4D Lung Tumor Segmentation via Shape Prior and Motion Cues

Jung won Cha[1], Mohammad M. Farhangi[2], Neal Dunlap[3], and Amir Amini,[4] *IEEE Fellow*

[1,2,4]*Medical Imaging Lab, Department of Electrical and Computer Engineering,*
[3]*Department of Radiation Oncology, University of Louisville, Louisville, KY*

Abstract— **Lung tumor segmentation is important for therapy in the radiation treatment of patients with thoracic malignancies. In this paper, we describe a 4D image segmentation method based on graph-cuts optimization, shape prior and optical flow. Due to small size, the location, and low contrast between the tumor and the surrounding tissue, tumor segmentation in 3D+t is challenging. We performed 4D lung tumor segmentation in 5 patients, and in each case compared the results with the expert-delineated lung nodules. In each case, 4D image segmentation took approximately ten minutes on a PC with AMD Phenom II and 32GB of memory for segmenting tumor in five phases of lung CT data.**

I. Introduction

The ability to track lung nodules in 4D CT can be beneficial in improving the accuracy in radiation treatment plans. However, tumor segmentation is still challenging due to the surrounding tissues with little distinction between the Hounsfield unit of tumor and non-tumor areas. Many lung nodule segmentation techniques have been proposed in 3-D. Dehmeshki et al. [1] proposed a new region growing method showing the possibility to extract various types of lung nodules. Xu et al. [2] presented the segmentation of lung nodules using dynamic programming. Ye et al. [3] used the joint spatial-intensity-shape feature for shape prior in the graph-cuts method. Moreover, Linguraru et al. [4] presented the abdominal organs segmentation using the 4D graph-cuts with shape and location constraints.

In this paper, the lung nodules segmentation in 4D is performed by graph cuts using shape prior information, motion cues, and intensity information.

II. Methodology

A. Graph-cuts segmentation

$$Energy\,(A) = \lambda \cdot \sum_{p \in P} R_p(A_p) + \sum_{\{p,q\} \in N} c \cdot B_{p,q}\,(A_p, A_q)$$

$$(1)$$

Given a 3D image volume, a graph$=< V, E >$ can be constructed with V as a set of nodes comprising image voxels and E as a set of edges which connect the nodes [5]. A set N of all node pairs $\{p, q\}$ of neighboring nodes in the set of image nodes P. To find an optimal boundary with a graph-cuts algorithm, two additional special terminal nodes are needed: these are called the source node, S, and the sink node, T. In the case of binary image segmentation, the set of labels $A_p = \{0, 1\}$ correspond to the object, tumor region, and the background, non-tumor region, respectively. Terminals correspond to the set of labels that can be assigned to nodes. The set E in the graph has two types of edges: n-link and t-link. The n-links connect pairs of neighboring nodes such as $\{p, q\}$. This neighboring system can be 6-(3D), and 26-(3D) neighborhood in the 3D image volume. The cost of n-links corresponds to similarity between the nodes. These values are normally obtained from the node interaction term $B_{p,q}$ in (1). T-links connect pixels with terminals (labels). The cost of a t-link connecting a terminal and a node corresponds to a value for assigning the corresponding label to the node. This cost is usually derived from the regional term R_p in (1). The regional term R_p assumes that the individual penalties for assigning voxel p to "object" and "background", correspondingly $\lambda \cdot R_p("obj")$ and $\lambda \cdot R_p("bkg")$ are given. For example, R_p represents how well a voxel p is assigned to the object and background intensity group, e.g., through using a probability density function expressed by histogram [5]: $R_p\,(A_p) = -lnP\,(I_p|Ap)$. The weight on the t-link $\{p, S\}$ can be calculated as $-\lambda \cdot lnP(I_p \mid "bkg")$ and on the t-link $\{p, T\}$, $-\lambda \cdot lnP(I_p \mid "obj")$ is the edge weight. The node interaction term $B_{p,q}$ contains the "boundary" properties of segmentation, and can increase as a function of similarity between p and q. When p and q have low contrast, $B_{p,q}$ is large, while very different intensities between p and q result in $B_{p,q} \approx 0$ as may be deduced from $B_{p,q}\,(A_p, A_q) = exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \delta_{A_p \neq A_q}$. In order to find the minimum s/t cut which separates two image regions, object and background, the sum of edge weights along the cut is optimized through the max-flow algorithm. For the weight setting of this algorithm, the formula in [6] are used. And we can find the minimum energy cut through the graph which cuts the edges so that the "object" is separated from the "background". In the case of binary labeled segmentation (object, background), the standard max-flow algorithm [5], [7], [8] is sufficient.

B. Active Shape Model (ASM) with Signed Distance Function

Active shape model (ASM) have been used for several medical image segmentation tasks. In its standard form, ASM models the shape boundary through principle components of landmark coordinates in a training set. However, defining landmarks and determining all correspondences requires huge human efforts and is always subject to errors especially in 3D shapes. Despite computational advantages, this representation suffers from low resolution because the space between neighboring landmarks should be interpolated and is not the real boundary. Thus, we decided to choose an implicit shape representation in which the shape boundary is represented as the zero level set of a higher dimensional surface. One standard surface function used in the literature [9] is the signed distance function ϕ. This function shows the distance between every voxel (x, y, z) and the nearest voxel on the boundary. Depending on whether the voxel (x, y, z) appears inside or outside the curve, $\phi(x, y, z)$ will be negative or positive respectively. Before extracting shape statistics we should align shapes to capture the statistics of like-parts of the shapes. Finding the exact alignment in general is a very challenging problem. However, using the signed distance function makes our analysis robust to slight misalignments [9]. So, we roughly align shapes by matching their centroids. In addition, other shape variations like scale and curvature could be captured in eigenvectors extracted by Principle Component Analysis (PCA). Unlike classical ASM, as in [9], we extract principle components of signed distance functions instead of landmark vectors and capture the most important variations in training shapes. The training set consists of training signed distance functions wrapped into the columns of training matrix $T = \{\phi_1, \phi_2, \cdots, \phi_n\}$. We start the shape statistic extraction by computing the mean of all signed distance functions: $\bar{\phi} = \sum_{i=1}^{n} \phi_i$. Consequently, shape variations are computed by PCA. More specifically, we construct the residual matrix by subtracting the mean signed distance function from each shape instance (columns of training matrix): $R = \{\phi_1 - \bar{\phi}, \phi_2 - \bar{\phi}, \cdots, \phi_n - \bar{\phi}\}$. The covariance matrix $\frac{RR^T}{n}$ is decomposed into its eigenvalues and eigenvectors using the singular value decomposition: $UDU^T = \frac{1}{n}RR^T$. U is a matrix whose columns represent eigenvectors; the direction of variations in the training set. Having eigenvectors, a new shape ϕ can be approximated by $\phi \simeq U_k w + \bar{\phi}$. U_k is the first k column of U, and $w \in \mathbb{R}^k$ is the weight vector of each entry which corresponds to one element of eigenvector and is referred to as the model parameters.

$$w = U_k^T \left(\phi - \bar{\phi}\right) \qquad (2)$$

In our algorithm, the ϕ is computed from segmented image produced by the graph-cut algorithm, and this equation refines the segmented image by transferring the signed dis-tance function to the range of variation in training database. Another constraint is also put on w wherein each w_i should lie in the range $(-2\lambda_i, 2\lambda_i)$ where λ_i is the ith ($i = 1 \sim k$) eigenvalue.

C. Motion estimation by Optical Flow

The optical flow provides a good balance between computational costs and theoretical flexibility, and finds voxel correspondences between two images by computing motion fields. With I denoting the intensity of the image voxels, the constant intensity assumption yields a differential equation commonly referred to as the optical flow constraint equation (also known as the Horn-Schunck optical flow constraint) [10]. This imposes that the gray value of a voxel should not change after displacement: $I(\xi + W) = I(\xi)$. ξ represents a function of space and time, $\xi := (x, y, z, t)^T$, and W is the displacement vector field between two images, $W := (u, v, w, 1)^T$. This yields the optical flow constraint equation: $I_x u + I_y v + I_z w + I_t = 0$. To solve for the full motion field, we need the additional constraint called the smoothness constraint, which in this paper is the sum of the squares of the Laplacians of each coordinate component. Also we approximate the Laplacians of u as $\bigtriangledown^2 u \approx \bar{u} - u$. v, and w are followed this as well. To minimize the optical flow constraint and the smoothness constraint, we can compute a new set of motion fields $\left(u^{n+1}, v^{n+1}, w^{n+1}\right)$ iteratively from the estimated derivative as follows

$$u^{n+1} = \bar{u}^n - \frac{I_x \left[I_x \bar{u}^n + I_y \bar{v}^n + I_z \bar{w}^n + I_t\right]}{\left(\alpha^2 + I_x^2 + I_y^2 + I_z^2\right)}. \qquad (3)$$

$\bar{u}^n, \bar{v}^n$, and $\bar{w}^n$ are the local average of the motion field and α is the weighting factor. v^{n+1} and w^{n+1} can also be computed the same way as (3). Moreover, using the optical flow method to compute motion fields through phases during the respiration cycle, we accomplish the 4D lung nodule segmentation in remarkably effective computing time.

D. 4D tumor segmentation via Graph-cut Optimization Including Prior Trained Shape and Motion

The graph-cuts has been widely used for image segmentation due to its advantage of global optimization and computational speed. However, the intensity-based character of graph-cuts makes difficult for the object and background which have low contrast to be segmented. As an example of difficulty with this approach, a low-contrast tumor may not be extracted by the graph-cut from surrounding tissues having similar intensities as the tumor. To solve this problem, we need shape prior information, which helps the segmentation for the tumor in the lung. Equation (4) is the modified version of (1), which is the energy function of graph-cuts using the shape prior from ASM [9], [11]. The set of shape

prior information is trained by ASM strategies.

$$Energy\,(A) = \sum_{p \in P} \left(\lambda \cdot R_p(A_p) + \alpha \cdot S_p(A_p) \right)$$
$$+ \sum_{\{p,q\} \in N} c \cdot B_{p,q}\,(A_p, A_q) \qquad (4)$$

The shape prior term S_p assumes that the signed distance function which is the distance from voxel p to the boundary of the shape. λ, α and c are the weights for the regional term, shape prior term, and boundary term, respectively. Xinjian Chen et al. [11] have introduced the segmentation method called IGC-ASM which stands for iterative graph-cuts and ASM. They use the landmark during the PCA analysis and perform 3D anatomical segmentation of organs (liver, left and right kidneys, and spleen) and foot bones.

For the training step, the input is n shapes of the tumor ground-truth from patient data. We use centroids of mask to translate them until they are aligned with a reference mask, and we obtain signed distance functions from every aligned shape. These signed distance functions are stored as shape prior in a matrix. In order to compute the mean shape and eigenvectors, we apply PCA analysis on the training data. For the test step, we are given 4D lung data sets each of which consists of five phases during the respiration cycle (from the phase10 which is max-inhale to the phase50 which is max-exhale). The tumor would be extracted from each phase, so we will have five 3D segmented tumor volumes. During the test step, on each phase the graph-cuts and ASM (with signed distance function) is performed. To run the algorithm, the initialized mean shape from the training step is used as the shape prior. Moreover, the method need two user defined point selections. One is the Volume-of-Interest (VOI) selection for a location on the center of the tumor. The original 3D volume is cropped around this center of the tumor as a cube which has the same size of the initialized mean shape cube. This selection is necessary for the shape prior, and this cropping makes computing time much faster. Another choice is the selection of seed points for the graph-cuts – seed points for object (the tumor having the mid gray intensities) and other seed points for background (the non-tumor having very low gray intensities). These seed point selections are essential for calculating the weight for the t-link in the graph-cuts. The PCA analysis refines the shape through several iterations to find feasible fitting with the segmented image until the energy value become stable. In order to extend this segmentation to 4D, we use the optical flow for motion estimation between successive phase (e.g. phase10-phase20, phase20-phase30,...). Motion fields that we obtained from the optical flow warp the mean shape on a point by point basis. After getting the warped mean shape, we perform 4D segmentation via graph cuts and the shape refinement via PCA analysis.
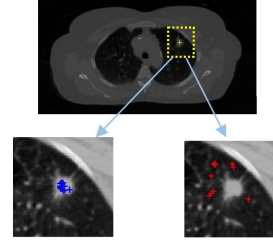

Fig. 1: **Top: The 2D axial slice from 3D image volume at the phase10 of certain patient data. The yellow dot square shows the VOI with the yellow cross indicating the defined center of tumor by user-selection on each slice. Bottom: The cropped slice from 3D volume for the selection of seed points. The blue crosses are seed points for object (tumor region), and the red crosses are seed points for background (non-tumor region).**

III. DATA

Following IRB approval, seven 4D lung CT data were collected with $3mm$ slice thickness on a Philips Brilliance Big Bore CT scanner at the James Graham Brown Cancer Center at the University of Louisville. For training to find the mean shape, we obtained the tumor ground-truth from seven data sets, and for testing, five separate data were used. The size of 4D data ranged from $512 \times 512 \times 94$ to $512 \times 512 \times 141$ through five phases. The size of the VOI is $61 \times 61 \times 61$ for all cases. The run times of 4D segmentation for each case took less than ten minutes on an **AMD Phenom II** and 32GB of memory.

IV. RESULTS AND VALIDATIONS

We performed 4D lung tumor segmentation for five patients having the solitary lung nodule. Inputs are 4D lung data sets, parameter settings, the point selection of the approximate center of tumor region, and the seed point selection of the object and the background. We will have segmentation results on the '3D through phases' which is 4D, and the change of the refined shape obtained by shape refinement on each phase. The cropped cube centered at the user-defined point is used as an input image volume for the segmentation algorithm shown on Fig. 1. The size of this cube should have the same as the size of the cube which has the initialized mean shape. At the phase10, segmentation is performed through several iterations for shape refinement to find optimal refined shape. The first row of Fig. 2(a) shows refinement of the shape during iteration, and the second row of Fig. 2(a) shows the change of segmentation results during iteration for shape refinement. After the fourth iteration, the energy function (Fig. 2(b)) stabilizes, which means that more iterations will not be needed. The shape is the signed distance function, but in order to show the shape more clearly, the shape in all figures in this paper is expressed as binary masks. Now we get the segmentation of tumor on the phase10. We calculate the motion field between

the phase10 and the phase20 via the optical flow method as shown on Fig. 3. Applying the motion field, we warp the shape. The same approach to 3D tumor segmentation on the phase10 is performed again phase20, phase30, phase40, and phase50. Finally we get five 3D segmented tumor volume through phases as shown on Fig. 4, and five 3D refined shapes as well. As we can see Fig. 4 and Fig. 5, we realize that through phases the size and the shape of tumor doesn't change much. The left of Fig. 5 shows the box plot of the tumor volume which indicates that which range of the tumor size for each patient has.
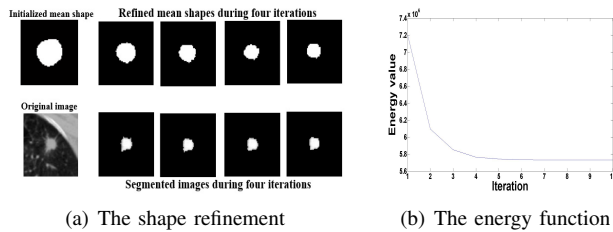

(a) The shape refinement (b) The energy function

Fig. 2: **2D axial slices at the center of 3D VOI on the phase10 through the shape refinement during four iterations (a) First row: refined shapes, second row: segmented images, (b) The plot of the energy function during ten iterations.**

To validate our result, we use the error metric, Dice Similarity Coefficient (DSC) $= \frac{2 \times |A \bigcap B|}{|A| + |B|}$. From validations in our results of 4D lung tumor segmentation with the tumor ground-truth, DSC shows around $0.7 \sim 0.8$ as shown on the right of Fig. 5. Although the DSC appears smaller than expected the reasons are as follows: tumor size is very small, so even a small non-matched voxels between the segmented image and the ground truth can affect the change of DSC.

V. CONCLUSION

Image segmentation of 4D tumor volume is a difficult task. Our 4D lung tumor segmentation via graph-cuts and signed distance ASM is a feasible solution to this problem. Signed distance ASM provides the shape prior information which prevents the unwanted surrounding tissue from being part of the segmented object. Moreover, the optical flow method extends 3D segmentation to 4D by warping the mean shape prior through time.
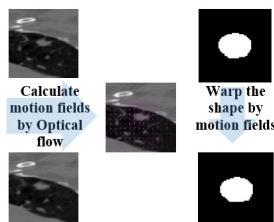

Fig. 3: **The motion field from the phase10 to the phase20 on the cropped 2D axial slice on the phase10, and the mean shape and the warped mean shape.**
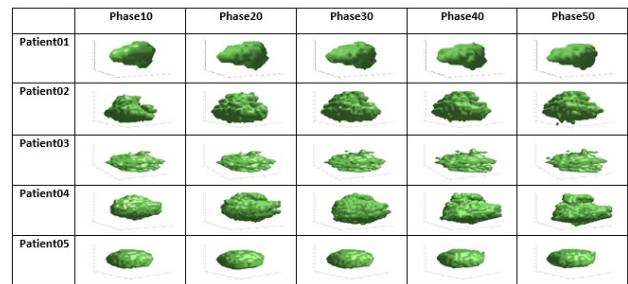

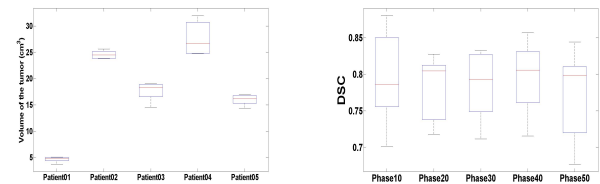Fig. 4: **3D visualization of tumors for five patients through five phases.**

Fig. 5: **Left: Box plots of the tumor volume for different patients, Right: Box plots of the Dice Similarity Coefficient through phases.**

REFERENCES

[1] J. Dehmeshki, H. Amin, M. Valdivieso, and X. Ye, "Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 4, pp. 467–480, 2008.

[2] N. Xu, N. Ahuja, and R. Bansal, "Automated lung nodule segmentation using dynamic programming and EM-based classification," in *Medical Imaging 2002*. International Society for Optics and Photonics, 2002, pp. 666–676.

[3] X. Ye, G. Beddoe, and G. Slabaugh, "Graph cut-based automatic segmentation of lung nodules using shape, intensity, and spatial features," in *Workshop on Pulmonary Image Analysis, MICCAI*, 2009.

[4] M. G. Linguraru, J. A. Pura, V. Pamulapati, and R. M. Summers, "Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT," *Medical image analysis*, vol. 16, no. 4, pp. 904–914, 2012.

[5] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 9, pp. 1124–1137, Sept 2004.

[6] J. Cha, B. Wang, N. Dunlap, and A. Amini, "3D segmentation of lung CT data with graph-cuts: analysis of parameter sensitivities," in *SPIE Medical Imaging, San Diego, CA*, 2016.

[7] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 11, pp. 1222–1239, 2001.

[8] V. Kolmogorov and R. Zabin, "What energy functions can be minimized via graph cuts?" *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 2, pp. 147–159, 2004.

[9] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Computer Vision and Pattern Recognition, Proceedings*, vol. 1, 2000, pp. 316–323.

[10] B. K. Horn and B. G. Schunck, "Determining optical flow," in *1981 Technical symposium east*. International Society for Optics and Photonics, 1981, pp. 319–331.

[11] X. Chen and U. Bagci, "3D automatic anatomy segmentation based on iterative graph-cut-ASM," *Medical physics*, vol. 38, no. 8, pp. 4610–4622, 2011.