content, choice of gray-value window width and level, blurred object borders due to partial volume effects [arrow A in Fig. 15(e)], etc., which influence operator performance. Due to a lack of ground truth, it is unclear which of the six segmentation results is closer to “truth”. However, as can be seen in Fig. 15, the proposed method reduces interobserver variability, because it only requires the user to specify the approximate lymph node center point. In addition, because the method works in 3D, the vessel in close proximity to the lymph node [arrow B in Fig. 15(e)] is correctly excluded in all three semiautomatic segmentations.

Analyzing individual refinement cases leads to the following conclusions. First, letting the user refine a segmentation can introduce operator variability. For example, low resolution scans of lymph nodes will likely introduce a high inter- and intraobserver variability, because it can be difficult for the user to locate the correct lymph node boundary. However, as the experiment depicted in Fig. 15 suggests, variation due to subjective interpretation in manual segmentation will also be higher in such cases. Second, for almost all cases, the average Dice coefficient increases with refinement and has values larger than 0.8. However, for small lymph nodes, the Dice coefficient can remain low, even after refinement, because even

a 'small' segmentation inaccuracy will impact the Dice coefficient more than for larger lymph nodes.

VI. CONCLUSION

The assessment of lymph nodes plays an important role in diagnosis, monitoring, and treatment of diseases, like cancer, sarcoidosis, or histoplasmosis. In this paper, we presented a graph-based approach for the computer-aided 3D segmentation of lymph nodes. Segmentation of lymph nodes in volumetric CT images is a challenging task due to partial volume effects, neighboring structures with similar intensity profiles and potentially inhomogeneous appearance. The presented approach is able to deal with common segmentation problems by utilizing a weighted edge and a region homogeneity term in the developed cost function. In the majority of cases investigated, the initial segmentation method delivered results that did not require further processing. For lymph nodes with irregular shape or large variability in density, the developed refinement method can be utilized to adjust the segmentation result, which takes a few seconds of user interaction. Compared to manual or semiautomated slice-by-slice segmentation approaches, the developed method is fast and offers similar flexibility when dealing with lymph nodes that are difficult to segment. Future work will focus on developing a fully automated lymph node detection method that can be utilized in combination with the proposed segmentation framework.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Geoffrey McLennan for the support of this work. He was involved in data collection, trained and supervised image analysts, and inspected the independent reference standard. This work was supported in part by NIH grant R01-EB004640.

^{a)}Author to whom correspondence should be addressed. Electronic mail: reinhard-beichel@uiowa.edu

¹L. Schwartz, J. Bogaerts, R. Ford, L. Shankar, P. Therasse, S. Gwyther, and E. Eisenhauer, "Evaluation of lymph nodes with RECIST 1.1," *Eur. J. Cancer* **45**(2), 261–267 (2009).

²E. Eisenhauer, P. Therasse, J. Bogaerts, L. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, and J. Verweij, "New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)," *Eur. J. Cancer* **45**(2), 228–247 (2009).

³L. H. Schwartz, M. Mazumdar, W. Brown, A. Smith, and D. M. Panicek, "Variability in response assessment in solid tumors: Effect of number of lesions chosen for measurement," *Clin. Cancer Res.* **9**(12), 4318–4323 (2003).

⁴J. J. Erasmus, G. W. Gladish, L. Broemeling, B. S. Sabloff, M. T. Truong, R. S. Herbst, and R. F. Munden, "Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: Implications for assessment of tumor response," *J. Clin. Oncol.* **21**(13), 2574–2582 (2003).

⁵M. Fabel, H. von Tengg-Kobligh, F. L. Giesel, L. Bornemann, V. Dicken, A. Kopp-Schneider, C. Moser, S. Delorme, and H.-U. Kauczor, "Semi-automated volumetric analysis of lymph node metastases in patients with malignant melanoma stage III/IV-A feasibility study," *Eur. Radiol.* **18**(6), 1114–1122 (2008).

⁶J. Cordes, J. Dornheim, B. Preim, I. Hertel, and G. Strauss, "Pre-operative segmentation of neck CT datasets for the planning of neck dissections," *Proc. SPIE* **6144**(2), 6144–6163 (2006).

⁷J. Yan, T. Zhuang, B. Zhao, and L. H. Schwartz, "Lymph node segmentation from CT images using fast marching method," *Comput. Med. Imaging Graph.* **28**, 33–38 (2004).

⁸J. Dornheim, H. Seim, B. Preim, I. Hertel, and G. Strauss, "Segmentation of neck lymph nodes in CT datasets with stable 3D mass-spring models segmentation of neck lymph nodes," *Acad. Radiol.* **14**(11), 1389–1399 (2007).

⁹D. Maleike, M. Fabel, R. Tetzlaff, H. von Tengg-Kobligh, T. Heimann, H. Meinzer, and I. Wolf, "Lymph node segmentation on CT images by a shape model guided deformable surface method," *Proc. SPIE* **6914**, 69141S-1–69141S-8 (2008).

¹⁰K. Lu, S. A. Merritt, and W. E. Higgins, "Extraction and visualization of the central chest lymph-node stations," *Proc. SPIE* **6915**, 69151B-1–69151B-15 (2008).

¹¹G. Unal, G. Slabaugh, A. Ess, A. Yezzi, T. Fang, J. Tyan, M. R. R. Krieg, R. Seethamraju, M. Harisinghani, and R. Weissleder, "Semi-automatic lymph node segmentation in LN-MRI," in *IEEE International Conference on Image Processing (ICIP)* (IEEE, 2006), pp. 77–80.

¹²Y. Wang and R. Beichel, "Graph-based segmentation of lymph nodes in CT data," in *Proceedings of International Symposium on Visual Computing 2010*, Lecture Notes in Computer Science Vol. 6454 (2010), pp. 312–321.

¹³K. Li, X. Wu, D. Chen, and M. Sonka, "Efficient optimal surface detection: Theory, implementation and experimental validation," *Proc. SPIE* **5370**, 620–627 (2004).

¹⁴K. Li, X. Wu, D. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images—A graph-theoretic approach," *IEEE Trans. Pattern Anal. Mach. Intell.* **28**, 119–134 (2006).

¹⁵B. Reitinger, A. Bornik, and R. Beichel, "Efficient volume measurement using voxelization," in *Proceedings of the 19th Spring Conference on Computer Graphics* (ACM, New York, NY, 2003), pp. 47–54 (ACM).

¹⁶W. Barrett and E. Mortensen, "Interactive live-wire boundary extraction," *Med. Image Anal.* **1**(4), 331–341 (1997).

¹⁷L. Dice, "Measures of the amount of ecologic association between species," *Ecology* **26**(3), 297–302 (1945).