Computers in Biology and Medicine 76 (2016) 143–153

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

An approach for reducing the error rate in automated lung segmentation

Gurman Gill [a,b], Reinhard R. Beichel [a,b,c,*]

[a] Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242, United States
[b] The Iowa Institute for Biomedical Imaging, The University of Iowa, Iowa City, IA 52242, United States
[c] Department of Internal Medicine, The University of Iowa, Iowa City, IA 52242, United States

ARTICLE INFO

Article history:
Received 15 March 2016
Received in revised form
21 June 2016
Accepted 22 June 2016

Keywords:
Lung segmentation
Segmentation fusion
Classification
Computed tomography

ABSTRACT

Robust lung segmentation is challenging, especially when tens of thousands of lung CT scans need to be processed, as required by large multi-center studies. The goal of this work was to develop and assess a method for the fusion of segmentation results from two different methods to generate lung segmentations that have a lower failure rate than individual input segmentations. As basis for the fusion approach, lung segmentations generated with a region growing and model-based approach were utilized. The fusion result was generated by comparing input segmentations and selectively combining them using a trained classification system. The method was evaluated on a diverse set of 204 CT scans of normal and diseased lungs. The fusion approach resulted in a Dice coefficient of 0.9855 ± 0.0106 and showed a statistically significant improvement compared to both input segmentation methods. In addition, the failure rate at different segmentation accuracy levels was assessed. For example, when requiring that lung segmentations must have a Dice coefficient of better than 0.97, the fusion approach had a failure rate of 6.13%. In contrast, the failure rate for region growing and model-based methods was 18.14% and 15.69%, respectively. Therefore, the proposed method improves the quality of the lung segmentations, which is important for subsequent quantitative analysis of lungs. Also, to enable a comparison with other methods, results on the LOLA11 challenge test set are reported.

1. Introduction

Lung segmentation is one of the first processing steps in computer-aided quantitative lung image analysis. For high throughput applications with tens of thousands of data sets to be analyzed—as required by large multi-center trials—fully automated lung segmentation approaches with high robustness and low error rate are imperative to minimize the need for human intervention (i.e., manual correction). This is especially important when segmenting lungs with lung disease.

A number of papers describing lung segmentation algorithms have been published, and a comprehensive review can be found in [1]. Basically, methods developed can be grouped into three categories given below:

(a) Simple, low complexity methods like region growing [2,3], which are based on simple assumptions (e.g., density range of lung tissue). These methods typically work well for normal lungs, but may fail in the case of diseased lungs or imaging artefacts. An advantage of such methods is the low computational complexity.
(b) Advanced, more robust algorithms that try to overcome the problems of category (a) and typically show higher computational complexity. Examples in this category include approaches based on registration [4], lung shape models [5,6] and advanced threshold-based segmentations utilizing adaptive border matching [7] or texture features [8].
(c) Hybrid approaches that try to use a method in category (b) only if a result produced with method in category (a) is classified as failed based on some heuristics (e.g., assumptions about lung volume). Representatives in this category are the work of Rikxoort et al. [9] and Mansoor et al. [10]. The main motivation behind such approaches is to take advantage of the low computational complexity of methods in category (a), but with the optional performance of more advanced methods in category (b). The behavior of methods in group (c) depends on whether the heuristics for switching to a method in category (b) work or not. Furthermore, with increasing computing power combined with lower hardware costs, hybrid methods

* Corresponding author at: Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242, United States.
E-mail address: reinhard-beichel@uiowa.edu (R.R. Beichel).

http://dx.doi.org/10.1016/j.compbiomed.2016.06.022
0010-4825/

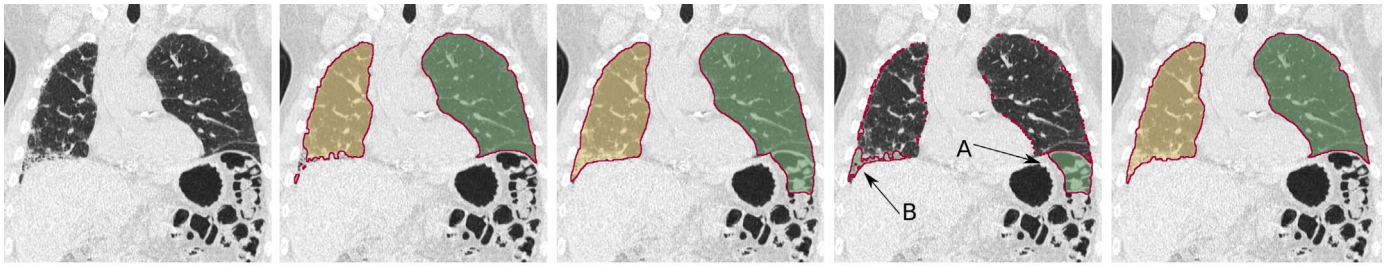

Fig. 1. Comparison of lung segmentation methods applied to a CT scan of a lung with IPF. (a) Coronal slice of the CT scan. (b) Region growing segmentation result. (c) Model-based segmentation result. (d) Difference volume between the segmentations in (b) and (c); arrows indicate components corresponding to (A) leak into colon and (B) lung tissue affected by IPF. (e) Result of the fusion approach in which component (A) is rejected and component (B) is accepted.

may become less attractive, because as computational costs become less important, more advanced methods can be utilized routinely.

All methods in these categories have different pros and cons and are based on different design assumptions that might or might not hold. In the case of pathological lungs, it is expected that the likelihood of failure of methods in category (b) is lower than for the ones in (a), but despite all the efforts, they can (locally) fail too. For example, Fig. 1(a) depicts a coronal CT cross-section of a lung with idiopathic pulmonary fibrosis (IPF). Corresponding segmentations of a region growing and model-based [6] method are shown in Figs. 1(b) and (c), respectively. As can be seen in the difference image of both segmentations (Fig. 1(d)), both methods show local segmentation errors due to different reasons like a violation of the assumption of a specific lung density range (Fig. 1 (b)) or problems with model initialization (Fig. 1(c)).

In this paper, we propose a segmentation fusion approach based on a classification framework, which selectively combines (components of) two independently generated lung segmentations to form a new segmentation result with no or reduced errors. The idea behind this approach is to take advantage of the strength of both methods, but without including their errors. In our case, the two segmentations are generated by a region growing and robust active shape model (RASM) based method [6] (Section 2). Compared to other lung segmentation approaches, it does not rely on a fallback method [9,10] where a more complex segmentation approach is chosen if the output of a simple region growing method is classified as being incorrect, nor does it simply combine segmentation results with a logic OR operation [8,10]. Instead, our approach follows a more flexible approach that can selectively combine components of both lung segmentation results, as demonstrated in Fig. 1(e). We assess fusion performance on a diverse set of 204 lung CT scans and provide a comparison to the performance of both input lung segmentations. Also, the fusion method can be easily adapted to different input segmentation methods by retraining of the classification system.

2. Selecting suitable input segmentation methods and prior work

In this section, we discuss the general requirements for selecting suitable input methods for our segmentation fusion approach and introduce the two segmentation approaches utilized in this paper.

2.1. Considerations and requirements

The overarching assumption of deploying a fusion approach is that existing lung segmentation methods are—to a certain degree—imperfect. Thus, algorithms can and will fail, especially when applied to a large number of medical data sets, as is the case in large multi-site studies (e.g., COPDGene[1]). The aim of the presented framework is to improve segmentation accuracy and reduce the failure rate by utilizing a segmentation fusion approach on two base segmentations. We assume that the base algorithms A and B are suited for lung segmentation and show already good performance, but will still fail in a number of cases. We note that such segmentation methods rarely produce complete failures (e.g., segmenting the air surrounding the patient instead of the lung). Typically, failures occur locally and are limited (e.g, leakage into colon, including the trachea, excluding a tumor, etc.). Instead of selecting method A or B, and having to deal with frequent occurring errors by time-consuming manual editing, the idea is to use both segmentation results selectively to produce a new segmentation C with no or reduced errors (i.e., lower error rate at a required accuracy level).

The ideal set of candidates for producing input segmentations A and B has non-overlapping weaknesses and strengths, resulting in (local) disagreement between methods. Fig. 2 provides several examples for a region growing and model-based lung segmentation approach that will be utilized in this paper. Differences in generated lung masks result in local volume components of disagreement (Fig. 2d), which can have many causes. Typically, they are caused by assumptions that methods make. As can be seen in Fig. 2, both input segmentation methods show non-overlapping weaknesses, and thus, are suited for a fusion approach.

Given the results of two lung segmentation algorithms A and B, we assume that if both methods label a voxel as lung tissue, then the likelihood of the voxel representing lung is high. Therefore, it will be labeled as lung by our fusion method. For components of disagreement, a trained classifier is utilized to individually decide which components should be added to the volume of mutual agreement between both methods, resulting in the final output segmentation of the algorithm. Note that classification is performed on components of disagreement (i.e., volume chunk). Therefore, all voxels of the volume chunk will receive the same label by the classifier.

2.2. Method A—region growing based segmentation

The region growing segmentation $\mathcal{V}_{RG}$ is obtained using a threshold of -500 HU. The seeds for region growing are identified automatically as follows. Let s_x, s_y, and s_z denote the size of a CT data set in x-, y-, and z-direction, respectively. First, initial seeds are placed. For the left lung, two initial seeds are generated at $\left(\frac{1}{3}s_x, \frac{1}{2}s_y, \frac{2}{3}s_z\right)$ and $\left(\frac{1}{3}s_x, \frac{2}{3}s_y, \frac{1}{2}s_z\right)$. For the right lung, two initial seeds are placed at $\left(\frac{2}{3}s_x, \frac{1}{2}s_y, \frac{2}{3}s_z\right)$ and $\left(\frac{2}{3}s_x, \frac{2}{3}s_y, \frac{1}{2}s_z\right)$. Second, near

[1] http://www.copdgene.org