**BioMedical Engineering OnLine**

**RESEARCH**                                                                    **Open Access**

# Head and neck lymph node region delineation with image registration

Chia-Chi Teng[1*], Linda G Shapiro[2], Ira J Kalet[2,3]

* Correspondence:
ccteng@byu.edu
[1]School of Technology, Brigham Young University, Provo, UT, USA

## Abstract

**Background:** The success of radiation therapy depends critically on accurately delineating the target volume, which is the region of known or suspected disease in a patient. Methods that can compute a contour set defining a target volume on a set of patient images will contribute greatly to the success of radiation therapy and dramatically reduce the workload of radiation oncologists, who currently draw the target by hand on the images using simple computer drawing tools. The most challenging part of this process is to estimate where there is microscopic spread of disease.

**Methods:** Given a set of reference CT images with "gold standard" lymph node regions drawn by the experts, we are proposing an image registration based method that could automatically contour the cervical lymph code levels for patients receiving radiation therapy. We are also proposing a method that could help us identify the reference models which could potentially produce the best results.

**Results:** The computer generated lymph node regions are evaluated quantitatively and qualitatively.

**Conclusions:** Although not conforming to clinical criteria, the results suggest the technique has promise.

## Background

Malignant tumors in the head and neck represent a great epidemiological problem in western countries. Head and neck cancer accounts for approximately 3% of all cancer cases reported in the United State, or roughly 50,000 cases per year [1]. Due to the tumor position, the risk of developing lymph node metastases in the neck region is very high. Radiation therapy is used as part of the treatment in a majority of the cases. Therefore a fast and effective system for creating a conformal radiation treatment for enlarged (i.e. potentially malignant) lymph nodes is essential.

Computerized tomography (CT) scanning is commonly used for conformal radiation treatment. The scan is performed with the patient set in the treatment position, immobilized using custom devices, thereby minimizing movement of the treatment target. Radiation oncologists have adopted definitions for the various components of the target volume, in order to achieve some uniformity and facilitate the conduct of inter institutional clinical trials [2,3]. The Gross Target Volume (GTV) is the visible and palpable tumor mass. Although it can usually be seen on images (CT and MR), it is normally difficult to automatically identify with existing image processing

techniques. To date it is still usually hand drawn by clinicians using a computer software drawing tool. The Clinical Target Volume (CTV) includes the locations of microscopic local and regional spread, which usually means the GTV plus the lymph node regions around it. Microscopic disease cannot currently be imaged by any existing technique. Even the nodes themselves are often hard to identify in the images. The task of delineating these nodal regions, which is also usually done by the clinicians, is very time consuming. Figure 1 shows how these target volumes are related to each other [4].

Creating the 3D CTV is a critical part of the 3D radiation treatment and Intensity-Modulated Radiation Therapy (IMRT) as the success of radiotherapy depends on the accuracy of the CTV. A conformal IMRT plan with accurately drawn CTV can avoid critical anatomic structures and maximize radiation dosage. As 3D conformal radiotherapy and IMRT become the state of the art, the process of CTV delineation is more important than ever. This process currently also requires radiation oncologists to manually draw the 2D target contours on axial CT slices. It is tedious, time consuming and can be the bottle neck to make IMRT available to more patients. As imaging based cervical lymph node region classification is developed, it is possible to design a system that can identify critical anatomic structures and contour CTV by segmenting patients' CT images with little or no user interaction. Software tools that automate the segmentation of critical structures and contouring of target volume is crucial to the success of implementing a fast and effective radiation treatment planning system as it can dramatically decrease the planning effort for radiation oncologists and increase the availability of IMRT to more patients. The objective of this study is to create a prototype system which is capable of generating a patient's head and neck CTV contours from his CT scan. This paper summarizes our previous work [5-8] and presents a complete system with more comprehensive results.
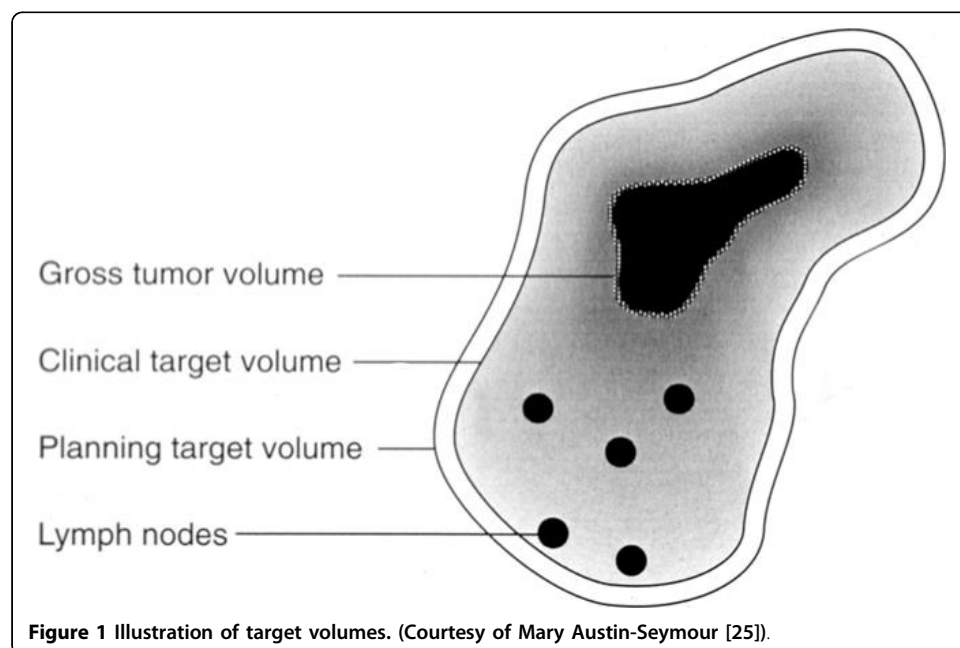


**Figure 1 Illustration of target volumes. (Courtesy of Mary Austin-Seymour [25])**.

### Imaging-based lymph node regions

The neck has an extensive lymphatic network [9]. In fact, more than one third of the body's total number of lymph nodes resides in the extracranial head and neck. Cervical lymph nodes are divided into regions or 'levels' that are described by their anatomic location [10]. Although this traditional classification was decided using surgical landmarks, translation into an imaging-based nodal classification is feasible.

Automatic segmentation of cervical lymph nodes remains to be an open problem, researchers [11,12] are actively working on techniques to segment the lymph nodes for diagnosis or surgery planning. However, in the context of radiation therapy planning, the exact contours of lymph nodes are not as important as the lymph node regions including the surrounding tissue which make up the CTV. Studies have been conducted to create an imaging-based classification for the lymph node levels of the neck that can be accepted by clinicians and easily used by radiologists [4,13-17]. Anatomic landmarks were chosen to create a consistent nodal classification similar to the clinically-based classifications. Radiologists must be able to identify the pertinent anatomic landmarks such as the bottom of the hyoid bone, the back edge of the submandibular gland, and the back edge of the sternocleidomastoid muscle. The Radiation Therapy Oncology Group (RTOG) [18] has also published guidelines for CT-based delineation of lymph node levels in the neck and the anatomic boundaries for delineation.

Automatic delineation of lymph node region can reduce physicians' manual CTV contouring time even though the results are not sufficiently accurate for clinical use directly [19,20]. Atlas-based segmentation is used in most of the state of the art research [17,21,22] and commercial tools [23,24] for automatic delineation of lymph node levels in head and neck CT. These methods tend to yield better results when the atlas is more anatomically similar with the target subjects. The method and database (CT images) used to construct an unbiased atlas is critical to the success of the segmentation [25,26]. However, the high anatomical variability in post-operative head and neck CT images makes it very difficult to construct a mean image and atlas that works well for all patients. We proposed an alternative approach which uses a collection of CT images with contoured CTV from previously treated patients as reference models [6], and a method to identify reference subjects whose anatomic structures share similar properties or features for a given target [8]. Using previously treated patients or canonical models with the most similar head and neck anatomy as references, an image registration process can segment lymph node regions more accurately for a target patient based on known contours in the reference models.

Recent studies evaluated some of the state of the art atlas-based segmentation tools listed above by comparing the automatically delineated head and neck lymph node region contours and volumes against the ones drawn by physicians [27-29]. In addition to qualitative assessment by physicians, statistics measures such as sensitivity and specificity or Dice similarity coefficient were commonly used as quantitative assessment. We also proposed an alternative quantitative evaluation using Hausdorff surface distance measure which maybe more clinically relevant than the statistical metrics [6].

Given the set of post-op head and neck cancer patients, a series of 2D contours were manually delineated for each of the lymph node levels on axial CT images; which build up to 3D volumes. Using an image registration technique, these expert drawn lymph

node regions are used as reference models and templates to project the lymph node regions in another target image which are compared to the expert drawn contours in the target image, i.e. the "gold standard" or "ground truth". Instead of the atlas-based approach or choosing one patient as the reference model, we will determine criteria for choosing one or more similar reference models which can produce optimal results. Traditional 3D shape retrieval systems [30-32] mostly experiment with artificial models and focus mainly on classifying 3D models of very different shapes. While these experimental systems can match models of the same classes to a certain degree of success, they usually fall short of distinguishing the finer details of objects within a class. Using 3D medical images to find similarity among a known set of patients is becoming a research subject of interest in many medical domains. Ruiz *et al.* [33] use a shape-based similarity measure to find similar craniosynostosis patients for intervention planning. We developed a method to find similar head and neck cancer patients for radiation planning. The similarity of head and neck anatomy between patients is based not only on shape features of structures, such as outer body volume, mandible, and hyoid, but also on their relative locations. These types of medical-image-based problems are very domain specific, and are not solved by the traditional shape-based retrieval system.

Recent reviews of 3D shape matching techniques were done by Iyer [34] and Tangelder [35]. A majority of the 3D shape matching systems use feature-based methods, which compare geometric and topological properties of 3D shapes. Methods using features or distributions work reasonably well in classifying objects of different shapes, but they do not discriminate between objects of the same class such as the head and neck anatomy of different patients. The matching process is usually done by computing a distance between feature vectors representing the different objects. Most systems do not give many details on the distance measurements or their comparison methods, although they usually imply a Euclidian vector space model and use either a simple (weighted) Euclidean distance or a city-block ($L_1$ Minkowski) distance.

## Methods

This prototype system is designed to take a cancer patient's head and neck CT images as input and use image registration techniques to produce projections of lymph node regions as output, which can be used to produce a CTV for the radiation treatment plan. The system can be divided into the following major components:

- Segmentation,
- Retrieval of similar reference models,
- Image registration.

Figure 2 is the flow chart which shows how these components are linked. The offline process on the left creates a database **DB** of CT scans from prototypical reference patients on which experts have drawn contours that denote the lymph regions. These reference images {$d_i$} are segmented offline to extract 3D volumes of landmark anatomical structures such as the mandible and hyoid [5]. 3D meshes and geometric properties of these 3D volumes are also stored in the database.

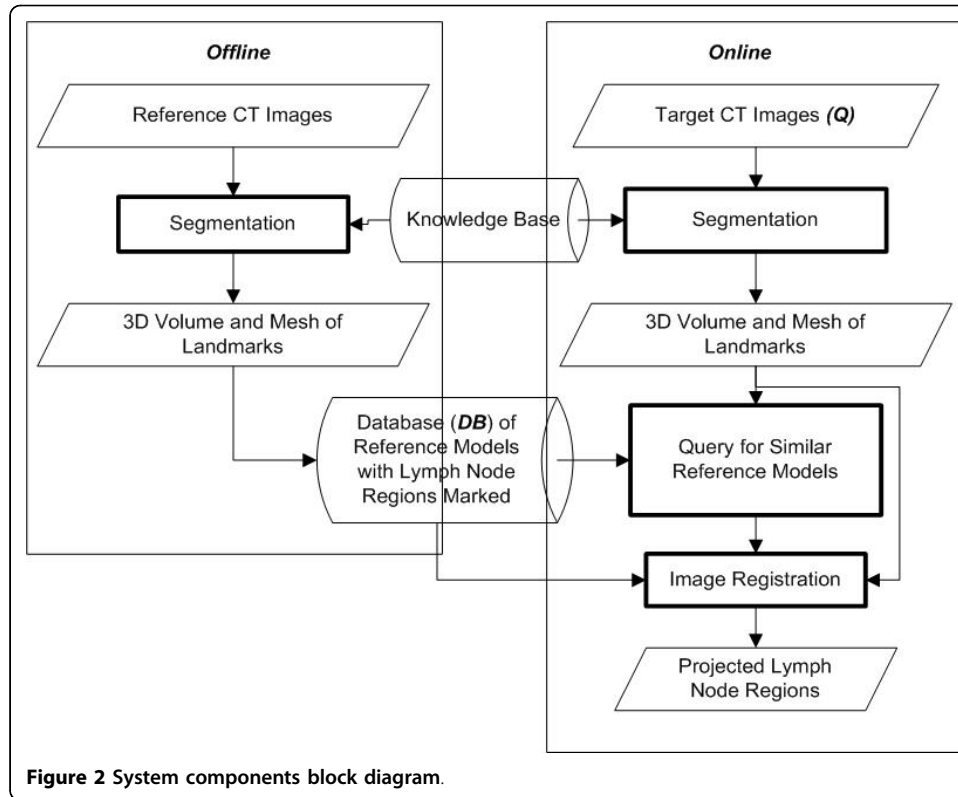**Figure 2 System components block diagram**.

## Image registration

Given a reference model and a target patient's CT data set, we can use image registration methods to align the sets of CT images. Image registration is commonly used in medical imaging applications. It is essentially a process of finding a geometric transformation **g** between two sets of images, which maps a point **x** in one image-based coordinate system to **g(x)** in the other. By assuming the head and neck anatomy has similar characteristics between a specific target patient and a reference subject, we can use image registration methods to transform a region from the reference image set to the target image set.

Mattes and Haynor [36] implemented a multi-resolution non-rigid (deformable) image registration method using B-splines and mutual information. The transformation of a point $\mathbf{x} = [x, y, z]^T$ in the reference image coordinate system to the test image coordinate system is defined by a $3 \times 3$- homogeneous rotation matrix $\mathbf{R}$, a 3-element transformation vector $\mathbf{T}$ and a deformation term $\mathbf{D}(\mathbf{x}|\boldsymbol{\delta})$:

$$\mathbf{g}(\mathbf{x} \mid \boldsymbol{\mu}) = \mathbf{R}(\mathbf{x} - \mathbf{x}_C) - (\mathbf{T} - \mathbf{x}_C) + \mathbf{D}(\mathbf{x} \mid \boldsymbol{\delta}) \tag{1}$$

where $\mathbf{x}_C$ is the center of the reference volume. A rigid body transformation defined by $\mathbf{R}$ and $\mathbf{T}$ was first calculated and used as the initial transformation for the deformation process. The deformation term $\mathbf{D}(\mathbf{x}|\boldsymbol{\delta})$ gives an x-, y-, and z- offset for each given **x**. Hence the transformation parameter vector μ becomes

$$\boldsymbol{\mu} = \{\gamma, \theta, \phi, t_x, t_y, t_z; \boldsymbol{\delta}_j\} \tag{2}$$

The first three parameters $\gamma$, $\theta$, $\phi$ are the roll-pitch-yaw Euler angles of R. The translation vector **T** is defined by $[t_x, t_y, t_z]^{\mathrm{T}}$. **T** and **R** together define the rigid body transformation.

The parameter $\boldsymbol{\delta}_j$ is the set of the deformation coefficients. The deformation was modeled on cubic B-splines [37] because of their computational efficiency (via separability in multidimensional expression), smoothness, and local control. The deformation is defined on a sparse, regular grid of control points $\boldsymbol{\lambda}_j$, each having associated x-, y-, and z-components of the deformation. The resolution

$$\boldsymbol{\rho} = \mid \rho_x, \rho_y, \rho_z \mid \tag{3}$$

of the deformation determines the spacing of the grid and can be anisotropic. Mattes uses control points on a uniform grid with spacing

$$\Delta\boldsymbol{\rho} = [\Delta\rho_x, \Delta\rho_y, \Delta\rho_z]^{\mathrm{T}} = \left[ \frac{q_x - 1}{\rho_x - 1}, \frac{q_y - 1}{\rho_y - 1}, \frac{q_z - 1}{\rho_z - 1} \right]^{T} \tag{4}$$

where $q_x$, $q_y$, and $q_z$ are the dimensions of the reference image.

The deformation at any point $\mathbf{x} = [x, y, z]^{T}$ in the reference image is interpolated using a cubic B-spline convolution:

$$\mathbf{D}(\mathbf{x} \mid \boldsymbol{\delta}) = \sum_{j} \boldsymbol{\delta}_j \boldsymbol{\beta}^{(3)}\left( \frac{\mathbf{x} - \boldsymbol{\lambda}_j}{\Delta\boldsymbol{\rho}} \right) \tag{5}$$

By displacing the control points, intermediate deformation values are computed by cubic spline interpolation between them.

Because contract enhancement is often used in head and neck CT scans, image intensity range and distribution may vary in different data sets. A mutual information based registration method such as Mattes' can work with images in different range. However, the high anatomical variability in cancer patients' head and neck CT images particularly contributed by the surgical resections, pure voxel intensity based registration such as Mattes' method described in previous section does not always produce satisfactory results. A new method [6] is developed which integrates landmark based information with intensity scheme; hence CT data set need to be preprocessed to extract landmark information.

### Segmentation and landmark correspondence

Fully automatic segmentation in the neck region is particularly difficult, because many soft tissue anatomic entities are small in size and similar in density. Furthermore, they can be directly adjacent to each other or only divided by fascial layers that are not visible in CT images. The relative locations between anatomic entities can vary in different axial locations. Little work has been done specifically for the neck images; few exceptions include the work of Krugar *et al.* [10] who implemented a semi-automatic system to segment neck CT images for pre-operative planning of neck dissections; and the work of Cordes *et al.* [38] who developed NeckVision system for neck dissection planning. We implemented an automatic segmentation method [5] designed to locate anatomic structures in the neck that are relevant to the lymph node region boundaries

including cervical spine, mandible, hyoid, jugular veins, and carotid arteries. This method is motivated by a knowledge-based technique [39] using consists of constraint-based dynamic thresholding, negative shape constraints to rule out infeasible segmentation, and progressive landmarking that takes advantage of the different degrees of certainty of successful identification of each structure.

The drawback of this 2D thresholding approach is the difficulty in determining the optimal threshold, especially in the head and neck region where many structures of similar density are crowded in a tight space. Our proposed method eliminates the need of finding the optimal threshold by combining the 2D thresholding with a 3D active contour procedure (by ITK, www.itk.org). A sub-optimal threshold that only produces partial structure on some axial slices can grow into a 3D volume through 3D active contouring. The 2D regions produced by the knowledge-based dynamic thresholding method can be used as the initial region or "seed" to eliminate the need of user input. In the context of finding relevant landmarks, it is not necessary to fully segment certain structures such as the blood vessels and their branches as we are only interested in sections lateral to the hyoid. This lax requirement circumvents the need of perfect segmentation and optimal active contouring parameters. These selected structures usually have clear contour allowing successful segmentation within the section of axial slices of interests.

The anatomical structures of interests are segmented in the order according to the reliability of successful detection. In addition to domain knowledge of gray tone range, size and shape, each structure is also associated with location relative to other structures that can be found prior to itself. For each structure, we first run the dynamic thresholding process to find 2D regions in axial slices according to domain knowledge. Note that some axial slices may yield successful 2D segmentation results and others do not, we then the 3D active contouring to build a 3D model using the partial 2D segmentation result as the initial seed. Also note that the active contouring may not grow the structure perfectly to its full extent, but sufficient to provide landmark information. This two step process is then repeated for the next structure. 3D surface meshes of those structures are constructed for both the reference and target subjects. The meshes can be used as landmarks to improve the alignment and to measure similarity between the subjects. Figure 3 shows examples of segmentation results on selected subjects.

We choose to use these 3D anatomic structure surfaces as landmarks because virtually no 1D (points) or 2D (lines) anatomic features that can be used as landmarks are defined in the neck. By using Shelton's method [40] of finding surface mesh correspondence, we can estimate correspondence between the surfaces of the landmark anatomic structures of the reference and target subjects. The following energy function for which smaller values indicate better correspondences is defined to evaluate possible correspondence relations:

$$E(C) = E_{\mathrm{sim}}(C) + \alpha E_{\mathrm{str}}(C) + \beta E_{\mathrm{pri}}(C) \qquad (6)$$

where **C** is the function that maps points on surface **A** to matching points on surface **B**, $\alpha$ and $\beta$ are weight parameters, $E_{\mathrm{sim}}$ is the similarity term which measures how closely **C** matches points on **A** to points on **B**, $E_{\mathrm{str}}$ is the structural term that minimize the distortion of surface **A**, and $E_{\mathrm{pri}}$ is the "prior information" term which ensures **C**
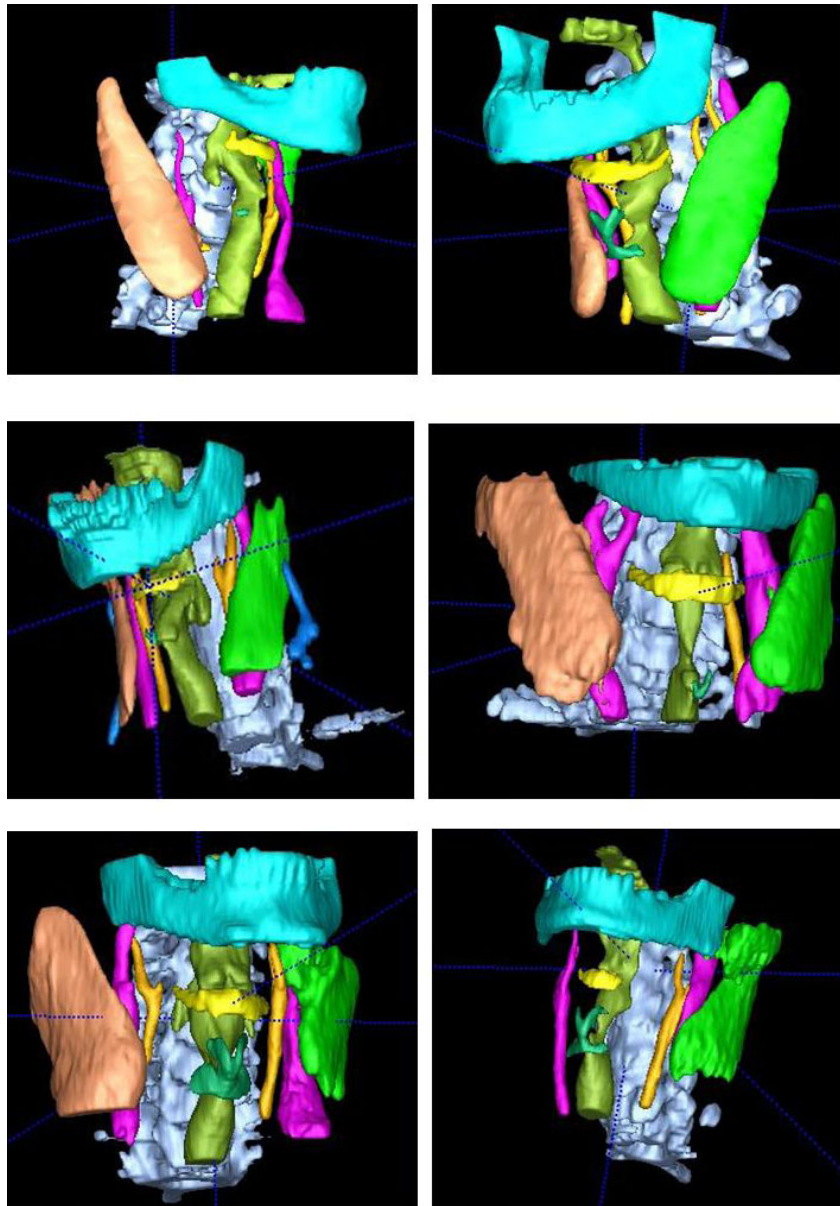
**Figure 3 Examples of automatic segmentation results for selected subjects: (a) cervical spine, (b) respiratory tract, (c) mandible, (d) hyoid, (e) thyroid cartilage, (f) jugular vein, (g) common carotid artery, and (h) sternocleidomastoid muscle**.

represent a plausible deformation. Initial values of $\alpha$ and $\beta$ are set to 0.001 and 0.0001 respectively.

Let $v_k$ be the set of landmark points sampled from the mandible, hyoid and other surface meshes of the reference image set or surface **A**, and $\varpi_k$ be their corresponding locations on surface $C(\mathbf{A})$ or the transformed surface which matches the test image set and minimize the energy function. The deformation $\zeta$ at those landmark points is simply

$$\zeta_k = \varpi_k - \upsilon_k. \tag{7}$$

## Image registration with landmark correspondence

Given the anatomic structure surface correspondence between the reference and target data, we developed a method to incorporate these landmarks into the image registration process. Instead of initializing the deformations at the control points to zeros or random numbers as in the Mattes method, we can use the landmark correspondence to initialize or adjust the deformations at the control points at each of the multi-resolution stages. The deformation control points are set to a uniform grid

$$\boldsymbol{\lambda}_j = [l\Delta\rho_x, m\Delta\rho_y, n\Delta\rho_z]^T, \tag{8}$$

where $0 \leq l \leq \rho_x$, $0 \leq m \leq \rho_y$, $0 \leq n \leq \rho_z$, and the corresponding deformation values $\mathbf{D}(\boldsymbol{\lambda}_j)$ are either initially set to zero or calculated from the deformation coefficients $\boldsymbol{\delta}_j$ of the previous iteration at a lower resolution of control points as in equation (5).

Given $\mathbf{v}_k$ and $\varpi_k$ as sets of corresponding landmarks in the reference and target images described in equation (7), the deformation of each control point that has landmark points in close proximity is modified to the deformation of the closest landmark point as follows

$$D'(\boldsymbol{\lambda}_j) = \begin{cases} \zeta_k & \begin{aligned} &if \mid \lambda_{jx} - v_{kx} \mid < \frac{1}{2}\Delta\rho_x \vee \\ &\mid \lambda_{jy} - v_{ky} \mid < \frac{1}{2}\Delta\rho_x \vee \mid \lambda_{jx} - v_{kx} \mid < \frac{1}{2}\Delta\rho_z \end{aligned} \\ D(\boldsymbol{\lambda}_j) & otherwise \end{cases} \tag{9}$$

where $\mathbf{v}_k$ is the closest landmark point to $\boldsymbol{\lambda}_j$ in the reference image set, and $\zeta_k$ is the deformation at $\mathbf{v}_k$ obtained from the surface correspondence in equation (7). A new set of deformation coefficients $\boldsymbol{\delta}$ is then set to the spline coefficients of the new grid of deformation values $\mathbf{D}'(\boldsymbol{\lambda})$. While the new $\mathbf{D}'$ might not be initially smooth, the following mutual information registration will mitigate the side effect. Finally the transformation parameter vector $\boldsymbol{\mu}$ is input to the optimizer for alignment.

## Similarity measurement based on anatomical structure geometry

As we need to compare similarity between anatomic structures from different patients, we do so by measuring the errors between structure surfaces using the 3D Hausdorff distance [41]. Given two surface meshes, $S_R$ and $S_T$, the distance between a point $p_R$ belonging to $S_R$ and the mesh $S_T$ can be defined as follows:

$$d(p_R, S_T) = \min_{p \in S_T} \left\| p_R - p \right\|. \tag{10}$$

We first align meshes $S_R$ and $S_T$ with the Iterative Closest Point (ICP) rigid body registration [42] so they are roughly in the same 3D coordinate. Given the 3D point sets $P_R = \{p_i\}$ containing the n vertices of $S_R$, the registration process will produce a transformation matrix $\boldsymbol{T}$ which minimizes the function

$$D(S_R, S_T) = \sum_{i=1}^{n} d(Tp_i, S_T). \tag{11}$$

The transformed reference mesh $TS_R$ consists of vertices $\{T_{pi}\}$, and the Hausdorff distance between $TS_R$ and $S_T$ is given by

$$d_h(TS_R, S_T) = \max_{p \in S_R} d(Tp, S_T). \tag{12}$$

A database $DB$ of CT scans is created from prototypical reference subjects on which experts have drawn "ground truth" contours that denote the lymph node regions. These reference data sets $\{d_i\}$ are segmented offline to extract 3D volumes of identifiable structures, including the mandible, hyoid, jugular veins and the outer body contour that are relevant to the boundaries for the lymph node regions [5]. We use these landmark anatomic structures to rapidly produce a distance metric between a target CT scan as query $Q$ and each data set $d$ in $DB$. The feature vectors that we use to compare two CT scans include three kinds of features: 1) simple numeric 3D regional properties of these structures, such as volume and extents, 2) vector properties or the relative location between structures and 3) shape properties or the surface meshes of these structures. The feature vector consists of the following properties,

- volume and extents of the overall head and neck region,
- surface meshes of the mandible and outer body contour,
- 3D centroid difference vector between mandible and hyoid,
- 2D centroid difference vectors between hyoid and jugular veins, and between hyoid and spinal cord on the axial slice at the centroid of the hyoid,
- normalized centroid locations of the hyoid and the mandible within the region.

Given feature vectors $F_d$ and $F_Q$ for model $d$ and query $Q$ in the feature vector space $R^N$, the following weighed Euclidean distance is used as the distance measure:
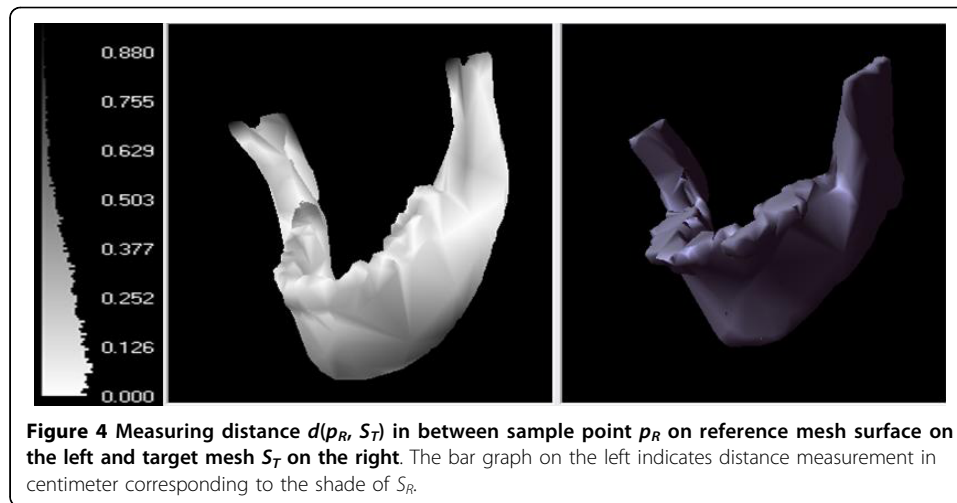
$$D_F(F_d, F_Q) = \left[ \sum_{i=1}^{N} w_i d_i(F_{di}, F_{Qi})^2 \right]^{\frac{1}{2}} \tag{13}$$

where $w_i$ is the weight parameter, $d_i$ is the distance function for feature $i$,

$$d_i(f_1, f_2) = \begin{cases} d_h(Tf_1, f_2) & \text{for meshes} \\ |f_1 - f_2| & \text{otherwise} \end{cases} \tag{14}$$

$d_h$ is the Hausdorff distance defined in equation (12), and $T$ is the ICP registration transformation matrix. The weight parameters range from 10 to 0.1 from heavy to light in the order of: 1) hyoid locations (normalized and relative to other structures), 2) mandible distance (between two models), 3) vertical distances between skull base, mandible and, hyoid, 4) head and neck outer contour and volume, and so on.

The distance between mandible meshes of two subjects is one major discriminating feature of the proposed distance measure. Figure 4 shows the measurement of point to surface distance $d$ as in Equation (10), from which the directional Hausdorff distance $d_h$ between the reference mandible surface mesh to the target mesh is derived. The mesh on the left with shading indicates the distance from a given point on the surface to the mesh on the right.

**Figure 4 Measuring distance $d(p_R, S_T)$ in between sample point $p_R$ on reference mesh surface on the left and target mesh $S_T$ on the right**. The bar graph on the left indicates distance measurement in centimeter corresponding to the shade of $S_R$.

## Results and discussion

CT images are acquired from the radiation therapy treatment planning system. Head and neck cancer patients who did not have large scale surgical resection are the primary candidates. All the images used in the experiments are CT scans performed at the University of Washington Medical Center using a General Electric CT scanner. Head and neck CT images were selected in which all of the slices are 512 × 512 pixels in axial dimension; the distance between slices varies between 1.25 mm and 3.75 mm for each image set. Selected lymph node regions are drawn as 2D contours on axial CT slices of twenty subjects chosen by resident physicians in University of Washington Radiation Oncology Department. These lymph node regions are used as "ground truth" or "gold standard" as we evaluate the computer generated lymph node regions. A simple automated segmentation pre-process removes the bed and immobilization devices on each CT images. The experiment was conducted on personal computers with Pentium 4 processors and 2 GB RAM.

### Image registration

We ran the image registration experiment with the twenty sets of CT images chosen by the oncologists, each of which is used as a reference and a target data set to align with every other image set. The maximum resolution of the deformation control points as defined in Eq. (3) is [15,15,11], or 2475 control points. The pairing results in 20 × (20-1) = 380 total registration tests. We used the results of the Mattes method as the baseline to compare with the results of the new method.

Out of the 380 total tests, the Mattes method failed to converge in 13 cases while the proposed method succeeded in all cases. The average time to complete the registration improved by approximately 20%. Tables 1 and 2 compares the success rate and convergence time of the two image registration method out of 380 inter-subject cases.

3D volumes and surface meshes of the lymph node regions are built from the 2D "ground truth" contours as well as the computer generated lymph node regions. We can quantitatively evaluate the results of the image registration process with statistical metric such as spatial overlap index and surface distance measure. Given $V_T$ as the ground true volume and $V_C$ as the computer generated volume, we can use Dice

**Table 1 Success rate comparison between Mattes and proposed method**

|  | Total cases | Successful registration cases | Success rate (%) |
|---|---|---|---|
| Mattes method | 380 | 367 | 96.57% |
| Proposed method using landmark correspondence | 380 | 380 | 100.00% |

**Table 2 Time of convergence comparison between Mattes and proposed method**

|  | Average time of convergence | Standard deviation time of convergence | Cases faster |
|---|---|---|---|
| Mattes method | 32 minutes | 6 minutes | 326/380 |
| Proposed method using landmark correspondence | 26 minutes | 5 minutes | 54/380 |

similarity coefficient (DSC) to measure the spatial overlap between two segmentations [43], where

$$DSC(V_C, V_T) = 2 \frac{(V_C \cap V_T)}{V_C + V_T}. \tag{15}$$

Table 3 shows examples of the statistical evaluation of the automatically delineated contours. While statistical metrics give certain validation, we suggest that Hausdorff surface distance error may be more clinically relevant as it represents the "worst case scenario" or often a critical error. Given two 3D surface meshes: reference lymph node region transformed into target space as the projected region and the corresponding "ground truth" region of the target subject; we measure the Hausdorff distance and mean distance between these meshes by sampling the surface points. The Hausdorff distance between the projected lymph node region $N$ of reference subject $S_R$ and the expert drawn region for a given target subject $S_T$ can be defined as $D_H(TS_R, S_T, N)$, where $T$ is transformation resulted from the image registration of $S_R$ and $S_T$.

Figure 5 shows an example of qualitative comparison of image registration results using the Mattes' method without landmark information and the proposed method which incorporates landmark correspondence. Rows A-E show selected axial CT slices in the neighborhood of the hyoid in various data sets from superior to inferior location. Column 1 shows slices from the reference subject, and column 4 shows slices from the target subject. Columns 2 and 3 show transformed reference images which were re-sampled to match or align with the target images using the transformation function $\mathbf{g}(\mathbf{x}|\mu)$ of equation (1) produced from the image registration procedure. Column 2 is the result of Mattes' image registration method without using landmark information, and column 3 is the result of the proposed method using landmark correspondence. The images in column 3 match better qualitatively to column 4 in several

**Table 3 Example of volume overlap analysis with Dice similarity index (DSC)**

| Level | Side | Average DSC | Standard deviation |
|---|---|---|---|
| 1B | Left | 0.74 | 0.11 |
| 1B | Right | 0.76 | 0.10 |
| 2 | Left | 0.71 | 0.12 |
| 2 | Right | 0.72 | 0.12 |

ways. First of all, for example, the outer body contours is closer to the ones of target CT. Second, the superior-inferior location of the hyoid bone and the mandible in column 3 match more closely to the ones in column 4; such as the hyoid bone is in slices (rows) C-E (and beyond) of column 2, but in slices A¬D of column 4, and the inferior boundary of the mandible ends in slice D of column 2, but continues beyond slice E in columns 3 and 4. This is important because both hyoid and mandible represent superior-inferior boundaries to adjacent lymph node levels. Third, the proximity of the cervical spine in column 3 appears to be more closely matched to column 4. Even though certain soft tissue structures appear distorted in column 3; in the context of delineating lymph node regions, it is relatively less important compared to matching the location of the dominant structures defining the boundaries.

Figure 6 shows comparisons of image registration results from the Mattes method and the proposed method incorporating the landmark correspondence. The horizontal axis represents the Hausdorff or mean distance between projected lymph node regions based on the reference model using image registration results from the Mattes method and corresponding expert drawn lymph node regions of the target subject. The vertical axis represents matching results using the proposed registration method. The diagonal dotted lines represent the points where the two measures are equal. The figures show overall improvement using the proposed registration method. Tables 4 and 5 compare the mean and standard deviation of results from two methods. The new landmark-based method improved the average Hausdorff distance by as much as 25%, and the average mean distance by as much as 42%.

There is currently no quantitative standard defining clinical acceptability for automatically delineated target contours, whether in statistical error or maximum surface distance error. However, as the results for two sample lymph node region shown in Tables 4 and 5, the average distances (error) between the generated CTV and ground truth are under 1 cm for the cases tested; the Hausdorff distances are still in the 3 cm range. Methods are being investigated to reduce the maximum error as part of the future work for this continuing project. One of which is to port the registration module to a supercomputer platform and increase to resolution of the control points during the deformation process. Another is to merge the results of projected CTV from multiple similar reference models and possibly alleviate regions that contribute the worst Hausdorff distance.

Figure 7 shows an example of the projected lymph node region contour based on the result of the image registration. The projected 3D lymph node regions are shown at the lower left 3D view. They are also overlaid on the target CT images in axial, sagittal, and coronal views. The radiation oncologists can make adjustments to the suggested contours as needed.

While the proposed method improves the overall results, it does not decrease the Hausdorff or mean distance in every case. One reason is that some of the "ground truth" contours deviate from the image-based classification because physicians were sometimes influenced by their clinical judgment as they outline the contours [20]. We also observed that for cases which already produce small Hausdorff or mean distance, the improvement or regression tends to be negligible. As the collection of reference data sets grow, it becomes critically important to locate the candidates which can potentially produce the best results based on the similarity measurement.
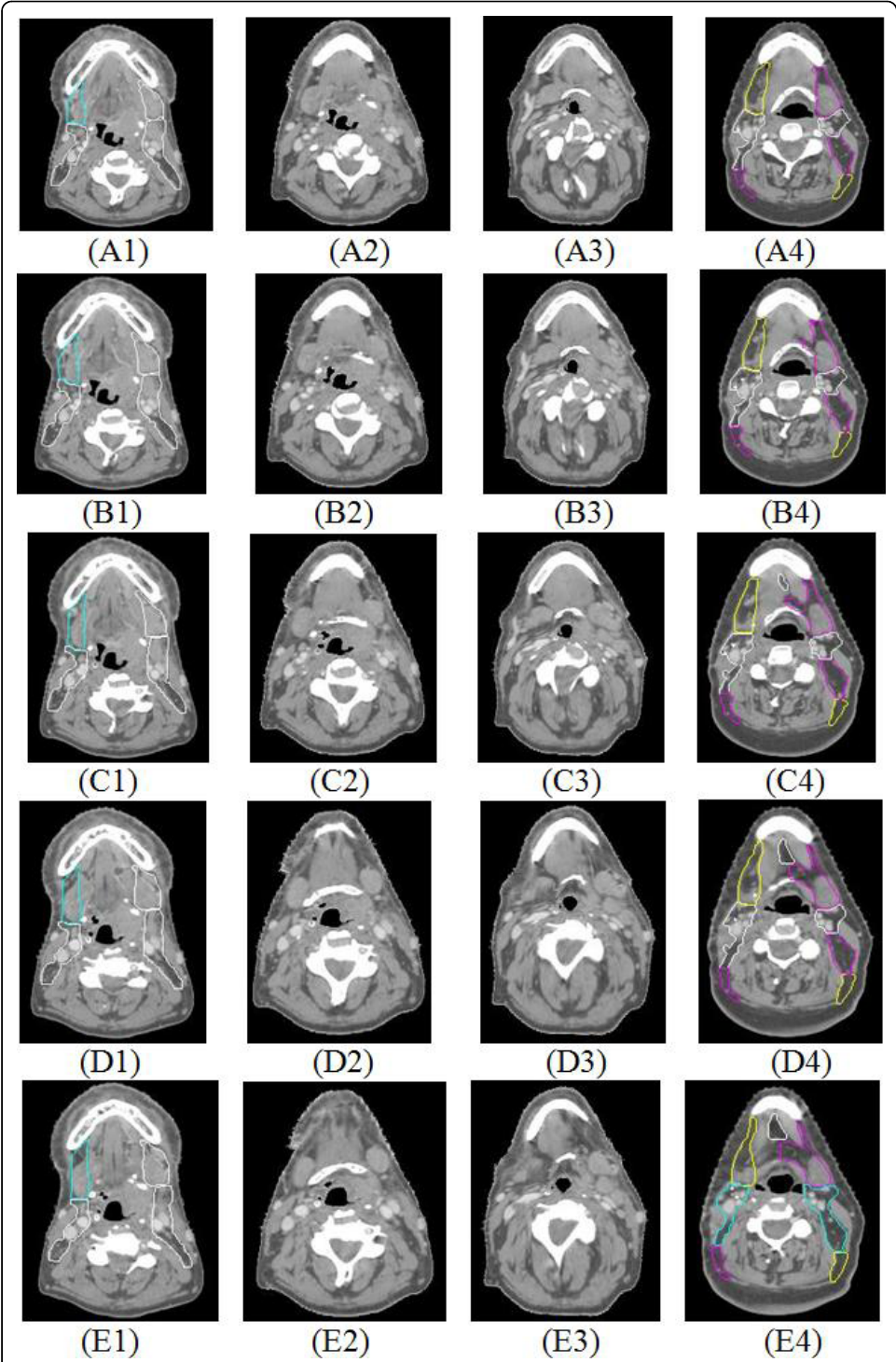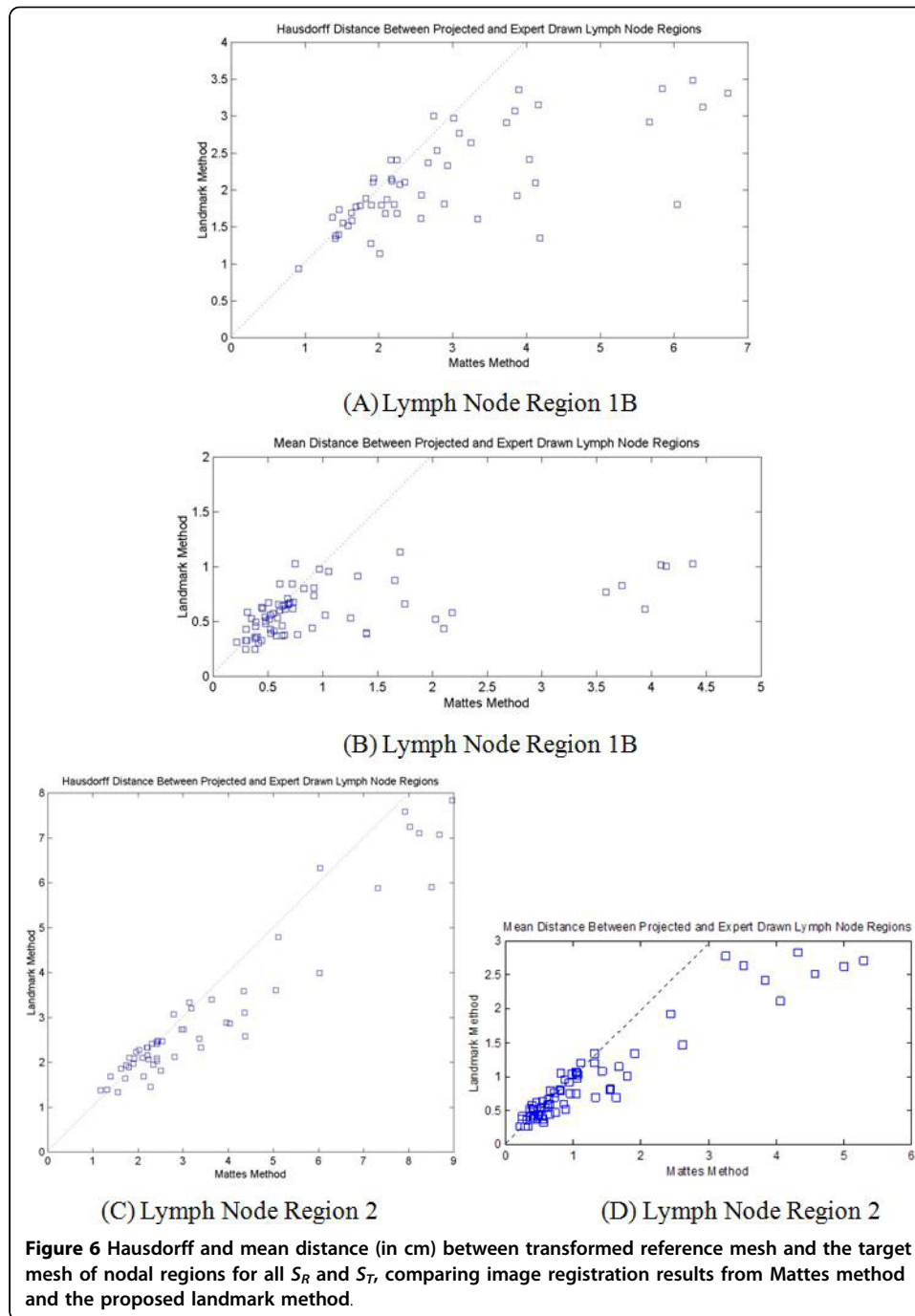
**Figure 5 Comparison of results from the image registration methods with and without using landmark correspondence**. Rows A-E show selected axial CT slices in the neighborhood of the hyoid in various data sets from superior to inferior. Column 1 shows slices from the reference subject, column 4 from the target subject, column 2 is the result of Mattes' image registration method, and column 3 is the result of the new method using landmark correspondence.

**Figure 6 Hausdorff and mean distance (in cm) between transformed reference mesh and the target mesh of nodal regions for all $S_R$ and $S_T$, comparing image registration results from Mattes method and the proposed landmark method**.

### Similarity measurement

For the twenty selected subjects with expert drawn "ground truth" lymph node regions, each of which is used as target and reference, we measure the feature-vector-space distance between each target-reference pairs. For each target subjects, we rank all of the reference subjects by two separate measurements:

- Ranking $R_I$ ranks by Hausdorff distance $D_H$ between projected lymph node region and corresponding expert drawn region of target subject. The smallest value has the highest rank. This represents the ranking of similarity according to the image

**Table 4 $D_H(TS_R, S_T, L)$ for all $S_R, S_T$. Hausdorff distance (in cm) between transformed reference mesh and target mesh for nodal levels ($L$) 1B and 2**

| Level | Algorithm | Average | Standard deviation |
|---|---|---|---|
| 1B | Mattes method | 2.85 | 1.44 |
| 1B | Landmark enhanced | 2.12 | 0.64 |
| 2 | Mattes method | 3.48 | 2.15 |
| 2 | Landmark enhanced | 3.07 | 1.75 |

**Table 5 Mean distance (in cm) between transformed reference mesh and target mesh for nodal level 1B and 2**

| Level | Algorithm | Average | Standard deviation |
|---|---|---|---|
| 1B | Mattes method | 1.02 | 1.01 |
| 1B | Landmark enhanced | 0.59 | 0.21 |
| 2 | Mattes method | 1.21 | 1.20 |
| 2 | Landmark enhanced | 0.91 | 0.68 |

registration results, assuming that the more similar reference image aligns better with the target image. Note that there may be different rankings for each lymph node region as the Hausdorff distance results vary. These rankings are used as the baseline.

- Ranking $R_F$ ranks by distance measure in the feature vector space $D_F$ in Equation (13). The smallest value has the highest rank. This represents the ranking of similarity according to the geometrical features.
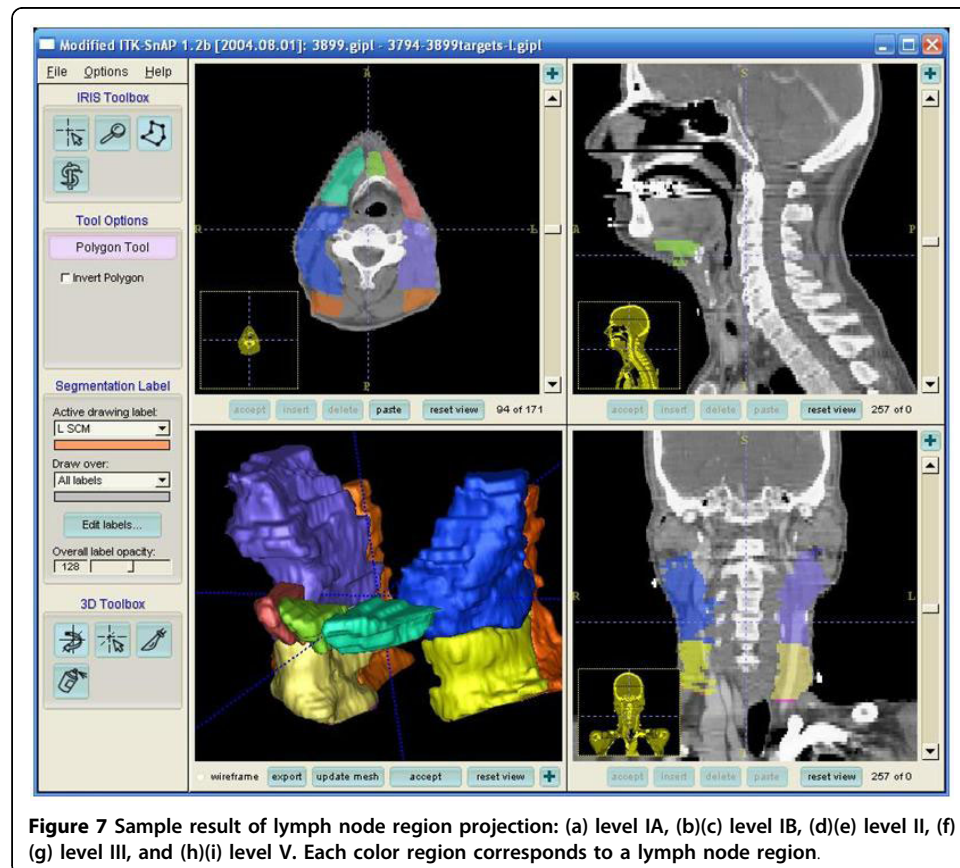


**Figure 7 Sample result of lymph node region projection: (a) level IA, (b)(c) level IB, (d)(e) level II, (f)(g) level III, and (h)(i) level V. Each color region corresponds to a lymph node region**.

**Table 6 Probabilities of the most similar subject based on geometric features matching subjects that align best with target**

|  | $x = 1$ | $x = 2$ | $x = 3$ | $x > 3$ |
|---|---|---|---|---|
| $P(R_I(x, S_T) = R_F(1, S_T) \mid R_F(1, S_T))$, where $R_F(1, S_T)$ is the most geometrically similar reference subject | 81% | 11% | 4% | 4% |

To validate the similarity ranking based on geometrical features $R_F$, we compare it to the ranking based on image registration results for each target subject $R_I$. We define $R(i, S_T)$ as the $i$th reference subject in the ranking for the target subject $S_T$. For 81% of the instances, $R_F(1, S_T)$ is the same as $R_I(1, S_T)$, meaning that 81% of the most similar reference subjects according to geometrical features matches the one that align best with the target according to image registration. In other words, the probability $P(R_I(1, S_T) = R_F(1, S_T) \mid R_F(1, S_T)) = 0.81$. Additionally, $R_F(1, S_T)$ has 96% chance of matching one of the top three subjects in $R_I$. Table 6 summarizes the probabilities of the most similar subject based on geometric features matching subjects that align best with the target.

As we previously defined $D_H(TS_R, S_T, n)$ as the Hausdorff distance between the projected lymph node region $n$ based reference subject $S_R$ and the expert drawn for a given target subject $S_T$. The Hausdorff distance based on the most similar reference subject in the feature space $R_F(1, S_T)$ becomes $D_H(TR_F(1, S_T), S_T, n)$. Table 7 compares the image registration results of the most similar reference subject to results of all 380 test cases listed in previous section. The average $D_H$ of 1.28 cm is approximately 50% improvement over the average of all test cases.
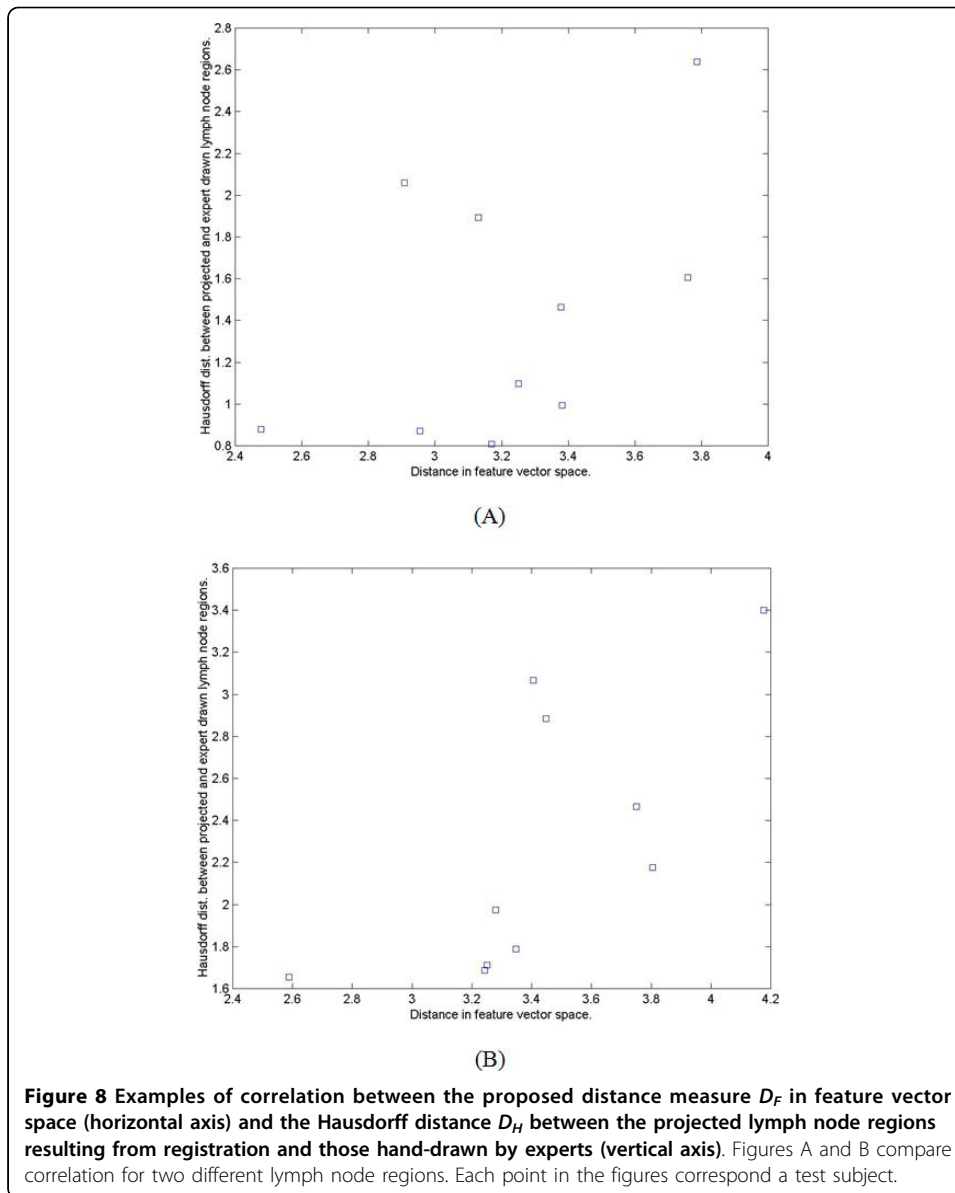
Figure 8 shows examples of correlation between the proposed distance measure and the result of 3D deformable image registration for selected target patients. The vertical axis shows the Hausdorff distance $D_H$ between the transformed 3D mesh of the reference model lymph nodal region and the corresponding mesh of the test model using the transformation produced by the 3D deformable image registration. The horizontal axis represents the distance measurement, or the weighed Euclidian distance $D_F$ between the test and reference model properties in the feature vector space. The left-most point in each graph represent the most similar reference subjects $R_F(1, S_T)$ which matches to one of the subjects that align well with the target, i.e. with the shortest Hausdorff distance $D_H(TR_F(1, S_T), S_T, n)$ compares to others. Table 8 also shows statistical analysis of the correlation coefficients of $D_F$ and $D_H$ for all the target subjects $S_T$.

**Qualitative evaluation**

The projected lymph node regions can be evaluated qualitatively by superimposed on the target subject's CT images. Figure 9 compares projected lymph node regions to

**Table 7 Compare image registration results based on most similar reference subject and results based on all test cases**
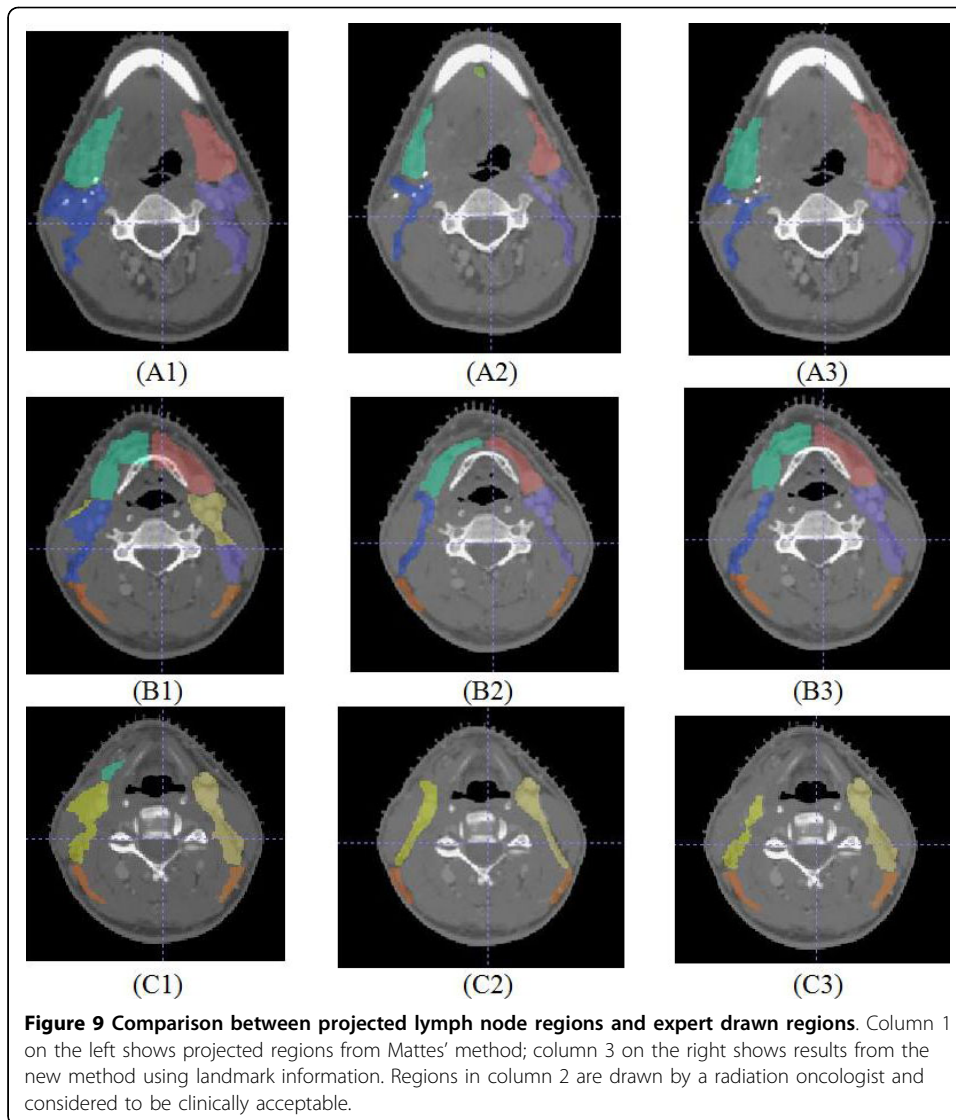
|  | Average | Standard deviation |
|---|---|---|
| $D_H(TR_F(1, S_T), S_T, n)$ for all $S_T$ and $n$, based on most similar reference subjects of each target subject. | 1.28 | 0.31 |
| $D_H(TS_R, S_T, n)$ for all $S_R$, $S_T$ and $n$, based on all combinations of reference and target subjects. | 2.59 | 0.90 |

**Figure 8 Examples of correlation between the proposed distance measure $D_F$ in feature vector space (horizontal axis) and the Hausdorff distance $D_H$ between the projected lymph node regions resulting from registration and those hand-drawn by experts (vertical axis)**. Figures A and B compare correlation for two different lymph node regions. Each point in the figures correspond a test subject.

expert drawn regions, each region is color coded. Rows 1-3 are sample CT slices from superior to inferior positions of the same target subject. Column 1 on the left shows projected regions from Mattes' method; column 3 on the right shows results from the new method using landmark information. Regions in column 2 are drawn by a radiation oncologist. These projected lymph regions are reviewed by the radiation oncologist and considered to be clinically acceptable. The results from the Mattes' method are more generous in certain areas covering muscle tissues. Although it may be considered harmless today, it can be less desirable as the precise lymph node region contours becomes more important in the future.

**Table 8 Correlation coefficients of $D_F$ and $D_H$ for all the target subjects $S_T$**

|  | Average | Standard deviation |
| --- | --- | --- |
| Correlation_coefficient($D_H$, $D_F$) for all $S_T$. | 0.61 | 0.19 |

**Figure 9 Comparison between projected lymph node regions and expert drawn regions**. Column 1 on the left shows projected regions from Mattes' method; column 3 on the right shows results from the new method using landmark information. Regions in column 2 are drawn by a radiation oncologist and considered to be clinically acceptable.

## Conclusion

This software system is the first of its kind that attempts to automate the process of identifying Clinical Target Volume for head and neck cancer radiation treatment by projecting lymph node regions using inter-subject image registration techniques. We developed a image registration technique using landmark correspondences in conjunction with a voxel based mutual information method, along with a patient similarity measurement using the 3D geometrical relationship between anatomic structures in addition to 3D shape descriptors of the structures. The image registration technique provides a way to potentially automate the lymph node region contouring process for radiation therapy planning. While the alignment of the projected lymph node contours on the target image are close enough to suggest the technique has promise, the results do not conform to clinical criteria. However, this system could drastically reduce the time needed for a physician to draw a fully-detailed target volume as an intermediate step that suggests an initial target volume upon which the physician can adjust and refine before it is used for therapy.

While the new image registration technique improves the overall result of the target volume projection, the results discussed in previous section shows that it is perhaps more important to identify the reference models which can potentially produce the best alignment. As the similarity measurement is validated by the target volume projection, twenty data sets with "ground truth" information is not quite enough to provide a variety of subjects with different anatomic features. We will expand the test data set to refine and improve the method, while continuing to investigate other potential features that can be incorporated in the similarity measure for better results, including more anatomic structures and their relationships. In addition to using one reference model to project the CTV for a test subject as in the current method, we will investigate the possibility of using multiple similar reference models and merge their CTV projects to further reduce the maximum error.

These software tools will need to be evaluated at a clinical environment for head and neck cancer patients who undergo 3D conformal radiation therapy where the computer generated target contours can be compared to the expert drawn contours of more target subjects. The method can also be generalized for treatment of other types of cancer. Furthermore, this work can potentially be integrated with other research work such as predictions of the regional lymphatic involvement of head and neck squamous cell carcinoma based on primary tumor location and T-stage [44]. By using the lymphatic prediction, we can delineate proper target volumes for given primary tumor location and T-stage.

**Author details**
[1]School of Technology, Brigham Young University, Provo, UT, USA. [2]Department of Computer Science and Engineering, University of Washington, Seattle, WA, USA. [3]Department of Radiation Oncology, University of Washington, Seattle, WA, USA.

**Authors' contributions**
CT carried out most of the design and implementation. LS participated in its design. IK conceived of the study, and participated in its design and coordination with the clinicians. All authors read and approved the final manuscript.

**Competing interests**
The authors declare that they have no competing interests.

**References**
1.  Greenlee RT, Murray T, Bolden S, Wingo PA: **Cancer Statistics.** *CA-A Cancer J Clin* 2000, **50**:7-33.
2.  International Commission on Radiation Units and Measurements: **Prescribing, Recording and Reporting Photon Beam Therapy.** Bethesda, MD, International Commission on Radiation Units and Measurements, Report 50 1993.
3.  International Commission on Radiation Units and Measurements: **Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU Report 50).** Bethesda, MD, International Commission on Radiation Units and Measurements, Report 62 1999.
4.  Chao KSC, Wippold FJ, Ozyigit F, Tran BN, Dempsey JF: **Determination and Delineation of Nodal Target volumes for Head-and-Neck Cancer Based on Patterns of Failure in Patients Receiving Definitive and Postoperative IMRT.** *Int Journal of Radiation Oncology Biol Phys* 2002, **53**:1174-1184.
5.  Teng C, Shapiro LG, Kalet IJ: **Automatic segmentation of neck CT images.** *Proc. 19th IEEE International Symposium on Computer-Based Medical Systems* 2006, 442-445.
6.  Teng C, Shapiro LG, Kalet IJ: **Head and neck lymph node region delineation using a hybrid image registration method.** *Proc. 2006 IEEE International Symposium on Biomedical imaging* 2006, 462-465.
7.  Teng C, Austin-Seymour MM, Barker J, Kalet I, Shapiro LG: **Head and neck lymph node region delineation with 3D CT image registration.** *Proc. American Medical informatics Assoc Fall Symp* 2002, 767-771.
8.  Teng C, Shapiro LG, Kalet I, Rutter C, Nurani R: **Head and neck cancer patient similarity based on anatomical structural geometry.** *Proc 2007 IEEE international Symposium on Biomedical imaging* 2007, 1140-1143.
9.  Williams DWW: **An Imager's Guide to Normal Neck Anatomy.** *Seminars in Ultrasound, CT and MR* 1997, **18**:157-181.
10. Kruger A, Tietjen C, Hintze J, Preim B, Hertel I, Strau G: **Interactive Visualization for Neck Dissection Planning.** *Proc of the IEEE/Eurographics Symposium on Visualization (EuroVis)* 2005, 295-302.
11. Dornheim J, Seim H, Preim B, Hertel I, Strauss G: **Segmentation of Neck Lymph Nodes in CT Dataset with Stable 3D Mass-Spring Models.** *Academic Radiology* 2007, **14**:1389-1399.

12. Maleike D, Fabel M, Tetzlaff R, Tengg-Kobligk H, Heimann T, Meinzer HP, Wolf I: **Lymph node segmentation on CT images by a shape model guided deformable surface method.** *Proc SPIE Medical Imaging* 2008, **6914**.
13. Som PM, Curtin HD, Mancuso A: **An imaging based classification for the cervical nodes designed as an adjunct to recent clinical based nodal classification.** *Archives of Otolaryngol-Head and Neck Surery* 1999, **125**:388-396.
14. Som PM, Curtin HD, Mancuso A: **Imaging-based classification for evaluation of neck metastatic adenopathy.** *American Journal of Roentgenology* 2000, **174**:837-844.
15. Gregoire V, Scaliet P, Ang KK: **Clinical target volumes in conformal and intensity modulated radiation therapy.** 2004, 85.
16. Chao C, Apisarnthanarax S, Ozyigit G: **Practical essentials of intensity modulated radiation therapy.** 2004, 113.
17. Commowick O, Gregoire V, Malandain G: **Atalas-based delineation of lymph node levels in head and neck computed tomography images.** *Radiotherapy and Oncology* 2008, **87(2)**:281-289.
18. Radiation Therapy Oncology Group: [http://www.rtog.org/atlases/hnatlas/main.html], Last accessed on May 24, 2010.
19. Reed VK, Woodward WA, Zhang L, Strom EA, Perkins GH, Tereffe W, Oh JL, Yu TK, Bedrosian I, Whitman GJ, Buchholz TA, Dong L: **Automatic segmentation of whole-breast using atlas approach and deformable image registration.** *Int J Radiat Oncol Bio Phys* 2009, **73**:1493-1500.
20. Chao KSC, Bhide S, Chen H, Asper J, Bush S, Franklin G, Kavadi V, Liengswangwong V, Gordon W, Raben A, Strasser J, Koprowski C, Frank S, Chronowski G, Ahamad A, Malyapa R, Zhang L, Dong L: **Reduce in variation and improve efficiency of target volume delineation by a computer-assisted system using a deformable image registration approach.** *Int J Radiat Oncol Biol Phys* 2007, **68**:1512-1521.
21. Gorthi S, Duay V, Houhou N, Cuadra MB, Schick U, Becker Minerva, Allal AS, Thiran JP: **Segmentation of Head and Neck Lymph Node Regions for Radiotherapy Planning Using Active Contour-Based Atlas Registration.** *IEEE Journal of Selected Topics in Signal Processing* 2009, **3**:135-147.
22. Han X, Hoogeman M, Levendag P, Hibbard L, Teguh D, Voet P, Cowen A, Wolf T: **Atlas-based auto-segmentation of head and neck CT images.** *Medical image Computing and Computer-Assisted Intervention - MICCAI* 2008, 434-441.
23. VolecityAI, Velocity Medical Solutions: [http://www.velocitymedical.com/velocityai/index.shtml].
24. MIMcontouring Advanced, MIMvista: [http://www.mimvista.com/products/mimcontouringadvanced].
25. Joshi S, Davis B, Jomier M, Gerig G: **Unbiased diffeomorphic atlas construction for computational anatomy.** *Neuroimage* 2004, **23**:S151-S160.
26. De Craene M, duBois d' Aische A, Macq B, Warfield SK: **Multi-subject registration for unbiased statistical atlas construction.** *Medical image computing and computer assisted intervention LNCS* Springer 2004, **3217**:655-662.
27. Sims R, Isambert A, Gregoire V, Bidault F, Fresco L, Sage J, Mills J, Bourhis J, Lefkopoulos D, Commowick O, Benkebil M, Malandian G: **A pre-clinical assessment of an atlas-based automatic segmentation tool for the head and neck.** *Radiation and Oncology* 2009, **93**:474-478.
28. Stapleford LJ, Lawson JD, Perkins C, Edelman S, Davis L, McDonald MW, Waller A, Schreibmann E, Fox T: **Evaluation of automatic atlas-based lymph node segmentation for head-and-neck cancer.** *Int JRadiat Oncol Biol Phys* **77(3)**:959-66.
29. Tsuji SY, Hwang A, Weinberg V, Yom SS, Quivey JM, Xia P: **Dosimetric evaluation of automatic segmentation for adaptive IMRT for head-and-neck cancer.** *Int J Radiat Oncol Biol Phys* **77(3)**:707-14.
30. Min P, Halderman JA, Kazhdan M, Funkhouser TA: **Early experiences with a 3D model search engine.** *Web3D Symposium* 2003, 7-18.
31. Chen DY, Tian XP, Shen YT, Ouhyoung M: **On visual similarity based 3D model retrieval.** *Computer Graphics Forum (EG 2003 Proceedings)* 2003, **22**:223-232.
32. Tangelder JWH, Veltkamp RC: **Polyhedral model retrieval using weighted point sets.** *International Journal of Image and Graphics* 2003, **3**:209-229.
33. Ruiz-Correa S, Sze RW, Starr JR, Lin HJ, Speltz ML, Cunningham ML, Hing AV: **New scaphocephaly severity indices of sagittal craniosynostosis: A comparative study with cranial index quantifications.** *Clef Palate Craniofacial Journal* 2006, **43**:211-221.
34. Iyer N, Jayanti S, Lou K, Kalyanaraman Y, Ramani K: **Three dimensional shape searching: State-of-the-art review and future trends.** *Computer-Aided Design* 2005, **37**:509-530.
35. Tangelder JWH, Veltkamp RC: **A survey of content based 3D shape retrieval methods.** *Multimedia Tools and Applications* 2008, **39**:441-471.
36. Mattes D, Haynor DR, Vesselle H, Lewellen TK, Eubank W: **PET-CT Image Registration in the Chest Using Free-form Deformations.** *IEEE Transaction on Medical Imaging* 2003, **22**:120-128.
37. Thevenaz P, Unser M: **Spline pyramids for intermodal image registration using mutual information.** *Proceeding of the SPIE* 1997, **3169**:236-247.
38. Cordes J, Dornheim J, Preim B, Hertel I, Strauss G: **Pre-operative Segmentation of Neck CT Datasets for the Planning of Neck Dissections.** *Proc. SPiE Medical imaging* 2006.
39. Kobashi M, Shapiro LG: **Knowledge-based organ identification from CT images.** *Pattern Recognition* 1995, **28**:475-491.
40. Shelton CR: **Morphable Surface Models.** *International Journal of Computer Vision* 2000, **38**:75-91.
41. Aspert N, Santa-Cruz D, Ebrahimi T: **MESH: Measuring Errors between Surfaces Using the Hausdorff Distance.** *Proc. of the IEEE International Conference in Multimedia and Expo (ICME) 2002* 2002, 705-70.
42. Besl PJ, McKay ND: **A method for registration of 3-d shapes.** *IEEE Trans PAMI* 1992, **14**:239-256.
43. Zou KH, Warfield SK, Bharatha A, Tempany CMC, Kaus MR, Haker SJ, Wells WM, Jolesz FA, Kikinis R: **Statiscal validation of image segmentation quality based on a spatial overlap index.** *Acad Radiol* 2004, **11**:178-189.
44. Benson N, Whipple M, Kalet IJ: **A Markov Model Approach to Predicting Regional Tumor Spread in the Lymphatic System of the Head and Neck.** *Proc of the American Medical informatics Assoc Fall Symposium* 2006, 31-35.