AUTOMATIC WHOLE HEART SEGMENTATION FOR RADIOTHERAPY IN CT VOLUMES

X. Zhao, Y. Wang
Polytechnic Institute of NYU
Electrical and Computer Engineering
Brooklyn, NY

G. Jozsef
New York University School of Medicine
Radiation Oncology
New York, NY

ABSTRACT

Automatic segmentation of the whole heart in computed tomography (CT) image is crucial for efficient treatment planning of thoracic radiotherapy. In this paper, we propose a fully automatic method for whole heart segmentation of thoracic CT images. The heart is first localized in the images using a two-stage global-to-local registration scheme. A robust active shape model (Robust ASM) is proposed to reduce outliers due to blurry boundary and similar appearance. A cubic profile demonstrated better performance than the original line profile. A mean point-to-surface error of 2.87 mm was measured based on 38 images among which 26 are training data and 12 are testing data. The averaged Dice index is 0.88.

Index Terms—Whole Heart Segmentation, CT image, Radiotherapy

1. INTRODUCTION

Computed tomography (CT) has been widely used for treatment planning in radiotherapy because of its ability to reveal anatomic structures in an accurate and noninvasive way. Accurate and consistent delineation of the heart in CT images is critical for breast (lung) cancer radiotherapy. Currently, this is done manually by physicians in a slice-by-slice manner. The procedure is not only time-consuming but also inconsistent among different physicians. Automatic segmentation is increasingly important as the resolution of the CT scanner becomes higher and the amount of data become larger.

Although CT image usually has high resolution and signal to noise ratio, fully automatic segmentation of the heart is still a challenging problem due to similar intensity of neighboring organs and patient-specific variations. Also, the respiration motion and heart beating lead to blurry boundary, especially when non-gated imaging is used. Because of those difficulties, the use of the prior knowledge has been extensively investigated for medical image segmentation problem.

Active shape model (ASM) [1] is a widely used method for medical image segmentation because of its robustness and easy adaption. In the ASM, the segmentation task becomes finding the optimal model coefficients that minimize the mismatch between the image and learned appearance profiles at the shape boundary. Fritz et al. [2] applied ASM to left ventricle segmentation in cardiac multi-slice CT. It is a semi-automatic method requiring a single manual click point. Ecabert et al. [3] used Generalized Hough transform to automate the initialization and employed a deformable adaptation constrained by the piecewise affine transform to segment heart chambers. Zheng et al. [4] presented a segmentation method based on marginal space learning and steerable features for multi-chambers segmentation. Funka-Lea et al. [5] proposed a method to segment the entire heart using graph-cuts. In this paper, we proposed a whole heart segmentation method for radiotherapy, which combines a multi-stage heart localization scheme and a robust ASM. The overall segmentation workflow is shown in Fig. 1.

2. ROBUST ACTIVE SHAPE MODEL

In the ASM, shape is represented by $n$ landmark points. The landmark points are determined either manually or automatically in a set of $N$ training images. Shape vector $x_i$ for the $i$th training data is the concatenation of all landmark points $x_i = (x_{i1}, \cdots, x_{in})^T$.

Principal component analysis (PCA) is applied to aligned $X = [x_1, \cdots, x_N]$ to get the eigenvectors (eigen shapes). We only keep the eigenvectors corresponding to the largest $k$ eigenvalues $\lambda_i$ that account for 95% percent of the total energy. Without considering alignment, a shape can be approximated by the corresponding eigenvectors $\Phi, x \approx \bar{x} + \Phi b$, where $\bar{x}$ is the mean shape and $b$ is the shape deviation.
where $\bar{x}$ is the mean shape. $b$ can be computed by projecting onto the eigenvectors $b = \Phi^T(x - \bar{x})$.

Before PCA, the shapes should be aligned by translating, rotating and scaling to make the shape model independent of the size, position, and orientation of the object. To fit the model to a given new image, the transform $\mathcal{T}$ between the model points and current shape $x$ in the image coordinate system is estimated. The current shape $x$ is mapped to model coordinate system before projecting to the eigen shapes $b = \Phi^T(\mathcal{T}(x) - \bar{x})$. Projected points are mapped back to image coordinate system using $\mathcal{T}^{-1}$.

An implicit assumption of the ASM is that the transform $\mathcal{T}$ can be reliably estimated, which is not true when the shape update vector $\delta x$ contains large amount of outliers. Outliers would also distort model parameters when projecting point coordinates to the eigen shapes. In CT images, the inferior part of the heart attaches to the diaphragm (and liver), has similar image intensity as the heart tissue. Also because the CT image for radiotherapy planning is non-gated, the boundary between structures are blurry due to heart beating and respiration. Therefore, the boundary point searching is not always reliable especially in the bottom part of the heart. We propose a Robust ASM (RASM) method using RANSAC (RANdom SAmple and Consensus) scheme [6] to robustly estimate the transformation between model points and deformed points. In each RANSAC iteration, $m$ randomly selected points from all $n$ points are used to estimate the transform $\mathcal{T}$. A consensus set $C$ is identified as the points satisfying $||\mathcal{T}(x) - \bar{x}|| \leq d_{\text{ransac}}$, where $d_{\text{ransac}}$ is a distance threshold beyond which the point is considered as an outlier. $\mathcal{T}$ is then re-estimated by $x \in C$. The final transform $\mathcal{T}$ selected by RANSAC is the one that gives the largest consensus set. Model parameter $b$ is updated by

$$b = \Phi^T(\mathcal{T}(x^*) - \bar{x})$$

where

$$x^* = [x^*_l, l = 1, 2, \ldots, n], x^*_l = \begin{cases} x_l, & \text{if } x_l \in C \\
\mathcal{T}^{-1}(\bar{x}_l), & \text{otherwise} \end{cases}$$

3. SHAPE MODEL

To build the shape model from a set of training data, correspondent landmark points $x$ need to be identified for each training data. We used a volume-to-volume registration scheme to propagate landmarks and correspondences to each manually segmented heart atlas.

3.1. Atlas registration

To align all training heart atlases, a reference heart atlas is chosen and all other heart atlases are registered to the reference atlas based on a restricted affine transform. The transformation matrix of the restricted affine transform is defined as

$$\mathcal{T}_{\text{heart}} = T \cdot S \cdot R_z$$

where $T$ is the translation matrix, $S$ the scale matrix, $R_z$ the rotation matrix about $z$-axis.

Since all patients are positioned along the $z$-axis during the CT image acquisition, we only consider the rotation about $z$-axis to simplify the registration. Signed distance function (SDF) of the heart atlas is used for registration. This resulted in a more accurate result than registering the binary atlas. Squared difference is used as the similarity measure. A mean heart atlas is computed by

$$\overline{s} = \frac{1}{N} \sum_{i=1}^{N} \mathcal{T}_{\text{heart},i}(S_i).$$

3.2. Shape correspondence

One of our challenges of the heart segmentation is that manual heart atlases are often incomplete, where the superior part of the heart is inconsistent in different images. Physician usually pays less attention to this part during the manual delineation as this region is less important for breast radiotherapy.

In order to find correspondences of inconsistent atlases, we manually selected an intermediate index for each image. The intermediate index is a $z$-axis coordinate less than which the atlas is considered belonging to the superior heart, and beyond which the atlas is considered as the inferior heart. A two-part heart model is built by using planes parallel to xy-plane intersecting with the mean atlas $S$. The number of slices in the superior (inferior) part of heart is $N_{\text{sup}} (N_{\text{inf}})$. An uniform angular sampling is used to generate landmarks on the contour in each intersection slice. The number of samples in each slice is proportional to the contour perimeter. We chose $N_{\text{sup}} = N_{\text{inf}} = 10$ and 50 samples on the longest contour based on the complexity of the heart surface to represent. The model has a total 918 points and 1832 faces. To propagate sampled landmark points in the mean atlas to training atlases, each aligned training atlas is sampled by the same number of slices and angles as for the mean atlas. The sampled points is transformed back to the original image by $\mathcal{T}_{\text{heart}}^{-1}$, resulting the landmark points $x_l$ for image $I_l$. Since the heart has a regular shape, sampled landmark points in different training atlas based on aforementioned method are considered correspondences.

4. APPEARANCE MODEL

In the original ASM paper, line profile perpendicular to the surface is used as the appearance feature. The mean and the covariance matrix of line profiles in all training images are computed. During search, the point with the most close profile to the mean profile is chosen. In this study, we used a cubic profile with $k$ points in each dimension and a sampling spacing $d_s$ oriented along the normal direction of the surface at each landmark point. The squared difference between cubic profiles is used as the search criterion. Our experiments showed that cubic profile gave better segmentation result than line profile in the original ASM.

5. INITIALIZATION

Although ASM is less sensitive to the initial pose compared with other locally deformable methods, e.g. Active Contour,
a good initialization is essential to get stable and accurate segmentation result.

We used a global-to-local registration scheme to estimate the heart pose. Given the transformation $T_{\text{heart},i}$ between the reference atlas and each training atlas, the mean image is defined as $\bar{I} = \frac{1}{N} \sum_{i=1}^{N} T_{\text{heart},i}(I_i)$. First, a restricted affine transformation $T_{\text{lung}}$ (sec. 3.1) between the mean image $\bar{I}$ and a new image $I_0$ is estimated based on the centroid, orientation and range of the lung atlases of mean and new image. The lung atlas was estimated using thresholding and connected component analysis inside the chest cavity.

In the second step, a transform $T_{\text{local}}$ between $\bar{I}$ and $I_0$ is estimated based on image-intensity based registration. A binary mask image that cover $\bar{S}$ and extend to a distance $d_{\text{mask}}$ is used as the registration mask to consider the image difference only within the mask. This mask helps us to focus only on the heart alignment by excluding other non-relevant region. In practice, we found the registration result is not sensitive to $d_{\text{mask}}$ ranging from 20-80mm, and we chose $d_{\text{mask}} = 40\text{mm}$ for all experiments. The overall transformation for heart initialization is $T_{\text{init}} = T_{\text{local}} \cdot T_{\text{lung}}$.

### 6. DATA AND RESULTS

#### 6.1. Image Data

38 noncontrast enhanced 3-D CT scans acquired from 38 prone-positioned breast cancer patients after segmental mastectomy were used for evaluation. All scans were acquired in New York University Radiation Oncology department under the trial 05-181 for the purpose of identifying optimal individual positioning for breast treatment. Images have a resolution of $512 \times 512$ pixel in plane and slices number between 80 – 110. The in-plane pixel sizes range from 1.10mm to 1.27mm and the slice thickness are all 3.75mm. Manual segmentation is done by physicians using Varian’s Eclipse treatment planning system.

#### 6.2. Experiment

Out of the 38 subjects, 26 were used as training data and 12 were used as testing data. To evaluate the segmentation result, Dice index and mean distance error (MDE) from the landmarks of manual segmentation to the final deformed mesh surface are used as evaluation measurements. For the superior heart, manual segmentations are inconsistent since there is no clear boundary among the atriums, aorta and pulmonary artery. To better evaluate our segmentation result, we first determine the slice below which manual and auto segmentation both present. The Dice and distance measures are only computed in the region below this slice for fair comparison. Fig. 2 showed the error measures after initialization and final results using different methods. For the final result, we measured averaged Dice index of 0.88 and MDE of 2.87mm. First 26 points in Fig. 2 are training data and the last 12 points are testing data. Fig. 3 shows segmentation results of two patients. From the first row, ASM without RANSAC (blue) generated irregular shape due to noisy deformation. Line profile (cyan) can not locate the correct inferior boundary of heart. The second row is the worst case, where the heart had an unusual concave shape which cannot represented well by training data.

#### 6.2.1. RANSAC evaluation

In [7], a different scheme of using RANSAC in ASM framework was proposed, to be denoted RASM2. They applied RANSAC to find a consensus set of points to estimate the eigen coefficients $\mathbf{b}$, while in our method the inlier found by RANSAC for transformation estimation is used to compute $\mathbf{b}$. Box plots of the registration by different methods are shown in Fig. 4, where ASM refers to original ASM with cubic profile. RASM had significant gain compared with original ASM.

---

Fig. 2. Dice index and MDE of initialization (green), ASM with cubic profile (blue), RASM with line profile (cyan) and RASM with cubic profile (red).

Fig. 3. (a)-(c) showed 3-view (transverse, coronal and sagittal) results of patient #13 (MDE 2.37mm, (d)-(f) showed results of patient #21 (MDE 4.38mm, worst case)). Green line is the manual contour, red line is the result of RASM and cubic profile, blue and cyan are results by ASM with cubic profile and RASM with line profile (PCA), respectively.
Fig. 4. Averaged registration error for ASM without RANSAC and with different schemes. The upper and lower edges of the blue box are the 25th and 75th percentiles. Black lines are extreme values. Red line is the median value.

Fig. 5. Effect of profile models on the segmentation result (MDE from 4.41 mm to 2.87 mm), especially for the worst cases where original ASM totally failed. Our method also outperformed RASM2 (MDE 3.04 mm) as showed in Fig. 4.

6.2.2. Profile model evaluation
We compared the result using cubic profile and line profile appearance model used in the original ASM model. For the line profile, the length of the profile is evaluated at $l = 11$, which gave best performance. The sampling spacing is fixed at 1 mm. For the cubic profile, the cubic length is evaluated at $k = 5$ and the sampling spacing $d_c = 2$ mm. We used the squared difference for the cubic profile matching. For the line profile, we compared both Mahalanobis distance and coefficients difference based on PCA decomposition of the covariance matrix of the line profile as discussed in [8]. All experiments in this section used RANSAC for parameter estimation. As shown in Fig. 5, the block matching based method outperformed line profiles (6.15 mm using Mahalanobis distance and 3.44 mm for PCA based distance measure).

We also compared proposed method with Graph cut (GC) and Region growing (RG) method. A small region inside the heart is selected as the foreground seed for GC and RG, the region far away from the estimated initial heart is used as background seed for GC. The foreground and background intensity distributions are calculated from training images. From Fig. 6, our method had significant improvement compared with GC and RG. The gain mainly comes from the global shape constraint, which is absent in GC and RG.

7. DISCUSSION AND CONCLUSIONS
Although the segmentation of cardiac structures in CT images has been extensively studied, there are only a few studies on whole heart segmentation. The work by Funka-Lea et al. [5] using Graph cut on whole heart had an MDE of 5.5 mm. Chen et al. [9] reported a 6.4% error in terms of area difference using a classification based method for whole heart segmentation. [3] and [4] segmented four chambers of the contrast enhanced CT images and reported MDE $\sim 1$ mm. Our study, on the other hand, focused on whole heart segmentation in noncontrast enhanced CT image.

In this paper, a fully automatic whole heart segmentation method is proposed. A mean point-to-surface error of 2.87 mm was measured on 38 images. The proposed method would facilitate automatic treatment planning for thoracic radiotherapy.

8. REFERENCES