

# HEAD AND NECK LYMPH NODE REGION DELINEATION USING A HYBRID IMAGE REGISTRATION METHOD

*Chia-Chi Teng<sup>1</sup>, Linda G. Shapiro<sup>1,2</sup>, Ira Kalet<sup>2,3</sup>*

<sup>1</sup> Department of Electrical Engineering, <sup>2</sup> Department of Computer Science

<sup>3</sup> Department of Radiation Oncology, University of Washington

## ABSTRACT

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's biomedical images will contribute greatly to the success of radiation therapy and drastically reduce the workload of radiation oncologists, who currently often draw the targets by hand on images using simple computer drawing tools. We are developing methods for automatically selecting and adapting standardized regions of tumor spread based on the location of lymph node regions in a standard or reference case, using image registration techniques. Previously available image registration techniques (deformable transformations computed using mutual information [5]) appear promising and can be supplemented by utilizing landmark correspondences in the optimization process to come closer to achieving a clinically acceptable match.

## 1. INTRODUCTION

With the rapid development of conformal radiation therapy and Intensity Modulated Radiation Therapy (IMRT) systems in the field of Radiation Oncology, it is now possible to deliver a precise dose of radiation to irregularly shaped target (tumor) volumes. A 3-dimensional target volume needs to be defined to facilitate a treatment plan, and the success of the treatment depends on knowing the exact extend of the target volume in each patient.

Radiation oncologists have adopted definitions for various components of the target volume. The Gross Target Volume (GTV) is the visible and palpable tumor mass usually visible on images (CT and MR). It is not automatically identifiable with existing image processing techniques. 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 clinical technique. Even the nodes themselves are often hard to identify in the images. The task of delineating these nodal regions, which is usually done by clinicians, is quite time consuming. Clinicians often elect to perform less aggressive, non-conforming treatment, because they do not have the time to draw the outlines of the nodal regions and CTV, even if they are confident of which node groups are likely to have disease to treat.

Inter-subject image registration methods have been the subject of extensive study in many areas of biomedical imaging applications, for example the brain mapping. We are proposing an image registration method that maps predefined head and neck nodal regions of reference (canonical) models to target patients and suggests where the CTV might be in a target image set.

## 2. IMAGE BASED NODAL CLASSIFICATION

Cervical lymph nodes are divided into regions or "levels" that are described by their anatomical locations. Traditional classification has used surgical landmarks or other physical assessment criteria, but more recently image-based classifications have been proposed [1][2] that provide a more consistently reproducible nodal staging model. Figure 1 shows an example of nodal region contours in a CT image.

## 3. LANDMARK CORRESPONDENCE

Anatomical landmarks are commonly utilized in image registration methods, which use landmark points to match the image properties in different image sets and bring them into alignment. They are also often used in combination with an entirely different registration basis, such as brain mapping. For example, Christensen et al. [3] use interactive methods to locate anatomical landmarks that define the Frankfort Horizontal Plane, the Median Sagittal Plane, and the Coronal Plane; they then use those planes to estimate the rigid registration before running non-rigid registration with a different method (elastic model). We have chosen the mandible and hyoid bones as our landmarks because of

the proximity to the nodal regions of interest, high contrast, and consistent location, size and shape.

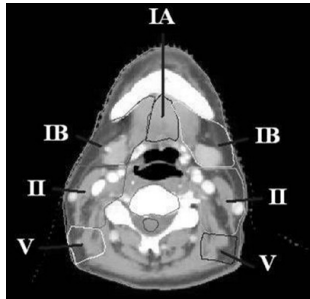


Fig. 1: Level system of lymph node classification.

### 3.1. Segmentation

Two common problems make segmenting the mandibles a non-trivial task: (1) dental implants may cause metal artifacts in CT images; (2) the teeth may be in contact and/or overlap with the maxillary. Figure 2 shows an example of such problems, where maxillary and mandible may be segmented into one structure.

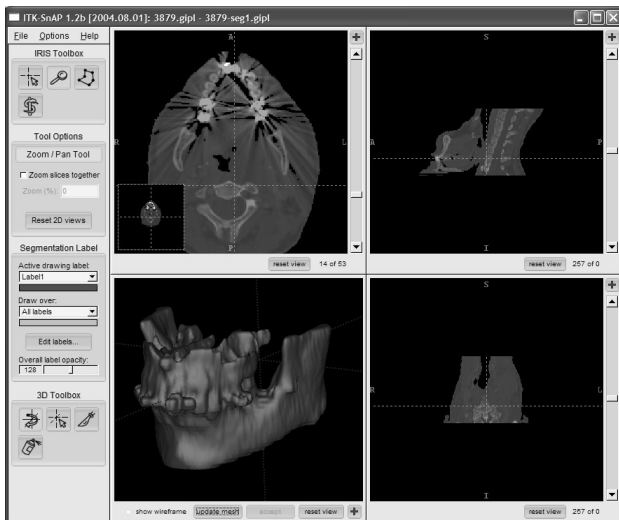


Fig. 2: Example of a noisy CT image data set with metal artifact from dental implants, and an undesirable segmentation result of mandible.

We chose to use a 3-dimensional active contour (snake) method to segment mandible and hyoid bones. The Insight Toolkit (ITK [www.itk.org](http://www.itk.org)) provides a flexible 3-D snake utility. With proper parameters for the threshold, propagation (balloon force) velocity, and curvature term, it provides improved results for mandible and hyoid bone segmentation. Figure 3 shows a better mandible segmentation from the same data set.

### 3.2. Surface-based warping

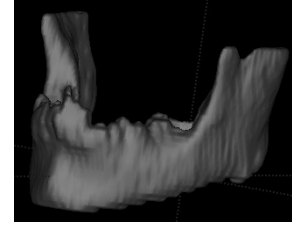


Fig. 3: An improved segmentation of the mandible from the same noisy CT image data set as shown in Fig. 2.

Once we have mandible and hyoid bone surface meshes from both the reference and target images, we can obtain correspondences between points on reference and target meshes using Shelton's algorithm [4].

Let  $\mathbf{v}_k$  be the set of landmark points sampled from the mandible and hyoid surface meshes of the reference image set and  $\mathbf{w}_k$  be their corresponding locations in the test image set resulted from the surface warping. The deformation at those landmark points is simply

$$\boldsymbol{\zeta}_k = \mathbf{w}_k - \mathbf{v}_k \quad (1)$$

## 4. IMAGE REGISTRATION

### 4.1. Previous work

Mattes and Haynor [5] 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  $\mathbf{D}(\mathbf{x}|\boldsymbol{\delta})$  [6]:

$$\mathbf{g}(\mathbf{x}|\boldsymbol{\mu}) = \mathbf{R}(\mathbf{x} - \mathbf{x}_C) - \mathbf{T}(\mathbf{x} - \mathbf{x}_C) + \mathbf{D}(\mathbf{x}|\boldsymbol{\delta}) \quad (2)$$

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  $\mathbf{x}$ . Hence the transformation parameter vector  $\boldsymbol{\mu}$  becomes [6]

$$\boldsymbol{\mu} = \{\gamma, \theta, \phi, t_x, t_y, t_z, \boldsymbol{\delta}_f\} \quad (3)$$

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

The parameter  $\boldsymbol{\delta}_f$  is the set of the deformation coefficients. The deformation was modeled on cubic B-splines [10] 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  $\lambda_j$ , each having associated x-, y-, and z- components of the deformation. The resolution

$$\rho = [\rho_x, \rho_y, \rho_z] \quad (4)$$

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

$$\Delta \rho = [\Delta \rho_x, \Delta \rho_y, \Delta \rho_z]^T = \left[ \frac{q_x - 1}{\rho_x - 1}, \frac{q_y - 1}{\rho_y - 1}, \frac{q_z - 1}{\rho_z - 1} \right]^T \quad (5)$$

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:

$$\mathbf{D}(\mathbf{x}|\boldsymbol{\delta}) = \sum_j \boldsymbol{\delta}_j \beta^{(3)}\left(\frac{\mathbf{x} - \lambda_j}{\Delta \rho}\right) \quad (6)$$

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

#### 4.2. Using landmark correspondences

Instead of initializing the deformations at the control points to zero, we propose to use the landmark correspondences to initialize the deformations at the control points at each of the multi-resolution stages. The deformation control points are set to a uniform grid

$$\lambda_j = [l\Delta\rho_x, m\Delta\rho_y, n\Delta\rho_z]^T, \quad \text{where } 0 \leq l \leq \rho_x, 0 \leq m \leq \rho_y, 0 \leq n \leq \rho_z. \quad (7)$$

and the corresponding deformation values  $\mathbf{D}(\lambda_j)$  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 Eq. (6).

Then 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

$$D'(\lambda_j) = \begin{cases} \zeta_k & \text{if } |\lambda_{jx} - v_{kx}| < \frac{1}{2}\Delta\rho_x \ \& \\ & |\lambda_{jy} - v_{ky}| < \frac{1}{2}\Delta\rho_y \ \& \ |\lambda_{jx} - v_{kx}| < \frac{1}{2}\Delta\rho_z \\ D(\lambda_j) & \text{otherwise} \end{cases} \quad (8)$$

where  $v_k$  is the closest landmark point to  $\lambda_j$  in the reference image set, and  $\zeta_k$  is the deformation at  $v_k$  obtained from the surface correspondence in Eq. (1). A new set of deformation coefficients  $\boldsymbol{\delta}_j$  is then set to the spline coefficients of the new grid of deformation values  $\mathbf{D}'(\lambda_j)$  [5]. Finally the transformation parameter vector  $\mu$  is input to the optimizer for alignment.

## 5. EXPERIMENT AND RESULTS

The test images are CT scans performed at the University of Washington Medical Center using a General Electric CT scanner. A simple segmentation process was performed to remove the bed and immobilization devices from the CT images before they were used in the segmentation and image registration processes.

Nine sets of CT images in which all of the slices are 512×512 pixels are used; the distance between slices varies between 1.25 mm and 3.75 mm for each image set. A subset of slices was selected from the base of skull to the thoracic inlet for each image set, which is most relevant to the anatomy around the cervical lymph node regions. Each image set is used as both reference and target data.

### 5.1. Qualitative evaluation

We can apply the final transformation  $\mu$  to resample the target image data. The resulting image can be compared with the reference image visually to assess the alignment qualitatively.

We compared results from Mattes' image registration method and our proposed hybrid method combining Mattes' algorithm with landmark correspondences. Figure 4 shows a sample result, where the hybrid method exhibits much better alignment visually.

### 5.2. Quantitative evaluation

The result of the image registration process can be evaluated quantitatively by measuring the 3-D Hausdorff distance between the two 3-dimensional meshes: reference target volume transformed into target space and target volume. Table 1 shows some of the results using Mattes' method. Table 2 shows the results using the new method utilizing landmark correspondences.

## 6. CONCLUSIONS

The proposed hybrid image registration method demonstrates the benefit of utilizing landmark information in the mutual information optimization method. Figure 4 shows improved anatomical alignment, and Table 2 shows

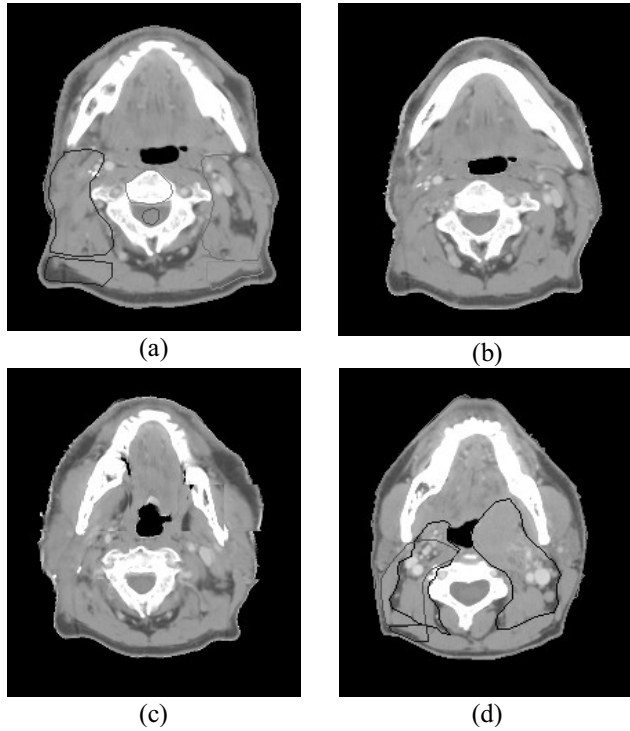


Fig. 4: Example of image registration results, showing one slice from the 3-dimensional image set (a) Reference image. (b) Warp reference image to target image using Mattes' method. (c) Warp reference image to target image using Mattes' method with landmark correspondence. (d) Target image.

Ref Test	case 3	case 4	case 5	case 6	case 7
case 3	n/a	1.334	0.938	1.2	1.453
case 4	1.443	n/a	0.876	1.913	2.047
case 5	0.911	0.996	n/a	2.209	1.649
case 6	1.249	1.368	1.074	n/a	1.389
case 7	0.804	1.044	0.76	1.274	n/a

Table 1. Results of Mattes' method: Hausdorff distance (in cm) between transformed reference mesh to target mesh for nodal region 1B-right.

Ref Test	case 3	case 4	case 5	case 6	case 7
case 3	n/a	1.296	0.967	1.523	1.175
case 4	1.06	n/a	0.875	1.448	1.079
case 5	1.106	1.474	n/a	1.119	1.323
case 6	0.998	1.91	1.014	n/a	1.142
case 7	0.87	1.127	0.967	1.061	n/a

Table 2. Results of the new method incorporating landmark information: Hausdorff distance (in cm) between transformed reference mesh to target mesh for nodal region 1B-right. Compare to Table 1.

decreased error measurement comparing to Table 1 in most cases. In cases where the new method did not improve the result, it is mainly because the nodal region contours drawn by the oncologists are affected by the physicians' clinical judgment, rather than purely following the image based classification. Although the technique shows promise, the results do not yet conform to clinical criteria.

We will continue this work by collecting more reference models, including clinical patient data with different anatomical characteristics, or various tumor stages, and analyze the results to study the effects of those different characteristics. In addition to hyoid and mandible, more anatomical structures will be used to generate landmark correspondences, for example, cervical spine and sternocleidomastoid muscle. Landmark information will also be used in the optimization process as part of the criteria for convergence. Further, integration with a radiation therapy planning system [7] will enable evaluation in a clinical setting by more clinicians on more cases.

## 7. REFERENCES

- [1] P.M. Som, H.D. Curtin, A. Mancuso, "An imaging based classification for the cervical nodes designed as an adjunct to recent clinical based nodal classification," *Archives of Otolaryngol-Head and Neck Surgery.*, vol. 125, pp. 388-396, Apr. 1999.
- [2] K.S.C. Chao, F.J. Wippold, G. Ozyigit, B.N. Tran, and J.F. Dempsey, "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. J. of Radiation Oncology Biol. Phys.*, vol. 53, pp. 1174-1184, 2002.
- [3] G.E. Christensen, A.A. Kane, J.L. Marsh, and M.W. Vannier, "Synthesis of an individual cranial atlas with dysmorphic shape," *Mathematical methods in biomedical image analysis*, pp. 309-318, 1996.
- [4] C.R. Shelton, "Morphable Surface Models," *International Journal of Computer Vision*, vol. 38, pp. 75-91, 2000.
- [5] D. Mattes, D.R. Haynor, H. Vesselle, T.K. Lewellen, and W. Eubank, "PET-CT Image Registration in the Chest Using Free-form Deformations," *IEEE Transaction on Medical Imaging*, vol. 22, pp. 120-128, Jan. 2003.
- [6] N. Aspert, D. Santa-Cruz, and T. Ebrahimi. "MESH: Measuring Errors Between Surfaces Using the Hausdorff Distance," *Proc. of the IEEE Int. Conference in Multimedia and Expo (ICME) 2002*, vol. 1, pp. 705-70. Aug. 2002.
- [7] I.J. Kalet, J.P. Jacky, M.M. Austin-Seymour, S.M. Hummel, K.J. Sullivan, J.M. Unger, "Prism: A new approach to radiotherapy planning software," *Int. J. of Radiation Oncology, Biology and Physics*, vol. 36, pp. 451-461, 1996.