The Third International Workshop on Pulmonary Image Analysis

Beijing, September 20, 2010

EDITED BY

Matthew Brown
Marleen de Bruijne
Kai Ding
Bram van Ginneken
Atilla Kiraly
Jan-Martin Kuhnigk
Jamie McClelland
Kensaku Mori
Joseph Reinhardt

ISBN-13: 978-1-4537-7600-1

Preface

These are the proceedings of the third edition of the International Workshop on Pulmonary Image Analysis held on September 20, 2010 at MICCAI 2010 in Beijing. The two first events, organized also as workshops in conjunction with the MICCAI conferences in 2008 in New York and in 2009 in London, were very successful in bringing together researchers in pulmonary image analysis to discuss new developments in this growing field. For this edition, we received a large number of high quality papers that received a detailed review by two to four reviewers from the organizing team or from guest reviewers. A total of eight papers have been accepted for oral presentation at the workshop, in sessions on segmentation, classification and quantification, and registration. Six papers and one software were selected for poster presentation. Together these papers cover a wide range of topics within the field of pulmonary image analysis. We are also very happy that Eric Hoffman, Professor of Radiology, Medicine and Biomedical Engineering at the University of Iowa and director of the Iowa Comprehensive Lung Imaging Center, has agreed to give the keynote lecture of the third International Workshop on Pulmonary Image Analysis. Dr Hoffman was involved in the development of CT in the very early days and continues to lead forefront research in unravelling the mechanisms of pulmonary ventilation and lung diseases.

We would like to take this opportunity to thank the MICCAI 2010 organizers for their organizational support and their willingness to host this event, and all the reviewers for helping us with the paper selection. We acknowledge the generous contributions of MedQIA, VIDA Diagnostics, and MeVis Medical Solutions which helped make this event possible.

Matthew Brown
Marleen de Bruijne
Kai Ding
Bram van Ginneken
Atilla Kiraly
Jan-Martin Kuhnigk
Jamie R. McClelland
Kensaku Mori
Joseph M. Reinhardt

Organization

Organizers and Program Committee

Matthew Brown *(UCLA Radiological Sciences, Los Angeles, USA)*

Marleen de Bruijne *(Department of Computer Science, University of Copenhagen, Denmark; Biomedical Imaging Group Rotterdam, Erasmus MC - University Medical Center Rotterdam, The Netherlands)*

Kai Ding *(Department of Biomedical Engineering, The University of Iowa, Iowa City, USA)*

Bram van Ginneken *(Radboud University Nijmegen, The Netherlands)*

Atilla Kiraly *(Siemens Corporate Research, Princeton, USA)*

Jan-Martin Kuhnigk *(Fraunhofer MEVIS, Bremen, Germany)*

Jamie McClelland *(Centre for Medical Image Computing, University College London, London, United Kingdom)*

Kensaku Mori *(Department of Media Science, Graduate School of Information Science, Nagoya University, Japan)*

Joseph Reinhardt *(Department of Biomedical Engineering, The University of Iowa, Iowa City, USA)*

Reviewers

Matthew Brown	Bram van Ginneken	Jamie R. McClelland	Michael Schmidt
Marleen de Bruijne	André Homeyer	Jan H. Moltz	Lauge Sørensen
Kunlin Cao	Atilla Kiraly	Kensaku Mori	Dörte van Straaten
Kai Ding	Jan-Martin Kuhnigk	Joseph M. Reinhardt	Youbing Yin
Kaifang Du	Pechin Lo	Eva van Rikxoort	

Sponsoring

This workshop is sponsored in part by contributions from MedQIA, MeVis Medical Solutions, and VIDA Diagnostics.

Table of Contents

Oral Presentations

Poster Presentations

Software Demonstrations

Oral Presentations

Fast and Memory Efficient Segmentation of Lung Tumors Using Graph Cuts

Nicolas Lermé[1,2], François Malgouyres[1], and Jean-Marie Rocchisani[3,4]

(1) LAGA CNRS UMR 7539, (2) LIPN CNRS UMR 7030, (3) SMBH
Université Paris 13 –Avenue J.B. Clément
93430 Villetaneuse - France
(4) Hôpital Avicenne, 93009 Bobigny - France nicolas.lerme@lipn.univ-paris13.fr,
malgouy@math.univ-paris13.fr,
jean-marie.rocchisani@univ-paris13.fr

Abstract. In medical imaging, segmenting accurately lung tumors remains a quite challenging task when they are directly in contact with healthy tissues. In this paper, we address the problem of extracting interactively these tumors with graph cuts. The originality of this work consists in (1) reducing input graphs to decrease drastically memory consumption when segmenting a large volume of data and (2) introducing a novel energy formulation to inhibit the propagation of the object seeds. We detail our strategy to achieve relevant segmentations of lung tumors and compare our results to hand made segmentations provided by an expert. Comprehensive experiments show how our method can give solutions near from ground truth in a fast and memory efficient way.

Keywords: segmentation, lung tumor, graph cut, reduction.

1 Introduction

Since last years, accurate measurements of lung tumors sizes has become a challenging task for staging and assessing tumor response to treatments or its progression. Revised RECIST criterions, largely used by radiologists, are based on the measurement of one diameter on a few number of lesions [24], and suffer from a lack of reproducibility [22]. Alternatively, tumor volumetry has been proposed to overcome those difficulties in order to improve the staging of nodules [5], the evaluation of tumor aggressiveness [18], tumor response to chemotherapy [3,26] or to radiotherapy [16] and the progression rate of tumors [18] or metastases [15]. Moreover, it becomes a necessary tool for the automatic screening of lung nodules on CT scans, and is currently on evaluation on ongoing trials [23]. Several methods have been proposed to deal with the different kind of objects to segment. Nodules are homogeneous spheroid of small size. Masses and tumors have larger sizes and irregular shapes, and may be necrotic. All may be connected to some extent to vessels, to the pleura wall, or to the mediastinum. To tackle this issue, methods make often use of morphological operators [9,10,17]. A classification of those methods can be found for instance in [21] and [5].

Among semi-automatic approaches of segmentation based on level-sets and (geodesic) active contours, graph cuts have become in few years a leading method since the introduction of a fast maximum-flow/minimum-cut algorithm [2]. In contrast to other

methods, graph cuts have the ability to solve quickly a wide range of problems in computer graphics such as image segmentation [1], while achieving a global minimum of the energy function.

Recently, Ye *et al.* have used this technique for automatically segmenting lung nodules using a volumetric shape index [25]. Since nodules have presumably an elliptical shape, they can select the appropriate range of index values for segmenting nodules. However, segmenting lung tumors of various shapes is a much more difficult task. Tumors might indeed be connected to healthy tissues and it is not possible to distinguish the tumor and the healthy tissues by only using simple features like the gray levels. The correct segmentation can therefore only be achieved thanks to the interaction of an expert. To our knowledge, this is the first paper to tackle this problem using graph cuts. In this paper, we propose a semi-interactive graph cut-based method for segmenting lung tumors. An overview of the approach is given on Figure 1. First, we

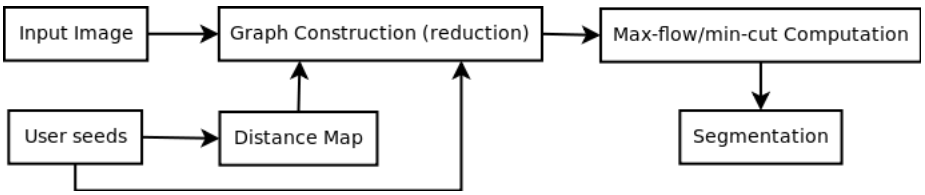


Fig. 1: Flow diagram of our approach.

compute a distance map from the object seeds for lowering the "seeds propagation". Then during the graph construction, we reduce the input graph by deciding locally which nodes are really useful for the minimum-cut computation according to [12]. Typically, the nodes are located around the contours of the object to segment. Finally, we compute the minimum-cut and get the final solution.

The rest of this paper is organized as follows. In section 2, we review the graph-cuts framework. We detail our strategy for reducing graphs in section 3 while we introduce our novel energy formulation in section 4. Finally, we validate experimentally our algorithm on several CT images in section 5.

2 Graph cuts framework

Let us first review the graph cuts framework. In this setting, an image I is a function defined over a finite discrete set $\mathcal{P} \subset \mathbb{Z}^d$ $(d > 0)$ that maps each point $p \in \mathcal{P}$ to a value $I(p)$. Usually, $\mathcal{P}$ correspond to a square when $d = 2$, a cube when $d = 3$ and a cube during a time interval when $d = 4$. A binary segmentation of the image is defined by a mapping u that assigns to each element of $\mathcal{P}$ the value 0 for the background and 1 for the object. We write $u \in \{0, 1\}^{\mathcal{P}}$.

In [1], Boykov and Jolly showed that the image segmentation problem can be efficiently solved by minimizing a Markov Random Field of the form:

$$E(u) = \beta \cdot \sum_{p \in \mathcal{P}} E_p(u_p) + \sum_{\substack{p,q \in \mathcal{P} \\ q \in \mathcal{N}(p)}} E_{p,q}(u_p, u_q), \tag{1}$$

among $u \in \{0,1\}^{\mathcal{P}}$ and for $\beta > 0$. The neighborhood system $\mathcal{N}(p)$ is in practice either

$$\begin{array}{ll} \mathcal{N}_0(p) = \{q : \sum_{i=1}^{d} |q_i - p_i| = 1\} & \forall p \in \mathcal{P}, \text{ or} \\ \mathcal{N}_1(p) = \{q : |q_i - p_i| \leq 1,\ \forall 1 \leq i \leq d\} & \forall p \in \mathcal{P}, \end{array}$$

where p_i denote the ith coordinate of the point p and $|.|$ denotes the modulus. (in this paper $|.|$ also denotes the cardinality of a set, the notations will note not be ambiguous once in context). The above neighborhood systems correspond to the classical 4-connectivity and 8-connectivity when $d = 2$. Beside on the border of the image/volume, we have for any d and any $p \in \mathcal{P}$: $|\mathcal{N}_0(p)| = (2d)$ and $|\mathcal{N}_1(p)| = 3^d - 1$. In practice, larger neighborhood systems (i.e. $\mathcal{N}_1$) yield better results but increase running time and memory consumption. Typically, we have $|\mathcal{E}_n| \sim |\mathcal{P}|.|\mathcal{N}|$, where $|.|$ denotes cardinality. In the sequel, the terms "connectivity 0" and "connectivity 1" will denote the use of a $\mathcal{N}_0$ and $\mathcal{N}_1$ neighborhood, respectively.

As usually, the region term $E_p(.)$ in (1) favors the belonging of each pixel/voxel to either the background or to the object. It is deduced from the input data, an object seed $\mathcal{O}$ and a background seed $\mathcal{B}$. The regularity term $E_{p,q}(.)$ penalizes neighboring pixels p and q having different labels. The weight of the penalization depends on the difference $|I(p) - I(q)|$ and favors boundaries located at pixels/voxels with a strong gradient. Generally speaking, the definition of E_p and $E_{p,q}$ depends on the considered application.

According to [8], the minimizer of the energy (1) corresponds to a minimum-cut in a graph that can be efficiently computed by the algorithm proposed in [2]. In this context, the directed weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, c)$ consists of a set of nodes $\mathcal{V} = \mathcal{P} \cup \{s, t\}$, a set of edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ and a positive weighting function $c : \mathcal{E} \rightarrow \mathbb{R}^+$ defining the edge capacity. Notice that two special nodes are distinguished from $\mathcal{V}$: the source node s ("object" terminal) and the sink node t ("background" terminal). After the computation of the minimum cut we set $u_p = 1$, if p is connected to s and $u_p = 0$ otherwise. Moreover, the set of edges $\mathcal{E}$ is split into two disjoint sets $\mathcal{E}_n$ and $\mathcal{E}_t$ denoting respectively n-links and t-links. The t-links are the edges connecting the terminal nodes s or t to the pixels/voxels and the n-links are the edges connecting pixels/voxels.

3 Reducing graphs

To obtain high-resolution output, graph cuts must build huge graphs containing several billions of nodes and even more edges. Such graphs may sometimes do not fit in central memory. To solve this issue, some authors have recently proposed heuristics [13,14,20,7]. However, these algorithms can easily get trapped in local minima of the energy. Also, these algorithms often fail to recover details. This is a real drawback since thin structures like blood vessels or nodules are ubiquitous in medical imaging. The only exact alternative is [11], but it has not been developed for the purpose of image segmentation.

Thus, segmenting high-resolution data using graph cuts require a prohibitive amount of memory. For instance, the maximum-flow algorithm described in [2] allocates

$24|\mathcal{P}| + 14|\mathcal{E}_n|$ bytes [1]. Table 1 shows that for a fixed amount of RAM, the maximum volume size decreases quickly as dimension d increases. Nevertheless, as showed in a

	Connectivity 0	Connectivity 1
2D	6426	4459
3D	319	219
4D	68	45

Table 1: Maximum size of a square image for which the graph fits in 2GB of RAM.

previous paper [12], most of the nodes in the graph are useless during the maximum-flow computation. They are indeed not traversed by any flow. Then, one would like to extract the smallest possible graph $\mathcal{G}' = (\mathcal{V}', \mathcal{E}', c)$ from $\mathcal{G}$ while keeping a minimum cut u' identical (or very close) to u. In other words, we want to minimize $|\mathcal{V}'|$ under the constraint that $u \simeq u'$. In fact, this is an ideal optimization problem which we will not try to solve, because the method for determining $\mathcal{G}'$ also needs to be (very) fast. We will rather consider heuristics aiming at that goal.

First, let us introduce some definitions before describing our method for building $\mathcal{G}'$. In accordance with the graph construction given in [8], we consider (without loss of generality) that a node is linked to at most one terminal:

$$(s,p) \in \mathcal{E}_t \Rightarrow (p,t) \notin \mathcal{E}_t, \qquad \forall p \in \mathcal{P}.$$

We also summarize the capacities on the t-links connected to any node $p \in \mathcal{P}$:

$$c(p) = c(s,p) - c(p,t).$$

Let us consider a square window B of size $(2r+1)$ $(r > 0)$ centered at the origin. We denote by $\widetilde{B}_p$ the translation of B at a point $p \in \mathcal{P}$: $\widetilde{B}_p = \{b + p \mid b \in B\}$. For $Z \subset \mathcal{P}$, we also denote by $\widetilde{Z}_B = \bigcup_{p \in Z} \widetilde{B}_p$ the dilation of Z by B.

The intuitive idea for building $\mathcal{G}'$ is the following: removing the nodes in any $Z \subset \mathcal{P}$ such that pixels/voxels in Z are not directly connected to the sink t and the flow that might come into the region $\widetilde{Z}_B \setminus Z$ suffices to saturate the edges located around $\widetilde{Z}_B$ (see Figure 2). Building such sets Z is done by testing each pixel p of Z. Thus, the nodes in $\mathcal{G}'$ are typically located around the contours of the object to segment. Assuming

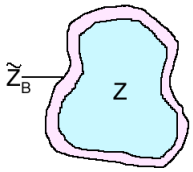


Fig. 2: Principle of the reduction. The nodes from Z are removed because every node $p \in Z$ satisfy (2). Remaining nodes are typically located in the narrow band $\widetilde{Z}_B \setminus Z$.

[1] This corresponds to the max-flow algorithm v2.2 freely available at `http://www.cs.cornell.edu/People/vnk/software.html`

that all capacities on n-links are smaller than one (which remains true for all the energy models in segmentation), we use a more conservative condition for testing each individual pixel $p \in Z$ [12]:

$$\begin{cases} \left(\forall q \in \widetilde{B}_p, c(q) \geq \delta\right) & \text{or} \\ \left(\forall q \in \widetilde{B}_p, c(q) \leq -\delta\right), \end{cases} \tag{2}$$

where $\delta = \frac{P(B)}{(2r+1)^2-1}$, with

$$P(B) = \max(|\{(p,q) : p \in B,\ q \notin B \text{ and } p \in \mathcal{N}(q)\}|, \\ |\{(p,q) : p \in B,\ q \notin B \text{ and } q \in \mathcal{N}(p)\}|).$$

For any p satisfying (2), p is only connected to s (respectively t) and the flow that might come in (respectively come out) through t-links in $\widetilde{B}_p \setminus \{p\}$ suffices to saturate the n-links going out (respectively going in) of $\widetilde{B}_p$. The pixel/voxel p is not needed and can be removed from $\mathcal{G}$. The subgraph $\mathcal{G}'$ is now fully determined by the set of nodes

$$\mathcal{V}' = \{p \ \in \ \mathcal{P} \text{ not satisfying (2)}\} \cup \{s,t\}.$$

Experiments presented in [12] confirm the intuitive dependence between the reduction rates and the parameters of the model. For instance, the capacities $c(q)$ are obtained by multiplying a quantity by the parameter β of (1). Looking at (2), it is straightforward to see that the test is satisfied on a smaller set of pixels/voxels if β decreases. In fact, β small corresponds to a strong regularization. In such a situation, we need a larger window radius to obtain a smaller δ. The latter results in wide bands around the object contours. Conversely, this results in narrow bands around the object contours when β is large. The result of such a reduction is illustrated in Figure 3. In our experiments, we always take $\beta = 3$ and $r = 1$. Additionally, the condition (2) can be tested through an easy to implement "non-optimized" algorithm with a worst-case complexity of $O(|B|)$. However for large window radii, such an algorithm cannot handle images of large size and large dimension d. Decomposing (2) along the dimensions d speed up significantly the previous algorithm. This yields a test whose computation is of complexity $O(1)$

(a) Image and seeds (b) Graph $\mathcal{G}'$ (c) Segmentation

Fig. 3: Illustration of the reduction for segmenting a CT image ($r = 1$). Light gray pixels correspond to the nodes belonging to $\mathcal{G}'$ (middle). Object and background seeds are superimposed on the original image (left). On the right image, the segmentation is superimposed in blue.

(except for image borders). In particular, its complexity is independent of the window radius. Finally, we have both theoretical and empirical evidence suggesting that this reduction scheme provides an exact solution (see [12] for details).

4 Energy function

The most famous graph cut-based energy model for image segmentation was proposed by Boykov and Jolly in [1] (see below). Total Variation-based models have also been proposed (see [19]). To obtain good results, those models require the colors in the object to be different from the colors of the background. This requirement is not efficient when segmenting lung tumors in CT images, because tumors and healthy tissues appear in the same range of intensities. Moreover, in many cases (and in our experiments), the tumor is attached to the healthy tissues and the corresponding area of the image has a uniform color.

To solve this issue, we propose to add in our energy a prior on the location of the tumor. The prior is obtained from the location of the object seeds. This leads to a modification of the original Boykov/Jolly's energy model [1]. We take the same regularity criterion:

$$E_{p,q}(u_p, u_q) = g(p,q) \cdot |u_p - u_q| \qquad \text{and} \qquad g(p,q) = \frac{1}{d_e(p,q)} \cdot exp\Big(- \frac{|I(p) - I(q)|^2}{2\sigma^2} \Big),$$

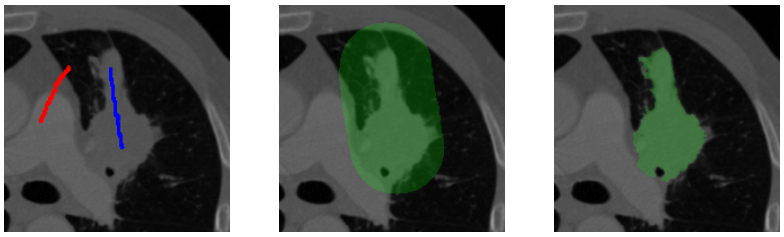
where d_e is the Euclidean distance between p and q, I is the original image and $\sigma > 0$. The region term is defined in Table 2. The sets $\mathcal{O}$ and $\mathcal{B}$ correspond respectively to

	$p \in A_{\sigma_A}$	$p \notin A_{\sigma_A}$
$E_p(u_p = \text{"}bkg\text{"})$	$-log\ \big[Pr(I(p) \mid p \in \mathcal{O}) \times exp\Big(- (\frac{d(p,\mathcal{O})}{\sigma_a})^2 \Big)\big]$	$+\infty$
$E_p(u_p = \text{"}obj\text{"})$	$-log\ \big[Pr(I(p) \mid p \in \mathcal{B})\big]$	0

Table 2: Definition of the region term.

object and background seeds provided by the user, the probability distributions are estimated according to [1], $d(p, \mathcal{O})$ is a distance function between the point $p \in \mathcal{P}$ and the set $\mathcal{O} \subset \mathcal{P}$ and $\sigma_a > 0$ is a parameter. The parameter σ_a controls how far the object seeds propagate from their location and then defines an area of influence A_{σ_a}. Beyond this area, the nodes are only linked to the background terminal with a large weight. This ensures both that the algorithm categorizes them as background pixels or voxels and that the capacity of the corresponding t-link is sufficiently high for removing the node from the graph. Although the parameter σ_A is an important parameter that impact the way of positioning the seeds in the image, we always take $\sigma_a = 10$ in our experiments.

The main difference between the proposed energy and [1] lies in the distance term. The function d is defined as $d(p, \mathcal{O}) = \min\{dist(p,q) \mid q \in \mathcal{O}\}$, where $dist$ denotes the distance between two points. We have made two attempts for $dist$:



(a) Image and seeds (b) Euclidean distance (c) Geodesic distance

Fig. 4: Area of influence for an Euclidean and a geodesic distance. Here, we set $\sigma_a = 40$.

- The Euclidean distance. In this case, the distance between a set and a point is efficiently computed with the algorithm described in [4]. We mostly use it for the purpose of illustration.
- The geodesic distance is according to the graph metric where the distance between a node $p \in \mathcal{P}$ and a node $q \in \mathcal{P}$ is defined as:

$$dist(p,q) = \begin{cases} \sqrt{(I(p) - I(q))^2 + |p - q|^2} & \text{if } q \in \mathcal{N}(p), \\ 0 & \text{otherwise.} \end{cases}$$

  In this latter case, the distance transform is computed with [6].

The area of influence for the above two metrics is displayed through an example on Figure 4. The green color corresponds to the region where the exponential in the region term is greater than some $\epsilon \simeq 0$. Observe how the geodesic distance better sticks to the tumor boundaries than the Euclidean distance. In particular, it only has a limited overflow on the healthy tissues.

5 Experimental results

In this section, we present experiments for segmenting a set of ten 3D CT images consisting both of nodules, masses and tumors (see Table 3 for more information). Each volume has a size of $512 \times 512 \times 50$ except T8 which has a size of $512 \times 512 \times 316$. All experiments are performed in connectivity 1. Objects to segment may present a very different contrast with their surrounding structures among the images. Since the parameter σ is contrast-sensitive, we are constrained to use different values for this parameter. For example, when the average gradient around the object become lower, we need a smaller σ. Then, the edges around the object are more likely to belong to the minimum-cut because they become cheaper to cut. The automatic tuning of σ is left for future work. In this setting, we use $\sigma = 0.2$ for all images except for T8 where $\sigma = 2$ and T7 where $\sigma = 0.05$. Note that a sub-volume is automatically extracted for all images (except for T8 where the border is sufficiently high to encompass the whole volume) by considering an extra border of 60 pixels around the object seeds for speeding up the segmentation.

First, we evaluate our algorithm with hand made segmentations provided by an expert, for all CT images. Table 4 contains statistics on the differences between the

segmentation and the ground truth. We use several evaluation measures [2]. Table 4 shows promosing results. For all images, we always get a Dice Coefficient greater than 70% while having a mean maximum distance less than 20mm between the ground truth and the segmentation.

We also evaluate our method in a qualitative manner. Figure 7 shows the segmentations obtained at equally spaced z-values for images T1, T8 and T9 (see Figure 5). For illustrating the propagation of seeds, the seeds in the Figure were chosen on equally spaced on z but for different values. Thus, one can observe how the seeds propagate around object seeds, avoiding us to mark every slice. Compared to the ground truth, the segmentation of T1 is very close, while the segmentations obtained for T8 and T9 differ sligthly. This also illustrate the difficulty to extract tumors/masses with a large connection to healthy tissues and the ability of our method to segment such objects.

Secondly, we compare the performance of standard graph cuts against our method in terms of speed and memory consumption (see Table 5) for segmenting the CT images using the same set of seeds and parameters as previously. Experiments were performed on an Athlon Dual Core 6000+ 3GHz with 2GB RAM. Times are averaged over 10 runs. Table 5 also indicates the proportion of object seeds with respect to the tumor volume in the ground truth. This provides an objective measure of the interaction for assessing the effort required by the user for positioning the seeds. The results obtained show that our method performs a little bit faster using 7 to 500x less memory while getting exactly the same solution. Note that a relatively small amount of seeds is necessary for segmenting all images.

Generally, the segmentation time depends both on the image size and the skill of the user for positioning the seeds not too far from the contours of the object to segment. The segmentation accuracy also depend directly on the seeds location but additionnal corrections can be done quickly if some parts of the object are uncorrectly labeled. The computation of the distance map, the building of the graph and the computation of the minimum-cut take only few seconds. Thus, our method demonstrates its ability to segment lung tumors quickly without requiring much effort if it is supported by a good graphical user interface.

Fig. 5: Overall context of lung tumors T1 (left), T8 (middle) and T9 (right).

[2] A detailed view of these measures is available at `http://lts08.bigr.nl/about.php`

Tumor	Type	Resolutions (x,y,z)	Description
T1	Mass	$0.68 \times 0.68 \times 3$	Mass of the upper right lobe (CT)
T2	Nodule	$0.70 \times 0.70 \times 1$	Nodule of the right apex (CT)
T3	Nodule	$0.68 \times 0.68 \times 3$	Nodule of the lower right lobe (CT)
T4	Tumor	$1.17 \times 1.17 \times 1.5$	Marge left hilar tumor inducing a peripheral atelectasia (CT)
T5	Tumor	$1.17 \times 1.17 \times 1.5$	Same as T4 (dosimetric CT scanner)
T6	Mass	$0.77 \times 0.77 \times 1.25$	Mass of the lower left lobe appended to the pleura (CT)
T7	Mass	$0.69 \times 0.69 \times 1.25$	Same as T6, after four months of treatment (CT)
T8	Tumor	$0.63 \times 0.63 \times 1$	Large left hilar tumor and peripheral atelectasia, before treatment (contrast enhanced CT)
T9	Tumor	$0.70 \times 0.70 \times 1$	Same as T8, after chemo-radiotherapy (CE-CT)
T10	Mass	$1.17 \times 1.17 \times 1.5$	Right hilar lymph node mass

Table 3: Characteristics of images containing lung tumors. Resolutions are given in millimeters.

Tumor	Dice Coefficient (%)	Volume Overlap (%)	Volume Difference (%)	Average Surface Distance (mm)	RMS Surface Distance (mm)	Maximum Surface Distance (mm)
T1	90.97	83.45	7.39	0.86	0.92	4.42
T2	80.95	67.99	4.98	1.25	1.54	6.63
T3	72.95	57.42	15.76	1.26	1.50	6.87
T4	71.33	55.44	42.31	3.30	4.01	14.34
T5	80.53	67.41	29.22	3.63	4.55	16.56
T6	86.63	76.42	18.02	1.30	1.49	5.90
T7	82.49	70.21	22.28	1.34	1.56	5.16
T8	89.25	80.59	9.59	1.20	1.47	9.32
T9	72.66	57.07	34.17	1.75	2.09	7.36
T10	74.04	58.79	41.09	4.97	5.55	15.99
Average	80.18	67.47	22.48	2.08	2.46	9.25

Table 4: Comparison between our method and the segmentations provided by the expert.

Tumor	Standard graph cuts		Our method		Amount of object seeds (%)
	Time	Memory	Time	Memory	
T1	4.08	472.34	2.00	24.71	2.52
T2	4.89	573.05	2.71	83.42	2.71
T3	4.90	580.78	2.87	83.42	2.47
T4	5.34	729.72	2.14	37.07	12.99
T5	5.36	737.41	2.10	36.37	10.22
T6	10.18	1476.30	3.80	37.07	3.15
T7	5.16	544.74	3.21	83.42	8.86
T8	MP	43091.25	68.93	83.42	2.45
T9	4.24	496.79	2.39	37.07	8.01
T10	10.36	1151.74	5.42	125.13	9.19

Table 5: Speed (secs) an memory usage (Mb) for our method and the graph cuts without reduction. The label MP means there is not enough memory for allocating the graph.

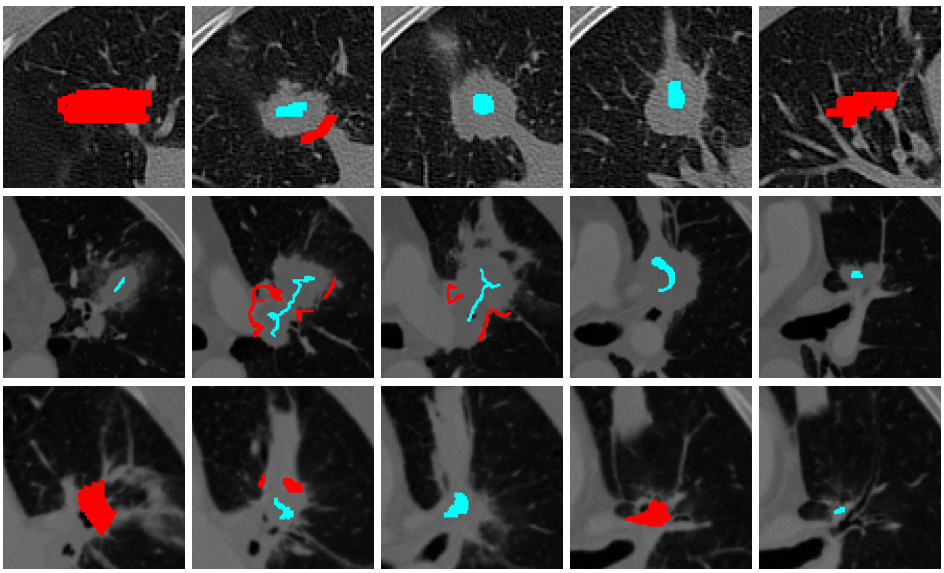


Fig. 6: Seeds location for segmenting lung tumors T1 (top row), T8 (middle row) and T9 (bottom row). Object seeds (light blue) and background seeds (red) are superimposed on successive slices of the original image.

References

1. Y. Boykov and M-P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *ICCV*, volume 1, pages 105–112, 2001.
2. Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on PAMI*, 26(9):1124–1137, 2004.

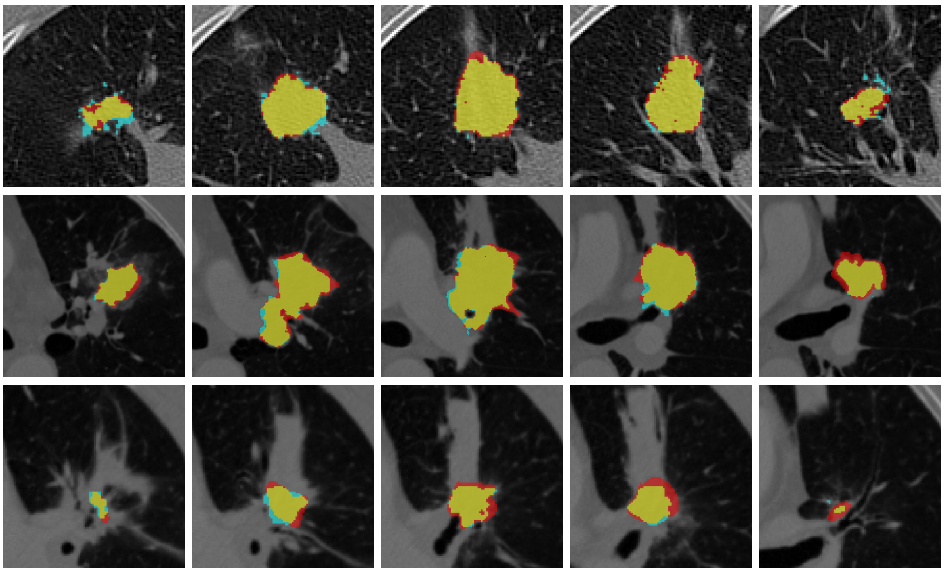


Fig. 7: Segmentation of lung tumors T1 (top row), T8 (middle row) and T9 (bottom row). Ground truth (red) and segmentation (cyan) are superimposed on the original image. Yellow color corresponds to the intersection.

3. A. J. Buckler, J. L. Mulshine, R. Gottlieb, B. Zhao, P. D. Mozley, and L. Schwartz. The use of volumetric CT as an imaging biomarker in lung cancer. *Academic Radiology*, 17(1):100–106, January 2010.
4. P.F. Felzenszwalb and Huttenlocher D. P. Distance transforms of sampled functions. Technical report, Cornell Computing and Information Science, 2004.
5. M. A. Gavrielides, L. M. Kinnard, K. J. Myers, and N. Petrick. Noncalcified lung nodules: Volumetric assessment with thoracic CT. *Radiology*, 251(1):26–37, April 2009.
6. L. Ikonen. Pixel queue algorithm for geodesic distance transforms. In *Discrete Geometry for Computer Imagery*, volume 3429, pages 228–239, April 2005.
7. P. Kohli, V. Lempitsky, and C. Rother. Uncertainty driven multi-scale energy minimization. Technical report, april 2010.
8. V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Transactions on PAMI*, 26(2):147–159, 2004.
9. W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, and C. I. Henschke. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. *IEEE Transactions on Medical Imaging*, 22(10):1259–1274, October 2003.
10. J-M. Kuhnigk, V. Dicken, L. Bornemann, A. Bakai, D. Wormanns, S. Krass, and H-O. Peitgen. Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans. *IEEE Transactions on Medical Imaging*, 25(4):417–434, April 2006.
11. V. Lempitsky and Y. Boykov. Global optimization for shape fitting. In *IEEE, Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007.
12. N. Lermé, F. Malgouyres, and L. Létocart. Reducing graphs in graph cut segmentation. In *To appear in ICIP*, 2010.
13. Y. Li, J. Sun, CK. Tang, and HY. Shum. Lazy Snapping. *ACM Transactions on Graphics*, 23(3):303–308, 2004.

14. H. Lombaert, Y.Y. Sun, L. Grady, and C.Y. Xu. A multilevel banded graph cuts method for fast image segmentation. In *ICCV*, volume 1, pages 259–265, 2005.
15. K. Marten, F. Auer, S. Schmidt, E. J. Rummeny, and C. Engelke. Automated CT volumetry of pulmonary metastases: The effect of a reduced growth threshold and target lesion number on the reliability of therapy response assessment using RECIST criteria. *European Radiology*, 17(10):2561–2571, October 2007.
16. M. F. McNitt-Gray, L. M. Bidaut, S. G. Armato, C. R. Meyer, M. A. Gavrielides, C. Fenimore, G. McLennan, N. Petrick, B. Zhao, A. P. Reeves, R. Beichel, H-J. Kim, and L. Kinnard. Computed tomography assessment of response to therapy: Tumor volume change measurement, truth data, and error. *Translational Oncology*, 2(4):216–222, December 2009.
17. J. H. Moltz, Bornemann L., Kuhnigk J-M., Dicken V., Peitgen E., S. Meier, H. Bolte, M. Fabel, H. C. Bauknecht, M. Hittinger, A. Kiessling, M. Pusken, and H. O. Peitgen. Advanced segmentation techniques for lung nodules, liver metastases, and enlarged lymph nodes in CT scans. *Selected Topics in Signal Processing*, Volume: 3 , Issue: 1:122–134, 2009.
18. L. E. Quint, J. Cheng, M. Schipper, A. C. Chang, and G. Kalemkerian. Lung lesion doubling times: Values and variability based on method of volume determination. *Clinical Radiology*, 63(1):41–48, January 2008.
19. F. Ranchin, A. Chambolle, and F. Dibos. Total variation and graph cuts approaches for variational segmentation. In *Proceedings of SSVM*, pages 743–753, June 2007.
20. A.K. Sinop and L. Grady. Accurate banded graph cut segmentation of thin structures using laplacian pyramids. In *MICCAI*, volume 9, pages 896–903, 2006.
21. I. Sluimer, A. Schilham, M. Prokop, and B. van Ginneken. Computer analysis of computed tomography scans of the lung: a survey. *IEEE Transactions on Medical Imaging*, 25(4):385–405, April 2006.
22. C. Suzuki, M. R. Torkzad, H. Jacobsson, G. Aström, A. Sundin, T. Hatschek, H. Fujii, and L. Blomqvist. Interobserver and intraobserver variability in the response evaluation of cancer therapy according to RECIST and WHO-criteria. *Acta Oncologica*, 49(4):509–514, May 2010.
23. R. J. van Klaveren, M. Oudkerk, M. Prokop, E. T. Scholten, K. Nackaerts, R. Vernhout, C. A. van Iersel, K. A. van den Bergh, S. van 't Westeinde, C. van der Aalst, E. Thunnissen, D. M. Xu, Y. Wang, Zhao Y., H. A. Gietema, B. J. de Hoop, H. J. Groen, G. H. de Bock, P. van Ooijen, C. Weenink, J. Verschakelen, J. W. Lammers, W. Timens, D. Willebrand, A. Vink, W. Mali, and H. J. de Konin. Management of lung nodules detected by volume CT scanning. *The New England Journal of Medicine*, 361(23):2221–2229, December 2009.
24. E. L. van Persijn van Meerten, H. Gelderblom, and J. L. Bloem. RECIST revised: Implications for the radiologist. a review article on the modified RECIST guideline. *European Radiology*, 20(6):1456–1467, June 2010.
25. 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.
26. B. Zhao, L. H. Schwartz, C. S. Moskowitz, M. S. Ginsberg, N. A. Rizvi, and M. G. Kris. Lung cancer: Computerized quantification of tumor response–initial results. *Radiology*, 241(3):892–898, December 2006.

Bronchial region extraction from 3D chest CT image by voxel classification based on local intensity structure

T. Kitasaka[1] ⋆, H. Yano[2], M. Feuerstein[3], and K. Mori[2,4]

[1] School of Information Science, Aichi Institute of Technology,
kitasaka@aitech.ac.jp,
[2] Graduate School of Information Science, Nagoya University, Japan,
[3] Computer Aided Medical Procedures (CAMP), Technische Universitat Munchen, Germany,
[4] Information and Communications Headquarters, Nagoya University, Japan.

Abstract. This paper presents a method for extracting bronchial regions from 3D chest CT images by voxel classification based on local intensity structure. Most of previous methods are based on trace of bronchial tree structures by region growing algorithms, so that it fails when the bronchial lumen is interrupted by abnormals such as tumor. Thus, we focus on detecting candidate voxels which have bronchus-like intensity structure and selecting appropriate candidates from them, instead of tracing the bronchial tree, int bronchial region extraction. Two types of tube enhancement filters are employed in our algorithm. One is designed to enhance bronchus-like intensity structure where low intensity regions are surrounded by higher intensity regions and cross section of a bronchus along its running direction forms the circular shape. The other filter is utilized to enhance line structures based on the Hessian matrix analysis. The experimental results for eight cases of 3D chest CT images showed that the accuracy of the proposed method improved by 10 % compared with a previous method, while FPs increased relatively.

1 Introduction

In diagnosis and treatment of chest diseases, precise recognition of the bronchus is important. Lung cancer and chronic obstructive pulmonary disease (COPD) are representative diseases in the chest. These diseases strongly correlate or affect bronchial morphology. In clinical practice, 3D chest X-ray CT images are used for inspection for such diseases, since the bronchus is clearly depicted in

⋆ The authors thank our colleagues for suggestions and advices. Parts of this research were supported by the Grant-In-Aid for Scientific Research from the Ministry of Education (MEXT), Japan Society for the Promotion of Science (JSPS), the Japan Society for Promotion of Science, and the Grant-In-Aid for Cancer Research from the Ministry of Health and Welfare. The authors also thank to Dr. Hiroshi Natori, Dr. Masaki Mori, Dr. Hirotsugu Takabatake, and Dr. Yoshinori Hasegawa for their medical advices.

comparison with any other medical imaging devices. Therefore, a computer aided diagnosis or treatment system of the chest should recognize bronchial structure before measuring precise features of suspicious regions from chest CT images. Many studies on bronchus recognition for automated detection and analysis of suspicious regions observed in the bronchus or lung have been reported [1–11].

Mori et al. [1], Kitasaka et al. [2], Feuerstein et al. [3], Tschirren et al. [4], Schlathoelter et al. [5], Lo et al. [6] uses region growing or front propagation methods to segment and restructure a bronchial region. There are two main issues of the methods based on region growing algorithms; (a) leakage issue and (b) interruption of growing. First, leakages beyond the bronchial wall occur at the areas where intensity contrast between the bronchial wall and bronchial lumen becomes low, e.g. due to the partial volume effect. Second, when there is interruption such as a tumor or intensity gap caused by heart beat, bronchial regions beyond them cannot be extracted at all. Contrastly, the methods based on classification that distinguishes voxels in the bronchus from those in others using several bronchial features [8–11]. These methods use typical features of the bronchus, such as (a) low intensity at the bronchial lumen, (b) relatively higher intensity at the bronchial wall than that of bronchial lumen, (c) the bronchus forms tube structure, and (d) the radious of a bronchial branch do not change rapidly along its running direction. Lo et al. [11] introduced a method for extracting bronchial regions from CT volume based on voxel classification approach. In this method voxels of an input CT volume are classified into airway voxels or others based on an airway appearance model. Then, false positives (FPs) are removed by using the measure of vessel orientation similarity. We also propose a novel method based on bronchial feature classification by developing features of bronchial tube structure in order to cope with the second issue described above. Our method focuses on development of new bronchial feature as a preceding step of utilization of a statistical bronchial atlas such as an airway appearance model Lo et al. used.

We explain our features and processing procedures in 2, conduct the experiment in 3 and add discussion in 4.

2 Method

2.1 Overview

We utilize the following image features of bronchi that can be seen on CT images: (1) low intensity in luminal area, (2) relatively higher intensity in the bronchial wall region than that in luminal area, and (3) tube-like structure. Thus, we extract bronchial regions by classifying each voxel based on above features. In the preprocessing step, we apply a LoG filter to an input CT volume to enhance bronchial wall edges. Then, two types of tube enhancement filters are employed to detect voxels showing bronchus-like structures. The first filter enhances voxels forming line structures based on the Hessian analysis [12]. The second filter enhances voxels having bronchus-like intensity structure that (a) low intensity lumen is surrounded by higher intensity wall and (b) a cross section of a bronchus

along its running direction forms the circular shape. Then, we detect candidate voxels of bronchus regions based on outputs of above two filters. Since the above procedure extracts only voxels around center of bronchus regions, bronchus regions are extracted by the postprocessing step that recovers bronchus regions.

2.2 Preprocessing

We apply a LoG filter of standard deviation σ to sharpen an input volume. Let $\mathbf{F}$ and $\mathbf{G}$ be the input and output images. A LoG filter with σ is represented by LoG_σ. Our sharpening operation can be expressed as

$$\mathbf{G} = \mathbf{F} - wLoG_\sigma(\mathbf{F}), \tag{1}$$

where w is a weight to control the effect of sharpening operation. Bronchial walls, especially bronchial walls of thinner bronchial branches, are enhanced with suppressing noise.

2.3 Line enhancement filter

We enhance luminal region of bronchi by performing the Hessian matrix analysis method proposed in [12]. From eigenvalues of the Hessian matrix computed for each voxel, λ_1, λ_2 and λ_3 ($\lambda_1 \leq \lambda_2 \leq \lambda_3$), a lineness measure $S^{line}_{\sigma_h}(\lambda_1, \lambda_2, \lambda_3)$ for scale σ_h is calculated by

$$S^{line}_{\sigma_h}(\lambda_1, \lambda_2, \lambda_3) = \begin{cases} |\lambda_3| \cdot \psi(\lambda_2; \lambda_3) \cdot \omega(\lambda_1; \lambda_2) & (\lambda_3 \geq \lambda_2 > 0) \\ 0 & otherwise \end{cases} \tag{2}$$

where

$$\psi(\lambda_s, \lambda_t) = \begin{cases} (\frac{\lambda_s}{\lambda_t})^\gamma & (\lambda_t \geq \lambda_s > 0) \\ 0 & otherwise \end{cases} \tag{3}$$

and

$$\omega(\lambda_s, \lambda_t) = \begin{cases} (1 + \frac{\lambda_s}{|\lambda_t|})^\gamma & (\lambda_t \geq \lambda_s > 0) \\ (1 - \alpha\frac{\lambda_s}{|\lambda_t|})^\gamma & (-\frac{\lambda_t}{\alpha} < \lambda_s < 0). \\ 0 & otherwise \end{cases} \tag{4}$$

$S^{line}_{\sigma_h}$ is calculated by changing scale σ_h. Output of each voxel $\mathbf{p}$ by this filter, $l_h(\mathbf{p})$, is the largest response at certain σ_h. Figure 1 shows examples of outputs of the line enhancement filter with and without the preprocessing. Thinner bronchial vessels are well enhanced by performing the preprocessing as shown in the left figure. However, other structures are enhanced as well. We extract bronchial candidate voxels satisfying $l_h(\mathbf{p}) > T_l$. For an extracted voxel $\mathbf{p}$, we calculate the running direction $\mathbf{c}$ and radius R of the bronchial branch to which the voxel $\mathbf{p}$ belongs as the eigen vector $\mathbf{e}_1$ corresponding to λ_1 and $4\sigma_h$, respectively.

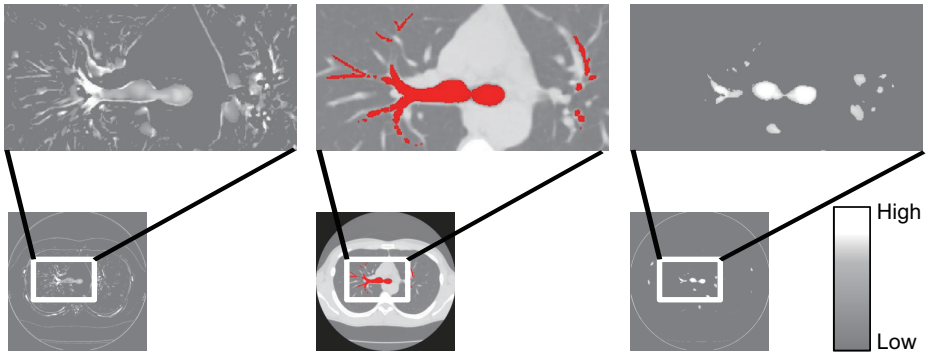


Fig. 1. Examples of line enhancement filter w and w/o LoG filter. (left) Result with LoG filter, (middle) bronchus region manually inputted in red overlaid on a CT slice and (right) result without LoG filter. Thinner bronchial vessels are well enhanced in the left figure compared with the right figure, while other structures are enhanced as well.

2.4 Tube extraction filter

To extract voxels having bronchus-like intensity structure (low intensity areas are surrounded by higher intensity areas), we employ a radial reach filter (RRF) [13]. RRF scans intensity profiles radially from a voxel of interest on a cross section. If a voxel currently being checked does not satisfy given conditions in the scan for a certain direction, we visit neighboring voxel along the direction. This process is iterated until we encounter a voxel satisfying the conditions. If the voxels found in all scans form a pre-defined shape (in our case, circular shape), the pixel of interest is extracted.

We apply the RRF by changing the normal of a cross section for each candidate voxel extracted in 2.3. If a voxel $\mathbf{p}$ does not satisfy any of following four conditions,

Condition (1) $f(\mathbf{p}) < T_1$,
Condition (2) $f(\mathbf{s_k}) - f(\mathbf{p}) > T_2 \quad (k = 1, 2, \cdots, D)$,
Condition (3) More than T_D% of all scans satisfies the condition (2),
Condition (4) $V[l_k] < T_V$,

the voxel $\mathbf{p}$ is removed from candidate voxel set. In these conditions $f(\mathbf{p})$ is the intensity at a voxel $\mathbf{p}$. $\mathbf{s_k}$ is the voxel satisfying the Condition (2) along a scan direction $\mathbf{d_k}$. The distance between $\mathbf{s_k}$ and $\mathbf{p}$ is represented by l_k. $\mathbf{s_k}$ is then represented as $\mathbf{s_k} = \mathbf{p} + l_k\mathbf{d_k}$, where $\mathbf{d_k} = (\cos\psi, \sin\psi)^T$, $\psi = 2\pi(k-1)/D$, $l_k < L_{max}$, and D represents the number of scans. $V[x]$ denotes the variance of x. The Condition (1) checks whether $f(\mathbf{p})$ is sufficiently low as the bronchial lumen. The Conditions (2) and (3) decide whether $\mathbf{p}$ is surrounded by higher intensity. Final Condition (4) checks whether higher intensity voxels surrounding $\mathbf{p}$ forms the circular shape. The normal vector $\mathbf{n}$ of a cross section whose $V[l_k]$ is minimum is estimated as the running direction of a bronchus at $\mathbf{p}$.

2.5 Postprocessing

The candidate voxels obtained so far include many false positives. Thus, we eliminate FPs as follows. First, we check whether two estimations of the running direction of a bronchus, $\mathbf{c}$ and $\mathbf{n}$, are consistent by

$$\frac{|\mathbf{n} \cdot \mathbf{c}|}{\|\mathbf{n}\|\|\mathbf{c}\|} \geq T_a. \tag{5}$$

If a candidate voxel does not hold above equation, the voxel is removed from candidate voxel set. Then, a connected component whose number of voxels is less than T_c is removed as well. Since the bronchial candidates are detected at only central parts of bronchial tube structure, we recover the thickness of each tube based on estimated radius R. However, branching parts of thick tubes cannot be recovered well. Such parts may not be detected as bronchial tube structure because their shape on cross sections do not show circular one. Thus, by unifying the roughly extracted region by the simple region growing method [1] which extracts relatively thick branches up to subsegmental lobes at least and the recovered region, the bronchial region is finally extracted.

3 Experiments

We applied the proposed method to eight cases of 3D chest CT images. Specifications of the CT images are; 512 $\times$ 512 pixels in a slice, 209 - 751 slices in an image, 0.506 - 0.684 $[mm]$ in pixel spacing, and 0.401-1.25 $[mm]$ in slice spacing, respectively. We evaluated the extraction results by comparing with ground truth data which are generated manually by an author. The following three indices were used for evaluation: (a) Branch Detectability (BD) : percentage of the number of extracted branches out of bronchial branches manually detected, (b) True Positives (TP) : Number of extracted voxels which were true bronchial voxels, and (c) False Positives (FP) : Number of extracted voxels which were not bronchial voxels actually.

Parameters used in the experiment are shown in Table 1. They were set experimentally by searching adequate values one by one. Here, parameters for the tube extraction filter, T_2, L_{max}, T_V and T_D, were gradually changed according to the estimated radius of a branch. Figures 2-4 and Table 2 show examples of the extracted bronchial regions and evaluated results by BD and FP in comparison with the previous (adaptive branch tracing) method [3]. In Figs. 2-4, ground truth data, TP and FP regions are shown in red, yellow and blue, respectively. As shown in Table 2, the number of extracted branches and BD by the proposed method are better than those by the adaptive branch tracing method, in paticular, for Cases 7 and 8 which have a tumor inside the bronchial lumen. On the other hand, the amount of FPs of the proposed method increased relatively than that of the adaptive branch tracing method, though it reduced significantly in cases 1 and 3.

Table 1. Parameter setting used in the experiment.

σ	0.5	w	0.5 ($LoG(\mathbf{p}) \leq 0$) 0.05 ($LoG(\mathbf{p}) > 0$)	T_l	250	T_a	0.9	T_c	500
T_1	-700 [HU]	N	8	D	72				
$R > 10[mm]$			400 [HU]		R [mm]		2.0		100[%]
$R > 5[mm]$		T_2	250 [HU]	L_{max}	R [mm]	T_V	1.0	T_D	100[%]
$R > 3[mm]$			100 [HU]		5 [mm]		1.0		90[%]
otherwise			50 [HU]		5 [mm]		1.0		90[%]

Table 2. Comparison of results by the proposed and adaptive branch tracing (ABT) [3] methods. The term "BD" indicates the branch detectability.

Case	Number of branches	Number of extracted branches	BD of proposed method [%]	BD of ABT method [%] (Number of extracted branches)	FP of proposed method [mm^3]	FP of ABT method [mm^3]
1	385	260	67.5	69.9 (269)	1093.4	4883.2
2	331	229	69.2	63.7 (211)	2274.0	1575.7
3	422	323	76.5	55.7 (235)	3569.0	6609.4
4	490	409	83.5	75.5 (370)	16477.9	9101.3
5	236	142	60.2	46.6 (110)	402.3	230.9
6	511	322	63.0	65.6 (335)	1089.2	1615.5
7	171	126	73.7	23.4 (40)	1678.2	650.1
8	548	322	58.8	40.9 (224)	2203.5	629.7
ave	386.8	266.6	68.9	58.0(224.3)	3598.4	3162.0

4 Discussion

As shown in Figs. 2 and 3, the proposed method can extract bronchial regions beyond the interruption by a tumor. This is because our method detects candidate voxels which have bronchus-like intensity structure and select appropriate candidates from them, instead of tracing the bronchial tree as in previous methods [1, 3]. However, there are some FPs caused by other structure such as the esophagus and areas under furcation parts of blood vessels, because these FPs show partially similar intensity structure with the bronchial region. In particular for areas under furcation parts of blood vessels, since their intensity is slightly lower than lung parenchyma they seem to be surrounded by higher intensity objects. As the side effect of the sharpning operation, this intensity structure is emphasized as well. In Case 4, strong influence of the side effect causes large FPs. To reduce FPs more, we should take tree structural constraints into account in addition to our bronchus-like intensity structure measures. Utilization of a statistical bronchial altas will give good constraints on our method. Adoption of algorithms based on global optimization such as a graph cut algorithm may also

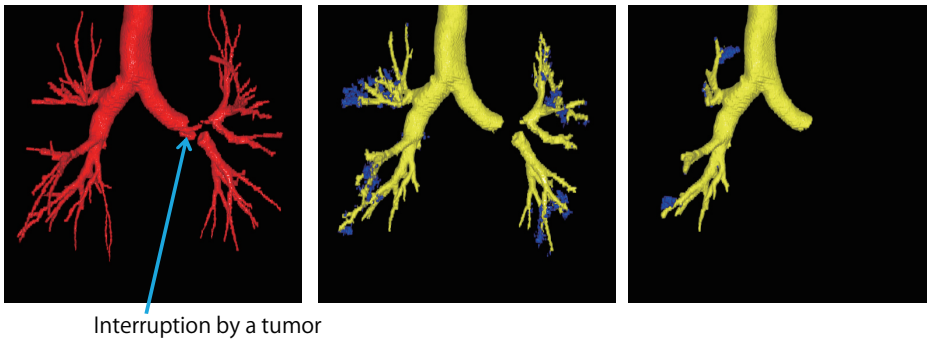


Fig. 2. Volume rendered results of the ground truth (left), the proposed method (middle) and the adaptive branch tracing method [3] (right) for Case 7. Yellow region indicates TP while blue region shows FP. Though the bronchial lumen is interrupted by a tumor indicated by arrow, the proposed method can extract the lumen regions beyond the tumor.

improve the accuracy of the proposed method. Furthermore, extraction parameter tuning, e.g. by machine learning approach, is one of future works since it also affects the accuracy of the proposed method.

5 Conclusions

This paper presented a method for extracting bronchial regions from 3D chest CT images by voxel classification based on local intensity structure. The proposed method detected candidate voxels which have bronchus-like intensity structure and selected appropriate candidates from them, instead of tracing the bronchial tree. The experimental results for eight cases of 3D chest CT images showed that the accuracy of the proposed method improved by 10 % compared with the adaptive branch tracing method, while FPs increased relatively. We also confirmed that the proposed method could extract bronchial regions beyond the interruption by a tumor.

References

1. K. Mori, J. Hasegawa, J. Toriwaki, H. Anno and K. Katada, "Automated Extraction and Visualization of Bronchus from 3D CT Images of Lung," Proc of 1st CVRMIed'95, pp. 542-548, 1995.
2. T. Kitasaka, K. Mori, J. Hasegawa and J. Toriwaki, " A Method for Extraction of Bronchus Regions from 3D Chest X-ray Images by Analyzing Structural Features of the Bronchus ", Forma, vol.17, pp.321-338, 2002.
3. M. Feuerstein, T. Kitasaka and K. Mori, "Adaptive Branch Tracing and Image Sharpening for Airway Tree Extraction in 3-D Chest CT," Proc. of Second International Workshop on Pulmonary Image Analysis, ISBN: 978-1448680894, pp. 273-284, 2009.

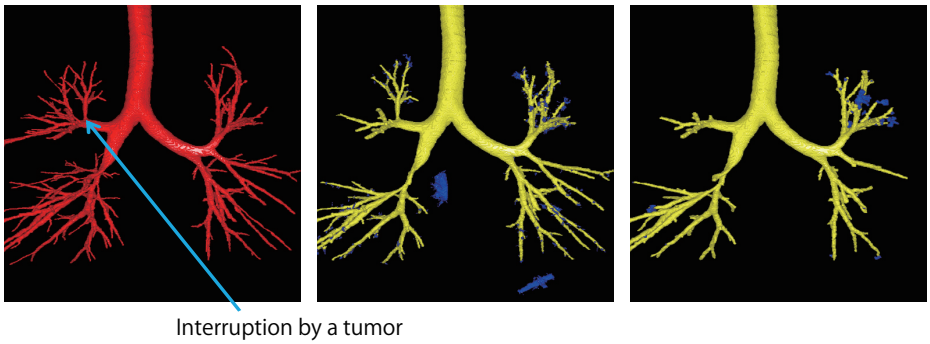


Fig. 3. Volume rendered results of the ground truth (left), the proposed method (middle) and the adaptive branch tracing method [3] (right) for Case 8. Yellow region indicates TP while blue region shows FP. Though the bronchial lumen is also interrupted by a tumor indicated by arrow, the proposed method can extract the lumen regions beyond the tumor.

4. J. Tschirren, E. A. Hoffman, G. McLennan and M. Sonka, "Intrathoracic airway trees: segmentation and airway morphology analysis from low-dose CT scans", IEEE Transactions on Medical Imaging, vol.24, no.12, pp.1529-1539, 2005.
5. T. Schlathoelter, C. Lorenz, I. C. Carlsena, S. Renischa and T. Deschamps, '' Simultaneous segmentation and tree reconstruction of the airways for virtual bronchoscopy", Proc. of SPIE on Medical Imaging, vol.4684, pp.103-113, 2002.
6. P. Lo and M. de Bruijnea, "Voxel classification based airway tree segmentation", Proc. of SPIE on Medical Imaging, vol.6914, pp.69141K-1-12, 2008.
7. D. Aykac, E. A. Hoffman, Member, G. McLennan and J. M. Reinhardt, S. Member, "Segmentation and Analysis of the Human Airway Tree From Three-Dimensional X-Ray CT Images", IEEE Transactions on Medical Imaging, vol.22, no.8, pp.940-950, 2003.
8. A. Szymczak, A. Stillman and A. Tannenbaum, "3D Models of Airway Trees from CT Scans: A Point-Based Approach", http://www.mines.edu/aszymcza/
9. C. Bauer, H. Bischof and R. Beichel, "Segmentation of Airways Based on Gradient Vector Flow", Proc. of Second International Workshop on Pulmonary Image Analysis, pp.192-201, 2009.
10. C. Bauer, T. Pock, H. Bischof and R. Beichel, "Airway Tree Reconstruction Based on Tube Detection", Proc. of Second International Workshop on Pulmonary Image Analysis, pp.203-213, 2009.
11. P. Lo, J. Sporring, H. Ashraf, J. JH. Pedersen and M. de Bruijne, "Vessel-guided airway tree segmentation: A voxel classification approach, " Medical Image Analysis, vol. 14, pp. 527-538, 2010.
12. Y. Sato, C. F. Westin, A. Bhalerao, S. Nakajima, N. Shiraga, S. Tamura and R. Kikinis, "Tissue classification based on 3D local intensity structures for volume rendering, " IEEE Transactions on Visualization and Computer Graphics, vol.6, no.2, pp. 160-180, 2000.
13. Y. Sato, S. Kaneko, Y. Niwa and K. Yamamoto, "Robust object detection using a radial reach filter (RRF), " Systems and Computers in Japan, vol. 35, no. 10, pp. 63-73, 2004.

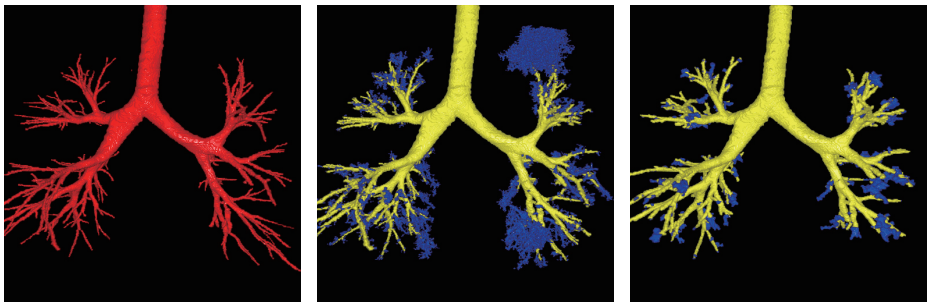


Fig. 4. Volume rendered results of the ground truth (left), the proposed method (middle) and the adaptive branch tracing method [3] (right) for Case 4. Yellow region indicates TP while blue region shows FP. The proposed method produced larger volume of FP than the adaptive branch tracing method, although BD becames much better.

Model-Based Segmentation of Pathological Lungs in Volumetric CT Data

Shanhui Sun[1,5], Geoffrey McLennan[2,3,4,5], Eric A. Hoffman[3,2,4,5], and Reinhard Beichel[1,2,5]

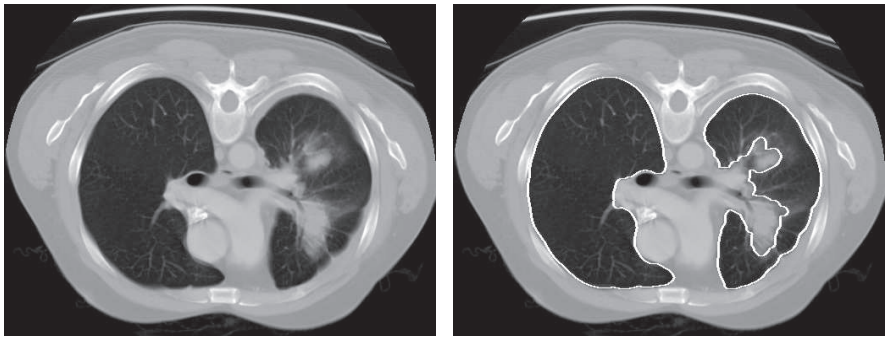
[1] Dept. of Electrical and Computer Engineering
[2] Dept. of Internal Medicine
[3] Dept. of Radiology
[4] Dept. of Biomedical Engineering
[5] The Iowa Institute for Biomedical Imaging
The University of Iowa, Iowa City, IA 52242, USA
reinhard-beichel@uiowa.edu

Abstract. Segmentation of pathological lungs is a non-trivial problem. We present a new approach for the segmentation of lungs with high-density pathologies like lung cancer. Our method consists of two main processing steps. First, a novel robust active shape model matching method is applied to roughly segment the outline of the lungs. Second, an optimal surface finding approach is utilized to refine the initial segmentation result. Left and right lungs are segmented separately. A comparison to an independent reference on eleven abnormal (lung cancer) and nine normal test cases resulted in an average Dice coefficient of 0.9741 and 0.9758, respectively. Our algorithm was specifically designed for general-purpose computation on graphics processing units (GPGPU) and requires on average 116 seconds for segmenting a left or right lung.

1 Introduction

Many computer-aided lung image analysis methods require the segmentation of lung tissue in an initial processing step. In the case of normal lungs imaged with X-ray computed Tomography (CT), lung segmentation is a rather simple task because of the large density difference between air-filled lung tissue and surrounding tissues. Many algorithms can be found in the literature that deal with the segmentation of normal lungs [1–5]. In the case of lungs with high density pathology (e.g., cancer, pneumonia, etc.) as shown in Fig. 1(a), segmentation becomes a non-trivial problem and conventional algorithms fail to deliver correct segmentation results (Fig. 1(b)). Pathological cases are frequently occurring in clinical practice and constitute the crucial cases for computer-aided diagnosis and treatment planning/monitoring.

There is scant literature about robust lung segmentation methods. Sluimer et al. propose a segmentation by registration scheme in which scans of normal lungs are elastically registered to a scan with pathology [6]. While delivering promising results, not all pathological cases could be handled successfully [6].



(a) (b)

Fig. 1. Lung segmentation. (a) Axial CT image showing normal right and pathological (cancer) left lung tissue and (b) corresponding segmentation result generated with a standard lung segmentation method.

In addition, the authors also identified the need to reduce the time needed by the registration algorithm from three hours to a clinically more acceptable processing time [6]. To address this problem, a hybrid lung segmentation method was presented in a recent publication [7]. The idea is to first use a fast conventional lung segmentation algorithm followed by a segmentation error detection step. If errors are detected, the more complex algorithm is utilized. An adaptive border marching algorithm was presented in [8] to include juxtapleural nodules in lung segmentations. Larger areas of under-segmentation were reported in hilar and pulmonary consolidation regions [8]. A combination of a Bézier surface model for the side walls of lungs and a conventional lung segmentation technique was proposed in [9] to deal with lesions located at the lung border. An approach for the robust segmentation of lung parenchyma based on the curvature of ribs was presented in [10]. The method is based on an adaptive thresholding scheme and utilizes a comparison of the curvature of the lung boundary to the curvature of the ribs to select thresholds. Because lung pathologies like cancer can have density values similar to other tissues surrounding the lung, the method will likely produce errors in such cases.

In this paper we present a new approach for the segmentation of lungs with high-density lung disease (e.g., cancer, fibrosis, pneumonias, etc.) that addresses limitations of existing methods like robustness or processing speed. Our approach is based on a lung model and utilizes a novel robust model matching method. To be suitable for routine application, the time required for lung segmentation must be low. The model-based 3D segmentation of lungs is particular challenging, because of the size of lungs and the amount of image data to be processed. Our approach to robust lung segmentation addresses this issue—the algorithm is specifically designed to take advantage of general-purpose computation on graphics processing hardware, which reduces the execution time considerably.

2 Methods

In the following Sections, we describe the individual processing steps of our robust lung segmentation approach.

2.1 Lung Model Generation

To represent lung shapes, a model describing the mean shape and variation around the mean was generated. For this purpose, $n = 41$ segmented normal lung CT data sets imaged at total lung capacity (TLC) were used, and a set of corresponding landmarks $\{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_m\}$ with $m = 2562$ were identified on all lung shapes by means of minimum description length (MDL) approach [11]. All n landmark sets were aligned into a common coordinate frame by using procrustes analysis, resulting in a mean shape vector $\bar{\mathbf{x}}$. For each learning shape, a shape vector $\mathbf{x}_i$ with $i = 1, 2, \ldots, m$ was generated by concatenating the coordinates: $\mathbf{x}_i = [x_{i,1}, y_{i,1}, z_{i,1}, x_{i,2}, y_{i,2}, z_{i,2}, \ldots, x_{i,m}, y_{i,m}, z_{i,m}]^T$. A principal component analysis (PCA) was applied to generate a point distribution model (PDM) [12]. The above outlined process was applied to left and right lungs, yielding independent left and right lung PDMs. An instance of a left or right lung shape can be generated from the corresponding PDM by the linear model: $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b}$, where $\mathbf{P}$ denotes the shape eigenvector matrix and $\mathbf{b}$ represents the shape coefficients.

2.2 Robust Active Shape Model Matching

a) Standard Matching Scheme The PDM described in the previous Section can be used for lung segmentation by matching the model to the target structure. This could be accomplished by utilizing a standard active shape model (ASM) matching framework [12]. First, an instance of the shape model (e.g., mean shape) is generated and placed in proximity to the target structure in the image volume. Second, to match the model, shape points are updated by searching from the current landmark position along a profile of length l_{ASM} perpendicular to the model surface. To identify suitable update points $\mathbf{y}$, we use the following cost function $c_i = g_{mag_i}$, and update point locations with $\boldsymbol{n}_i \cdot \boldsymbol{g}_{dir_i} < 0$ are ignored. c_i represents the cost of the i-th column element, and the associated sampled gradient magnitude, gradient direction, and normal vector are denoted as g_{mag_i}, $\boldsymbol{g}_{dir_i}$, and $\boldsymbol{n}_i$, respectively. The gradient calculation is based on Gaussian derivatives with a standard deviation of σ_{ASM}, and the calculation of $\boldsymbol{g}_{dir}$ and g_{mag} is done for each voxel of the volume before the model matching is started. These pre-calculated gradient values are then used to interpolate gradient vectors during model matching. If a gradient value outside the volume is required during model matching, a value closest to the boundary is utilized. In case that no new update point can be found, the old position is used.

Once all shape points are updated, pose parameters are adjusted to map the updated shape points to the mean model. For this purpose, a procrustes alignment step is used to estimate transformation matrix $\mathbf{T}$, which consists of

scaling, rotation, and translation parameters, by minimizing $(\mathbf{T}[\mathbf{y}] - \bar{\mathbf{x}})^T(\mathbf{T}[\mathbf{y}] - \bar{\mathbf{x}})$. Model parameters $\mathbf{b}$ are updated by using

$$\mathbf{b} = \mathbf{P}^T(\mathbf{T}[\mathbf{y}] - \bar{\mathbf{x}}) \; . \tag{1}$$

The outlined process is repeated until the model converges.

We are especially interested in the segmentation of pathological lungs that contain areas of lung disease with high density (e.g., pneumonia, cancer, etc.). Thus, it is very likely that some update points are found during the model matching procedure that do not represent lung surface and belong to an area of transition between normal and diseased lung tissue. Consequently, the standard matching approach will fail, since it is a least squares optimization procedure that is not suitable to handle outliers. Therefore, a robust shape model matching approach is needed.

b) Robust Matching Approach The basic idea behind robust ASM matching is to only use inlier components of $\mathbf{y}$ to update model parameters. In this context, Rogers et al. investigated M-estimators and random sampling-based robust parameter estimation techniques for 2D ASM matching [13]. It is well known that the effectiveness of M-estimators strongly depends on the selection of the weighting function and its parameters. Usually, this selection is not trivial, and the optimal selection might change from case to case. Random sampling techniques try to find a subset of inliers by evaluating a number of randomly sampled subsets of update points. Such approaches work well, if the required subset of inliers is quite small. In case of large ASM models, this strategy can lead to suboptimal results, because a small set of inliers might not be representative enough to describe a complex lung shape (many landmark points), and thus can negatively impact the matching result. For our application, it is desirable to use as many inliers as possible for the model update.

In the following, we present a new robust ASM matching algorithm that extends the standard matching approach (Section 2.2.a) by an outlier detection step. For this purpose, we use a robust PCA coefficient estimation scheme that builds on the work of Storer et al. [14]. The method presented in [14] was designed for robust image reconstruction and targets a pre-defined number of inliers. In this paper, we propose a novel voting scheme that does not require to specify a targeted number of inliers. Our method consists of two processing steps. First, normal shape patterns of landmark subsets are learned. Second, these patterns are then utilized during ASM matching to identify and reject outliers.

i) Offline learning Corresponding landmark points of all learning shapes are partitioned randomly into k shape subsets of approximately equal size. This process is repeated l-times, resulting in a set of subsets: $\Omega = \{\omega_{i,j} | i \in 1, 2, \ldots, l, \;\; j \in 1, 2, \ldots, k\}$. Note that corresponding landmark points of all n learning data sets are always assigned to the same subset. Consequently, each subset $\omega_{i,j}$ consists of n subset samples. For each subset $\omega_{i,j}$, a mean shape $\bar{\mathbf{x}}_{\omega_{i,j}}$ is calculated by using procrustes analysis, and all shapes of the subset are aligned. The subset shapes are then converted to shape vectors by concatenating

their x-, y-, and z-components. By means of PCA, the corresponding eigenvectors $\mathbf{P}_{\omega_{i,j}}$ are calculated. $\bar{\mathbf{x}}_{\omega_{i,j}}$ and $\mathbf{P}_{\omega_{i,j}}$ are stored and will be utilized for robust ASM matching.

ii) Robust matching To robustly match the model to image data, we must identify outliers in the update point vector $\mathbf{y}$ in each iteration. This is accomplished by analyzing subset combinations of $\mathbf{y}$ and by utilizing a voting scheme. Let $\mathbf{y}_{\omega_{i,j}}$ represent the components of $\mathbf{y}$ that are corresponding to the landmark points that constitute subset $\omega_{i,j}$. For each subset, a reconstruction error $e_{\omega_{i,j}} = \left\| \mathbf{T}_{\omega_{i,j}}[\mathbf{y}_{\omega_{i,j}}] - \left[\bar{\mathbf{x}}_{\omega_{i,j}} + \mathbf{P}_{\omega_{i,j}} \tilde{\mathbf{b}}_{\omega_{i,j}} \right] \right\|$ is calculated, where $\mathbf{T}_{\omega_{i,j}}$ is a transformation matrix that aligns $\mathbf{y}_{\omega_{i,j}}$ to the corresponding mean $\bar{\mathbf{x}}_{\omega_{i,j}}$. The vector $\tilde{\mathbf{b}}_{\omega_{i,j}}$ is derived from $\mathbf{b}_{\omega_{i,j}} = \mathbf{P}^T_{\omega_{i,j}} \left[\mathbf{T}_{\omega_{i,j}}[\mathbf{y}_{\omega_{i,j}}] - \bar{\mathbf{x}}_{\omega_{i,j}} \right]$ by limiting the components of the parameter vector $\mathbf{b}_{\omega_{i,j}}$ to a value of $\pm\xi$-times the square-root of the corresponding eigenvalue.

A large reconstruction error $e_{\omega_{i,j}}$ is an indication that subset $\omega_{i,j}$ is very likely contaminated by one or more outliers. To identify the outliers, the reconstruction error $e_{\omega_{i,j}}$ is interpreted as a vote, which is casted for all update points that are included in the subset $\mathbf{y}_{\omega_{i,j}}$. This voting process is carried out for all subsets $\omega_{i,j} \in \Omega$. The casted votes are collected in a matrix $\mathbf{V}_{err}$ of size $m \times l$, in which rows correspond to shape points in $\mathbf{y}$. After all votes are casted, $\mathbf{V}_{err}$ is analyzed to detect outliers. First, to increase robustness, a rank order statistics filter is applied to each row; the values are sorted, and the g-lowest value is selected to represent the filter result. The filtering step reduces $\mathbf{V}_{err}$ to a vector $\mathbf{v}_{err} = [v_1, v_2, \ldots, v_m]^T$ and helps to reject accidental occurring point constellations that contain outliers, which are similar to constellations of inlier points. Second, a threshold δ is derived from $\mathbf{v}_{err}$ by analyzing the distribution of vector components v_i: $\delta = \mu + \beta\sigma$ with $\mu = median_{i \in 1,2,\ldots,m}\{v_i\}$ and $\sigma = \sqrt{1/m \sum_{i=1}^{m} (v_i - \mu)^2}$, where β represents a constant. If needed, this step can be replaced by a more advanced mean shift-based analysis step, similar to the approach reported in [15]. Third, the threshold is applied to $\mathbf{v}_{err}$ to yield a selection vector: $\mathbf{p}_{sel} = [p_1, p_2, \ldots, p_m]^T$ with

$$p_i = \begin{cases} 1 : v_i < \delta \\ 0 : v_i \geq \delta \end{cases} \tag{2}$$

to discriminate between inliers ($p_i = 1$) and outliers ($p_i = 0$). Once the inliers are identified, an update of the model parameter vector is calculated by utilizing a modified version of Eq. 1 which takes $\mathbf{p}_{sel}$ into account. An advantage of our outlier detection algorithm is that it is well suited for parallel processing. For example, we utilize a general-purpose computing on graphics processing units (GPGPU) implementation.

For our application, we used the following parameters: $l_{ASM} = \pm 40$ mm, $k = 200$, $l = 60$, $\xi = 2$, $g = 10$, and $\beta = 1.3$. To update the robust ASM, a gradient image was calculated based on Gaussian derivatives with a standard deviation of $\sigma_{ASM} = 4$. The maximum gradient position along the search profile was used to calculate updates for shape points. The maximum number of iterations during

model matching was set to 100, and model matching was stopped if the average of the shape point movement was below 0.04 mm.

2.3 Model Constrained Optimal Surface Finding

Depending on the training data utilized for model building, the model might not be able do describe smaller local shape variations. To capture this information, we generate the final lung segmentation by applying a global optimal surface finding method [16]. The algorithm transforms the segmentation problem into a graph optimization problem, which is solved by means of a maximum-flow algorithm [16]. Thus, an edge-weighted directed graph is built, and weights derived from the volume are assigned to the graph edges to reflect local image properties. For this purpose, the final ASM mesh is utilized. Since the ASM vertices are sparse, the mesh is restructured by adding triangles, before the graph is built. For graph generation, columns along the surface normal of each vertex (search profile) are generated. The length l_p of the profile is utilized to constrain the segmentation to the proximity of the initial ASM segmentation. In addition, a surface smoothness constraint Δ is incorporated into the graph as described in [16]. For segmentation we use the following cost function:

$$c_i = \begin{cases} g_{max} & if\ \boldsymbol{n}_i \cdot \boldsymbol{g}_{dir\,i} < 0 \\ g_{max} - g_{mag_i} & otherwise \end{cases}, \tag{3}$$

where c_i represents the cost of the i-th column element and g_{max} the maximum gradient magnitude of the volume. The gradient calculation is based on Gaussian derivatives with a standard deviation of σ_g. We utilize the optimal surface finding in an iterative fashion, and the used sequence of values for σ_g and Δ was $\{6.0, 3.0, 1.0, 0.5\}$ and $\{10, 8, 5, 2\}$, respectively. For the search profile, $l_p = \pm 10$ voxel was selected.

3 Case Studies

To evaluate the performance of our method, 20 segmentations were performed on 10 multidetector computed tomography (MDCT) scans (Fig. 2). For each data set, left and right lungs were segmented, and either the left and/or right lung contained diseased lung tissue with a significant higher density compared to normal tissue. The image size varied from $512 \times 512 \times 458$ to $512 \times 512 \times 572$ voxel. The in-plane resolution of the images ranged from 0.580×0.580 to 0.809×0.809 mm and the slice thickness from 0.6 to 0.7 mm. To generate an independent reference standard, a reference segmentation of left and right lungs was generated by an expert for all 10 scans. This process took several hours per case.

The initialization of the model was done manually by placing the corresponding left or right mean shape model with a fixed size in the individual data sets. For each method investigated in this paper, the same starting position and initial shape was used.

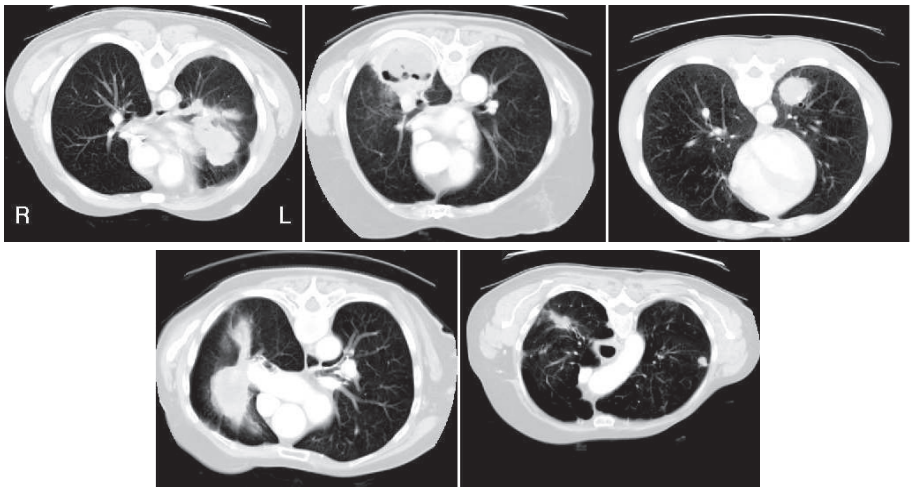
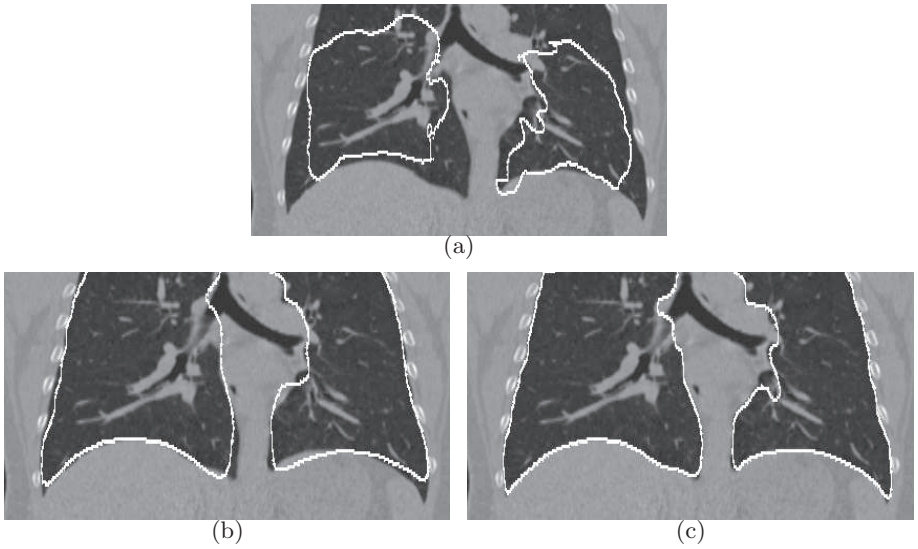


Fig. 2. Examples of lung data sets with high-density pathology that were utilized for evaluation of our approach. From left to right and top to bottom: hdl01, hdl03, hdl04, hdl07, and hdl10.



(a) (b) (c)

Fig. 3. Segmentation of an incomplete lung CT data set; the top portion was not scanned. (a) Standard ASM. (b) Robust ASM. (c) Combination of robust ASM and optimal surface finding. Note that the standard and robust ASMs are not aware of the spatial extent of the data set, because of the clamping of gradient values to the boundary (Section 2.2.a). Surfaces outside of the data set were clipped after the segmentation process was completed.

Segmentation results obtained with a standard ASM, our robust ASM, and the proposed approach (robust ASM and optimal surface finding (OSF) combined) are summarized in Table 1. To quantify segmentation performance, the Dice coefficient $D(S,R) = 2(|S \cap R)/(|S| + |R|)$ was calculated, where S represents the segmentation and R the reference.

On average, 115 seconds and 116 seconds were required by our robust lung segmentation method for right and left lung segmentation, respectively. The average time required for the intrinsic robust ASM matching process was 24 seconds. All computations were performed on a standard workstation equipped with a NVIDIA Tesla C1060 Computing Processor.

To assess the robustness of our approach, we applied our method to a CT data set where the top of the lung was not imaged (Fig. 3(b) and 3(c)). For comparison, the standard ASM result is depicted in Fig. 3(a). Figs. 1 and 4 allow us to compare our segmentation approach with a clinically used lung segmentation method.

4 Discussion and Conclusion

The experiments presented in Section 3 demonstrate the ability of our lung segmentation method to successfully deal with high-density lung pathology (Fig. 4) or other disturbances (Fig. 3). Our robust ASM matching clearly outperforms the standard ASM approach (Table 1). Even on normal lungs, the standard ASM delivers inferior performance. Pathologies like lung masses degrade matching performance even further, as demonstrated by the results in Table 1. These results are not surprising, because (standard) ASM matching is a least squares optimization, which is sensitive to outliers. Since the model is only roughly placed in proximity to the lung during initialization, all obstacles between model and target structure can cause problems.

The optimal surface finding step after robust ASM segmentation allows us to refine the robust ASM result. This step is needed, because our shape model was built from only 41 data sets, and consequently, some local shape variations cannot be explained by the model. In our final segmentation results, major deviation from the reference were observed in regions where airways and pulmonary vessels enter/leave the lung. Even for experts, it is hard to segment this area consistently.

So far, we have tested our robust lung segmentation method on lung cancer cases. In the future, we plan to significantly enlarge our training and test data set to include instances of fibrosis, pneumonia, etc. Our approach is currently limited to scans acquired at TLC. We plan to address this issue by developing a 4D lung model.

A method for the segmentation of lungs with high-density pathology was presented and evaluated on lung cancer cases. The robustness of our approach was demonstrated in 20 experiments, and a low segmentation error was achieved in cases with and without high-density pathology. A core component of our

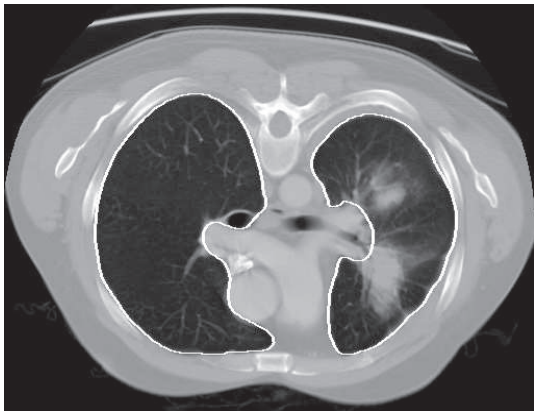


Fig. 4. Robust lung segmentation. Example shown in Fig. 1(a) segmented with the proposed robust lung segmentation approach.

method is a novel robust ASM matching approach that is well suitable for large shape models and is general applicable.

Our method for robust lung segmentation opens up new avenues for computer-aided lung image analysis. For example, segmentation of diseased lungs and segmentation of the diseased tissue itself are related problems. Thus, we expect that our method will be of significant benefit for the quantification of lung diseases.

Conflict of interest statement: Eric Hoffman and Geoffrey McLennan are founders and share holders of VIDA Diagnostics, a company which is commercializing lung image analysis software.

References

1. Armato, S.G., Sensakovic, W.F.: Automated lung segmentation for thoracic CT. Acad. Radiol. **11**(9) (2004) 1011–1021
2. Leader, J.K., Zheng, B., Rogers, R.M., Sciurba, F.C., Perez, A., Chapman, B.E., Patel, S., Fuhrman, C.R., Gur, D.: Automated lung segmentation in X-ray computed tomography. Acad. Radiol. **10**(11) (2003) 1224–1236
3. Silva, A., Silva, J.S., Santos, B.S., Ferreira, C.: Fast pulmonary contour extraction in X-ray CT images: a methodology and quality assessment. In: Proc. SPIE (Medical Imaging). Volume 4321. (2001) 216–224
4. Hu, S., Hoffman, E.A., Reinhardt, J.M.: Automatic lung segmentation for accurate quantitation of volumetric X-ray CT image. IEEE Trans. Med. Imag. **20**(6) (2001) 490–498
5. Hoffman, E.A., Sinak, L.J., Robb, R.A., Ritman, E.L.: Noninvasive quantitative imaging of shape and volume of lungs. Journal of Applied Physiology **54**(5) (1983) 1414–1421
6. Sluimer, I., Prokop, M., van Ginneken, B.: Toward automated segmentation of the pathological lung in CT. IEEE Trans. Med. Imag. **24**(8) (2005) 1025–1038

Case	Dice Coefficient					
	Right Lung			Left Lung		
	ASM	RASM	RASM+OSF	ASM	RASM	RASM+OSF
hdl01	0.9274	0.9549	0.9806	**0.8438**	**0.9301**	**0.9685**
hdl02	**0.8399**	**0.9436**	**0.9753**	0.8422	0.9453	0.9746
hdl03	**0.7456**	**0.9098**	**0.9606**	0.8505	0.9364	0.9724
hdl04	0.9071	0.9557	0.9761	**0.8473**	**0.9372**	**0.9686**
hdl05	**0.8368**	**0.9356**	**0.9707**	0.8490	0.9446	0.9741
hdl06	**0.9117**	**0.9600**	**0.9820**	0.8652	0.9440	0.9798
hdl07	**0.8117**	**0.9338**	**0.9824**	0.8007	0.9222	0.9751
hdl08	0.9315	0.9499	0.9800	**0.7836**	**0.9349**	**0.9768**
hdl09	**0.8810**	**0.9639**	**0.9785**	0.7571	0.9346	0.9699
hdl10	**0.9197**	**0.9425**	**0.9742**	**0.8505**	**0.9395**	**0.9775**
mean	0.8712	0.9450	0.9760	0.8290	0.9369	0.9737

Table 1. Lung segmentation results for standard ASM (ASM), robust ASM (RASM), and proposed (RASM+OSF) approach. Bold numbers represent cases with high-density pathology.

7. van Rikxoort, E., de Hoop, B., Viergever, M., Prokop, M., van Ginneken, B.: Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. Medical Physics **36** (2009) 2934–2947
8. Pu, J., Roos, J., Yi, C.A., Napel, S., Rubin, G.D., Paik, D.S.: Adaptive border marching algorithm: Automatic lung segmentation on chest CT images. Computerized Medical Imaging and Graphics **32**(6) (2008) 452 – 462
9. Kitasaka, T., Mori, K., Hasegawa, J., Toriwaki, J.: Lung area extraction from 3D chest X-ray CT images using a shape model generated by a variable Bézier surface. Systems and Computers in Japan **34**(4) (2003) 60–71
10. Prasad, M.N., Brown, M.S., Ahmad, S., Abtin, F., Allen, J., da Costa, I., Kim, H.J., McNitt-Gray, M.F., Goldin, J.G.: Automatic segmentation of lung parenchyma in the presence of diseases based on curvature of ribs. Academic Radiology **15**(9) (2008) 1173 – 1180
11. Heimann, T., Wolf, I., Williams, T., Meinzer, H.P.: 3D active shape models using gradient descent optimization of description length. In: In Proc. IPMI. Volume 3565., Springer, Heidelberg (2005) 566–577
12. Cootes, T.F., Cooper, D., Taylor, C.J., Graham, J.: Active shape models - their training and application. Computer Vision and Image Understanding **61**(1) (1995) 38–59
13. Rogers, M., Graham, J.: Robust active shape model search. In: In Proceedings of the European Conference on Computer Vision, Springer (2002) 517–530
14. Storer, M., Roth, P., Urschler, M., Bischof, H.: Fast-robust PCA. In: 16th Scandinavian Conference on Image Analysis. (2009) 430–439
15. Beichel, R., Bischof, H., Leberl, F., Sonka, M.: Robust active appearance models and their application to medical image analysis. IEEE Trans. Med. Imag. **24**(9) (2005) 1151–1169
16. Li, K., Wu, X., Chen, D., Sonka, M.: Optimal surface segmentation in volumetric images - a graph-theoretic approach. IEEE Trans. Pattern Anal. Mach. Intell. **28**(1) (2006) 119–134

Multi-classifier semi-supervised classification of tuberculosis patterns on chest CT scans

E. M. van Rikxoort[1], M. Galperin-Aizenberg[1], J.G. Goldin[1], T.T.J.P. Kockelkorn[2], B. van Ginneken[2,3], M.S. Brown[1]

[1]Department of Radiological Sciences, University of California Los Angeles, United States
[2]Image Sciences Institute, University Medical Center Utrecht, The Netherlands
[3]Diagnostic Image Analysis Group, Radboud University Nijmegen Medical Centre, The Netherlands

Abstract. Classification of different textures present in chest CT scans of patients with pulmonary tuberculosis (TB) is of crucial importance for the success of ongoing vaccine and drug testing trials. In this paper, a new multi-classifier semi-supervised method (MCSS) is proposed that is trained with a small set of labeled examples and improves classification performance by sampling interesting samples from unlabeled scans based on uncertainty among a pool of classifiers. The interesting samples are added to the small labeled set with a label assigned by 'expert' classifiers. MCSS is applied to 20 scans of patients with proven TB for which a reference standard was obtained by a consensus reading. Another set of 35 scans was used without manual labels. The performance of MCSS is compared to conventional supervised classification and two other semi-supervised methods and shown to outperform all other methods.

1 Introduction

Pulmonary tuberculosis (TB) is a major cause of morbidity and mortality world wide with 9.4 million new cases and 1.8 million deaths reported in 2008 [1]. Computed Tomography (CT) imaging is the most sensitive imaging technique for monitoring lung disease and can be used both for detecting and evaluating the progression of TB. On chest CT scans, TB presents as a wide variety of textural abnormalities. Quantifying the extent of TB is hard and time consuming even for expert radiologists. In addition, TB is most frequent in regions in the world where not many expert radiologists are available. Therefore, the development of computer aided diagnosis systems for detecting and quantifying TB is of crucial importance for the success of ongoing vaccine, drug testing, and screening programs. Some research toward detection and quantification of TB from chest radiographs has been performed, e.g. [2], however, no previous work on automatic quantification of TB from chest CT scans is available.

The large amount of data, combined with the difficulty of obtaining expert annotations, makes the problem of quantification of TB an excellent candidate for semi-supervised learning approaches. Semi-supervised learning is a popular technique in pattern recognition in which the performance of a classifier is improved by learning from unlabeled data, next to labeled data. There are a plethora of semi-supervised

methods available, an overview can be found in [3]. Two common paradigms are self-training and co-training. In self-training, a classifier is first trained with a small amount of labeled data. The trained classifier is then applied to the unlabeled data and the samples for which the classifier is most confident about the label are added to the labeled set. This process is iterated several times. A problem of self-training is that it mainly enforces already known knowledge and errors in the classifier. Co-training [4] requires the feature set to be divided into two sets that are conditionally independent. Using these two feature sets, two classifiers are trained, and each classifier teaches the other with unlabeled examples of which it is sure of the label. The assumption underlying co-training that conditionally independent feature sets exist is a limiting factor for many applications. Therefore, several studies have been performed applying so-called multiview learning, in which multiple models are trained using the same set of labeled data, e.g. [5, 6].

In this paper, a variation of multiview learning is proposed in which an active learning based uncertainty sampling strategy to find interesting samples in the unlabeled data to be added to the labeled set is used. Instead of the human experts used in active learning, a set of three classifiers which can be regarded 'expert' classifiers are employed. The main contributions of this paper are the uncertainty based selection method of unlabeled samples for semi-supervised learning and the application of semi-supervised learning to a multi-class problem. A similar approach coined 'tri-training' was proposed for two-class problems in [6]. In tri-training, three classifiers are trained. If two classifiers agree on the label, the label is added to the training dataset of the third classifier not taking into account the label or confidence of the third classifier. A disadvantage of tri-training is that if the models in tri-training are not sufficiently different, the method degenerates to single-classifier self-training.

Multi-classifier semi-supervised classification (MCSS) is applied to the classification of TB patterns in 20 scans from 20 different patients enrolled in a vaccine-testing trial. For all 20 scans manual annotations were obtained by a consensus reading of two expert radiologists. Scans of 35 additional patients were used as unlabeled data. The results of MCSS are compared to conventional supervised classification, multi-classifier self-training, and multiview learning with majority voting in a cross-validation procedure.

2 Materials

In this paper, 55 scans from 55 different patients with smear positive TB were used. All scans are low dose CT (30mAs at 120 kV), reconstructed to 512 $\times$ 512 matrices with a pixel size of 0.7 $\times$ 0.7 mm and a slice spacing and slice thickness of 1 mm. As a preprocessing step, the lungs in all scans were automatically segmented [7]. Next, the lungs in all scans were automatically divided into small volumes of interest (VOIs) with roughly similar texture [8], which was shown to outperform square regions of interest. On average 4170 VOIs with an average volume of 1.2 ml per VOI were produced per scan. For both the manual labeling and automatic method these volumes will be classified instead of single voxels.

A team of expert radiologists predefined a set of five textures that can be found in chest CT scans of patients with TB: normal lung, consolidations, nodules (cavitated and non-cavitated), TB fibrosis, and (small) airway disease. Examples of each texture class are shown in Figure 1. It can be appreciated that especially in the classes nodules and airway disease there is some variation in the appearance of the lung parenchyma. For 20 out of the 55 scans, manual annotations were obtained by a consensus reading

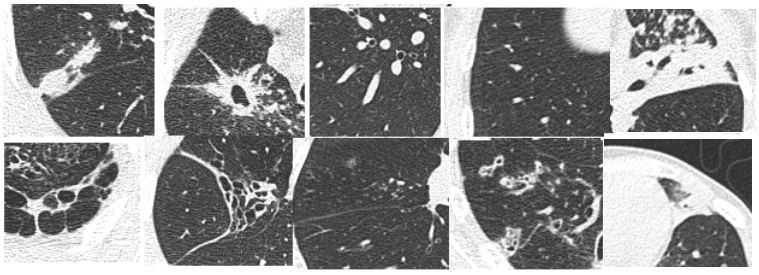


Fig. 1. Example patches of the different texture categories used. The first two images in the top row show examples of nodules (NOD), the second two images show normal lung regions (NL), the last image shows a consolidation (CONS). In the bottom row, the first two images show TB fibrosis (TBF), the second two images are examples of airway disease (AIR), an finally the last image is another example of a consolidation. It can be appreciated that even within the classes textures vary.

as follows: first, one expert radiologist annotated all 20 scans. Next, the second expert radiologist joined and together they went over all annotated regions. In case of doubt a consensus reading was performed. Since annotating is very time consuming and not all patterns occur in each scan, the radiologists were instructed to freely annotate in each scan classical examples of the available textures. Counters of the lung volume annotated were shown to the radiologist. In total, the first radiologist spent 8.5 hours annotating the scans. Consensus reading took another 3 hours. In total 2148 volumes were annotated after consensus reading. The division of the labeled VOIs over the different texture classes was: normal lung 423, consolidations 19, nodules 362, TB fibrosis 304, and airway disease 1040.

3 Methods

All supervised methods consist of a training phase, in which the classifier is trained, and a test phase in which the trained classifier is applied to test data. The difference between the different methods described in this section is the training phases, the outputs of all systems is a trained classifier, which is applied to test data to label each VOI in the test scans.

This section consists of four parts. First, the features and classifiers used for all methods are provided. In the second part, a conventional supervised texture classification system is described followed by a description of the proposed multi-classifier semi-supervised classification method in the third part. Finally, the semi-supervised methods implemented for comparison are briefly introduced.

Features and classifiers For each VOI, the first four statistical moments (kurtosis, skew, mean, standard deviation) of a set of 14 features on four scales were used, resulting in 224 features in total. The image features used were the output of Gaussian filters up to and including second order derivatives (L, L_x, L_y, L_z, L_{xx}, L_{xy}, L_{xz}, L_{yy}, L_{yz}, L_{zz}),the gradient (L_i), and the eigenvalues of the Hessian matrix ($|\lambda_0| \geq |\lambda_1| \geq |\lambda_2|$). All features were calculated on scales 1, 2, 4, and 8. Classification

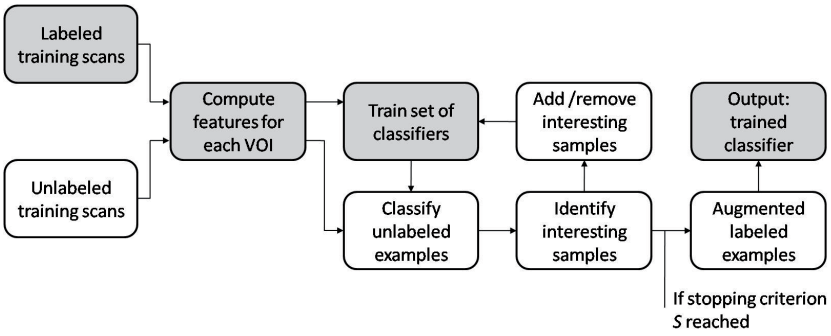


Fig. 2. Flowchart of multi-classifier semi-supervised (MCSS) segmentation. Steps indicated in gray are the steps performed during conventional supervised classification.

was preceded by a principal component analysis (PCA) retaining 95% of variance for the purpose of dimensionality reduction.

For training purposes, the initial labeled set was randomly sub sampled to contain equal class sizes for all 5 classes. However, since for one class, consolidations, the number of samples is substantially lower than for the others, this class is not taken into account during subsampling. As a result, all classes contain 304 labeled examples except consolidations, which contains 19 labeled examples. All 2148 labeled VOIs will be used for evaluation in a leave-one-patient-out cross validation procedure.

For conventional supervised texture classification a support vector machine (SVM) classifier with radial basis kernel functions was used as a classifier. The settings of the SVM were determined on the initial labeled set using cross validation [9]. For MCSS and other semi-supervised methods a linear discriminant classifier (LDC) and k-NN classifier (KNN) with k equal to 7 were added to the pool of classifiers. The final classifier after all semi-supervised methods is the SVM trained with the extended labeled set.

Conventional supervised texture classification (CVS) In the training phase of CVS, the set of features is computed for all VOIs which have been labeled. Based on the features and the known output labels, the SVM classifier is trained to be able to assign a label to previously unseen VOIs. In the flowchart in Figure 2 the steps to train a classifier in a conventional supervised texture classification system are indicated in gray.

For a test scan, the feature vector is calculated for each VOI and the trained SVM classifier assigns a label to each feature vector.

Multi-classifier semi-supervised texture classification (MCSS) The goal of semi-supervised classification is to extend the initial labeled set using unlabeled examples to increase classifier performance. A flowchart of the proposed semi-supervised method is provided in Figure 2. Globally, the procedure of MCSS is as follows: given a set of labeled VOIs, L, and a set of unlabeled VOIs, U, a pool of n classifiers $C_i,\ i = 1, \ldots, n$ is trained using L. The trained classifiers C_i are applied to all samples s_u of U. Inspecting the posterior probabilities p of each C_i for each s_u, 'interesting' s_u are identified that should be removed from U and added to L with the label assigned by the pool of classifiers. This process is iterated until a stopping criterion S is reached.

When S is reached, the extended labeled set L is used to train a final classifier. The key point of this scheme is the identification of 'interesting' samples in the unlabeled data.

We propose to use a paradigm that is used in active learning to identify interesting samples: if there is uncertainty about the label of a sample, this is an informative sample and an expert opinion should be obtained. In MCSS we view the set of classifiers used as individual experts. Uncertainty is defined in two ways: if one of the classifiers is unsure about the label to assign but the other classifiers agree on the label with high confidence, the sample is added to the labeled set with the label of the agreeing classifiers. Or secondly, uncertainty about the label exists if two classifiers agree with a high confidence on the label of a sample but the third classifier is confident about another label. In this case the sample is added to the labeled set with the label of the agreeing classifiers. The labeled example is added to the training dataset of all classifiers, which leads to a combination of majority voting (all but one classifiers agree with high confidence) and uncertainty sampling (only if the remaining classifier is unsure or disagrees on the label). The rationale behind this approach is twofold. First, samples for which uncertainty exist are the most informative since they change the classifier as opposed to samples for which all classifiers agree. And second, the high agreement between the two 'expert' classifiers makes it more likely that they made the right decision. MCSS is implemented by setting three parameters: the posterior probability p at which a classifier is confident, $p > p_c$, and the posterior probabilities between which a classifier is unsure of its label, $p_{ul} < p < p_{uh}$. For classification of TB textures, p_c was set to 0.6 and p_{ul} and p_{uh} were set to 0.2 and 0.5, respectively.

Once all unlabeled samples have been processed, the classifiers are retrained with the extended labeled set and the process is iterated. Due to the unbalanced appearance of the different structures in the data, a pruning step is performed in each iteration of MCSS; the classes are pruned to be of equal size, the size being the size of the one but smallest class. Since all classifiers used are provided new examples, they are refined in each iteration and therefore different results will be obtained after each iteration. It is important to note that the classifiers used in this scheme should be diverse since if their labeling of unlabeled samples is identical, no interesting features can be identified. The final parameter to be set is a stopping criterion S for iterating this process. In this paper, S was set to no more interesting samples being identified for at least 2 of the classes. As a final classifier, any of the classifiers used can be trained or a combination of them can be used. In this paper the SVM classifier trained with the final extended labeled set is used as a final classifier.

It is obvious that if the class label assigned by the agreeing classifiers is correct, the training data is augmented with a sample with a valid label. Otherwise, a noisy label will be added to the training dataset. In [6] it is shown that if the amount of newly added examples is sufficient, the increase in classification noise can be compensated for.

Semi-supervised methods implemented for comparison For comparison, two other semi-supervised classification methods were implemented: multi-classifier self training (MCST) in which a label is added when all classifiers agree with high confidence, and multiview majority voting (MMV), in which a label is added if at least two classifiers agree with high confidence. For MCST and MMV the same p_c as for MCSS was used.

4 Experiments and Results

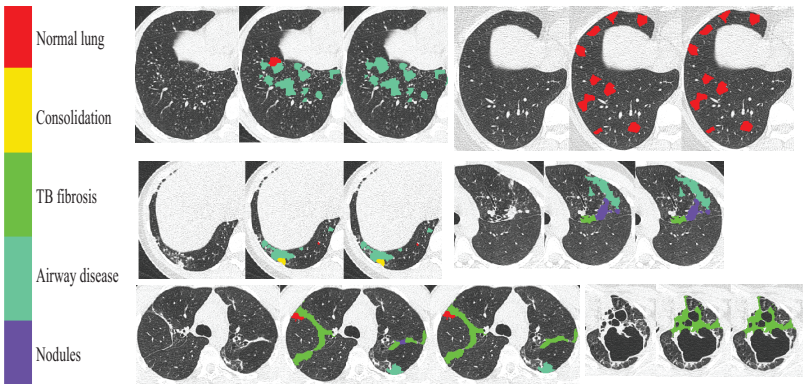


Fig. 3. Example output of MCSS and the ground truth for several slices of different scans. For each group of 3 slices, the left slice is the original slice, the middle slice is the output of MCSS, and the last slice shows the ground truth. Two wrong assignments can be seen in these examples: the top left example shows a VOI that was labeled as normal lung by MCSS and as airway disease by the observers. The left images in the bottom row show an example of fibrosis mislabeled as lesion, which is the most common mistake of MCSS.

All experiments were performed in a leave-one-patient out cross validation procedure. On average, MCSS performed 5 iterations, adding 6024 samples to the labeled set in total. 1477 samples were added with the label nodules, 50 samples were added to the consolidations class, and 1499 labels were added to the other three classes. The equal number for the last three classes is due to the pruning that is performed after each iteration. To show the validity of the classifiers used as experts during semi-supervised learning, we calculated their accuracy on the initial labeled set for those labels for which $p > p_c$. The LDC classifies 93% of the samples with $p > p_c$ with an accuracy of 0.83. For the KNN classifier 66% of the samples has $p > p_c$ with an overall accuracy of 0.81, the SVM classifier classifies 88% of the samples with $p > p_c$ with an accuracy of 0.85.

Figure 3 shows for several scans an original slice, the result of MCSS, and the manual labeling. Table 1 provides the confusion matrices for CVS, MCSS, MCST, and MMV. The accuracies for the different classes are provided next to the confusion matrix, both in total and per class. Next to overall improvement in performance, it can be seen that especially classes with a low accuracy (consolidations and TB fibrosis) in the CVS get boosted when applying MCSS. In addition, MCSS performs better than the other semi-supervised methods, not only in total but also for every class separately; there is no class for which one of the other methods performs better.

Table 1. Confusion matrices and accuracies for CVS, MCSS, MCST, and MMV. The rows depict the ground truth, columns the output of the automatic methods.

		CVS				acc	MCSS					acc
		N	**C**	**N**	**F**	**A**	0.816	**N**	**C**	**N**	**F**	**A**	0.863
GT	**NL**	405	0	0	4	14	0.957	418	0	0	0	5	0.988
	CONS	0	11	7	0	1	0.579	0	17	2	0	0	0.895
	NOD	5	2	270	75	10	0.746	3	0	302	54	3	0.834
	TBF	8	0	87	185	24	0.609	2	0	51	235	16	0.773
	AIR	74	0	32	51	883	0.849	70	0	42	45	883	0.849
		MCST					acc	**MMV**					acc
		N	**C**	**N**	**F**	**A**	0.831	**N**	**C**	**N**	**F**	**A**	0.837
GT	**NL**	409	0	0	1	13	0.967	416	0	0	1	6	0.983
	CONS	0	15	3	0	1	0.789	0	17	2	0	0	0.895
	NOD	6	0	282	65	9	0.779	4	0	275	73	10	0.760
	TBF	8	0	59	214	23	0.704	11	0	58	214	21	0.701
	AIR	99	0	29	47	865	0.831	88	0	30	44	878	0.844

5 Conclusion & Discussion

This paper presents a multi-classifier semi-supervised approach (MCSS) for classification of TB textures on chest CT scans. The problem of TB classification is highly appropriate for a semi-supervised approach due to the difficulty of obtaining manual annotations. The main contribution of this paper is the selection of interesting unlabeled samples to add to the labeled set based on uncertainty sampling. The proposed method performs well for the task of classification of TB textures and outperforms conventional supervised classification as well as other well-known semi-supervised methods. The proposed method of identifying interesting samples from unlabeled data is especially fit for multi-class problems since classifiers are more often unsure in these cases.

A limitation of this study is the relatively small number of scans that were manually annotated. In addition, due to the difficulty of the task, the the observers only indicated classical examples of the different textures. This potentially makes the classification task easier and boosts the performance of the automatic classification. To improve the manual annotations for future work, an active learning step will be introduced in which observers are asked to annotate VOIs for which the classifier is unsure.

Several parameters have to be set in MCSS. The influence of the setting of these parameters has not been studied in this paper. In general, the higher p_c, the less samples will be added to the labeled data but the more confident the classifiers are about the assigned label. For the application of TB classification setting p_c above 0.90 leads to only sampling normal VOIs from the unlabeled data. The setting of p_{ul} and p_{uh} in this paper were based on the fact that a five class classification task was performed; a posterior probability between 0.2 and 0.5 in a five class problem indicates that the posterior probabilities are relatively spread over at least 3 classes.

For any semi-supervised method it is important that the samples added to the labeled set have correct labels. Since it is unavoidable that also samples with incorrect labels are added to the labeled set, it is important to have a large pool of unlabeled data. In this paper a set of 35 unlabeled scans was used. Due to the variation in appearance of the abnormalities and the difference in their prior probability (e.g. consolidations vs.

airway disease) future work includes extending the set of unlabeled scans used during MCSS.

To conclude, a multi-classifier semi-supervised classification method was presented that was applied to the classification of TB texture patterns. The proposed method was shown to be able to increase classification performance compared to conventional supervised classification by adding unlabeled samples to the labeled dataset and outperform other semi-supervised methods implemented for comparison.

References

1. World Health Organization: WHO Report 2009: Global tuberculosis control, Epidemiology, Strategy, financing (2009)
2. Arzhaeva, Y., Hogeweg, L., de Jong, P., Viergever, M., van Ginneken, B.: Global and local multi-valued dissimilarity-based classification: application to computer-aided detection of tuberculosis. In: Medical Image Computing and Computer-Assisted Intervention. Volume 5762 of Lecture Notes in Computer Science. (2009) 724–731
3. Chapelle, O., Schölkopf, B., Zien, A., eds.: Semi-Supervised Learning. MIT Press, Cambridge, MA (2006)
4. Blum, A., Mitchell, T.: Combining labeled and unlabeled data with co-training. In: Proceedings of the eleventh annual conference on Computational learning theory. (1998) 92–100
5. Zhou, Y., Goldman, S.: Democratic co-learning. In: Proceedings of the 16th IEEE International Conference on Tools with Artificial Intelligence. (2004) 594–602
6. Zhou, Z.H., Li-Yueh, M.: Tri-Training: Exploiting unlabeled data using three classifiers. IEEE Transactions on Knowledge and Data Engineering **17**(11) (2005) 1529–1541
7. van Rikxoort, E.M., de Hoop, B., Viergever, M.A., Prokop, M., van Ginneken, B.: Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. Medical Physics **36**(7) (2009) 2934–2947
8. Kockelkorn, T.T.J.P., de Jong, P.A., Gietema, H.A., Grutters, J.C., Prokop, M., van Ginneken, B.: Interactive annotation of textures in thoracic CT scans. In: SPIE Medical Imaging. Volume 7624. (2010) 76240X1–76240X8
9. Chang, C., Lin, C.: LIBSVM: a library for support vector machines (2001) Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

Comparison of Intensity- and Jacobian-Based Estimates of Lung Regional Ventilation

Kai Ding[1], Kunlin Cao[2], Ryan E. Amelon[1], Gary E. Christensen[2]
Madhavan L. Raghavan[1], and Joseph M. Reinhardt[1⋆]

[1] Department of Biomedical Engineering
[2] Department of Electrical and Computer Engineering
The University of Iowa, Iowa City, IA 52242

{kai-ding, kunlin-cao, ryan-amelon, gary-christensen
ml-raghavan ,joe-reinhardt}@uiowa.edu

Abstract. Regional ventilation is the measurement of pulmonary function on a local, or regional level. The measurement of pulmonary function may be useful as a planning tool during radiation therapy (RT) planning, may be useful for tracking the progression of toxicity to nearby normal tissue during RT and can be used to evaluate the effectiveness of a treatment post-therapy. In this paper, we show that an intensity-based regional ventilation measure can be derived from a Jacobian-based measure by relaxing the assumption that there is no tissue volume change. We compare intensity-based and Jacobian-based measures of regional ventilation to xenon-CT (Xe-CT) measures of specific ventilation. The results show that the Jacobian-based measure correlates better (average $r^2 = 0.80$) with Xe-CT-based measures of specific ventilation than the intensity-based measure (average $r^2 = 0.54$). The difference between the intensity-based measure and the Jacobian-based measure of regional ventilation is linearly related to the tissue volume difference between scans (average $r^2 = 0.86$).

1 Introduction

Regional ventilation is the term used to characterize the volume of air per unit time that enters or exits the lung on a local, or regional, level. Since the primary function of the lung is gas exchange, ventilation can be interpreted as an index of lung function. Injury and disease processes can alter lung function on a global and/or a local level. Recent advances in multi-detector-row CT (MDCT), 4DCT respiratory gating methods, and image processing techniques enable us to study pulmonary function at the regional level with high resolution anatomical information compared to other methods. MDCT can be used to acquire multiple static breath-hold CT images of the lung taken at different lung volumes, or 4DCT images of the lung reconstructed at different respiratory phases with proper respiratory control. Image registration can be applied to this data to

⋆ J. M. Reinhardt is shareholder in VIDA Diagnostics, Inc.

estimate a deformation field that transforms the lung from one volume configuration to the other. This deformation field can be analyzed to estimate local lung tissue expansion, calculate voxel-by-voxel intensity change, ventilation, and make biomechanical measurements [1–5].

An important emerging application of these methods is the detection of pulmonary function change in subjects undergoing radiation therapy (RT) for lung cancer. During RT, treatment is commonly limited to sub-therapeutic doses due to unintended toxicity to normal lung tissue. Reducing the frequency of occurrence and magnitude of normal lung function loss may benefit from treatment plans that incorporate relationships between regional and functional based lung information and the radiation dose. Measurement of pulmonary function may be useful as a planning tool during RT planning, may be useful for tracking the progression of toxicity to nearby normal tissue during RT, and can be used to evaluate the effectiveness of a treatment post-therapy [6, 7].

The physiologic significance of the registration-based measures of respiratory function can be established by comparing to more conventional measurements, such as nuclear medicine or contrast wash-in/wash-out studies with CT or MR. Xenon-enhanced CT (Xe-CT) measures regional ventilation by observing the gas wash-in and wash-out rate on serial CT images [8]. Xe-CT imaging has the advantage of high temporal resolution and anatomical information. Although it comes along with limited axial coverage, it can be used to compare with registration-based measures of regional lung function in animal studies for validation purpose.

This paper describes two measures to estimate regional ventilation from image registration of CT images: the intensity based and the Jacobian based measures. We show that the intensity based regional ventilation measure can be derived from Jacobian based measure by making the assumption that there is no tissue volume change. They are evaluated by comparison with Xe-CT estimated ventilation. These results may provide insight into which measures may best estimate regional ventilation using image registration of respiratory-gated CT images.

2 Material and methods

2.1 Data Acquisition

Three sheep were anesthetized and mechanically ventilated during experiments. The 4DCT images were acquired with the animals in the supine position using the dynamic imaging protocol and images were reconstructed retrospectively. The 0% (EE) and 100% (EI) inspiration phases were used for ventilation measurements. Twelve contiguous axial locations and approximately 45 breaths for the Xe-CT studies were selected from the whole lung volumetric scan performed near end-expiration (EE_0 to EE_{44}). The animal was not moved between scans. The respiratory rate (RR) for three animals ranges from 15 to 18 breaths per minute. Both the 4DCT and the Xe-CT images were reconstructed using a ma-

trix of 512 by 512 pixels. The in-plane pixel spacing is approximately 0.5 mm × 0.5 mm.

2.2 Image Registration

The tissue volume and vesselness measure preserving nonrigid registration (TVP) algorithm is used to estimate transforms EI to EE and EE_0 to EE. The algorithm minimizes the sum of squared tissue volume difference (SSTVD) [9] and vesselness measure difference (SSVMD), utilizing the rich image intensity information and natural anatomic landmarks provided by the vessels. This method has been shown to be effective at registering across lung CT images with high accuracy [10, 11].

Let I_1 and I_2 represent two 3D image volumes to be registered. The vector $\mathbf{x}$ defines the voxel coordinate within an image. The algorithm find the optimal transformation $\mathbf{h}$ that maps the template image I_1 to the target image I_2 by minimizing the cost function

$$C_{\mathrm{TOTAL}} = \rho \int_{\mathbf{\Omega}} \left[V_2(\mathbf{x}) - V_1(\mathbf{h}(\mathbf{x}))\right]^2 d\mathbf{x} + \chi \int_{\mathbf{\Omega}} \left[F_2(\mathbf{x}) - F_1(\mathbf{h}(\mathbf{x}))\right]^2 d\mathbf{x}. \quad (1)$$

where Ω is the union domain of the images I_1 and I_2. The first integral of the cost function defines the SSTVD cost and the second integral of the cost function defines the SSVMD cost.

The SSTVD cost assumes that the measured Hounsfield units (HU) in the lung CT images is a function of tissue and air content. Following the findings by Hoffman et. al [12], from the CT value of a given voxel, the tissue volume can be estimated as

$$V(\mathbf{x}) = \nu(\mathbf{x}) \frac{I(\mathbf{x}) - HU_{air}}{HU_{tissue} - HU_{air}} = \nu(\mathbf{x})\beta(I(\mathbf{x})), \quad (2)$$

and the air volume can be estimated as

$$V'(\mathbf{x}) = \nu(\mathbf{x}) \frac{HU_{tissue} - I(\mathbf{x})}{HU_{tissue} - HU_{air}} = \nu(\mathbf{x})\alpha(I(\mathbf{x})), \quad (3)$$

where $\nu(\mathbf{x})$ denotes the volume of voxel $\mathbf{x}$ and $I(\mathbf{x})$ is the intensity of a voxel at position $\mathbf{x}$. HU_{air} and HU_{tissue} refer to the intensity of air and tissue, respectively. In this work, we assume that air is -1000 HU and tissue is 0 HU. $\alpha(I(\mathbf{x}))$ and $\beta(I(\mathbf{x}))$ are introduced for notational simplicity, and $\alpha(I(\mathbf{x})) + \beta(I(\mathbf{x})) = 1$.

Given (2), we can then define the SSTVD cost:

$$C_{\mathrm{SSTVD}} = \int_{\mathbf{\Omega}} \left[\nu_2(\mathbf{x})\beta(I_2(\mathbf{x})) - \nu_1(\mathbf{h}(\mathbf{x}))\beta(I_1(\mathbf{h}(\mathbf{x})))\right]^2 d\mathbf{x} \quad (4)$$

With the warping function $\mathbf{h}(\mathbf{x})$, $I_1(\mathbf{h}(\mathbf{x}))$ can be interpolated from the template image. $\nu_1(\mathbf{h}(\mathbf{x}))$ can be calculated from the Jacobian $J(\mathbf{x})$ of the deformation as $\nu_1(\mathbf{h}(\mathbf{x})) = J(\mathbf{x})\nu_2(\mathbf{x})$.

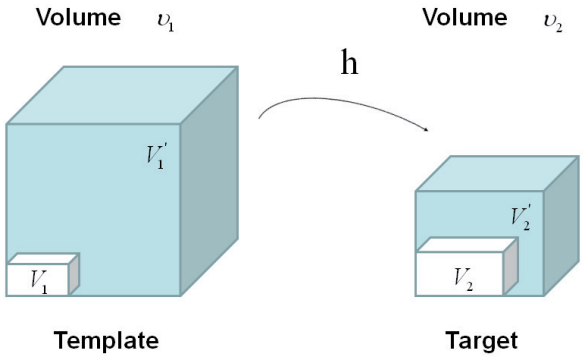


Fig. 1. Example of a given voxel under deformation $\mathbf{h}(\mathbf{x})$ from template image to target image. V_1 and V_2 are tissue volumes. V_1' and V_2' are air volumes.

Figure 1 shows an example of a given cube under deformation $\mathbf{h}$ from template image to target image. The total cube volumes are ν_1 and ν_2. The total cube volume can be decomposed into the tissue volume and air volume based on their individual intensity. The small white volume inside the cube represents the tissue volume V_1 and V_2. The air volume is represented by V_1' and V_2' in blue. As the ratio of air to tissue decreases, the CT intensity of a voxel increases. The intensities of the cubes in the template image and the target image are I_1 and I_2 and are a function of the ratio of air to tissue content of the cube.

As the blood vessels branch to small diameters, the raw grayscale information from vessel voxels provide almost no contribution to guide the intensity-based registration. To better utilize the information of blood vessel locations, we use the vesselness measure (VM) based on the eigenvalues of the Hessian matrix of image intensity. Frangi's vesselness function [13] is defined as

$$F(\lambda) = \begin{cases} (1 - e^{-\frac{R_A^2}{2\alpha^2}}) \cdot e^{\frac{-R_B^2}{2\beta^2}} \cdot (1 - e^{-\frac{S^2}{2\gamma^2}}) & \text{if } \lambda_2 < 0 \text{ and } \lambda_3 < 0 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

with

$$R_A = \frac{|\lambda_2|}{|\lambda_3|}, \qquad R_B = \frac{|\lambda_1|}{\sqrt{|\lambda_2 \lambda_3|}}, \qquad S = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}, \tag{6}$$

where R_A distinguishes between plate-like and tubular structures, R_B accounts for the deviation from a blob-like structure, and S differentiates between tubular structure and noise. α, β, γ control the sensitivity of the vesselness measure. The vesselness measure is rescaled to [0, 1] and can be considered as a probability-like estimate of vesselness features. For this study, $\alpha = 0.5$, $\beta = 0.5$, and $\gamma = 5$. The transformation $\mathbf{h}(\mathbf{x})$ is a cubic B-splines transform. Note that the Jacobian value must be positive here, which can be achieved by using displacement constraints on the control nodes [14]. The total cost in equation 1 is optimized using a limited-memory, quasi-Newton minimization method with bounds (L-BFGS-B) algorithm.

2.3 Regional Ventilation Measures from Image Registration

After we obtain the optimal warping function, we can calculate the regional ventilation, which is equal to the difference in local air volume change per unit time. Therefore, the specific ventilation (SV) is equal to specific air volume change per unit time.

Specific air volume change by corrected Jacobian (SACJ): Applying the same assumptions (2) and (3) used in the SSTVD cost function, we have

$$SACJ = \frac{V_1'(\mathbf{h}(\mathbf{x})) - V_2'(\mathbf{x})}{V_2'(\mathbf{x})} \tag{7}$$

$$= \frac{\nu_1(\mathbf{h}(\mathbf{x}))\alpha(I_1(\mathbf{h}(\mathbf{x})))}{\nu_2(\mathbf{x})\alpha(I_2(\mathbf{x}))} - 1 \tag{8}$$

Given a warping function $\mathbf{h}(\mathbf{x})$, $I_1(\mathbf{h}(\mathbf{x}))$ can be interpolated from the template image. $\nu_1(\mathbf{h}(\mathbf{x}))$ can be calculated from the Jacobian $J(\mathbf{x})$ of the deformation as $\nu_1(\mathbf{h}(\mathbf{x})) = J(\mathbf{x})\nu_2(\mathbf{x})$. Therefore, the specific air volume change is then

$$SACJ = J(\mathbf{x})\frac{\alpha(I_1(\mathbf{h}(\mathbf{x})))}{\alpha(I_2(\mathbf{x}))} - 1 \tag{9}$$

$$= J(\mathbf{x})\frac{HU_{tissue} - I_1(\mathbf{h}(\mathbf{x}))}{HU_{tissue} - I_2(\mathbf{x})} - 1 \tag{10}$$

$$= J(\mathbf{x})\frac{I_1(\mathbf{h}(\mathbf{x}))}{I_2(\mathbf{x})} - 1. \tag{11}$$

The correction factor $\frac{I_1(\mathbf{h}(\mathbf{x}))}{I_2(\mathbf{x})}$ above depends on the voxel intensity.

Specific air volume change by intensity (SAI): The intensity based measure of regional air volume change SAI can be derived from the SACJ. Now we introduce another assumption that the tissue volume is preserved, or equivalently, that the tissue volume difference $\Delta V(\mathbf{x}) = V_1(\mathbf{h}(\mathbf{x})) - V_2(\mathbf{x}) = 0$. Under this assumption,$V_1(\mathbf{h}(\mathbf{x})) = V_2(\mathbf{x})$ and we have

$$\nu_1(\mathbf{h}(\mathbf{x}))\beta(I_1(\mathbf{h}(\mathbf{x}))) = \nu_2(\mathbf{x})\beta(I_2(\mathbf{x})), \tag{12}$$

and

$$\nu_1(\mathbf{h}(\mathbf{x})) = \nu_2(\mathbf{x})\frac{\beta(I_2(\mathbf{x}))}{\beta(I_1(\mathbf{h}(\mathbf{x})))}, \tag{13}$$

Since $\nu_1(\mathbf{h}(\mathbf{x})) = J(\mathbf{x})\nu_2(\mathbf{x})$, with above equation, we have

$$J(\mathbf{x}) = \frac{\beta(I_2(\mathbf{x}))}{\beta(I_1(\mathbf{h}(\mathbf{x})))} \tag{14}$$

$$= \frac{I_2(\mathbf{x}) - HU_{air}}{I_1(\mathbf{h}(\mathbf{x})) - HU_{air}}. \tag{15}$$

Substituting the above equation into equation 10 with assumption that air is -1000 HU and tissue is 0 HU, then

$$SAI = 1000\frac{I_1(\mathbf{h}(\mathbf{x})) - I_2(\mathbf{x})}{I_2(\mathbf{x})(I_1(\mathbf{h}(\mathbf{x})) + 1000)} \tag{16}$$

which is exactly the equation as described in Simon [15], Guerrero et al. [1] and Fuld et al. [16].

Difference of specific air volume change (DSA) and difference of tissue volume (DT): To investigate the relationship between the measurements of specific air volume changes and the tissue volume change, we also measure the difference between equation (11) and equation (16) by comparing the difference of specific air volume change (DSA) between SACJ and SAI, and the difference of tissue volume (DT) as:

$$DSA = SACJ - SAI \tag{17}$$

$$\begin{aligned} DT &= V_1(\mathbf{h}(\mathbf{x})) - V_2(\mathbf{x}) && (18)\\ &= J(\mathbf{x})\nu_2(\mathbf{x})\beta(I_1(\mathbf{h}(\mathbf{x}))) - \nu_2(\mathbf{x})\beta(I_2(\mathbf{x})) && (19)\\ &= \nu_2(\mathbf{x})\frac{J(\mathbf{x})(I_1(\mathbf{h}(\mathbf{x})) - HU_{air}) - (I_2(\mathbf{x}) - HU_{air})}{HU_{tissue} - HU_{air}} && (20)\\ &= \nu_2(\mathbf{x})\frac{J(\mathbf{x})(I_1(\mathbf{h}(\mathbf{x})) + 1000) - (I_2(\mathbf{x}) + 1000)}{1000}. && (21) \end{aligned}$$

In this study, the absolute values of DT and DSA are used in analysis.

2.4 Compare Registration Regional Ventilation Measures to Xe-CT Estimated Ventilation

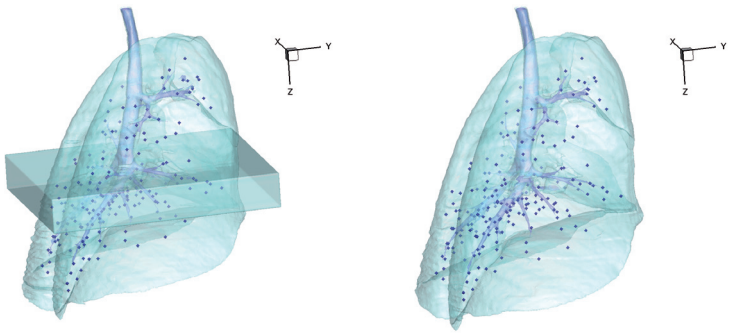
The Xe-CT estimate of SV is computed in the coordinates of the EE_0 using Pulmonary Analysis Software Suite 11.0 (PASS) [17] at the original image size of 0.5 mm × 0.5 mm × 2.4 mm voxels. Overlapping 1 × 8 regions of interest (ROI) are defined in the lung region on each 2D slice. All the images including the EE, EI, EE_0 and SV map are then resampled to a voxel size of 1 mm × 1 mm × 1 mm. After preprocessing, EI is registered to EE using the TVP for measuring the regional ventilation from these two phases in a 4DCT scan. The resulting transformation is used to estimate the SACJ and SAI. Then EE_0 is registered to EE using TVP as well to map the SV to the same coordinate system as that of the SACJ and SAI. A semi-automatic landmark system is used for landmark detection and annotation [18]. For each animal, after 200 anatomic landmarks are identified in the EE, the observer marks the locations of the voxels corresponding to the anatomic locations of the landmarks in the EI. For each landmark, the actual landmark position is compared to the registration-derived estimate of landmark position and the error is calculated. To compare

the regional ventilation measures, the corresponding region of Xe-CT image EE_0 in the EE is divided into about 100 cubes with size of 20 mm $\times$ 20 mm $\times$ 20 mm. We compare the average regional ventilation measures (SACJ and SAI) to the corresponding average SV measurement from Xe-CT images within each cube. The correlation coefficients between any two estimates (SACJ-SV or SAI-SV) are calculated by linear regression. To compare two correlation coefficients, the Fisher Z-transform of the r values is used and the level of significance is determined [19]. The relationship between the specific air volume change and difference of tissue volume is also studied in three animals by linear regression analysis.

3 Results

3.1 Registration Accuracy

Approximately 200 automatic identified landmarks within the lungs are used to compute the registration accuracy. The landmarks are uniformly distributed in the lung regions. Figure 2 shows an example of the distribution of the landmarks in on animal for both the EE and EI images. The coordinate of each landmark location is recorded for each image data set before and after registration. For all three animals, before registration, the average landmark distance is 6.6 mm with minimun 1.0 mm, maximum 14.6 mm and stand deviation 2.42 mm. After registration, the average landmark distance is 0.4 mm with minimum 0.1 mm, maximum 1.6 mm and stand deviation 0.29 mm.



(a) (b)

Fig. 2. 3D view of the landmarks in: (a) EE with EE_0 and (b) EI. The dark region below the carina in (a) is the EE_0 and the spheres are the automatically defined landmarks.

3.2 Registration Estimated Ventilation Compared to Xe-CT Estimated Ventilation

Figure 3(a) shows an example of the color-coded cubes of the regions where we average the registration estimated ventilation measures and the Xe-CT estimated SV and compare them. For each animal, the corresponding Xe-CT regions in the EE are divided into about 100 cubes. Figure 3(b) is the Xe-CT estimate of SV. Figure 3(c), (d) are the corresponding registration ventilation measures SACJ amd SAI. The regions with edema are excluded from the comparison. Figure 3(b) to (d) all show noticeable similar gradient in the ventral-dorsal direction. Notice that the color scales are different in each map and are set according to their ranges in Fig. 4.

Figure 4 shows the scatter plots between the registration ventilation measures and the Xe-CT ventilation SV with linear regression in all three animals. The SACJ column shows the stronger correlation (average $r^2 = 0.80$) than The intensity based measure SAI (average $r^2 = 0.54$) with the SV.

Table 1 shows the results of comparing the r values from SACJ vs. SV and SAI vs. SV. All three animals show that the correlation coefficient from SACJ vs. SV is significantly stronger than it from SAI vs. SV.

Figure 5 shows the scatter plots between DSA (the absolute difference of the value between the SACJ and SAI) and the DT (the absolute difference of the tissue volume) with linear regression in all three animals (average $r^2 = 0.86$). From the equation (11) and (16), we know that the SAI takes the assumption about no tissue volume change for a given voxel between the two volumes which may not be valid. Figure 5 shows that as the tissue volume change increases, the difference between the measures of regional ventilation from SACJ and SAI increases linearly in all animals.

Table 1. Comparison of ventilation measures between SACJ and SAI in small cube ROIs with size 20 mm × 20 mm × 20 mm

Animal	Correlation pair (with SV)	Correlation with SV (r value)	Number of samples	p value
A	SACJ	0.88	83	p<= 0.0001
	SAI	0.65		
B	SACJ	0.93	119	p<= $1.18e^{-6}$
	SAI	0.77		
C	SACJ	0.89	86	p<= 0.015
	SAI	0.78		

4 Discussion

We have described two measures to estimate regional ventilation from tissue volume and vesselness preserving image registration of CT images. The validity

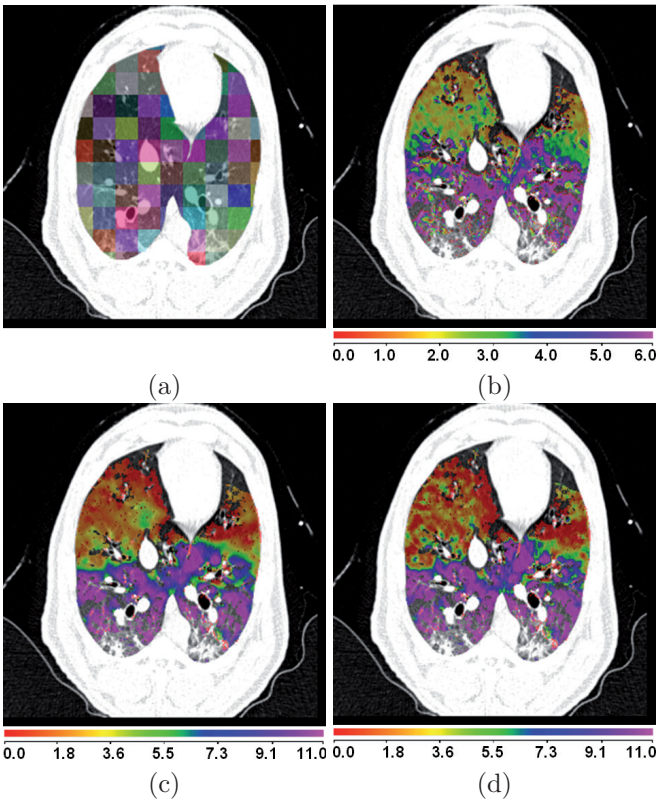


Fig. 3. (a): EE with color coded cubes showing the sample region. (b), (c), and (d): color map of the SV, SACJ and SAI.

and comparison of different measures for estimates of regional ventilation are evaluated by Xe-CT estimated ventilation. Individual regional ventilation measures are compared to Xe-CT estimates of ventilation by transforming them to the same coordinate system. The difference between two registration measures and their relationship with the tissue volume difference is analyzed using linear regression.

The tissue volume and vesselness preserving algorithm is used to register the EI to the EE for estimating ventilation measures. It is also used to register the EE_0 to the EE for comparing two ventilation measures to the Xe-CT based SV. About 200 anatomical landmarks are identified and annotated to evaluate the registration accuracy. The average landmark error is on the order of 1 mm after registration.

The ventilation measures SACJ and SAI are derived and the relationship between them is described. The SACJ is is based on the voxel air-tissue fraction theory of HU. With further assumption about no change in the tissue volume between the corresponding voxels in the template and target images, SAI can be

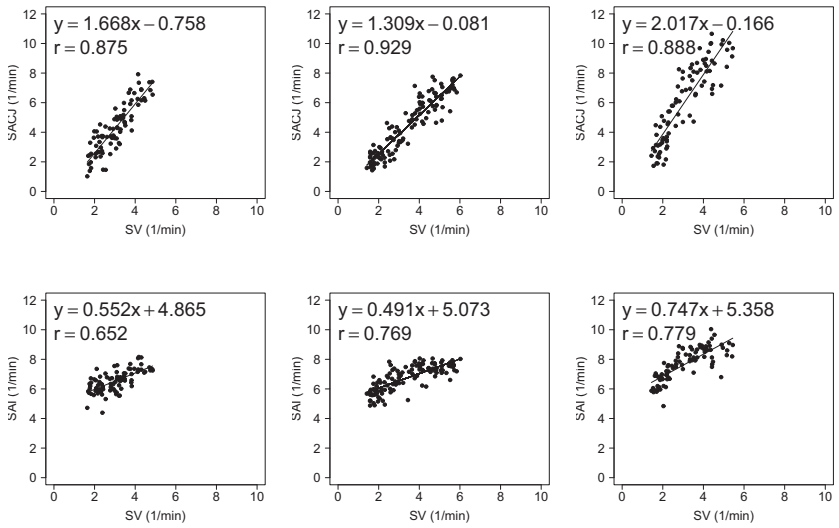


Fig. 4. Small cube ROIs with size 20 mm × 20 mm × 20 mm results for registration estimated ventilation measures compared to the Xe-CT estimated ventilation SV in scatter plot with linear regression in animals A, B, and C.

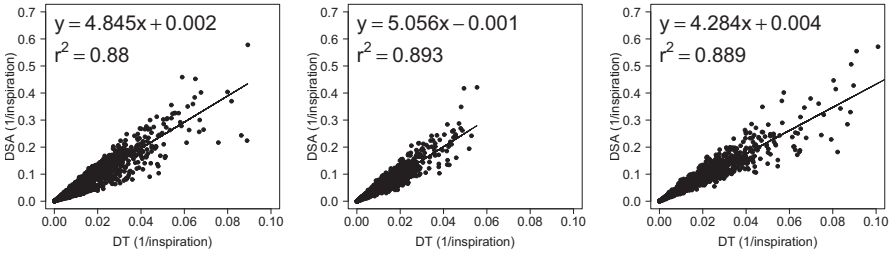


Fig. 5. (a) to (c): DSA (the absolute difference of the value between the SACJ and SAI) compared to DT (the absolute difference of the tissue volume) in animals A, B, and C.

calculated. Compared to SACJ which explicitly combines information both from the Jacobian and the intensity, SAI only uses the intensity information. SACJ has the most basic form for regional ventilation measure directly from the HU based voxel air-tissue fraction.

The two registration-based ventilation measures as well as the SV from Xe-CT are averaged and compared in predefined cubes. Averaging and comparing by 20 mm × 20 mm × 20 mm ROIs, the SACJ shows significantly higher correlation with Xe-CT based SV than the SAI in all three animals. By examining the

relationship between the DSA and DT, we see that the difference between SACJ and SAI may be due to the assumption of no tissue volume change (14) implicit in SAI. As the derivation in equations (11) and (16), to use SAI, the tissue volume change should be approximately zero. From Fig. 5, it is shown that while the tissue volume difference is usually small (less than 5%), regional ventilation measure SAI with the zero tissue volume change may introduce difference of more than 10% unit volume per inspiration comparing with the SACJ measure. For the ventilation measured over a minute, the DSA is about more than 1.7 unit volume per voxel (average $RR = 17.59$ breaths/min). Table 1 shows that the both the SACJ have significantly better correlation with SV than the SAI. This is consistent with the findings by Kabus et al. [4] who show that the Jacobian based ventilation has less error than the intensity based ventilation analysis using the segmented total lung volume as a global comparison.

The image registration algorithm used to find the transformation from EI to EE for measurement of regional ventilation produces accurate registrations by minimizing the tissue volume and vesselness measure difference between the template image and the target image. It would be interesting to compare different image registration algorithms and their effects on the registration-based ventilation measures. For example, if two registration algorithms achieve the similar landmark accuracy, the one does not preserve tissue volume change may show even larger difference in the SACJ and SAI measures than the results using TVP as described above.

In conclusion, with the same deformation field by the same image registration algorithm, a significant difference between the Jacobian based ventilation measures and the intensity based ventilation measure is found in a regional level using Xe-CT based ventilation measure SV. The ventilation measure by corrected Jacobian SACJ gives best correlation with Xe-CT based SV and the correlation is significantly higher than from the ventilation by intensity SAI indicating the ventilation measure by corrected Jacobian SACJ may be a better measure of regional lung ventilation from image registration of 4DCT images.

References

1. Guerrero, T., Sanders, K., Noyola-Martinez, J., Castillo, E., Zhang, Y., Tapia, R., Guerra, R., Borghero, Y., Komaki, R.: Quantification of regional ventilation from treatment planning CT. International Journal of Radiation Oncology*Biology*Physics **62**(3) (2005) 630 – 634
2. Reinhardt, J.M., Ding, K., Cao, K., Christensen, G.E., Hoffman, E.A., Bodas, S.V.: Registration-based estimates of local lung tissue expansion compared to xenon CT measures of specific ventilation. Medical Image Analysis **12**(6) (2008) 752 – 763 Special issue on information processing in medical imaging 2007.
3. Ding, K., Cao, K., Christensen, G.E., Hoffman, E.A., Reinhardt, J.M.: Registration-based regional lung mechanical analysis: Retrospectively reconstructed dynamic imaging versus static breath-hold image acquisition. Volume 7262., SPIE (2009) 72620D

4. Kabus, S., von Berg, J., Yamamoto, T., Opfer, R., Keall, P.J.: Lung ventilation estimation based on 4D-CT imaging. In: First International Workshop on Pulmonary Image Analysis, New York (2008) 73–81
5. Ding, K., Yin, Y., Cao, K., Christensen, G.E., Lin, C.L., Hoffman, E.A., Reinhardt, J.M.: Evaluation of lobar biomechanics during respiration using image registration. In: Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention 2009. Volume 5761. (2009) 739–746
6. Yaremko, B.P., Guerrero, T.M., Noyola-Martinez, J., Guerra, R., Lege, D.G., Nguyen, L.T., Balter, P.A., Cox, J.D., Komaki, R.: Reduction of normal lung irradiation in locally advanced non-small-cell lung cancer patients, using ventilation images for functional avoidance. International Journal of Radiation Oncology*Biology*Physics **68**(2) (2007) 562 – 571
7. Ding, K., Bayouth, J.E., Buatti, J.M., Christensen, G.E., Reinhardt, J.M.: 4dct-based measurement of changes in pulmonary function following a course of radiation therapy. Medical Physics **37**(3) (2010) 1261–1272
8. Chon, D., Simon, B.A., Beck, K.C., Shikata, H., Saba, O.I., Won, C., Hoffman, E.A.: Differences in regional wash-in and wash-out time constants for xenon-CT ventilation studies. Respiratory Physiology & Neurobiology **148**(1-2) (2005) 65 – 83
9. Yin, Y., Hoffman, E.A., Lin, C.L.: Mass preserving nonrigid registration of CT lung images using cubic B-spline. Medical Physics **36**(9) (2009) 4213–4222
10. Cao, K., Ding, K., Christensen, G.E., Reinhardt, J.M.: Tissue volume and vesselness measure preserving nonrigid registration of lung ct images. Volume 7623., SPIE (2010) 762309
11. Cao, K., Ding, K., Christensen, G.E., Raghavan, M.L., Amelon, R.E., Reinhardt, J.M.: Unifying Vascular Information in Intensity-Based Nonrigid Lung CT Registration. In Fischer, B., Dawant, B., Lorenz, C., eds.: Biomedical Image Registration, Lubeck (2010) 1–12
12. Hoffman, E.A., Ritman, E.L.: Effect of body orientation on regional lung expansion in dog and sloth. J Appl Physiol **59**(2) (1985) 481–491
13. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: MICCAI. Volume 1496. (1998) 130–137
14. Choi, Y., Lee, S.: Injectivity conditions of 2d and 3d uniform cubic b-spline functions. Graphical Models **62**(6) (2000) 411–427
15. Simon, B.A.: Non-invasive imaging of regional lung function using X-Ray computed tomography. Journal of Clinical Monitoring and Computing **16**(5) (2000) 433 – 442
16. Fuld, M.K., Easley, R.B., Saba, O.I., Chon, D., Reinhardt, J.M., Hoffman, E.A., Simon, B.A.: CT-measured regional specific volume change reflects regional ventilation in supine sheep. J Appl Physiol **104**(4) (2008) 1177–1184
17. Guo, J., Fuld, M.K., Alford, S.K., Reinhardt, J.M., Hoffman, E.A.: Pulmonary analysis software suite 9.0: Integrating quantitative measures of function with structural analyses. In: First International Workshop on Pulmonary Image Analysis, New York (2008) 283–292
18. Murphy, K., van Ginneken, B., Pluim, J., Klein, S., Staring, M.: Semi-automatic reference standard construction for quantitative evaluation of lung CT registration. In: Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention 2008. Volume 5242. (2008) 1006–1013
19. Papoulis, A., ed.: Probability and Statistics. Prentence Hall, Englewood Cliffs, NJ (1990)

Robust Matching of 3D Lung Vessel Trees

Dirk Smeets*, Pieter Bruyninckx, Johannes Keustermans,
Dirk Vandermeulen, and Paul Suetens

K.U.Leuven, Faculty of Engineering, ESAT/PSI
Medical Imaging Research Center, UZ Gasthuisberg,
Herestraat 49 bus 7003, B-3000 Leuven, Belgium

Abstract. In order to study ventilation or to extract other functional information of the lungs, intra-patient matching of scans at a different inspiration level is valuable as an examination tool. In this paper, a method for robust 3D tree matching is proposed as an independent registration method or as a guide for other, e.g. voxel-based, types of registration. For the first time, the 3D tree is represented by intrinsic matrices, reference frame independent descriptions containing the geodesic or Euclidean distance between each pair of detected bifurcations. Marginalization of point pair probabilities based on the intrinsic matrices provides soft assign correspondences between the two trees. This global correspondence model is combined with local bifurcation similarity models, based on the shape of the adjoining vessels and the local gray value distribution. As a proof of concept of this general matching approach, the method is applied for matching lung vessel trees acquired from CT images in the inhale and exhale phase of the respiratory cycle.

1 Introduction

The pulmonary ventilation, i.e. the inflow and outflow of air between the lungs and atmosphere, is the result of the movement of the diaphragm or the ribs leading to small pressure differences between the alveoli and the atmosphere. Quantification of pulmonary ventilation is a clinically important functional component in lung diagnosis. Pulmonary ventilation can be studied using several CT images in one breathing cycle (4D CT) [1]. In radiotherapy treatment, extraction of the lung deformation is important for correction of tumor motion, leading to a more accurate irradiation.

Matching, i.e. spatially aligning images (also referred as registration), inspiration and expiration scans is a challenging task because of the substantial, locally varying deformations during respiration [2]. To capture these deformations, a non-rigid registration is required, which can be image-based, surface-based and/or landmark-based. Generally, a non-rigid registration algorithm requires three components: a similarity measure, a transformation model and an optimization process.

* Corresponding author: dirk.smeets@uz.kuleuven.be

In image registration, common voxel similarity measures are sum of square differences (SSD), correlation coefficient (CC) and mutual information (MI) [3, 4]. To regularize the transformation, popular transformation models are elastic models, viscous fluid models, spline-based techniques, finite element models or optical flow techniques [3]. Voxel-similarity based techniques have been applied to lung matching in [5–7]. They have the advantage that dense and generally accurate correspondences are obtained. Disadvantages are the sensitivity to the initialization and the computational demands. Surface-based registration methods use a similarity measure that is a function of the distances between points on the two surfaces. Thin-plate splines are popular for regularization of the transformation. A combination of an voxel-similarity based and surface based registration for lung matching is presented in [8]. Generally, landmark-based non-rigid registration algorithms allow large deformations. Because of their sparsity, they are very efficient. In [9] bifurcations of lung vessels are first detected after which these landmarks are matched by comparing the 3D local shape context of each pair of bifurcations with the χ^2-distance.

In this paper, we present a robust registration method for matching vessel trees. After detecting candidate bifurcations (Sec. 2.1), correspondences are established by combining a global and a local bifurcation similarity model. For the novel global model (Sec. 2.3), the tree is represented by two reference frame independent, hence intrinsic, matrices: the Euclidean (rigid transformation invariant) and geodesic (isometric deformation invariant) distance matrix. Marginalization of point pair probabilities provides soft (not binary) correspondences between the detected bifurcations of the two trees. The local bifurcation similarity model reflects correspondences based on local intensities and is explained in Sec. 2.3. The optimization to obtain hard correspondences is done using a spectral decomposition technique (Sec. 2.4). The proof of concept of the general matching approach is shown in Sec. 3. Finally we draw some conclusions.

2 Method

The goal of the proposed matching framework is to establish correspondences between characteristic points in the image, independent of the reference frame (translation and rotation invariant) and, for characteristic points in structures that are assumed to deform isometrically, invariant for isometric deformations. Isometries are defined as distance-preserving isomorphisms between metric spaces, which generally means that structures only bend without stretching. The bifurcations, the splittings of the vessels in two or more parts, are chosen as characteristic points, although their detection is not the main goal of this paper. We then assume the vessel tree to deform isometrically. This assumption of nearly constant vessel length has been made before in a.o. [10] and [11].

2.1 Bifurcation detection

Bifurcations are locations where a blood vessel splits into two smaller vessels. A subset of the bifurcations that involves only major vessels can be detected in a robust way by analyzing the skeletonization of the segmented vessels.

As preprocessing, the lungs are segmented by keeping the largest connected component of all voxels with an intensity lower than -200 HU. A morphological closing operator includes the lung vessels to the segmentation and a subsequent erosion operation removes the ribs from the segmentation. Then, a rough segmentation of the major blood vessels within the lung is obtained by thresholding at -200 HU. Cavities in binary segmentation smaller than 10 voxels are closed. The skeletonization of the segmentation localizes the major bifurcations and is obtained through a 3D distance transformation by homotopic thinning [12]. Since bifurcations are locations where three vessels join, they are characterized as skeleton voxels having three different neighbors belonging to the skeleton. Afterwards, bifurcations are discarded that have a probability lower than $0.5/\sqrt{2\pi\sigma^2}$ according to a statistical intensity model. The vessel intensities I_k are assumed to be Gaussian distributed

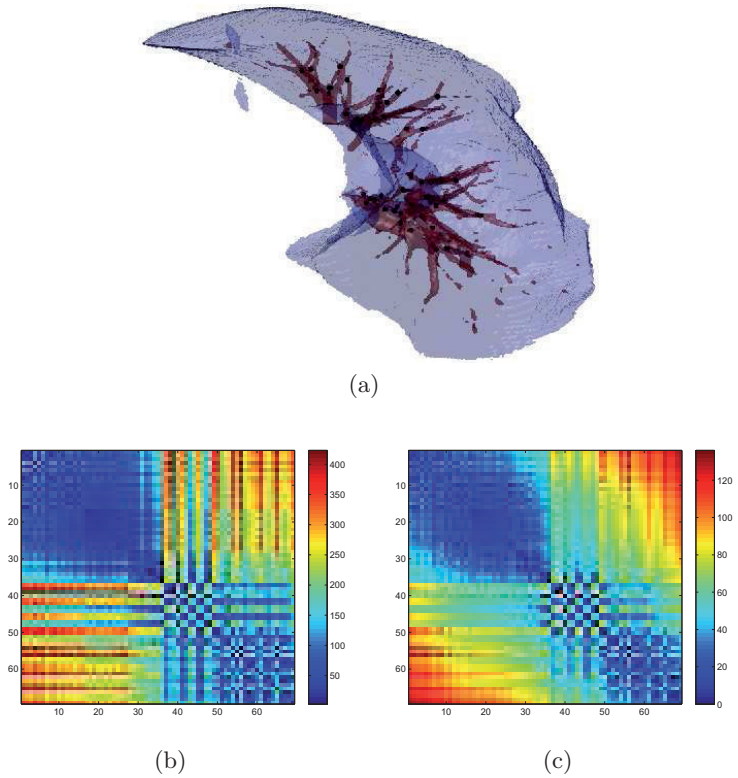
$$P(c_{vessel}|I_k) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(I_k-\mu)^2}{\sigma^2}\right), \tag{1}$$

with experimentally determined parameters ($\mu = -40$ HU and $\sigma = 65$ HU).

2.2 Global correspondence model

After the vessel bifurcations are detected, soft correspondences are established based on a global and a local correspondence model, both independent of rigid and isometric deformations of the considered vessel trees.

Intrinsic vessel tree representation. Each tree is intrinsically represented by a Euclidean distance matrix $E = [d_{\mathbb{R}^3,ij}]$ (EDM), containing Euclidean distances between each pair of bifurcations, and a geodesic distance matrix $G = [g_{ij}]$ (GDM). Each element g_{ij} corresponds to the geodesic distance between the bifurcations i and j. This distance is the distance between i and j along the vessels and is computed with the fast marching method in an image containing a soft segmented vessel tree using Eq. (1). Isometric vessel deformations, by definition, leave these geodesic distances unchanged. Therefore, the GDM is invariant to the bending of the vessels. On the other hand, the EDM is only invariant to rigid transformations (and, when normalized, invariant to scale variations) of the vessel tree. However, Euclidean distance computation is expected to be more robust against noise than geodesic distance computation since the error accumulates along the geodesic, i.e. the shortest path. Both the EDM and the GDM are symmetric and uniquely defined up to an arbitrary simultaneous permutation of their rows and columns due to the arbitrary sampling order of the bifurcations. An example of a lung vessel tree, represented by a GDM and a EDM, is shown in Fig. 1.



(a) (b) (c)

Fig. 1. A lung vessel tree with detected bifurcations (a) is represented by a geodesic (b) and a Euclidean (c) distance matrix, containing the geodesic or Euclidean distance between each pair of bifurcations.

Soft correspondences of the global model. A probabilistic framework is used to estimate the probability that a bifurcation i of one tree corresponds with a bifurcations k of the other tree. It is applied twice, once for the EDM and once for the GDM. We will explain the framework using the GDM.

Given two lung vessel trees, represented by GDMs, the probability that the pair of bifurcations (i, j) of the first tree G_1 correponds with the pair (k, l) of the second tree G_2 is assumed to be normally distributed,

$$P(C_{(i,j),(k,l)}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(g_{1,ij} - g_{2,kl})^2}{\sigma^2}\right), \tag{2}$$

with σ chosen to be 1 and $g_{1,ij}$ an element of the GDM representing the first tree. It reflects the assumption that geodesic distances between pairs of bifurcations are preserved, obeying an isometric deformation. The probability that a

bifurcation i corresponds with k is then given by

$$P(C_{i,k}) = \sum_j \sum_l P(C_{(i,j),(k,l)}) = m_{G,ik}, \tag{3}$$

from which the soft correspondence matrix, $M_G = [m_{G,ik}]$, is constructed.

The same procedure is followed to obtain an assignment matrix M_E based on the EDM as intrinsic tree representation. This matrix is expected to be more robust against noise, but varies under isometric deformations of the vessel tree, contrary to M_G.

2.3 Local correspondence model

Complementary to the global bifurcation similarity model, a local model based on intensities is implemented. In the computer vision literature a large number of local image descriptors are proposed, see for example [13].

In this paper, the n-SIFT (scale invariant feature transform) image descriptor is used [14], which summarizes a cubic voxel region centered at the feature location, in casu the bifurcation location. The cube is divided into 64 subregions, each using a 60 bin histogram to summarize the gradients of the voxels in the subregion. This results in a 3840-dimensional feature vector $\mathbf{f}$ by combining the histograms in a vector after weighting by a Gaussian centered at the feature location. The probability that a bifurcation i corresponds with k is then proportional to

$$P(C_{i,k}) \propto \exp(- \parallel \mathbf{f}_i - \mathbf{f}_k \parallel^2) = m_{L,ik}, \tag{4}$$

2.4 Spectral matching

The combined match matrix M_C is found as the pointwise product of the match matrices of the separate models, $m_{C,ik} = m_{G,ik}.m_{E,ik}.m_{L,ik}$ (i.e. product of the corresponding probabilities). To establish hard correspondences (one-to-one mapping), the algorithm proposed by Scott and Longuet-Higgins [15] is used. It applies a singular value decomposition to M_C ($M_C = U\Sigma V^T$) and computes the orthogonal matrix $M'_C = U\tilde{I}_n V^T$, with $\tilde{I}_n$ a pseudo-identity matrix. Two bifurcations i and k match if $m'_{C,ik}$ is both the greatest element in its row and the greatest in its column.

3 Experimental results

The described matching framework is evaluated using a publicly available 4D CT thorax dataset of the Léon Bérard Cancer Center & CREATIS lab (Lyon, France), called "POPI-model". It contains 10 3D volumes representing the different phases of the average breathing cycle, with an axial resolution of 0.976562 mm $\times$ 0.976562 mm and a slice thickness of 2 mm. Also, 3D landmarks, indicated by medical experts, and vector fields describing the motion of the lung, are available [16].

3.1 Matching manually indicated landmarks

First, the matching framework is evaluated using manually indicated landmarks to examine the performance independent of the bifurcation detection step. For this purpose, all 40 landmarks are used in one CT image (end-inhalation phase), while varying the number of landmarks in the other image. The result of this experiment, expressed as the percentage of correct correspondences in function of the percentage outliers of the total number of landmarks, is shown in Fig. 2 for different images in the respiratory cycle. Since not all landmarks are located in the vessels and therefore geodesic distances are not meaningful, M_G is not computed.

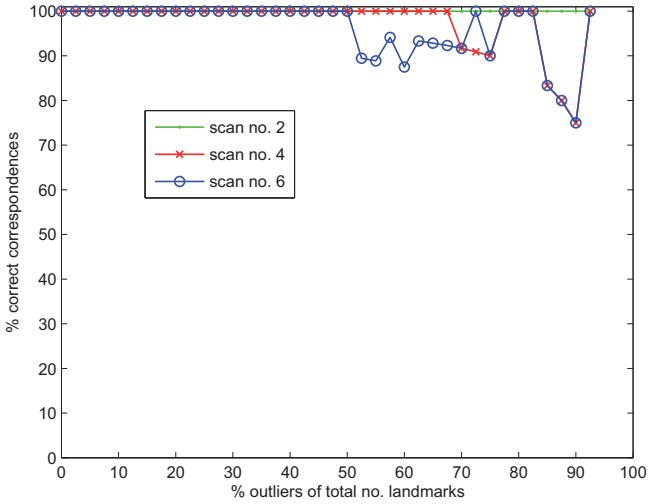


Fig. 2. Results of the framework for matching manually indicated landmarks expressed as the percentage of correct correspondences in function of the percentage outliers. The three curves correspond with different scans in the respiratory cycle, in which scan no. 1 is end-inhale and scan no. 6 end-exhale.

These results show that even for a large number of outliers good correspondences are still found. Consequently, it can be expected that not all bifurcations detected in one image must be present in the other image. Also, a small decrease in performance can be noticed when more lung deformation is present (end-exhale vs. end-inhale), probably because the Euclidean distance matrix (used in the global correspondence model) is not invariant for these deformations.

Second, the robustness against landmark positioning errors is examined. Therefore, we add uniform distributed noise in each direction to one set of landmarks and vary the maximum amplitude of the noise. Fig. 3 shows the percentage of correct correspondences in function of this maximum amplitude,

averaged over 15 runs per noise level. These results show that indication errors of more than 5 mm (± 5 voxels in x and y direction and 2.5 voxels in z direction) decrease the matching performance. It is therefore expected that the localization of bifurcations during automatic bifurcation detection must not be extremely accurate in order to still obtain good correspondences.

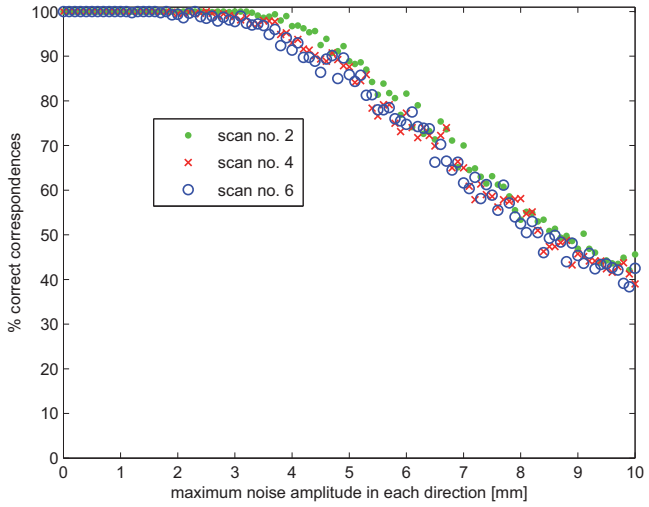


Fig. 3. Indication error dependence, expressed as the percentage of correct correspondences in function of the maximum amplitude of the added uniform noise. The three curves correspond with different scans in the respiratory cycle, in which scan no. 1 is end-inhale and scan no. 6 end-exhale.

3.2 Matching automatically detected bifurcations

Next, the framework is evaluated for the task of matching 3D vessel trees by means of automatically detected bifurcations. The result of matching the end-inhale with the end-exhale vessel tree is shown in Fig. 4, clarifying that most correspondence pairs are correct. It is also clear that simple thresholding has resulted in oversegmentation in one of the two images. This, however, did not affect the automatic matching.

The performance of matching automatically detected bifurcations is quantified using the dense deformation field that is available in the public dataset (POPI-model). This deformation field is obtained using a parametric image-based registration [17, 18, 16]. The average target registration error in the manual landmarks is 1.0 mm with a standard deviation of 0.5 mm.

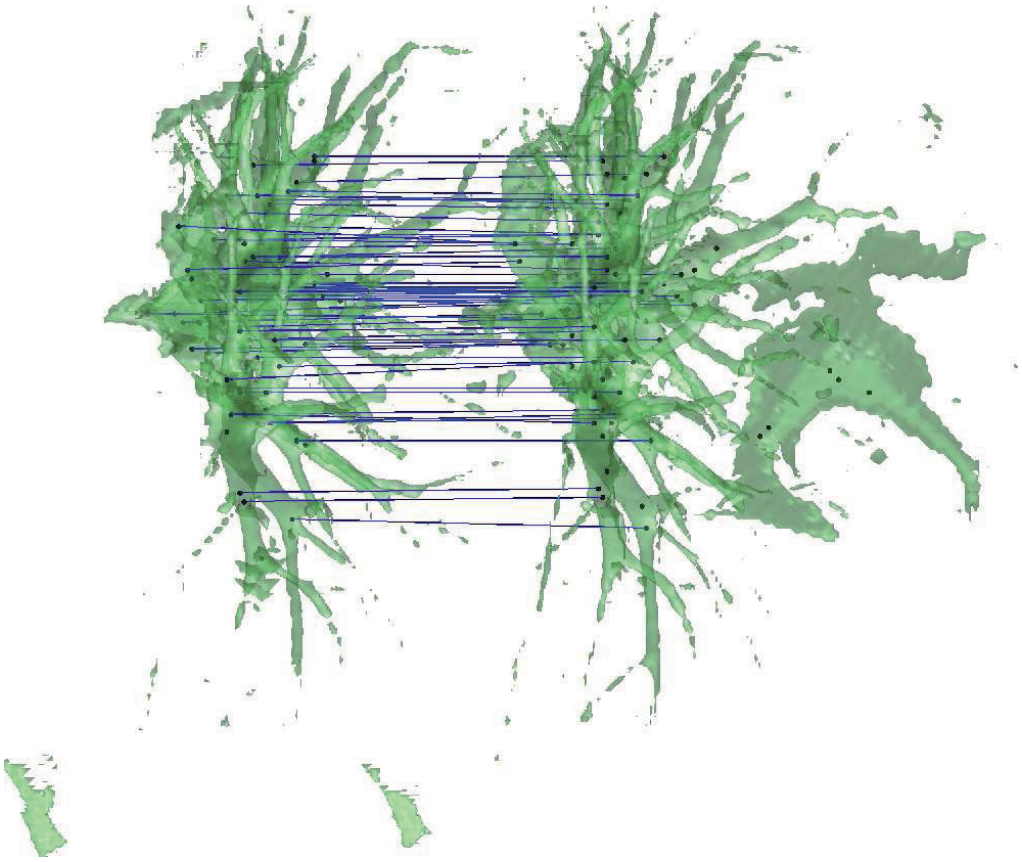


Fig. 4. Qualitative evaluation for matching 3D vessel trees by means of automatically detected bifurcations.

Fig. 5 illustrates the accuracy of the matching of the end-inhale scan with all other scans. It shows histograms of the registration error (in each direction and in absolute value) for all bifurcations. The average registration error is 2.31 mm, the median 1.84 mm and the maximum 24.0 mm. Only for 0.21% of the bifurcation matches the absolute registration error is larger than 1.5 cm, demonstrating the robustness of the matching algorithm.

4 Conclusion

A robust matching framework is proposed, combining two global and one local landmark similarity model. The results of matching manually indicated landmarks demonstrate the potential of the proposed method for matching landmarks in unimodal medical images independent of the translation and rotation between both images. A high number of outliers is allowed when the landmarks

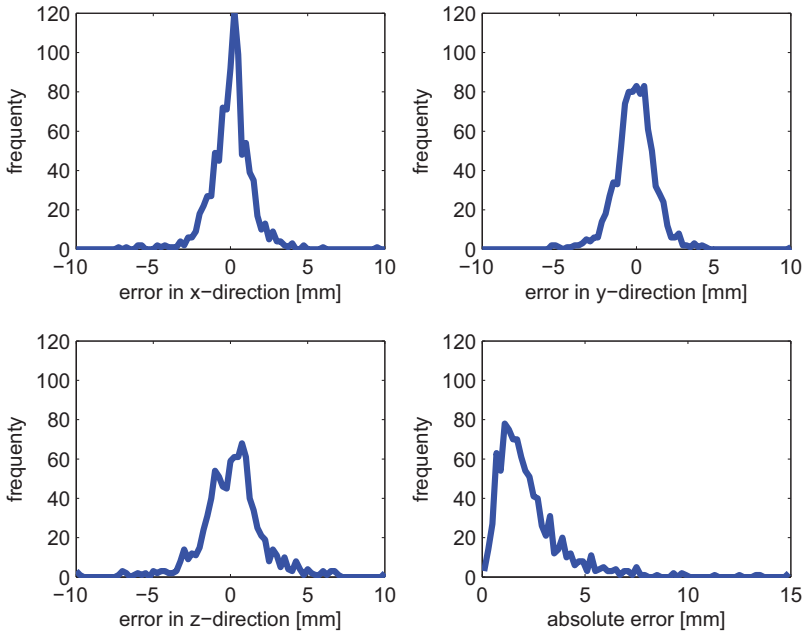


Fig. 5. Histograms of the registration error (in each direction and in absolute value) for all bifurcations give an impression of the matching accuracy.

in the image are well located. Moreover, we demonstrated that the matching framework can also be used for automatically detected landmarks, in this case lung bifurcations extracted from CT data.

As future work, we see some applications of the soft correspondences for robust initialization of an iterative non-rigid, e.g. voxel-based, registration method.

References

1. Guerrero, T., Sanders, K., Castillo, E., Zhang, Y., Bidaut, L., Pan, T., Komaki, R.: Dynamic ventilation imaging from four-dimensional computed tomography. Phys. Med. Biol. **51** (2006) 777–791
2. Sluimer, I., Schilham, A., Prokop, M., van Ginneken, B.: Computer analysis of computed tomography scans of the lung: a survey. IEEE Trans. Med. Imaging **25**(4) (2006) 385–405
3. Crum, W.R., Hartkens, T., Hill, D.L.G.: Non-rigid image registration: theory and practice. Br J Radiol **77 Spec No 2** (2004) S140–53
4. Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A.: Mutual information based registration of medical images: A survey. IEEE Trans. Med. Imaging **22**(8) (2003) 986–1004

5. Dougherty, L., Asmuth, J.C., Gefter, W.B.: Alignment of ct lung volumes with an optical flow method. Academic Radiology **10**(3) (2003) 249 – 254
6. Stancanello, J., Berna, E., Cavedon, C., Francescon, P., Loeckx, D., Cerveri, P., Ferrigno, G., Baselli, G.: Preliminary study on the use of nonrigid registration for thoraco-abdominal radiosurgery. Med Phys. **32**(12) (2005) 3777–85
7. Loeckx, D.: Automated Nonrigid Intra-Patient Image Registration Using B-Splines. PhD thesis, K.U.Leuven (2006)
8. Kaus, M., Netsch, T., Kabus, S., Pekar, V., McNutt, T., Fischer, B.: Estimation of organ motion from 4D CT for 4D radiation therapy planning of lung cancer. In: MICCAI. Volume 3217 of LNCS. (2004) 1017–1024
9. Hilsmann, A., Vik, T., Kaus, M., Franks, K., Bissonette, J.P., Purdie, T., Beziak, A., Aach, T.: Deformable 4DCT lung registration with vessel bifurcations. In: CARS. (2007)
10. Klabunde, R.E.: Determinants of resistance to flow. http://www.cvphysiology.com/Hemodynamics/H003.htm (august 2008)
11. Groher, M., Zikic, D., Navab, N.: Deformable 2D-3D registration of vascular structures in a one view scenario. IEEE Trans Med Imaging **28**(6) (2009) 847–60
12. Selle, D.: Analyse von gefässstrukturen in medizinischen schichtdatensätzen für die computergestützte operationsplanung. PhD thesis, Aachen (2000)
13. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. **27**(10) (2005) 1615–1630
14. Cheung, W., Hamarneh, G.: n-sift: n-dimensional scale invariant feature transform. Trans. Img. Proc. **18**(9) (2009) 2012–2021
15. Scott, G.L., Longuet-Higgins, H.C.: An algorithm for associating the features of two images. Proc. R. Soc. Lond. B **244** (1991) 21–26
16. Vandemeulebroucke, J., Sarrut, D., Clarysse, P.: The POPI-model, a point-validated pixel-based breathing thorax model. In: Conference on the Use of Computers in Radiation Therapy. (2007)
17. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: Application to breast mr images. IEEE Transactions on Medical Imaging **18** (1999) 712–721
18. Delhay, B., Clarysse, P., Magnin, I.E.: Locally adapted spatio-temporal deformation model for dense motion estimation in periodic cardiac image sequences. In: FIMH'07: Proceedings of the 4th international conference on Functional imaging and modeling of the heart, Berlin, Heidelberg, Springer-Verlag (2007) 393–402

Prediction-driven Respiratory Motion Atlas Formation for 4D Image-guided Radiation Therapy in Lung

Xiaoxiao Liu[1], Bradley C. Davis[2], Marc Niethammer[1], Stephen M. Pizer[1], Gig S. Mageras[3]

[1] Department of Computer Science, University of North Carolina at Chapel Hill, NC
[2] Kitware, Inc., Clifton Park, NY
[3] Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY

Abstract. Respiratory motion challenges lung radiation therapy with uncertainties of the location of important anatomical structures in the thorax. To capture the trajectory of the motion, dense image matching methods and learning-based motion prediction methods have been commonly used. However, both methods have limitations. Serious motion artifacts in treatment-guidance images, such as streak artifacts in respiration-correlated cone-beam CT, challenge the intensity-based image matching; the learning-based prediction methods require consistency between the training data for planning and the data for treatment. This paper proposes a prediction-driven motion atlas framework for motion estimation with artifact-laden images, using a Fréchet-mean-image matching scheme that is softly constrained by deformation predictions. In this framework, all the respiration phase-stamped images within a breathing cycle are diffeomorphically deformed to their Fréchet mean. The iterative optimization is driven by both intensity matching forces and the prediction forces trained from patient-specific planning images. The effectiveness of the framework is demonstrated with computational phantom and real cone-beam CT images.

1 Introduction

Positional uncertainties caused by the respiratory motion have been shown to have a large impact on radiation dose [1]. Accurate respiratory motion estimation is necessary for removing the motion-induced uncertainties. Cone-beam CT (CBCT) exploits flat panel technology to integrate the imaging system directly into the treatment accelerator, allowing acquisition of a volumetric image at each respiratory phase in the treatment position and eliminating the need for marker implantation [2–4]. A major limitation of CBCT systems is image degradation caused by respiration-induced motion, which compromises tumor and organ-at-risk localization for guiding radiation treatment of cancer. The work presented here is part of a larger project to correct respiratory motion-induced artifacts

in CBCT scans [4]. Briefly, CBCT projection images are sorted into subsets according to a respiration signal and reconstructed to obtain a set of low-quality respiration-correlated CBCT images. Application of nonrigid registration deforms each of the respiration-correclted CBCT (RC-CBCT) images to a chosen reference image in the set; combining all images yields a single high-quality CBCT image with reduced blurring and motion artifacts. However, degradation of image quality resulting from the sparse projections for each subset in the filtered back-projection reconstructions imposes serious limitations on non-rigid image registrations. The fourth dimension of the "4D" notion in this paper refers to the respiration phase during one breathing cycle.

Two different categories of methods have been used to capture the respiratory motion trajectory. One is ordinary intensity-based image matching or tracking. Various non-linear dense image registrations can be used to calculate the spatial changes of each voxel between images by matching their intensity profiles [5, 6]. However, the image registration could easily get trapped in local minima when imaging artifacts are present and thus tends to over-fit to those artifacts.

The other category is linear motion modeling with surrogate signals. Recognizing the hysteresis of respiration, various external and internal surrogate signals have been used for motion modeling and prediction in lung. The diaphragm position of the lung has been used as a navigator of the image deformation and used for motion prediction for CBCT-guided radiation therapy [7, 4]. Recently, the shape of the lung has been used as an advanced surrogate for motion prediction, in which the so-called *shape-correlated deformation statistics* (SCDS) reveals the maximum linear correlations between the shape surrogates and the image deformations [8, 9]. The SCDS is trained on a prior respiration-correlated CT (RCCT) set acquired for treatment planning purposes. The RCCT consists of a sequence of volumetric CT images over the breathing cycle.The underlying assumption of any surrogate-involved learning-based model that uses a prior image set is that the correlation between the surrogate and the underlying motion are the same for both planning/training time and treatment/test time. The assumption simplifies the complicated breathing system and justifies the estimation by directly incorporating the training information. However, the correlations between the surrogate and the spatial deformation cannot be exactly the same, especially for cancer patients who have difficulty in stabilizing their breathing over time. Besides, noise kept in the SCDS trained from few planning phases tend to result in local region prediction errors.

To fully utilize both categories of methods while avoiding their limitations, in this paper we combine the intensity information (from the treatment RC-CBCT) and the SCDS-predictions (from the planning RCCT) into a unified framework for improved motion estimation. On the one hand, the motion prediction can help regularize the intensity matching from over-fitting. On the other hand, meaningful image features can be utilized to reduce prediction errors.

A respiratory motion atlas formation method driven by a combination of prediction matching forces and image matching forces is developed in this paper. A respiratory motion atlas contains an atlas image and the dense image

deformations that transform each phase-stamped image in the breathing cycle to the atlas image. Instead of an image at an arbitrary time point, a Fréchet mean image that takes the minimum total amount of transformations to match all images is computed and used as the atlas image for increased robustness. The deformations predicted from a shape-correlated deformation statistics (SCDS) model are used as a soft constraint during the optimization. The balancing force between the intensity force and the prediction force can be adjusted via a weighting factor, selected upon the credibility of the training statistics and the quality of the treatment images.

The rest of the paper is organized as follows. Section 2 introduces the framework of the proposed prediction-driven deformation atlas formation. Specifically, section 2.1 introduces the Fréchet mean image formation method that is used to obtain the respiratory motion atlas from a 4D image sequence; section 2.2 presents the techniques that are used to apply trained SCDS to predict motion from a noisy 4D image sequence; Section 2.3 introduces the prediction-driven atlas formation by integrating the SCDS-predictions into the Fréchet mean formation framework. Experimental results are presented in Section 3.

2 Methodology

2.1 Respiratory Fréchet mean image formation

To quantify the breathing motion from a 4D image sequence, non-linear dense image registrations are often used to compute the spatial changes for each voxel in the image. The breathing motion can be quantified by the non-linear deformations that match each phase-stamped image in the breathing cycle to an atlas image. The atlas image, together with the deformations, form the respiratory motion atlas for this patient. For the respiration-correlated CT (RCCT) that is used at the planning time in IGRT, the high-contrast and good image resolution enables intensity-based registration methods to accurately capture the spatial changes over the breathing cycle.

There are several aspects to be considered in choosing a proper atlas image. First of all, due to the large anatomical variations between patients, it is more practical for the atlas image to be patient-specific. Secondly, to be used for motion prediction, the conformation of the atlas image should also be stable over time or at least stable between the planning time and the treatment or target time. The end expiration (EE) phase is commonly used as the reference phase for registration due to its relatively stable repeatability. However, computationally, a smaller total amount of deformations is preferred for better image registration accuracy and efficiency. Further, in order to carry out statistical analysis on the deformations (for SCDS calculation in Section 2.2), an "averaged" atlas will help to get tightest statistical distribution.

A Fréchet mean image has the property that it minimizes the sum of squared distances on the Riemannian manifold of diffeomorphic transformations to a group of images (see Figure 1). It represents an averaged spatial configuration

of that group [10]. Therefore, the Fréchet mean image of the breathing sequence well satisfies the aforementioned criteria and is used in this paper as the atlas image.

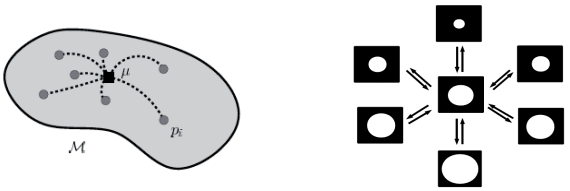


Fig. 1: Fréchet image mechanism: a) The filled circles represent individual points $\mathbf{p}^i$ on the Riemannian manifold $\mathcal{M}$. The Fréchet mean (filled square) is the point μ on the manifold that minimizes the sum of squared distances to the observations. Distances are measured along the manifold; b) Iterative Fréchet mean image construction framework illustrated on images of spheres with varying radius. The mean image in the middle minimizes the sum of squared deformation distances required to match all input images.

Given a group of phase-stamped images, geometric changes over time are represented as the action of a group of diffeomorphisms on images. Let $\mathrm{Diff}_V(\Omega)$ be the group of diffeomorphisms that are isotropic to the identity. An element $\phi : \Omega \to \Omega$ in $\mathrm{Diff}_V(\Omega)$ deforms an image I to the image $I \circ \phi$. The geodesic distance between a pair of images on the manifold is defined by diffeomorphic matching:

$$d^2(I^F, I^M) = \operatorname{argmin} \int_0^1 ||v_t||_V^2 dt + \frac{1}{\sigma^2} ||I^M \circ \phi - I^F||_{L2}^2, \tag{1}$$

subject to $\phi(x) = x + \int_0^1 v_t dt$, where $t \in [0, 1]$. The first term defines a metric on the space of diffeomorphisms that are generated by integrating velocity fields v. These diffeomorphisms are used to deform a moving image I^M (image at $t = 0$) to match a fixed image I^F (image at $t = 1$). The second term penalizes residual image dissimilarity. The parameter σ controls the relative weight of these terms.

The Fréchet mean $\hat{I}$ is the image that requires the least amount of deformation to map onto the group of input images I^i: $\hat{I} = \operatorname{argmin} \sum_{i=1}^N d(I, I^i)^2$. Combined with geodesic distance definition (1), the optimization problem can be summarized as

$$\hat{I}, \hat{\phi^i} = \underset{I, \hat{\phi^i} \in \mathbb{I} \times \mathrm{Diff}_V(\Omega)^N}{\operatorname{argmin}} \sum_{i=1}^N \left[\int_0^1 ||v_t^i||_V^2 dt + \frac{1}{\sigma^2} ||I - I^i \circ \phi^i||_{L2}^2 \right],$$

subject to $\phi_0^i = Id$, $\phi^i(x) = x + \int_0^1 v_t^i(\phi_t^i(x)) dt$.

Initialized with identity transformations, an iterative optimization updates the Fréchet mean image at each iteration, and the deformations that transform all the phases to the Fréchet mean are optimized at the same time.

2.2 SCDS-based prediction

The shape-correlated deformation statistics (SCDS) model has been shown to effectively reveal the patient-specific linear correlations between the shape surrogates and the image deformations [8, 9]. In this method, the shape of the lung is used as an internal surrogate signal to navigate the dense image deformation by linear regression. The shape models are extracted using lung surface segmentation from the images followed by application of an entropy-based particle system [11] to obtain a group-wise surface correspondence over the phases within the breathing cycle. The SCDS model trained from the planning images is used to predict the motion of the images via extracted shape surrogates.

In previous papers [8, 9], the SCDS model used the EE phase as the image atlas. In this paper, we improve the tightness of the SCDS by using the Fréchet mean as the image atlas. Also we use deformable segmentation techniques to apply the SCDS for motion prediction to CBCT images with image artifacts. The resulting predicted deformations are going to be used as soft constraints in the overall prediction-driven atlas framework described in Section 2.3.

Probabilistic deformation segmentation To apply this method to CBCT images, robustly extracting the lung boundaries against the streak intensity artifacts is important. We have developed a posterior probability optimization scheme to calculate the models that fit into the target images while staying in the trained shape space. The optimization is described in

$$\mathbf{log}\ p(\mathbf{q}^i|J^i) = \underset{q^i}{\text{argmax}}[\mathbf{log}\ p(J^i|\mathbf{q}^i) + \mathbf{log}\ p(\mathbf{q}^i)], \tag{2}$$

where $\mathbf{q}^i$ is the deformable surface mesh (to distinguish the shapes $\mathbf{p}^i$ in training) segmented from the CBCT image J^i (to distinguish the training image I^i). The image match term or the likelihood term is the summation of a second-order gradient magnitude measured on the surface of the model, indicating how well the model fits to the boundaries. The prior term is measured by Mahalanobis distance of the model from the training mean in the trained shape space.

SCDS-prediction for 4D CBCT In the motion atlas, the deformation of the whole dense deformation field is represented by the displacement vector field (DVF) u as the result of the diffeomorphic transformation from each phase image to the atlas image. To distinguish the planning data and the treatment data, in this paper ϕ^i denotes the deformations calculated from RCCT images I^i and φ^i denotes the deformations calculated from RC-CBCT images J^i. It follows that

$\mathbf{u}_\varphi$ denotes the DVFs from RCCTs and $\mathbf{u}_\varphi$ denotes the DVFs from RC-CBCT. In summary, three major steps are carried out to estimate the image deformations in 4D CBCT.

1. The DVFs $\mathbf{u}_\phi^i$ of RCCTs are obtained by the intensity-based Fréchet mean formation method introduced in the last section. The surface models of the lung are extracted from each CT phase images. The linear correlation $\mathbf{C}$ that maps a shape surrogate $\mathbf{p}^i$ to its corresponding image deformation $\mathbf{u}_\phi^i$ is calculated such that $\mathbf{u}_\phi = \mathbf{C} \cdot \mathbf{p} + \epsilon$, where ϵ is the regression error.
2. The lung shape $\mathbf{q}^i$ is segmented from the CBCT images J^i using the posterior probability optimization scheme.
3. Deformations of each time-point of the CBCT sequence $\mathbf{u}_\varphi^i$ are calculated by $\mathbf{u}_\varphi^i = \mathbf{C} \cdot \mathbf{q}^i$.

After the motion prediction, an atlas image can be calculated by averaging all the intensity images after warping them using the predicted deformations.

2.3 Prediction-driven respiratory motion atlas formation

The SCDS-prediction method is a learning-based approach that is resistant to image artifacts. However, as mentioned in the introduction, its accuracy also depends on the *correlation consistency* condition of the motion between the training data and the target data. For lung cancer patients who themselves have difficulties to strictly regularize their breathing patterns or as a result of anatomical changes such as tumor growth, the correlation between the surrogate lung and the overall image deformation can not be exactly the same. Besides, the linear correlation regression results contain statistics errors due to the high dimensional low sample size (HDLSS) problem.

To increase the prediction robustness of the method, image intensity features can be used to adapt to the motion variations between the training and the testing data. Despite the CBCT artifacts, there are still many intensity features (besides the lung boundaries) useful for guiding the image registration, such as the bony rib cage, the bronchial structures, and the tumor region itself.

A prediction-driven deformation atlas formation method, optimized by the combination of prediction constraints and image matching forces is presented here to improve the motion estimation accuracy and robustness. The deformation predictions are used as soft constraints in the iterative Fréchet mean image optimization, as follows:

$$\hat{J}, \hat{\varphi}^i = \underset{J,\hat{\varphi^i}}{\operatorname{argmin}} \sum_1^N \left[\int_0^1 ||v_t^i||_V^2 dt + \frac{1}{\sigma_1^2} ||J - J^i \circ \varphi^i||_{L2}^2 + \frac{1}{\sigma_2^2} d_R^2(\varphi^i, \phi^i(q^i)) \right],$$

subject to $\varphi^i = x + \int_0^1 v_t^i dt$, where J^i denotes the CBCT image at phase i, $\hat{J}$ is the atlas image and $\mathbf{q}^i$ is the lung shape segmented from J^i using the deformable segmentation method introduced in Section 2.2. The distance between the varying deformation φ^i and the prediction $\phi(q^i)$ is measured via the Riemannian

manifold metric d_R, which is defined as the minimum of the integral over all piecewise smooth curves that connect ψ_1 and ψ_2. This distance can be alternatively computed by $d_R(\psi_1 \circ \psi_2^{-1}, id)$, where id is the identity transformation.

To simplify the computation and to directly use the resulting deformation represented by DVFs (deformations from the atlas image) computed from the SCDS-prediction, an Euclidean approximation of the square Riemannian distance is given by $d_R^2(\varphi^i, \phi(q^i)) \approx ||\mathbf{u}_{\varphi^i} - \mathbf{u}_{\phi(\mathbf{q}^i)}||_{L^2}^2$, where the SCDS-predicted deformation $\mathbf{u}_{\phi(\mathbf{q}^i)}$ is computed by linear mapping $\mathbf{u}_{\phi(\mathbf{q}^i)} = \mathbf{C} \cdot \mathbf{q}^i$. The Euclidean deformation space is only an approximation of the Riemannian deformation manifold. However, when deformations are not very large, the Euclidean space can be thought of as the tangent plane of the Riemannian manifold at the Fréchet mean and thus the linear approximation is sufficient.

The balancing force between the prediction and the noisy intensity profile can be adjusted via the weighting factors σ_1 and σ_2, selected upon the credibility of the training statistics and the quality of the treatment images. To have equal influences from both the intensity and the prediction, the weighting factors should make the two energy terms have the same order of magnitude. The energy term of the prediction is treated as an extra feature channel.

Computationally, this extra channel itself is a three-dimensional-vector channel and takes three times the storage as the image intensity. In comparison to the intensity-based atlas method, the prediction-driven atlas method takes more time to compute the gradient for the extra prediction constraints. On the other hand, due to the prediction constraints fewer iterations are typically required for convergence.

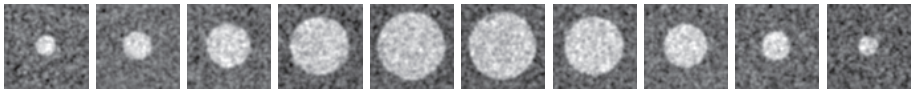
3 Experimental results

3.1 Breathing spheres

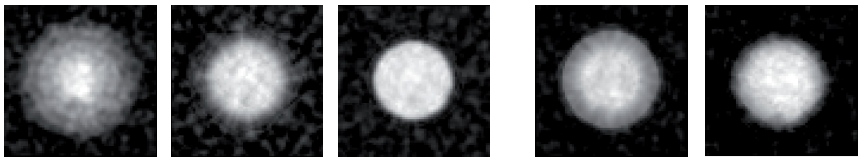
We started with some simulation data to test the prediction-driven atlas formation method. A sequence of sphere images with varying radii were designed to mimic the breathing scenario. The radii follow a sinusoidal curve to simulate the breathing pattern of a volume enlarging process followed by a volume shrinking. The surface points on the spheres (not shown) are used as the shape surrogate to carry out the SCDS motion prediction. With the same dateset, Gaussian noise is added to create the test data, as shown in Figure 2a.

The underlying correlation between the shape surrogate, the surface point set sampling on the spheres with group-wise correspondence, and the image deformation are the same for the training and the testing data, since the only difference between the two datasets is the added Gaussian noise. It is shown that the intensity-based atlas formation method (Figure 2c) tends to over-fit the noise, while the SCDS-prediction (Figure 2d) that is determined by the training SCDS shows resistance to the noise and gives better results.

To simulate variations in the correlations between the training set and the test set, the correlation coefficients between the shape surrogate (surface point

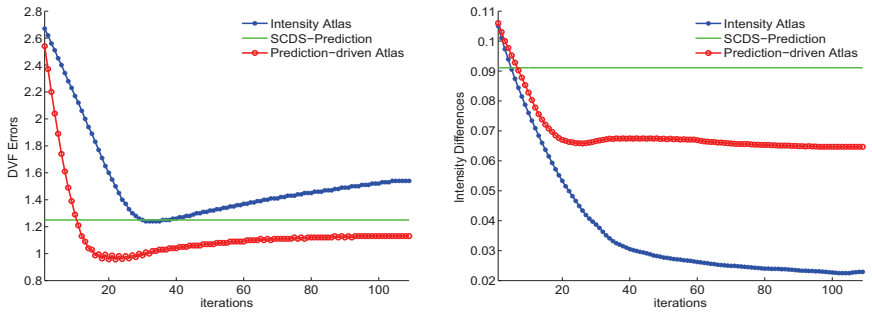


(a) noisy input images



(b) Euclidean (c) Fréchet (d) SCDS (e) SCDS (f) pred-driven

Fig. 2: **Noisy breathing spheres test**: a) The noisy test image sequence. b) The Euclidean mean of the noisy test sequence. c) The Fréchet mean image of the test data using the intensity-based atlas formation method. d) The resulting atlas image of SCDS motion prediction. **Test with correlation perturbations**: e) The resulting atlas image of SCDS motion prediction with correlation perturbations; f) The prediction-driven atlas image with correlation perturbations.



(a) DVF (image deformation) error (b) Intensity differences

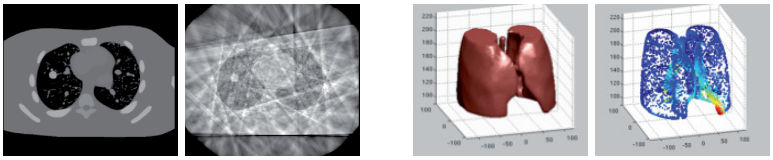
Fig. 3: Noisy spheres with correlation perturbations: a) Average displacement vector field error per voxel at each iteration. b) Average image intensity difference per voxel at each iteration.

set of the sphere) and the image deformations were perturbed randomly by 0.15 at maximum. The SCDS prediction is influenced by the artificial perturbation and produces errors mostly visible at the edge of the atlas sphere (Figure 2e). On the other hand, it is shown that the prediction-driven deformation is able to balance between the intensity force and the prediction force thus getting the best result (Figure 2f). Errors and intensity energies are shown at each iteration step in Figure 3 for a detailed investigation. The intensity-based matching over-fits the noise: intensity differences diminish with iteration but DVF error increases.

Since the SCDS-prediction is directly computed without iterative optimizations, it is shown as the constant value line for comparison. The best estimation result (i.e., lowest DVF error) is achieved by the prediction-driven atlas formation method.

3.2 NCAT data

4D Nurbs-based Cardiac-Torso (NCAT) phantom thorax RCCTs were produced [12] at 10 phases sampled in one breathing cycle. A corresponding RC-CBCT sequence was simulated from the NCAT RCCTs using the protocol of a gantry-mounted KV on-board imaging system [3] that is used in patient radiation therapy guidance. An example image pair is given in Figure 4. Note the strong artifacts in the phase-sorted CBCT caused by the sparse and unevenly spaced projections. The image grid for each image is $512 \times 512 \times 100$, with voxel size of 0.742 mm $\times$ 0.742 mm $\times$ 1.52 mm. Figure 4 shows an example of the NCAT CT and CBCT image and the surface lung shape model.

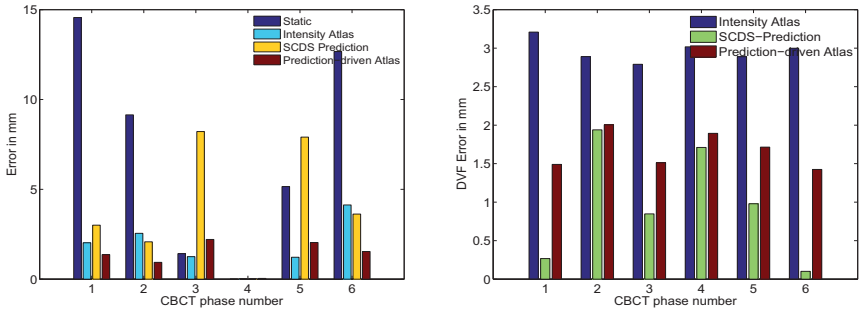


(a) NCAT RC-CBCT (b) Lung shape model

Fig. 4: a) An axial slice of a NCAT RCCT image at EE phase and its corresponding RC-CBCT image. b) The surface mesh representation of the shape of lungs extracted from NCAT RCCTs. The color shows the magnitude of the spatial variation of each point during the breathing cycle.

To test the prediction-driven atlas formation method, inconsistency of the breathing correlation patterns between the training data and the test data were simulated by adding random perturbations into the SCDS-predicted DVFs.The NCAT data set in Figure 5 has a maximum of 2.0 cm diaphragm motion and 0.5 cm anterior-posterior motion. The intensity-based atlas formation results, the SCDS motion prediction results and the prediction-driven atlas formation results are compared in terms of the center of gravity (COG) location errors of the tumor region (Figure 5a). The intensity-based atlas performs better than the SCDS-prediction in terms of the tumor region estimation. The reason is that the tumor region after the CBCT reconstruction has a quite strong contrast respect to its surrounding tissue despite the global streak artifacts. The overall DVF errors of the three methods are compared in Figure 5b. The SCDS-prediction performs better than the intensity-based atlas method. The prediction-driven atlas method shows a compromised overall DVF estimation between the other

two as a result of the combination. These measurements are confirmed from the visual comparison of the atlas images in Figure 6. The prediction-driven atlas (Figure 6c) has less global signal-to-noise ratio (SNR) compared to the intensity atlas (Figure 6a) and has a sharper tumor boundary than the SCDS-prediction atlas (Figure 6b).



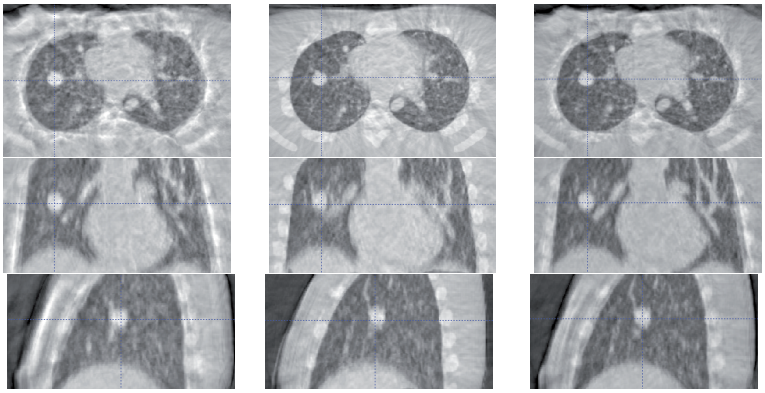
(a) Tumor COG error (b) DVF (image deformation) error

Fig. 5: a) Tumor COG estimation errors, with the static measurement indicating the mobility of the tumor. The fourth phase is used as the base phase to propagate the tumor contour to other phases. b) DVF estimation error per voxel. The ground truth DVF is obtained by linear interpolation from the RCCT DVFs.

3.3 Patient data

RCCT data sets are provided by a 8-slice scanner, acquiring repeat CT images for a complete respiratory cycle at each couch position while recording patient respiration. The CT images are retrospectively sorted to produce a series of 3D images at 10 respiratory phases. The RC-CBCT scans of five-minute duration are acquired using a slowing-rotating gantry-mounted KV on-board imaging system. The scan produces 3D images at 6 respiratory time points. The image grid for each image is $196 \times 196 \times 100$, with voxel size of 1.52 mm $\times$ 1.52 mm $\times$ 1.52 mm.

Motion estimation results are evaluated on a mock tumor region as shown in Figure 7. Manual segmentations are provided for each CBCT image for error measurements. The manual tumor contour of the fourth phase image (the end-expiration phase) is propagated to all the other phases. The three methods are compared in terms of the COG location errors, see Figure 7. The average COG errors of the 5 phases for the three approaches are 3.5 mm, 2.3 mm and 1.7 mm respectively. The prediction-driven atlas method outperforms the other two approaches in this patient.



(a) Intensity-based (b) SCDS-prediction (c) Prediction-driven

Fig. 6: Atlas image comparison of NCAT data set tests. The axial (the first row), coronal (the second row) and sagittal (the third row) slices are compared. a) The Fréchet mean atlas image of the CBCT sequence. b) The atlas image from the SCDS-prediction results. c) The atlas image from the prediction-driven atlas formation method.

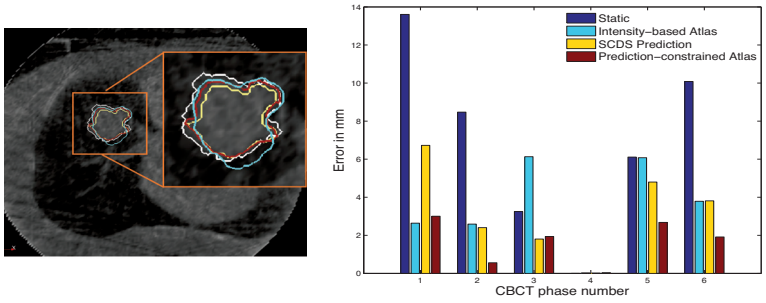


Fig. 7: Left: Axial slices of tumor contours at the fifth phase from the three methods, with the same color legend used in the bar plot on the right. The manual segmentation is shown in white. Right: The comparison of tumor COG errors.

4 Conclusion

The prediction-driven atlas formation framework has the advantage of utilizing high-contrast intensity information from CBCT while being constrained by the shape-correlated prediction results. The overall image deformation result is a compromise between the pure-intensity-based atlas formation method and the SCDS-prediction results, and the structures that have relatively higher contrast contribute to a more accurate local motion estimations. With our preliminary studies on both simulated phantom data and cone-beam CT data, the

prediction-driven atlas method is shown to be more robust for modeling and estimating sophisticated respiratory motion in lung than both the intensity-based atlas method and the learning-based SCDS prediction. More comprehensive validations on patient data are needed to quantify the robustness of the method.

Acknowledgments This project was supported by Award Number R01CA126993 from the National Cancer Institute.

References

1. Minohara, S., Kanai, T., Endo, M., Noda, K., Kanazawa, M.: Respiratory gated irradiation system for heavy-ion radiotherapy. Int J Radiat Oncol Biol Phys **47**(4) (Jul 2000) 1097–103
2. Jaffray, D., Siewerdsen, J., Wong, J., Martinez, A.: Flat-panel cone-beam computed tomography for image-guided radiation therapy. Int J Radiat Oncol Biol Phys **53** (2002) 1337– 1349
3. Kriminski, S.A., Lovelock, D.M., Seshan, V.E., Ali, I., Munro, P., Amols, H.I., Fuks, Z., Bilsky, M., Yamada, Y.: Comparison of kilovoltage cone-beam computed tomography with megavoltage projection pairs for paraspinal radiosurgery patient alignment and position verification. Int J Radiat Oncol Biol Phys **71**(5) (2008) 1572–80
4. Zhang, Q., Hu, Y.C., Liu, F., Goodman, K., Rosenzweig, K.E., Mageras, G.S.: Correction of motion artifacts in cone-beam ct using a patient-specific respiratory motion model. Medical physics **37**(6) (2010) 2901–2910
5. Reinhardt, J., Christensen, G., Hoffman, E., Ding, K., Cao, K.: Registration-Derived estimates of local lung expansion as surrogates for regional ventilation. Information Processing in Medical Imaging (2007) 763–774
6. Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Schulz, B., Handels, H.: Generation of a mean motion model of the lung using 4D CT image data. Eurographics Workshop on Visual Computing for Biomedicine (2008) 69–76
7. Zhang, Q., Pevsner, A., Hertanto, A., Hu, Y., Rosenzweig, K., Ling, C., Mageras, G.: A patient-specific respiratory model of anatomical motion for radiation treatment planning. Medical Physics **34** (2007) 4772–4781
8. Liu, X., Saboo, R.R., Pizer, S.M., Mageras, G.S.: A shape-navigated image deformation model for 4D lung respiratory motion estimation. IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI) (2009) 875–878
9. Liu, X., Oguz, I., Pizer, S.M., Mageras, G.S.: Shape-correlated deformation statistics for respiratory motion prediction in 4D lung. SPIE Medical Imaging **7625** (2010)
10. Davis, B.C., Fletcher, P.T., Bullitt, E., Joshi, S.C.: Population shape regression from random design data. IEEE International Conference of Computer Vision (ICCV'07) (2007) 1–7
11. Oguz, I., Cates, J., Fletcher, T., Whitaker, R., Cool, D., Aylward, S., Styner, M.: Cortical correspondence using entropy-based particle systems and local features. IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI) (2008) 1637–1640
12. Segars, W., Lalush, D., Tsui, B.: Modeling respiratory mechanics in the MCAT and spline-based MCAT phantoms. IEEE Transactions on Nuclear Science **48**(1) (2001) 89–97

A Clinical Feasibility Study on Respiratory Sorted Megavoltage Cone Beam CT

Mingqing Chen, R. Alfredo Siochi

Department of Electrical and Computer Engineering
Department of Radiation Oncology
The University of Iowa

Abstract. In order to improve gated radiotherapy for lung cancer, we explore the use of projection data from localization Mega-voltage cone-beam computed tomography (MVCBCT) scans to reconstruct respiratory correlated images. The position of the ipsilateral hemi-diaphragm apex (IHDA), automatically detected in raw data, is used for projection sorting. Two algorithms, the Feldkamp, Davis and Kress and the simultaneous algebraic reconstruction technique are implemented for incomplete projection reconstruction. The tumor and the IHDA are manually identified on full exhale and full inhale phases of MVCBCT images. 16 MVCBCT scans and 2 planning CT scans of a non-small cell lung cancer patient are analyzed. Results show that it is feasible to use sorted projection reconstruction to determine relationship between the tumor and the diaphragm motion and observe tumor volume change due to effect of radiotherapy.

Keywords: diaphragm, respiratory motion, cone beam CT, reconstruction

1 Introduction

Respiratory motion is a major concern in radiotherapy for lung tumors. In our clinic, a strain gauge system is used during the planning CT scan to record the relative respiratory phase and create 4DCT images that allow us to determine the phases to be used for treatment. The strain gauge system is considered an external surrogate of tumor motion and provides respiratory gating signals to turn the treatment beam on or off. However, tumor motion may vary from one fraction to the next, and may be different from the motion at the time of the 4DCT. Because this variation can cause tumor underdose or overexpose normal tissue, the strain gauge should be calibrated to tumor motion.

Mega voltage cone beam CT (MVCBCT) is a promising technique for daily imaging of patients [20, 22]. It only provides one static image reconstructed from all the raw images projected from different gantry angles. For static organs, inter-fractional changes are observable. But intra-fractional changes such as tumor and diaphragm motion are averaged over all projections, which makes it difficult to observe the tumor size and shape before each treatment fraction. Utilizing

MVCBCT projection images is a promising solution. Reitz et al. [23] projected 3D tumor volume onto projections to verify the tumor position prior to radiotherapy. Siochi developed a semi-automatic method to determine respiratory-like motion of a tungsten pin on projection images with minimal manual input [25]. Unlike these objects with clearly defined edges, the tumor boundary is not visible in many views. Methods that work well with fluoroscopic images [6,18,28,29] may not be suitable for the application of MVCBCT projection images. Alternatively, the diaphragm has sudden changes in intensity in most views and its motion is a more accurate internal surrogate for tumor motion [4, 27]. The strain gauge could be calibrated to tumor motion indirectly using the diaphragm. We previously developed a dynamic Hough transform (DHT) method based on Lappas' work [15,16]. It was able to detect the ipsilateral hemi-diaphragm apex (IHDA) positions in all MVCBCT projection views within a minute [5] (The IHDA is the most superior point in all projection views of one hemi-diaphragm). Compared to the method introduced by Sonke et al. [26], which projects the diaphragm in 2D CBCT projection images to 1D to extract relative respiratory phase, our previous method detects the actual 3D anatomical position of the IHDA. The algorithm is not robust in a few lateral views when both hemi-diaphragms are overlapping each other (figure 3a). Manual tuning of DHT parameters may correct the ambiguous recognition. It is also required to establish the relationship between diaphragm motion and tumor motion.

In this work, we further investigated the feasibility of identifying the tumor in MVCBCT images reconstructed from respiratory phase sorted data. First, a more robust method for IHDA detection especially for problems in lateral views is presented. Secondly, projection images are sorted to full inhale (FI) and full exhale (FE) phases based on the IHDA position. Volumetric 3D images are reconstructed from sorted projections. The tumor boundary and the IHDA position in each 3D image set are manually identified. Tumor size, tumor centroid and the diaphragm to tumor motion ratio (DTMR) are quantified.

2 Materials and Methods

2.1 MVCBCT Data

MVCBCT projection data are acquired as the gantry rotates. The standard protocol for rotation is 200 degrees from -90° to 110°, with one projection per degree. The images are acquired with an amorphous Si flat panel electronic portal imaging device (EPID) with a detector pixel size of $0.4 \times 0.4mm^2$. It has a source to axis distance (SAD) of 100 cm, a source to imager distance (SID) of 145 cm, and an imaging field of view (FOV) of $27.4 \times 27.4cm^2$. The EPID has a resolution of 1024×1024, with a pixel spacing of $0.4mm$. MVCBCT raw data are taken from one NSCLC patient who was treated with five fractions each week for almost 8 weeks. The imaging protocol used a 10MU exposure, with an acquisition time of 55s. MVCBCT scans performed on each Thursday and Friday and two planning CTs are studied. The first planning CT was taken six days

before the first treatment day (a Thursday) while the second one was taken on the Thursday of the 4th week.

2.2 IHDA Detection in Projected Images

The Hough transform (HT) is a robust object detection technique specified for prior known parametric shapes [1,7,13]. It is based on global evidence accumulation in shape parameter space. Over the years, different variations of HT for motion detection in image sequences were developed, including global optimization based approaches such as velocity HT [21] and dynamic HT (DHT) [15,16], and tracking based approaches [11,12,19].

Our previous method for IHDA detection can be found in [5], which employs a Hough transformation for all ROIs of the projection images. The hemidiaphragm in MVCBCT projections is modeled by two parabolas connected at their common vertex (x_0, y_0), with focal lengths of $4a_1$ and $4a_2$, with representation as:

$$\begin{aligned} f(x) &= a_1(x - x_0)^2 + y_0, \text{ if } x \leqslant x_0 \\ &= a_2(x - x_0)^2 + y_0, \text{ if } x > x_0 \end{aligned} \tag{1}$$

A five dimensional (5D) Hough accumulator array $I(x_0, y_0, a_1, a_2, t)$ is built with each cell representing a two-parabola contour, where t is the index of projection image. It is then reduced into the 3D parameter space of $I(x_0, y_0, t)$ by maximal intensity projection (MIP) for computation efficiency, with two associated arrays $a_1(x_0, y_0, t)$ and $a_2(x_0, y_0, t)$ recording the corresponding a_1 and a_2.The optimal motion trajectory of the IHDA over all images is detected by dynamic programming.

The method presented here employs a frame-by-frame point tracking technique in Hough arrays. The program starts to track through an entire MVCBCT raw image sequence from four frames with the IHDA identified by the user. The IHDA position is then determined frame by frame iteratively. In each iteration, as the IHDA position of the current frame is determined, a joint probability function is evaluated for every local maxima of Hough space in the next frame. The maxima with the largest probability is picked. The joint probability function is:

$$\begin{aligned} f(x_0^{t+1}, y_0^{t+1}, a_1^{t+1}, a_2^{t+1}) &= f_x(x_0^{t+1}|x_0^t) f_y(y_0^{t+1}|y_0^t) \\ & f_a(a_1^{t+1}|a_1^t) f_a(a_2^{t+1}|a_2^t) f_{cor}(x_0^{t+1}, y_0^{t+1}|x_0^t, y_0^t) \end{aligned} \tag{2}$$

Where superscripts t and $t+1$ represent the current and next frame, respectively. The joint probability function has five components, each measuring the conditional probability of parameters in the next frame using the current frame as its input. Functions f_x, f_y, f_a measure probability based on parabola

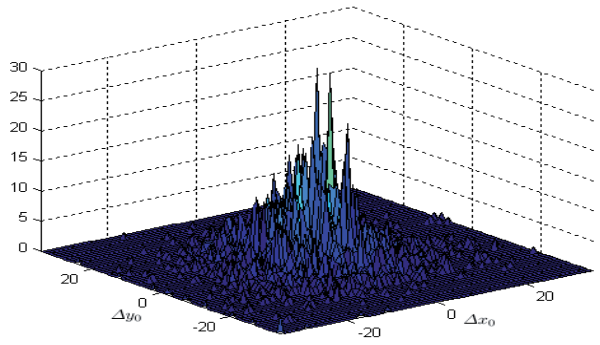


Fig. 1: Distribution of IHDA positions in 21 MVCBCT projection sets represented by Δx_0 and Δy_0

parameters x_0, y_0, a_1 and a_2 respectively. f_{cor} determines the probability of a local maxima candidate in frame $t+1$ based on the normalized cross correlation (NCC) [3]. The NCC is computed between two image patches in Hough space, one of which consists of the neighboring pixels of the local maxima candidate in frame $t+1$, while the other is the group of neighboring pixels of the IHDA in frame t.

The conditional probability of the five components are assumed to be Gaussian distributions. The mean values for functions f_x, f_y, f_a are the parabola parameters of the current frame and for f_{cor} , it is 1.0, which represents two identical images.

$$f_x(x_0^{t+1}, x_0^t) = k_x e^{-\frac{(x_0^{t+1}-x_0^t)^2}{2\sigma_x^2}} \tag{3}$$

$$f_y(y_0^{t+1}, y_0^t) = k_y e^{-\frac{(y_0^{t+1}-y_0^t)^2}{2\sigma_y^2}} \tag{4}$$

$$f_a(a^{t+1}, a^t) = k_a e^{-\frac{(a^{t+1}-a^t)^2}{2\sigma_a^2}} \tag{5}$$

$$f_{cor} = k_{cor} e^{-\frac{(NCC(x_0^{t+1}, y_0^{t+1}, x_0^t, y_0^t)-1)^2}{2\sigma_{cor}^2}} \tag{6}$$

The Gaussian window σ is derived from a statistical analysis of the manual IHDA identification results on 40 MVCBCT scans from 6 patients. Figure 1 shows the distribution of IHDA displacements, which is similar to a 2D Gaussian distribution.

The IHDA positions in 2D projections are converted to the IHDA positions in the patient coordinate system based on an interpolated ray tracing method [25]. The IHDA position in the patient is used for respiratory sorted reconstruction.

2.3 Respiratory Correlated MVCBCT

Current MVCBCT imaging uses all the projections to reconstruct a static image. By selecting projections belonging to the same respiratory phase, it is possible to reconstruct MVCBCT images of moving tumors with multiple phases. In this study the FE and FI phases are reconstructed for each MVCBCT scan. The superior-inferior position of the IHDA, provided by the previous step, is scaled from 0 to 100 to represent normalized respiratory phases. The ideal phase bin for the FE phase should be 0 to $phasewindow$, and $(100 - phasewindow)$ to 100 for the FI phase. But for real respiratory traces, especially for irregular breathing, inclusion of 0 or 100 will only encompass a small amount of projections. Both factors should be considered for proper phase window and level values. An exhaustive search strategy is applied to find these values. For the FI phase, it aims to find a phase window and level that maximize the averaged respiratory amplitude of projections within the phase window. For the FE phase, the same procedure is used for minimization of the amplitude.

The default CBCT reconstruction algorithm by Feldkamp, Davis, and Kress (FDK) [8], is fast and produces good images, but an inadequate number of projections results in strong aliasing artifacts. Algorithms derived from Algebraic reconstruction techniques (ART) [10] are more robust for incomplete projections but require much more computation time due to iterative re-projection and back-projection [14]. Hardware acceleration based on graphic cards made it feasible to accomplish cone beam ART reconstruction within a clinical time limit [24]. In this work, both methods are investigated to evaluate the feasibility of tumor contouring without considering reconstruction speed. MVCBCT projection matrices, which were derived during MVCBCT geometric calibration [20], are used for voxel-driven projection computations in FDK and ray-driven models in ART [9]. A Bessel-Kraiser filter [17] is used as the interpolation kernel for ART, since it has less aliasing artifacts than bilinear or Gaussian kernels for ray-driven back-projection.

The IHDA position and tumor boundary in both FE and FI images are identified by a researcher without prior knowledge of the patient using the Pinnacle treatment planning software. The volume and centroid of the tumor and the DTMR are then calculated. These parameters are also quantified on the planning CT for comparison. Figure 4 shows one example of manual contouring of the tumor using Pinnacle.

3 Results and Discussions

3.1 IHDA Detection

For each MVCBCT scan, automatic IHDA detection is compared to the independent manual identification in the raw data, which is considered as ground truth. Root mean square (RMS) errors are calculated and compared with those of the DHT method on 19 images from 6 patients. In figure 2, DHT has a large

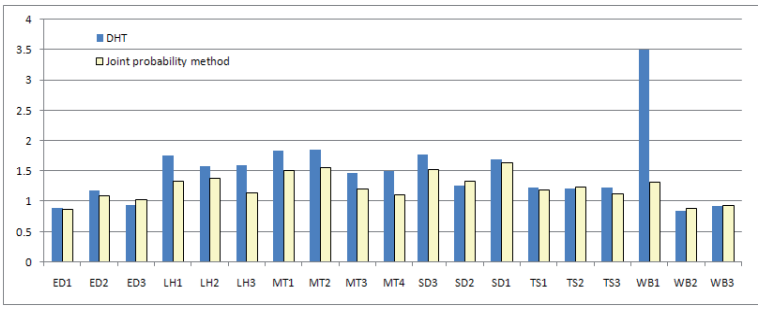
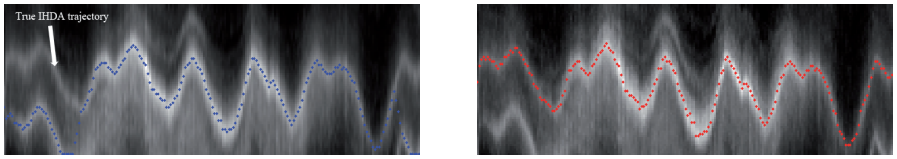


Fig. 2: Comparison between DHT and new point tracking method on RMS error (in mm) in 19 MVCBCT scans.



(a) (b)

Fig. 3: 2D representation of Hough accumulator array. Each column represents a 1D Hough image of a projection frame. One example of IHDA trajectory detected by (a) DHT with inter-column constraint 5; (b) The new point tracking method.

error for one image due to the erroneous detection of the contra-lateral hemi-diaphragm (CHD) in a number of frames. The new method corrects this problem and reduces RMS errors. Of the 19 images, 14 have improved accuracy due to reduced cases of CHD detection and improved accuracy in IHDA localization. For the other 5 images, both of the methods are accurate enough and the difference in the RMS is small, and can be considered a statistical fluctuation. Overall, the new method presented in this work achieves a mean RMS error of $1.228mm$ with a standard deviation of $0.220mm$, compared to the DHT result of $1.341mm \pm 0.640mm$.

For the DHT method, using a global optimization strategy seems to be an advantage over the new method. It aims to maximize an energy function composed of Hough values and motion penalties on motion trajectories by using a dynamic programming (DP) technique. However, the existence of the CHD in some lateral views makes the global optimization strategy error-prone. Although the hard inter-frame constraint used in DP restricts the sudden jump of the IHDA position from one hemi-diaphragm to the other, the optimization strategy tends to find stronger features in Hough space (Figure 3a). In our experiment, finding the most probable Hough peak with a given variation in position, parabola curvature and image intensity in a neighborhood continuously over the image sequence is more likely to track the correct hemi-diaphragm (figure 3b).

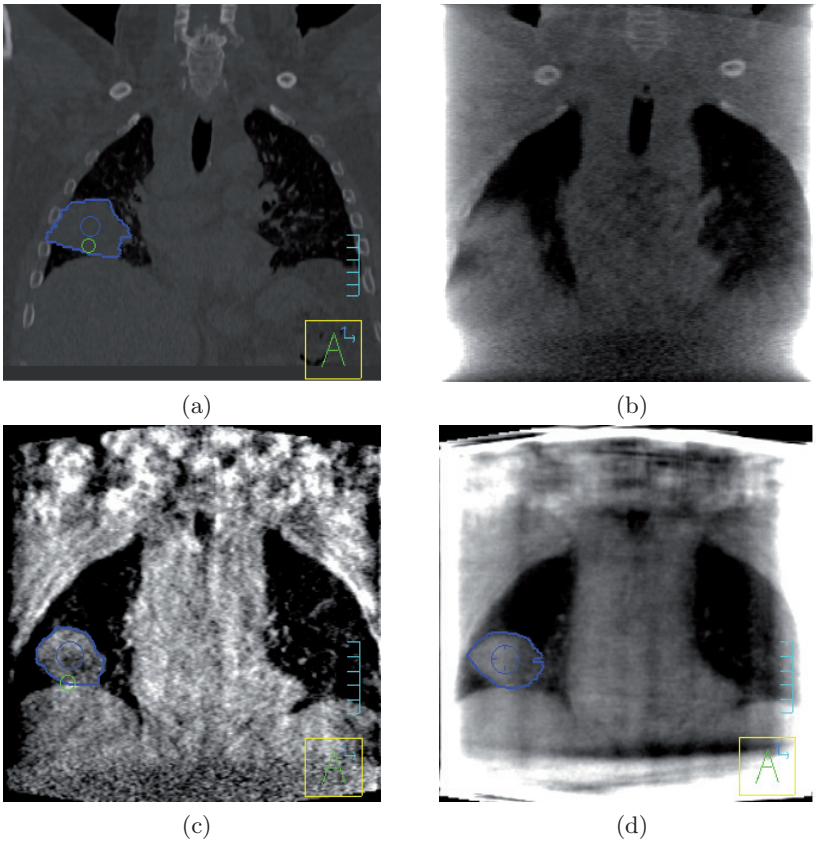
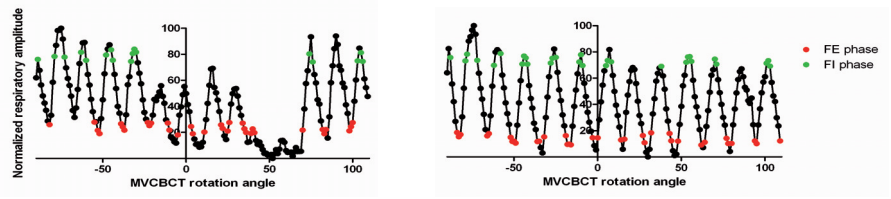


Fig. 4: Image reconstruction example slice in coronal view: (a) FE phase of planning CT; (b) FDK-MVCBCT reconstructed from all the projections; (c) FE phase of FDK-MVCBCT; (d) FE phase of ART-MVCBCT. Tumor contour is displayed as blue

3.2 Respiratory Correlated Reconstruction

Figure 4a shows one coronal slice of a 4D planning CT. The MVCBCT has a much lower contrast to noise ratio (CNR). In figure 4b a MVCBCT image reconstructed using all the projection data is shown. Static tissues such as shoulders are visible, but objects in motion, such as ribs, hemi-diaphragms and the tumor are very blurred. In the example image of the respiratory sorted reconstruction, the tumor boundary and diaphragm are better defined, regardless of the reconstruction method: FDK (figure 4c) or ART (figure 4d). Reconstruction artifacts existed in both FDK-MVCBCT and ART-MVCBCT, further degrading the image quality. Both FDK and ART have truncation artifacts at the border of the imaging FOV. The intensities of organs outside the FOV contribute to the



(a) (b)

Fig. 5: Normalized respiratory phase for (a) irregular breath (b) regular breath

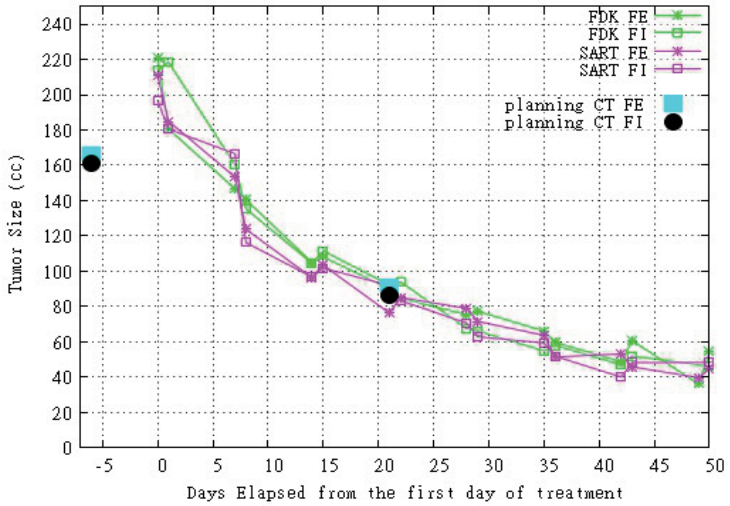


Fig. 6: Tumor volume plotted against elapsed days from first fraction

voxel at the FOV border during back-projection. For FDK, the streak artifact is observable, which is due to incomplete projections.

The patient received audio-coaching during the treatment in order to encourage regular breathing. While this makes the gated RT treatment more reliable, it can also help improve the reconstruction of limited projection data by avoiding large gaps of missing projections. However there are some exceptions when the patient is unable to maintain the regularity during the scan. In our study, the IHDA respiratory phase in one of 16 MVCBCT images shows a very irregular pattern(figure 5a), where the respiratory amplitudes for some maxima are much smaller than those of the other maxima. The exhaustive search strategy had a hard time for this image, since no projections are within the FI phase window from frame 60 to 160 (figure 5a). The image reconstructed in this projection set is severely degraded, making it impossible to contour the tumor correctly. For reference, figure 5b shows a regular breathing pattern.

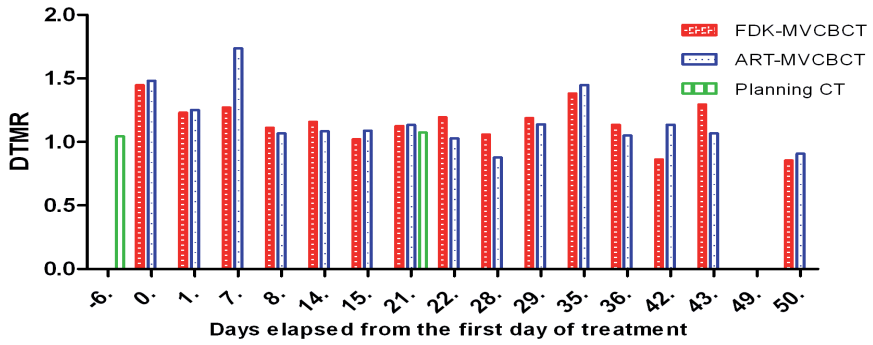


Fig. 7: DTMR value for 15 MVCBCT and 2 planning CT (MVCBCT with irregular breathing is excluded)

The radiation effect on tumor volume over the course of radiotherapy is plotted in figure 6. Images reconstructed by FDK and ART are both studied. The effect of radiotherapy shows an exponential pattern of volume shrinkage. The average discrepancy in tumor volume between FDK and ART methods is 9.38%, with a standard deviation of 7.2%. The discrepancy between the two methods may be due to human subjective error and improper setting of display window and level, which is shown to affect the object size measurement [2]. Although small errors existed between FDK and ART in size measurement, both methods agree well for the purpose of monitoring the trend of tumor change. Tumor volumes derived from two 4D planning CT scans are also shown in the graph, which can be considered as a gold standard. For the planning CT taken 6 days before the start of radiotherapy treatments, the tumor size is much smaller than the tumor size in both FDK-MVCBCT and ART-MVCBCT images taken minutes before the first radiation treatment. This may be due to tumor growth during the 6 day interim between imaging and treatment, since there is no radiation given during the 6 day interim and the tumor was particularly aggressive. The tumor size contoured in the second planning CT agrees well with the MVCBCT taken on the same day.

DTMR values are plotted in figure 7, which is computed from the IHDA motion and tumor motion displayed in figure 8. The measured motion is generally consistent between FDK-MVCBCT and ART-MVCBCT. The measured IHDA motion ranges from 0.78cm to 2.17cm, with an average and standard deviation of the difference between FDK and SART methods of $0.058cm \pm 0.052cm$. The tumor motion ranges from 0.8cm to 1.82cm, with a deviation between the two methods of $0.131cm \pm 0.112cm$. Comparing tumor and IHDA motion to the slice thickness in the superior-inferior direction, which is 1.071mm and 3.0mm for the MVCBCT and 4DCT respectively, variation by one voxel in the IHDA identification may induce a 10% DTMR change. This error is consistent with the average

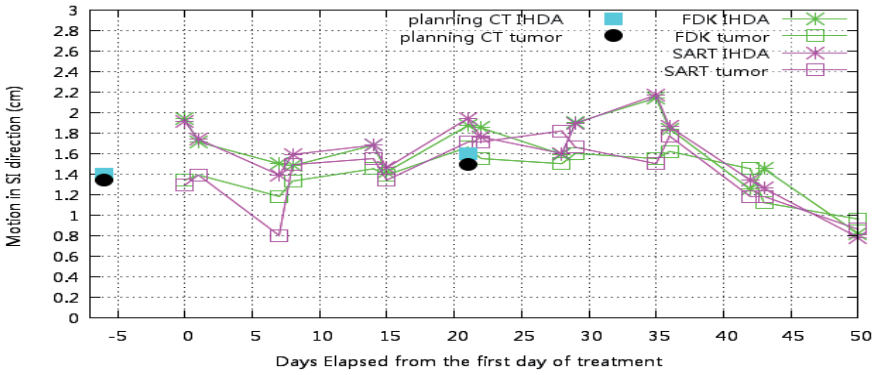


Fig. 8: Motion of IHDA and tumor centroid for 15 MVCBCT and 2 planning CT (MVCBCT with irregular breathing is excluded)

difference between FDK and ART measurements. There are a few cases with a significant exception such as the one on day 7. The motion of both the IHDA and the tumor centroid in this scan is relatively small. The division between the two measured motions to compute the DTMR further increases the inaccuracy, which is part of the reason for the larger error. DTMR from the planning 4DCT taken 6 days before the first treatment day is much smaller than that of the MVCBCT, possibly due to discrepancies in tumor volume and/or natural variations in breathing patterns from one day to the next. For the planning 4D CT taken on the 4th week, the DTMR is close to that of the MVCBCT. Typically most images have a DTMR larger than 1, which is consistent with lung expansion during inhale. For the few cases where DTMR is smaller than 1, they may be due to manual identification and contouring errors, but it is also possible that the patient was breathing more with their chest than with their diaphragm.

In this study we have verified the feasibility of assessing tumor response based on respiratory sorted MVCBCT derived tumor volumes. Parameters measured for the tumor contour and the IHDA in MVCBCT images are close to those of the planning CT and are in a reasonable range, except for a few cases in the first week of the MVCBCT, probably due to the time delay between 4DCT imaging and the first treatment fraction. The comparison of the DTMR for planning CT data and MVCBCT data also shows that it is feasible to use limited projection reconstruction to determine the tumor and diaphragm motion relationship. Although the result is only based on one patient and it is hard to draw a conclusion about lung volume measurement accuracy, it shows the feasibility of observing trends in tumor size changes and measuring the DTMR to establish a relationship between tumor and diaphragm motion.

Reconstruction artifacts induced by incomplete projections and truncations are the main reasons that affect tumor contour accuracy. The accuracy of using respiratory sorted MVCBCT reconstruction to measure object size and motion

still needs to be assessed. It will be very helpful to image an object with prior knowledge about its density, shape and size, in order to quantify the reconstruction error. Our future study will focus on phantom design and quantification of the reconstruction accuracy.

References

1. Ballard, D.: Generalizing the hough transform to detect arbitrary shapes. Pattern recognition 13(2), 111–122 (1981)
2. Baxter, B., J., S.: Factors affecting the measurement of size and ct number in computed tomography. Investigative Radiology 16(4), 337–41 (1981)
3. Briechle, K., Hanebeck, U.: Template matching using fast normalized cross correlation. vol. 4387, pp. 95–102 (2001), proceedings of SPIE
4. Cervino, L.I., Chao, A.K., Sandhu, A., Jiang, S.B.: The diaphragm as an anatomic surrogate for lung tumor motion. Phys Med Biol 54(11), 3529–41 (2009)
5. Chen, M., Siochi, R.: Diaphragm motion quantification in megavoltage cone-beam ct projection images. Medical Physics 37, 2312–20 (2010)
6. Cui, Y., Dy, J., Sharp, G., Alexander, B., Jiang, S.: Multiple template-based fluoroscopic tracking of lung tumor mass without implanted fiducial markers. Physics in medicine and biology 52, 6229–6242 (2007)
7. Duda, R., Hart, P.: Use of the hough transformation to detect lines and curves in pictures (1972)
8. Feldkamp, L., Davis, L., Kress, J.: Practical cone-beam algorithm. Journal of the Optical Society of America A 1(6), 612–619 (1984)
9. Galigekere, R., Wiesent, K., Holdsworth, D.: Cone-beam reprojection using projection-matrices. IEEE transactions on medical imaging 22(10), 1202–1214 (2003)
10. Gordon, R., Bender, R., Herman, G.: Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography* 1. Journal of theoretical Biology 29(3), 471–481 (1970)
11. Greenspan, M., Shang, L., Jasiobedzki, P.: Efficient tracking with the bounded hough transform. vol. 1. IEEE Computer Society; 1999 (2004)
12. Hills, M., Pridmore, T., Mills, S.: Object tracking through a hough space. pp. 53–56. London; Institution of Electrical Engineers; 1999
13. Hough, P.: Method and means for recognizing complex patterns (1962)
14. Jouvelot, P., Gifford, D.: Algebraic reconstruction of types and effects. p. 310. ACM (1991), proceedings of the 18th ACM SIGPLAN-SIGACT symposium on Principles of programming languages
15. Lappas, P., Carter, J., Damper, R.: Robust evidence-based object tracking. Pattern Recognition Letters 23(1-3), 253–260 (2002)
16. Lappas, P., Damper, R., Carter, J.: Object tracking by energy maximization. Soft Computing-A Fusion of Foundations, Methodologies and Applications 10(1), 20–26 (2006)
17. Lewitt, R.: Multidimensional digital image representations using generalized kaiser-bessel window functions. Journal of the Optical Society of America A 7(10), 1834–1846 (1990)
18. Lin, T., Cervino, L.I., Tang, X., Vasconcelos, N., Jiang, S.B.: Fluoroscopic tumor tracking for image-guided lung cancer radiotherapy. Phys Med Biol 54(4), 981–92 (2009)

19. Mills, S., Pridmore, T., Hills, M.: Tracking in a hough space with the extended kalman filter. p. 173?82. Citeseer (2003), british Machine Vision Conference
20. Morin, O., Gillis, A., Chen, J., Aubin, M., Bucci, M., Roach III, M., Pouliot, J.: Megavoltage cone-beam ct: System description and clinical applications. Medical Dosimetry 31(1), 51–61 (2006)
21. Nash, J., Carter, J., Nixon, M.: Dynamic feature extraction via the velocity hough transform. Pattern Recognition Letters 18(10), 1035–1047 (1997)
22. Pouliot, J., Bani-Hashemi, A.: Low-dose megavoltage cone-beam ct for radiation therapy. International Journal of Radiation Oncology* Biology* Physics 61(2), 552–560 (2005)
23. Reitz, B., Gayou, O., Parda, D., Miften, M.: Monitoring tumor motion with on-line mega-voltage cone-beam computed tomography imaging in a cine mode. Physics in Medicine and Biology 53(4), 823–836 (2008)
24. Schiwietz, T., Bose, S., Maltz, J., Westermann, R.: A fast and high-quality cone beam reconstruction pipeline using the gpu. Medical Imaging 2007: Physics of Medical Imaging 6510 (2007)
25. Siochi, R.: Deriving motion from megavoltage localization cone beam computed tomography scans. Physics in Medicine and Biology 54, 4195 (2009)
26. Sonke, J., Zijp, L., Remeijer, P., van Herk, M.: Respiratory correlated cone beam ct. Medical Physics 32(4), 1176–86 (2005)
27. Vedam, S., Kini, V., Keall, P., Ramakrishnan, V., Mostafavi, H., Mohan, R.: Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker. Medical Physics 30, 505 (2003)
28. Xu, Q., Hamilton, R.J., Schowengerdt, R.A., Alexander, B., Jiang, S.B.: Lung tumor tracking in fluoroscopic video based on optical flow. Med Phys 35(12), 5351–9 (2008)
29. Xu, Q., Hamilton, R., Schowengerdt, R., Jiang, S.: A deformable lung tumor tracking method in fluoroscopic video using active shape models: a feasibility study. Physics in Medicine and Biology 52(17), 5277–5294 (2007)

Poster Presentations

A Fully Automated Approach to Segmentation and Registration of 3D Medical Image Data for Pulmonary Diagnosis

Alvin Ihsani[1], Jan Modersitzki[2], and Troy Farncombe[3]

[1]Computing and Software, McMaster University, Ontario, Canada
[2]Institute for Mathematics, University of Lübeck, Lübeck, Germany
[3]Radiology, McMaster University, Ontario, Canada
ihsania@mcmaster.ca

Abstract. Molecular imaging is an important tool that has found widespread use in the diagnosis and observation of various diseases and has more recently been used in areas such as drug development in order to facilitate the observation and analysis of newly developed drugs. The amounts of data in drug development experiments may be very large due to the involvement of both spatial and temporal information of medical images. Imaging techniques can facilitate the analysis of this data by automating information extraction and providing meaningful results. We propose a fully automated approach to pulmonary diagnosis in the context of drug development experiments using image segmentation and registration techniques. In particular, we propose a modification of the Chan-Vese approach for a stable segmentation of the lungs which then serves as a starting point in a spatial alignment process. Our results demonstrate the feasibility and potential of the proposed approach.

Keywords: image segmentation, image registration, automated diagnosis

1 Introduction

Medical imaging is a relatively new technology that has found widespread use in the observation and diagnosis of various diseases and symptoms. More recently, this technology has penetrated such areas as drug development in order to facilitate the observation of the effects of newly developed drugs on various subjects.

Drug development experiments employing medical imaging techniques (such as CT and SPECT) may produce large amounts of data since spatial and temporal information spanning over a few weeks or months (e.g. 8 months) may be utilized to derive an informed conclusion of the effectiveness of the drug and its short and long term effects. The scans may be taken in regular intervals over the duration of the treatment and therefore corrections may need to be made to the data due to

– the position of the subject changing in consecutive scans,
– different machine artifacts over the period of treatment,
– artifacts introduced by motion such as breathing, heart-beats or other involuntary motions and
– physiological differences in the subject over the period of treatment.

When the amount of data in such experiments is large the automatic correction of the aforementioned artifacts becomes indispensable in order to minimize the time and resources needed to process the information and provide meaningful results.

This paper focuses on the automated correction of motion artifacts introduced by the placement of the subject in the scanning machine in consecutive scans.

1.1 Problem Description

The objective of the project presented in this paper is to facilitate the observation of the effects of a newly developed drug on the lungs of a subject over the period of treatment.

The approach taken by the McMaster Pre-clinical Imaging Facility to observe the effects of the treatment on the lungs is by tracking the development of the air flow and blood circulation in the lungs using SPECT (Single Photon Emission Computed Tomography) and CT (Computed Tomography) data acquired using a combined SPECT–CT scanner. This approach is commonly taken when looking for such diseases as cancer or pulmonary embolism [5,9,10].

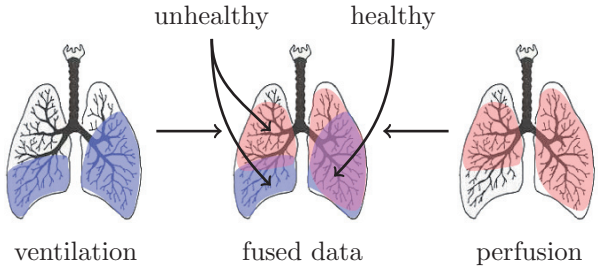


Fig. 1. Visualization of the diffusion of the perfusion contrast agent (right) showing blood-flow, ventilation contrast agent (left) showing air-flow and fused data (centre) showing potentially unhealthy regions in the lungs.

In order to track air flow in the lungs a ventilation contrast agent (Tc-99m Technegas) is inhaled by the subject and its diffusion is tracked using a SPECT scanner. Similarly, a perfusion contrast agent (Tc-99m labeled macro-aggregate albumin) is injected in the bloodstream and its diffusion is tracked using the same scanning machine [5,9]. The air and blood diffusion data, "fused" together, provide information about healthy and unhealthy regions in the lungs [5,9,10]

as shown in Figure 1. As an example, cancer can be identified by regions in the lungs where the blood flow is high but air flow is poor.

The fusion of the ventilation and perfusion SPECT data and its interpretation is outside the scope of this paper. The problem treated here deals with the automatic (or unsupervised) alignment of the ventilation and perfusion SPECT data, since these are taken more than two hours apart, so that the data is ready for fusion and interpretation.

SPECT data provides only information about the diffusion of a contrast agent in the lungs and this may not at all me indicative of the shape of the lungs. Also, the structural differences between ventilation and perfusion SPECT may be large (see red and blue areas in Figure 1). This complicates the task of aligning the SPECT data meaningfully.

Using a combined SPECT–CT scanner, CT and SPECT data can be acquired almost concurrently (two minutes apart) and output a pre-aligned pair of ventilation (or perfusion) SPECT and CT. Therefore, to obtain the desired alignment between the two SPECT data, the corresponding pre-aligned CT data are used since they contain consistent information about the shape of the lungs and exhibit more structural consistency over different scans.

Since SPECT data contains information about the lungs then the CT data pair must be aligned using information from the chest region (i.e. lungs, heart, ribs and other proximal tissues), which in this paper are referred to as the *target region*. The intent is to remove far away objects or organs (head, limbs, muzzle holder etc.) that might obstruct the proper alignment of the chest between the two CT data. The *target region* is found using a novel segmentation technique which is a central step in the correct alignment of the ventilation and perfusion CT data.

1.2 A High-Level Description of the Alignment Process

At a high-level the alignment process designed in this project aims to automatically align the perfusion and ventilation SPECT by finding a transformation which aligns the corresponding perfusion and ventilation CT using the the *target region*. The alignment process is broken down into the following steps.

1. Find an approximate region of the lungs in the perfusion CT data using the perfusion SPECT data. Since the location of the lungs in the perfusion CT data is not known the perfusion SPECT can provide some "hints" due to the fact that it only contains information about the blood-flow inside the lungs.
2. Use a segmentation technique, which uses the region found in the previous step as an initial guess, to extract the lungs from the perfusion CT and lead to the *target region*.
3. Finally, a rigid model multi-level image registration method is used to align the *target region* in the perfusion CT to the ventilation CT. The resulting transformation is used to align the perfusion and ventilation SPECT data.

Figure 2 shows a sketch of the aforementioned steps.

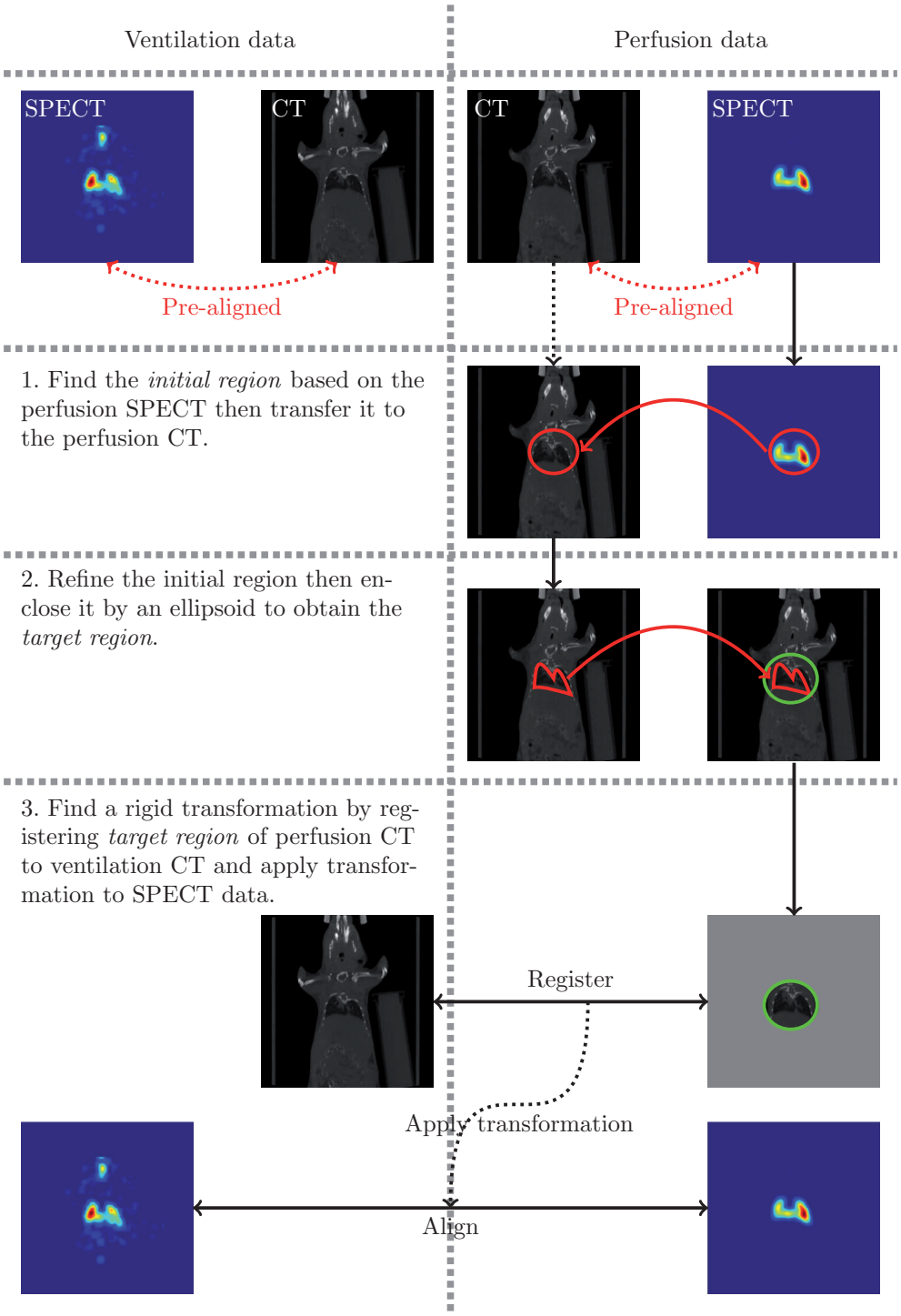


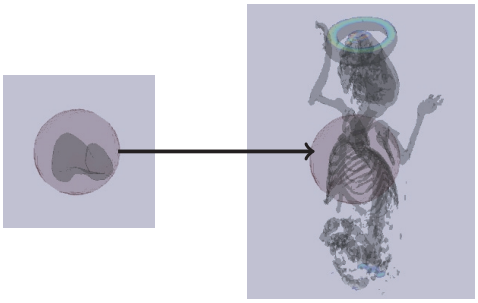
Fig. 2. Diagram of the perfusion/ventilation alignment process.

2 A Detailed View of the Methods Used in the Alignment Process

This section describes the reasons behind each step of the alignment process as well as some of the mathematical details of the segmentation and registration methods.

2.1 Ellipsoid Fitting: Finding the Initial Region

In the first step of the alignment process, the aim is to find an approximate location where the lungs are situated in the perfusion CT scan. Because the region where the lungs may be situated is narrowed down the initial guess for the segmentation step is *stabilized*. Initially, the alignment process "does not know" where the lungs are in the perfusion CT, but because the corresponding perfusion SPECT contains only information from inside and closely around the lungs the SPECT can be used to indicate an initial region of the lungs in the CT as shown in Figure 3.



(a) (b)

Fig. 3. A typical result of the initial region (a) on the perfusion SPECT and (b) on the perfusion CT.

In this step, the initial region is described by a smooth ellipsoid since it is an easy shape to describe mathematically. The shape of the initial region is not of importance since it is modified in the segmentation step.

While a simple thresholding of the perfusion CT data may seem sufficient at first this complicates the problem of finding a region of interest that does not include structures such as the arms and neck. For instance, a simple thresholding (e.g. at -300 HU) would include "fatty" structures in the arms of the mouse in addition to the lungs rendering the initial guess for the *target region* too large and making the result of the following segmentation step "less reliable". In addition, it is a challenging task to find an appropriate thresholding value and a considerable amount of time needs to be spent analyzing each sample of data. In order to avoid these challenges, the ellipsoidal region bases its initial guess on the perfusion SPECT data which contains only lung information however small this may be.

The ellipsoid used in this project is three-dimensional and is described by the function

$$W(\omega, x) = \begin{cases} 1 & R < 0.5 \\ -\frac{\cos(2\pi R)}{2} + \frac{1}{2} & 0.5 \leq R \leq 1 \\ 0 & R > 1 \end{cases} \tag{1}$$

with

$$R = \sqrt{\left(\frac{x_1 - \omega_4}{\omega_1}\right)^2 + \left(\frac{x_2 - \omega_5}{\omega_2}\right)^2 + \left(\frac{x_3 - \omega_6}{\omega_3}\right)^2}. \tag{2}$$

This defines an ellipsoid parameterized by ω whose "density" increases smoothly towards the centre with $0 \leq W(\omega, x) \leq 1$.

The aim is to find the smallest ellipsoid enclosing a thresholded version of the perfusion SPECT data as shown in Figure 3(a). Considering that the highest concentration of contrast agent is exhibited in the inside of the lungs, the threshold value is chosen so as to remove small nonzero values outside the lungs (or "miscounts").

Minimizing the objective function

$$F(\omega) = ||\omega_1^2 + \omega_2^2 + \omega_3^2||_2 + \alpha \int_\Omega u(x)(1 - W(\omega, x))dx \tag{3}$$

the smallest ellipsoid enclosing the thresholded perfusion SPECT data

$$u(x) = \begin{cases} 1 & g(x) \geq \text{threshold} \\ 0 & g(x) < \text{threshold} \end{cases} \tag{4}$$

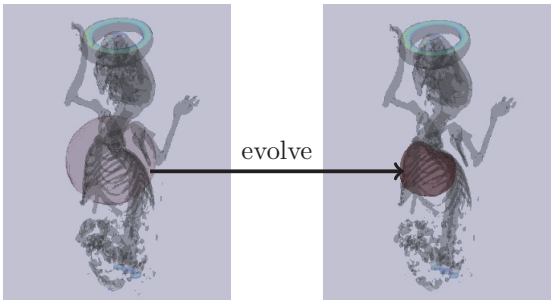
is found where $g : \Omega \to \mathbb{R}$ is the image and $\Omega \subset \mathbb{R}^d$ is the image domain with $d = 3$. The parameter $\alpha > 0$ is chosen empirically. In equation (3), the first term aims to minimize the radii of the ellipsoid while the second term aims to include as much data as possible inside the ellipsoid. The minimizer of equation (3) is found using the Gauss-Newton method available in FAIR [4].

2.2 Lung Segmentation: Refining the Initial Region

The initial region obtained in the step described in the previous section may be so large that it may enclose organs or objects of the perfusion CT data which might obstruct proper alignment of the chest. Conversely, the initial region may be so small (due to poor circulation in the lungs) that the information it contains may be insufficient to properly align the perfusion and ventilation CT. For these reasons, the initial region must be refined to a shape that is comparable to the size of the lungs as they appear in the perfusion CT (see Figure 4). The extraction of the lungs (or *target region*) can be performed using a segmentation method.

While a large variety of segmentation methods exist, including region growing, histogram-based, model-based and edge detection methods [1,6], a segmentation method based on the Chan-Vese model [13] was chosen. All of the aforementioned methods may be used for lung segmentation, however, the choice for

this project was based on the fact that a "stable" initial guess for the *target region* is obtained from the perfusion SPECT data in the step described in the previous section (i.e. a guess which lies either mostly inside the lungs or very proximally to their surface) and also on the requirement of avoiding the supervision of the segmentation and registration steps. Region growing methods may also have provided the desired result, however, this has not been tested.



(a) (b)

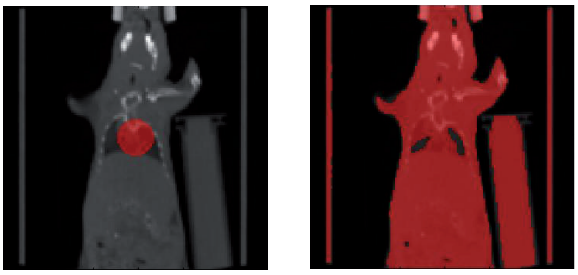
Fig. 4. The evolution from (a) the initial region to (b) the *target region* closely enclosing the lungs in the perfusion CT data.

A Novel Segmentation Method The Chan-Vese method [13] seemed to better fit the needs of this project since the image needs to be partitioned into only two regions (the lungs and the "rest") and therefore a simple two phase piecewise constant model is sufficient. However, the Chan-Vese approach [13] did not provide the desired result (see Figure 5) and the energy functional was modified to

$$E(\phi) = \mu \int_{\Omega} |\nabla H(\phi(x))| dx + \nu \left| \mathbf{V} - \int_{\Omega} H(\phi(x)) dx \right| + \lambda \int_{\Omega} ||g(x) - c||^2 H(\phi(x)) dx. \tag{5}$$

Here, $\mu, \nu, \lambda > 0$ are scalars which are chosen empirically, $\Omega \subset \mathbb{R}^d$ is the image domain with $d = 3$, $g : \Omega \to \mathbb{R}$ is the image, $\phi : \mathbb{R}^d \to \mathbb{R}$ is the level-set function [11], H is a smooth approximation to the Heaviside function as proposed in [13,14], $\mathbf{V}$ is a constant representing a rough estimate of the volume of the lungs and c is a constant representing an estimated average intensity value (in this case in Hounsfield Units) inside the lungs in the CT data.

In equation (5), the first term aims to minimize the segmenting surface defined as $\Gamma = \{\mathbf{x} | \phi(\mathbf{x}) = 0\}$ [7,8,13], the second term asks that the volume inside Γ be as close as possible to the estimated volume $\mathbf{V}$ and the third term asks that the region of the perfusion CT data g inside Γ be as close as possible to the estimated average intensity value c. The functional in (5) is minimized using a variational approach and the same discretization scheme proposed in [13].



(a) (b)

Fig. 5. Coronal slices of 3D data showing (a) a typical initial guess and (b) a typical result of the Chan-Vese segmentation method with segmented area shown in red.

The main difference between our functional and that proposed by Chan and Vese in [13] is that our functional is only affected by intensity values of the data inside the segmenting surface while the functional proposed in [13] is affected by values both inside and outside the segmenting surface.

As a concluding step in the refinement of the initial region, an ellipsoid is fit on the segmented region using the same approach described in Section 2.1. The motivation behind this additional step is to stabilize the registration. The less information is used from the perfusion CT the more sensitive to noise the registration becomes and also other regions in the ventilation CT (e.g. fatty tissues in the arms) may match the intensity values in the lungs. The ellipsoid enclosing the segmented region is a simple way to include closely surrounding tissues, specifically the ribs, which may help in providing more information for the rigid registration due to their relatively rigid structure. This final ellipsoid will be referred to as the *target region*.

2.3 Registration: Aligning the Ventilation and Perfusion Data

In the last step of the alignment process, the aim is to use a registration method to find a rigid transformation that aligns the perfusion and ventilation CT by using the *target region*. For the purposes of this project (in small animal imaging) a rigid registration model is sufficient and was also recommended by the staff at the McMaster Pre-clinical Imaging Facility, however, for larger animals and humans an affine or elastic registration model may have to be used.

In particular, the objective functional

$$\mathcal{J}[y](\nu) = \frac{1}{2}\int_{\Omega} (\mathcal{T}(y(\nu, x)) - \mathcal{R}(x))^2 dx \tag{6}$$

was minimized in order to find a rigid transformation y parameterized by ν (allowing only for rotation and translation) that increases the similarity between the *target region* weighted perfusion CT data, $\mathcal{T}$, and the ventilation CT data, $\mathcal{R}$. The reason for the spatial filtering is to remove unwanted organs or objects (head, legs, muzzle holder etc.) in the perfusion CT data that might obstruct registration of the chest area between perfusion and ventilation CT.

The registration of the CT data is multi-level [2,3], where the two images have been registered at multiple gradually increasing resolutions. Multi-level image registration is an effective strategy to avoid local minima [2]. Functional (6) was minimized using the Gauss-Newton method available in FAIR [4].

Finally, after obtaining the optimal parameters ν, the rigid transformation is applied to the perfusion SPECT for alignment to the ventilation SPECT.

3 Testing the Alignment Process

Materials The materials provided by our clinical partners for this project consist of SPECT and CT samples of a mouse taken over a period of six weeks for a total of 24 samples. For each week two pairs of pre-aligned data were provided, namely ventilation SPECT and CT and perfusion SPECT and CT at a resolution of $256 \times 256 \times 256$ voxels. The size of each sample is $(58.88 \times 58.88 \times 58.88) mm^3$.

Experiment Setting To test the alignment process the ground truth was simulated on real-life data using a set of 60 randomly generated transformations with mean zero for both rotation and translation and standard deviation $\sigma_{rotation} = 0.3rad$ and $\sigma_{translation} = 7mm$. The simulated transformations are roughly 20 times larger than those found in real-life data. The results of the segmentation and alignment process for the real-life data were also visually inspected for correctness.

The perfusion SPECT data, after some observation, was thresholded at intensity value above 300 counts. The parameter α in the ellipsoid fitting objective function (3) was set to $\alpha = 1$.

The parameters for the minimization of the novel segmentation functional (5) were found empirically and set to $\mu = 0.5 \cdot 255^2$, $\nu = 10^5$, $\lambda = 1$ with a fixed time-step $\tau = 0.046$. The segmentation was performed at $16 \times 16 \times 16$ voxels.

The resolutions used for the multi-level registration are $2^i \times 2^i \times 2^i$ voxels with $i = \{3, 4, 5, 6\}$.

Alignment Process Results The results obtained for the simulated set of transformations are summarized in Table 1 in terms of the absolute error between the simulated transformation and the result of the alignment process showing both the mean and standard deviation for translation and rotation.

In order to get a better idea of the "reliability" of the alignment process it is also important to look at the *target regions* obtained for each week. Table 2 summarizes these results.

The alignment process takes on average 5 iterations (0.5 seconds) for the ellipsoid fitting steps, 15 iterations (20 seconds) for segmentation and 3–6 iterations (3 minutes and 20 seconds) for the multilevel registration step. The machine used is a standard PC with 2 GB of RAM and a 2 GHz AMD Turion X2 processor running in Matlab.

	Mean		
Rotation (rad)	$-2.0447 \cdot 10^{-5}$	$-3.5597 \cdot 10^{-5}$	$-2.2869 \cdot 10^{-5}$
Translation (mm)	$-7.9440 \cdot 10^{-5}$	$-2.4414 \cdot 10^{-4}$	$-1.2205 \cdot 10^{-3}$

	Standard Deviation		
Rotation (rad)	$2.7496 \cdot 10^{-5}$	$3.3461 \cdot 10^{-5}$	$3.4985 \cdot 10^{-5}$
Translation (mm)	$1.2506 \cdot 10^{-3}$	$1.3136 \cdot 10^{-3}$	$1.4992 \cdot 10^{-3}$

Table 1. Absolute error between the simulated transformations and alignment process results summarized in terms of mean and standard deviation.

Radii (mm)	Week					
	1	2	3	4	5	6
ω_1	6.4212	4.8325	4.1047	4.2579	2.4686	2.0770
ω_2	7.1103	5.5471	6.1601	4.4618	3.5114	2.0770
ω_3	5.8149	3.5433	3.4566	3.0731	2.1633	0.7248

Table 2. The target region radii may vary for each week depending on the perfusion SPECT thresholding value (here ≥ 300 counts).

4 Discussion and Conclusions

The novel segmentation method proposed in this paper has shown to be effective for the purposes of lung segmentation and has provided the desired results. The objective functional (5) is not convex since the result of the segmentation depends on the starting guess, however, by stabilizing the starting guess using the ellipsoid fitting method described in Section 2.1 the result of the segmentation has shown to be very consistent (see Figure 6).

Furthermore, the overall alignment process has shown to be reliable since the absolute error in the resulting transformation was in the order of 10^{-3} for both rotation (rad) and translation (mm) given that the target region changes considerably in each week (see Table 2) and the smallest cell size used for registration is $\approx 0.92mm$ at the edge. The down-sampling of the data for both segmentation and registration is done to decrease computation time but also to test the sensitivity of the overall alignment process against varying resolutions. This has shown to work quite well for very "rough" samples of data as presented here as well as for finer data without changing parameters.

The novel segmentation method proposed in this paper achieves the desired result using only two additional pieces of information, namely an estimated lung volume and average intensity value in the lungs. While it is hard to estimate these two parameters accurately, from the tests carried out so far it seems the segmentation method will provide the desired result even for very rough estimates. For instance, in this project the lung volume was estimated as one-sixteenth of the total volume of the data which is a very rough estimate. In reality, the lung volume may be estimated using statistical data based on species, age, weight and other factors. In addition, the average Hounsfield unit values used for the lungs

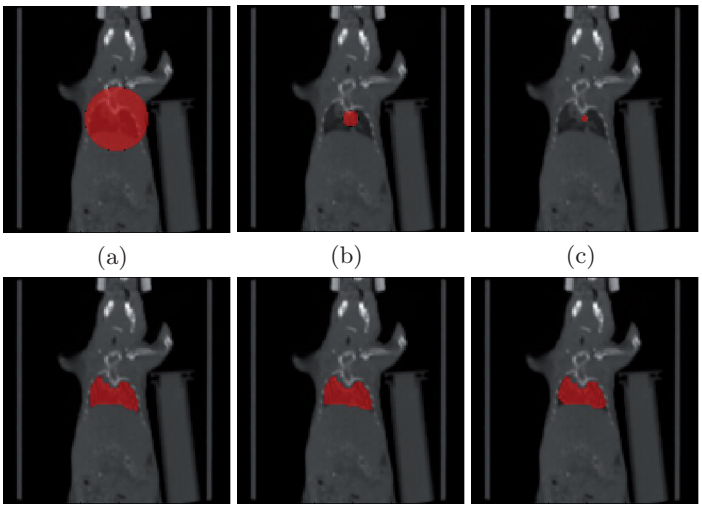


Fig. 6. Typical initial guesses 6(a)–6(c) and their respective results 6(d)–6(f) shown in coronal slices of 3D real-life data. The area in red represents the region inside the segmenting surface.

were set between -100 HU and -400 HU but the segmentation method still provides the desired result even though the actual average value is -150 HU. The segmentation parameters μ, ν and λ in equation (5) remain unchanged for the samples of data presented here. In fact the parameter μ has not been changed for other (synthetic) data sets as well, however, the parameters β and λ may need to be tweaked very slightly from the values presented above to obtain the desired result.

One limitation of the alignment process is the thresholding of the perfusion SPECT data for ellipsoid fitting. In the future this step will be substituted by the Chan-Vese segmentation model which, after some preliminary tests, has shown to provide much better results without the need of estimating a threshold value. Furthermore, the resulting level-set function from the Chan-Vese segmentation method can be used as the initial guess for the novel segmentation functional (5).

Other considerations for the improvement of the alignment process include improving the numerical optimization scheme of the segmentation step, performing multi-scale segmentation of the perfusion CT data and using non-rigid registration models.

References

1. D. L. Pham and C. Xu and J. L. Prince. A Survey of Current Methods in Medical Image Segmentation. In *Annual Review of Biomedical Engineering*, volume 2, pages 315–338. 2000.
2. E. Haber and J. Modersitzki. A Multilevel Method for Image Registration. *SIAM J. Sci. Comput.*, 27(5):1594–1607, 2006.

3. J. Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, 2004.
4. J. Modersitzki. *FAIR: Flexible Algorithms for Image Registration*. SIAM, 2009.
5. J. Palmer and U. Bitzén and B. Jonson and M. Bajc. Comprehensive Ventilation/Perfusion SPECT. *The Journal of Nuclear Medicine*, 42:1288–1294, 2001.
6. J. S. Suri and D. Wilson and S. Laxminarayan. *Handbook of Biomedical Image Analysis: Volume 1: Segmentation Models Part A (Topics in Biomedical Engineering International Book Series)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.
7. L. Ambrosio and V. Tortorelli. Approximation of Functionals Depending on Jumps by Elliptic Functionals via Γ-Convergence. volume 28, n.3, pages 213–221, 1998.
8. L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. 2000.
9. M. Bajc and B. Jonson. Ventilation/perfusion SPECT an essential but underrated method for diagnosis of pulmonary embolism and other diseases. *European Journal of Nuclear Medicine and Molecular Imaging*, 2009.
10. O. Schillaci. Hybrid SPECT/CT: a new era for SPECT imaging? *European Journal of Nuclear Medical Molecular Imaging*, 32(5):521–524, 2005.
11. S. Osher and J. A. Sethian. Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations. *Journal of Computational Physics*, 79:12–49, 1988.
12. T. F. Chan and L. A. Vese. A level set algorithm for minimizing the Mumford-Shah functional in image processing. In *VLSM '01: Proceedings of the IEEE Workshop on Variational and Level Set Methods (VLSM'01)*, page 161, Washington, DC, USA, 2001. IEEE Computer Society.
13. T. F. Chan and L. A. Vese. Active Contours Without Edges. *IEEE Transactions on Image Processing*, 10(2):, 2001.
14. Zhao, Hong-Kai and Chan, T. and Merriman, B. and Osher, S. A variational level set approach to multiphase motion. *J. Comput. Phys.*, 127(1):179–195, 1996.

Airway Tree Segmentation By Removing Paths of Leakage

Gang Song[1], Nicholas Tustison[1], James C. Gee[1],

Penn Image Computing and Science Laboratory (PICSL), Dept. of of Radiology, University of Pennsylvania, Philadelphia, PA, USA,

Abstract. This paper describes a new method for airway segmentation from CT images. We propose a form of adjusted image gradients for thin airways segmentation and apply it in multi-stencil fast marching. A graph of airway path segments is constructed from the arrival time function of fast marching. Instead of detecting leakage during segmentation, our method verifies each path segment in a separate step, using a novel leakage cost function defined on the whole path. Our scheme of path removal can be viewed as complementary post-processing for existing region growing methods. Experiments show that the proposed method can remove leakage regions while keeping most thin airways.

1 Introduction

Segmentation of airway trees from computed tomography images is critical for various clinical applications involving pulmonary diseases. Diameters of fourth generation airway in a typical CT image are about two or three voxels wide. The limitation of imaging resolution and noise lead to the inhomogeneity of image intensities inside airway walls and also the blurring effects around airway walls. These factors make the balance between detecting leakage and extending thin airways very critical in airway segmentation

Various algorithms have been proposed in the literature. Schlathölter et al used level set methods for a simultaneous segmentation and tree reconstruction framework [1]. The authors proposed several heuristic rules to detect leakage in the growing regions. Tschirren et al proposed to keep an active region of cylinder shape [2]. By tracking the orientation of active cylinder, the active region was extended to next possible airway location. A multi-threshold approach was adopted in [3] to increase robustness in growing airways trees. Recent work from Christian et al [4] used gradient vector flow to guide growing direction. The work of [5] focused on extending thin airways by computing the shortest paths inside a search sphere from end points in the initial segmentation.

Many of these algorithms have shown successful segmentation of the bronchi and trachea. However, for the third and higher generations in the airway tree, current segmentation results still have room for improvement, indicated by a recent evaluation on 15 airway segmentation algorithms [6].

In this paper we explore a new approach in airway segmentation. Instead of mixing airway segmentation and leakage detection at every iteration as in

[4, 5], we divide this problem into a hypothesis generation of thin airway paths and a post processing procedure of removing leakage path candidates. For the purpose of generating as many hypotheses as possible, we propose a novel speed function for thin airways. To exclude leakage regions, we propose a novel cost function defined on the whole path candidate. Such a scheme is more flexible when evaluating the whole path and can be viewed as complementary to current region growing methods.

2 Methods

Our method includes two steps: path candidate generation and path candidate removal. In the first step, we generate all initial path candidates in the segmentation using the fast marching method. A new formulation of adjusted image gradients is proposed to compute the speed image. Path candidates are then extracted by back tracing in the arrival time function. Next, those paths which may contain the leakage regions are further removed from the initial segmentation using our novel leakage cost function.

2.1 Initial Segmentation Using Speed Function on Adjusted Gradients

We use fast marching as the first step to generate the initial airway segmentation. The arrival time is modeled by the Eikonal equation $|\nabla T|F = 1$, where F is the speed function of the propagation front and T is the arrival time. Given the 3D image volume I, one form of the gradient based speed function is $F = e^{-\beta|\nabla I|}$ used in [1]; β is a scalar weighting coefficient.

By definition ∇I usually has high values at the locations close to the boundary of airways, both inside and outside the airway. Thus fast marching will have a low speed along the boundary inside airways. For the trachea and bronchi, this is not a practical problem since the airway boundary takes a small portion in the whole volume. However, for thin airways (like fourth order airways), where the diameter of the airway is about two voxels wide, the speed F would be low along whole thin airway segment and this would prevent fast marching from successfully extending further (See Fig. 1b).

Using $|\nabla I|$ directly is not suitable for front propagation along thin airways with partial volume effect. One way to deal with this problem is to interpolate image to a higher resolution, or to assign a special label for airway boundaries in fast marching. Here we propose a different solution by adjusting the definition of image gradients. The ideal gradient for airway segmentation should be defined only on the wall of airways (brighter voxels) rather than inside airways (darker voxels), such that the speed inside airways are high while low outside. To formalize this notion, the following modified gradient is proposed:

$$g(x) = \begin{cases} |\nabla I(x)| & \text{if } |I\left(x + \frac{1}{2}\boldsymbol{n}_x\right) - I(x)| < |I(x) - I\left(x - \frac{1}{2}\boldsymbol{n}_x\right)|; \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

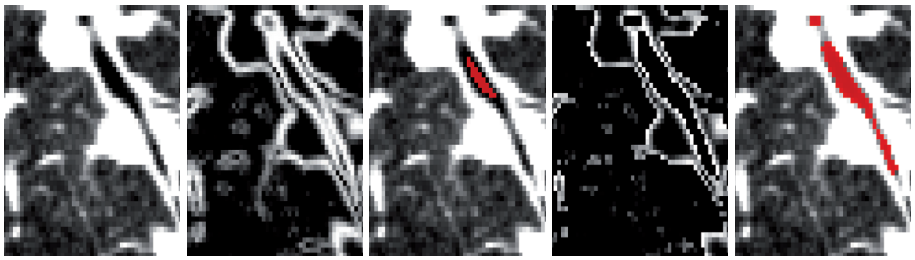
$\boldsymbol{n}_x$ is the unit direction of the gradient: $\boldsymbol{n}_x = \frac{\nabla I(x)}{|\nabla I(x)|}$ and is measured in voxel units. Note that this definition is not for leakage detection, but for generating all possible airway segmentation.

With this gradient definition, the corresponding front propagation equation becomes $F = e^{-\beta g(x)}$. A seed point x_0 in the trachea region is selected as the initial condition $T(x_0) = 0$. A comparison of results using $g(x)$ and $\nabla I(x)$ is shown in Fig.1. The fast marching gives the correct segmentation using the proposed $g(x)$ on this thin airway segment.

Another issue in solving front propagation for thin airways is the choice of connectivities. Many airway segments are extended along diagonal directions. However, Sethian's original fast marching method (FMM, [7]) is not accurate along these directions as it computes the derivative using 6-connectivity. In order to adopt the 26-connectivity in fast marching, we applied the multi-stencil fast marching method (MSFM, [8]) to solve the equation in practice. Compared to FMM, MSFM solves the equation along several predefined stencils to cover the entire 26 neighboring locations. For those stencils that are not aligned with the natural coordinate system, the equation is derived using directional derivatives. Details of MSFM can be found in [8]. A 26-connectivity is especially important for fast marching in thin airways when one location has no 6-connectivity neighbors in the airway.

2.2 Path Candidate Generation from Initial Segmentation

Given the arrival time function T and a threshold t_0, the initial segmentation is obtained as $S = \{x|T(x) < t_0\}$. S contains both airway tree segments and leakage regions. A nice property of the function T is that the propagation path from any given location to the seed location x_0 can be traced back using the gradient of T. Our motivation is to cluster voxels in S into different path segments. In



(a) (b) (c) (d) (e)

Fig. 1. Example of the proposed adjusted image gradients. (a) A region containing a thin airway (b) Gradients computed as $|\nabla I|$. Note that it has high gradients inside the airway, which prevent fast marching. (c) Airway segmentation using (b) to compute speed image and MSFM. (d) Adjusted gradients computed using $g(x)$. Only the gradients in the bright airway wall are preserved. (e) Airway segmentation using adjusted gradients in (d) to compute speed image and MSFM.

turn, leakage can be detected on voxels in each path segment as a whole, instead of on each voxel individually.

More formally, a graph G is built to describe the structure of the binary mask S. Each node in G is a cluster of voxels in S, corresponding to part of the airway or to part of the leakage. We apply an approach similar to [1] and [9] to get the graph G.

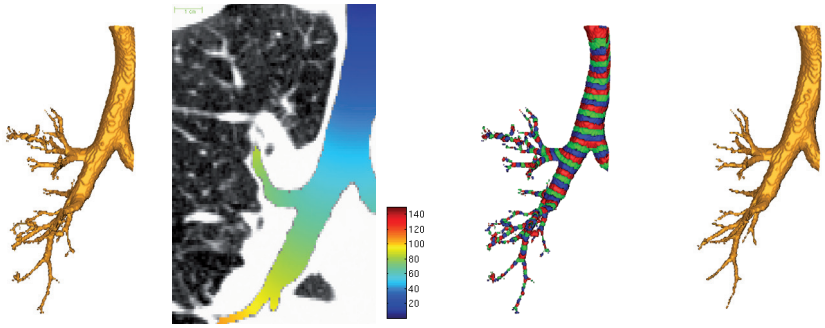
First a distance field, $D(x)$, to the initial seed point x_0 is computed in the domain S. Then each continuous value $D(x)$ is discretized into an integer j such that $D_0(x) = j$ if $jh \leq D(x) < (j+1)h$, in which h is the bin width for discretization. By assigning a node n to one connected component of the same integer value in $D_0(x)$, a graph G is constructed from S. Neighboring connected components are connected in the graph. The bin width h controls the shape of each node in the airways so that a node is roughly a tube-like structure. An example of the graph construction using discretization is illustrated in Fig.2.

The propagation from initial seed point x_0 to a voxel x_1 in S, denoted as $C_{x_0}^{x_1}$, can be traced by solving ordinary differential equation $\frac{dC(t)}{dt} = -\frac{\nabla D}{|\nabla D|}$ with boundary condition $C(t_1) = x_1$ and $C(0) = x_0$. We define the path from x_0 to x_1 on the graph G as the series of nodes that intersect with $C(t)$:

$$P_{x_0}^{x_1} = \{n | n \in G, n \cap C_{x_0}^{x_1} \neq \varnothing\} \tag{2}$$

2.3 Leakage Removal Using Cost Function on Path Nodes

Each path segment node n in G is a candidate for leakage removal. We consider three properties of the path segment: its volume, its vesselness measurement and how it is separated from the background in the image. The first two properties



(a) (b) (c) (d)

Fig. 2. Path graph construction. Only the airway tree in the right lung is shown. (a) Initial segmentation S using fast marching with adjusted gradients. (b) Distance transform D inside S with color bar shown on the right. (c) Graph G using discretization on D. Nodes $\{n\}$ are colored by red, green and blue. Each node n has the same discretization value in D_0 and belong to the same connected component. (d) Final segmentation of airway tree using our path segment removal approach.

have been investigated in the literature [1, 4, 5]. In this paper we use Frangi's vesselness definition [10] on dark tubes to measure each voxel in n. Those voxels of vesselness smaller than t_{vessel} are removed from node n. The nodes of a volume greater than a threshold t_{vol} are also removed. These two properties make sure that path segment n is of a tube-like shape, however, it may still contain leakage areas (see Fig.3 for an example).

We propose a novel leakage cost function to estimate how each path node can be separated from background. For each node n, its medial axis is computed as the segment direction . The local coordinate system use this direction as z axis of the segment. The surrounding region of n (obtained from image dilation operation) is then divided into k sections by different angles uniformly distributed in the x-y plane. Fig.3 shows an example of $k = 8$ sections in different colors. If the node has leaked into the background from some direction, the average gradient in that section would be low.

The leakage cost on path n, $S(n)$, is defined as the minimum of average gradients in all sections:

$$S(n) = \min_i \frac{1}{|N_i(n)|} \sum_{x \in N_i(n)} g(x) \tag{3}$$

where $g(x)$ is the adjusted gradient in section 2.1, $N_i(n)$ is the neighborhood of n in the i-th angle. The nodes of $S(n)$ lower than a threshold t_{path} are considered as leakage and are removed from G. The final airway segmentation is the connected component in G that contains most remaining nodes.

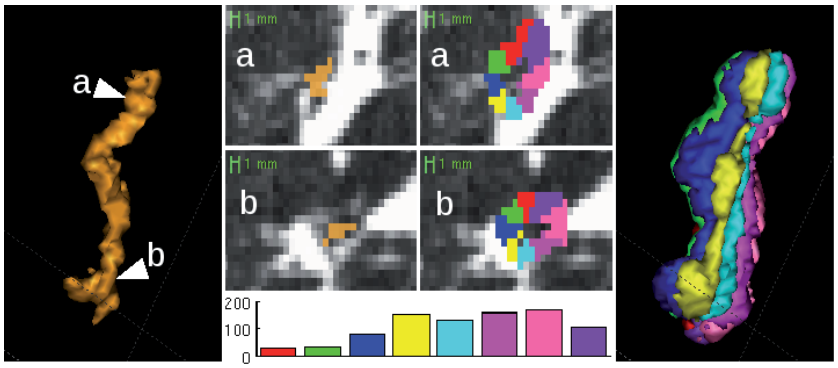


Fig. 3. Example of the proposed path leakage cost function on a node of leakage. *Left*: shape of the node is similar to airway, but is not part of airway. *First and second rows in the middle*: on left) one 2D slice with the path n shown in brown; on right) $k = 8$ sections of the neighborhood of n in different directions, each section $N_i(n)$ shown in different colors. *Last row in the middle*: bar plots of $\frac{1}{|N_i(n)|} \sum_{x \in N_i(n)} g(x)$ with each i shown in corresponding colors. *Right*: surface rendering of the $k = 8$ sections. Location a has an leakage in the upward direction; however location b is difficult to segment locally. Our proposed cost $S(n)$ is defined as the minimum of the $k = 8$ values (the red bar here), which is significantly lower than other bars.

3 Results

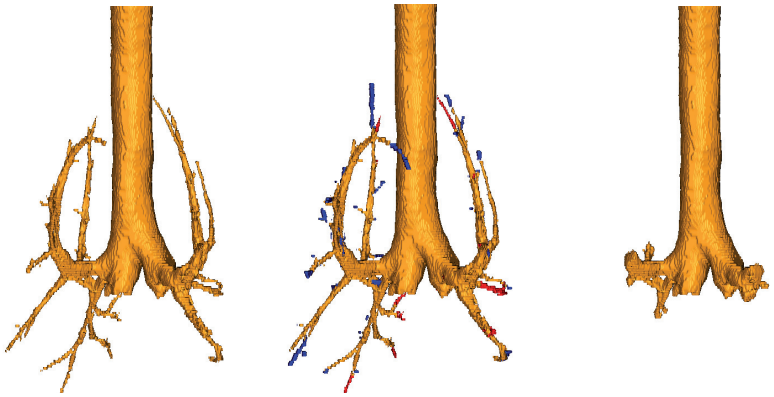
The proposed method was applied to extract airways in four canine specimens at 15 months after pulmonary lobe surgery (part of lobes were removed). The images were taken under forced inflationary pressure of 30 cm H_2O. All images have a slice thickness of 1.25 mm and in-plane resolution of 0.46×0.46 mm^2. The airways in these volumes were semi-supervised extracted and manually modified as the ground truth for evaluation.

In all our experiments, parameter β used in computing speed function was set to 0.05. The initial segmentation threshold t was 200. The distance field $D(x)$ was discretized by bin width $h = 2$. The value of h is related to the length of the node in practice. We tested h from 2 to 4 and got similar results.. We applied the multi-scale strategy when computing Frangi's vesselness, using four scales $s = 0.5, 1, 2, 3$ (see the discussion of scale s in [10]). For each path node n, its surrounding neighborhood was obtained by computing the dilation of radius= 3 and was further divided to $k = 8$ sections along the segment direction. The vesselness threshold was 30 and the volume threshold was 1000. The final segmentation was given by removing all path nodes whose path leakage cost $S(n)$ was lower than 10.

Our method requires a seed point in the trachea region as the initial condition for fast marching. While such points could be easily manually selected, we further automated the whole process by applying Hough transform on circles. Assuming that the trachea was roughly aligned with z axis in image volume, we scanned each slice along z axis from the top and the seed point x_0 was identified as the center of the first detected circle.

All final segmentation results S were evaluated against the manually labeled airway trees S_g. Using $|\cdot|$ to denote the number of voxels in a set, we computed the recall rate as $\frac{|S \cup S_g|}{|S|}$, the false alarm rate as $\frac{|S - S_g|}{|S|}$ and the missing rate as $\frac{|S_g - S|}{|S_g|}$. As trachea and bronchi are relatively easy to segment using region growing approaches and our main interest was to extract thin airways, these regions were excluded in computing both S and S_g. Another reason is that the trachea and bronchi might take up to 95% volume in the whole airways; thus it should be excluded as a huge bias in evaluating thin airways. An example of our final segmentation is illustrated in Fig.4. Missing regions $S_g - S$ are colored in blue and false alarm regions $S - S_g$ in red (in Fig.4b).

On the four canine specimens, our algorithm got the average recall rate of 96.8% and the missing rate is 8.2%. Most missing regions are extended from the end segments of 2 to 3 voxels wide. The false alarm rate is 3.2%, but we observed that many of the false alarm regions are due to the human label errors. These are not leakage regions in segmentation, which shows the good performance of our path elimination approach in preventing leakage while keeping most airways.



(a) (b) (c)

Fig. 4. Surface rendering of an example result using the proposed method. (a) Segmented airways using our approach. (b) Comparison with ground truth. The blue regions are labeled in ground truth but missing in our results. The red regions are labeled in our results but missing in ground truth. The brown regions exist in both ground truth and our results. (c) Example of the trachea and bronchi region, which is excluded when computing the rates.

4 Discussion

In this paper we proposed a new method for extracting airway trees from 3D computed tomography images. We focused our method on removing leakage regions while still segmenting most thin airways. Compared to existing region growing methods, our method has a unique path removal procedure to exclude potential leakage. A graph of path candidates is constructed from the arrival time using the fast marching method. The leakage regions are identified from all path segment candidates using our proposed leakage cost function. A similar idea of checking different directions was proposed in [11]. They used gradients in different radial directions from one voxel to track vessels. In comparison, our cost function is defined on the node of a whole path segment for leakage removal. Each node is a higher level structure, which is more robust to noise and has a more flexible definition. Furthermore, this is not a linear measurement which means it can not be represented as an integration along the path. Thus we do not use Dijkstra's algorithm, which was used in [5].

The second contribution in this paper is that we proposed a form of adjusted gradient in computing speed image. We also apply multi-stencil fast marching for a 26 connectivity neighborhood in thin airway segmentation. We show that thin airways of 2 to 3 voxels wide can be extracted by our adjusted image gradients and utilizing the connectivities along diagonal directions in fast marching.

It should be noted that our approach is not contradictory to existing region growing methods. By separating the leakage detection as a post processing step, our method can take the advantage of current work on region growing while

reducing leakage. The result from one approach can be used as input to the other. For example, the missing rate in our experiment may be further improved using the method in [5] in the stage of generating path candidates.

The number of datasets used in the evaluation is limited. A larger evaluation data set is needed in our future work. Also, we used manual segmentation as our ground truth. An alternative for constructing the reference is combining results from different algorithms as in [6]. It would also be interesting to test our algorithm on the output of other algorithms to verify the effectiveness of post-processing.

References

1. Schlathölter, T., Lorenz, C., Carlsen, I., Renisch, S., Deschamps, T.: Simultaneous segmentation and tree reconstruction of the airways for virtual bronchoscopy. In: SPIE International Symposium on Medical Imaging. (2002) 103–113
2. Tschirren, J., Hoffman, E., McLennan, G., Sonka, M.: Intrathoracic airway trees: segmentation and airway morphology analysis from low-dose ct scans. Medical Imaging, IEEE Transactions on **24**(12) (dec. 2005) 1529 –1539
3. van Ginneken, B., Baggerman, W., van Rikxoort, E.: Robust segmentation and anatomical labeling of the airway tree from thoracic ct scans. In: Medical Image Computing and Computer-Assisted Intervention. (2008) 219–226
4. Bauer, C., Bischof, H., Beichel, R.: Segmentation of airways based on gradient vector flow. In: The Second International Workshop on Pulmonary Image Analysis. (2009) 191–201
5. Lo, P., Sporring, J., Pedersen, J.J., Bruijne, M.: Airway tree extraction with locally optimal paths. In: Medical Image Computing and Computer-Assisted Intervention. (2009) 51–58
6. Lo, P., van Ginneken, B., Reinhardt, J., de Bruijne, M.: Extraction of airways from ct (exact'09). In: Second International Workshop on Pulmonary Image Analysis. (2009) 175–189
7. Sethian, J.A.: Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science. 2 edn. Cambridge University Press (June 1999)
8. Hassouna, M.S., Farag, A.A.: Multistencils fast marching methods: A highly accurate solution to the eikonal equation on cartesian domains. IEEE Transactions on Pattern Analysis and Machine Intelligence **29** (2007) 1563–1574
9. Hassouna, M.S., Farag, A.A.: Robust centerline extraction framework using level sets. In: Computer Vision and Pattern Recognition (CVPR'05). (2005) 458–465
10. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: Medical Image Computing and Computer-Assisted Intervention. (1998) 130–137
11. Wink, O., Niessen, W., Viergever, M.: Fast delineation and visualization of vessels in 3-d angiographic images. IEEE Transactions on Medical Imaging, **19**(4) (april 2000) 337 –346

Evaluation of Four-dimensional (4D) Computed Tomography (CT) Pulmonary Ventilation Imaging by Comparison with Single Photon Emission Computed Tomography (SPECT) Scans for a Lung Cancer Patient

Tokihiro Yamamoto[1], Sven Kabus[2], Jens von Berg[2], Cristian Lorenz[2], Michael L. Goris[3], Billy W. Loo, Jr.[1] and Paul J. Keall[1]

[1] Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA, USA
Tokihiro@stanford.edu
[2] Department of Digital Imaging, Philips Research Europe, Hamburg, Germany
[3] Division of Nuclear Medicine, Stanford University School of Medicine, Stanford, CA, USA

Abstract. The goal of this study was to present a method for physiologic validation and show results for a comparison of the 4D-CT ventilation and SPECT ventilation or perfusion. A 4D-CT and a SPECT ventilation/perfusion (V/Q) (V_{SPECT} and Q_{SPECT}) scans were acquired for a lung cancer patient. Two 4D-CT ventilation images were created using two metrics: Hounsfield unit (HU)-change ($V_{4D\text{-}CT}^{HU}$) and Jacobian determinant ($V_{4D\text{-}CT}^{Jac}$), and were compared with SPECT. The 4D-CT ventilation images were of higher resolution than SPECT. Areas of visual agreement and disagreement were identified. V_{SPECT} suffered from central airway depositions. Quantitatively, the correlation coefficients with V_{SPECT} (Q_{SPECT}) were 0.030 (0.357) for $V_{4D\text{-}CT}^{HU}$, and 0.184 (0.479) for $V_{4D\text{-}CT}^{Jac}$. The physiologic accuracy of 4D-CT ventilation imaging has been evaluated on a human subject. The weak correlations with SPECT could be due to limitations of the 4D-CT and/or SPECT techniques. Further investigation and patient studies are needed.

Keywords: Lung, functional imaging, ventilation, four-dimensional (4D) computed tomography (CT), single photon emission CT (SPECT).

1 Introduction

Imaging techniques of regional pulmonary function would further our understanding of pathophysiological characteristics of pulmonary diseases and also could be used for functional avoidance in lung cancer radiotherapy. Regional pulmonary function can be represented by ventilation or perfusion. There are several techniques for ventilation imaging which includes nuclear medicine imaging (the current standard of care) [1], hyperpolarized gas magnetic resonance (MR) imaging [2], and Xe gas computed tomography (CT) imaging [3]. These techniques have drawbacks such as

low resolution, high cost, long scan time, and/or low accessibility. Ventilation images can be created by a novel four-dimensional (4D) CT-based technique in three steps, by (1) acquisition of 4D-CT scans, (2) spatial voxel-wise mapping of different respiratory phases of 4D-CT images using deformable image registration (DIR), and (3) quantitative analysis of the resultant displacement vector fields (DVFs) for computing the ventilation metric. Guerrero et al. have originally developed this method using paired breath-hold CT images [4]. They used DIR to map the peak-exhale CT image to the peak-inhale image, and calculated the change in the fraction of air per voxel as regional ventilation based on the theory proposed by Simon [5]. Several investigators have applied this approach to 4D-CT images with some modifications [6-10]. The 4D-CT ventilation can be considered as 'free' information, because 4D-CT scans are in routine use for lung cancer radiotherapy in many centers, and ventilation computation involves only image processing (i.e., DIR) and analysis. Moreover, 4D-CT ventilation imaging has high resolution, low cost, short scan time, and high accessibility. There are various DIR algorithms and two ventilation metrics that have been and could be used for 4D-CT ventilation imaging [4, 6-10]. Variability in the DIR results between algorithms has been reported [11, 12], which could potentially impact 4D-CT ventilation imaging. To date there has been limited physiologic validation of 4D-CT ventilation imaging. Reinhardt et al. demonstrated a reasonably high correlation (average, $R^2 = 0.73$) between the average ventilation values in 4 mm thick sub-regions determined by the 4D-CT ventilation and Xe-CT for 5 anesthetized sheep [8]. Yamamoto et al. compared the 4D-CT ventilation in emphysematous (i.e., known low signal regions) and non-emphysematous lung regions for 12 patients, and demonstrated significantly lower ventilation in emphysema ($p < 0.001$) [13]. More recently, Castillo et al. compared the 4D-CT ventilation and single photon emission CT (SPECT) ventilation for seven patients, however demonstrated low Dice similarity coefficients (i.e., < 0.4 on average) for the functional lung segments. More thorough validation is clearly needed for human subjects. The goal of this study was to present a method for physiologic validation and show quantitative results for a comparison of the 4D-CT ventilation and SPECT ventilation or perfusion in a lung cancer patient.

2 Methods and Materials

First, we acquired a 4D-CT and a SPECT ventilation/perfusion (V/Q) (V_{SPECT} and Q_{SPECT}) scans for a lung cancer patient. Second, we created 4D-CT ventilation images. Two different 4D-CT ventilation images were created using two metrics: Hounsfield unit (HU)-change ($V_{\text{4D-CT}}^{\text{HU}}$) [4, 5, 7, 10, 13] and Jacobian determinant ($V_{\text{4D-CT}}^{\text{Jac}}$) [7, 8, 13]. Finally, we quantified the physiologic accuracy of 4D-CT ventilation imaging by comparison with V_{SPECT} and Q_{SPECT}. Here we define the physiologic accuracy as the voxel-based correlation between the 4D-CT ventilation and SPECT. Fig. 1 shows a schematic diagram for this process. $V_{\text{4D-CT}}^{\text{HU}}$ and $V_{\text{4D-CT}}^{\text{Jac}}$ were created in three steps, by (1) acquisition of 4D-CT scans, (2) spatial voxel-wise mapping of the peak-exhale 4D-CT image to the peak-inhale image using DIR,

and (3) quantitative analysis of the resultant DVFs for computing the ventilation metric (see the section 2.2 for more details). We also performed a SPECT/CT V/Q scan (see the section 2.3). Each of V and Q scans consisted of a low-dose CT acquisition followed by a SPECT acquisition. For the 4D-CT ventilation to SPECT comparison, the CT image of SPECT was aligned with the peak-exhale 4D-CT image using rigid image registration. The derived translations and rotations were applied to the SPECT images to align with the 4D-CT ventilation images, assuming that the patient did not move between the SPECT and CT scans.

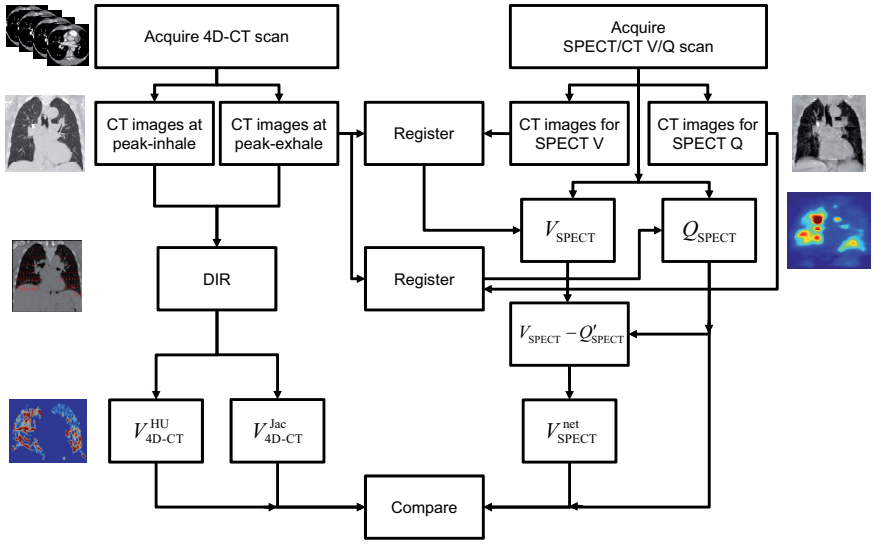


Fig. 1. A schematic diagram for creating 4D-CT ventilation images, acquiring SPECT/CT V/Q images, and comparing these images. For 4D-CT ventilation imaging, deformable image registration (DIR) was performed for spatial voxel-wise mapping of the peak-exhale 4D-CT image to the peak-inhale image. Resultant displacement vector fields were analyzed to create ventilation images using two different metrics: Hounsfield unit (HU)-change ($V_{4D\text{-}CT}^{HU}$) and Jacobian determinant of deformation ($V_{4D\text{-}CT}^{Jac}$). For SPECT/CT scans, the Q scan (Q_{SPECT}) was acquired first, followed by the V scan (V_{SPECT}). For the 4D-CT ventilation to SPECT comparison, the CT image of SPECT was aligned with the peak-exhale 4D-CT image using rigid image registration. The derived translations and rotations were applied to the SPECT images to align with the 4D-CT ventilation images. The activity-corrected SPECT Q image (Q'_{SPECT}) was subtracted from the V image to determine a net V (V_{SPECT}^{net}).

2.1 Patient

This study was a prospective analysis approved by Stanford University's Institutional Review Board. We studied a radiotherapy patient (66 year-old male) who had a stage IIB lung cancer with a tumor mass in the left upper lobe extending into the left hilum. The patient had a smoking history of 25 pack-years and severe chronic obstructive lung disease (COPD) (GOLD class III).

2.2 4D-CT Pulmonary Ventilation Imaging

4D-CT Scan Acquisition. The 4D-CT ventilation images were created in three steps as follows. The first step was the acquisition of a 4D-CT scan. Four-dimensional CT images are created by acquiring oversampled CT data simultaneously with a respiratory trace and by reconstructing a number of 3D-CT data sets correlated with a given respiratory phase range [14]. The 4D-CT scan was acquired on the GE Discovery ST multislice positron emission CT (PET)/CT scanner (GE Medical Systems, Waukesha, WI) in cine mode with the Varian Real-time Position Management (RPM) system (Varian Medical Systems, Palo Alto, CA) to record patient respiratory traces. The patient was in a supine position. The CT data were acquired at multiple couch positions covering the entire lung for a cine duration that is a little longer than the estimated respiratory period at each position. Audiovisual (AV) biofeedback [15] has been used to improve the patient breathing and to minimize the imaging artifacts. Scan parameters were set as follows: 120 kVp, approximately 100 mAs per slice, 0.5 sec gantry rotation, 0.45 sec cine interval, 512 × 512 matrix with a pixel size of 1 mm × 1 mm, and 2.5 mm slice thickness, as used clinically in our department. The GE Advantage 4D software was used to create a 4D-CT image set by sorting raw 4D-CT slices correlated with the RPM data into 10 respiratory phase-based bins (i.e., 0% to 90% at 10% intervals). We used paired 4D-CT images at the peak-exhale (60% for the current patient) and peak-inhale (0%) phases for ventilation computation. More details on the 4D-CT acquisition using the GE scanner with Advantage 4D have been described by Rietzel et al. [14]. The 4D-CT scan was acquired before radiotherapy is started.

Deformable Image Registration (DIR). The second step was DIR for spatial voxel-wise mapping of the peak-exhale 4D-CT image to the peak-inhale image, deriving a DVF. To estimate regional ventilation, it is necessary to establish correspondence of the individual respiratory phases. A common method for this task is elastic registration which tries to find a DVF such that one phase image (i.e., peak-inhale 4D-CT image) is transformed into another phase (i.e., peak-exhale). The resulting DVF points from the domain of the peak-exhale image into the domain of the peak-inhale image and therefore can be used for ventilation assessment. Elastic registration minimizes both a certain similarity measure D and a regularizing term S. By adding a regularizing term, the registration problem is well-posed. For D, we chose the sum of squared differences. For S, an elastic regularizer based on the Navier-Lamé equation was employed. The elastic regularizer assumes that the underlying images can be characterized as an elastic and compressible material. Its properties are modeled by the so-called Lamé constants, λ and μ. Based on calculus of variations, we arrive at a system of non-linear partial differential equations to be solved, i.e.,

$$\mu\Delta u + (\mu + \lambda)\nabla \cdot \nabla u = \nabla T_u (R - T_u), \quad (1)$$

where T_u and R correspond to the displaced peak-inhale phase and the peak-exhale phase, respectively. For the discretization of the formula (1), finite differences in conjunction with Neumann boundary conditions have been chosen. The resulting

system of linear equations consists on one hand of a sparse, symmetric and highly structured matrix arising from the regularizer and, on the other hand, of a so-called force vector corresponding to the similarity measure. By nature, the larger the contrast of misaligned image structures is, the larger the modulus of the force vector is. Therefore, bone structures would be perfectly aligned in CT images, whereas soft tissue may not. This holds particularly for the lung-rib interface with on the one side the parenchyma following the motion, and on the other side the ribs staying in place or even moving in opposite direction. To circumvent misalignment of parenchymal structures, we added a simple masking of the force vector. For every voxel with a HU value greater than 0 in the reference phase, the force vector was set to 0 for this voxel location. This results in a lung deformation which is not influenced by misalignment of the rib cage.

Ventilation Computation. The final step was quantitative analysis of the resultant DVFs for computing the ventilation metric. Ventilation can be estimated for each voxel based on the established correspondence between the peak-exhale and peak-inhale 4D-CT images. In this study, we investigated two metrics exploiting: (1) HU-change in corresponding parenchymal structures ($V_{\text{4D-CT}}^{\text{HU}}$) [4, 5, 7, 10, 13], or (2) local volume changes of the computed DVF, i.e., Jacobian determinant ($V_{\text{4D-CT}}^{\text{Jac}}$) [7, 8, 13]. For $V_{\text{4D-CT}}^{\text{HU}}$, we presumed that regional ventilation can be represented by the relative change in fraction of air as originally proposed by Simon [5]. The $V_{\text{4D-CT}}^{\text{HU}}$ value in the voxel at location (x) of the peak-exhale image is given by,

$$V_{\text{4D-CT}}^{\text{HU}}(x) = 1000\,\frac{HU_{\text{inhale}}\{x+u(x)\} - HU_{\text{exhale}}(x)}{HU_{\text{exhale}}(x)\left[HU_{\text{inhale}}\{x+u(x)\}+1000\right]}, \qquad (2)$$

where $HU(x)$ is the HU value and $u(x)$ is the displacement vector mapping the voxel at location (x) of the peak-exhale image to the corresponding location of the peak-inhale image. Given the influence of the statistical noise, the CT images have been smoothed with a Gaussian filter kernel with a variance, σ^2 = 1.5 mm^2 before computing $V_{\text{4D-CT}}^{\text{HU}}$. For $V_{\text{4D-CT}}^{\text{Jac}}$, we presumed that regional ventilation can be represented by a Jacobian determinant of deformation. The $V_{\text{4D-CT}}^{\text{Jac}}$ value in the voxel at location (x) is given by,

$$V_{\text{4D-CT}}^{\text{Jac}}(x) = \det\left[\nabla\{x+u(x)\}\right] - 1\,. \qquad (3)$$

For both $V_{\text{4D-CT}}^{\text{HU}}$ and $V_{\text{4D-CT}}^{\text{Jac}}$, positive and negative values indicate local expansion and contraction, respectively.

2.3 SPECT/CT V/Q Scan

The SPECT/CT V/Q scan consisted of four parts: (1) CT for SPECT Q (Q_{SPECT}), (2) Q_{SPECT}, (3) CT for SPECT V (V_{SPECT}), and (4) V_{SPECT}. The SPECT/CT V/Q scan was acquired on the GE Infinia Hawkeye SPECT/CT scanner. For Q_{SPECT}, 1 mCi of technetium-99m-labeled macroaggregated albumin (^{99m}Tc-MAA) was injected intravenously. For V_{SPECT}, technetium-99m-labeled diethylenetriamine pentaacetate (^{99m}Tc-DTPA) was aerosolized. The patient's nose was clamped and an inhalation of ^{99m}Tc-DTPA aerosols was given through a mouth piece of a nebulizer until the count

rate reached three times higher than the original count rate recorded after Q_{SPECT}. The SPECT scans were acquired in a 64 × 64 matrix with a pixel size of 8.8 mm × 8.8 mm with 60 projections over 360° and durations of 25 s per projection for Q_{SPECT}, and 30 s for V_{SPECT}. The slice thickness was 8.8 mm. The low-dose CT scans were also acquired at 120 kVp and 2.5 mAs. The SPECT images were reconstructed with attenuation correction using the CT images. For the SPECT/CT scan, the patient was in a supine position on a flat Lucite board placed on top of a concave couch in order to reduce registration errors between 4D-CT and SPECT images. Both the SPECT and CT scans were acquired under free breathing as it took long. The SPECT/CT scan was acquired on the same day as 4D-CT and hence the patient was expected to have the same functional status.

2.4 Statistical Analysis

The $V_{4D\text{-}CT}^{HU}$ and $V_{4D\text{-}CT}^{Jac}$ images were compared to V_{SPECT} and Q_{SPECT}. Here we assumed V_{SPECT} as ground truth. Q_{SPECT} was also used for the comparison as surrogate of V_{SPECT}, given that there was concern for central airway depositions of aerosol particles in V_{SPECT} for severe COPD patients [16]. The current patient had severe COPD as described above in the section 2.1. Strong correlations between ventilation and perfusion have been demonstrated [17, 18] (e.g., 0.89 [17]), and therefore perfusion can be used as surrogate. For the comparison, the CT image of SPECT was aligned with the peak-exhale 4D-CT image using rigid image registration (Fig. 1). The derived translations and rotations were applied to the SPECT images to align with the 4D-CT ventilation images. Given that V_{SPECT} was acquired after Q_{SPECT} and ^{99m}Tc-MAA activity remained on V_{SPECT}, the activity-corrected Q_{SPECT} (i.e., Q'_{SPECT}) was subtracted from V_{SPECT} to determine a net V (i.e., V_{SPECT}^{net}). The activity correction for Q_{SPECT} was simply based on an exponential decrease with a physical half-life of ^{99m}Tc (i.e., 6 h) and a biological half-life of MAA (i.e., 5 h [19]). After registration, the Spearman's rank correlation coefficients were determined to quantify the correlation between $V_{4D\text{-}CT}^{HU}$ or $V_{4D\text{-}CT}^{Jac}$ and V_{SPECT}^{net} or Q_{SPECT}. For the correlation analysis, the 4D-CT ventilation images were downsampled to the SPECT image resolution.

3 Results

Fig. 2 shows example images of 4D-CT, $V_{4D\text{-}CT}^{HU}$, $V_{4D\text{-}CT}^{Jac}$, V_{SPECT}^{net} and Q_{SPECT}. The V and Q images are shown with a scale from the 10th percentile value to the 90th percentile value. The percentile values were determined for the V or Q values inside the lungs. The $V_{4D\text{-}CT}$ images provided a higher resolution than V_{SPECT}^{net} or Q_{SPECT}. In the comparison between $V_{4D\text{-}CT}^{HU}$ or $V_{4D\text{-}CT}^{Jac}$ and V_{SPECT}^{net}, there were several agreements in some regions and many disagreements throughout the lungs. Overall, all V images showed higher ventilation in the right lung than in the left lung, which would be due to a large tumor mass near the left hilum compressing large bronchi. For $V_{4D\text{-}CT}^{HU}$, we found reasonable agreements in the right lower region at mid-anterior (low ventilation) and in the right upper region at mid-posterior and posterior

(high ventilation) for example. However, disagreements were more obvious and were found throughout the lungs, e.g., $V_{4D\text{-}CT}^{HU}$ showed moderate ventilation in the left lung at anterior and mid-anterior, where V_{SPECT}^{net} showed high ventilation. In addition, we found disagreements in periphery, where $V_{4D\text{-}CT}^{HU}$ showed higher ventilation than V_{SPECT}^{net}. Presumably these disagreements are due to central airway depositions of aerosol particles in V_{SPECT}^{net} as observed at mid-posterior, which would underrepresent peripheral ventilation. The Spearman voxel-based correlation coefficient between $V_{4D\text{-}CT}^{HU}$ and V_{SPECT}^{net} was 0.030. Similarly $V_{4D\text{-}CT}^{Jac}$ also had many disagreements with V_{SPECT}^{net} and the correlation was 0.184. Fig. 3 shows the voxel-based relationships between $V_{4D\text{-}CT}^{HU}$ or $V_{4D\text{-}CT}^{Jac}$ and V_{SPECT}^{net} or Q_{SPECT}. There were considerable probabilities of negative 4D-CT ventilation values that are not supposed to appear during inhalation in principle, as observed by Christensen et al. [9]. This leaves open the question of DIR algorithm errors.

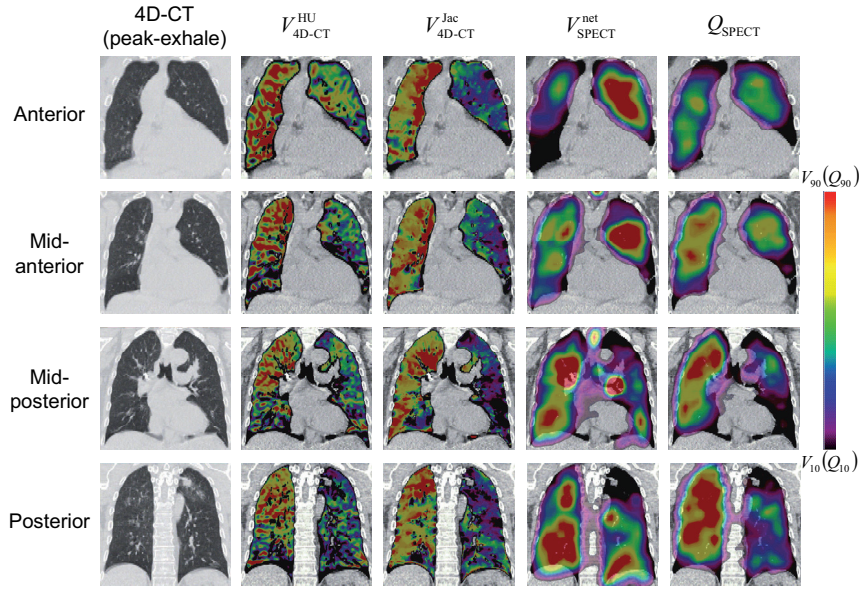


Fig. 2. Example images at the same coronal level of peak-exhale 4D-CT, 4D-CT ventilation derived by the Hounsfield unit (HU) metric ($V_{4D\text{-}CT}^{HU}$) or the Jacobian metric ($V_{4D\text{-}CT}^{Jac}$), the net SPECT ventilation (V_{SPECT}^{net}), and perfusion (Q_{SPECT}). Each row represents different coronal level. The ventilation and perfusion images are shown with a scale from the 10th percentile value (V_{10} or Q_{10}) to the 90th percentile value (V_{90} or Q_{90}). The xth percentile value is the ventilation (perfusion) value at which x% of the total lung volume has lower value. Note that the $V_{4D\text{-}CT}$ values outside the lungs are masked out, while V_{SPECT}^{net} and Q_{SPECT} are not masked out to show central airway depositions of aerosol particles in the V_{SPECT}^{net} image.

In the comparison between $V_{4D\text{-}CT}^{HU}$ or $V_{4D\text{-}CT}^{Jac}$ and Q_{SPECT}, there were both agreements and disagreements as with V_{SPECT}^{net}, however there were clearly more agreements for Q_{SPECT} compared to V_{SPECT}^{net} from visual inspection (Fig. 2). The Q_{SPECT}

image also showed higher perfusion in the right lung than in the left lung similarly to the V images, which would be due to a large tumor compressing large vessels. For $V_{4D\text{-}CT}^{HU}$, we found reasonable agreements in the left upper region at anterior and mid-anterior (moderate ventilation) and in the right upper region at mid-posterior (high ventilation) for example. Disagreements were found in the right lung at anterior for example, where $V_{4D\text{-}CT}^{HU}$ showed high ventilation, however Q_{SPECT} showed low to moderate ventilation. The Spearman voxel-based correlation coefficient between $V_{4D\text{-}CT}^{HU}$ and Q_{SPECT} was 0.357 (Fig. 3). $V_{4D\text{-}CT}^{Jac}$ also demonstrated better agreements with Q_{SPECT} than with V_{SPECT}^{net} and the correlation coefficient was 0.479.

Two different ventilation metrics led to spatially variant 4D-CT ventilation images ($V_{4D\text{-}CT}^{HU}$ and $V_{4D\text{-}CT}^{Jac}$) (Fig. 2). From visual inspection, there were good agreements in the right upper region at mid-posterior (high ventilation) and in the left lung at posterior (low ventilation). However, $V_{4D\text{-}CT}^{HU}$ showed low ventilation in the right lower region at mid-anterior and mid-posterior, where $V_{4D\text{-}CT}^{Jac}$ showed high ventilation in contrast. The Spearman voxel-based correlation coefficient between $V_{4D\text{-}CT}^{HU}$ and $V_{4D\text{-}CT}^{Jac}$ was 0.459. Also, $V_{4D\text{-}CT}^{HU}$ demonstrated a more heterogeneous distribution with a coefficient of variation (i.e., the ratio of the standard deviation to the mean) of 0.160 than $V_{4D\text{-}CT}^{Jac}$ with 0.072, which is reflected in more scattered distributions in Fig. 3.

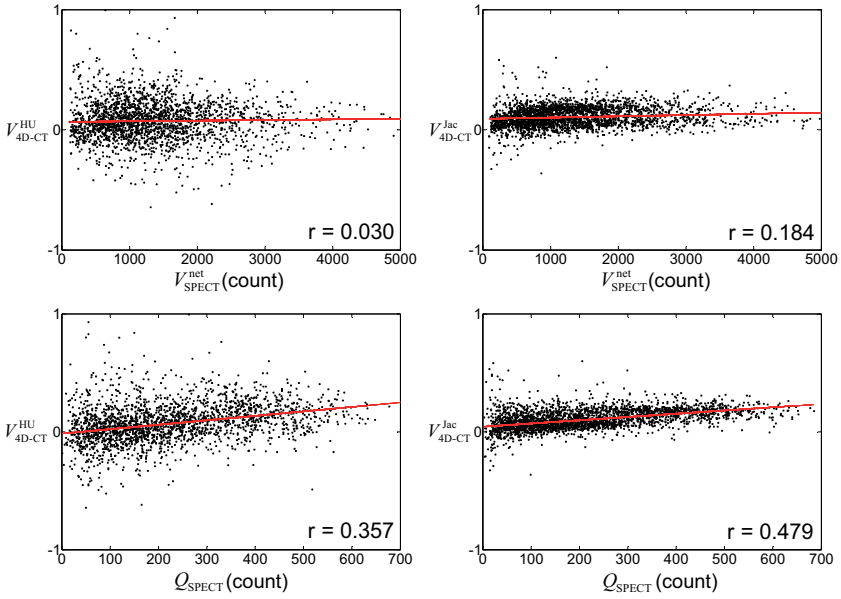


Fig. 3. Four-dimensional (4D) computed tomography (CT) ventilation vs. single photon emission CT (SPECT) ventilation or perfusion. The Spearman correlation coefficients and least square regression lines are also shown.

4 Discussion

The correlations between the 4D-CT ventilation and SPECT ventilation (assumed ground truth) were weak, which is consistent with Castillo et al. [6]. The weak correlations could be due to limitations of the 4D-CT and/or SPECT techniques. The limitations of the 4D-CT ventilation imaging technique include (1) a potential bias in the HU metric, (2) the need for the physiologic validation, and (3) the artifacts in 4D-CT images. There is a possibility that the HU metric underestimates the actual HU change, given that the DIR algorithm used in this study minimizes the sum of squared differences, however this effect would be small considering that it is also based on the elastic regularizer. Our recent study demonstrated significantly lower ventilation in emphysematous lung regions than in non-emphysematous regions ($p < 0.001$), which indicates a potential of 4D-CT ventilation imaging to achieve the high physiologic accuracy [13]. However, this study focused on the global accuracy only. Artifacts are observed in the 4D-CT images at an alarmingly high frequency (i.e., 90%) [20], which could impact 4D-CT ventilation imaging. Although we used AV biofeedback to minimize the artifacts, several artifacts still remained in the images. There are several more strategies to improve 4D-CT, including a wide-area detector CT scanner with a greater coverage [21], improved sorting methods [22, 23], and post-scan image processing [24, 25]. The limitations of the SPECT ventilation scan with ^{99m}Tc-DTPA aerosols include central airway depositions in severe COPD patients [16]. That was the case in the current study as the patient had severe COPD. Central deposition limits the uniform penetration of aerosol particles to periphery, yielding non-representative SPECT ventilation [26, 27]. In a recent study, an ultrafine dispersion of technetium-99m-labeled carbon (i.e., Technegas) has been found to be superior to ^{99m}Tc-DTPA for COPD patients [26]. In addition to these limitations of the 4D-CT and SPECT techniques, several differences between these two techniques might lead to the weak correlation. The differences include the spatial resolution, patient positioning and respiratory phase. For the difference in the resolution, appropriate smoothing for creating a 4D-CT ventilation image may result in an improved correlation. The patient received the 4D-CT and SPECT scans separately on the different scanners, where there was a considerable difference in positioning even though we tried to minimize it. We would not be able to disregard this difference unless a SPECT/CT scanner with the 4D-CT capability is used to acquire both 4D-CT and SPECT scans. The difference in the respiratory phase is another issue, i.e., the 4D-CT ventilation was based on the peak-exhale and peak-inhale phases only, however the SPECT ventilation was acquired under free breathing. Integrating the 4D-CT ventilation created from all the 4D-CT images over a full respiratory cycle would more appropriately represent free breathing and may improve a correlation.

The correlations between the 4D-CT ventilation and SPECT perfusion were much higher than those with the SPECT ventilation. However, they were still weak, which could be due to a potential difference between ventilation and perfusion in addition to limitations of the 4D-CT and/or SPECT techniques as described above. It has been demonstrated that regional ventilation is tightly matched to perfusion in large laboratory animals under normal conditions [17, 18], even though its mechanisms

remain unresolved. Altemeier found a strong correlation between ventilation and perfusion in pigs with an average correlation coefficient of 0.89 [17]. There are several potential causes degrading this matching, which include pathological conditions and mechanical ventilation [28]. Mismatch between ventilation and perfusion occurs mainly due to pulmonary vascular disease such as pulmonary embolism and hypoplasia/aplasia of pulmonary artery [19]. In this study, the patient had no clear evidence of such diseases, however had a lung cancer and COPD that might degrade the ventilation and perfusion matching. Another difference may come from the effect of posture and gravity on the relationship between ventilation and perfusion [29, 30]. Recently, Petersson et al. investigated this effect using the SPECT V/Q scan for 7 healthy volunteers in a supine posture [29]. They found more uniform V/Q ratios close to 1 in the superior-inferior direction compared to an upright posture, however there were slight increases up to approximately 1.3 in the inferior regions.

The 4D-CT ventilation imaging technique has higher spatial and temporal resolution, lower cost, shorter scan time, and/or higher accessibility compared with other techniques (including nuclear medicine, hyperpolarized gas MR, and Xe-CT). The 4D-CT ventilation can be considered as 'free' information, because 4D-CT scans are in routine use for lung cancer radiotherapy in many centers, and ventilation computation involves only image processing (DIR) and analysis. Furthermore, 4D-CT ventilation imaging has an inherent advantage, i.e., perfect registration to the original CT images, which enables high-resolution anatomic correlation. Once validated, 4D-CT ventilation images could be used in lung cancer radiotherapy to avoid high-functional lung regions from radiation that may allow pulmonary toxicity reduction [31, 32]. Also, the functional response to radiotherapy treatment could be investigated with 4D-CT ventilation imaging.

5 Conclusion

The physiologic accuracy of 4D-CT ventilation imaging has been quantitatively evaluated on a human subject by comparison with the SPECT/CT V/Q scan. The correlation between the 4D-CT ventilation and SPECT V was very weak, where the SPECT V showed significant central airway depositions of aerosol particles. The correlation between the 4D-CT ventilation and SPECT Q was much higher, which indicates a potential of 4D-CT ventilation imaging to achieve the high physiologic accuracy. A further study is necessary to investigate the differences between the 4D-CT ventilation and SPECT V/Q and to improve the accuracy through acquiring more patient data including non-COPD patients.

Acknowledgments. Julian Hong, Lindee Burton, and Jayesh Patel at Stanford University provided significant assistance with the patient accrual process and the study coordination. Dr. Guoqiang Cui at Stanford assisted with AV biofeedback for the 4D-CT scan.

References

1. Suga, K.: Technical and analytical advances in pulmonary ventilation SPECT with xenon-133 gas and Tc-99m-Technegas. Ann Nucl Med **16** (2002) 303-310
2. Fain, S.B., Korosec, F.R., Holmes, J.H., O'Halloran, R., Sorkness, R.L., Grist, T.M.: Functional lung imaging using hyperpolarized gas MRI. J Magn Reson Imaging **25** (2007) 910-923
3. Simon, B.A.: Regional ventilation and lung mechanics using X-Ray CT. Acad Radiol **12** (2005) 1414-1422
4. Guerrero, T., Sanders, K., Noyola-Martinez, J., Castillo, E., Zhang, Y., Tapia, R., Guerra, R., Borghero, Y., Komaki, R.: Quantification of regional ventilation from treatment planning CT. Int J Radiat Oncol Biol Phys **62** (2005) 630-634
5. Simon, B.A.: Non-invasive imaging of regional lung function using x-ray computed tomography. J Clin Monit Comput **16** (2000) 433-442
6. Castillo, R., Castillo, E., Martinez, J., Guerrero, T.: Ventilation from four-dimensional computed tomography: density versus Jacobian methods. Phys Med Biol **55** (2010) 4661-4685
7. Kabus, S., von Berg, J., Yamamoto, T., Opfer, R., Keall, P.J.: Lung ventilation estimation based on 4D-CT imaging. Proc. of the First International Workshop on Pulmonary Image Analysis, MICCAI 2008 (2008) 73-81
8. Reinhardt, J.M., Ding, K., Cao, K., Christensen, G.E., Hoffman, E.A., Bodas, S.V.: Registration-based estimates of local lung tissue expansion compared to xenon CT measures of specific ventilation. Med Image Anal **12** (2008) 752-763
9. Christensen, G.E., Song, J.H., Lu, W., El Naqa, I., Low, D.A.: Tracking lung tissue motion and expansion/compression with inverse consistent image registration and spirometry. Med Phys **34** (2007) 2155-2163
10. Guerrero, T., Sanders, K., Castillo, E., Zhang, Y., Bidaut, L., Pan, T., Komaki, R.: Dynamic ventilation imaging from four-dimensional computed tomography. Phys Med Biol **51** (2006) 777-791
11. Brock, K.K.: Results of a Multi-Institution Deformable Registration Accuracy Study (MIDRAS). Int J Radiat Oncol Biol Phys **76** (2009) 583-596
12. Kashani, R., Hub, M., Balter, J.M., Kessler, M.L., Dong, L., Zhang, L., Xing, L., Xie, Y., Hawkes, D., Schnabel, J.A., McClelland, J., Joshi, S., Chen, Q., Lu, W.: Objective assessment of deformable image registration in radiotherapy: a multi-institution study. Med Phys **35** (2008) 5944-5953
13. Yamamoto, T., Kabus, S., von Berg, J., Klinder, T., Blaffert, T., Lorenz, C., Keall, P.J.: Physiological validation of 4D-CT-based ventilation imaging in patients with chronic obstructive pulmonary disease (COPD). Med Phys **36** (2009) 2821 (Abstract)
14. Rietzel, E., Pan, T., Chen, G.T.: Four-dimensional computed tomography: image formation and clinical protocol. Med Phys **32** (2005) 874-889
15. Venkat, R.B., Sawant, A., Suh, Y., George, R., Keall, P.J.: Development and preliminary evaluation of a prototype audiovisual biofeedback device incorporating a patient-specific guiding waveform. Phys Med Biol **53** (2008) N197-N208
16. Magnant, J., Vecellio, L., de Monte, M., Grimbert, D., Valat, C., Boissinot, E., Guilloteau, D., Lemarie, E., Diot, P.: Comparative analysis of different scintigraphic approaches to assess pulmonary ventilation. J Aerosol Med **19** (2006) 148-159

17. Altemeier, W.A., Robertson, H.T., Glenny, R.W.: Pulmonary gas-exchange analysis by using simultaneous deposition of aerosolized and injected microspheres. J Appl Physiol **85** (1998) 2344-2351
18. Melsom, M.N., Kramer-Johansen, J., Flatebo, T., Muller, C., Nicolaysen, G.: Distribution of pulmonary ventilation and perfusion measured simultaneously in awake goats. Acta Physiol Scand **159** (1997) 199-208
19. Ziessman, H.A., O'Malley, J.P., Thrall, J.H.: Nuclear Medicine: The Requisites. A Mosby Title (2005)
20. Yamamoto, T., Langner, U., Loo, B.W., Jr., Shen, J., Keall, P.J.: Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients. Int J Radiat Oncol Biol Phys **72** (2008) 1250-1258
21. Mori, S., Endo, M., Tsunoo, T., Kandatsu, S., Tanada, S., Aradate, H., Saito, Y., Miyazaki, H., Satoh, K., Matsushita, S., Kusakabe, M.: Physical performance evaluation of a 256-slice CT-scanner for four-dimensional imaging. Med Phys **31** (2004) 1348-1356
22. Rietzel, E., Chen, G.T.: Improving retrospective sorting of 4D computed tomography data. Med Phys **33** (2006) 377-379
23. Lu, W., Parikh, P.J., Hubenschmidt, J.P., Bradley, J.D., Low, D.A.: A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4D CT. Med Phys **33** (2006) 2964-2974
24. Ehrhardt, J., Werner, R., Saring, D., Frenzel, T., Lu, W., Low, D., Handels, H.: An optical flow based method for improved reconstruction of 4D CT data sets acquired during free breathing. Med Phys **34** (2007) 711-721
25. Schreibmann, E., Chen, G.T., Xing, L.: Image interpolation in 4D CT using a BSpline deformable registration model. Int J Radiat Oncol Biol Phys **64** (2006) 1537-1550
26. Jogi, J., Jonson, B., Ekberg, M., Bajc, M.: Ventilation-perfusion SPECT with 99mTc-DTPA versus Technegas: a head-to-head study in obstructive and nonobstructive disease. J Nucl Med **51** (2010) 735-741
27. Agnew, J.E., Francis, R.A., Pavia, D., Clarke, S.W.: Quantitative comparison of 99Tcm-aerosol and 81Krm ventilation images. Clin Phys Physiol Meas **3** (1982) 21-30
28. Glenny, R.W.: Determinants of regional ventilation and blood flow in the lung. Intensive Care Med **35** (2009) 1833-1842
29. Petersson, J., Rohdin, M., Sanchez-Crespo, A., Nyren, S., Jacobsson, H., Larsson, S.A., Lindahl, S.G., Linnarsson, D., Neradilek, B., Polissar, N.L., Glenny, R.W., Mure, M.: Regional lung blood flow and ventilation in upright humans studied with quantitative SPECT. Respir Physiol Neurobiol **166** (2009) 54-60
30. Musch, G., Layfield, J.D., Harris, R.S., Melo, M.F., Winkler, T., Callahan, R.J., Fischman, A.J., Venegas, J.G.: Topographical distribution of pulmonary perfusion and ventilation, assessed by PET in supine and prone humans. J Appl Physiol **93** (2002) 1841-1851
31. Yamamoto, T., Kabus, S., von Berg, J., Lorenz, C., Keall, P.J.: Impact of four-dimensional computed tomography pulmonary ventilation imaging-based functional avoidance for lung cancer radiotherapy. Int J Radiat Oncol Biol Phys **In press** (2010)
32. Yaremko, B.P., Guerrero, T.M., Noyola-Martinez, J., Guerra, R., Lege, D.G., Nguyen, L.T., Balter, P.A., Cox, J.D., Komaki, R.: Reduction of normal lung irradiation in locally advanced non-small-cell lung cancer patients, using ventilation images for functional avoidance. Int J Radiat Oncol Biol Phys **68** (2007) 562-571

Computer Aided Detection for Pneumoconiosis Screening on Digital Chest Radiographs

Horace Xu[1], Xiaodong Tao[2], Ramasubramanian Sundararajan[3], Weizhong Yan[2], Pavan Annangi[3], Xiwen Sun[4], Ling Mao[4]

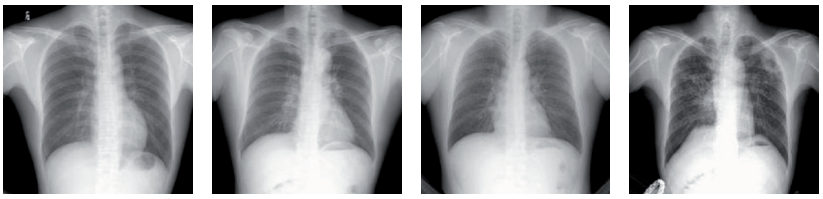
[1] GE Global Research Shanghai, China
[2] GE Global Research Niskayuna, USA
[3] GE Global Research Bangalore, India
[4] Shanghai Pulmonary Hospital, China

Abstract: This paper presents a computer aided detection scheme on digital chest radiographs for pneumoconiosis screening. The scheme involves several medical image processing and analysis technologies, i.e. lung segmentation algorithm using the active shape model, image enhancement and features extraction from lung regions, feature down-selection by correlation analysis and clustering method, classification on lung regions by multi-scale classifiers support vector machine, and finally, chest based reporting out on the classification probabilities. Experiments are conducted on our digital chest x-ray database, and the result shows a good classification performance for screening application in clinic.

Keywords: **C**omputer aided detection (CAD), active shape model (ASM), support vector machine (SVM), pneumoconiosis, digital radiography (DR).

1 Introduction

Pneumoconiosis is a lung disease caused by a long-term inhalation of operative dust, such as coal, asbestos and silica, and the local tissue reaction to the accumulated dust particles [1]. Chest radiography is the most practical tool according to the guidelines of China and International Labor Organizer (ILO) standard for pneumoconiosis diagnosis, which is based on the assessment of opacities on chest X-Ray (CXR) films and comparison with standard films. The opacities shape can be round or irregular, and the number of opacities in each lung zones can be described by a discrete score from 0/-, 0/0, 0/1, 1/0, 1/1, 1/2, 2/1, 2/2, 2/3, 3/2, 3/3, and 3/+, corresponding to twelve possible categories, which is named profusion level. Based on the perfusion level of the opacities on a patient's CXR, a stage between 0 and III is assigned to the patient to indicate the severity of the disease [2]. Figure 1 shows the X-ray images of normal chest and pneumoconiosis chests in different stage.



(a) (b) (c) (d)

Fig. 1. Normal and pneumoconiosis chest X-ray images. (a) normal, (b) stage I, (c) stage II, (d) stage III.

In China, pneumoconiosis holds ~90% occupational diseases. The number of prevalence cases of pneumoconiosis is 638,234, up to the end of 2008, while there are over 610,000 suspicious cases, reported by Ministry of Health of the Peoples Republic of China [3]. However, the diagnosis of pneumoconiosis, which is based on the assessment of subtle opacities on chest X-rays films and comparison with standard films, is a challenge job in clinic. The diagnosis process is not only time consuming, subjective in staging, high experience depended, but also big variant in intra-observer and inter-observer. The radiologists' study in China shows that the intra-observer variance is ~12.5%, and inter-observer variance is ~22.5%[1], which will cause different compensation to patients as well as the treatment.

In this paper, we present a computer aided detection (CAD) tool to detect the pneumoconiosis in digital Posterior-Anterior (PA) chest radiographs by image analysis automatically, which follows the process of lung segmentation, feature extraction in lung fields and sub-zones, classification and chest based report out. The CAD tool will help the radiologists in screening application as the second view, which will improve the work flow in mass screening, reduce the work load in x-ray images reading, and keep a consistent accuracy in pneumoconiosis diagnosis.

2 Lung Segmentation and Sub-division

Segmenting the lungs from digital chest x-ray image is to get the lung fields for the following automatic analysis. Also, the automatic segmentation of lung fields from chest radiographs can be a useful tool for the morphology analysis based detection of abnormalities.

2.1 Active Shape Model Segmentation

Lung field segmentation on chest radiographs has attracted a number of researchers. Rule based schemes have been investigated by Ginneken and Romeny [4]. Ginneken [5] proposed a new hybrid segmentation scheme that combined the strengths of a rule-based approach and of a pixel classification approach. Vittitoe [6] developed a pixel classifier for the identification of lung regions using Markov random field modeling. Ginneken et al.[7] put forth a lung field algorithm based on ASM method with optimal features. Iglesias [8] investigated lung segmentation based on the detection of oriented edges and active contours models. As the efficiency and

accuracy are required in the segmentation process, ASM algorithm is employed to delineate the lung fields in this paper.

The ASM algorithm could be mainly divided into two steps, training and segmentation.

In the training step, a procedure is to be followed to build up the model, i.e. labeling lung contour landmarks on training images; aligning manual landmark shapes by scaling, rotating and translating; calculating the statistics of a set of aligned shapes; constructing the shape model.

The second step is to delineate the lung contour for an input image. The procedure includes initializing the process with the average shape obtained from training phase; searching the optimal point along the direction perpendicular to the contour for each landmark in the initial contour; finding the best fitted lung field shape and update the parameters and the shape model; repetition until the process converges.

Here, we extract a seed region of lung by histogram equalization and Ostu thresholding, and transform the average shape model with scaling and translating before searching the lung contour. The initial location of lung and the transfer of the average shape improve the accuracy of ASM in lung segmentation. Figure 2 shows the result of lung field segmentation with ASM algorithm.

Based on the segmentation result, we calculate couple of parameters as morphology analysis, i.e. the symmetry of left/right lung fields by comparing their areas, the costophrenic angle which illuminates pleural effusion, and the cardiothoracic ratio to detect cardiomegaly.

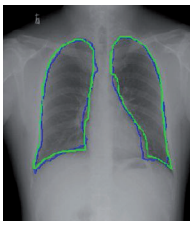


Fig. 2. The result of lung field segmentation using ASM algorithm. (Green contour is the ground truth.)

2.2 Subdividing the Lung Fields into Six Zones

According to the domain knowledge of pneumoconiosis diagnosis, the texture features are extracted from multiple regions to better capture the features of the diffuse abnormalities. On the basis of ILO guidance and radiologists' diagnostic measurement, six non-overlapping regions are subdivided in the lung fields, as shown in Figure 3. Each region will be analyzed with a separate classifier that is built by the features extracted solely from this region. The goal is able to capture the feature of abnormal variations in that particular part; in principle, this will reduce the influence of superimposed anatomical structures, which tend to make the abnormalities difficult to distinguish.

Given a lung field mask image, each lung field is subdivided into three zones - the upper, middle and lower zone. Right lung field subdivision is done by calculating the apex and diaphragm position and computing horizontal lines that divide the field into three parts of equal height. However, as the large variability of left lung shapes leads to difficulty in obtaining diaphragm position, the left lung field subdivision differs from the right lung field. Left lung field is subdivided by calculating the apex and bottom position and computing horizontal lines that divide the field into three parts of equal height. The outer point of the diaphragm is the bottom outer point of each lung field, and the inner point is found by tracing inward along the lung contour till there is a maximum change in curvature.

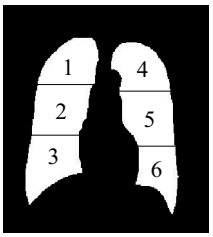


Fig. 3. The lung fields are subdivided into six regions.

3 Feature Extraction and Selection

Once a region of interest (ROI) – each of 6 zones and two lung fields – is segmented, the next task is to characterize it in terms of a set of features. Here, we are following two procedures to achieve the goal. First, we extract three types of features from each ROI, totally 247 features. Second, we select a set of features for the following classification by correlation analysis and clustering method.

3.1 Feature Extraction

1. Image enhancement by difference filtering

To enhance various size and various degree of opacities in pneumoconiosis chest, a multi-scale difference filter bank $L_n^{\theta}(x, y, d)$ is developed to improve the image contrast, where n, θ, and d denote the order of difference, the orientation in which the difference is computed and the difference scale, respectively.

2. Image intensity based feature

We extract a set of 6 features based on the histogram of intensity values – mean, standard deviation, skewness, kurtosis, energy and entropy.

Apart from calculating these on the original ROI, we also extract these features after applying difference filters on the image, where, two scales {1,2}, nine orientations {0˚,30˚,45˚,60˚,90˚,120˚,135˚,150˚,180˚} and the first and second order are used.

3. Co-occurrence matrix based feature

We extract a set of 5 features based on the gray level co-occurrence matrix computed for the ROI, namely energy, entropy, local homogeneity, correlation and inertia. The co-occurrence matrix allows us to capture the level of similarity and dissimilarity among adjacent pixels in an ROI. Thus, an ROI with opacity will contain adjacent pixels with similarly high intensities, whereas a normal ROI will not contain such adjacent pixels. Computing these features for various orientations captures this information for various types of adjacency. In this paper, we use 4 orientations {0˚,45˚,90˚,135˚} to construct the co-occurrence matrix.

4. Frequency domain based feature

We also extract a set of 5 features based on a frequency domain transformation of the ROI. These features are the mean, root mean square (RMS), first moment of power spectrum (M1), horizontal filter and vertical filter.

In total, there are 247 features extracted as described above. Given the number of data points used for model building, this number is too large and needs to be pruned. We shall discuss the methods used for feature down-selection in the following subsection.

3.2 Feature Selection

We select a subset of features for classification by analyzing similarities (such as correlation) and eliminating redundancy. The objective is to retain the features that achieve maximum compression of the data with minimum information loss. However, the real measure of the selection method is simply whether the set of features provides a good enough description of the ROIs to allow for a good model to be built to classify ROI into twelve categories of profusion level, which is described in section 1. Therefore, we build classifiers with feature subsets of various sizes, and choose the smallest feature set where the classifier performance is acceptable.

Consider a data set with m rows (each representing an ROI) and n columns (each representing a feature that describes the ROI). We can now create an n*n matrix of values, where the value in cell (i, j) represents the similarity (absolute value of correlation) between feature *i* and feature *j* in the dataset.

If we consider each feature as a "point" in some space and the appropriate entry in this matrix as the similarity between pairs of points, we can cluster these points to arrive at groups of similar features. We perform clustering of features using a method known as hierarchical clustering [9]. First, features are grouped into a dendrogram, which is a tree where elements close to each other are clubbed together into branches. Depending upon the criterion provided by the user (number of clusters, distance between clusters etc), one can split the tree into branches and thus arrive at clusters. We can use this method to arrive at groups of similar features from the given dataset. Once these groups are given, we choose one feature from each group as a representative.

4 Classification

From machine learning point of view, pneumoconiosis detection is a binary classification problem, where classification is to determine a subject has or not has pneumoconiosis. This is the purpose of screening for workers under high risk of dust. On the other hand, pneumoconiosis staging, to determine the stage (or severity) of a patient's pneumoconiosis, is a multi-class classification problem. This can be regarded as the activity for quantitative diagnosis. This stage information coming out of pneumoconiosis staging is crucial because the determined stage is used as both legal evidence for workman's accident compensation and the means for selecting appropriate treatments.

4.1 Classification for Pneumoconiosis Screening

Normally, the task of building a classifier based on the available data can be viewed in terms of two major sub-tasks: finding the optimal level of complexity for the classifier, and building the classifier at the specified level of complexity.

For screening application, we build up 8 classifiers for left/right lungs, and 6 zones by sub-division. The methodology used for building the models is described in [10], and is outlined here:

1. Search through the space of possible parameter values that describe classifier complexity.

2. For each possible set of parameter values, build and evaluate the performance of the classifier. We perform this evaluation using the k-fold cross validation method, whereby a dataset is divided into k parts, *k-1* of which are used for training and the k^{th} is used for validation. This is done in k different ways, and the average performance across these validation cycles is reported as the performance of the classifier at that level of complexity.

3. Pick the parameter values where the best performance is achieved and build the final classifier using those values. The final classifier performance is evaluated using a method known as leave-one-out validation, which works the same way as k-fold validation but with k equal to the number of rows in the dataset. The leave-one-out performance is considered to be a good measure of the generalization ability of a classifier [11].

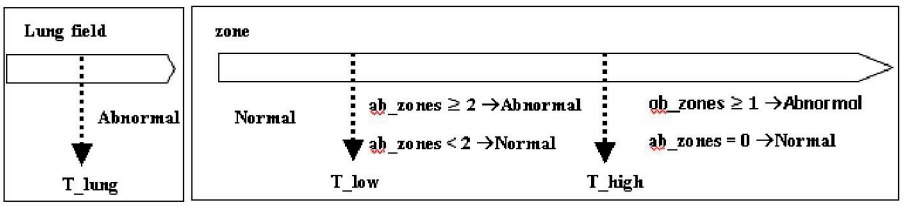


Fig. 4. The strategy of chest classification from two levels.

Once models for lung and zone level ROIs have been built, we evolve the chest based final report by leveraging the domain knowledge of pneumoconiosis screening,

for example, the classification probability of one lung ROI is larger than a threshold, the chest will labeled as abnormal, but for zone ROIs, the final conclusion is generated from how many abnormal zones and how much the abnormal zone's probability. The strategy is shown in Figure 4. The thresholds are optimized by training in database.

For all the individual classifiers, we use Support Vector Machines (SVM), a machine technique pioneered by Vladimir Vapnik [11] and Nello Cristianini [12], and now extensively used in a number of applications. An SVM attempts to construct a hyperplane in a high dimensional space that has the largest distance (maximal margin) to the examples in the training set close to the hyperplane. Complexity is scaled by transforming the original features using an appropriate kernel function.

A key design choice to be made in building SVMs is the selection of the kernel to compute the similarity between two examples. We use the Radial Basis Function (RBF) kernel in our case, and control the complexity by varying the gamma(γ) parameter. The other parameter we control is the penalty for the error term C. These two parameters determine the complexity of the classifier being built. In order to choose the right level of complexity as defined by these two parameters, we measured the performance on unseen examples of an SVM model built at a particular complexity level using the k-fold validation technique.

4.2 Classification Study for Pneumoconiosis Staging

The pneumoconiosis detection problem has been actively studied. However, work on pneumoconiosis staging is very sparse, in spite of its importance.

Since the pneumoconiosis stages (0, I, II, III) are discrete categories ordered by disease severity, to be more exact, the pneumoconiosis staging problem is an ordinal classification problem. Ordinal classification is a special machine-learning task that falls between classification and regression [13]. Unlike nominal classification, where the target classes are un-ordered and classification performance is measured by classification accuracy, for ordinal classification both accuracy and the distance between the actual and predicted classes are relevant in evaluating classification performance. Hence, an important design strategy in developing ordinal classifiers is to leverage the ordering information to minimize the distance between actual and predicted classes.

In literature there are several approaches towards addressing ordinal classification problems. The simplest approach would be to solve ordinal problems by treating them as un-ordered nominal classification problems. This "passive" approach ignores the ordering information that could be used to improve the classification performance. More sophistical approaches have bee proposed in recent years, for example, [14-17].

In this paper, we propose using SVM for pneumoconiosis staging. In our model, the inputs are the selected features and outputs are the pneumoconiosis stages. To address the ordinal classification problem, we explore several design strategies, including categorical, ordinal, and regression.

5 Experiments and Results

In this section, we introduce the experiments by applying the algorithms described above on a database of digital chest X-Ray images comprising normal and pneumoconiosis chest x-ray images.

Normal chest images and pneumoconiosis chest images were collected for building up the database from clinic. The normal chest images are categorized by gender and age, while the pneumoconiosis chest images are categorized by stage, from 0+ to III.

A total of 427 images were collected, of which 252 were of normal subjects and 175 were of patients afflicted with pneumoconiosis.

A lung segmentation result is shown in Figure 2. Ground truth for segmentation validation was available on 400 images from this database. The performance of the algorithm was evaluated on the basis of DICE similarity coefficient (DSC) between the segmented lung mask and the ground-truth mask. DICE coefficient is computed as twice the area overlap (logical AND) of the two masks divided by the sum of the areas of the masks (logical OR). On the 400 cases tested, 89% of the images have > 85% DSC, and about 96% cases have DSC > 80%.

Of the 247 features extracted at each resolution, a smaller subset was selected for classification. At the lung level, the correlation threshold-based selection method was employed with a threshold of 80% to select 38 features. At the lung zone level, the hierarchical clustering method was used to select between 12 ~ 14 features for each zone. 2 models were built at the lung level (one for each lung) and 6 at the zone level (one for each zone). The feature list and the parameters of model are provided in Table 1 and 2 as example.

Table 1. Lung models

Lung	Left	Right
Log2 (C)	5	5
Log2 (Gamma)	-5	-7
Variables	'0_0_0_Mean'	
	'0_0_0_Skewness'	
	'0_0_0_Kurtosis'	
	'0_0_0_Entropy'	
	'2_45_1_Entropy'	
	……	

Table 2. Right lung zone models

Zone	Right Upper	Right Middle	Right Lower
Log2 (C)	15	1	5
Log2 (Gamma)	-9	-1	-3
Variables	2_0_2_Mean	0_0_0_Energy	1_180_2_Mean

	2_150_1_Energy	Deg90_Energy	2_150_1_Energy
	1_180_2_Mean	1_120_2_Skewness	0_0_0_Mean
	0_0_0_Mean	2_180_2_Entropy	Deg90_ Correlation
	…		

The outputs of these models were combined to get the final result for the chest according to the strategy in Figure 4. And the thresholds we used in the experiment are below.

Threshold of lung level is 0.5;

Threshold of zone level is 0.64;

Low threshold of sum probability of chest is 0.43;

High threshold of sum probability of chest is 0.54.

We used the LIBSVM package to build and test the SVM classifiers that were used in the ensemble. This package uses a variant of the Sequential Minimal Optimization Algorithm [18, 19] to build the SVM. An RBF kernel was used in the SVM, and the and C parameters which determine classifier complexity were both varied in the range $[2^{-15}; 2^{15}]$ in order to arrive at the model with optimal complexity and good generalization. Table 3 gives a summary of the results in the database.

Table 3. Experimental results in the database.

	Sensitivity	Specificity	Accuracy
Test result	99.5%	92.7%	95.5%

6 Conclusions and Discussion

Pneumoconiosis detection, especially for coal workers, has attracted many researchers in the past decades [20-22]. The computer scheme we present here improved the lung segmentation accuracy in chest radiography, extracted the features created a feature vector by feature down-selection, built up multiple classifiers on lung fields and subdivision zones, and reported out based on patient's chest according to the clinic domain knowledge.

Our future job is to build up multi-class classifiers on each 6 zones of lung to get quantitative diagnosis for pneumoconiosis staging, which is one of the most challenge tasks in clinic. Moreover, we plan to train the parameters in larger training set or test the computer scheme in clinic data set, which is to optimize the feature vector, avoid overtraining issue, and make the algorithms more robust.

Currently the tool is expected to provide value in occupational disease detection, including mass screening and quantitative staging applications. However, it should be possible to extend it to interstitial diseases with diffuse character, such as lung tuberculosis.

References

1. Li, Dehong: Physicians textbook of occupational diseases pneumoconiosis. Publishing House of Peoples Daily, Beijing (2004)
2. ILO: Guidelines for the use of the ilo international classification of radiographs of pneumoconiosis - revised edition. International Labor Office, (2000)
3. Xinhua News Agency: Report on Ministry of Health of the Peoples Republic of China. Beijing (2009)
4. Ginneken, BV., Romeny BM, Viergever MA.: Computer-aided diagnosis in chest radiography: a survey. IEEE Transctions on Medical Imaging 20(12): 1228-1241 (2001)
5. Ginneken, BV.: Computer-aided diagnosis in chest radiography. Medical Physics, 2001; 28(6): 1144-1150.
6. Vittitoe, NF., Vargas-Voracek, R. and Floyd, CE: Identification of lung regions in chest radiographs using Markov random field modeling. Medical Physics 25(6): 976-985, 1998
7. Ginneken, BV., Alejandro F. Frangi, Joes J. Staal: Active shape model segmentation with optimal features. IEEE Transctions on Medical Imaging 21(8): 924-933, 2002
8. Iglesias, I., Souto, M., Alegria, A.M.: Lung segmentation on postero-anterior digital chest radiographs using active contours. Lecture Notes in Computer Science 3138: 538-546, 2004
9. Duda, Richard O., Hart, Peter E. and Stork, David G.: Pattern Classification. Wiley-Interscience (2000)
10. Hsu, Chih-Wei, Chang, Chih-Chung and Chih-Jen Lin: A practical guide to support vector classification. http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html
11. Vladimir Vapnik: Statistical Learning Theory. John Wiley (1998)
12. Nello Cristianini and John Shawe-Taylor: An Introduction to Support Vector Machines. Cambridge University Press (2000)
13. Kramer, S., Widmer, G., Pfahringer, B. and DeGroeve, M.: Prediction of ordinal classes using regression trees. Fundamenta Informaticae (2001)
14. McCullagh, P.: Regression model for ordinal data. Journal of the Royal Statistical Society Series B (42) 109-142 (1980)
15. Potharst, J.R. and Bioch, J.C.: Decision trees for ordinal classification. Intelligent Data Analysis, Vol. 4(2), pp. 97-112 (2000)
16. Chu & Keerthi: New approaches to support vector ordinal regression. Proc. 22nd International Conference on Machine Learning, Bonn, Germany (2005)
17. Mathieson, M. J.: Ordinal models for neural networks, in Neural Networks in Financial Engineering. World Scientific, Singapore (1995)
18. Platt, J. C.: Sequential minimal optimization: A fast algorithm for training support vector machines. Tech. Rep. MSR-TR-98-14, Microsoft Research, 1998.
19. Fan, R.-E., Chen, P.-H. and Lin, C.-J.: Working set selection using second order information for training svm. Journal of Machine Learning Research, vol. 6, pp. 1889~1918, 2005.
20. Kruger, R., Thompson, W. and Turner, A.: Computer diagnosis of pneumoconiosis. IEEE Trans. Syst., Man, Cybern., vol. SMC-4, 1974.
21. Turner, A., Kruger, R. and Thompson, W.: Automated computer screening of chest radiographs for pneumoconiosis. Investigation Radiology, vol. 11, pp. 258--266, 1976.
22. Hiroshi Kondo and Takaharu Kouda: Computer-aided diagnosis for pnemoconiosis using neural network. The 14th IEEE Symposium on Computer-Based Medical Systems (CMBS'01), 2001.

Training a Computer Aided Detection System with Simulated Lung Nodules in Chest Radiographs

Peter R. Snoeren, Geert J.S. Litjens, Bram van Ginneken, and Nico Karssemeijer

Radboud University Nijmegen Medical Centre
Nijmegen, The Netherlands
P.Snoeren@rad.umcn.nl

Abstract. This paper addresses the hypothesis that artificially implanted lung nodules from computed tomography exams (CT exams) into chest radiographs can improve the performance of a computer aided detection system (CAD system). Twenty-four three-dimensional lung nodules were segmented and projected in five directions, mimicking 120 X-rayed nodules. The CAD system was tested by fivefold cross validation on a publicly available database. The results were evaluated by a free-response receiver operating characteristic analysis (FROC). It was found that the performance of the CAD system trained with simulated nodules comes close to the performance of state of the art CAD systems that are trained with real nodules. The CAD system trained with real nodules did improve by adding simulated nodules, but only when there were few real nodules used for training.

1 Introduction

Chest radiography is often the first study performed on symptomatic patients and on asymptomatic patients in screening. While it is a cost-effective imaging tool it has the drawback that images are hard to interpret, even for trained radiologists. After some decades of development computer aided detection (CAD) is becoming a widespread tool in medical diagnosis to assist radiologists. It is obvious that a radiologist in training improves when he has diagnosed more radiographs with feedback, with different kinds of diseases, with abnormalities on different locations, and under different circumstances. That is not very different for CAD systems. A CAD system's performance also improves when training data are richer in content and amount. A CAD system must first of all be well-designed, but it also needs to be trained with many normal and abnormal images before it can be applied successfully. Many papers have been written about the design of CAD systems [1,2,3,4], in most cases silently assuming that training with a large training database is possible. Unfortunately though, large training databases are rare. This paper describes the results of a feasibility study with artificially implanted lung nodules to increase the size of training databases.

Stratum	Number
Obvious	12
Relatively obvious	38
Subtle	48
Very subtle	24
Extremely subtle	19
Total	141

Table 1. Subtlety rating of abnormal JSRT images.

A small number of three-dimensional lung nodules were segmented in CT volumes. Projection of these volumes in different directions mimics X-rayed CT nodules that can be implanted in chest radiographs. This way a training database can be extended with abnormalities at will. Simulated nodules can for example be placed at locations with few training data, e.g., at locations that are partly obscured by other organs, for separately training the CAD system. They can be systematically varied in size and/or contrast by scaling or placed on the edges of ribs or just in between. To prove the concept, a single contrast and the real projected size of the simulated nodules were used in this paper. Furthermore, the location variation is limited by locations of contralateral real nodules on the same radiograph.

The following describes the segmentation of CT nodules, the construction of simulated nodules, the implanting of these simulated nodules in chest radiographs, and the CAD system that is used to test the hypothesis that implanted nodules can in principle improve CAD systems. An FROC analysis is used to objectively evaluate the CAD system with differently enriched training data.

2 Materials and Methods

2.1 Databases

Chest Radiography Database. The publicly available database of the Scientific Committee of the Japanese Society of Radiological Technology (JSRT, Shiraishi et al. [5]) was used for training and testing the CAD system. The performance of CAD systems on this database often serves as a benchmark for comparison between CAD systems. At present, it is the only publicly available database of chest radiographs of moderately large size with normals and proven abnormals (lung nodules). The JSRT database consists of 247 Posterior-Anterior PA chest radiographs, digitized at a spatial resolution of 175 μm/pixel and 12 bit gray levels. 154 images contain exactly one pulmonary lung nodule; the other 93 images contain no lung nodules. Additionally, a publicly available database of lung field segmentations of the JSRT-images was used. This database consists of manually segmented unobscured lung fields (only the parts that are unobscured by the diaphragm or heart), heart, and clavicles (van Ginneken et al. [6]). Thirteen images with lung nodules that were not part of the unobscured lung fields

(a) JSRT nodules

Stratum	Number
$< 10\,mm$	33
$\geq 10\,mm$ and $< 20\,mm$	82
$\geq 20\,mm$	26
Total	141

(b) Simulated nodules

Stratum	Number
$< 10\,mm$	6
$\geq 10\,mm$ and $< 20\,mm$	68
$\geq 20\,mm$	43
Total	117

Table 2. Diameter of Lung Nodules. The total number of implanted CT nodules is 117. The diameter of simulated nodules is defined by the diameter of a disk with 95% of the total 'mass'.

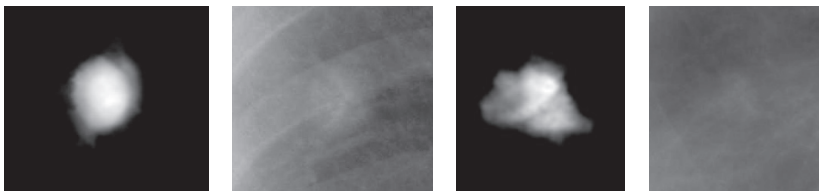
were removed from the dataset. Before features were computed the images were downsampled to $350\,\mu m$/pixel to speed up processing. Table 1 gives the subtlety rating of the abnormal JSRT images and Table 2(a) gives the effective diameters of the nodules in those images.

Computed Tomography Database. CT volumes in an anonymized database from a lung cancer screening study were used. To minimize influence of inaccurate segmentation in our simulations, only volumes with nodules that are not in contact with large blood vessels were selected from this database. Twenty-four lung nodules were annotated by a radiologist.

The CT volumes were acquired on a 16-slice Philips Brilliance scanner with a slice thickness of $1\,mm$ and a slice spacing of $0.7\,mm$. The in-plane resolution was between $0.6\,mm$ and $0.8\,mm$.

2.2 Simulated Nodules from Computed Tomography

X-rayed CT Nodules. Twenty-four CT nodules were segmented using a method by Kostis et al. [7]. This is basically a region growing algorithm with a refinement by morphological operators. It leads to segmented nodule templates that are rescaled to the resolution of the radiographs and placed in a box of air ($-1000\,HU$). Subsequently, nodules were projected along five randomly oriented axes using simple orthogonal ray casting. This leads to 120 two-dimensional CT intensity profiles. Three simulated nodules that looked unrealistic were not used, resulting in 117 simulated nodules. Table 2(b) gives the sizes of the 117 simulated nodules. The last step is a conversion from the projected CT units to radiograph units. This is done by a calibration obtained from studies with both a CT exam and a radiography exam [8]. The calibration function is determined by measuring the intensities in corresponding regions of the radiographs and the projected CTs. The intensities were measured in the heart, the spine, and in the lungfields. A calibration function was constructed by fitting a third-order polynomial to these data. It will not completely undo the effects of all post-processing steps that manifactures apply, like edge detection or unsharp masking, but it was adequate for the purpose of this study.



(a) (b) (c) (d)

Fig. 1. Examples of Implanted Nodules. Fig.s 1(a) and 1(c) depict two simulated nodules and Fig.s 1(b) and 1(d) depict parts of chest radiographs with these nodules implanted.

Implanting X-rayed CT Nodules. The CAD system is trained with both normal and abnormal images. We implanted exactly one simulated nodule in the abnormal images. The location was chosen in the contralateral lung field of the real nodule. (Remember that each abnormal image in the JSRT database contains exactly one real nodule.) Hence, the total number of available lung nodules for training (real and implanted) doubles. The vertical locations of the real and implanted nodules are chosen approximately the same and also their horizontal distances from the vertebral column are chosen approximately the same. The reason for this careful choice of placement is that we test the hypothesis that adding artificially implanted lung nodules to a training database improves the performance of a CAD system. If locations were carelessly chosen, e.g., by random placement, we might have impaired the CAD performance beforehand.

After projection of the CT nodules there is still one degree of freedom, namely the orientation, which is chosen to be random.

141 simulated lung nodules are sampled from the set of 117 simulated nodules with replacement and assigned to the 141 JSRT chest radiographs with a real lung nodule. For a wide range of pixel values the pixel value is approximately proportional to the thickness of tissue in film-screen radiography. Implanting nodules is therefore nothing more than adding pixel values. Figure 1 depicts two examples of simulated nodules.

2.3 CAD system

Local Direction Features. Detection is mainly based on two *local direction features* features: the first measures whether the number of local gradients pointing to a common center is statistically larger than the expected number for random gradients; the second measures whether local gradients are rotation symmetric.

If many gradient vectors are directed to a certain location this indicates that a lung nodule may be present. To quantify such events we define the statistic

$$s_{i,j} = \begin{cases} 1 - p_{i,j}, & \text{if “hit”, (pixel } j \text{ is oriented to pixel } i) \\ -p_{i,j} & \text{if “miss”.} \end{cases}$$

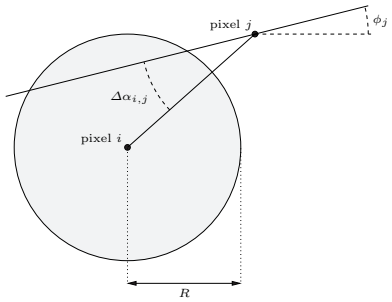


Fig. 2. Hit/Miss. A pixel j is oriented to pixel i when the line through $\mathbf{r}_j$ with orientation ϕ_j crosses a disk with radius R centered at pixel i.

A pixel j is oriented to pixel i, i.e., we speak of a "hit", when the line through the location of pixel j, denoted by $\mathbf{r}_j$, with orientation $\phi_j \in [0, \pi]$ crosses a disk with radius R centered at $\mathbf{r}_i$ (see Figure 2). This happens when

$$\sin \Delta\alpha_{i,j} \leq \frac{R}{\|\mathbf{r}_j - \mathbf{r}_i\|}, \tag{1}$$

where $\Delta\alpha_{i,j} \in [0, \pi/2]$ is the acute angle between a line with orientation ϕ_j and the line through $\mathbf{r}_i$ and $\mathbf{r}_j$. The prior probability that a line through $\mathbf{r}_j$ is oriented to pixel i is given by $p_{i,j}$. Assuming that the prior probability of orientations ϕ_j is uniform on $[0, \pi]$, we have

$$p_{i,j} = \begin{cases} \frac{2}{\pi} \arcsin \frac{R}{\|\mathbf{r}_j - \mathbf{r}_i\|}, & \text{if } \|\mathbf{r}_j - \mathbf{r}_i\| \geq R, \\ 1 & \text{otherwise.} \end{cases} \tag{2}$$

A sum is computed by

$$S_i = \sum_{j \in \mathcal{N}_i} s_{i,j}$$

where an annular neighborhood $\mathcal{N}_i$ with outer radius D and inner radius $D_{\min}$ is defined as

$$\mathcal{N}_i = \left\{ \forall\, \mathbf{r}_j \in \mathbb{R}^2 \,|\, R < D_{\min} \leq \|\mathbf{r}_j - \mathbf{r}_i\| \leq D \right\}.$$

The first feature, the *concentration feature*, is defined as

$$g_1 = \frac{S_i}{\operatorname{std}\{S_i\}} = \frac{\sum_{j \in \mathcal{N}_i} s_{i,j}}{\sqrt{\sum_{j \in \mathcal{N}_i} p_{i,j}\,(1 - p_{i,j})}}. \tag{3}$$

The normalization $\operatorname{std}\{S_i\}$ in the denominator is computed by assuming that all orientations at $j \in \mathcal{N}_i$ are uniformly distributed and statistically independent. For random orientations this feature is zero, and only when orientations systematically have a common center at i (or the opposite, not one is oriented to i) it deviates significantly from zero.

(a) X-location (b) Y-location (c) Heart border (d) Lung border

Fig. 3. Location Features within the Unobscured Lungs Fields.

The second feature takes the symmetricity of orientation patterns into account. Each neighborhood $\mathcal{N}_i$ is divided into K directional bins, i.e., like a pie ($K = 24$ is used). For each bin the statistic S_i is computed as before, but now with $\mathcal{N}_i$ replaced by the region of one piece of the pie. When $S_i > 0$ for all bins, there is evidence that the pattern of orientations is symmetric around pixel i. The second feature, the *concentration symmetricity feature*, is defined as

$$g_2 = \frac{n_+ - K_i}{\sqrt{K_i/4}}, \tag{4}$$

where $K_i \leq K$ is the number of bins at pixel i, and $n_+ \in [0, K_i]$ is the number of bins with positive S_i. The number of bins is variable as pixels close to the breast or image boundary are excluded if they do not contain enough pixels. The denominator is the standard deviation of n_+ when all orientations are uniformly distributed and statistically independent, i.e., when n_+ is a binomial random variate (with sample size $K = 24$ and probability $p = 1/2$ that bin i has a positive statistic S_i).

Note that when the summations above are computed for a certain upper radius D, it takes little effort to compute the summations for a somewhat larger upper radius. This means that g_1 and g_2 can be computed efficiently for all encountered radii D in a rectangular grid of pixels. Instead of using g_1 and g_2 for fixed upper radius D, we use the multi-scale features after computing them as a function of D

$$G_1 = \max_D \{g_1(D)\} \tag{5}$$

and

$$G_2 = \max_D \{g_2(D)\}. \tag{6}$$

Before determining the maxima the functions $g_i(D)$ are smoothed.

Location Features. Because nodules are not uniformly distributed over the lungs and, more importantly, because the local direction features suffer from the vessels near the heart, four *location features* are added: the relative X- and Y-coordinates within the pixel's unobscured lung field (computed by the cumulative area horizontally or vertically); the relative distance to the heart; and the relative

distance to the border of the unobscured lung fields (see Figure 3). The four location features are normalized between zero and one.

Classifier. We used a three-layer feedforward neural network with sigmoid transfer functions and one hidden layer to classify feature vectors. The number of input nodes was equal to the length of the feature vector. The number of nodes in the hidden layer was five and one output node was used. The network is trained by the backpropagation algorithm.

The neural network output was computed on regularly spaced grid locations (the grid spacing was $2.8\,mm$) in all testing images and thresholded local maxima of the resulting output images were further analyzed. A local maximum of an output image was considered a true-positive (TP) when its location was less than $2.5\,cm$ from the center of any annotated lesion, otherwise the local maximum was considered a false-positive (FP). This criterion was proposed by Hardie et al. [4].

In most CAD systems a second stage is employed in which candidate nodules detected in the first stage are segmented and classified with more features. This study is limited to investigating the use of simulated nodules in the initial detection stage.

Training and Testing. Lung nodule images in the JSRT database are classified into six strata according to the degrees of subtlety shown, ranging from "obvious" to "extremely subtle" and "no lung nodules (normals)" (taken from Shiraishi et al. [5]). The numbers of images per degree of subtlety are given in Table 1. Stratified fivefold cross-validation was used to partition the dataset into training sets and testing sets.

Let us call the set with testing images S_{20} (20% of all normal and abnormal JSRT images) and the disjoint set with training images S_{80} (80% of all normal and abnormal JSRT images). The pixel classifier is trained with the set S_{80} (with existing and implanted lung nodules) and tested with the disjoint set S_{20} (only with existing lung nodules). This is repeated with circulating sets S_{20} until all images have been tested, i.e., five times. The five sets S_{80} are randomized once and remain the same in all experiments thereafter. The only thing that differs between experiments are the regions of interest (ROI), comprising the set of pixels used for training. When a nodule, irrespectively whether it is a real or a simulated nodule, is not to be used in a certain experiment then it is removed from the ROI. This is done by removing all pixels in a disk with a diameter of $4\,cm$ centered at the nodule's location from the ROI. When N_{real} real nodules and $N_{implanted}$ simulated nodules are to be used for training then all nodules are blanked, except the real nodules in the first N_{real} images and the simulated nodules in the following $N_{implanted}$ images. Of course, the ROIs of the testing sets S_{20} are not manipulated at all. For experiments with a constant of $N_{real} + N_{implanted}$ lung nodules in the training set, the only thing that may be varied is the relative amount of implanted lung nodules in the training sets,

i.e., the blanked nodules; The sets of training and test images remain the same under all conditions.

3 Experimental Results

Location Features and Local Direction Features. The main experiment is performed with different numbers of real nodules and simulated nodules. For this experiment all six features are included. The results, evaluated by a free-response receiver operating characteristic analysis (FROC; [9]), are given in Figure 4.

In the next section a control experiment is performed to test whether these results can be explained by the way the simulated nodules were implanted.

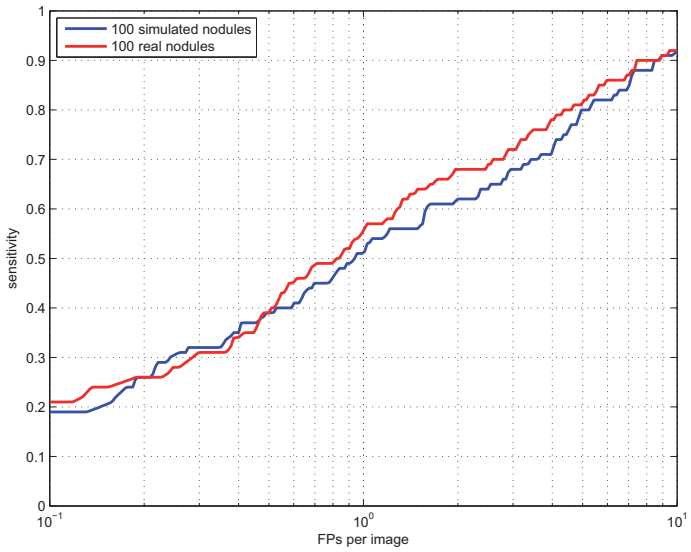
Location Features. To show that the previous results cannot be explained by the locations of the implanted lung nodules alone, a control experiment is performed with only location features. The results are given in the FROC curves of Figure 5. Evidently, the curves lie much lower than the curves in Figure 4(a) (e.g., at one FP per image the sensitivity drops from 0.56 to 0.14 for the real nodules). This proves that the previous results cannot be explained by location features alone. In other words, the morphology of the simulated nodules adds something essential.

The slight difference between the two curves can be explained as follows. Although the relative locations of the implanted nodules are the same as the relative locations of the real nodules, the distances to the heart and the distances to the unobscured lung border, also two location features, are not.

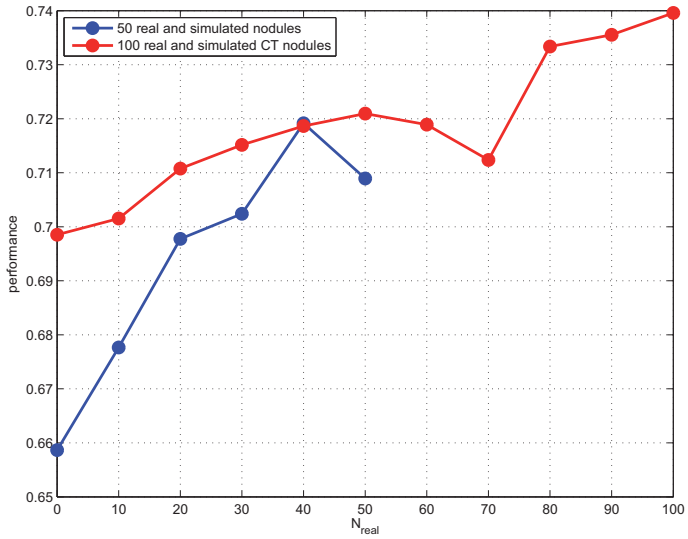
4 Discussion

Conclusion. This paper presents the preliminary results of a study with implanted lung nodules for the purpose of training CAD systems. The hypothesis that a CAD system improves with artificially added X-rayed CT nodules is confirmed. Even when no real nodules are in the training set, the performance of the CAD system is reasonably high, certainly if one considers that only twenty-four three-dimensional CT nodules form the basis of the implants. The sensitivity of the CAD system trained with 100 simulated nodules was 0.51 at one FP per image. The best results reported to date were obtained by Hardie et al. [4]. With the same criterion for TPs, they found a sensitivity of 0.56 at one FP per image (tenfold cross validation on the JSRT database). The CAD system they used was however trained with more than 100 abnormal images, but more importantly it included an extra classifier for candidate lung nodules, an extra step that was omitted here.

Future Work. In an extended CAD scheme the candidate nodules are further processed in a subsequent, second classifier. It will be investigated whether such a CAD schemes will increase the performance (which is already relatively high).



(a) FROC curves



(b) Performance with $N_{real} + N_{implanted} = 50$ and 100

Fig. 4. FROC analysis of a CAD system with Location Features and Local Direction Features. Figure 4(a) shows the FROC curves when the CAD system is either trained with either $N_{real} = 100$ real nodules or with $N_{implanted} = 100$ simulated nodules. Figure 4(b) shows aggregated results for other numbers of real and simulated nodules by integrals over the FROC curves with respect of $\log_{10}$ (FPs per image) from 0.1 to 4 FPs per image. Notably, the red points at $N_{real} = 0$ and 100 are the aggregate performances of Figure 4(a).

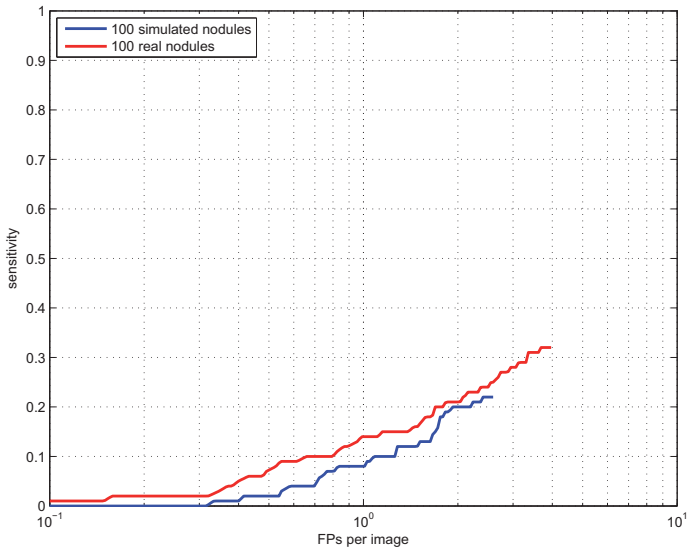


Fig. 5. FROC analysis of a CAD system with Location Features only. The CAD system is either trained with $N_{real} = 100$ real nodules or with $N_{implanted} = 100$ simulated nodules.

Probably more than twenty-four CT nodules (the number used in the current study) are required to capture the more subtle properties that nodules may possess.

The locations of implants were realistic locations in this work, because it is our believe that the location of a nodule is critical for its detection. This assumption will be investigated. If it is true that locations of implants are critical for training then a generic, smooth distribution of nodule locations may be constructed from a database of real nodules.

It would be very interesting to see whether CAD results improve when chest radiographs are split in different parts, and each trained with a different classifier. The first thing that comes in mind are the occluded regions, the heart, the clavicles, and the diaphragm, but one can also use different classifiers for half-occluded on non-occluded nodules by the ribs. When implanted nodules cannot be distinguished from real nodules then availability of trained data is not an issue anymore.

4 References

1. Schilham, A., Ginneken, B.v., Loog, M.: A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database. Medical Image Analysis **10**(2) (2006) 247–258

2. Shiraishi, J., Abe, H., Li, F., Engelmann, R., Macmahon, H., Doi, K.: Computer-aided diagnosis for the detection and classification of lung cancers on chest radiographs: ROC analysis of radiologists' performance. Academic Radiology **13**(8) (2006) 995–1003
3. Freedman, M., Chung, S., Osicka, T., Lure, F., Xu, X.W., Lin, J., Zhao, H., Zhang, R.: Computer-aided detection of lung cancer on chest radiographs: effect of machine CAD false-positive locations on radiologists' behavior. Volume 4684., SPIE (2002) 1311–1319
4. Hardie, R., Rogers, S., Wilson, T., Rogers, A.: Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs. Medical Image Analysis **12**(3) (2008) 240–258
5. Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K., Matsui, M., Fujita, H., Kodera, Y., Doi, K.: Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. American Journal of Roentgenology **174**(1) (2000) 71–74
6. Ginneken, B.v., Stegmann, M., Loog, M.: Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. Medical Image Analysis **10**(1) (2006) 19–40
7. Kostis, W., Reeves, A., Yankelevitz, D., Henschke, C.: Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. IEEE Transactions on Medical Imaging **22**(10) (October 2003) 1259–1274
8. Litjens, G., Hogeweg, L., Ginneken, B.v.: Simulation of nodules and interstitial disease in chest radiographs using CT templates. In: Proceedings MICCAI 2010. Lecture Notes in Computer Science, Springer (2010)
9. Miller, H.: The froc curve: A representation of the observer's performance for the method of free response. The Journal of the Acoustical Society of America **46**(6B) (1969) 1473–1476

Parametric and Non-Parametric Nodule Models: Design and Evaluation

Amal A. Farag, James Graham, Aly A. Farag, Salwa Elshazly and Robert Falk*
Department of Electrical and Computer Engineering, University of Louisville
(*) Medical Imaging Division, Jewish Hospital, Louisville, KY, USA

Abstract. Lung nodule modeling quality defines the success of lung nodule detection. This paper presents a novel method for generating lung nodules using variational level sets to obtain the shape properties of real nodules to form an average model template per nodule type. The texture information used for filling the nodules is based on a devised approach that uses the probability density of the radial distance of each nodule to obtain the maximum and minimum Hounsfield density (HU). There are two main categories that lung nodule models fall within; parametric and non-parametric. The performance of the new nodule templates will be evaluated during the detection step and compared with the use of parametric templates and another non-parametric Active Appearance model to explain the advantages and/or disadvantages of using parametric vs. non-parametric models as well as which variation of non-parametric template design, i.e., shape based or shape-texture based yields better results in the overall detection process.

Keywords: Nodule modeling, Sensitivity and Specificity of CAD systems, Data-driven nodule models, Level Sets

1 Introduction

This paper focuses on modeling of the lung nodules which appear in low dose computer tomography (LDCT) of the human chest. In the past two decades numerous screening studies in Europe, Japan and the US have been conducted for studying the enhancements of early detection of lung cancer using CT vs. X-ray and for studying the correlation of early detection and possible enhancement in lung cancer related mortality. Suffices to say, lung cancer is a major problem worldwide [1]. The survival of lung cancer is strongly dependent on diagnosis [2]. Research studies to reach an optimal detection rate for early detection of lung cancer, is the hope for improved survival rate [3]-[5].

Machine learning and computer vision methodologies have been used for image analysis of low dose CT (LDCT) of the chest (e.g., [2] [5]-[8]). A computer-aided (CAD) system for interpretation of nodules is formed of four major steps: scan filtering to remove acquisition artifacts; segmentation to isolate the lung tissue from the rest of the chest region; nodule detection to isolate candidate nodules; and nodule classification which categorizes detected nodules into possible pathologies. The literature is rich in approaches to segment the lung from the rest of the chest tissues, but the majority of the nodule modeling methods are based on parametric descriptions

of the nodules (e.g., in 2D circular or semicircular models are used, while in 3D volumes spherical or hemispherical models are used [5]). This paper will focus on the third component of the CAD system used for early screening of lung cancer; more specifically how to properly model lung nodules by using the actual data information to create the models/templates that will improve detection rate. The approach of nodule detection hinges mainly on proper modeling of nodule templates and much less on the computational approach to carry out the detection. Extensive surveys on automatic lung nodule segmentation and detection may be found in [8][9]. The long history of work in the communications literature (e.g., matched filtering) and computer vision literature (e.g., active shape and active appearance models) has demonstrated the value of proper modeling of the objects to be detected. This is especially the case in biomedical applications as the anomalies are often camouflaged or occluded by anatomical structures and the limitations of the scanning device or protocols. To the best of our knowledge none of the studies on lung screening conducted worldwide has resulted in identifiable databases of nodules listing their types and pathologies. Therefore, the need is persistent for reliable nodule models based on the actual scans; this is one of the goals of the authors of this paper.

This paper will investigate two main topics: First, whether parametric template models are more effective than non-parametric models in terms of sensitivity and specificity. We use the term data-driven for the non-parametric templates since the raw data information is used. Second, we study the effect of shape and texture on the non-parametric models. Shape only models were obtained by co-registering the contours of an ensemble of nodules obtained by a level set approach. The shape and texture models were obtained by active appearance models (AAM) (e.g., [17]). Our focus in particular, is on four nodule types (e.g., Kostis et al. [6]) that possesses discriminatory features of shape and to some extent texture.

The closest related work to this paper are: 1) Kostis et al. [6] provides a description of four major types of lung nodules based on identifiable landmarks which will enable automatic annotation in our methods. 2) Lee et al. [5] established an empirical relationship, or behavior, for the intensity (or Hounsfield Units) of the nodules as a function of the radial distance from the centroid of the nodule; this is beneficial for texture/intensity/gray level estimation of the inside of a nodule shape. 3) Farag et al. [7][10][11] established a parametric form for the relationship between the radial distance and the Hounsfield units in Lee's work; this is very useful for estimating the intensity of a nodule model given the statistics of an ensemble and the size (radius of the bounding box containing the nodule). Other works include: the use of 3D Markov random models [13] to compute the optimal states over the cells from the relationship of the neighboring cells and Bayesian voxel labeling [14] which labels the image according to three categories; anatomy, pathology and miscellaneous. The probability that the voxel belongs to each class is computed and a decision is made. The approaches described here assume parametric shape representations of the nodules, whither circles, spheres, etc. or fused ellipsoids to represent the nodule and its outer surroundings. Parametric modeling of the nodules (e.g. circles and/or semicircles) will not capture the shape variations in the lung nodules. Likewise, assuming a uniform (or binary) HU for the nodule doesn't resemble reality. Shape and texture models have shown great promise in a number of computer vision and biomedical imaging analysis applications (e.g. [15][16]). To the

best of our knowledge Farag et al. in 2009 [17] is the first reported attempt at employing shape and appearance modeling into the problem of automatic detection and segmentation of lung nodules in LDCT scanning. This paper is organized as follows: section 2 describes the level sets shape based approach for obtaining a mean template representation for each nodule type and how these templates are filled with texture. The parametric and AAM methods are not described in great details in this paper. Section 3 discusses performance evaluation; and section 4 concludes the paper.

2 Nodule Modeling

This section will examine the process of nodule modeling and simulation using an ensemble of nodules identified by radiologists. The level sets method for generating nodule models is the main focus of this section, while the parametric [7][11] nodule modeling method and data-driven using AAM [17] to model the nodules in terms of both shape and intensity method will briefly be described.

2.1 Pulmonary Nodule Definitions

In radiology, a pulmonary nodule is a mass in the lung usually spherical in shape; however it can be distorted by surrounding anatomical structures such as the pleural surface. This paper uses the classification of Kostis et al. [6], which groups nodules into four categories (Fig. 1): *well-circumscribed* where the nodule is located centrally in the lung without being connected to vasculature; *vascularized* where the nodule has significant connection(s) to the neighboring vessels while located centrally in the lung; *juxta-pleural* where a significant portion of the nodule is connected to the pleural surface; and *pleural tail* where the nodule is near the pleural surface, connected by a thin structure; in all of these types there is no limitations on size or distribution in the lung tissue. These definitions are used in our approach. These nodule types are characterized mainly by shape, and location with respect to the anatomy of the chest; the appearance of an individual nodule may not hold too much discrimination. The ELCAP [12] database provides the LDCT scans where the center location of the nodule was specified by radiologists. We constructed a database of nodules using a semi-automatic method of cropping and categorizing each nodule into one of the four types.

2.2 Lung Nodule Model Simulation: Level Sets Approach

The goal of the modeling process is twofold: generate a template for each nodule type to be used in the detection process, and establish a procedure for simulation of nodules to be used for nodule classification. Our focus in this paper is on generating an overall mean template per type that best represents the characteristics of the real nodules. A database of nodules was constructed using the Early Lung Cancer Action Program (ELCAP) lung scans datasets [12]. As described in Farag, et al. 2009 [7][11], the probability density function of the Hounsfield (HU) vs. radial distance

distribution of the ELCAP dataset had an exponentially decaying form concentrated in distances of about 10 pixels from the centroid of the nodules (i.e., about 5mm); thus cropping boxes of size 21x21 pixels were used to obtain the nodule ensemble, given their location in the LDCT scans. These nodules where then classified into one of the four corresponding categories described in sec. 2.1, constructing a nodule database that contains variations in intensity distribution, shape/structural information and directional variability which the cropped regions, within the determined bounding-box, maintain. A sub-database of 96 nodules (24 nodules per type) is used in both the level sets approach and the AAM method to generate a mean nodule template for each type that depicts shape or shape and texture information of the nodules (Fig.1).

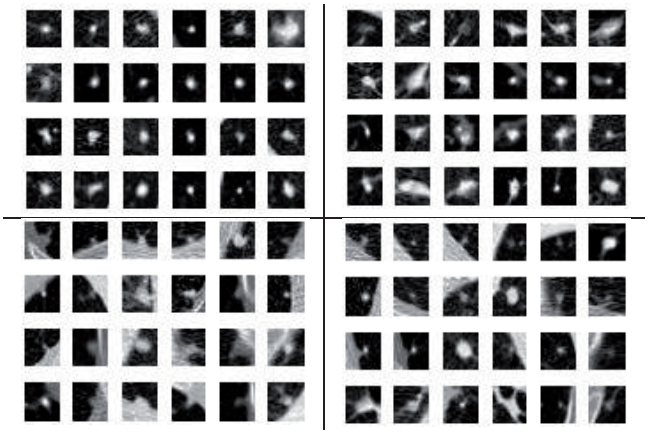


Fig. 1. *An ensemble of 24 nodules from the well-circumscribed (upper left), vascular (upper right), juxta-pleural (lower left) and pleural-tail (lower right) nodule types.*

The Procrustes registration-based AAM approach [15] for lung nodule modeling [17], required manual annotation by trained experts of the 96 nodules to employ Procrustes registration to obtain co-registered nodules. A combinational shape and texture AAM algorithm was used in [17] to generate a mean template that contained both shape and texture information of the lung nodules shown in Fig. 2.

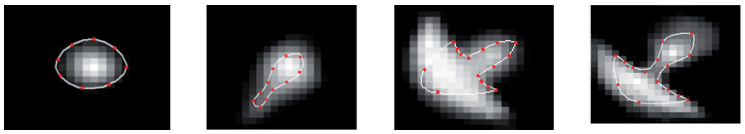


Fig. 2. *Nodule models generated using the Procrustes based AAM method. From left to right: Well-circumscribed, vascular, juxta-pleural, and pleural-tail nodule types.*

A second sub-database of 96 nodules, different from that in Fig. 1, was used to generate another set of four mean nodules using the AAM approach, the sub-database and mean nodules generated are depicted in figures 3 and 4, respectively.

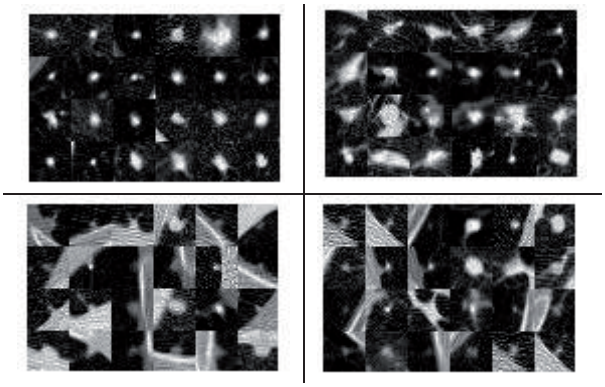


Fig. 3. *Second ensemble of 24 nodules from the well-circumscribed (upper left), vascular (upper right), juxta-pleural (lower left) and pleural-tail (lower right) nodule types.*

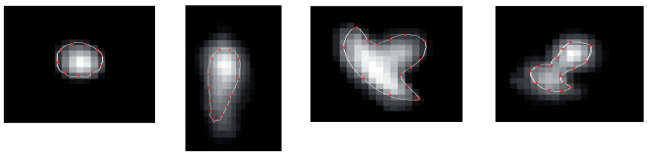


Fig. 4. *Nodule models generated using the Procrustes based AAM method. From left to right: Well-circumscribed, vascular, juxta-pleural, and pleural-tail nodule types.*

The usage of variational Level sets (e.g., [18]) in this paper replaces the step of manual annotation performed in the AAM method by a semi-automatic approach that generates the contours of the lung nodules depicting the shape information. These contours are then co-registered using the Procrustes method; the contour boundary points are obtained and used for registration. The Level sets approach eliminates sources of errors that can arise with manual annotation since only placing the seed point or points in the region of the nodule centorid is manually performed. Also, the elasticity of this variational approach addresses the issue of shape variations that can arise and handles these changes accordingly. In this paper, we used the approach of Abdelmunim and Farag, 2007 ([18]) for rigid and elastic shape representation via Level sets. The mean shape templates generated from the contours is shown in Fig. 5. Given two shapes represented by the vector functions Φ1 and Φ2, a transformation *A* with scales, rotation and translation is to be calculated to transform the first object to the second. The following dissimilarity measures the difference between the vector and the other scaled one:

$$r = SR\Phi_1(X) - \Phi_2(A) \tag{1}$$

The following energy is formulated as a sum of squared differences

$$E = \int_{\Omega} \delta_{\in} r^T r d\Omega \tag{2}$$

where the delta is an indicator function with value 1 inside the shape and zero otherwise. The two shapes are aligned by minimizing the energy function using various approaches including the gradient dissent method. The training shapes in our case the contours of the lung nodules are jointly registered with an evolving mean

shape to find the corresponding global transformations $A_1,\ldots,A_n$. The dissimilarity measure is used as follows:

$$r_i = S_i\Phi_M - \Phi_i(A_i) \tag{3}$$

The energy function will be:

$$E(\Phi_M,\Phi_1,\ldots,\Phi_n) = \sum_{i=1}^{n}\int_{\Omega}\delta_{\in} r_i^T r_i d\Omega \tag{4}$$

The shape model is a function of the registered training shapes:

$$\Phi_p = \Phi_M + \sum_{i=1}^{n} w_i(\Phi_i^t - \Phi_M) \tag{5}$$

The shape parameters ***w,*** need also to minimize the energy function to process the registration.

Once the mean shapes were generated estimation of texture information was required, thus using a synthetic approach described in [5][7] and [10] for parametric templates was performed. The probability density of the intensity for the nodules is the same as that in [7] since the ELCAP database is used in both.

The intensity of a nodule model was found to be estimated by the following equations [7][10]:

$$q(r) = q_{max}e^{-(r/\rho)^2}, \qquad 0 \le r \le R \;) \tag{6}$$

$$\rho = R\left(\ln(q_{max}) - \ln(q_{min})\right)^{-1/2} \tag{7}$$

where R is the radius of the circle interior to the bounding box containing the nodule model (mean shape). The parameters q_{min} and q_{max} are the lower and upper bounds of the intensity (Hounsfield Units) in the probability density function of each nodule type; which is estimated from the ensemble of nodules (e.g., [7][11]). Fig. 6 shows the templates generated based on the mean shape and the empirical forms of the intensity (Eq. 6 and 7).

Fig. 5. *Average shape of the nodules in Fig. 1. From left to right: Well-circumscribed, vascular, juxta-pleural and pleural-tail.*

Fig. 6. *The nodule templates resulting from the intensity equations Eq. 6 and 7 and the mean shapes in Fig.4. From left to right: Well-circumscribed, vascular, juxta-pleural, and pleural-tail nodule types shape modeling process.*

A set of parametric templates (e.g. circular and semi-circular) were generated using equations for circular and semi-circular shapes for various radii of empirical nodules and filled with texture information using equations 6 and 7. Template size

was the same 21x21 pixels bounding box area, similar to the AAM and level sets approaches. The same probability density of the nodule intensity observed for the nodule database for q_{min} and q_{max} were used once more and the templates were filled using the same technique implemented for the level sets approach. Circular or isotropic templates are defined in terms of radius and gray level distribution as a circular symmetric Gaussian function while the semi-circular template have an additional orientation parameter. Fig. 7 shows as set of isotropic and non-isotropic templates generated by this approach.

Fig.7. *An ensemble of generated circular and semi-circular templates with various orientations (adopted from* [7]).

2.3 Lung Nodule Modeling Summary

In this paper we examined three approaches for generating lung nodule models or templates. Of these methods two are non-parametric and data-driven, while the other approach was parametric as shown in Fig. 7. Parametric nodule models and the non-parametric nodule models, based on shape only, may be used in template matching in a binary form (after we segment the original lung images to generate a binary image) or in gray scale form applied straight to the original lung images (after we remove the non-lung tissues in the segmentation step preceding the detection). In this paper, as indicated above, the intensity information in the shape models (using level sets) as well as the parametric models were obtained using Eq. 6 and 7.

We should point out that the Procrustes approach was used in obtaining the mean shape (generated from level sets) and the mean shape and texture (generated from the AAM approach). Other methods may be used to carry out these co-registrations and may lend enhanced efficiencies with respect to orientations, shapes, etc. in the ensemble. The overall mean shapes of each nodule type from the two approaches showed extraordinary resemblance as seen in Fig 2 and Fig. 5.

In the next section we evaluate the performance using the nodule templates from the methods described above for the detection of candidate nodules.

3 Results

The Early Lung Cancer Action Program (ELCAP) public database [12] was used in this paper for nodule modeling, classification and detection. The database contains 50 sets of low-dose CT lung scans taken at a single breath-hold with slice thickness 1.25 mm. The locations of the 397 nodules are provided by the radiologists, where 39.12% are juxta-pleural nodules, 13.95% are vascularized nodules, 31.29% are well-circumscribed nodules and 15.65% are pleural-tail nodules. In the detection stage all slices containing nodules are used unlike the modeling stage where only 96 nodules of

the total 397 were used for generating the templates. Due to the mean template per nodule type is found, in the cases of the data-driven models, we use the slices from the modeling stage in the detection process as well since the templates are not biased towards any particular one nodule in Fig. 1. There are numerous methods for performing lung nodule detection in the literature; our goal for this paper is to depict the effectiveness of our nodule modeling, thus a generic implementation of template matching with the normalized cross-correlation (NCC) as the similarity measure is executed. The nodule model (template) is swept across the scan (2D slices or the 3D volume) in a raster fashion and a similarity measure is calculated between the intensity (or HU) of the template and the region of the CT data underneath, each template produces a binary image that represents candidate nodule locations, the four images are xored together to obtain one image that depict overall candidate nodule location. We use the widely known form of the NCC in the literature for the normalized cross-correlation of a template, *t(x,y)* with a sub-image *f(x,y)*:

$$NCC = \frac{1}{n-1}\sum_{x,y}\frac{(f(x,y)-\bar{f})(t(x,y)-\bar{t})}{\sigma_f \sigma_t}, \qquad (8)$$

where *n* is the number of pixels in template *t(x,y)* and sub-image *f(x,y)* which are normalized by subtracting their means and dividing by their standard deviations. The probability density functions (pdf) of nodule and non-nodule pixels are computed using the normalized cross correlation coefficients resulting from templates with varying orientations.

The NCC behavior with the data-driven nodule models takes the same general shape as with the parametric nodules except the distribution function decays faster as we approach a value of 0.5. Setting a suitable threshold for the NCC is important as lower thresholds will increase detection rate but increases the false positives, and vice versa. Various methods can be used for an optimal threshold including modeling the normalized histogram as two classes, nodules and non-nodules and a Bayesian approach may be devised to select an optimal threshold. In this paper we set the threshold to a NCC of 0.5 for the sake of comparison between the three methods.

For nodule recognition (i.e., deciding the pathology of the nodule), features from the detected nodules need to be compared with pathological counterparts. This issue will not be considered in this paper as its significance really depends on the availability of pathological nodule database which is under construction by this research group. Also, the validation of pathology requires three human experts at least (two independent reader and a third to decide on opposing decisions).

The overall sensitivity and specificity was computed using equations 9 and 10:

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives} \qquad (9)$$

$$Specificty = \frac{True\ Negatives}{True\ Negatives + False\ positives} \qquad (10)$$

True positive rate refers to the number of actual nodules that are detected as nodules while false negatives are the number of nodules that were not detected as nodules. Thus sensitivity depicts how well the detection was able to recognize nodules from other lung features using the desired designed templates (parametric and data-driven). True negatives are the number of nodules that are truly not nodules while false positives are the number of non-nodules that were detected as nodules.

Specificity rate is more subjective in its computation since it depends on how the true negatives and false positive rates are computed during detection. The specificity represents the negative rate that is correctly identified.

Table 1 depicts the overall sensitivity and specificity results of using the templates generated by the parametric method, the Level-sets algorithm and AAM method using the sub-database in Fig. 1 and Fig. 3, centered with respect to the x-axis (i.e., zero orientation). While Table 2 depicts the overall sensitivity and specificity results when the templates are rotated from 0^o to 360^o with step-size 90^o. From Tables 1 and 2 several conclusions can be drawn, first the data-driven method using the Level-sets to generate template that depict overall shape only which are then filled, as described in section 2.2, yields comparable results to that of the parametric templates in terms of sensitivity and slightly higher specificity results. Second, rotation of the templates improved sensitivity in only the parametric and level-set based template approaches and overall specificity slightly decreased for all approaches. Third, template generated by the AAM approach provides better results than the parametric, level sets method and the AAM mean templates depicted in Fig 4. Overall, the AAM algorithm using either set of mean templates generated from sub-database 1 or 3 yield better results in terms of both sensitivity and specificity. Thus, data-driven models are more robust and an enhanced method of lung nodule modeling over the use of parametric templates, since the actual data is used in modeling and generating mean templates to represent each nodule type. Also, shape and texture based approaches give a more accurate and precise representation to the true nodule that provides improved detection results. So, shape information alone does not suffice, both shape and texture information is required.

Table 1: *Overall sensitivity and specificity of level sets, parametric, AAM using dataset 1 and AAM using dataset 2 without accounting for template orientation.*

Algorithm	Sensitivity	Specificity
Parametric Approach with template radius 10 and single orientation for semi-circular template	72.16%	97.12%
Level Sets Approach using nodule contours from dataset 1and no orientation	72.16%	98.11%
AAM Approach using dataset 1 and no orientation	85.22%	97.81%
AAM Approach using dataset 2 and no orientation	83.51%	98.36%

Table 2: *Overall sensitivity and specificity of level sets, parametric, AAM using dataset 1 and AAM using dataset 2 templates averaging a number of orientations for the templates*.

Algorithm	Sensitivity	Specificity
Parametric Approach with template radius 10 and orientation 0^o-360^o with step-size 90^o for semi-circular template	78.01%	96.41%
Level Sets Approach using nodule contours from dataset 1and orientation 0o-360o with step-size 90o	76.98%	97.63%
AAM Approach using dataset 1and orientation 0o-360o with step-size 90o	86.94%	96.51%
AAM Approach using dataset 2and orientation 0o-360o with step-size 90o	83.51%	97.40%

Further studies were conducted using the AAM based approach on both sub-datasets used for template modeling. The number of annotation points necessary for proper registration was found to be a function of how many was necessary to depict the main discriminatory shape information and withholds substantial texture information commonly found in each nodule for that particular type (i.e. the 24 nodules used per type). Also, 24 nodules per type were used in the modeling of each nodule type but if fewer nodules were used what will be the effect on the generated mean nodule templates formulated and used in the detection process was examined.

Table 3 depicts the results obtained when the first 16 and 8 nodules from each of the sub-databases are used. From the table it is seen that overall sensitivity and specificity using half or one-third of the nodules in figure 1 results in overall similar sensitivity and specificity results while in the case of using the sub-database in figure 3 results in reduced sensitivity as the number of nodules used for modeling decrease. Overall we found that depending on how well the lung nodules are annotated and which nodules are represented in the sub-database effected the generation of the mean templates per type; i.e. if nodules used in the modeling depicted a majority of the nodules in the larger database and annotated well then the overall mean templates generated yielded improved sensitivity and specificity results, if the nodules were not adequately annotated the mean templates generated not always gave improvements in sensitivity and/or specificity likewise if the nodules in the modeling database did not depict a vast majority of those in the original database then detection rates reflected that.

The results are expected to be further enhanced using larger ensemble sizes than the 24 per nodule types which we used in our experiments. Likewise, involvement of several radiologists to create the ensemble may also lead to further improvements.

Table 3: Overall sensitivity and specificity of AAM approach using 8 and 16 nodules from datasets 1 and 2 for mean template modeling, respectively.

Algorithm	Sensitivity	Specificity
AAM Approach using 16 nodules from dataset 1 for modeling and no orientation	85.57%	97.84%
AAM Approach using 8 nodules from dataset 1 for modeling and no orientation	84.88%	97.99%
AAM Approach using 16 nodules from dataset 2 for modeling and no orientation	83.16%	98.42%
AAM Approach using 8 nodules from dataset 2 for modeling and no orientation	77.32%	98.57%

4 Conclusions and Extensions

In this paper, a data-driven approach using level sets was devised to model and simulate typical lung nodules. The modeling procedure of parametric and non-parametric template models was examined and used for nodule detection. The effect of template shape and texture on detection of different nodules types was studied.

From our extensive experimentation we can concluded that the data-driven AAM algorithm for lung nodule modeling yielded an overall higher sensitivity and specificity rate, yet, the Level sets approach showed instances of improvement for specificity and/or sensitivity over the usage of parametric templates. In the parametric case where we tested on all radii sizes between 1 and 20 pixels the sensitivity was higher but the specificity in comparison to the data driven nodule templates was still lower.

This paper has shown that approaches were both shape and texture information is simultaneously computed for modeling is more robust and an accurate approach than relying only on shape information for precise nodule descriptions. Current efforts are directed towards constructing and testing the data-driven modeling approach on a large clinical database and extending this work into the 3D space. The nodule databases will be made available to the research community in order to measure the enhancements made in the detection as well as recognition/classifications based on a common standard. Also, evaluating these templates to classify the nodules into designated pathologies (e.g. benign and malignant) is a key and important step that will be examined. Other algorithms of constructing shape and texture based models will be explored.

Therefore, we have established a systematic approach to model and simulate the lung nodules in LDCT scans which is applicable to any data protocol, and any nodule definition. This contribution is very crucial and may be the building block for all work on CAD systems applied to lung nodules; indeed, it may be also used for all similar approaches that generate templates to be detected in data of various types and formats. The main power of this approach is the fact that it is *data-driven;* hence, various attributes of the data may be incorporated in the template design.

Acknowledgements: Dr. Asem Ali and Ham Rara, members of the CVIP Lab, provided various insights and assistance in this research. The first author has been supported by a fellowship from the United States National Space and Aeronautics Agency, NASA. This research has been supported by the Kentucky Lung Cancer Program.

References

1. United States National Institute of Health www.nih.gov
2. B. Zaho, G. Gamsu, M.S. Ginsberg, L. Jiang, L.H. Schwartz, "Automatic Detection of small lung nodules on CT utilizing a local density maximum algorithm," Journal of Applied Clinical Medical Physics 4 (2003).
3. Armato, S. G. 3rd, Giger, M. L., Moran C. J., Blackburn, J. T., Doi, K., MacMahon H.: Computerized detection of pulmonary nodules on CT scans. Radio Graphics 19 pp.1303--1311 (1999).
4. S. Hu, E. A. Hoffman and J. M. Reinhardt, "Automatic lu1ng segmentation for accurate quantitation of volumetric X-ray CT images," IEEE Transactions on Medical Imaging, Vol. 20, pp. 490–498, 2001.

5. Y. Lee, T. Hara, H. Fujita, S. Itoh and T. Ishigaki, "Automated Detection of Pulmonary Nodules in Helical CT Images Based on an Improved Template-Matching Technique," IEEE Transactions on Medical Imaging, Vol. 20, 2001.
6. W. J. Kostis, A.P. Reeves, D. F. Yankelevitz and C. I. Henschke, "Three dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images," IEEE Transactions on Medical Imaging, Vol. 22, pp. 1259—1274, 2003.
7. Amal A. Farag, S.Y. Elhabian, S.A. Elshazly and A.A. Farag, "Quantification of Nodule Detection in Chest CT: A Clinical Investigation Based on the ELCAP Study," Proc. of International Workshop on Pulmonary Image Processing in conjunction with MICCAI-09, Sept. 2009, pp. 149-160.
8. I. Sluimer, A. Schilham, M. Prokop, and B. van Ginneken, "Computer Analysis of Computed Tomography Scans of the Lung: A Survey," IEEE Transactions on Medical Imaging, vol. 25, No. 4, pp. 385–405, April, 2006.
9. B. Ginneken, B. Romeny and M. Viergever, "Computer-Aided Diagnosis in Chest Radiography: A Survey," IEEE Transactions on Medical Imaging, Vol. 20, 2001.
10. A. Farag, A. El-Baz, G. L. Gimel'farb, R. Falk, M. Abou El-Ghar, T. Eldiasty, S. Elshazly, "Appearance Models for Robust Segmentation of Pulmonary Nodules in 3D LDCT Chest Images," Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'06), Copenhagen, Denmark, October 1-6, 2006, pp. 662-670
11. Amal Farag, Lung Nodule Modeling and Detection for Computerized Image Analysis of LDCT Imaging of the Chest, Master of Engineering Thesis, University of Louisville, April 2009.
12. ELCAP public lung image database - www.via.cornell.edu/databases/lungdb.html
13. H. Takizawa, S. Yamamoto, T. Matsumoto, Y. Tateno, T. Iinuma and M. Matsumoto, Recognition of lung nodules from X-ray CT images using 3d mrf models. In: Proceeding of International Congress of Computer Assisted Radiology and Surgery (2001), pp. 605–614
14. Mendonça, P.R.S., Bhotika, R., Zhao, F., Miller, J.V.: " Lung nodule detection via Bayesian voxel labeling," Int. Conf. on Information Proc. and Medical Imaging. (2007) 134–145
15. Edwards G.J. Taylor C.J. Cootes, T.F. Active apperance models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6), 2001.
16. I. Matthews, and S. Baker, "Active Appearance Models Revisited," Int. J. of Computer Vision, pp. 135-164, 2004.
17. Amal Farag, James Graham, Aly Farag and Robert Falk, "Lung Nodule Modeling – A Data-Driven Approach," 5th International Symposium on Visual Computing (ISVC'09), Nov. 30 – Dec. 2, 2009, Las Vegas, Nevada, USA.
18. H. Abdelmunim and A. A. Farag, "Shape Representation and Registration using Vector Distance Functions," Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07), Minneapolis, MN, June 18-23, 2007.

Software Demonstrations

Image Analysis Software Prototype for the Segmentation and Quantification of Lung Nodules based on the Cognition Network Technology®

M. Athelogou, J. Kim, M. Kietzmann, R. Korn
Definiens AG, Trappentreustrasse 1, 80339 Munich, Germany

Abstract. We present a software prototype for semi-automatic segmentation and quantification of lung nodules. The system is based on an image analysis platform, which enables users to develop different image analysis solutions for different data modalities. The technology behind the platform is the Definiens Cognition Network Technology®, which is a context-driven approach based on objects and knowledge. The presented system quantifies the diameters required by the Response Evaluation Criteria in Solid Tumors (RECIST) criteria and calculates the volume of the lung nodules.

Keywords: semi-automatic segmentation, objects, context, knowledge, volume.

1 Introduction

For Computer Aided Detection and Diagnosis systems (CAD) the image analysis is an important component. Fields of applications, which need extensive image data analysis, are for example automated tumor quantification by using CT, MRI, ultrasound data for mammography, detection of polyps in colonoscopy, detection and localization of ischemic areas in the brain, detection and volumetric measurements for liver and lung tumors, measurement of aneurisms in blood vessels [8]. Using analysis and reconstruction of 3D image data, three dimensional reconstructions of organs, vessels, muscles, and bones can be achieved. 3D data sets may be used for the reconstruction of individual anatomical models. The development of algorithms and methods for "multiorgan-extraction" is the goal of actual research and development [3, 4, and 7]. Most of the methods which are used for the development of the appropriate image analysis solutions are pixel based. Therefore they show strong limitations in using context and domain knowledge. The standard for performing quantification of lung nodules is RECIST [5]. The RECIST evaluation criteria require the measurement of nodule diameters. There are several initiatives like the Quantitative Imaging Biomarkers Alliance (QIBA), which evaluate the quantification of volumetric changes of nodules in CT data sets as a *standard* for the assessment of therapy response. There are also several works, which focus on the assessment of volumetric changes of lunge nodules (e.g., [9]). Some of them are using semi automated algorithms to calculate tumor volume and other parameters [6].

The software we present enables the measurement of the volume of lung nodules. The aim of the software demonstration is to show the advantages of context driven,

object and knowledge based versus pixel based algorithms concerning lesion measurements in CT series. For doing that, we use as example the quantification of lung nodules in such data sets.

2 Materials and Methods

The Cognition Network Technology® is a context-based approach based on objects and knowledge. The technology incorporates elements from semantic networks, description logics, and functional programming. It can be applied to different kinds of data [1, 2]. Such data are for example images of different dimensionality, magnification and resolution and from different acquisition modalities.

In pixel-based image analysis algorithms, objects of interest are extracted through a series of filters, including intensity thresholds, proximity, gradients, and edges. These mathematical filters are generally applied to a whole image or to "regions of interest" in images, evaluating pixels or pixel fields in relation to neighboring pixels. New images are produced, transforming "regions of interest" from the images into "objects of interest" that can be extracted with a simple threshold. Contents of biomedical image data are often very complex. In general, such problems cannot be solved with conventional procedures of applying a sequence of pixel filters combined with thresholds.

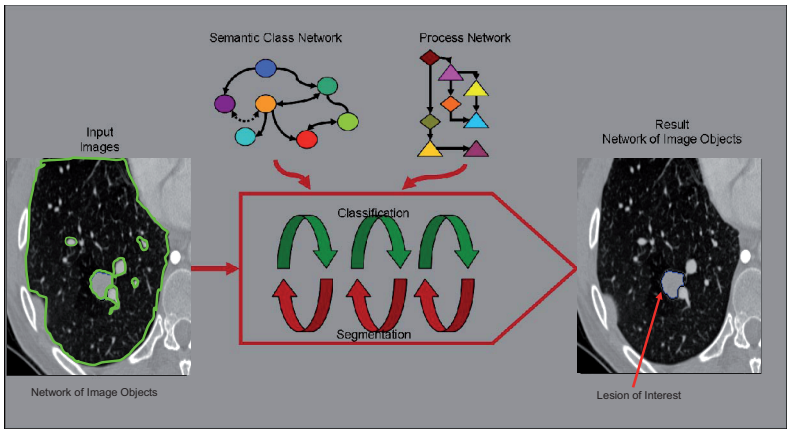


Fig. 1. Schematic representation of the Definiens Cognition Network Technology for image data analysis.

In contrast to such pixel-based approaches, Definiens Cognition Network Technology® creates not only the final product (objects of interest) but also automates the intermediate steps involved in the process. The process evolves during segmentation, rather than being predefined. Thus, semantic relationships are generated by classifying all image objects, including the intermediate objects. This image analysis technology invokes an evolutionary process comprised of alternating classification and segmentation which may be applied to different data modalities

synchronously (Fig. 1). Information about the image objects and their relationships gathered *during* the course of segmentation can then be used effectively for future purposes. Alternating segmentation processes may take place in a single image or image modality by using contextual information about objects drawn from the segmentation and classification of additional images or image modalities. Finally, a *network of objects* in the same or in different images of the same or of different modalities is created in order to understand the overall system. Subsequent data analysis relies on the development of these "*Cognition Networks*", consisting of objects and their mutual links. "*Cognition Network Language*" (CNL) is the corresponding Graphical User Interface meta language that allows efficient development of rule-based algorithms to create and process Cognition Networks (Fig. 1). CNL consists of four basic data structures: *Processes, Domains, Image Objects* and *Image Object Classes*. By selecting and parameterizing the Processes the particular processing algorithms are specified for a given programming step, whereas through the definition of a Domain the system is guided to the data structure that is going to be processed. The Processes define "what" and the Domains "where" processing takes place. The most important Domains are Pixel Level Domains for filtering and initial segmentation operations; Image Object Domains that specify objects based on location in the network, their classification, and/or its specific attributes for the object-oriented processing, and Image Object Relation Domain which allows the navigation in the image object network. Through the formulation of domains and semantics, CNL supports the modeling and usage of specific expert domain knowledge within rule sets. The platform allows for the analysis of an arbitrary number of image layers (channels) simultaneously.

By operating over the presentation of the structures of interest on different scales in the object network, measurements, such as morphology, embedding, distribution, composition, structure, intensities, relative location and co-localization are possible.

Therefore features are defined that describe the individual properties of objects, the relationships of objects to their neighborhoods, the mutual relationships of groups of objects in the same or in different images, image sets and modality related features and metadata, and image modality properties. The technology defines context-neutral and context-sensitive features.

Based on this technology an image analysis platform has been developed (Definiens Developer XD). The platform uses the concept of "maps". These maps may contain copies of the same or of different data modalities, such as X-ray images and excel tables [7]. Algorithms for segmentation and classification of data can be applied to different maps with the same image data in order to achieve the best segmentation results. Multiple data modalities may be stored and analyzed in maps synchronously. While each map can be analyzed independently, the system utilizes image analysis results from one image or from multiple images and image modalities to provide context for the analysis of objects in other images in an alternating, iterative process. In this way intermediate analysis results of objects in an image or in data of a certain modality may be used in the next analysis steps as context for the analysis of data in the same or in other data as image data of the same or different modalities. In this way, analysis of different objects of interest can be done in multiple images and image modalities simultaneously.

At the beginning of this analysis process, simple, knowledge-based "seed objects" are automatically segmented and classified for use in the initial data processing. These seed objects are then utilized in subsequent analysis steps in order to provide context

for the detection, segmentation and classification of other objects in the same image or in other images. As the technology enables multi-scale data analysis, multiple levels of information can be extracted from image data. They can also be analyzed in isolation from other objects, or in the context of other objects. Objects can be linked on demand and placed into object networks. Semantic objects, such as a lung lesion or the lungs itself, that are represented by multiple images may be linked to each other in multidimensional objects.

3 Results

We used Definiens Developer XD to develop a software prototype dedicated to the segmentation and quantification of lung nodules in CT image stacks. This "lung nodule" software prototype comprises a CNL Rule Set. As the CNL enables the automated measurement of object properties by using predefined and user defined object features, a detailed quantification of the nodule properties is made. Such features measure for example signal intensity, surface, volume, diameters, both as absolute and relative values. The system enables at least a detailed quantification of the nodules properties.
The major steps of the developed CNL Rule Set are the following: In the first step, the upper and lower threshold for the intensity of the lung nodule are calculated based on (a) the window settings and (b) the intensity distribution in the neighborhood of a seed point. Both the window settings and the seed point are provided by the user. Using the threshold values, an Image Object is created by region growing. If it is not isolated (touching the pleural wall or a vessel), the segmentation process is refined by using context Image Objects for the pleural wall and/or vessels.
Figure 2 shows three examples from the public database of the Lung Image Database Consortium (LIDC) with screen-shots of the annotation by a radiologist (left) and the segmentation result (right). Three types of lung nodules are depicted: isolated (top), juxtavascular (middle), and juxtapleural (bottom). The results of the volume quantification are stated below the corresponding screen-shots.

4 Discussion

The software prototype shows the ability of the Definiens Cognition Network Technology, the ability of CNL and the flexibility of the concerning image analysis

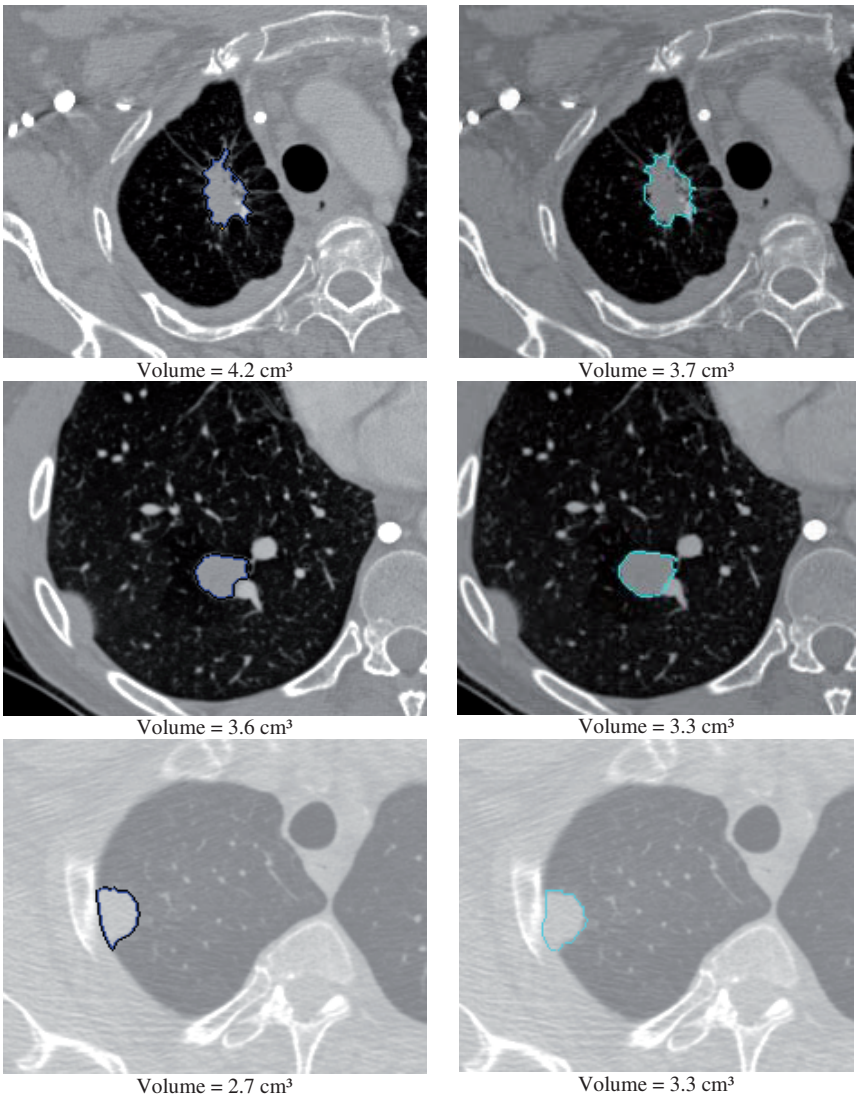


Fig. 2. Segmentation results for three examples from the LIDC database. The annotation of a radiologist and the corresponding segmentation result are shown on the left and the right, respectively.

platform for the development of tailored image analysis solutions. Such solutions can further be developed in order to fulfill the needs of the domain experts. Especially the solution for the lung nodules can be further developed in a CAD system and can support radiologists in their daily work. Although further evaluation of the system is needed, the first results are very promising. This kind of CAD prototype development may be used as an educational example for the development of further CAD systems as they are needed in different domains like radiology, pathology etc..

References

1. Athelogou, M., Schmidt, G., Schäpe, A., Baatz, M. and Binnig, G.: Cognition Network Technology – A Novel Multimodal Image Analysis Technique for Automatic Identification and Quantification of Biological Image Contents, in Imaging Cellular and Molecular Biological Functions. S. L. Shorte and F. Frischknecht (eds.), Springer (2007)

2. Athelogou, M., Schönmeyer, R., Schmidt, G., Schäpe, A., Baatz, M., Binnig, G.: Bildanalyse in Medizin und Biologie Beispiele und Anwendungen. Medizintechnik Life Science Engineering, E. Wintermantel, Suk-Woo (eds.), Springer (2008)

3. Schmidt, G., Athelogou, M., Schönmeyer, R., Korn, R., Binnig, G.: Cognition Network Technology for a fully automated 3D segmentation of liver. In Proc MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge. 125-33 (2007).

4. Herzog, P.: Innovative technology for automated lymph node analysis,. RSNA 2008, EH online

5. Eisenhauer, E.A., Therasse, P., Bogaerts, J., Schwartz L.H., Sargent, D., Ford, R.: New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 45:228-47 (2009).

6. Zhao, B., Schwartz, L.H., Moskowitz, C.S., Ginsberg, M.S., Rizvi, N.A., Kris, M.G.: Lung Cancer: Computerized Quantification of Tumor Response – Initial Results, Radiology: 241: 3 (2006)

7. Schönmeyer, R., Athelogou, M., Sittek, H., Ellenberg, P., Feehan, O., Schmidt, G., Binnig, G.: Cognition Network Technology prototype of a CAD system for mammography to assist radiologists by finding similar cases in a reference database. DOI: 10.1007/s11548-010-0486-8, Int J Comput Assist Radiol Surg (2010)

8. Gilbert, F. J., Lemke, H.: Computer Aided Diagnosis, BJR British Journal of Radiology Volume 78, Special Issue (2005)

9. A. Reeves et al. The VOLCANO'09 Challenge: Preliminary Results. Technical report. http://www.lungworkshop.org/2009/proc2009/353.pdf