Fall 2010

Graph-based segmentation of lymph nodes in CT data

Yao Wang
University of Iowa

GRAPH-BASED SEGMENTATION OF LYMPH NODES IN CT DATA

by

Yao Wang

A thesis submitted in partial fulfillment of the
requirements for the Master of Science
degree in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

December, 2010

Thesis Supervisor: Assistant Professor Reinhard Beichel

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

MASTER'S THESIS

This is to certify that the Master's thesis of

Yao Wang

has been approved by the Examining Committee for the thesis requirement for the Master of Science degree in Electrical and Computer Engineering at the December, 2010 graduation.

Thesis Committee: _____
Reinhard Beichel

_____
Milan Sonka

_____
Andreas Wahle

ACKNOWLEDGEMENTS

First of all, I would like to thank Dr. Reinhard Beichel for his patient guidance throughout my study at the University of Iowa.

During my study in this department, I have benefited from interactions with many faculty members. I would also like to thank Dr. Geoffrey McLennan for providing part of the experimental data. Dr. Xiaodong Wu contributed the source code for the optimal surface finding. Special thanks go to Dr. Milan Sonka and Dr. Andreas Wahle, who served on my defense committee. I would also sincerely thank my colleagues - Shanhui Sun, Yunlong Liu, Qi Song and Xin Dou for their help in my research.

Then I want to thank my parents. Their love crosses thousands of miles, touches my heart. Thanks to Weichen Gao, Panfang Hua, Xiayu Xu, Xiaoxin Fan, Limou Wang and Yunye Shi for their continuous support, especially the food. Thanks to Na Wang, Xue Wu, Hai Zhang, Chaozhi Li, Tammy Leung, Yan Lin and Xiaoyang Xin for their encouraging and comforting. Thanks to Zi Wang for his help and experience.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1
INTRODUCTION

1.1 Motivation

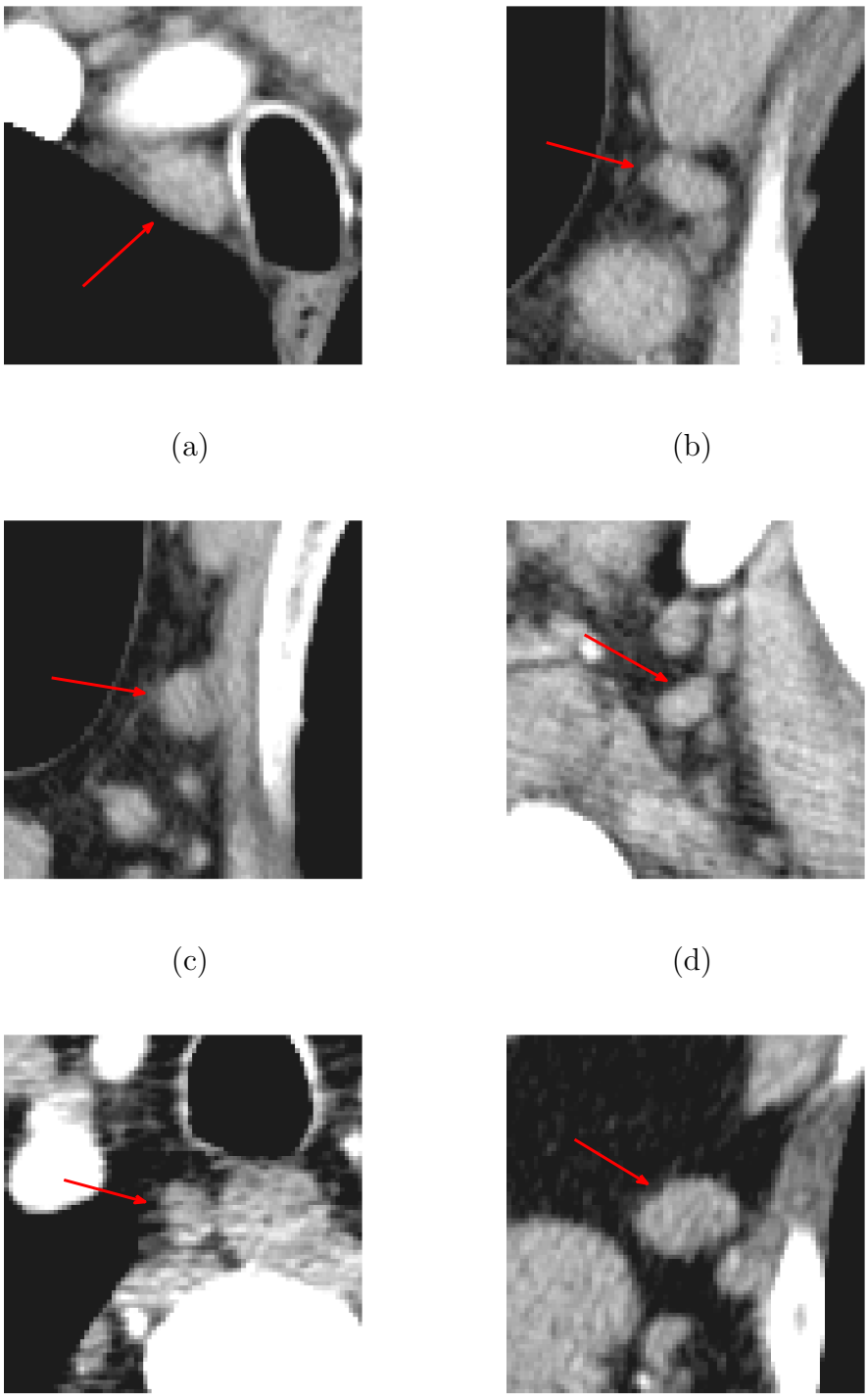
The lymphatic system of the human body is an important component of the immune system. Lymph nodes filter lymph fluid, which collects foreign substances throughout the body, and help dealing with viruses and bacteria. There are hundreds of lymph nodes in the human body [33]. Normal lymph nodes are small in size and slightly bean shaped. Schnyder and Gamsu's study showed that lymph nodes from healthy patients have a mean longest diameter of 5.5 ± 2.8 mm [28]. Einstein et al. considered normal sized lymph nodes to be less than 10 mm [13]. When diseases (e.g., cancer) affect lymph nodes, they can become inflamed or enlarged. For example, it is widely believed that a metastatic involvement of lymph node correlates with enlargement of node[7] [25] [15] [32]. It was shown by Carter et al. that the percentage of lymph node involvement has a linear relation with tumor diameter in breast cancer cases [8]. In another study, enlarged mediastinal lymph nodes with a short axis diameter of larger than 10 mm, were found to be malignant with a probability of 70% [2].

Lymph nodes size may change overtime due to disease progression or treatment. Hence the analysis of lymph node size over time is used to determine whether a treatment is successful or not [17]. Therefore, the assessment of the condition of lymph nodes over time is important for diagnosis, monitoring, and treatment of diseases, like cancer.

Multidetector computed tomography (CT) has become the primary lymph

node imaging modality in clinical routine and offers excellent spatial resolution for measuring lymph nodes [29]. In current clinical practice, lymph nodes are analyzed manually in volumetric CT data based on very rough measures of long and/or short axis length [31] [14] [11]. Such analysis requires to explicitly determine a lymph node axis, which is a source of potentially large errors. The accurate and reproducible measurement of lymph node is critical for determining response to therapy in clinical practice and informed research studies. A true 3D quantification of lymph nodes size promises to be more accurate and reproducible in longitudinal studies.

Lymph node segmentation in volumetric CT data is a challenging task due to partial volume effects, diffuse edges, neighboring structures with similar intensity profiles and potentially inhomogeneous density-values (Fig. 1.1). So far, computer-aided 3D segmentation approaches have not been utilized routinely in research or clinical practice, because currently available automated segmentation methods (Section 1.2) frequently fail to deliver usable results. It is generally accepted that a manual segmentation is too time consuming and introduces inter-operator variability. Thus, new segmentation approaches are required to facilitate the quantitative analysis of lymph nodes.



(a) (b)

(c) (d)

(e) (f)

Figure 1.1: Examples of cross-sectional CT images of lymph nodes.

1.2 Related Work

1.2.1 Lymph Node Segmentation Methods

Literature about the segmentation of lymph nodes is rare. Using a shape prior, Honea et al. [18] suggested a three-dimensional active surface model to deal with the lymph node segmentation problem in CT. The reported segmentation evaluation was merely based on phantom images and the running time of the segmentation was not provided. Lu et al. developed a live-wire-based semi-automatic system for segmenting central chest lymph nodes [23]. The system also includes a manual slice tracing method for refinement. Yan et al. proposed a fast marching approach to perform semi-automatic segmentation of lymph nodes in 2D CT data [35]. They suggested to use their algorithm on a slice-by-slice basis for the 3D segmentation of lymph nodes in volumetric CT image data. A stopping criteria was utilized to avoid 'boundary leaking'. This criteria was implemented by a bounding circle around the lymph node which is placed by the user. No quantitative evaluation results were provided in their paper. A manual segmentation approach for neck lymph node segmentation has been utilized to aid planning of neck dissections [9]. Dornheim et al. [12] proposed a 3D mass-spring model for the segmentation of neck lymph nodes in volumetric CT data. They used a mesh with a low number of triangles. When segmenting the same target by different users, the segmentation results were not consistent, especially for large lymph node cases. Their volumetric segmentation error ranged between 39% and 52%. Kitasaka et al. [20] utilized a 3-D minimum directional difference filter for extracting abdominal lymph nodes in CT data. A series of processing steps like region growing and several false positive reduction strategies were applied for segmentation. Their experiment was based on five volumetric CT images. The author reported

that the method could detect 57.0% of the enlarged lymph nodes with approximately 58 false positives per data sets. In [24], a deformable surface search was used in combination of statistical shape model for lymph node segmentation. Additionally, an application which allows manual interactive correction of errors was provided to help the algorithm in converging to the desired object contours. The method required the user to draw a rough contour inside a lymph node as initialization. For experiments, the method used 29 lymph nodes of 4 CT data sets. The computation time for their segmentation was 15s on a standard PC. In 6.9% out of 29 cases, the authors reported problems in getting a 'usable segmentation'.

In conclusion, the robustness of lymph node segmentation methods still needs to be further improved and it is desirable to have a 3D segmentation approach that requires only little user interaction.

1.2.2 Graph-based Image Segmentation

Since our method that we will introduce in Chapter 2 is based on a graph-based approach, we will give a short introduction to graph-based image segmentation method in this section.

Graph-based image segmentation plays an important role in image processing. Minimum cut/maximum flow algorithms were used by Greig et al. for binary image reconstruction [16]. Boykov et al. proposed an interactive segmentation algorithm for n-D image data based on minimum s-t cuts [5] [6] [4]. 'Seeds' representing object and background are initially identified in the images, which serve as hard constraints for the segmentation process. In [34] [21] [22], a globally optimal surface finding method

with a hard smoothness constraints was presented for segmentation of volumetric image . Our lymph node segmentation approach is based on this segmentation frame work, and in the following paragraph we generally describe the terrain-like single optimal surface finding method following the notation convention in [22].

First, a node-weighted directed graph $G = (V, E)$ is constructed in a volumetric image $\mathcal{I}$. Every node $V(x, y, z) \in V$ represents one voxel $\mathcal{I}(x, y, z) \in \mathcal{I}$. Let $\mathcal{N}(x, y)$ represents a surface in the volumetric image and $Col(x, y)$ represents a column, which is a set of nodes parallel to the **z**-axis. $\mathcal{N}(x, y)$ intersects each column at one voxel. A cost value $c(x, y, z)$ is computed for each node in $G = (V, E)$. $c(x, y, z)$ is an arbitrary real value. The smaller the value of $c(x, y, z)$ is, the higher the possibility that the desired surface contains the voxel $\mathcal{I}(x, y, z)$. An optimal surface is the surface with the minimum cost among all feasible surfaces in the 3D volume.

Fig. 1.2 presents a simple 2D example of how the optimal surface is detected in a 2×4 image. As can be seen from Fig. 1.2(a), the graph has only two columns. Each column consists of four nodes. The costs have already been assigned to each node. The goal of the optimal surface finding is to find the 'minimum cost path' from left to right. The edges are constructed as shown in Fig. 1.2(b). Every node $V(x, y, z)$ with $z > 0$ has a directed arc to the node immediately below, which is denoted as intracolumn arcs (ab, bc, cd, ef, fg and gh). For two adjacent columns, a directed arc is constructed from each node $V(x, y, z) \in Col(x, y)$ to node $V(x + 1, y, \max(0, z - \Delta)) \in Col(x + 1, y)$ (af, eb, bg, fc, ch, hd, dh, hd), which is denoted as intercolumn arcs. The adjacency of columns are defined by the neighborhood setting. For example, $Col(x, y)$ is adjacent to $Col(x + 1, y)$, $Col(x - 1, y)$, $Col(x, y + 1)$ and $Col(x, y - 1)$ under the 4-neighborhood setting. The maximum allowed vertical distance of an

intercolumn arc, which is defined as smoothness constraint Δ, equals to 1 in this example. In Fig. 1.2(c), a cost transformation is performed. The cost of each node is subtracted from the cost of node immediately below it. If the sum of the bottommost two nodes are less than 0, the bottommost two nodes are left unchanged. If not, the sum is increased by one and subtracted from any single one of the bottom most nodes. Then solving the optimal surface finding problem in this node-weighted graph $G = (V, E)$ is transformed into solving a minimum $s - t$ cut problem of an edge-weighted directed graph (Fig. 1.2(d)). By solving the $s - t$ cut, a minimum closed set is generated. The uppermost nodes of this minimum closed set represents the optimal surface in the graph.

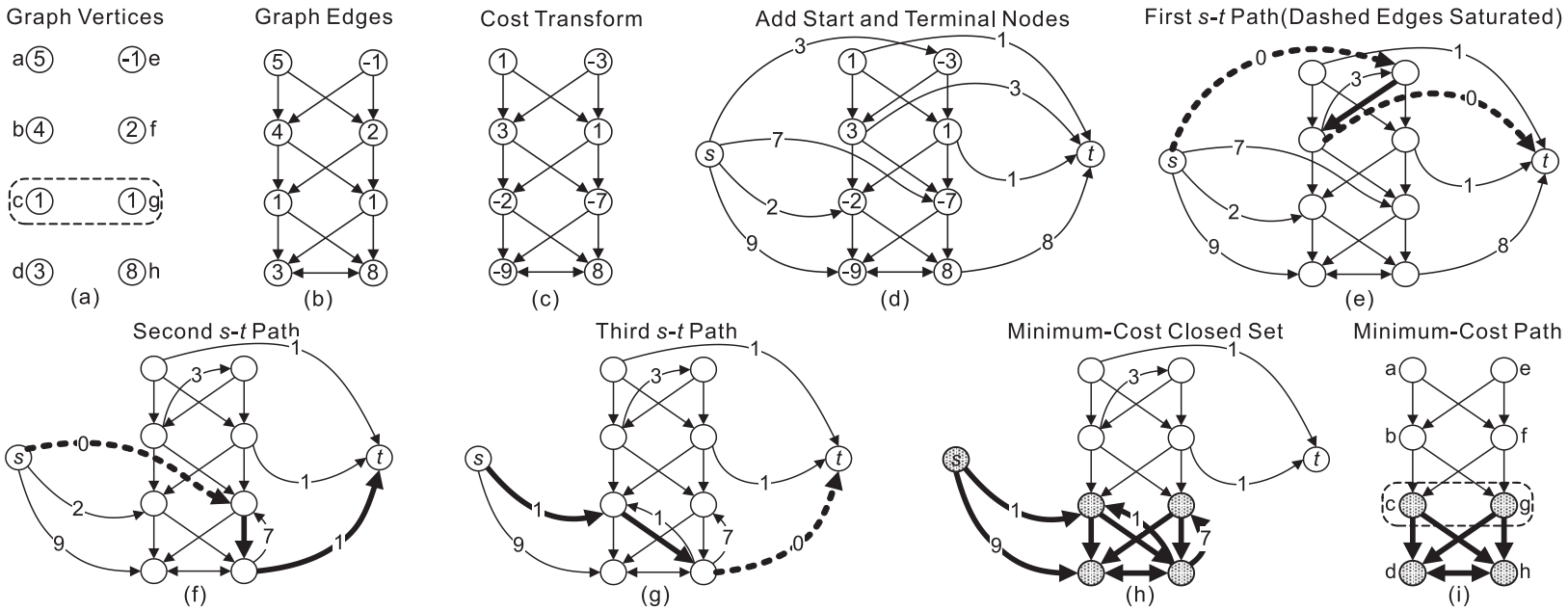


Figure 1.2: Simple 2D example of the minimum cost graph-search detection algorithm (taken from [30]).

1.3 Aims

The aims of this thesis are as follows. First, develop a lymph node segmentation method that a) requires little user interaction and can handle many of the obstacles common to lymph node segmentation in CT volumes and b) allows the user to quickly and interactively refine a segmentation result, if needed. Second, evaluate the method on a diverse collection of CT data sets to assess segmentation performance under different conditions like image resolution or lymph node size. The long-term goal behind this work is to aid physicians in the assessment of response to therapy in cancer patients by providing better methods for lymph node size measurement.

1.4 Organization of this Thesis

The thesis is organized as follows. In Chapter 2, the segmentation approach, consisting of a basic segmentation and a segmentation refinement method is presented. In Chapter 3, experiments performed on phantom and volumetric CT data are described, and validation results for volumetric CT data are presented. In Chapter 4, segmentation results are discussed in detail, and in Chapter 5, conclusions of this work are presented.

CHAPTER 2
METHODS

In this chapter we present a graph-based optimal surface finding approach for lymph node segmentation in volumetric CT data, which is based on a single optimal surface finding algorithm published in [21] [22], and introduces a method for refining a segmentation result, if needed. Our approach avoids common problems of lymph node segmentation (Fig. 2.1) by utilizing a cost function that consists of a weighted edge and a region homogeneity term in combination with a surface smoothness constraint. This cost function is based on the lymph nodes' shape and intensity features.

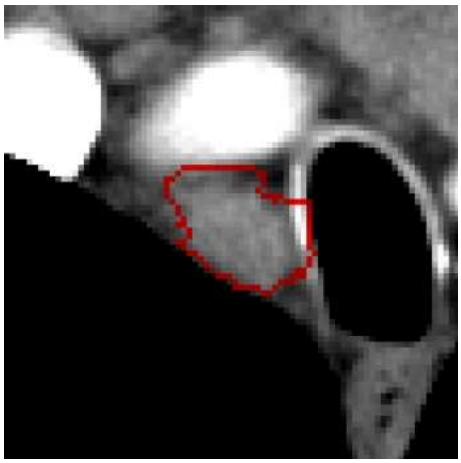


Figure 2.1: Example of a failed lymph node segmentation. In this case, an edge-based segmentation approach was utilized. Edges of neighboring structures like contrast enhanced vessels or airways are stronger and cause segmentation errors.

Our method consists of four main processing steps, as demonstrated in Fig. 2.2. 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 representation. Then the segmentation problem becomes a graph search (optimization) problem. For solving a graph search problem, we utilize the optimal surface finding framework introduced by Li et al. [22]. 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 and the result is converted from a graph representation into a surface mesh and labeled volume. Fourth, the user inspects the segmentation result and can utilize a provided tool to correct the segmentation, if needed. In the following sections, we describe our segmentation approach in detail. In Section 2.1, our preprocessing method is introduced. In Section 2.2, the graph construction process is explained. In Section 2.3, we show how the cost function is generated and applied to the graph. In Section 2.4, we explain how the segmentation result is obtained after solving the graph search problem. In Section 2.5, we describe our segmentation refinement method. In Section 2.6, we show an approach to select the smoothness constraint.

2.1 Preprocessing

Lymph nodes exhibit densities in CT with a range between -100 and 150 Hounsfield units (HU). In order to limit the CT density-values to this range, we

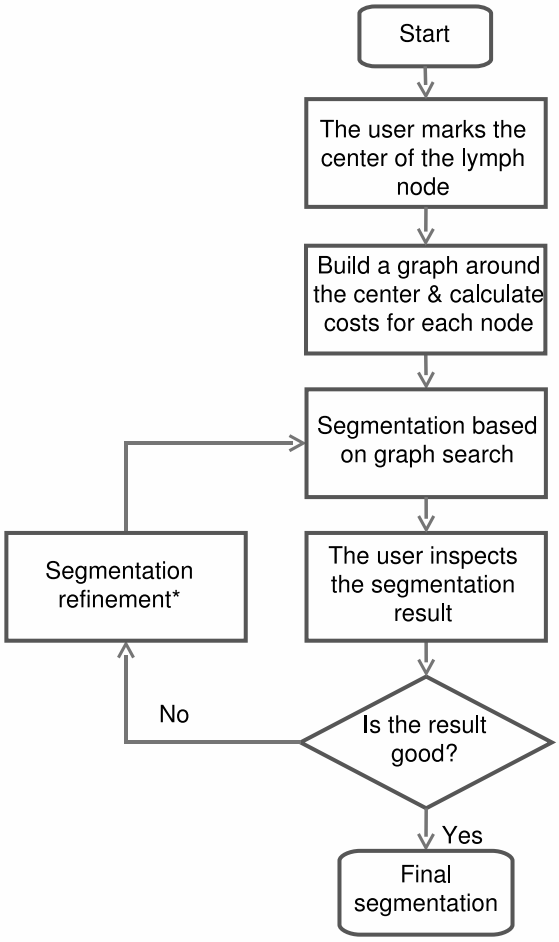


Figure 2.2: Flow chart of main processes. *The user can also undo the last refinement.

apply an intensity transfer function to the original volume data set:

$$f_{iw}(x,y,z) = \begin{cases} 150HU & \text{if } 150HU < f(x,y,z) \\ f(x,y,z) & \text{if } -100HU \leq f(x,y,z) \leq 150HU \\ -100HU & \text{if } f(x,y,z) < -100HU \end{cases}, \tag{2.1}$$

where $f(x, y, z)$ represents the density-value of a voxel at the coordinates x, y, and z. To reduce image noise, a $3 \times 3 \times 3$ median filter is applied to the transformed volume data set.

2.2 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. 2.3(a)) is built around $\mathbf{c}_k$ with radius r. The radius r is a constant and chosen to be larger than the maximally 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 $i \in n_v$, the volume f_{iw} is resampled along the line between center point $\mathbf{c}_k$ and vertex $\mathbf{p}_i$ in an equidistant fashion by using a linear interpolation function. The line between center point $\mathbf{c}_k$ and vertex $\mathbf{p}_i$ is denoted as column i. Note that $\mathbf{p}_i$ represents already a sample point, whereas $\mathbf{c}_k$ is not used as sample point. In addition, due to the definition of our cost function (Section 2.3), the first and last few nodes on each column will not be considered as a possible solution. The gray-value density samples on the line between $\mathbf{c}_k$ and $\mathbf{p}_i$ form the elements $g_i(j)$ of column i with $j \in [0, 1, \ldots, (n_e - 1)]$. 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 density sample point $g_i(j)$. The node $v_i(j)$ is the element j of column i in the graph. Because the goal of our algorithm is to find the minimum cost path in the graph, the element, belonging

to boundary, should be assigned a low cost value. The calculation of $c_i(j)$ is described in Section 2.3. Overall, there are $n_e \times n_v$ nodes in the graph. Once the node-weighted directed graph is constructed, it is converted to an an edge-weighted directed graph, as described in[22].

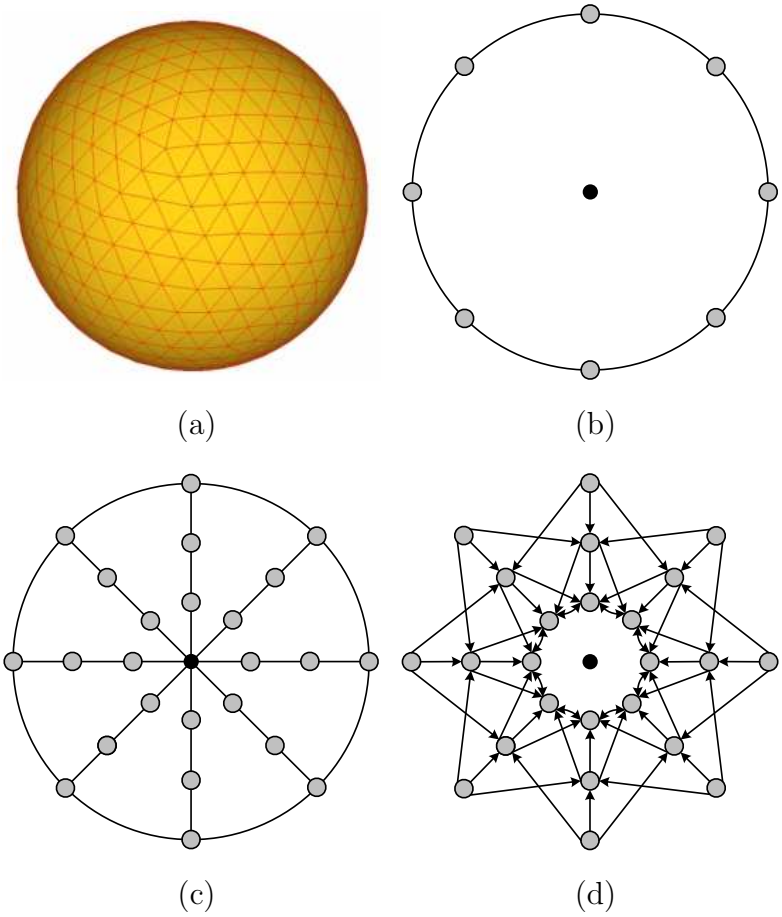


Figure 2.3: Graph generation for the optimal surface finding step. (a) Spherical triangle mesh utilized for graph building. (b)-(d) 2D illustration of the graph generation process.

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. This includes the set up of a surface smoothness constraint Δ between any

two adjacent columns to specify the maximum allowable change in columns for the surface [22]. Fig. 2.3 depicts a 2D example of the graph generation for $\Delta = 1$. By constructing the graph G as described above, we are able to utilize the surface detection algorithm for the segmentation of lymph nodes, which can have a spherical, elliptical, or slightly kidney-like shape.

The parameters used for constructing the mesh are presented in Section 3.1. The following paragraph outlines consideration used to select values for n_v and n_e. The number of vertices n_v is estimated based on the expected size (radius) of the lymph nodes. In our test data sets (Chapter 3), most of the lymph nodes have a radius of approximate 5 mm. For a spherical lymph node of radius of 5 mm, which is represented by a mesh with $n_v = 642$, the mean and standard deviation of the edge length is 0.75 ± 0.05 mm, and maximum edge length is 0.82 mm. Hence the mesh density on the lymph node surface is approximately 1 vertex per voxel, assuming that the voxel size is $0.7 \times 0.7 \times 0.7$ mm^3. If n_v is selected too small, the mesh will be sparse and unable to represent lymph node's surface accurately. On the otherhand, if n_v is chosen too large, the computation time will increase with no benefits regarding segmentation accuracy. For graph construction, we utilize r=20 mm to be able to segment larger lymph nodes, and n_e is set to 60 to have approximate two samples per voxel.

2.3 Cost Function

As we explained in Section 2.2, the cost $c_i(j)$ need to be assigned to the nodes $v_i(j) \in V$ in the node-weighted directed graph $G = (V, E)$. We define our cost

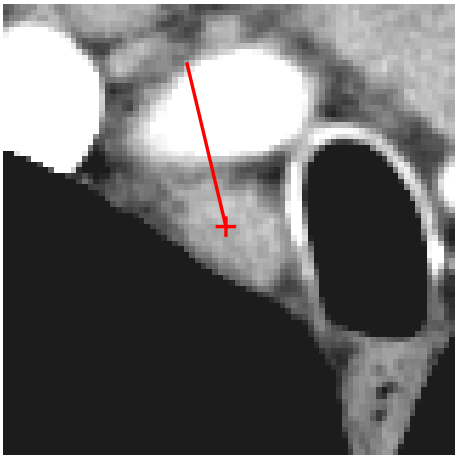


Figure 2.4: Cross-sectional CT image showing a lymph node. The line indicates a column of the graph structure G.

function as follows:

$$c_i(j) = (1 - w_{shape}(j)c_{edge_i}(j)) + \alpha c_{gh_i}(j), \tag{2.2}$$

where $c_{edge_i}(j)$ represents an edge term and $c_{gh_i}(j)$ a gray-value homogeneity cost term. $w_{shape}(j)$ is a global shape weight, which describes the lymph node's boundary distribution. It is derived from all the n_v columns of the graph structure. α is used to adjust the influence of the homogeneity term ralative to the edge term. Since nodes with a high probability of belonging to the lymph node surface should be assigned low costs and the maximum value of $w_{shape}(j)c_{edge_i}(j)$ is 1, we inverse it by using $1 - w_{shape}(j)c_{edge_i}(j)$. In the following paragraph, 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. 2.4. The line in Fig. 2.4 represents a column of the graph. The corresponding gray-value sample

points $g_i(j)$ are depicted in Fig. 2.5(a). The gray-value homogeneity term

$$c_{gh_i}(j) = \max_{a=0,1,\ldots,j}\{g_i(a)\} - \min_{a=0,1,\ldots,j}\{g_i(a)\} \tag{2.3}$$

measures the variation of gray-values along the path from column element $g_i(0)$ to $g_i(j)$ (Fig. 2.5(b)). Since the variation of gray-value in the lymph node is in a certain range, the larger the value for $c_{gh_i}(j)$, the more unlikely it is that the element j of column i belongs to the lymph node.

As mentioned before, the environment around lymph nodes and lymph nodes per se can vary considerably (Fig. 1.1). Hence an edge term that solely relies on edge magnitude is problematic, as demonstrated by the example shown in Fig. 2.1. To avoid this problem, an edge cost function $c_{edge_i}(j)$ that identifies potential edge locations, but doesn't directly utilize edge magnitude information, is generated. At first, the derivative of $c_{gh_i}(j)$ is calculated by using a central difference function:

$$c\prime_{gh_i}(j) = \frac{1}{3}\sum_{a=1}^{3}[c_{gh_i}(j+a) - c_{gh_i}(j-a)]. \tag{2.4}$$

Note that values for $c\prime_{gh_i}(j)$ are only generated for $j = 3, 4, \ldots, (n_e - 4)$ to avoid dealing with undefined border values. In addition, a lymph node boundary is unlikely to be located at the begin or the end of a column (Fig. 2.5(c)). Secondly, all local maximums of $c\prime_{gh_i}(j)$ are detected and the corresponding locations are stored in the set Λ_i. Thirdly, the edge term is calculated by using

$$c_{edge_i}(j) = \max_{a\in\Lambda_i}\{p(j,a)\} \tag{2.5}$$

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. 2.5(d)). Then the relative importance of possible edge locations is globally estimated by

$$w_{shape}(j) = \frac{\sum_{i=0}^{n_v-1} c_{edge_i}(j)}{\max_{j=0,1,\ldots,(n_e-1)}\{\sum_{i=0}^{n_v-1} c_{edge_i}(j)\}}. \tag{2.6}$$

The corresponding example plot is shown in Fig. 2.5(e). 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 very likely that the weight elements of $w_{shape}(j)$ have larger values in proximity of the real lymph node edge, and we can utilize $w_{shape}(j)$ to weight the edge cost function term $c_{edge_i}(j)$. This allows us to avoid problems as shown in Fig. 2.1.

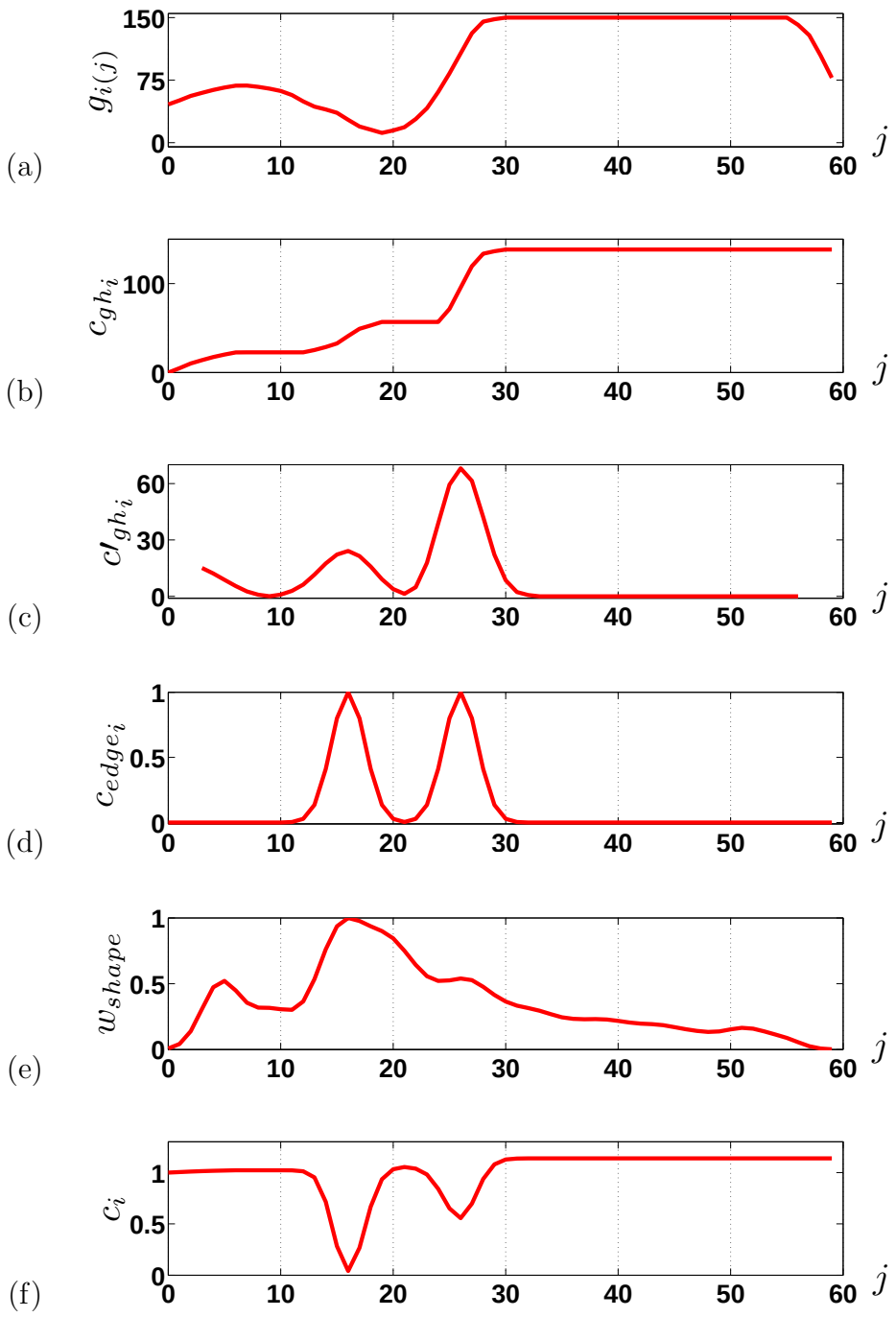


Figure 2.5: Cost function calculation. (a) Gray-value profile corresponding to the column shown in Fig. 2.4. (b) Gray-value homogeneity function and (c) its derivative. (d) Edge cost function. (e) Global shape weight. (f) Final cost function for the column shown in Fig. 2.5(a).

2.4 Lymph Node Segmentation

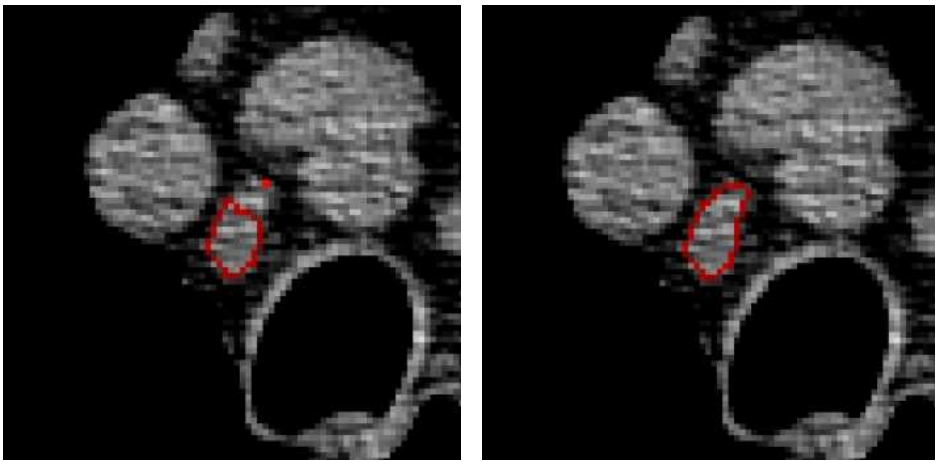
Once the graph G is generated and all costs are calculated, a maximum flow algorithm is used to solve the graph optimization problem [22], which runs in low degree polynomial time. The utilized surface finding framework guarantees to produce a globally optimal surface captured by our graph G according 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 of the same column found by graph search. Since no topology changes of the mesh structure are required, a mesh of the segmentation result can be generated. In addition, the mesh-based representation M of the segmentation results is converted into a volume-based representation S by using a voxelization method [26].

2.5 Segmentation Refinement

After the generation of the segmentation, the user can inspect the result and decide whether a refinement of the segmentation is required or not. If a refinement is needed, like in the case shown in Fig. 2.6(a), the user is required to specify a point (red dot) on the correct boundary and start refinement algorithm. The result of the refinement is shown in Fig. 2.6(b), and the error is successfully corrected. The refinement approach consists of four major steps, which are described below.

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

1. The user inspects the initial segmentation result and specifies a point $\mathbf{p}_u$ on the missed boundary.



(a) (b)

Figure 2.6: Refinement example. (a) The user identifies a point (red dot) on the missed (correct) lymph node boundary. (b) Refinement result.

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^* is denoted as user selected column.

3. Extract a gray-value profile (Fig. 2.7) around node j^* from column i^*, which is represented by a set of gray-values:

$$S(i^*, j^*) = \{g_{i^*}(j) | j \in [j^* - n_{re}, j^* - n_{re} + 1, \ldots, j^* + n_{re}]\}, \qquad (2.7)$$

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

Step II Search for similar neighboring columns by utilizing a breath-first-search-based region growing approach based on the mesh structure. As we have mentioned before, the mesh topology defines the neighborhood relation between columns. The region growing starts at column i^* and examines all adjacent columns using a similarity criteria. If a column adjacent to i^* pass the simi-

larity criteria, it is added to the region set, and all the untested neighbors of this column are tested in the next iteration. This process is repeated until all adjacent columns fail to pass the similarity criteria 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 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 $\Omega(i, j^*) = \{S(i, j^* - \gamma), S(i, j^* - \gamma + 1), \ldots, S(i, j^* + \gamma)\}$ (Fig. 2.8) for column i is produced, where $\gamma = \min\left\{\frac{d\left(v_{i^*}(j^*), v_i(j^*)\right)}{2}, \eta\right\}$ is a variable position offset to take the vertical changes of the surface between two columns into account. γ is utilized to constrain the search range for similar profiles on columns. If 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. η is a constant to constrain this similarity search on a column to a small range, and $d\left(v_{i^*}(j^*), v_i(j^*)\right)$ is the Euclidean distance between node $v_{i^*}(j^*)$ and $v_i(j^*)$.

Next, each gray-value profile $S(i, j) \in \Omega(i, j^*)$ of column i is compared to $S(i^*, j^*)$ by utilizing the following boolean function:

$$\phi(S(i^*, j^*), S(i, j)) = \left(\max_{a=-n_{re}, -n_{re}+1, \ldots, n_{re}} \left\{abs\left(g_{i^*}(j^* + a) - g_i(j + a)\right)\right\}\right) \leq \nu, \tag{2.8}$$

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 weights of the user selected column i^* and similar columns in the region set 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 given a very large cost c_{max}. Second, all the other columns which passed the similarity criteria are updated. For column i which passes the similarity criteria, the cost function is calculated by

$$c_i(j) = (1 - w_i(j)c_{edge_i}(j)) + \alpha c_{gh_i}(j), \tag{2.9}$$

where the original global shape weight $w_{shape}(j)$ is replaced by a new weighting function:

$$w_i(j) = e^{\frac{-(j-\tilde{j})^2}{2\tilde{\sigma}^2}}. \tag{2.10}$$

$\tilde{j}$ is determined as follows: If there are local maximums located on column i in the range of $[j^* - \gamma, j^* + \gamma]$, then $\tilde{j}$ equals to the corresponding location of the local maximum $c\prime_{gh_i}(j)$ closest to j^*. Otherwise, $\tilde{j}$ is set to j^*.

Step IV Re-run the optimal surface finding approach based on the modified cost, and generate a new segmentation result. After inspection of the new segmentation result, the user can 'undo' the performed refinement, if needed, so that the cost on each node and the segmentation result are set the same value as before the refinement. If another part of the new segmentation result requires refinement, the user can start a second refinement based on the newly generated result. Note, if the user selected column i^{*1} in the first refinement process is considered as a similar column in the second refinement process, the cost of nodes on

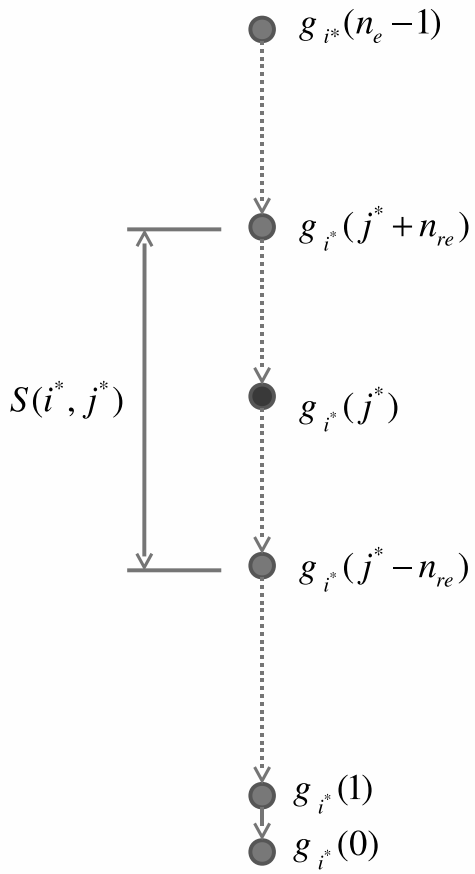


Figure 2.7: Extracting of a gray-value profile from the user selected column i^*.

this column i^{*1} will remain unchanged in the Step III of the second refinement process. The user can perform several refinement iterations until he/she is satisfied with the result.

2.6 Selection of the Smoothness Constraint Δ

The smoothness constraint (Section 1.2.2) is defined as the maximum allowed vertical distance of an intercolumn arc. For example, if we want to segment an ideal spherical structure and the center of the graph $\mathbf{c}_k$ has been put at the center of the

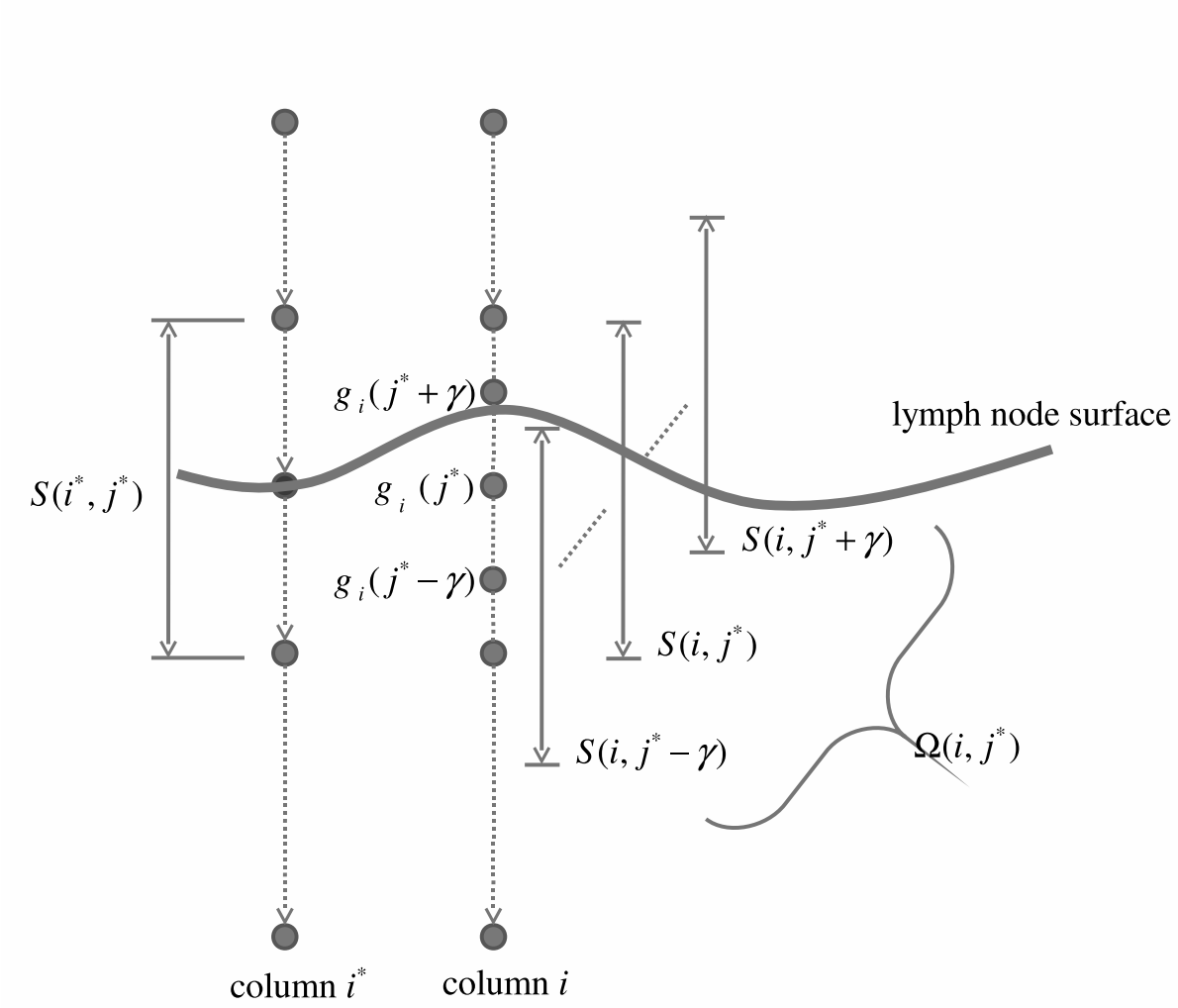


Figure 2.8: Search for similar neighboring columns.

sphere, then $\Delta = 0$ is adequate for the segmentation. But, if we want to segment a ellipsoidal structure, Δ should be chosen larger so that the maximum allowed vertical distance between two adjacent column is large enough to follow the surface. Hence, the selection of an adequate smoothness constraint 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. On the other hand, if a smoothness constrain is too large, the resulting surface can be noisy. In this section, we will introduce a method that allows to estimate the required smoothness constraint for ellipsoid structures.

The necessary smoothness constraint is depended on the maximum angle between two adjacent columns in our graph structure (Fig. 2.3(b)). The length of triangle edges of the mesh is approximately the same. Using (p, q) to represent edges of the triangular mesh, the nodes on column p are $v_p(j)$ and the nodes on column q are $v_p(j)$. In all our experiments, the spherical triangular mesh consisted of $n_v = 642$ vertices, and maximum angle of $\angle p\mathbf{c}_k q$ in our mesh is $\theta = 9.44\,^\circ$. The mean and standard deviation of $\angle p\mathbf{c}_k q$ is $8.64 \pm 0.56\,^\circ$.

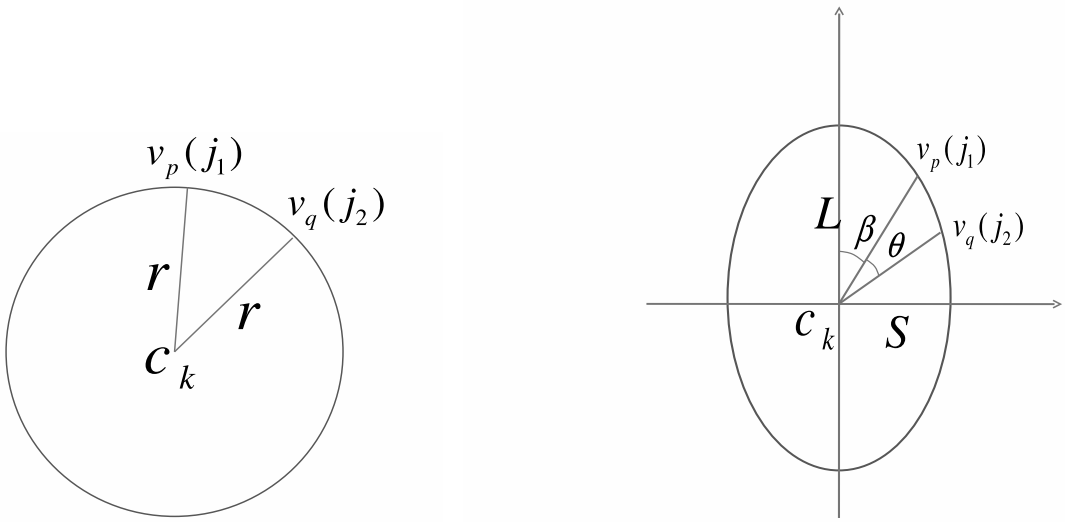
Column p, column q and center point $\mathbf{c}_k$ constitute a plane. Since some lymph nodes can have an ellipsoid like shape, we consider the case where cross section between the plane and the mesh of the segmentation result is an ellipse.

As shown in Fig. 2.9(a), if $v_p(j_1)$ and $v_q(j_2)$ are on the boundary of the cross section of a sphere, $v_p(j_1)$ and $v_q(j_2)$ have the same distance to the center $\mathbf{c}_k$, and $j_1 - j_2 = 0$. Hence, if we want to segment a sphere, we can choose the smoothness constraint $\Delta = 0$. However, if $v_p(j_1)$ and $v_q(j_2)$ are on the boundary of the cross section of an ellipsoid, the value of $j_1 - j_2$ is not the same for different edges. The maximum value of $j_1 - j_2$ is related to θ, the semi-long axis length L and semi-short

axis length S. If $L = mS$, where m is the ratio between L and S, then the length between $\mathbf{c}_k$ and a point $\mathbf{p}_i$ on the boundary of the cross section is

$$t = \frac{mS}{\sqrt{m^2 \sin^2 \beta + \cos^2 \beta}}, \tag{2.11}$$

where β is an angle between the semi-long axis and line segment $\mathbf{p}_i\mathbf{c}_k$.



(a) (b)

Figure 2.9: Intersection of a plane with a lymph node model. (a) Intersection of a plane with a sphere. (b) Intersection of a plane with an ellipsoid.

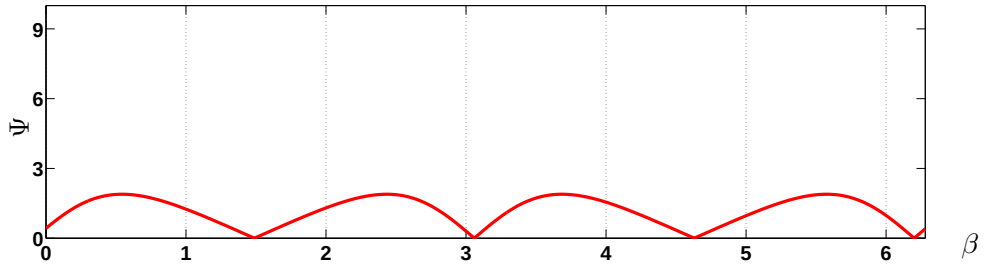
If t_1 is the length of line segment $v_p(j_1)\mathbf{c}_k$ and t_2 is the length of line segment $v_q(j_2)\mathbf{c}_k$ (as shown in Fig. 2.9(b)), the change of vertical length in adjacent columns will be $|t_1 - t_2|$. For our mesh structure, the maximum angle between these two line segments is θ. Hence

$$|t_1 - t_2| = |\frac{mS}{\sqrt{m^2 \sin^2 \beta + \cos^2 \beta}} - \frac{mS}{\sqrt{m^2 \sin^2(\beta + \theta) + \cos^2(\beta + \theta)}}|. \tag{2.12}$$

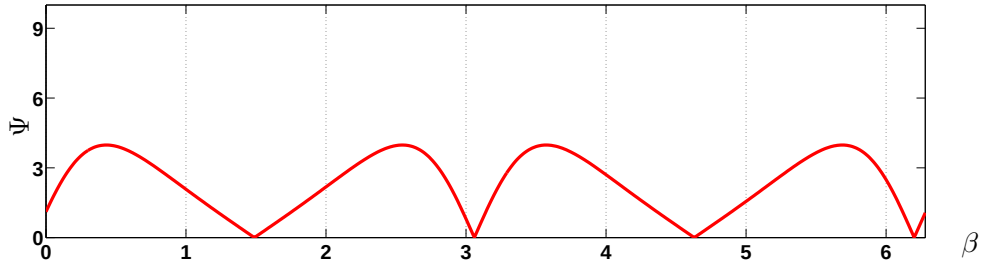
β is an angle as shown in Fig. 2.9(b). The vertical difference between two adjacent column in terms of node indices can be calculated with:

$$\Psi = \frac{n_e * |t_1 - t_2|}{r}, \tag{2.13}$$

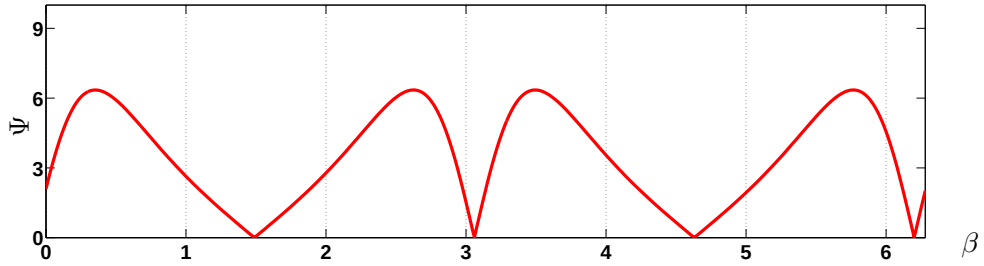
Ψ can be used to estimate the required Δ. Plots for Ψ are shown in Fig 2.10 for different m, if we utilize $r = 20$ mm, and a semi-short axis length $S = 15$ mm as an estimate of lymph node size. For example, for $m = 1.5$, the maximum value of Ψ is 3.8. Hence the smoothness constraint Δ for $m = 1.5$ can be assigned as 4. When $\beta = \pm\frac{\pi}{2} * n - \frac{\theta}{2}$, where $n = \{1, 2\}$, the length of t_1 is the same as the length of t_2 ($\Delta = 0$), so the plots show periodic minimums.



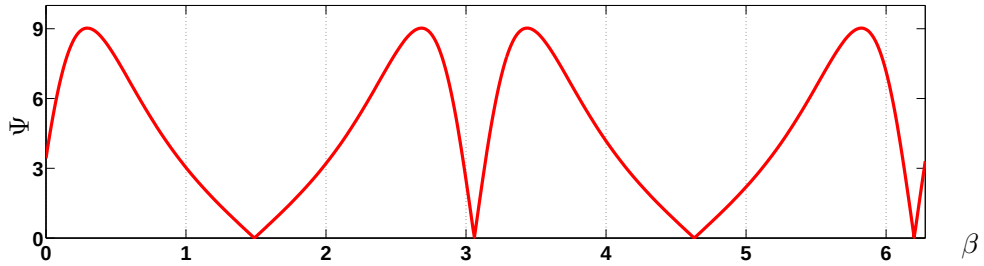
(a) $m = 1.25$



(b) $m = 1.5$



(c) $m = 1.75$



(d) $m = 2$

Figure 2.10: Plots of Ψ for $r = 20$ mm and $S = 15$ mm.

CHAPTER 3
EXPERIMENTS AND EVALUATION

3.1 Introduction

In this chapter, we will describe the performed experiments and present validation results. First, we will present the experiments performed on computer phantoms, to investigate the segmentation performance of our method on various geometrical structures. Second, we will demonstrate the impact of different center locations on our segmentation method. Third, we will describe the evaluation of our method on lymph nodes imaged with CT and present segmentation results.

3.1.1 Parameters

In all our experiments, we used the following parameters. For the graph generation, the column length was set to $r = 20$ mm, and the smoothness constraint $\Delta = 4$ was utilized. The spherical mesh consisted of $n_v = 642$ vertices. The number of elements per column was $n_e = 60$. For cost calculation, $\alpha = 0.001$ and $\sigma = 1.5$ were used. For the refinement approach, the following parameters were used: $n_{re} = 5$, $n_{max} = 6$, $\eta = 3$, $\nu = 30HU$, $c_{max} = 1000$, and $\tilde{\sigma} = 5$.

3.1.2 Quantitative Error Indices

In the following paragraph, we describe the utilized quantitative error indices.

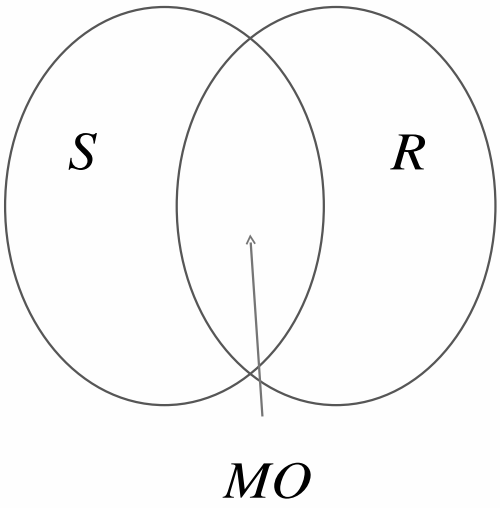


Figure 3.1: Mutual overlap region of segmented region S and ground truth R.

Dice Coefficient:

Dice coefficient, also known as the mutual overlap approach, is an validation method based on computing the overlap area between the ground truth and the segmentation result [10][19][3]. Let S and R be the binary volume of the segmentation result and the ground truth, respectively. MO is the area of their mutual overlap (Fig. 3.1), then Dice coefficient is defined as:

$$s_{dc} = \frac{2MO}{S+R}. \tag{3.1}$$

In our case, S is denoted as the volume-based representation of segmentation result converted from the mesh-based segmentation result M.

Signed and Unsigned Mean Border Positioning Error:

The signed or unsigned mean border positioning error is widely used in image segmen-

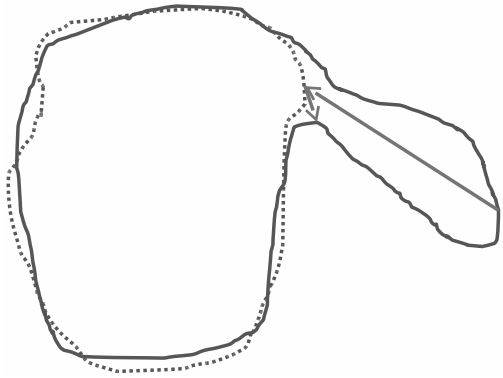


Figure 3.2: Mean border positioning error calculation. Segmented region: dashed line. Ground truth: solid line.

tation validation. Let S_b and R_b respectively denote the surface of the segmentation result and the ground truth. The mean border positioning error is defined as:

$$d_{mean} = \frac{\int_{x \in S_b} d_s(x, R_b)dx + \int_{x \in R_b} d_s(x, S_b)dx}{|S_b| + |R_b|}, \tag{3.2}$$

where $d_s(x, A) = \min_{x^* \in A} d(x, x^*)$ is the distance between a point x and a surface A (Fig. 3.2). $d(x, x^*)$ calculates the Euclidean distance between points in S_b and R_b. For unsigned mean border positioning error, $d(x, x^*)$ is always positive. However, for signed mean border positioning error, $d(x, x^*)$ is positive if the point of the segmentation result surface is located outside of the ground truth boundary, negative otherwise.

Compared to the dice coefficient, the unsigned mean border positioning error focuses on boundary information and can provide information about the bias of the segmentation boundary. However, because it takes the mean value of all positioning errors, a local segmentation error has a little influence on the overall result. Hence, for a complete and objective validation, we also consider the Hausdorff distance.

Hausdorff Distance:

Compared to mean border positioning error, the Hausdorff distance measures the maximum distance between the surface points of segmentation result and the ground truth. The Hausdorff distance [27] is calculated as follows:

$$H(S_b, R_b) = \max\{\max_{x \in S_b} d_s(x, R_b), \max_{x \in R_b} d_s(x, S_b)\}, \tag{3.3}$$

$d_s(x, A)$ is always positive when calculating Hausdorff distance.

3.2 Experiments on Phantom Data

To investigate the ability of our method to deal with different lymph node shapes, we test our method on two sets of computer-generated phantoms. The phantom data set model spherical, ellipsoidal and kidney shaped lymph node and have a data size of $80 \times 80 \times 80$ voxels. The phantom datasets were blurred and superimposed with heavy Gaussian noise.

Fig. 3.3 shows the cross-sections of a set of spherical and ellipsoidal phantom lymph node and their segmentation results. As can be seen in Fig. 3.3(c), the utilized smoothness constraint $\Delta = 4$ enables the correct segmentation of elongated ellipsoid structures. In addition, due to the smooth constraint in combination with the global shape weight, the algorithm is not influenced by nearby strong edge information (Fig. 3.3(e)). Fig. 3.3(f) shows the segmentation result of a phantom lymph node with a concave surface part. As can be seen in Fig. 3.3(f), our method is unable to segment very sharp corners, because of the smoothness constraint. However, such a

case is very unlikely to occur in real patient data.

In Fig. 3.4, the cross-sections of segmentation results of kidney shaped phantom data sets with different size and concavity are shown. As can be seen in Figs. 3.4(a) and 3.4(b), a segmentation error occurs when dealing with very elongated kidney shaped structures, because the cost term w_{shape} biases the segmentation towards a spherical shape. As can be seen from Fig. 3.4(b), week edge information at the ends of long axis of the kidney shaped lymph node amplifies this behavior. However, most of the abnormal lymph nodes are spherical shaped. In the case that such a shaped lymph node needs to be analyzed in practice, the segmentation refinement tool can be used to correct the error as shown in Figs. 3.5(a) and 3.5(b). The refinement result is generated using only one or two 'clicks' on the surface of the kidney shaped structure. In Fig. 3.5(b), the bottom part of the kidney shaped structure is attached to a sphere with the same gray-value range, which results in a local segmentation inaccuracy that was not corrected by the user.

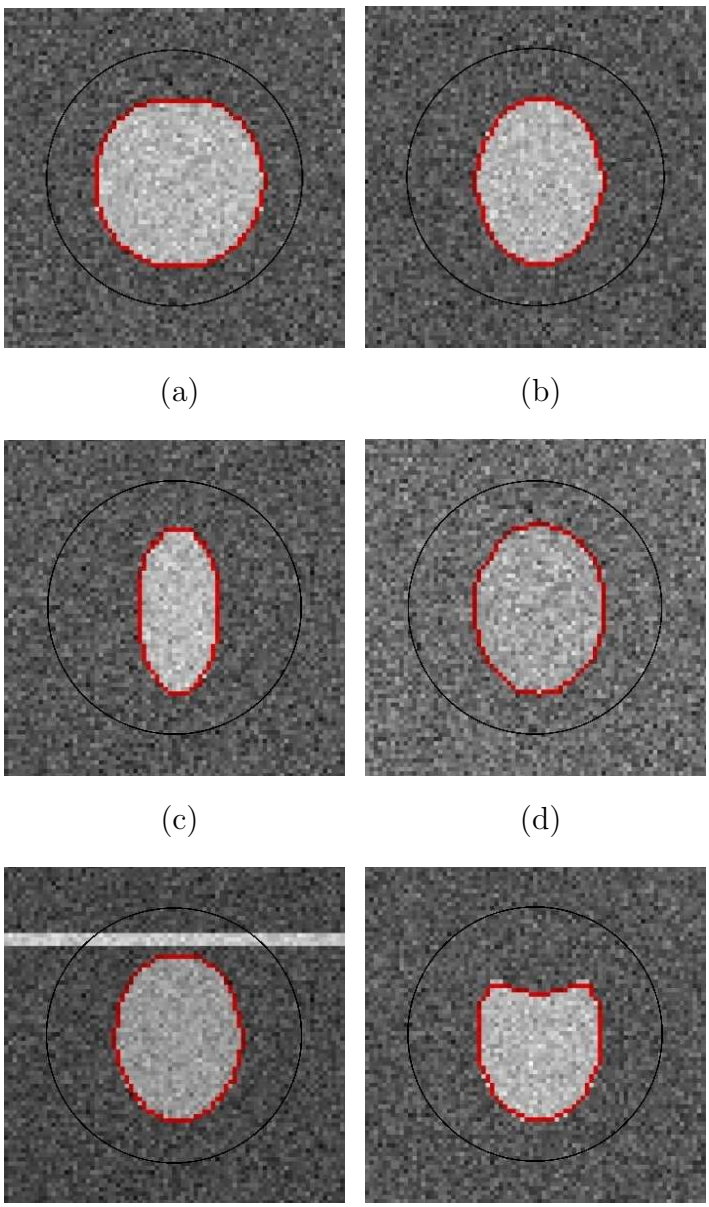


Figure 3.3: Segmentation results of spherical and ellipsoidal phantom data. The black circle indicates the boarder of the spherical graph utilized by our method. (a) Sphere with Gaussian noise and its segmentation result. (b) Ellipsoid with Gaussian noise and its segmentation result. (c) Elongated ellipsoid with Gaussian noise and its segmentation result. (d) Ellipsoid with severe Gaussian noise and its segmentation result. (e) Ellipsoid with strong edge nearby and its segmentation result. (f) Ellipsoid with a concave surface part and its segmentation result.

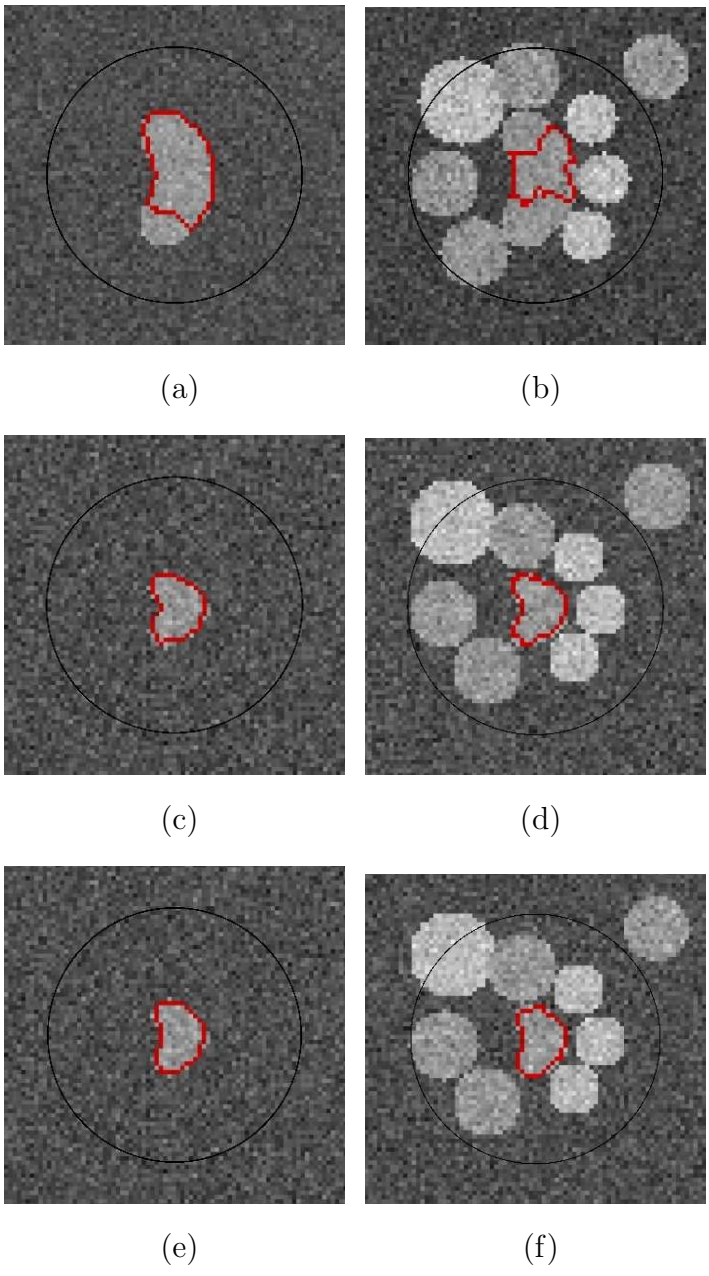
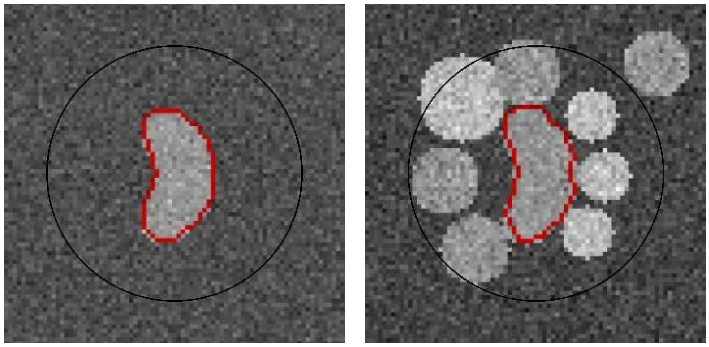


Figure 3.4: Kidney shaped phantom data. Segmentation results of kidney shaped phantom data with clean background (a, c, e) and complex background (b, d, f). The black circles indicates the boarder of the spherical graph utilized by our method.



(a) (b)

Figure 3.5: Refinement results of kidney shaped phantom data. The black circles indicates the boarder of the spherical graph utilized by our method. (a) Refinement result corresponding to Fig. 3.4(a). (b) Refinement result corresponding to Fig. 3.4(b).

3.3 Center Sensitivity

For the initial segmentation, the only input information provided by the user is the center position $\mathbf{c}_k$ of a lymph node to analyze. Obviously, placing $\mathbf{c}_k$ at different positions may lead to different segmentation results. The center position selected by different users might vary. Hence, we investigate the center sensitivity of our method, based on two case studies (Fig. 3.6). Case1 is an elliptical, slightly kidney shaped lymph node, whose long axis length is 14 mm and short axis length is 10 mm. Case2 is an approximately spherical lymph node, whose radius is 4.5 mm.

At first, we generate the center position $\mathbf{c}_k$ by calculating the mean position of all the voxels in the ground truth of the lymph node. Then the method is applied using a shifted center position along the $x-$, $y-$, and $z-$axis from the original center $\mathbf{c}_k$, respectively. Fig. 3.7 shows the gray-value profiles through the center positions for the two cases. The location 0 is the original center position $\mathbf{c}_k$ of the lymph node. Figs. 3.8 and 3.9 show the unsigned mean positioning error and Dice coefficient for

case1 and case2 when the center position is shifted along the $x-$, $y-$ and $z-$axis between the ground truth boundaries. It can be seen from these figures that the error measurement stay approximately in the same value range around the original center position. When the center is shifted towards the ground truth boundary, the segmentation error increases.

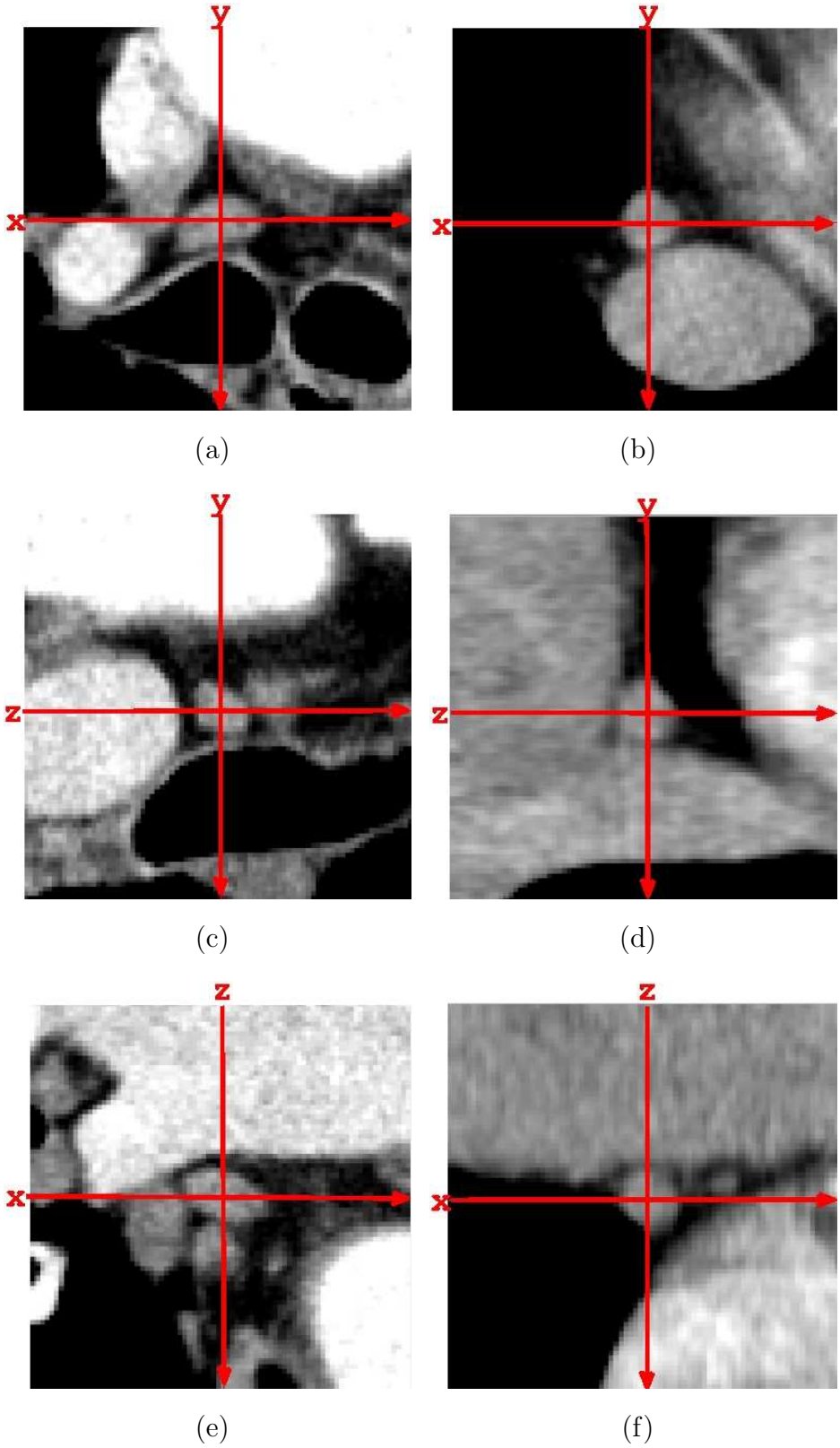


Figure 3.6: Two lymph node cases utilized for center sensitivity analysis. (a, c ,e) axial, coronal and sagittal cross-sections for case1. (b, d ,f) axial, coronal and sagittal cross sections for case2. In all images, the original center point $\mathbf{c}_k$ is located at center of the red cross.

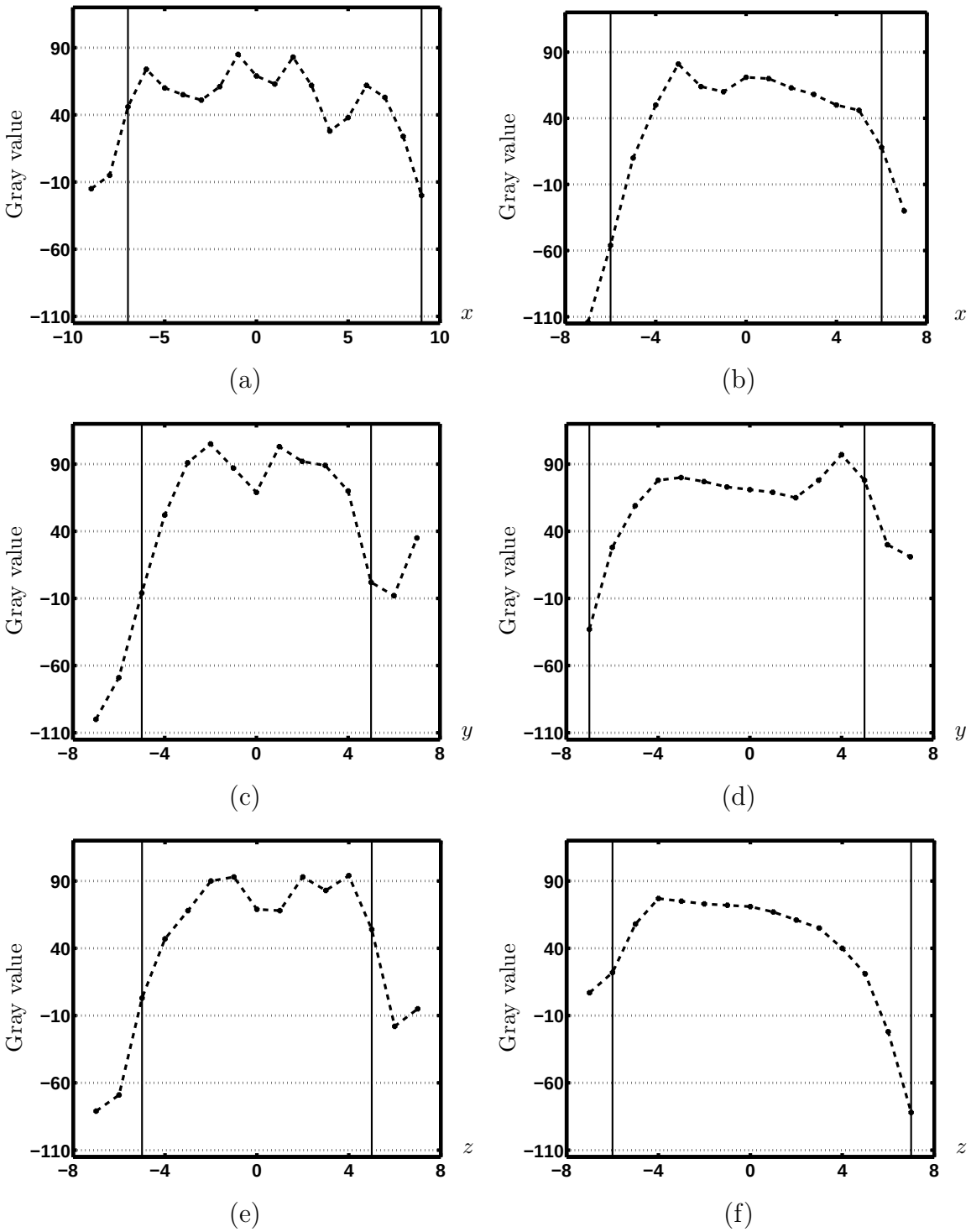


Figure 3.7: Gray-value profile plots for lymph nodes shown in Fig. 3.6. (a, c ,e) Gray-value plots for case1 along the $x-$, $y-$, and $z-$axis, respectively. (b, d ,f) Gray-value plot for case2 along the $x-$, $y-$, and $z-$axis, respectively. The vertical lines indicate the boundary location taken from the corresponding ground truth.

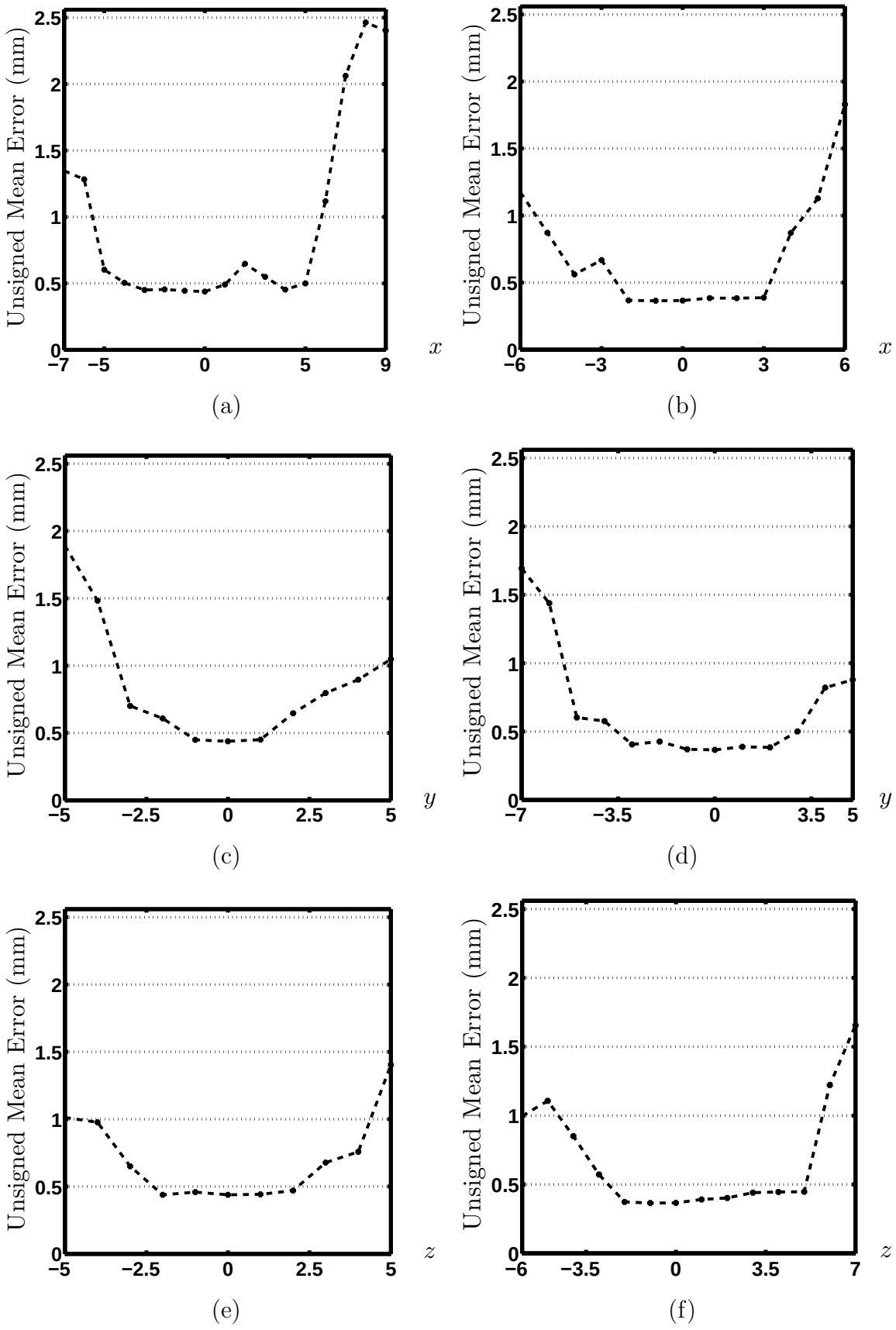


Figure 3.8: Unsigned mean positioning error plot for center sensitivity. Unsigned mean positioning error for case1 (a, c, e) and case2 (b, d ,f) along the $x-$, $y-$, and $z-$axis.

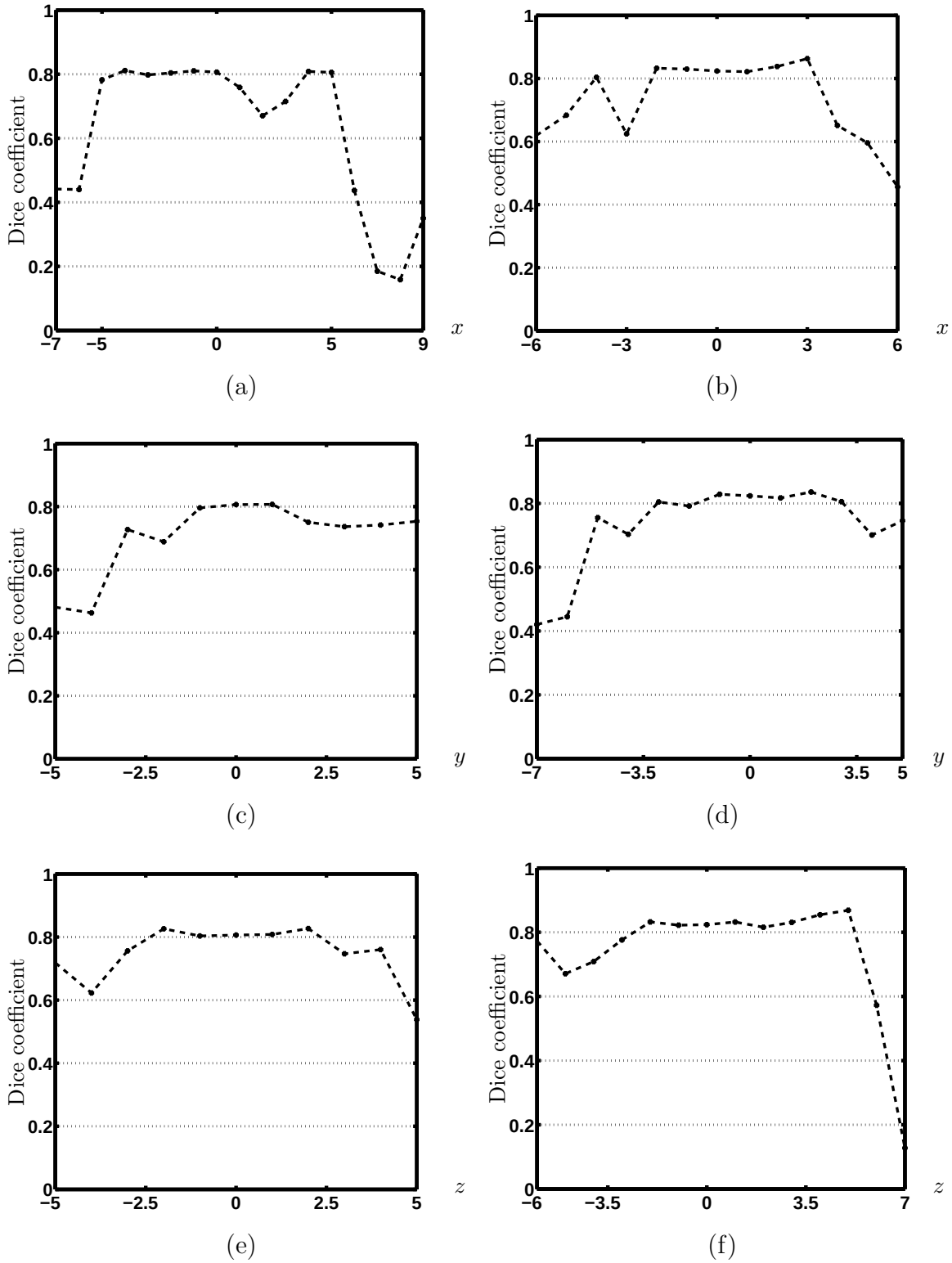


Figure 3.9: Dice coefficient plot for center sensitivity. Dice coefficient for case1 (a, c, e) and case2 (b, d ,f) along the $x-$, $y-$, and $z-$axis.

3.4 Evaluation on CT Data

3.4.1 CT Data Sets

For evaluation, we utilized 35 volumetric CT data sets, showing 111 mostly enlarged lymph nodes of the mediastinum, abdomen, head/neck, and axillary regions. Some of the CT data sets are contrast enhanced. There are three sets of volumetric CT scans. Test set 1 consists of 22 lymph nodes. Scans in test set 1 are from patients with large lung cancer masses. Test set 2 consists of 45 lymph nodes. The scans are taken from cancer patients and were acquired for radiation treatment planning. The resolution for test set 2 is lower compared to test set 1 (Tables 3.2 and 3.3). Test set 3 consists of 44 lymph nodes, and the CT images are from patients with lung nodules. Consequently, the lymph nodes in this set are smaller compared to the other two test sets. Imaging details about these CT data sets are summarized in Tables 3.2 to 3.4. Fig. 3.10 and Table 3.1 show the volume size of lymph nodes utilized for evaluation. The volume size calculation is based on the independent reference segmentation, which is described in the next section.

Table 3.1: Mean and Standard Deviation of volume Size for Each Data Set.

	Test Set 1	Test Set 2	Test Set 3
Mean and Standard Deviation (mm^3)	1056±1611	1621±2064	501±652

Table 3.2: Information about CT Data in the Test Set 1.

CT data set	Intra-slice resolution (mm)	Inter-slice resolution (mm)	Number of lymph node cases	Contrast enhancement (Y/N)
Data1	0.673828	3	10	Y
Data2	0.671875	0.5	2	Y
Data3	0.615234	0.6	10	Y

Table 3.3: Information about CT Data in the Test Set 2.

CT data set	Intra-slice resolution (mm)	Inter-slice resolution (mm)	Number of lymph node cases	Contrast enhancement (Y/N)
Data4	0.976562	2	8	Y
Data5	0.976562	2	9	Y
Data6	0.976562	2	9	Y
Data7	0.976562	2	8	Y
Data8	1.26953	2	3	N
Data9	0.976562	2	1	N
Data10	1.17188	2	1	N
Data11	0.976562	2	2	Y
Data12	0.976562	2	4	N

Table 3.4: Information about CT Data in the Test Set 3.

CT data set	Intra-slice resolution (mm)	Inter-slice resolution (mm)	Number of lymph node cases	Contrast enhancement (Y/N)
Data13	0.652344	0.75	2	Y
Data14	0.654297	1	1	Y
Data15	0.556641	1	3	Y
Data16	0.542969	1	2	Y
Data17	0.78125	0.625	1	N
Data18	0.703125	0.625	1	N
Data19	0.662109	1	1	Y
Data20	0.679688	1	3	Y
Data21	0.625	0.625	1	N
Data22	0.625	0.625	1	N
Data23	0.605468	0.625	1	N
Data24	0.703125	0.625	4	N
Data25	0.742188	0.625	3	N
Data26	0.703125	0.625	5	N
Data27	0.742188	0.625	1	N
Data28	0.703125	0.625	3	N
Data29	0.625	0.625	2	N
Data30	0.605469	0.700012	1	N
Data31	0.8125	1	1	N
Data32	0.625	0.625	3	N
Data33	0.703125	0.625	1	N
Data34	0.625	0.625	1	N
Data35	0.724609	0.75	2	N

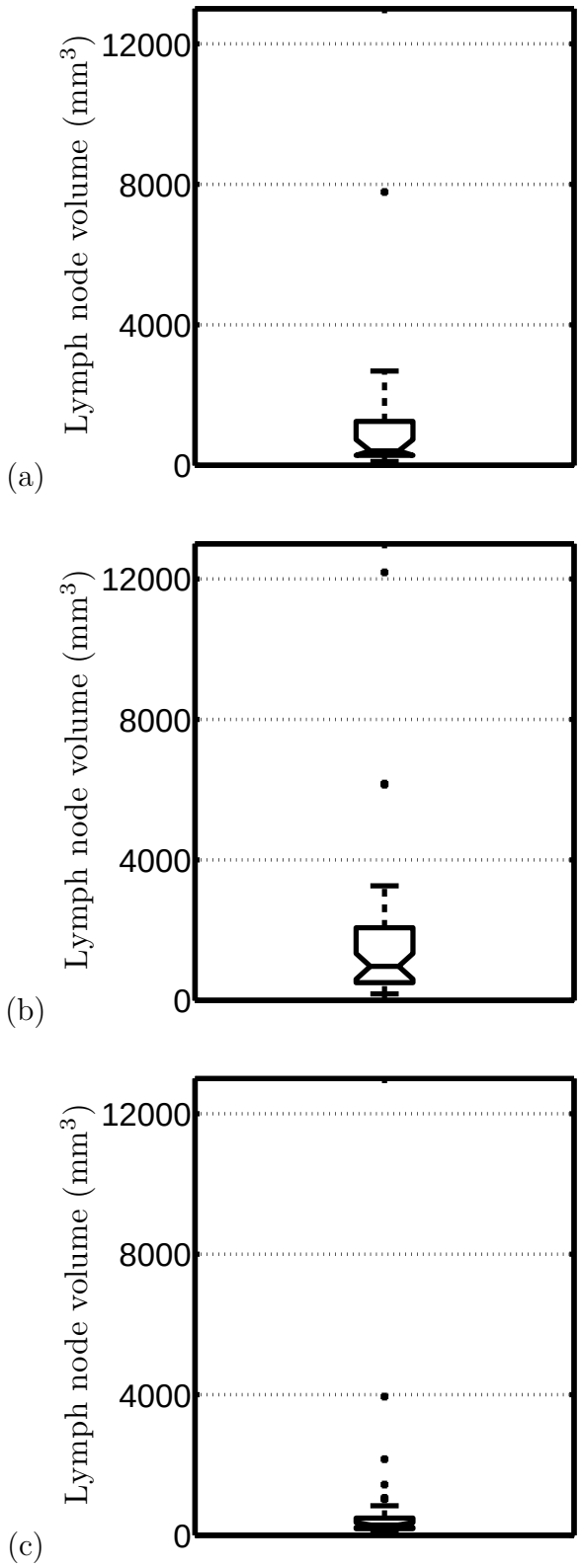


Figure 3.10: Boxplot of lymph node volume for (a) test set 1, (b) test set 2, and (c) test set 3.

3.4.2 Independent Reference

The independent reference was generated by an expert in a slice-by-slice fashion using a semiautomatic live wire [1] segmentation tool. The live-wire method was chosen to reduce time required by the user for generating the reference segmentation. A potential issue of this approach might be that object boundies are identified more consistently compared to completely manual segmentation. On the otherhand, the user has the option to influence the boundary location by using path-cooling technics. The process of this semiautomatic segmentation took approximately 10 minutes per lymph node. The independent reference was utilized for evaluation of the developed lymph node segmentation method.

3.4.3 Experimental Setup

All the segmentations were calculated on a workstation with a 2.40 GHz CPU. Data sets were resampled to isotropic voxel size. The user was asked to select the center of the lymph node, apply the algorithm and inspect the segmentation result. If the result was not correct, the user was required to refine the segmentation.

3.4.4 Results

In our experiments, the manual specification of the approximate centers of a lymph node required typically less than 20 seconds. On average, 5.3 seconds were required for the computation of a lymph node segmentation.

The validation results for the segmentation without refinement on the three test data sets are summarized in Table 3.5(a). Boxplots of validation results are shown

Table 3.5: Average and Standard Deviation of Validation Results on Test Sets.

(a) Before Refinement

Validation method	Result for test set 1	Result for test set 2	Result for test set 3
Intra-slice resolution (mm)	[0.6152, 0.6738]	[0.9766, 1.2695]	[0.5430, 0.8125]
Inter-slice resolution (mm)	[0.5000, 3.0000]	2.0000	[0.6250, 1.0000]
Signed mean border positioning error (mm)	0.0269±0.1776	0.4206±0.2164	-0.0880±0.2167
Unsigned mean border positioning error (mm)	0.5242±0.1485	0.8529±0.1759	0.5724±0.2350
Hausdorff distance (mm)	2.1245±1.0358	3.2588±1.0661	2.5842±1.5348
Dice coefficient	0.8447±0.0592	0.8276±0.0641	0.7747±0.0822

(b) After Refinement

Validation method	Result for test set 1	Result for test set 2	Result for test set 3
Signed mean border positioning error (mm)	0.0233±0.1705	0.3940±0.1889	0.0006±0.1457
Unsigned mean border positioning error (mm)	0.5176±0.1461	0.8241±0.1457	0.5011±0.1571
Hausdorff distance (mm)	2.0200±0.9964	3.0558±0.8990	1.9932±0.9433
Dice coefficient	0.8469±0.0607	0.8361±0.0584	0.8094±0.0700

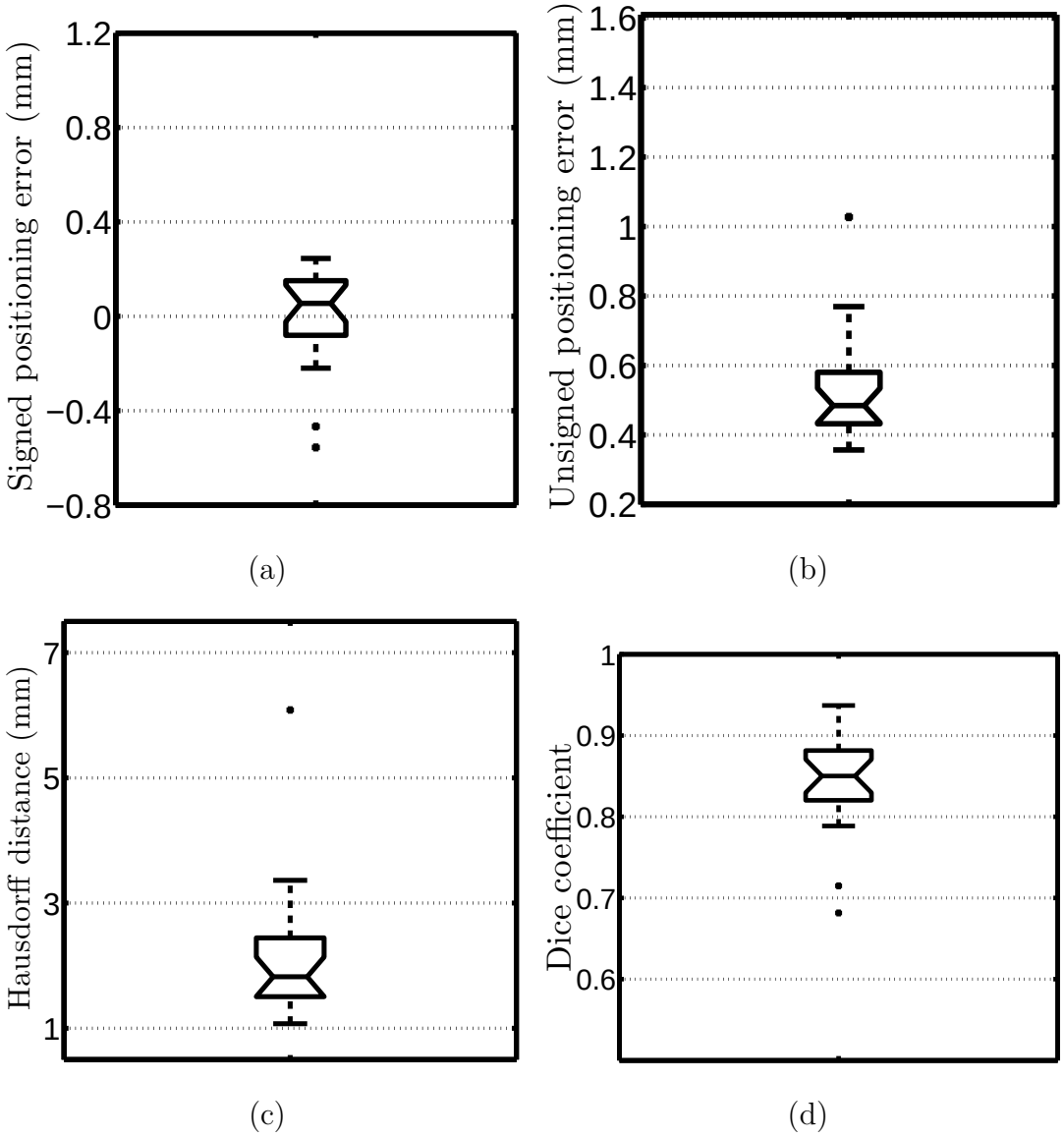


Figure 3.11: Validation results for Test Set 1. (a) Signed mean border positioning error. (b) Unsigned mean border positioning error. (c) Hausdorff distance. (d) Dice coefficient.

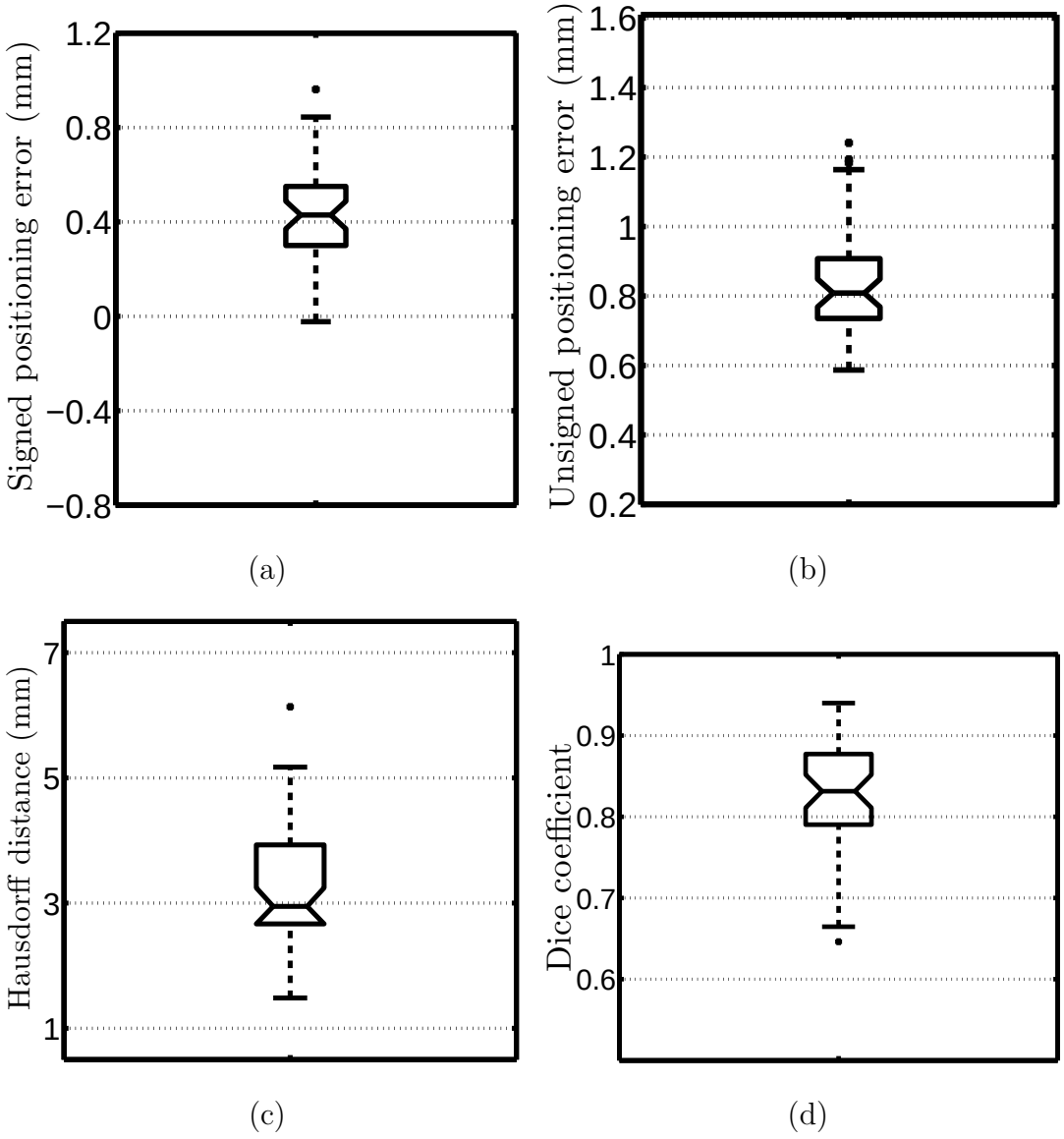


Figure 3.12: Validation results for Test Set 2 (a) Signed mean border positioning error. (b) Unsigned mean border positioning error. (c) Hausdorff distance. (d) Dice coefficient.

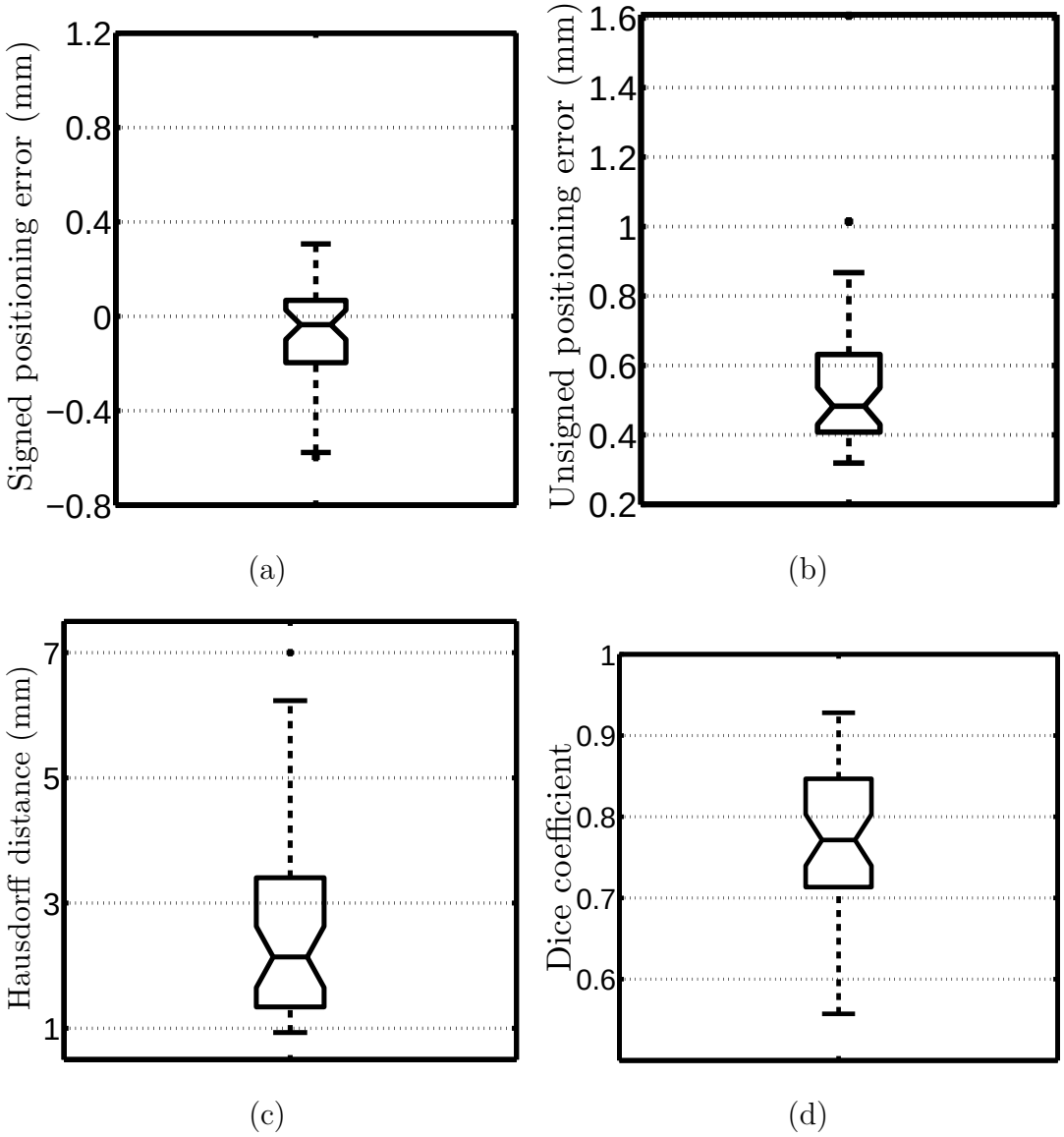


Figure 3.13: Validation results for Test Set 3 (a) Signed mean border positioning error. (b) Unsigned mean border positioning error. (c) Hausdorff distance. (d) Dice coefficient.

in Figs. 3.11 to 3.13, respectively. 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 0.4206 mm. The image resolution for test set 2 is low. The ideal resolution for lymph node segmentation is less than 1 mm in both intra-slice resolution and inter-slice resolution. For volumetric CT data with lower resolution, lymph node boundaries are getting blurred and the segmentation performance is impacted by the partial volume effect. 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 1 voxel size.

Generally, a refinement iteration ('click') can be done in less than 5 seconds of user interaction. For test set 1, 4 lymph node were refined. The average number of refinement 'clicks' was 1.25. For test set 2, 16 lymph nodes were refined. The average number of 'clicks' was 1.62. For the test set 3, 20 cases were refined, and the average number of refinement 'clicks' was 1.55. The average user interaction time required for refinement was approximately 10 seconds. Consequently, a lymph node can be segmented in less than a minute.

Table 3.5(b) also depicts the overall validation result after refinement. For those lymph nodes which needed refinement, Tables 3.6 to 3.8 show the average validation results before and after refinement for the three test sets. Corresponding boxplots are shown in Figs. 3.14 to 3.16. Plots of validation results before and after refinement for each case are shown in Fig. 3.18 to 3.20. The individual cases are sorted by their volume size from small to large (Fig. 3.17).

Table 3.6: Comparison of Validation Results Before and After Refinement on Test Set 1 for Cases that Needed Refinement (Intra-slice resolution: [0.6152, 0.6738] mm, inter-slice resolution: [0.5000, 3.0000] mm).

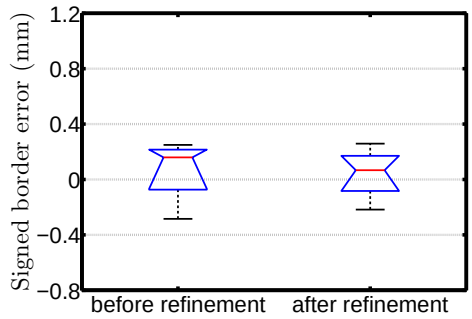
validation method	result before refinement	result after refinement
signed mean border positioning error (mm)	0.0710±0.2091	0.0439±0.1702
unsigned mean border positioning error (mm)	0.6070±0.0498	0.5709±0.0596
Hausdorff distance (mm)	2.6509±0.4301	2.1296±0.3263
Dice coefficient	0.8831±0.0218	0.8936±0.0197

Table 3.7: Comparison of Validation Results Before and After Refinement on Test Set 2 for Cases that Needed Refinement (Intra-slice resolution: [0.9766, 1.2695] mm, inter-slice resolution: 2.0000 mm).

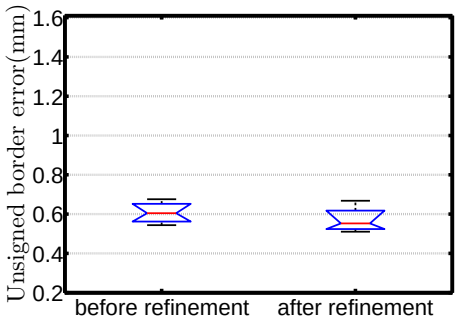
validation method	result before refinement	result after refinement
signed mean border positioning error (mm)	0.5595±0.2332	0.4845±0.2006
unsigned mean border positioning error (mm)	1.0005±0.1656	0.9197±0.1407
Hausdorff distance (mm)	4.0574±0.7422	3.4865±0.5737
Dice coefficient	0.7956±0.0791	0.8196±0.0737

Table 3.8: Comparison of Validation Results Before and After Refinement on Test Set 3 for Cases that Needed Refinement (Intra-slice resolution: [0.5430, 0.8125] mm, inter-slice resolution: [0.6250, 1.0000] mm).

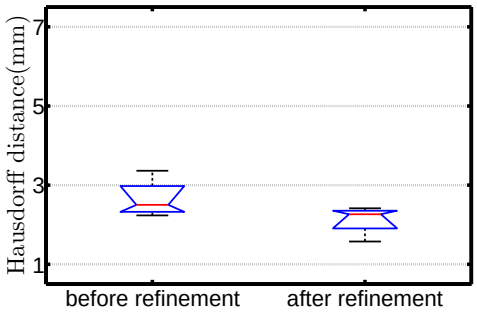
validation method	result before refinement	result after refinement
signed mean border positioning error (mm)	-0.1835±0.2569	0.0113±0.1609
unsigned mean border positioning error (mm)	0.7234±0.2448	0.5664±0.1635
Hausdorff distance (mm)	3.8493±1.2533	2.5492±0.8447
Dice coefficient	0.7360±0.0800	0.8124±0.0714



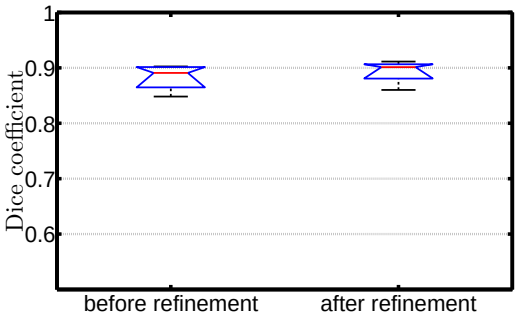
(a)



(b)

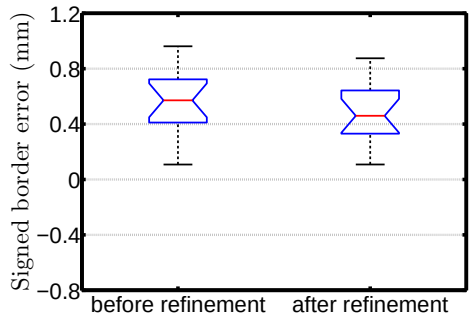


(c)

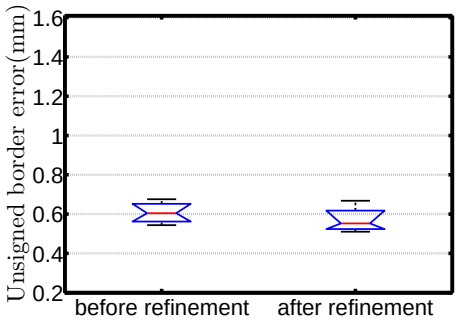


(d)

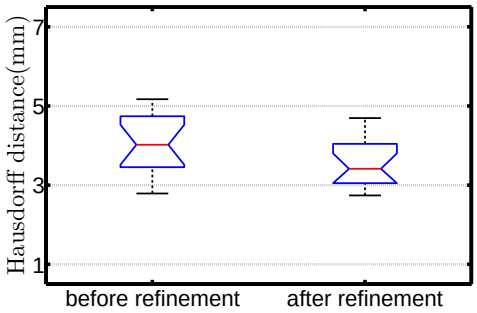
Figure 3.14: Boxplot for Test Set 1 before and after Refinement. (a) Signed mean border positioning error. (b) Unsigned mean border positioning error. (c) Hausdorff distance. (d) Dice coefficient.



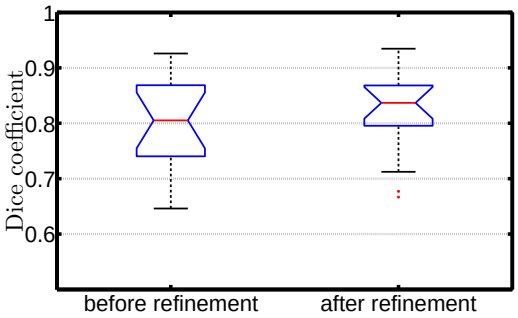
(a)



(b)

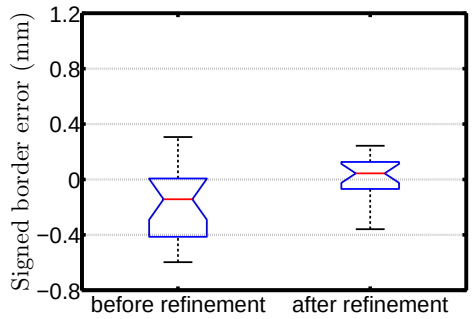


(c)

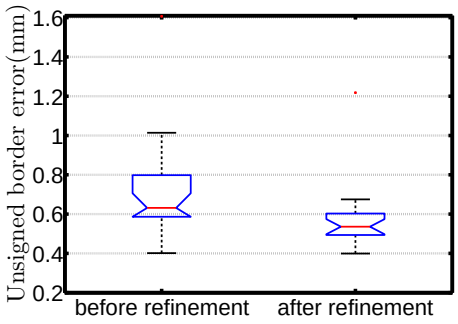


(d)

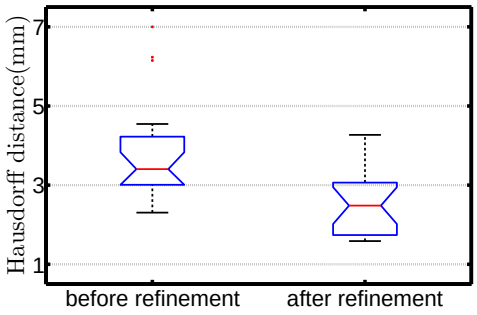
Figure 3.15: Boxplot for Test Set 2 before and after Refinement. (a) Signed mean border positioning error. (b) Unsigned mean border positioning error. (c) Hausdorff distance. (d) Dice coefficient.



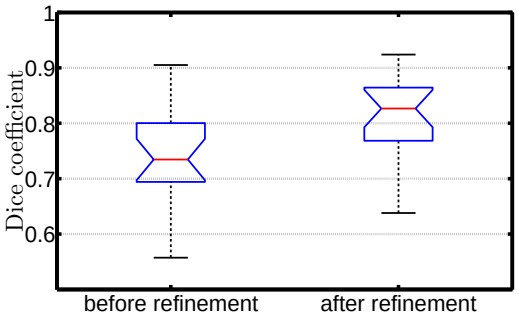
(a)



(b)

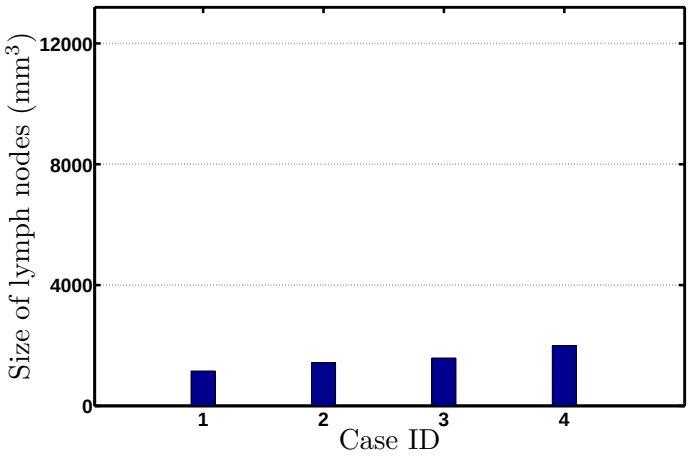


(c)

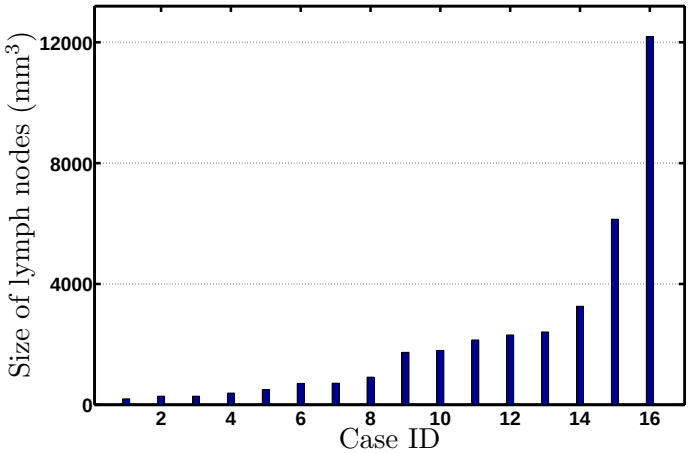


(d)

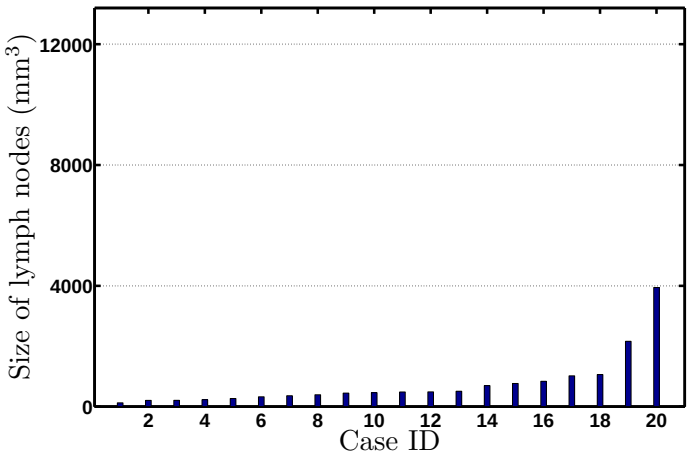
Figure 3.16: Boxplot for Test Set 3 before and after Refinement. (a) Signed mean border positioning error. (b) Unsigned mean border positioning error. (c) Hausdorff distance. (d) Dice coefficient.



(a)

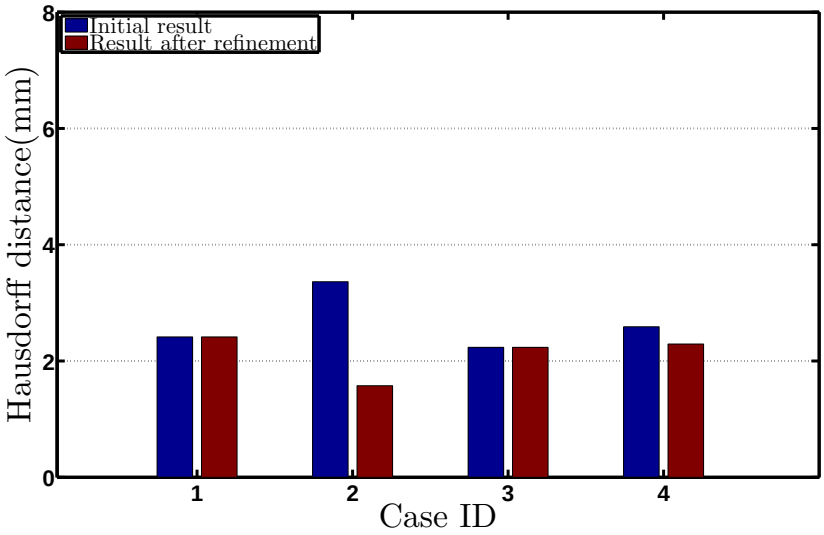


(b)

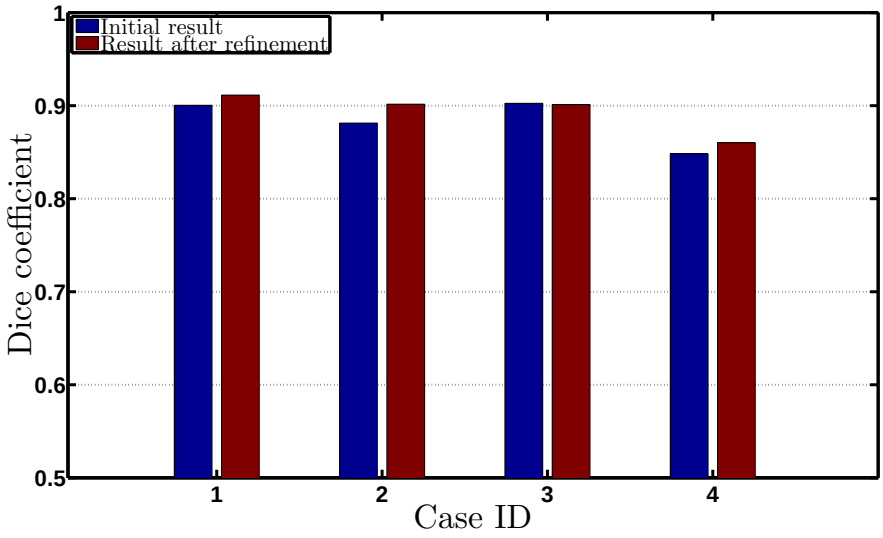


(c)

Figure 3.17: volume size of refined lymph nodes cases. (a) Test set 1. (b) Test set 2. (c) Test set 3.

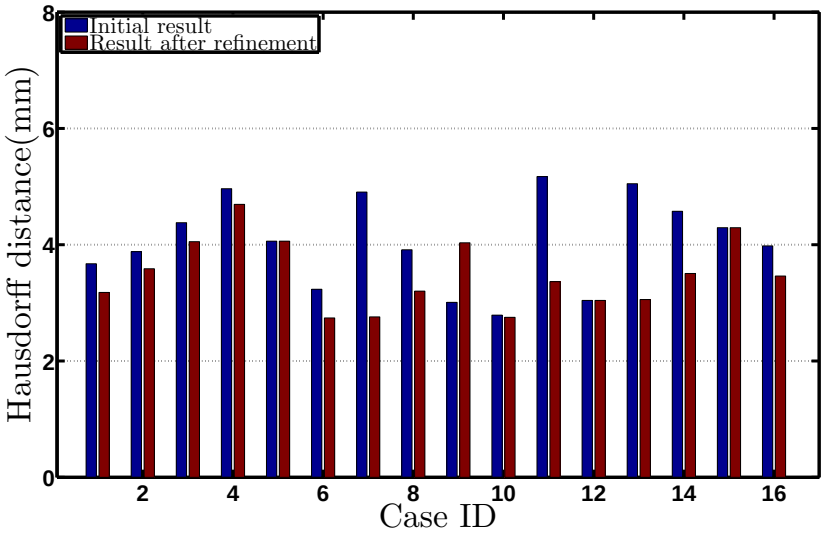


(a)

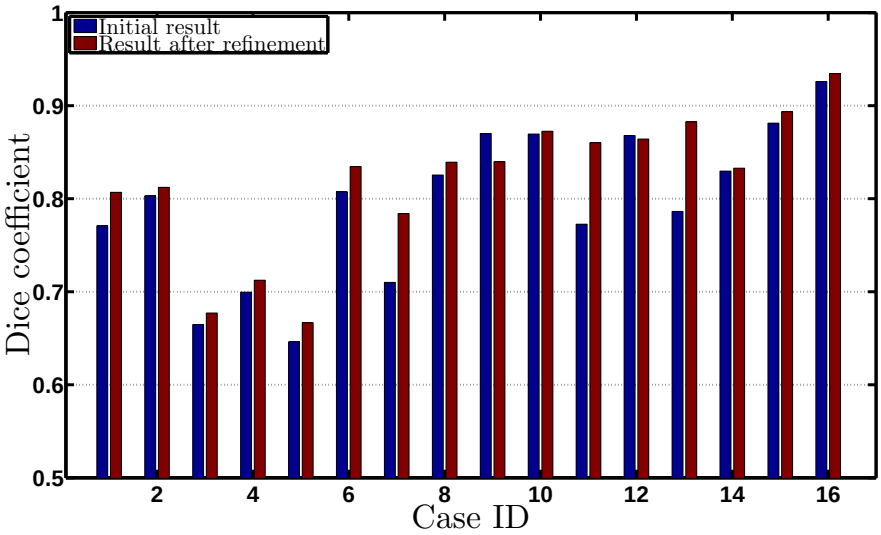


(b)

Figure 3.18: Validation results before and after refinement of test set 1 (a) Hausdorff distance (b) Dice coefficient.

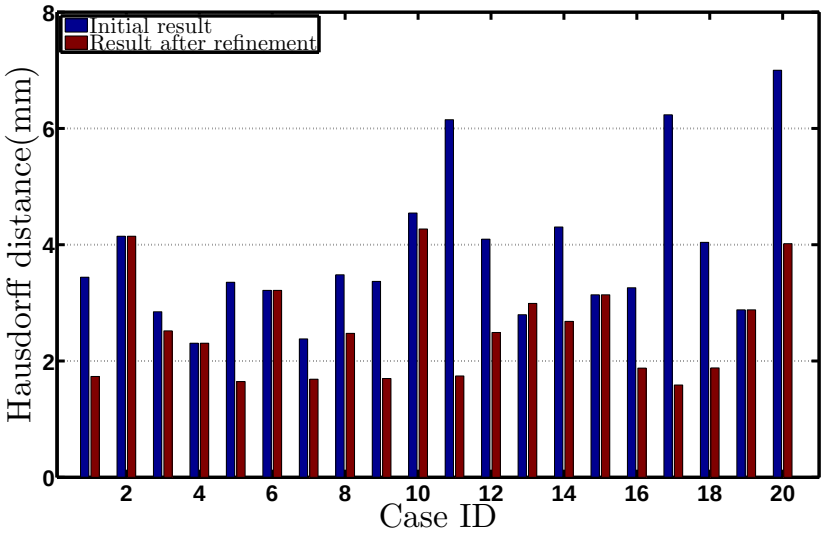


(a)

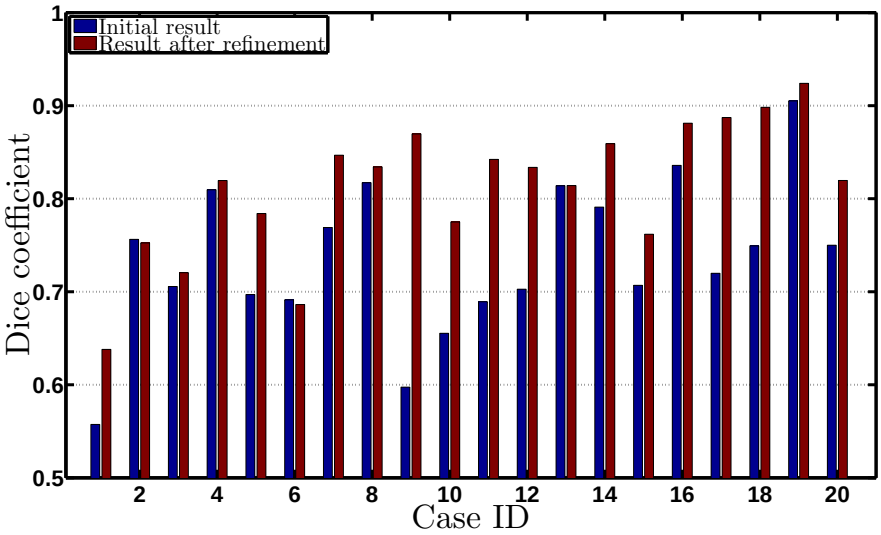


(b)

Figure 3.19: Validation results before and after refinement of test set 2 (a) Hausdorff distance (b) Dice coefficient.



(a)



(b)

Figure 3.20: Validation results before and after refinement of test set 3 (a) Hausdorff distance (b) Dice coefficient.

CHAPTER 4
DISCUSSION

In this section, we will discuss the performance of our method and analyze individual cases to demonstrate the properties of our method.

It is challenging to develop an algorithm that can perform correct lymph node segmentation in CT images under all kinds of conditions (imaging, anatomical variation, disease, etc.). Examples of the range of variation in shape, size, and imaging conditions are shown in Figs. 4.1(a, c, e, and g), and the segmentation results produced with our method are presented in Figs. 4.1(b, d, f, and h). Note that the results were produced without utilizing any segmentation refinement iterations. It can be seen that our method deals successfully with the variations in lymph node shape depicted in Fig. 4.1.

Traditional edge based segmentation methods fail when there are strong edges around a lymph node, as we have shown in Fig. 2.1. In our method, the influence of neighboring strong edges is reduced by utilizing a weighted edge and region homogeneity term in our cost function (Fig. 4.2(a)). Figs. 4.2 and 4.3 provide an comparison between our segmentation results (without refinement) and the ground truth. Both figures show mostly local differences between our segmentation results and the independent reference.

A comparison of test set 1 and 2 (Tables 3.5) 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 set 1 and 3 can mainly be observed for the Dice coefficient,

which is impacted by the size of the analyzed lymph nodes (Fig. 3.10).

In our cost function we assume that lymph nodes have a spherical or slightly ellipsoid shape. However, as can be seen in Fig. 4.4, some lymph nodes deviate from this assumption, which can cause segmentation errors. Our segmentation approach might try to provide a more spherical shaped result (e.g., Fig. 4.4(e)), in the case of elongated inhomogeneous lymph nodes. Also, noise and imaging artifacts can lead to similar issues.

If the initial segmentation contains errors, the user can utilize the developed segmentation refinement tool to correct the segmentation result, which takes only a few seconds of user interaction. Examples for initially incorrect lymph node segmentations and results of the refinement process are presented in Fig. 4.4.

In our evaluation on 111 lymph nodes, the user decided to refine 36 % (40 cases) of the segmentation results. Our segmentation refinement method is quite effective, as can be seen from Tables 3.6 to 3.8 and Figs. 3.14 to 3.16. For all the cases that needed to be corrected, less than 2 refinement iterations were required on average. Thus, the time needed to perform a lymph node segmentation (including refinement) is less than one minute. In comparison, generating the live wire based slice-by-slice reference segmentation did require 10 minutes per lymph node, which is one order of a magnitude more than the worst-case scenario when performing the segmentation with our approach.

Analyzing individual refinement cases based on the plots in Figs. 3.18 to 3.20 leads to the following conclusions. First, for almost all cases, segmentation performance measures improve with refinement. Second, letting the user refine a segmentation can introduce operator variability. For example, in Case9 (test set 2) the

Hausdorff distance and Dice coefficient get worse after the refinement (Figs. 3.19(a) and 3.19(b)). Cross-sectional images for this case are shown in Fig. 4.5(a), (c), and (e). Because the resolution of data set is low, it is difficult for the user to locate the correct lymph node boundary. Thus, low resolution scans of lymph nodes will likely introduce additional inter-and-intra observer variability and segmentation errors. Third, for almost all cases the average Dice coefficient increases with refinement and has values larger than 0.8, but for some cases the Dice coefficient stays very low, like for example case1 shown in Fig. 3.20. This case is also depicted in Fig. 4.5(b), (b) and (f). The size of this lymph node is very small (see Fig. 3.17). Consequently, a 'small' segmentation inaccuracy will impact the Dice coefficient more than for larger lymph nodes.

The developed segmentation refinement framework is quite intuitive to use and delivers expected results. One exception (case2 in Fig. 3.18) is depicted in Fig. 4.6. In this case an 'unexpected' additional refinement 'click' was required (Fig.4.6(e)), because the local refinement led to an alteration of a correctly segmented local region of the lymph node. The reason for this unexpected behavior is as follows. The lymph node is quite inhomogeneous and the gray-value pattern at the user selected refinement 'click' is similar to some profiles inside the lymph node. Because of the atypical shape of this lymph node and the utilized search for neighboring similar columns (Section 2.5, Step II), the cost functions for columns in a correctly segmented local region were updated. The error was corrected by providing a second refinement 'click', which changed the cost for the affected columns.

Overall, our approach to lymph node segmentation is quite powerful and allows the user produce segmentations that are suitable for quantification of lymph node size.

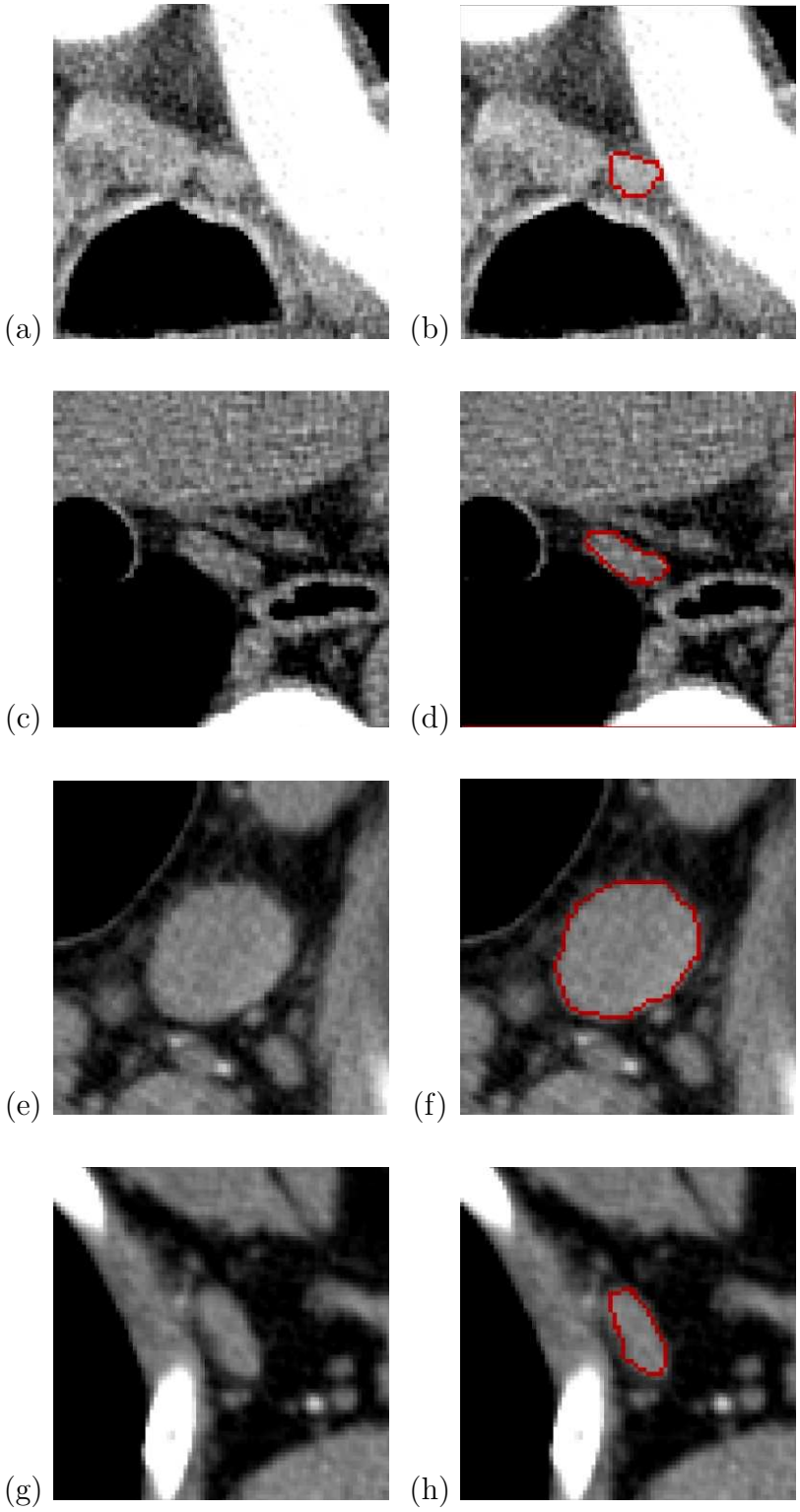


Figure 4.1: More examples of segmentation results. Examples depicting the variation in size, shape, and texture of lymph nodes (a, c, e and g). (b, d, f and h) Segmentation results generated with our approach. None of the results needed refinement by the user.

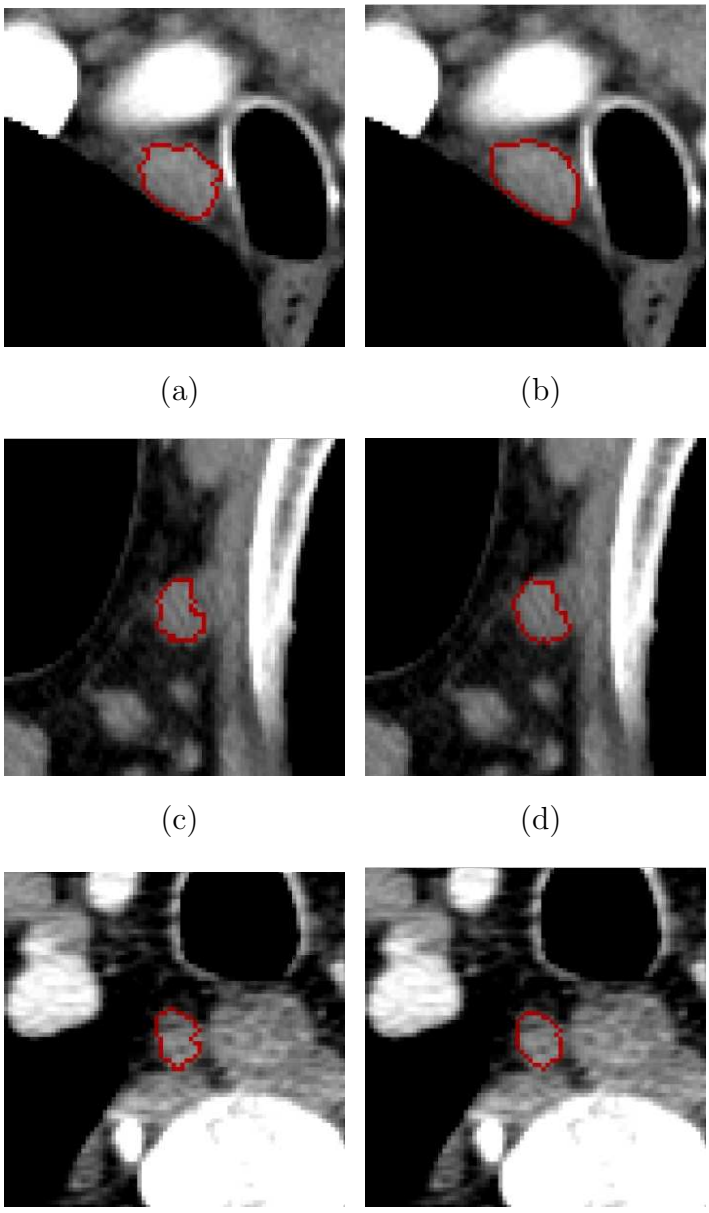
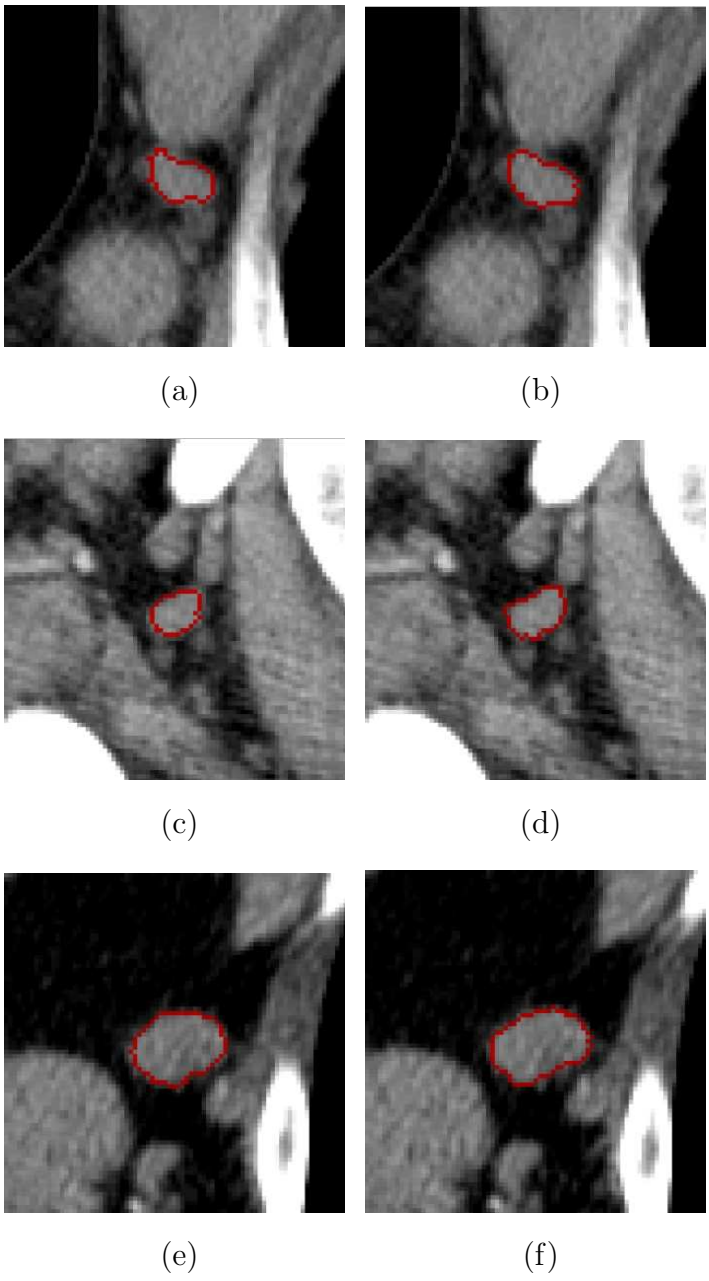


Figure 4.2: Segmentation results for corresponding cases in Fig. 1.1 and ground truth. (a, c, and e) Segmentation results for corresponding cases in Fig. 1.1 and (b, d, and f) ground truth.

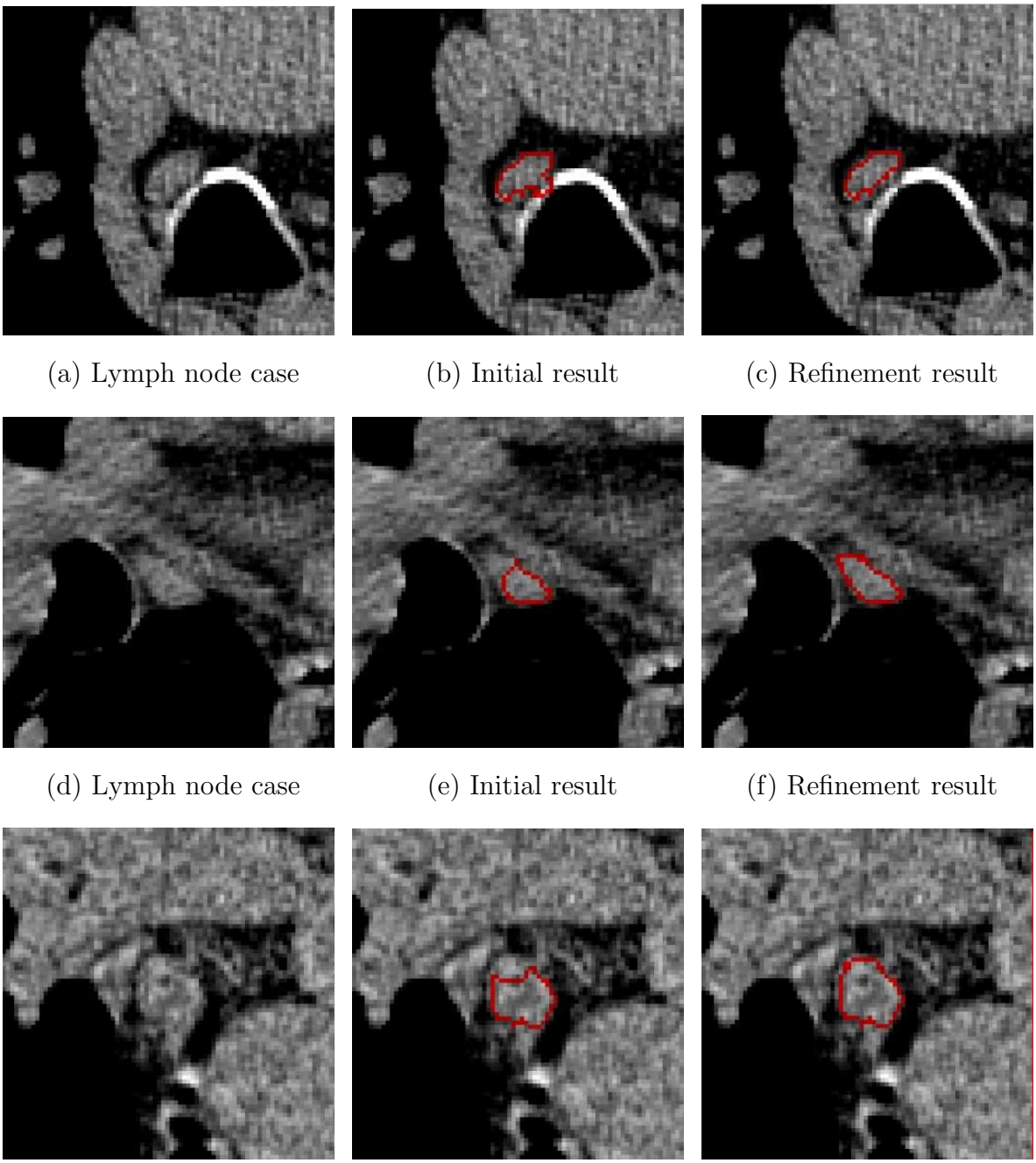


(a) (b)

(c) (d)

(e) (f)

Figure 4.3: Segmentation results for corresponding cases in Fig. 1.1 and ground truth. (a, c, and e) Segmentation results for corresponding cases in Fig. 1.1 and (b, d, and f) ground truth.



(a) Lymph node case (b) Initial result (c) Refinement result

(d) Lymph node case (e) Initial result (f) Refinement result

(g) Lymph node case (h) Initial result (i) Refinement result

Figure 4.4: Examples of basic segmentation and refinement results.

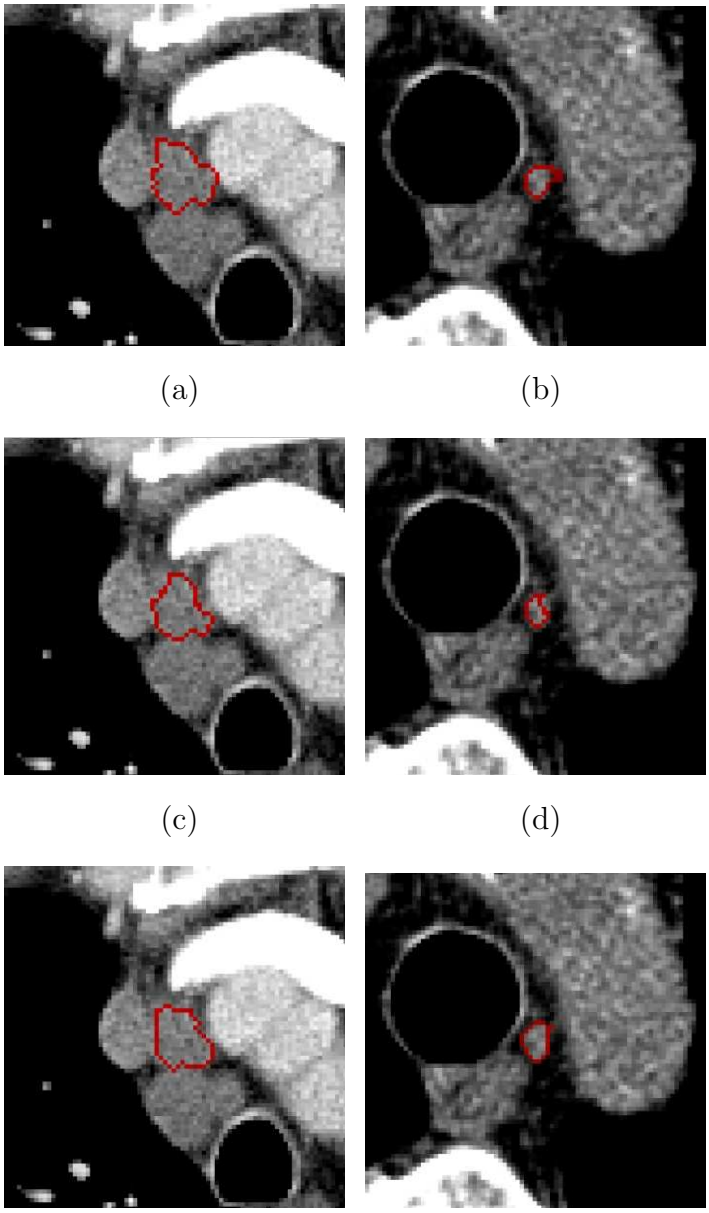
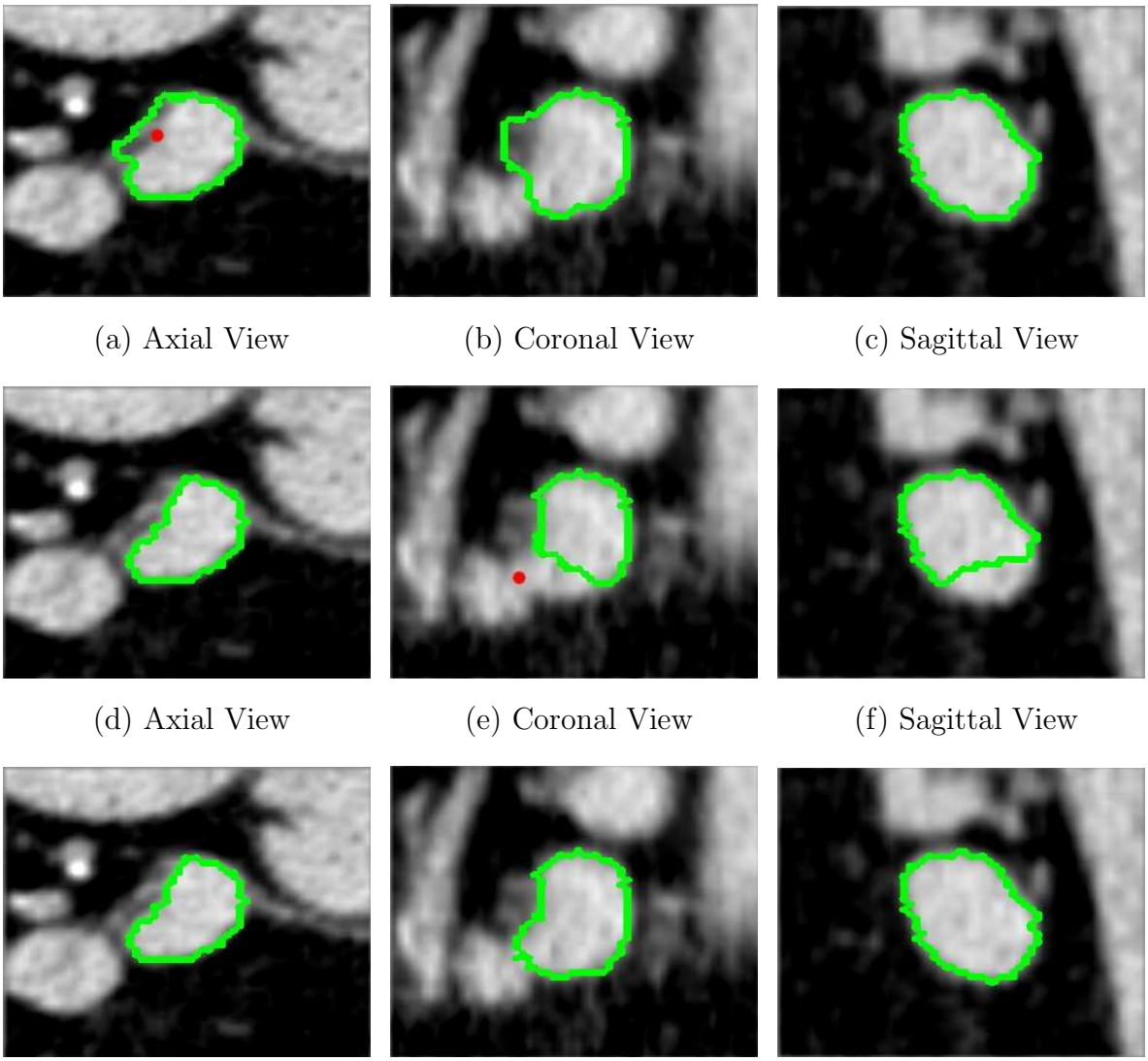


Figure 4.5: Segmentation results of lymph nodes before refinement (a and b), after refinement (c and d) and corresponding independent reference (e and f).



(a) Axial View (b) Coronal View (c) Sagittal View

(d) Axial View (e) Coronal View (f) Sagittal View

(g) Axial View (h) Coronal View (i) Sagittal View

Figure 4.6: An example of segmentation refinement. (a, b and c) Initial segmentation result. (d, e and f) Segmentation result after first refinement 'click'. (g, h and i) Segmentation result after second refinement 'click'. Red dots indicate the positions of the refinement 'clicks'.

CHAPTER 5
CONCLUSION AND FUTURE WORK

The quantitative assessment of lymph nodes plays an important role for diagnosis, monitoring, and treatment of diseases, like cancer. In this thesis, we presented a graph-based approach to enable the computer-aided 3D segmentation of lymph nodes. Segmentation of lymph node in volumetric CT images remains a challenging task in clinical routine. Shortcomings exist in currently available automated segmentation methods due to partial volume effects, diffuse edges, neighboring structures with similar intensity profiles and potentially inhomogeneous density-values. Our approach is able to deal with common segmentation problems by utilizing a weighted edge and region homogeneity term in our cost function. In the majority of cases investigation, the initial segmentation method delivered useful results. 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 only a few seconds of user interaction. The overall process of our segmentation approach, including locating the lymph node center, segmenting the lymph node, and refining the result, if needed, can be done in less than a minute. Compared to manual or semi-automated slice-by-slice segmentation approaches, which takes about ten minutes per case, our method is quite fast and offers similar flexibility when dealing with difficult lymph node segmentation.

There are two limitations of our approach. First, problems exist when dealing with low resolution volumetric CT data. Since the lymph nodes are quite small structures, lymph node boundaries are getting blurred because of the low resolution

and the segmentation performance is impacted by partial volume effects. Second, for some elongated inhomogeneous lymph nodes, our initial segmentation might provide a more spherical segmentation result, but this issue can be corrected in a refinement step.

At last, we point out some directions of research that can be studied in the future. First, since our approach is based on an approximate center position of lymph node, which needs to be specified by the user, it is desirable to develop a fully automated lymph node detection method that could be utilized for initialization of our method. Second, after each refinement iteration, our algorithm performs a new global graph search. Thus, it would be of interest to develop an approach that only requires a locally performed graph search to save computing time.

REFERENCES

[1] W.A. Barrett and E.N. Mortensen. Interactive live-wire boundary extraction. *Medical Image Analysis*, 1(4):331–341, 1997.

[2] E. Bollen, R. Goei, B.E. Hof-Grootenboer, C.W.M. Versteege, H.A. Engelshove, and R.J.S. Lamers. Interobserver variability and accuracy of computed tomographic assessment of nodal status in lung cancer. *The Annals of Thoracic Surgery*, 58(1):158–162, 1994.

[3] K.W. Bowyer. Validation of medical image analysis techniques. *Handbook of Medical Imaging*, 2:567–607, 2000.

[4] Y. Boykov and G. Funka-Lea. Graph cuts and efficient N-D image segmentation. *International Journal of Computer Vision*, 70(2):109–131, 2006.

[5] Y. Boykov and M.P. Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In *Proc. International Conference on Computer Vision (ICCV)*, volume 1935, pages 105–112, 2001.

[6] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, 2004.

[7] J.N. Bruneton, E. Caramella, M. Hery, D. Aubanel, JJ Manzino, and JL Picard. Axillary lymph node metastases in breast cancer: preoperative detection with US. *Radiology*, 158(2):325–326, 1986.

[8] C.L. Carter, C. Allen, and D.E. Henson. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer*, 63(1):181–187, 1989.

[9] J. Cordes, J. Dornheim, B. Preim, I. Hertel, and G. Strauss. Preoperative segmentation of neck CT datasets for the planning of neck dissections. In *Proc. SPIE Medical Imaging*, volume 6144, 2006.

[10] L.R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[11] R.E. Dorfman, M.B. Alpern, B.H. Gross, and M.A. Sandler. Upper abdominal lymph nodes: criteria for normal size determined with CT. *Radiology*, 180(2):319–322, 1991.

[12] J. 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. *Academic Radiology*, 14(11):1389–1399, 2007.

[13] D.M. Einstein, A.A. Singer, W.A. Chilcote, and R.K. Desai. Abdominal lymphadenopathy: spectrum of CT findings. *Radiographics*, 11(3):457–472, 1991.

[14] J. Feu, F. Tresserra, R. Fábregas, B. Navarro, P.J. Grases, J.C. Suris, A. Fernández-Cíd, and X. Alegret. Metastatic breast carcinoma in axillary lymph nodes: in vitro US detection. *Radiology*, 205(3):831–835, 1997.

[15] R. Glynne-Jones, A.R. Makepeace, M.F. Spittle, and W.R. Lees. A possible role for ultrasound of the axilla in staging primary breast cancer. *Clinical Oncology*, 2(1):35–38, 1990.

[16] D.M. Greig, B.T. Porteous, and A.H. Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B (Methodological)*, 51(2):271–279, 1989.

[17] D.M. Honea, Y. Ge, W.E. Snyder, P.F. Hemler, and D.J. Vining. Lymph node segmentation using active contours. In *Proceedings of SPIE*, volume 3034, pages 265–273, 1997.

[18] D.M. Honea and W.E. Snyder. Three-dimensional active surface approach to lymph node segmentation. In *Proceedings of SPIE*, volume 3661, pages 1003–1011, 1999.

[19] A. Hoover, G. Jean-Baptiste, X. Jiang, P.J. Flynn, H. Bunke, D.B. Goldgof, K. Bowyer, D.W. Eggert, A. Fitzgibbon, and R.B. Fisher. An experimental comparison of range image segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(7):673–689, 1996.

[20] T. Kitasaka, Y. Tsujimura, Y. Nakamura, K. Mori, Y. Suenaga, M. Ito, and S. Nawano. Automated extraction of lymph nodes from 3-D abdominal CT images using 3-D minimum directional difference filter. In *Proceedings of the 10th International Conference on Medical Image Computing and Computer-assisted Intervention*, pages 336–343, 2007.

[21] K. Li, X. Wu, D.Z. Chen, and M. Sonka. Efficient optimal surface detection: Theory, implementation and experimental validation. In *Proc. SPIE Medical Imaging*, volume 5370, pages 620–627, 2004.

[22] K. Li, X. Wu, D.Z. Chen, and M. Sonka. Optimal surface segmentation in volumetric images-a graph-theoretic approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 119–134, 2006.

[23] K. Lu, S.A. Merritt, and W.E. Higgins. Extraction and visualization of the central chest lymph-node stations. In *Proceedings of SPIE*, volume 6915, pages 69151B1–15, 2008.

[24] D. Maleike, M. Fabel, R. Tetzlaff, H. von Tengg-Kobligk, T. Heimann, H.P. Meinzer, and I. Wolf. Lymph node segmentation on CT images by a shape model guided deformable surface method. In *Proc. SPIE Medical Imaging*, volume 6914, 2008.

[25] M. Pamilo, M. Soiva, and EM Lavast. Real-time ultrasound, axillary mammography, and clinical examination in the detection of axillary lymph node metastases in breast cancer patients. *Journal of Ultrasound in Medicine*, 8(3):115–120, 1989.

[26] B. Reitinger, A. Bornik, and R. Beichel. Efficient volume measurement using voxelization. In *Proceedings of the 19th Spring Conference on Computer Graphics*, pages 47–54. ACM, 2003.

[27] G. Rote. Computing the minimum Hausdorff distance between two point sets on a line under translation. *Information Processing Letters*, 38(3):123–127, 1991.

[28] P.A. Schnyder and G. Gamsu. CT of the pretracheal retrocaval space. *American Journal of Roentgenology*, 136(2):303–308, 1981.

[29] L.H. Schwartza, J. Bogaertsb, R. Fordc, L. Shankard, P. Therassee, S. Gwytherf, and E.A. Eisenhauerg. Evaluation of lymph nodes with RECIST 1.1. *European Journal of Cancer*, 45:261–267, 2009.

[30] M. Sonka, V. Hlavac, and R. Boyle. Image processing, analysis, and machine vision. page 308, 1999.

[31] H.J. Steinkamp, M. Cornehl, N. Hosten, W. Pegios, T. Vogl, and R. Felix. Cervical lymphadenopathy: ratio of long-to-short axis diameter as a predictor of malignancy. *British Journal of Radiology*, 68(807):266–270, 1995.

[32] N. Tohnosu, K. Okuyama, Y. Koide, T. Kikuchi, T. Awano, H. Matsubara, T. Sano, H. Nakaichi, Y. Funami, K. Matsushita, et al. A comparison between ultrasonography and mammography, computed tomography and digital subtraction angiography for the detection of breast cancers. *Surgery Today*, 23(8):704–710, 1993.

[33] R. Warwick, P.L. Williams, M. Dyson, and L. Bannister. Gray's anatomy. *British Journal of Radiology*, 58(694):1029, 1985.

[34] X. Wu and D. Chen. Optimal net surface problems with applications. *Automata, Languages and Programming*, 2380:775–775, 2002.

[35] J. Yan, T. Zhuang, B. Zhao, and L.H. Schwartz. Lymph node segmentation from CT images using fast marching method. *Computerized Medical Imaging and Graphics*, 28(1-2):33–38, 2004.