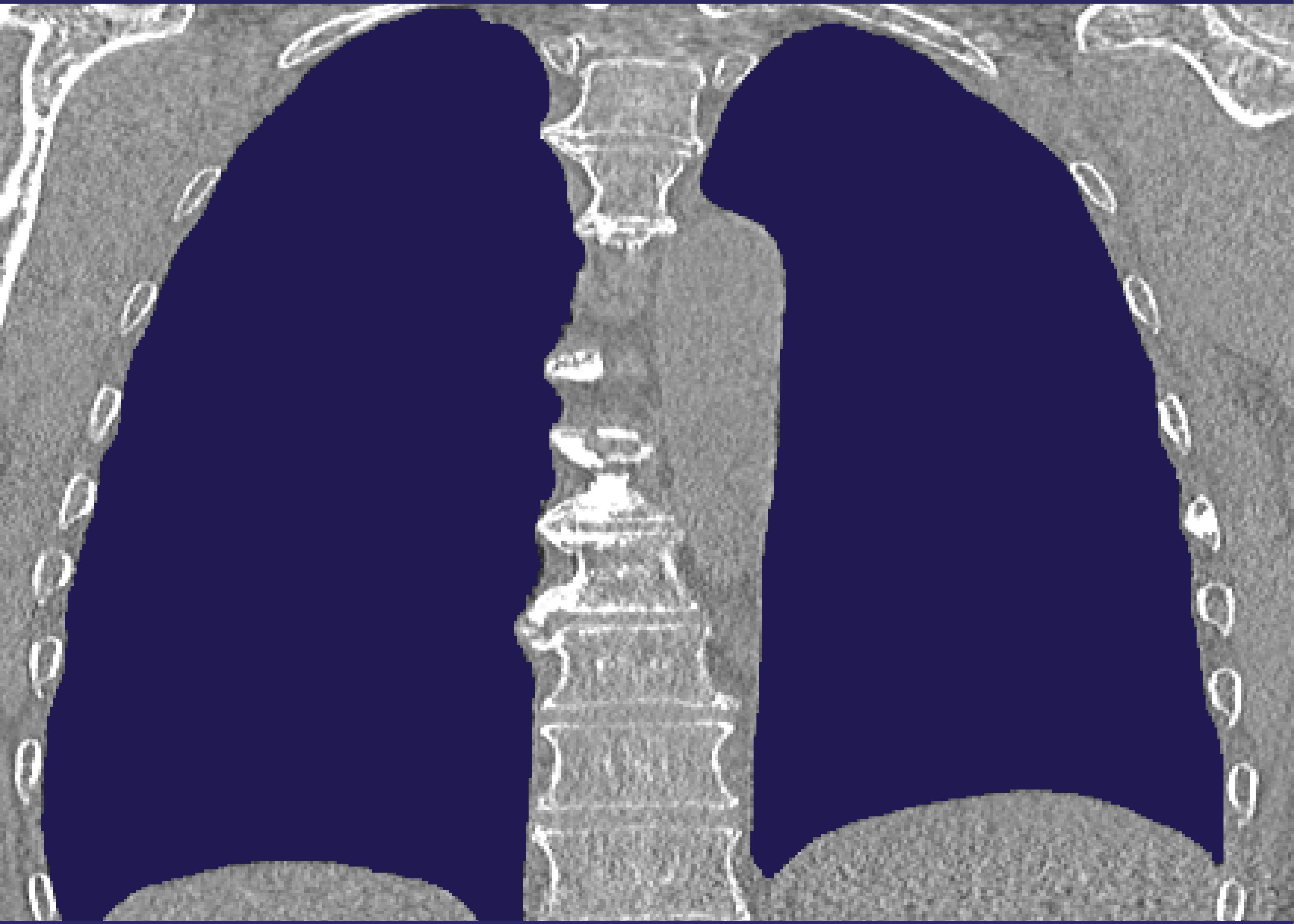
Lung Segmentation Using Multi-atlas Registration and Graph Cuts



Thomas Holm Sandberg
Master's thesis, 2014
Biomedical Engineering and Informatics

AALBORG UNIVERSITY
STUDENT REPORT

Lung Segmentation Using Multi-atlas Registration and Graph Cuts

Master's thesis
4th semester master, Biomedical Engineering and Informatics
Group 14gr1071
Contact: tsandb09@student.aau.dk

Author

Thomas Holm Sandberg
Stud.cand.polyt.
Aalborg University

Supervisors

Lasse Riis Østergaard
Associate Professor, M.Sc., Ph.D.
Aalborg University

Anne Sofie Korsager
Ph.D. student, M.Sc.
Aalborg University

Isabel Pino Pena
Ph.D. student, M.Sc.
Aalborg University

Page Numbers: 70 pages

Department of Health Science and Technology
Aalborg University
Fredrik Bajers Vej 7D2
DK-9220 Aalborg
Denmark

June 3, 2014

Abstract

Imaging of the lungs using computed tomography (CT) are used by clinicians to diagnose and study the progression of emphysema associated with patients suffering from chronic obstructive pulmonary disease (COPD). Risk factors for development of COPD are tobacco smoking and exposure to air pollution for longer durations, which may cause destruction of the alveoli in the lungs, leading to lung function decline by reducing the lungs ability to exchange gases. This is visible in chest CT imaging of the lungs, where density and texture of the lungs are altered by the destruction of lung parenchyma (lung tissue) caused by emphysema. Diagnosis of emphysema is made using pulmonary function tests (PFT), such as spirometry, which may be used to classify emphysema severity and classify emphysema into different stages. Recent studies show that lung function decline and emphysema severity may be predicted and classified using computational analysis of the lungs from chest CT scans. As a prerequisite for automated analysis of chest CT scans, the lungs must be segmented, which necessitates an automatic segmentation approach, as manual delineation can be tedious, time-consuming and subjective. This thesis present an automatic method for the segmentation of the lungs from chest CT scans based on multi-atlas registration and graph cuts.

The method combines global anatomical shape information, based on multi-atlas registration from a set of individual atlases, with intensity modeling and local voxel neighborhood modeling into a graph cut segmentation framework. A graph is constructed with nodes for each voxel and edge weights based on the spatial prior, intensity model and neighborhood model, and the optimal segmentation is obtained using graph cut optimization, equivalent to a maximum a posterior probability estimation or an energy minimization. The method has previously been applied in studies for segmentation of the hippocampus from MRI and segmentation of head and neck tissue from CT. The performance of the method will be evaluated on a dataset of 19 chest CT scans with ground truth segmentations serving as individual atlases and reference segmentations in a leave-one-out cross-validation experiment using the Dice similarity coefficient (DSC) as metric of performance and compared to a multi-atlas based segmentation method. The method will also be evaluated on a new dataset of high resolution CT (HRCT) scans of 10 healthy subjects and 10 patients with emphysema. The method will be evaluated by visual inspection and quantitatively, where manual ground truth segmentations are available.

The graph cut segmentation method performed significantly better compared to the multi-atlas segmentation method in the leave-one-out cross-validation experiment with a DSC of 0.9877 and 0.9858, respectively. The graph cut segmentation method successfully segmented the lungs on the HRCT dataset of the 10 healthy subjects and 10 patients with emphysema and showed better performance than the multi-atlas segmentation in terms of DSC where manual delineation was available. The graph cut method showed comparable performance to other studies on lung segmentation, however, more data from subjects with emphysema and other lung diseases are necessary for direct comparison and further evaluation of performance.

The graph cut segmentation method provides an accurate automatic segmentation of the lungs from chest CT scans, providing an objective segmentation method with no manual intervention needed and an alternative to the time-consuming and subjective manual segmentation. The segmentation may serve as the basis for further computational analysis of the lungs.

Resumé

Billedskanninger af lungerne ved brug af computertomografi (CT) bruges af klinikkere til diagnose og vurdering af sygdomsforløb af lungeemfysem hos patienter med kronisk obstruktiv lungesygdom (KOL) . Risikofaktorer for udvikling af KOL inkluderer længerevarende tobaksrygning og eksponering for luftforurening, hvilket kan ødelægge alveolerne i lungerne og føre til lungefunktionsnedsættelse pga. nedsat gasudveksling i lungerne. Dette kan ses i CT-skanninger af brystkassen, hvor billedintensitet og teksturen i lungerne ændres, som følge af ødelæggelse af lungevæv pga. lungeemfysem. Diagnose af lungeemfysem stilles vha. lungefunktionstests, eks. spirometry, som bruges til at klassificere graden af lungeemfysem og klassificere lungeemfysem i forskellige stadier. Nyere studier viser, at lungefunktionsnedsættelse og emfysem stadieinddeling kan forudsiges og klassificeres vha. computerbaseret analyse af lungerne fra CT-skanninger af brystkassen. En forudsætning for automatiseret analyse af CT-skanninger af brystskanninger er, at lungerne segmenteres, hvilket nødvendiggør en automatisk segmenteringsmetode, da manuelle indtegninger kan være anstrengende, tidskrævende og subjektive. Denne rapport præsenterer en automatisk metode til segmentering af lungerne fra CT-skanninger af brystkassen baseret på multi-atlas registrering og graph cuts.

Metoden kombinerer global, anatomisk shape information baseret på multi-atlas registrering fra et sæt individuelle atlas, med intensitetsmodellering og lokal voxel neighborhood-modellering i et graph cut segmenterings-framework. En graf konstrueres med knudepunkter for hver voxel og kantvægte baseret på den spatielle model, intensitetsmodellen og neighborhood-modellen, og den optimale segmentering opnås ved graph cut optimering, ækvivalent til en maximum a posterior probability estimering eller en energiminimering. Metoden har tidligere været anvendt i studier til segmentering af hippocampus fra MR samt segmentering af hoved- og nakkevæv fra CT. Metodens ydeevne vil blive evalueret på 19 CT-skanninger af brystkassen, med ground truth segmenteringer som individuelle atlas og reference segmenteringer, i et leave-one-out krydsvalideringsforsøg med Dice similarity koefficienten (DSC) som mål for ydeevne og sammenlignet med en multi-atlas baseret segmenteringsmetode. Metoden evalueres yderligere på et nyt datasæt med højopløselige CT-skanninger (HRCT) af 10 raske personer og 10 patienter med lungeemfysem. Metoden vil blive evalueret visuelt og kvantitativt på 1 rask person og 1 patient med lungeemfysem, hvor manuelle reference segmenteringer er tilgængelige.

Graph cut segmenteringsmetoden opnåede signifikant bedre resultater sammenlignet med multi-atlas segmenteringsmetoden i leave-one-out krydsvalideringsforsøget med en DSC på hhv. 0.9877 og 0.9858. Graph cut segmenteringsmetoden segmenterede lungerne i HRCT datasættet med 10 raske personer og 10 patienter med lungeemfysem og opnåede bedre ydeevne end multi-atlas segmenteringen mht. DSC ved den raske person og patienten med lungeemfysem, hvor manuelle indtegninger var tilgængelige. Graph cut segmenteringsmetoden opnåede tilsvarende resultater som i andre lungesegmenteringsstudier, men yderligere data fra patienter med lungeemfysem kræves for direkte sammenligning og yderligere evaluering af ydeevne.

Graph cut segmenteringsmetoden giver en præcis, automatiseret lungesegmentering fra CT-skanninger af brystkassen, vha. en objektiv segmenteringsmetode uden brug af manuel intervention, som alternativ til tidskrævende og subjektive manuelle segmenteringer. Denne segmentering kan danne grundlag for yderligere computerbaseret analyse af lungerne.

Contents

Preface

This thesis has been submitted in partial fulfillment of the requirements for the Master of Science (MSc) degree in Biomedical Engineering and Informatics at The Department of Health Science and Technology, Aalborg University. The work in this thesis has been carried out during the final semester of the master's degree programme in Biomedical Engineering and Informatics at Aalborg University from February - June 2014.

The thesis presents a method for automatic segmentation of the lungs from chest CT scans using multi-atlas registration and graph cuts.

References are cited using the IEEE citation style, e.g. [1] refers to the first reference in the bibliography provided in the end of the thesis. Figures, tables and equations are numbered according to the chapter in which they appear, e.g. the first figure in Chapter 2 is Figure 2.1.

Aalborg, June 2014.
Thomas Holm Sandberg

Chapter 1

Introduction

Medical imaging techniques produce images of the inside of the human body, often non-invasively, for clinical and/or research purposes. Medical imaging modalities include X-ray imaging, computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, endoscopy, functional magnetic resonance imaging (fMRI), positron emission tomography (PET) and many more. The resulting images are used by doctors and clinicians for diagnosis, treatment planning, disease progression monitoring and disease evaluation. Furthermore, research in disease studies can be carried out on large amounts of imaging data. Visual inspection of the images by radiologists and medical experts aid in diagnosis and evaluation of diseases such as tumors, where inspection of the tumor images are used to evaluate size and severity and delineation of the tumor is used for radiation therapy planning, or in the diagnosis of chronic obstructive pulmonary disease (COPD), where images of the lungs from a chest CT scan are used to evaluate emphysema severity due to breakdown of lung tissue.

Although humans are very good at recognizing objects of interest in an image and radiologists are experts at inspecting medical images for diagnostic purposes, delineation of anatomical structures, tumors or other diseased tissue can be tedious, time-consuming and subjective. Automatic (or semi-automatic) segmentation in medical imaging can therefore aid in identifying and delineating anatomical structures of interest and accurate segmentation is often a crucial prerequisite for further computational analysis of the images. The segmentation can be used to improve challenging 3D visualization for diagnostic purpose or guidance in image-guided surgery, or it can be used to analyze the image information within the segmentation. This analysis can be based on textural features or density measures and point out regions of interest with diseased or abnormal tissue as is the case in computer-aided diagnosis (CAD) systems, where the radiologist can be aided in diagnosis by receiving a second opinion from the decision supporting system. While radiologists excel in qualitative assessment of images, computers excel in quantitative assessment of images and computational analysis is therefore important in order to analyze large amounts of data and to aid radiologists and researchers with better quantitative, time-saving and objective measures.

Having an accurate segmentation of a structure of interest makes it possible to do further computational analysis such as: shape analysis, calculate volume change, analyze and classify texture, measure dynamic properties from continuos imaging and better radiation therapy treatment planning, hereby more accurately targeting a malignant tumor. Patients suffering from COPD experience small airway disease (obstructive bronchiolitis) and parenchymal (lung tissue) destruction (emphysema). Diagnosis of COPD is done through patient history and using pulmonary function tests (PFT) such as spirometry, which is the gold standard for COPD assessment [1]. However, CT scans are also used in the assessment of COPD and studies show the ability to predict greater lung function decline in male smokers from density analysis of CT images [2], and that emphysema severity classification using texture analysis of lung tissue correlates well with pulmonary function tests [3].

A first-step in these prediction and classification methods is an accurate segmentation of the structure of interest, the lungs. Although segmentation of a wide variety of anatomical structures has been studied extensively for several years, challenges still remain and no ‘one size fits all’ solution exists. Different segmentation methods and approaches apply to different tasks and depend on the problem at hand. Segmentation can be challenging due to low contrast between tissues, which cause insufficient boundaries between regions of interest (ROIs) and background, and/or noise in the image due to technical limitations of the imaging acquisition technique. Also, structural changes in pathological tissue can lead to deformation in anatomical structures and lead to changes in attenuation values in CT imaging, altering the shape and contrast of the ROI in the images which further complicates the segmentation task.

Several studies have attempted to accurately segment the lungs and pulmonary region from chest CT scans [4–7]. This thesis will investigate segmentation methods for segmenting the lungs from chest CT scans and propose a method for segmenting the lungs from chest CT scans in both healthy and emphysematous lungs. The perspective is to provide an accurate segmentation for further computational analysis of the lungs in healthy subjects and patients suffering from disease. The main focus will be on emphysematous lungs from patients with COPD, but the lung segmentation method could be applicable in analysis of other lung diseases, such as lung cancer. To provide the reader with the necessary background, the first part of this thesis will provide an introduction to:

- Imaging of the lungs
- Anatomy and physiology of the healthy and diseased lung
- Computer aided diagnosis
- Segmentation methods
- Previous work on lung segmentation

Following this, an outline of the thesis will be made, including chapters on the proposed segmentation method, data, results and discussion.

Chapter 2

Background

This chapter will provide the reader with the necessary background for the method chapter.

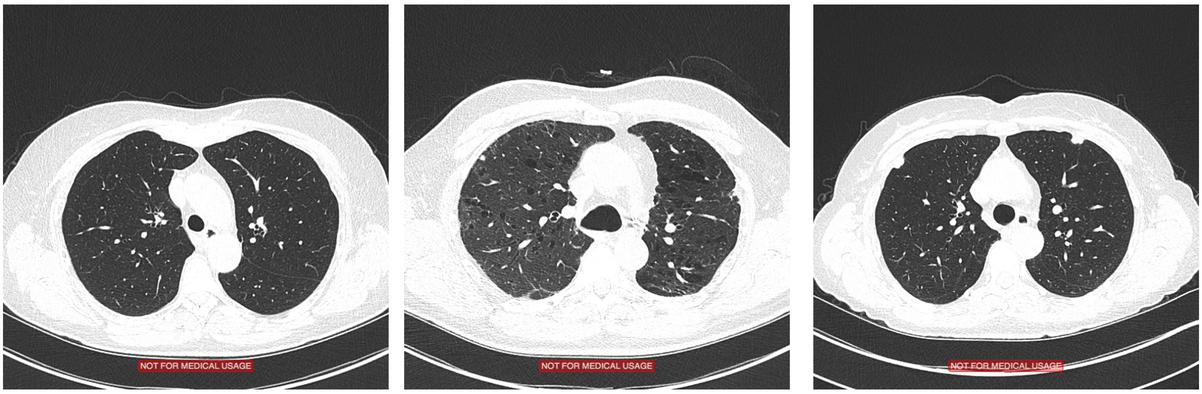
2.1 Imaging of the lungs

Computed tomography (CT) medical images are produced by a CT scanner where x-rays are projected through the subject. The CT scanner rotates the X-ray tubes around the subject in the scanner to obtain the tomographic images from the X-rays passing through the patient and detected by the detectors on the other side of the patient. Sir Godfrey Hounsfield invented the first commercial CT scanner in 1971, showing that cross-sectional image construction of the inside of the human body was possible by measuring the attenuation of X-rays projected through the body at various angles. It has since grown to become an important tool in everyday clinical practice for diagnostic purposes [8].

X-rays was discovered in 1895 by Wilhelm Conrad Röntgen and forms the basis for obtaining CT images. X-rays are electromagnetic radiation produced by fast deceleration of electrons [8]. The radiation energy depends on the electron velocity, which is accelerated between a cathode and anode in a X-ray tube. Important parameters of CT scanning are therefore the acceleration voltage, that determines the energy of the X-ray beam and the tube current, that determines the intensity of the radiation. Other parameters include the slice thickness and gantry tilt [8].

The number of photons in an X-ray beam decreases exponentially when going through an object and the attenuation is due to absorption and scattering. The is the reason why different tissue in the body are visible on the CT images and indeed the basis for why imaging of the inside of the body is possible, as different tissues have different attenuation coefficients. The Hounsfield unit (HU) is a scale of attenuation values used in CT imaging. The scale defines air as having the CT value of $-1,000$HU and water as having the CT value of 0HU [8]. Various tissue and organs of the body have different CT values and can therefore be distinguished on the CT images. Tissue or organs deviating from the normal CT values may indicate pathology and therefore, along with shape and structure alterations of organs, can be used for diagnostic purposes by radiologists.

The lungs have CT values below around -700HU as the lungs are primarily filled with air, which has a CT value of $-1,000$HU. However, structures within the lungs such as blood vessels, lymphs and connective tissue have higher CT values. The attenuation in the lungs can be altered by various pathology, including tumors, lung nodules, fibrosis, pleural wall thickening and emphysema, and CT scans may therefore be used for diagnosis in various lung diseases. For diagnosis of emphysema, high resolution computed tomography (HRCT) is especially reliable for detection of pathology of emphysema [9]. HRCT is a special thin-sectioning CT scanning technique with slice thickness around 1 mm, which makes it possible to detect and characterize small structures and tissue of the parenchyma and airways. Figure 2.1 shows HRCT scans from normal lungs, lungs with emphysema and lungs with pleural wall thickening.



(a) Normal **(b)** Emphysema **(c)** Pleural wall thickening

Fig. 2.1 CT scan of the lungs in subjects with: (a) normal lungs, (b) emphysematous lungs and (c) lungs with pleural wall thickening.

The emphysematous lungs of Figure 2.1 contains larger "black spots" than the normal lungs. This is due to air trapping and abnormal enlargement of the alveoli, which occur in patients with COPD and lead to the destruction of the alveolar walls. Inflammation and thickening of the bronchial walls may also be visible in HRCT of patients with COPD and emphysema [9]. The abnormal enlargements of the alveoli cause lower attenuation and thereby lower CT values, because of the destruction of lung tissue and more trapped air. In the image of the patient with pleural wall thickening it is visible how two bright "blobs" are present at the top pleural border of the lungs, which is due to the thickening of the wall.

HRCT of the lungs in patients with emphysema can aid in severity assessment of the emphysema and classification into the three major subtypes: centriacinar emphysema, panacinar emphysema, and distal acinar emphysema. Also, it may aid in classification of COPD into phenotypes such as primarily parenchymal disease or primarily airway disease [9].

Although diagnosis of COPD is done through patient history and spirometry tests, studies show the ability to predict lung function decline in smokers using CT images [2] and that emphysema severity classification from patterns in CT images correlates to spirometry tests [3]. This may lead to tools for computerized analysis and classification of COPD and emphysema from CT images used in screening trials, large scale studies or for individual assessment of patient scans.

2.2 Anatomy and physiology of the lungs

The function of the lungs is to oxygenate the blood in the body and transport gasses from the blood into the air of the surroundings, by making gas exchange between the alveoli and the blood. This happens by ventilation, pulling fresh oxygen-rich air into the lungs through the mouth (inspiration) and transporting oxygen-poor air with carbon dioxide through the mouth out into the surroundings (exhalation). This is accomplished by the movement of the chest cage and the diaphragm, which expand and constrict the intra-thoracic space, where the lungs reside [10].

The left and right lung are contained within the left and right pleural cavities. The pleura is a serous membrane, which wraps around each lung. The lungs are divided into distinct lobes, which are separated by deep fissures. The right lung has three lobes, while the left lung has two lobes. The lung follow the inner contours of the rib cage, which encloses the lungs. At the hilum, the medial surface of the lungs, the shape is more irregular, and the surface has grooves at the positions of the great vessels and the heart. As the heart is located to the left of the midline, the left lung has a larger impression than the right [11].

The upper airways of the lungs consist of the trachea and the right and left primary bronchi. The function of the trachea and primary bronchi are to convey air to and from the lungs. The trachea branches off to the right and left primary bronchi within the mediastinum and an internal ridge, the carina, separates the two bronchi. The mediastinum contains the structures, including the heart, greater heart vessels and esophagus, that sits between the right and the left lung and the sternum in the front and the vertebrae in the back. The hilum of the lung is a groove at the medial surface of each lung, which is the entry point for access to the lungs for the bronchi, pulmonary vessels, nerves and lymphatics. All these structures are anchored in a mesh of dense connective tissue. This entry point is called the root of the lung [11].

In Figure 2.2 the lungs including the structures of the mediastinum, airways and blood vessels are shown. The root of the lungs (the hilus / hilum) is where blood vessels and bronchi enter the lungs.

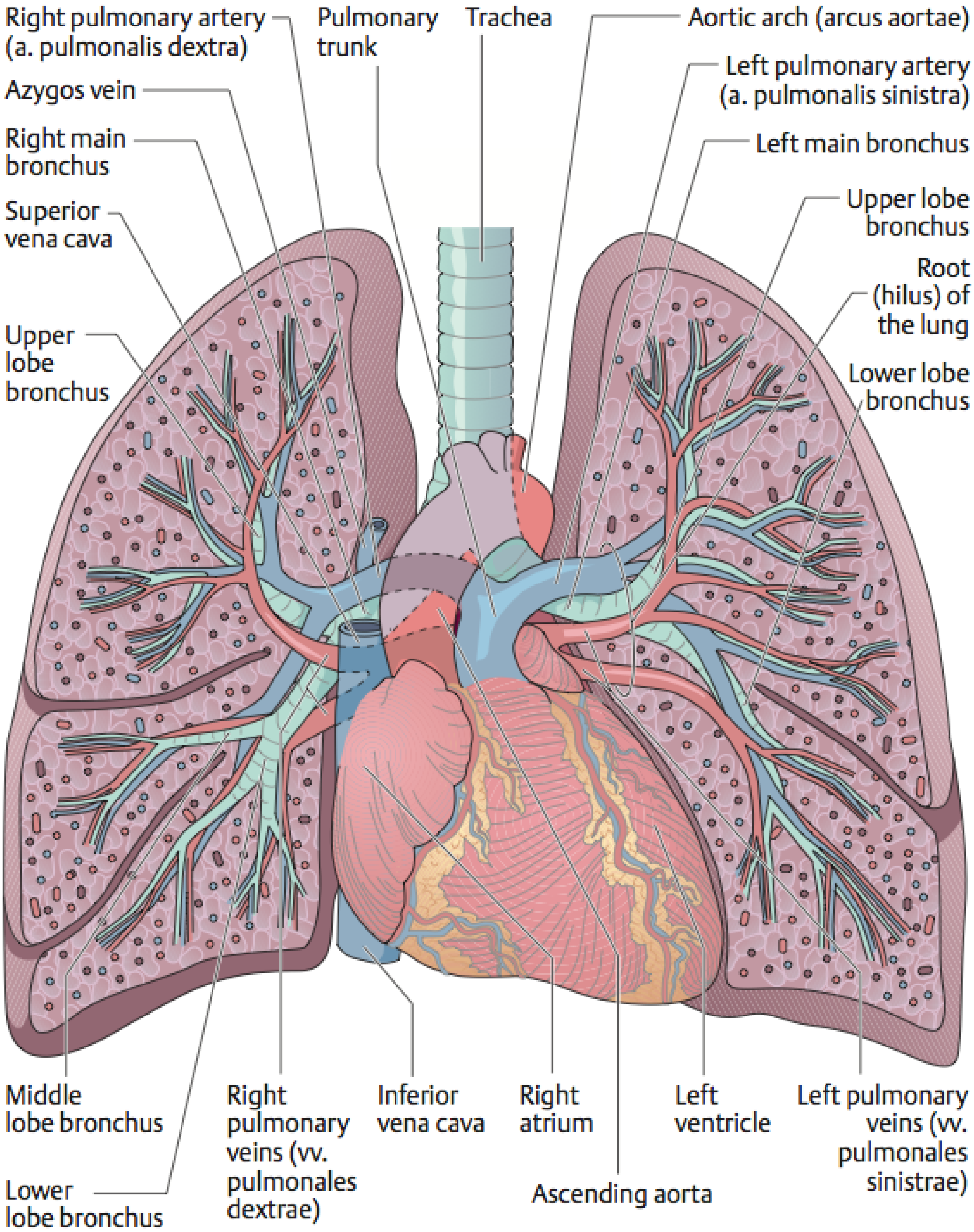


Fig. 2.2 The lungs with the airways and blood vessels as well as the mediastinum with heart, aorta and pulmonary blood vessels. Illustration from [10].

The primary bronchi are also called the extra-pulmonary bronchi, as they are outside the lungs. The branches within the lungs are called the intra-pulmonary bronchi collectively. The bronchi branches off into smaller and smaller structures until they end in the bronchioles, the terminal bronchioles and the smallest of the bronchioles, the respiratory bronchioles. Here they connect to the alveolar ducts and the alveoli responsible for gas exchange. The blood supply to the lungs enters at the hilum and branches along with the bronchi until they reach the smallest branches. Here, a network of capillaries surrounds each alveolus to allow for gas exchange [11].

In Figure 2.3 the lungs and its lobes, and the airways and its branches, are shown.

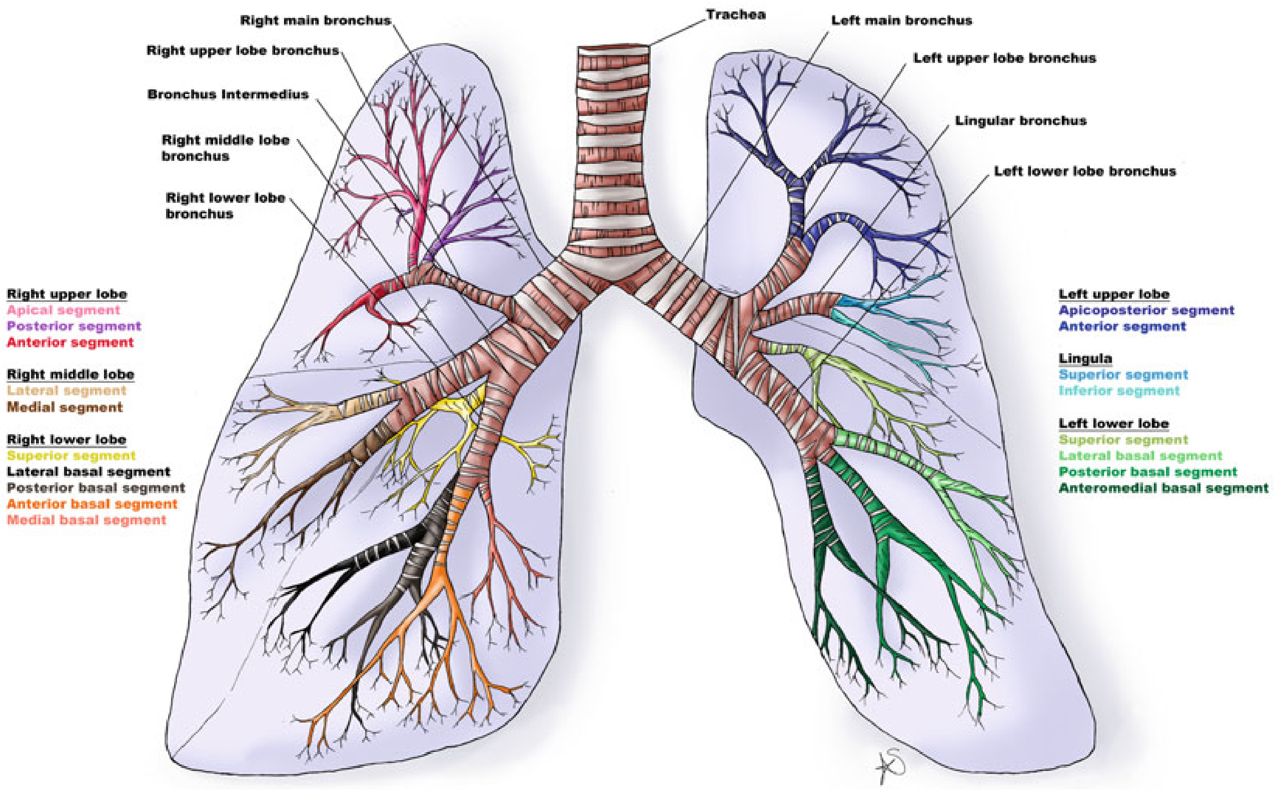


Fig. 2.3 Illustration of the lungs with its lobes and the airways and its branches. The upper airways transports air to and from the lungs through the trachea, the primary bronchi and branches off to the smallest respiratory bronchioles connecting to the alveoli, which does the gas exchange with the blood, transporting gasses to and from the blood. Illustration from [12].

The lung parenchyma refers to the functional lung tissue. This includes the tissue that makes up the alveoli, the functional gas exchange units of the lungs, the alveolar ducts and the respiratory bronchioles. However, other definitions of the parenchyma includes, besides the aforementioned tissue, also the bronchioles within the lungs as well as the blood vessels inside the lungs, thereby including all functional tissue within the lungs, excluding connective and supporting tissue [11, 13].

2.2.1 COPD

Chronic obstructive pulmonary disease (COPD) is a common respiratory disease characterized by persistent airflow limitation that usually progresses. It is the fourth leading cause of chronic morbidity and mortality in the United States and it is projected in 2020 to rank fifth of worldwide disease burdens [1]. COPD is a mixture of small airway diseases (obstructive bronchiolitis) and parenchymal destruction (emphysema). In obstructive bronchiolitis, the airways become inflamed and the walls thicken, which limits the flow of air. Emphysema is a permanent and abnormal enlargement of the alveoli, which leads to the destruction of the alveolar walls. COPD leads to structural changes and destruction of the lung parenchyma [1].

COPD is the result of exposure to various environmental particles over a longer period, most commonly tobacco smoking. The cigarette smoke and other inhaled particles cause inflammation, which is a normal response, but is modified in patients who develop COPD. This disrupts normal repair and defense mechanisms and lead to air trapping and limited airflow. The trapping of gas and flow limitations lead to a decrease in FEV_1 (forced expiratory volume in 1 second), which is a pulmonary function test (PFT) measuring lung function and used to diagnose COPD. The decrease is both caused by narrowing of airways and parenchymal destruction due to emphysema. Spirometry, the most common PFT, is required to make the diagnosis for COPD. A ratio between FEV_1 and FVC (forced vital capacity) of $\frac{\text{FEV}_1}{\text{FVC}} < 0.7$ after bronchodilator medication confirms presence of persistent airflow limitation and therefore COPD. Classification of COPD can be made into four GOLD (Global Initiative for Chronic Obstructive Lung Disease) stages: mild, moderate, severe and very severe. This is based on the $\frac{\text{FEV}_1}{\text{FVC}}$ ratio [1].

2.3 Computer-aided diagnosis

One of the goals of segmenting a structure of interest and recognizing tissue or objects of interest in medical images is to perform automated computational or quantitative analysis on these images supported by the segmentation. The analysis of anatomical structures or tissue can be used to recognize and categorize abnormal pathology in diseased subjects from that of healthy subjects. These systems are termed computer-aided diagnosis (CAD), which can assist medical professionals in the interpretation of medical images. A CAD system often consists of several steps, such as image processing, image feature analysis, and data classification. The output is commonly used as a ‘second opinion’ when assisting radiologists or other medical professionals interpretation of medical images and are in this way complementary to the expert before making the final decision [14]. CAD systems are part of the broader field of medical informatics and artificial intelligence in medicine and can be part of a clinical decision support system.

CAD systems for detection of breast cancer are part of daily clinical use in many hospitals in the United States and a lot of research has been made in the development of CAD systems for cancer detection in the breast, lungs and colon [14]. In [14], a CAD systems for aid in bone scintigraphy is developed by using temporal subtraction methods for whole-body bone scans. Here it is shown, that radiologists benefit from the CAD system by significantly increasing the sensitivity of the detection of interval changes in bone metastasis by the radiologists on scintigraphy images compared to not receiving the opinion from the CAD. Also, in [14] a CAD system based on artificial neural networks (ANN) are tested for diagnosis of lung nodules to distinguish between benign and malignant lung nodules on CT chest scans of the lungs. Here it is shown, that the radiologists receiving assistive opinions from the ANN CAD, incorporating both clinical parameters and image features, benefitted from the CAD information by improving differentiation of benign lung nodules from malignant.

Several studies have shown, that features derived from CT images correlates with patients diagnosed with COPD and features from CT images could therefore potentially be used as biomarkers for COPD and included in a CAD system [2, 3, 15–19]. Features derived from CT images can be based on density (percentage of voxels below a threshold), based on texture analysis (local binary patterns, texture measures from gray level run length matrices e.g. Haralick texture features, Gaussian Markov random fields) and registration-based (features from biomechanical lung analysis). Several of these features can be combined, and in addition combined with clinical parameters, to improve classification. Through classification, association and correlation between these features derived from CT images and pulmonary function tests, lung function decline, COPD severity and emphysema distribution has been shown.

In Figure 2.4 an example of an output from a classification algorithm based on Gaussian Markov random fields texture analysis is shown. Voxels in the CT scan of the lungs are classified according to lung parenchyma pathology and severity.

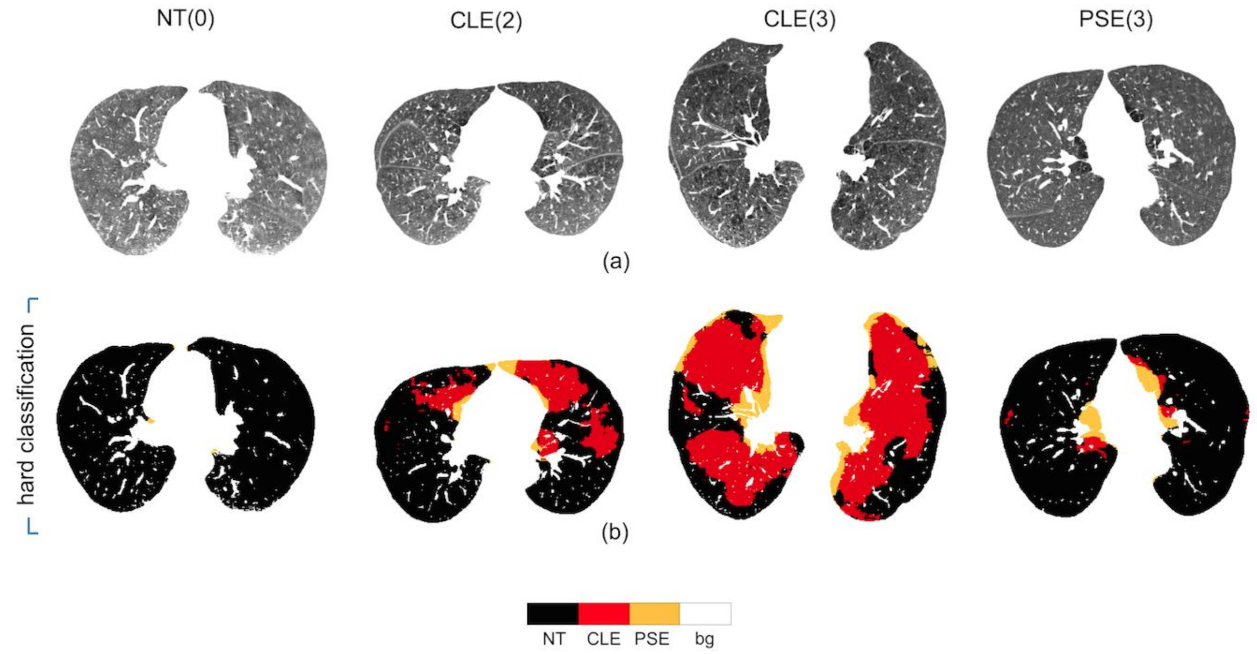


Fig. 2.4 Quantitative analysis of pulmonary emphysema using isotropic Gaussian Markov random fields for textural analysis. Original slices shown in (a). Hard classification results in (b). The patterns are: normal tissue (NT), centrilobular emphysema (CLE) and paraseptal emphysema (PSE). The clinical evaluation result is shown in the uppermost row including severity in parenthesis. Voxel classification color coding is shown in the bottom. Figure adapted from [18] and modified compared to original.

As described, CAD systems can provide beneficial added second opinion information to the radiologist, improving diagnosis and possibly improving patient outcome. In [14] it is argued that CAD systems will be implemented in hospital PACS systems in the near future. To perform analysis, quantification and classification, a well defined region for doing so is needed and therefore the following chapter will introduce medical image segmentation methods.

2.4 Segmentation methods

In medical image segmentation the task is to divide an image or a sequence of images into two or more disjoint and meaningful regions for a given purpose. The segmented regions can be different types of tissues or different anatomical structures and the images may be obtained by MRI, CT, ultrasound or other modalities. Traditionally, medical image segmentation is carried out by a radiologist or other medical experts, who manually delineate the structure of interest from an image. However, with the increased resolution of modern imaging modalities and increased amount of imaging data obtained from patients, the manual delineation becomes impractical, as the task is very time consuming and can also be subjective due to inter- and intraobserver variability, which has led to the need for objective, automated segmentation methods. The automated segmentation often serves as a crucial first-step for further computational analysis of the image data or for visualization purposes to assist a medical professional in diagnosis of diseases or during surgical interventions. Therefore, automated medical image segmentation methods are used in a broad range of applications, such as image guided surgery, surgical simulation, disease studies and therapy evaluation and have been applied to segment a whole range of anatomical structures, such as brain and brain structures, spine and vertebras, liver, kidney, colon, heart, lungs, prostate, etc. However, no universal method applies to all segmentation tasks and the method to use depends on the task at hand.

Medical image segmentation has been extensively studied for several decades and therefore many different segmentation methods have been developed [20–22]. Segmentation problems are often solved by combining two or more methods into a segmentation pipeline for an optimal solution. Different reviews and articles use different taxonomy and categorization of these segmentation methods, but we will adopt the categorization of [23]. In Figure 2.5 the segmentation methods are divided in two broad categories: primarily signal based and primarily model based. The boundaries, homogeneity and shape characterizes the object of interest to be segmented and influences the segmentation. The methods are often categorized according to their focus on these characteristics. The boundaries and homogeneity of an object are primarily image on signal based characteristics, whereas shape are often described by some kind of model, and therefore primarily model based. The signal based methods in Figure 2.5 are further divided into voxel based and region based methods. The voxel based methods includes thresholding and histogram based methods and only considers image information without incorporating prior knowledge about the object or interaction with neighboring voxels. The region based methods includes region growing, which primarily uses image signal information, but includes local information about neighboring voxels to form contiguous regions. The model based methods are subdivided into methods incorporating local priors and global priors. The local prior methods includes snakes and graph cuts, which incorporates prior shape knowledge about the structure of interest, but shape is only modeled locally by enforcing local smoothness and does not include a global prior knowledge about the overall shape of the structure of interest. Methods incorporating global priors includes active shape- and appearance models and atlas based methods, which incorporates knowledge about the global shape of the object of interest and is modeled from a number of reference shapes, similar to the structure of interest.

As seen in Figure 2.5, the number of parameters for the methods generally tends to increase from left to right, resulting in a high number of parameters for the model based methods. These parameters can be learned by use of supervised machine learning acting on annotated training cases. However, machine learning algorithms can be used in any of the segmentation methods and such are supportive to the automated segmentation by e.g. clustering of image features or performing statistical analysis.

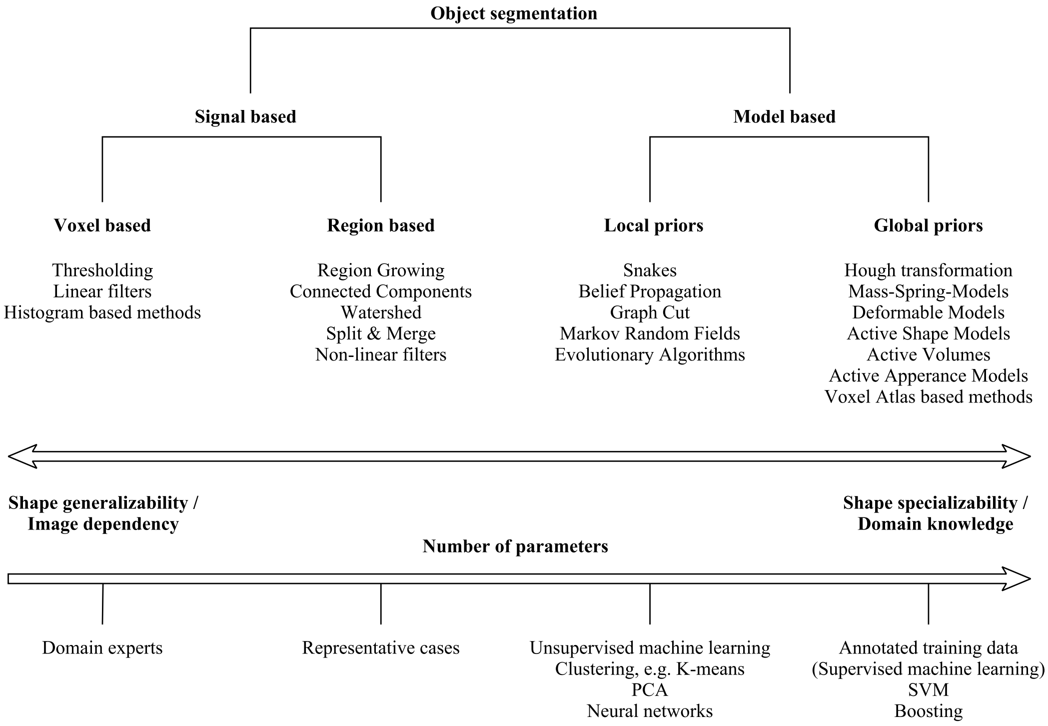


Fig. 2.5 Classification scheme of methods for segmenting objects in medical images adapted from [23].

]

2.5 Previous work on lung segmentation

Several studies have proposed methods for segmentation of the lungs from chest CT scans based on a variety of the segmentation methods described above, often incorporating more than one method. In [24] the lungs are segmented using an automated optimal threshold selection method and connected component analysis and divided into the right and left lung fields. The airways are segmented by a region growing process and removed from the lung segmentation. Finally, morphological opening and closing are used for smoothing of the segmentation. In this study only scans from healthy subjects were segmented.

In [5] an automatic lung segmentation method with error detection was developed. The segmentation pipeline consists of four major steps: firstly, the lungs are segmented by an approach very similar to the one described above. Afterwards, an automatic error detection algorithm is applied, where volume, ratio and shape consistency of the segmented lungs are examined and the errors are corrected. Then, if the lung segmentation are deemed to contain further errors, a multi-atlas lung segmentation are performed, by registering manually segmented atlases to the image and propagating the labels of the atlases to the image. As a final step, this segmentation is automatically checked for errors and if these exists, it is marked as needing manual intervention. In this study, it is argued, that the multi-atlas segmentation will likely produce better results in case of abnormal pathology in the lungs. Both healthy and diseased lungs were segmented.

In [7] a multi-shape graph cut with neighbor prior constraints are proposed for segmentation of healthy and diseased lungs from CT. The method is based on graph cut segmentation with an added energy term, that takes into account the multiple possible shapes of the target image. The shapes are formulated using a level set distribution model and the algorithm automatically chooses the optimal shape prior for the target image among multiple shape priors. Before the multi-shape graph cut segmentation, a rough body cavity and rough lung segmentation as well as trachea and main bronchi segmentation are performed using thresholding, region growing and morphological operations. The algorithm is tested on both healthy and diseased lungs and is able to segment lungs containing tumors, pleural effusion and emphysema because of the incorporated shape term in the graph cut energy function.

The aforementioned studies have tried to segment the lungs from chest CT scans using a combination of segmentation methods. In [24], a pure signal based approach was used, while [5] and [7] used a combination of both signal based methods and model based methods, combining thresholding and region growing as initial segmentation with further refinement using either multi-atlas segmentation or graph cut segmentation with statistical shape modeling. The use of local or global shape prior methods for lung segmentation has been applied in order to segment diseased lungs with abnormalities. This complicates the segmentation, as the attenuation in the lung tissue, and thereby the contrast in the image, is altered by the abnormal pathology. Also, the shape can be altered due to pathology and irregularities, compared to the healthy lung, such as tumors, nodules and pleural wall thickening, can be present due to abnormal pathology in the diseased lung.

2.6 Outline of the thesis

With the aim of providing an accurate segmentation for healthy lungs as well as emphysematous lungs in patients with COPD, this study will investigate a method for segmentation of the lungs from chest CT scans using multi-atlas registration and graph cuts. The method is similar to that of [25] and [26] used for segmentation of head and neck tissue from CT and segmentation of the hippocampus from MRI. The method will be compared to a multi-atlas segmentation without using graph cuts. Graph cut segmentation has the advantage of combining prior spatial knowledge from multi-atlas registration with intensity modeling and voxel neighborhood interaction modeling into a graph cut segmentation framework. The approach and methodology will be described in the following Chapter 3.

Two dataset are used in the thesis. The first dataset, named the EMPIRE10 dataset, contains individual atlases; a CT scan of a patient and a related binary mask of the lungs. The individual atlases will be used to do the two segmentation methods; the multi-atlas segmentation and the multi-atlas registration and graph cut segmentation. Both methods will be tested and evaluated on this dataset in a leave-one-out cross-validation experiment. The second dataset, named the Hospital of Vendsyssel dataset, contains HRCT scans of both healthy subjects and subjects with COPD. This dataset contains manual delineations of the lungs on select subjects. Both methods will be tested and evaluated on this dataset, quantitatively as well as qualitatively. The datasets are described in Chapter 4. Method and metrics for validation and evaluation of the two methods will be described in Section 3.6.

The results will be presented in Chapter 5 and a discussion of the method and results as well as a general discussion on performance, improvements and future work are made in Chapter 6.

Chapter 3

Methods

As described in Section 2.5 several previous studies have dealt with the segmentation of lungs from chest CT scans. One of the biggest challenges in segmenting lungs from chest CT scans is segmenting the lungs with abnormal pathologies in the case of patients with various pulmonary diseases. It becomes increasingly difficult when the anatomy, shape and signal varies from the healthy subjects and becomes something 'out of the ordinary'. It is therefore necessary to develop a method that takes these cases into account, if that is the intended use of the method, and to, as well as possible, account for the unknown by making the method robust and know its limitations. Therefore, the method should accordingly be validated and tested on representative data for its intended use and its limitations should be investigated. Another challenge in lung segmentation is the complexity of the segmentation method, which can result in high resource use and computational time.

3.1 Method overview

The method presented in this thesis will focus on automatic segmentation of healthy and emphysematous lungs in patients with COPD. The method should be applicable to these cases and will therefore be validated and tested on these data accordingly. Less emphasis will be put on the resource requirements and computational time. Possible future additions or improvements to the method could be extension to other pulmonary diseases and reduction of resource requirements and computational time. The goal of the method is to produce an accurate and robust automatic segmentation of the lungs from chest CT scans for further computational analysis and quantitative studies in healthy subjects and subjects with COPD.

The method will combine several well-known signal and model based methods. Multi-atlas registration in combination with other methods has shown good results in other studies where it has been applied for brain structure segmentation in [27] and for lung segmentation in [5]. Recently, graph cut segmentation has also shown good results in coronary lumen segmentation [28] and in segmentation of head and neck tissue in CT [25].

The approach in this thesis will follow the method of [25] and [26] used for head and neck tissue segmentation and hippocampus segmentation, respectively. These methods combine multi-atlas registration with graph cuts. This method will be applied to lung segmentation in healthy subjects and subjects with COPD. The presented method will add segmentation and elimination of the airways. An overview of the method is shown in Figure 3.6. The multi-atlas registration will provide a good spatial prior knowledge for the lung segmentation and the graph cut segmentation will refine and smooth the segmentation by classifying voxels according to intensity and neighborhood. The goal is to provide an accurate and robust segmentation method for healthy subject and subjects with COPD. The airway segmentation is necessary to eliminate these from the lung segmentation.

As the multi-atlas registration serves as a spatial prior knowledge model in the graph cut segmentation framework along with intensity modeling and neighborhood modeling, the graph cut segmentation method will be compared to a multi-atlas only segmentation. Using only the spatial prior model in the graph cut segmentation framework built from multi-atlas registration of individual atlases are equivalent to a multi-atlas segmentation with label fusion and majority voting and it is therefore relevant to investigate the performance of the multi-atlas registration and graph cut method compared to a multi-atlas segmentation only method. The procedure will be described in detail in the following sections.

The method will be validated on CT scans from both healthy subject and subject with COPD. The data will also be used as individual atlases for the multi-atlas registration. The data available are described in Chapter 4. The method will be validated with respect to various quantitative measures on volumetric overlap and accuracy described in Section 3.6.

In the following sections, the method for lung segmentation of healthy and emphysematous lungs will be explained in details and an introduction to the concepts used, such as registration, atlas segmentation and graph cut segmentation will be given.

3.2 Multi-atlas registration

The multi-atlas registration serves as the spatial prior model in the graph cut segmentation framework along with intensity modeling and neighborhood modeling. The multi-atlas registration can serve as a multi-atlas only segmentation by thresholding the probability map for a given target image built from individual atlases (majority voting). The multi-atlas registration and building of a target image specific probability map from individual atlases will be explained in the following and the incorporation of this spatial prior knowledge into the graph cut segmentation framework will follow.

3.2.1 Building a probability map

The registration of individual atlases to the target image to be segmented is done in order to create the target image specific probability map. The available individual atlases are all registered to the target image and the labels are averaged. The atlases available are described in Chapter 4. Each individual atlas consists of a CT volume (the intensity image) and a corresponding binary mask (the labeled image). The CT volume is the image resulting from a chest CT scan of a subject. The binary mask is the corresponding segmentation of the lungs from the CT volume of the subject. The registration process registers the CT volume to the target image to obtain a transformation that maps the atlas image to the target image. The resulting transformation is then applied to the binary mask of the atlas image resulting in a mapping of labels of the mask of the atlas to the target image. This is done for all individual atlases and a resulting probability map for the specific target image is obtained by averaging all transformed labels of the atlas to the target image. An overview of the procedure is shown in Figure 3.1.

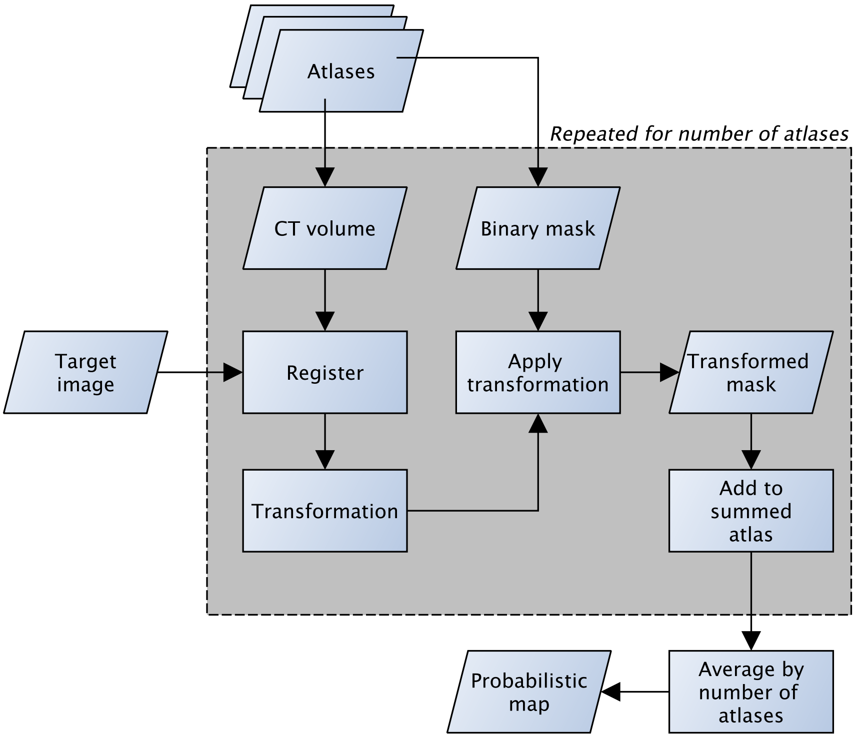


Fig. 3.1 Overview of the process of building a probability map.

The following section explains the process of aligning the individual atlases with the target image to be segmented in order to obtain the probability map: the registration process.

3.2.2 Registration

Registration of an image pair, or a set of images sequences, is the process of bringing these images into alignment, such as aligning atlases to a target image for segmentation or as in the case of the graph cut segmentation framework, to create a probability map for the spatial prior model [29]. To register the images, a geometric transformation which aligns the images must be estimated. To obtain this transformation, a cost function is constructed, which measures how much aligned (or misaligned) the image is as a function of the transformation. The goal is to minimize this cost function in order to obtain an optimal transformation. To measure how well the images are aligned, or how similar they are, a similarity measure is used. Most often, the similarity measure is based on image intensities and these methods are called voxel-based similarity measures.

Any point, or voxel, in the atlas image needs to be mapped to a point in the target image. Therefore, we need to find the optimal transformation $\mathbf{T}$ that maps any point $\mathbf{p}$ in the atlas to a point $\mathbf{p}'$ in the target:

$$\mathbf{T} : \mathbf{p} \mapsto \mathbf{p}' \tag{3.1}$$

To account for both a global transformation and local transformations or local deformations, Eq. 3.1 can be split in two transformations: a global and a local. The global transformation accounts for global scaling, rotation and translation, while the local transformation deals with local deformations between the images. The transformation then becomes:

$$\mathbf{T}(\mathbf{p}) = \mathbf{T}_{global}(\mathbf{p}) + \mathbf{T}_{local}(\mathbf{p}) \tag{3.2}$$

The local transformation are most commonly carried out by free-form deformations using cubic B-splines. A grid of equally spaced control points are laid out on the image and the local transformations are carried out on this grid. Voxel transformation between the control points is interpolated by computing the interpolation based on the positions of surrounding control points [30]. In Figure 3.2 an example of local deformations on a grid using B-splines are shown. When using B-splines for image registration, the registration can be done with a grid spacing varied in a coarse-to-fine manner in order to improve registration [29].

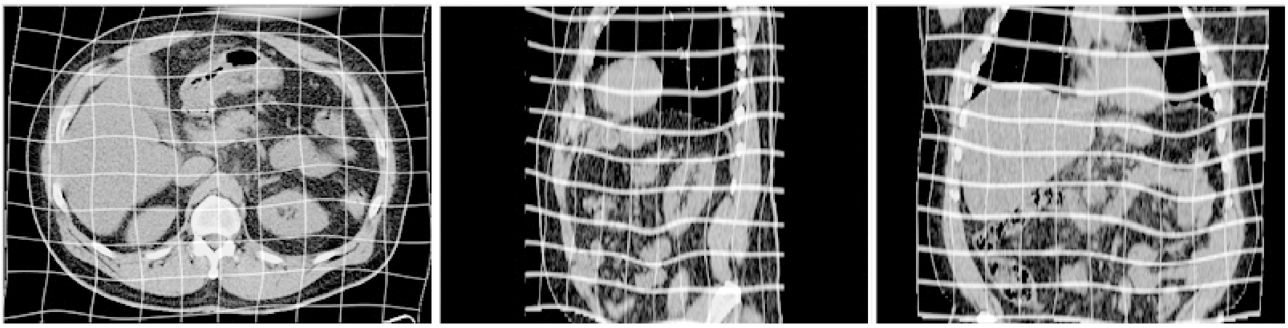


Fig. 3.2 Local deformation of liver CT dataset using B-splines. The gridlines shows the extent of deformation. Image from [31].

3.2.3 Similarity measure

A similarity measure is used in order to describe or quantify how well two images are aligned [29]. A voxel-based similarity measure is based on voxel intensities and the similarity measure defines the cost function, which describes the goodness of alignment. The simplest measure of similarity is the sum of squared differences (SSD), where each of the images corresponding voxels intensity differences are summed and squared and averaged by the number of voxels. A perfect alignment between two images would result in a SSD of zero. The assumption of SSD as a similarity measure is that the images are from the same modality and that the intensities of the images are the same.

Another similarity measure is the normalized cross correlation (CC). This measure can be used if there exists a linear relationship between the intensities of the two images. However, it is still restricted to images from the same modality. The normalized cross correlation similarity measure is expressed as:

$$\mathcal{S}_{CC} = \frac{\sum \left(I_A(\mathbf{p}) - \mu_A\right)\left(I_B(\mathbf{T}(\mathbf{p})) - \mu_B\right)}{\sqrt{\left(\sum I_A(\mathbf{p}) - \mu_A\right)^2 \left(\sum I_B(\mathbf{T}(\mathbf{p})) - \mu_B\right)^2}} \tag{3.3}$$

where I_A and I_B are the two images and μ_A and μ_B their corresponding average intensities. The normalized cross correlation is a sufficient similarity measure for registration of the images in our method as they are coming from the same modality (CT) and the intensities have not been contrast enhanced (non-linearly intensity differences). However, in cases with multi-modal images similarity measures such as normalized mutual information can be used.

3.2.4 Optimization

The registration can be formulated as an optimization problem, where the cost function ($\mathcal{C}$) to be optimized are based on the aforementioned similarity measure ($\mathcal{S}$) and a penalty function ($\mathcal{P}$) [29]:

$$\mathcal{C} = -\mathcal{S} + \lambda \mathcal{P} \tag{3.4}$$

As already mentioned, the cost of the similarity measure are based on the degree of alignment of the images as expressed by the intensity information in the images. The penalty function penalizes certain transformations and imposes constrains on the transformation. These constrains can include preservation of volume, amount of rigidity and penalty to folding. These constraints can be imposed to ensure plausible physiological deformations of tissue (images) during the registration. The λ parameter is a weighting parameter. Most often the cost function is minimized using gradient descent or variations hereof [29]. In Figure 3.3 the outline of a registration process is shown. The similarity measure is computed and optimized and the transformation is updated and applied to the source image (or the atlas image). This procedure is repeated until convergence and the final transformation is obtained.

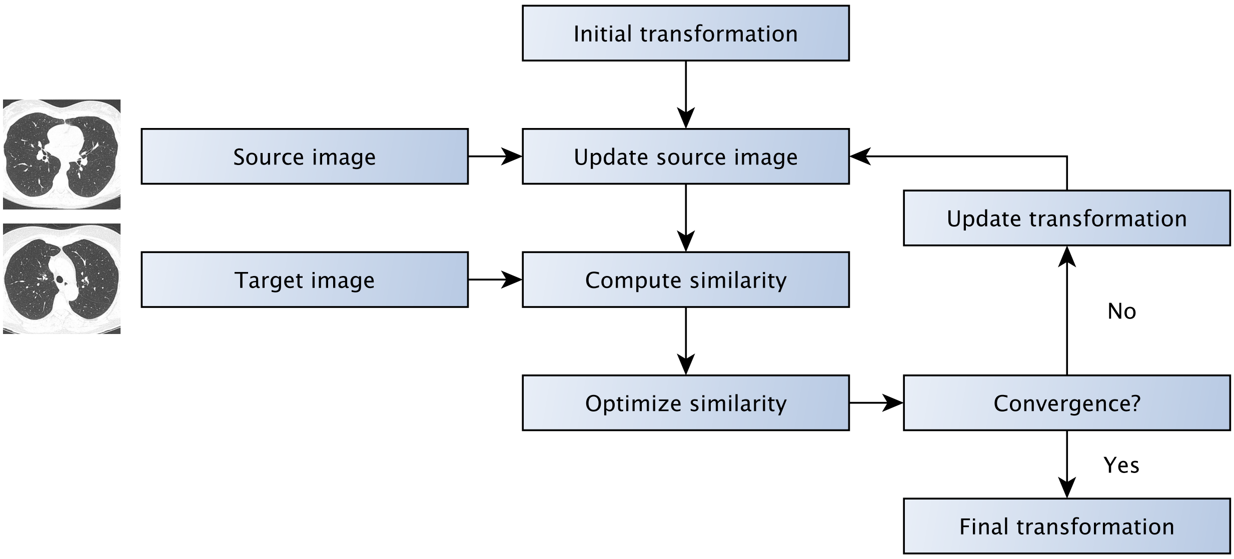


Fig. 3.3 Overview of the registration between two images. The optimization of the cost function and the updating of transformation and source image is repeated until convergence is reached based on the similarity measure between the images. Adapted from [29].

3.2.5 Label fusion

When transformation **T** for the intensity image of each individual atlas has been estimated, the transformation is applied to each of the labeled images. This results in N number of decisions or classifications of the lungs in the target image, one for each of the individual atlases. The final multi-atlas segmentation is obtained by decision fusion analogous to combining pattern classifiers. This is illustrated in the flowchart in Figure 3.4 where the combination of the classifiers produces the segmentation in a multi-atlas segmentation method:

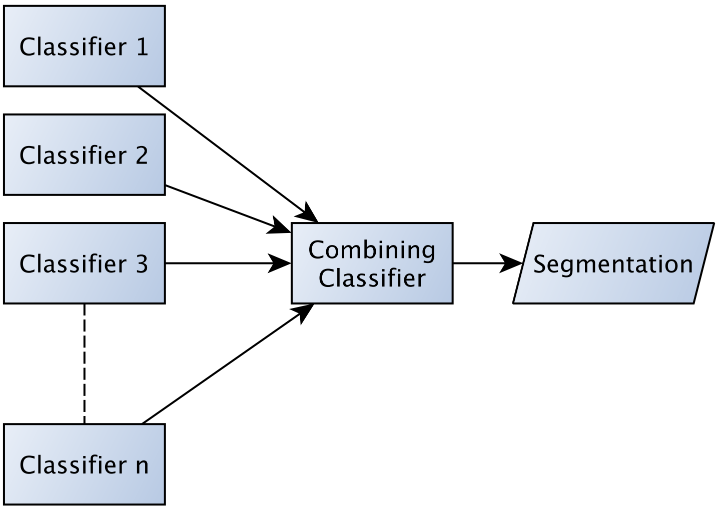


Fig. 3.4 Illustration of combining classifiers. The N number of classifiers are combined using a combining classifier and the resulting classification are the segmentation of the target image.

Various methods for combining classifiers exists. Common methods for multi-atlas registration and segmentation are majority voting and weighted voting. Majority voting requires, as the name implies, that the majority of classifiers agree on a classification. This essentially labels every voxel in the target image as lung tissue or non-lung tissue according to the majority of the registered label of a given voxel. This is therefore equivalent to thresholding the obtained target image specific probability map at > 0.5. Weighted voting assigns weights to each transformed atlas labels based on their similarity with the target image. This is in contrast with majority voting, where each atlas is assigned the same weight. Weighted voting methods such as the SIMPLE method [32] and the STAPLE method [33] has been shown to improve performance over majority voting. However, the majority voting method will be used for multi-atlas segmentation, but improved performance from weighted voting could be investigated in future work.

In Figure 3.5 the steps in creating a multi-atlas segmentation are shown. The multi-atlas segmentation will be compared to the multi-atlas registration and graph cut method, where the target specific probability map created from the individual atlases are used as spatial prior model in combination with intensity modeling and neighborhood modeling. The following section will describe the graph cut segmentation framework for the method.

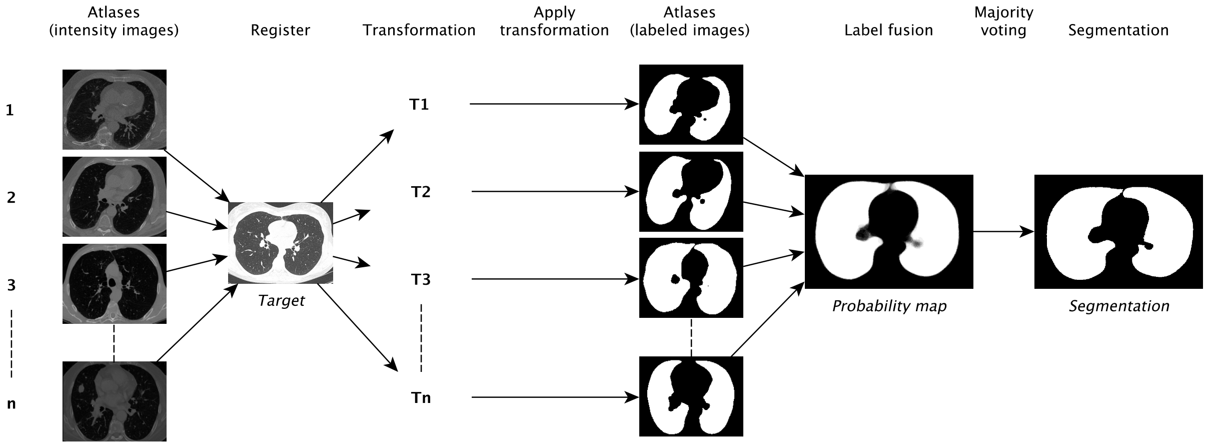


Fig. 3.5 Illustration of combining classifiers. The N number of classifiers are combined using a combining classifier and the resulting classification are the segmentation of the target image.

3.3 Graph Cut Segmentation Framework

The overall method is based on [25] and [26] and can be formulated as an a posteriori maximization problem (MAP), where the probability to maximize is the labeling of all voxels in the target image resulting in a segmentation of the target image into voxels labelled as foreground voxels (the lungs) and voxels labelled as background voxels (non-lung voxels). This involves formulating an energy function (derived from the a posteriori maximization problem) and constructing a graph based on the energy function and find the minimum cut of the graph, which is equivalent to maximizing the a posteriori probability. The energy function includes:

- a spatial prior energy term, based on a probability map of the target image built from individual atlases using multi-atlas registration (association potential)
- an intensity energy term, based on voxel classification and is the likelihood of a voxel belonging to some part of the image (association potential)
- a regularity energy term, based on penalization of intensity differences between neighboring voxels to obtain smooth segmentations (interaction potential)

The energy function and terms will be described in detail in the following section. In addition to the above mentioned parts of the method for segmentation of the lungs, the airways on the target image are segmented using a traditional region growing method and extracted from the lung segmentation. An overview of the complete segmentation method are shown in Figure 3.6.

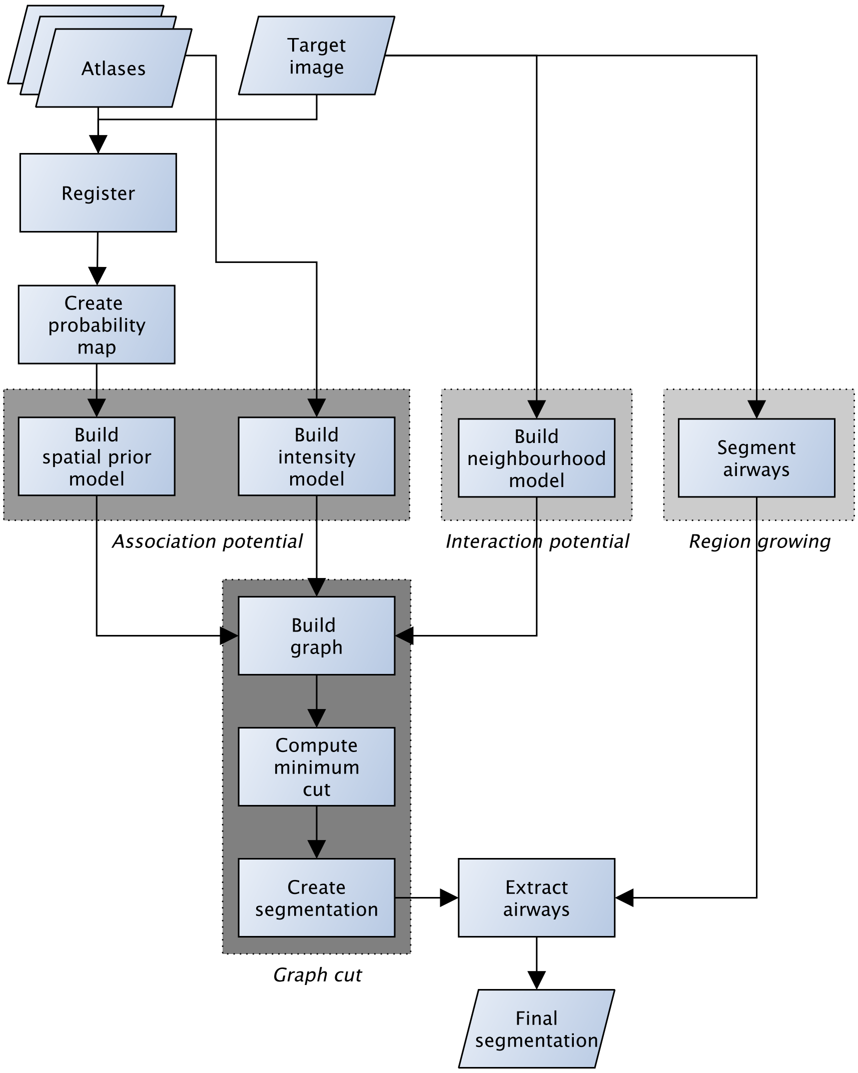


Fig. 3.6 Overview of the method.

3.3.1 Maximum-a-Posteriori (MAP) Estimation

As already mentioned, the segmentation of the lungs from chest CT scans, can be formulated as an a posteriori maximization problem (MAP). The lungs are to be assigned the label 1 and the remaining voxels (the background) are to be assigned the label 0, resulting in a binary segmentation. If we name the vector containing all labels for each voxel in the target image **f** and the image information **i**, the MAP becomes:

$$\hat{\mathbf{f}} = \underset{\mathbf{f}}{\operatorname{argmax}}\, p\left(\mathbf{f} \mid \mathbf{i}\right) = \frac{p\left(\mathbf{i} \mid \mathbf{f}\right) p\left(\mathbf{f}\right)}{p\left(\mathbf{i}\right)} \tag{3.5}$$

where $\hat{\mathbf{f}}$ is the vector of maximized label assignment probability to each voxel in the image given the information in the target image. In a binary segmentation, **f** is a vector containing labels $f_m \in 0, 1$ for each voxel m in the set $\mathcal{M}$ of voxels in the image.

The vector $\mathbf{i} = \{i_m : m = 1, ..., \mathcal{M}\}$ contains all intensity values i_m at all voxel locations $\mathcal{M}$. If we consider a binary segmentation of the lungs from chest CT images, a voxel in the lungs is likely to be surrounded other lung voxels in its immediate neighborhood. Therefore, there is a higher probability that neighboring voxels to a lung voxel are also voxels belonging to the lungs. However, based on a location of some lung voxel, we cannot infer much about the probability of voxels far away from that lung voxel, just from its location. We can therefore assume that the label f_m conditioned on the image intensity $\mathbf{i}$ depends only on the labels in its immediate neighborhood $n \in \mathcal{N}_m$. Eq. 3.5 can then be approximated as a discriminative random field according to [25] and [34]. The discriminative random field is a graphical model, where a graph denotes dependence between random variables. The random variables are the vertices of the graph and connection between vertices are called edges. For image segmentation, the voxels in the image are the vertices of the graph and the connection between these are the edges. The basics of a graph theoretic approach to image segmentation can be found in [34] and [35].

A clique potential is a subgraph and the number refers to the maximum connection between vertices in a graph. As the labeling only depends on voxels (vertices) in their immediate neighborhood, the discriminative random field contains one- and two-voxel clique potentials and the joint posterior probability $p(\mathbf{f} \mid \mathbf{i})$ of Eq. 3.5 can be approximated as:

$$p(\mathbf{f} \mid \mathbf{i}) \approx \frac{1}{Z}\exp\left[\sum_{n \in \mathcal{M}}\left(\lambda_1 A(f_m, \mathbf{i}) + \sum_{n \in \mathcal{N}_m} I(f_m, f_n, \mathbf{i})\right)\right] \tag{3.6}$$

The term $A(f_m, \mathbf{i})$ is called the association potential and the term $I(f_m, f_n, \mathbf{i})$ is called the interaction potential. The parameter λ_1 weights the association potential with respect to the interaction potential. Z is a normalization constant that can be disregarded, as it does not influence the MAP estimation.

As can be seen from Eq. 3.6 the probability of assigning labels to every single voxel in the image given the image information, $p(\mathbf{f} \mid \mathbf{i})$, is given by a summation over all voxels $\mathcal{M}$ in the image. The summation is weighted based on their probability based on intensity and spatial prior knowledge, the association potential, and their relationship to neighboring voxels, the interaction potential. The association and interaction potential will be explained in greater detail in the following sections.

3.3.2 Association potential

The association potential is based on two models: an intensity model and a spatial prior model. These models will be described in the following.

3.3.2.1 Spatial prior model

The spatial prior model is created using multi-atlas registration as described in Section 3.2. The individual atlases are registered to the target image and a probability map are created by averaging the number of atlases. This creates a probability map of the target image with probabilities for every voxel ranging from 0 to 1. Accordingly, the probability map is created as:

$$p_s(f_m) = \left(\frac{1}{J}\sum_{t_j \in \mathcal{T}} g_m^{t_j}\right)^{\lambda_2} \tag{3.7}$$

t_j is the jth atlas image from the set of J atlas images $\mathcal{T}$. $g_m^{t_j}$ is the label (0) or (1) at voxel location m of the jth atlas image registered to the target image. λ_2 is weight parameter that weights the influence between the intensity model and the spatial prior model. If the probability map of the target image alone was to be used for segmentation, thresholding it at > 0.5 would result in a multi-atlas majority voted segmentation of the target image as described in Section 3.2. The performance of this majority voted multi-atlas segmentation will be compared to that of the multi-atlas registration and graph cut method.

3.3.2.2 Intensity model

The intensity model describes the probability of a voxel belonging to either foreground or background depending on the intensity of the voxel. This probability is related to the histogram of the foreground and background regions. If there is good contrast between the region of interest, the foreground, and the background, the regions will be easier to distinguish from each other, as the foreground is well separated from the background in the histogram. This is often the case in healthy lungs, as the air-filled lungs have a low attenuation value in CT scans and the immediate surrounding tissue and bones have a high attenuation value in CT scans. Therefore, previous work [24] on segmentation of healthy lungs from chest CT scans has had success in lung segmentation based on thresholding because of the high contrast between the pulmonary region and surrounding tissue and bones as can be seen in Figure 3.7.

To estimate the probability distribution of the intensities, $p_{int}\left(i_m \mid f_m\right)$, the atlases are used. By sampling the intensities within the lung mask of the atlases and the intensities outside the lung mask, the probability distribution of the intensities can be estimated. To estimate the probability distribution, a Parzen window estimator with a Gaussian kernel is used:

$$p_{int}\left(i_m \mid f_m\right) = \frac{1}{K}\sum_{k\in K}\frac{1}{h\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{i_m - I_k}{h}\right)^2\right] \tag{3.8}$$

As the kernel of the probability density function is a Gaussian kernel, $p_{int}\left(i_m \mid f_m\right)$ becomes an average of normal densities centered at the samples [36], where h is the standard deviation of the Gaussian kernel. In [25] and [26] h was empirically determined and set to 50 and 30, respectively, and h will be set to 50 based on these studies. K is the number of samples and I_k the kth samples intensity.

The intensity probability distribution of Eq. 3.8 is calculated for both foreground and background. This gives essentially two histogram estimates and every voxels intensity is then checked wether it has the highest probability of occurring in the foreground or the background distribution. The intensity model uses this information to correct for possible inaccurate registrations from the multi-atlas registration, where some part of the lungs are not included in the spatial prior or in case of mis-registration resulting in some of the surrounding tissue voxels are included in the spatial prior.

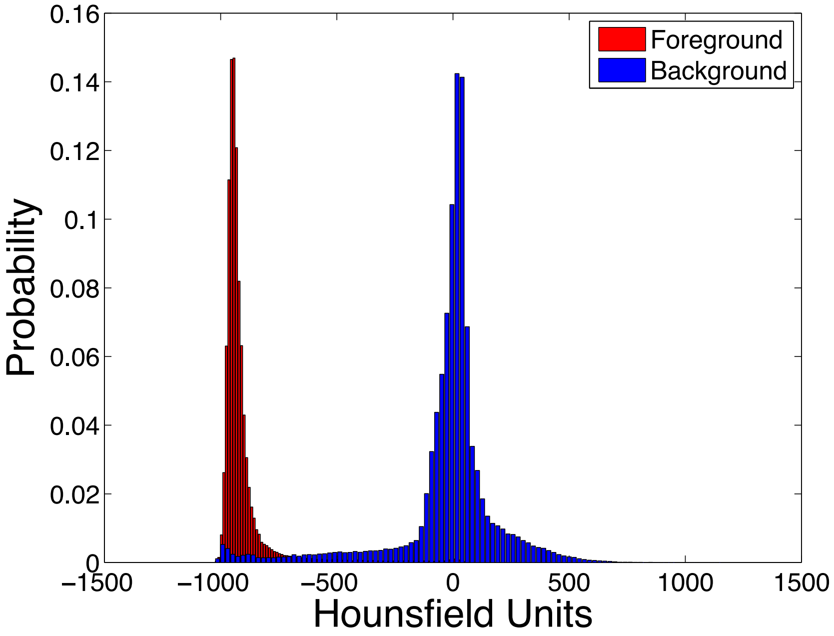


Fig. 3.7 Normalized histogram of the lungs (foreground) and surrounding structures (background).

3.3.3 Interaction potential

The interaction potential promotes a smooth segmentation at the borders of the segmentation [25]. The lungs are a relatively smooth structure and therefore the interaction potential promotes a smooth segmentation by penalizing rapid changes in intensities at the border of the segmentation. Thus, the interaction potential can be seen as a regularizing term and is expressed as:

$$I\left(f_m, f_n \mid \mathbf{i}\right) = \begin{cases} 0, & \text{if } f_m = f_n \\ \frac{1}{2}\exp\left(-\frac{(i_m - i_n)^2}{2\xi^2}\right) \cdot \frac{1}{\Delta(m,n)} & \text{if } f_m \neq f_n \end{cases}, \tag{3.9}$$

where $\Delta\left(m, n\right)$ is the Euclidean distance between voxels m and n. The regularization term penalizes different labels for neighboring voxels with an intensity difference of ξ. The neighborhood is a 3D 6-voxel neighborhood and ξ is set to 800, which is the average difference between foreground and background as can be seen in the histogram of Figure 3.7.

3.3.4 Graph Cut optimization

As mentioned earlier, the maximization of the maximum a posteriori probability (MAP) of Eq. 3.5 to find the optimal segmentation, is equivalent to minimizing an energy function derived from the MAP equation. This minimization can be done by using graph cut optimization, where a graph is constructed and the minimum cut of the graph is computed. In [26] a thorough description of the conversion from MAP estimation to energy function minimization is made, which will not be explained here. The following will explain the construction of a graph and the optimization by computing the minimum s-t cut on the graph resulting in the segmentation.

The graph $G =< V, E >$ consists of a set of nodes, or vertices, V with one for each voxel in the image as well as two special terminal nodes, which are called the source, s, and the sink, t [37]. Connecting the nodes are a set of edges E. If the edges connect to the terminal nodes they are called t-links and if they connect between two normal nodes they are called n-links. Each edge is assigned a non-negative weight, or cost, $w = (p,q)$. A $s-t$ cut C on the graph G divides the nodes into two disjoint subsets $C = \{S,T\}$. The cost of a cut is the sum of the edge weight and therefore the minimum cut problem is to find the minimum cost of all possible cuts. For segmentation, terminal t is considered the foreground label and terminal s is the background label and therefore separating the nodes results in a binary segmentation. In Figure 3.8, an image and a graph constructed from the image is shown together with the cut on the graph resulting in a segmentation.

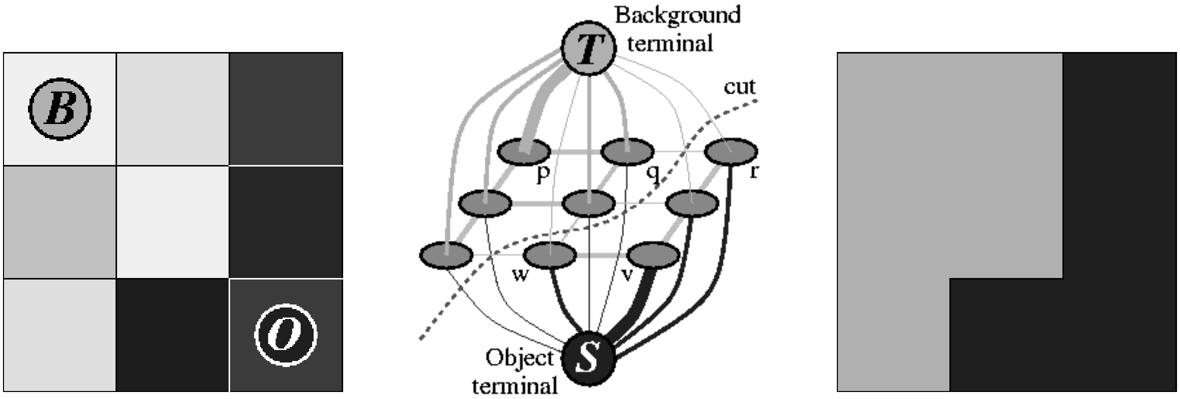


Fig. 3.8 A black and white image (left) with background pixels (B) and foreground pixels (O) are the basis of the constructed graph (middle) where a cut on the graph results in the segmentation (right). The cut is determined from the weights of the edges connected to the terminals and each pixel. Adapted from [37].

The weight of the edges of the t-links connecting all vertices, or voxels, to either foreground, t, or background, s, are determined by the association potential [26]. This is the probability of belonging to either foreground or background depending on the intensity model and spatial prior model as explained in earlier sections. The weight of the edges of the n-links, or the interaction between vertices, are determined by the interaction potential. This weight is based on the 3D 6-voxel neighborhood model as described in earlier.

If the edges are interpreted as being water pipes with capacities equal to their weights, the graph can be cut by finding the maximum flow of water from s to t [37]. This saturates a set of edges, which divides the set into two disjoint sets, and is equivalent to finding the minimum cut. Therefore, these cuts are called max-flow/min-cut. Several methods exists to find the minimum cut of a graph, but here the max-flow/min-cut method of [38], applied in [26] and [25], will be used.

3.4 Segmentation of large airways

As the airways are anatomical structures with in the lungs and very close to the lungs with very similar intensities as the lungs they need to be segmented in order to extract the airways, which might get segmented by the graph cut segmentation. The airway segmentation is not incorporated into the graph cut segmentation framework, but are extracted afterwards. The airways to extract are the trachea and primary bronchi. To segment the airways, an explosion controlled region growing algorithm with automatic seed selection similar to that of [5] is used. The seed point is automatically found in the top axial slices of a scan by searching for connected components with an average HU value below -950, having a minimum size of 50 mm^2, having a maximum size of 1225 mm^2 and mean x- and y-coordinates below a maximum distance of 30% of the x- and y-dimensions away from the image centre. The first 25 slices from the top are inspected slice-by-slice and the closest connected component to the image center is kept. In this connected component, the voxel with the lowest HU value are used as seed point for the explosion controlled region growing [39]. The region growing is initiated with an initial threshold. Then the threshold is increased slightly and the region growing is repeated. This is repeated until the size of the next grown airway volume has increased by at least a factor of two, which indicates that the region growing has been growing into the lung tissue. The increase in threshold is then lowered and the procedure is again repeated, but from the previous threshold value. This is repeated until an increase in threshold of 1 HU causes an explosion, and the region growing is stopped. This produces the final airway segmentation, which can be extracted from a possible over-segmentation from the graph cut segmentation of the airways.

3.5 Implementation

This section explains the implementation details of the method, including software and parameter values. The multi-atlas registration was implemented on a 8x2.6GHz Dual Core AMD Opteron server with 32GB RAM running the Ubuntu 8.04 LTS Linux distribution. The graph cut segmentation was implemented on a 4x3.50GHz Intel Core i7-4770K desktop computer with 16GB RAM running Microsoft Windows 8 Enterprise. Airway segmentation and all other MATLAB computation was implemented on a 2x2.66GHz Intel Core 2 Duo MacBook Pro laptop computer with 8GB RAM running Mac OS X 10.9.3.

3.5.1 Multi-atlas registration implementation

To create the target image specific probability map for the spatial prior model the individual atlases are registered to the target image using the Elastix software [40, 41]. Elastix is a toolbox for rigid and nonrigid registration of images and supports different similarity measures, optimization and interpolation methods, and multi-resolution schemes and has been used extensively in a range of research papers on medical image registration [42]. For registration of the atlases to build the probability map, an approach similar to that of [43] will be used. In this study, chest CT scans were registered to estimate the progression of pulmonary emphysema and are therefore suitable for our method. The full list of parameters for the Elastix toolbox registration used in the study of [43] and applied in for this method are publicly available at the Elastix website [44].

Two registration steps are performed. First an affine transformation are performed and afterwards a non-linear B-spline transformation. The registration is done in a multi-resolution framework using an image pyramid of 5 resolutions. Normalized cross correlation similarity measure is used in combination with a gradient descent optimizer. Following the registration of the atlases to the target image, the binary labels of the transformed atlases were averaged to obtain a probability map of the target image using MATLAB [45].

3.5.2 Graph cut segmentation implementation

The graph cut segmentation framework with the max-flow/min-cut method used in [25] and [26] has been made available for this project. The method has been implemented as a plugin for the medical imaging processing and visualization platform MeVisLab [46].

The parameter values for the graph cut segmentation was set as:

- $\lambda_1 = 0.5$, which weights the association potential with respect to the interaction potential.
- $\lambda_2 = 0.5$, which weights the influence between the intensity model and the spatial prior model of the association potential.

The parameters was based on that of [25] and [26].

3.5.3 Segmentation of airways implementation

The segmentation of the airways was done using the MATLAB software [45].

3.6 Validation

To evaluate the performance of the multi-atlas segmentation method and the multi-atlas registration and graph cut method a leave-one-out cross-validation approach was used. All but one of the individual atlases available are used to perform the registration and segmentation, while the one atlas left out are used to evaluate the resulting segmentation. The leave-one-out experiment are repeated, leaving out a new atlas in each iteration until all atlases have been used for evaluation. The labeled image of the atlas left out are used as the ground truth segmentation and the metric for segmentation performance evaluation is the Dice similarity coefficient (DSC). The DSC measures the overlap of two segmentations. The metric is normalized with respect to the segmentation volumes resulting in a measure ranging from 0 (no overlap) to 1 (total overlap). The DSC is defined as:

$$DSC = \frac{2\,|A \cap B|}{|A| + |B|} \tag{3.10}$$

The performance evaluation results in a DSC measure for each of the leave-one-out experiments for both methods. The performance of the methods are compared using the Wilcoxon signed rank test to test for statistical significant differences in DSC.

In addition to the leave-one-out cross-validation experiment, the methods will be applied to a new dataset not containing any atlases. This dataset contains manual delineations of the lungs for some of the scans. The performance of the methods will be evaluated with respect to the DSC where possible and qualitative assessment of the results will be presented.

3.7 Method summary

Two methods will be applied for segmenting the lungs from chest CT scans: a multi-atlas-based segmentation using majority voting and a multi-atlas registration and graph cut method. The multi-atlas segmentation is made by registering multiple individual and independent atlases to a target image and the segmentation is obtained by majority voting. The obtained probability map of the target image of this multi-atlas segmentation method will be used in a multi-atlas registration and graph cut method as a spatial prior model. The graph cut method will incorporate a spatial prior model and an intensity model (association potential) and a neighborhood model (interaction potential) into a graph cut segmentation framework. A graph will be constructed with edge weights according to the aforementioned models and a minimum cut on the graph will be calculated resulting in the segmentation. This is equivalent to calculating the maximum a posterior probability (MAP).

The performance of the two methods will be evaluated in terms of the Dice Similarity Coefficient (DSC). The methods will be evaluated in a leave-one-out cross-validation approach and on a new set of data with manual ground truth delineation of the lungs.

Chapter 4

Data

This chapter will describe the available data for multi-atlas registration and segmentation and for validation and performance evaluation. Two datasets are available; the EMPIRE10 dataset and the Hospital of Vendsyssel dataset. The EMPIRE10 dataset will be used for multi-atlas registration and segmentation and for the leave-one-out cross-validation experiment, as this dataset contains scans and binary masks from individual subjects, which serves as individual atlases. The Hospital of Vendsyssel dataset consists of HRCT scans of healthy subjects and patient with COPD. Manual delineation of some slices in the scan of one healthy subject and one patient with COPD are available, which will be used for performance evaluation. The rest of the scans does not contain delineations of the lungs and will therefore be used for qualitatively, visual inspection.

4.1 The EMPIRE10 data

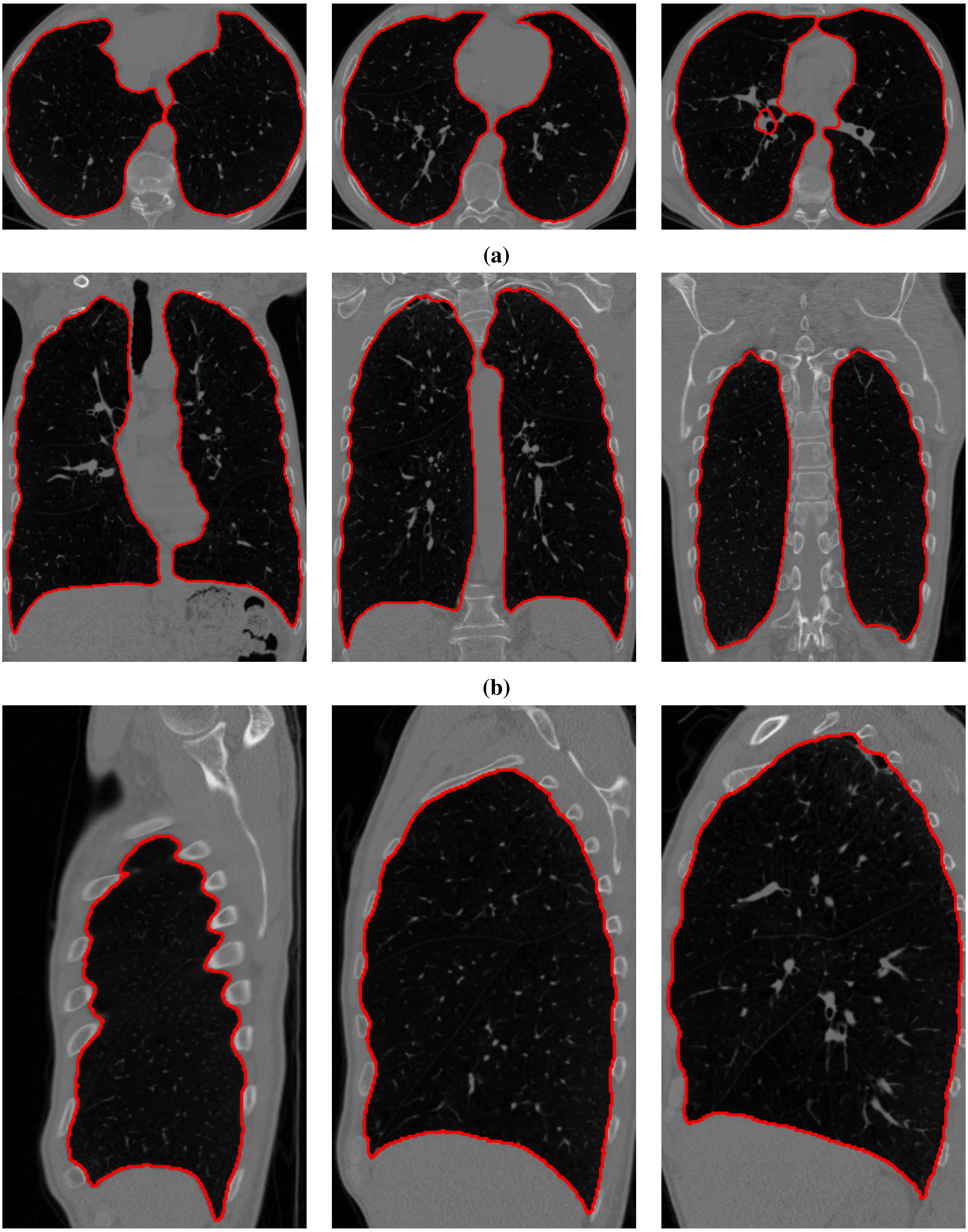
Data comes from the ***E**valuation of **M**ethods for **P**ulmonary **I**mage **Re**gistration 20**10*** (EMPIRE10) challenge, which is a registration challenge of thoracic CT data. The goal is to provide a platform for evaluation and comparison of registration algorithms [47].

The data consists of 30 pairs of chest CT scans from 30 different subjects scanned in the transverse (axial) plane. Each pair consists of a "fixed" (target) image and a "moving" (source) image, which encompasses two scans of the same subject to be registered in the registration challenge. In addition, for each fixed and moving image of each subject in the data set, a binary lung mask is included. The binary lung masks are automatically segmented [5] and have been visually checked and manually corrected if necessary. Therefore, for each of the 30 subjects there are 4 image series: one fixed scan and a matching mask, and one moving scan and a matching mask.

The scans are obtained from several different institutes and from a variety of different scanners. The scans are performed at different phases of the breathing cycle (inspiration, expiration, gating). The scans are a mix of subjects, who may be healthy or exhibit lung disease, but does not exhibit gross pathology. The data contains various voxel sizes and different slice spacings. In [47] a full description of the various sources and details of the data are available.

As the data comes from a variety of sources and of mixed quality, the data was inspected to secure good quality of the scans. Of the 30 different subjects, 19 subjects were deemed appropriate for use as atlases. As two scans including binary mask exists for each subject, only one of the scan and mask was used for atlases, while individual and independent atlases are to be used for the multi-atlas registration and segmentation. The 19 scans and binary masks will serve as individual atlases for the multi-atlas registration and segmentation.

In Figure 4.1, three transverse, coronal and sagittal slices from subject two from the EMPIRE10 dataset set are shown. The lung segmentation are shown as a red outline.



(a)

(b)

(c)

Fig. 4.1 Chest CT scan with delineation of the lungs in red. Images from the EMPIRE10 registration challenge dataset [47]. (a) Transverse slices, (b) coronal slices, and (c) sagittal slices.

The masks of the EMPIRE10 dataset segmented by the method in [5] are "cut" at the main bronchus, excluding the upper airways consisting of the trachea and the primary bronchi, while including the secondary bronchus and smaller airway branches descending to the lobes of the lungs. Figure 4.2 visualizes the segmentation of subject 2 from the EMPIRE10 dataset in a 3D rendering.

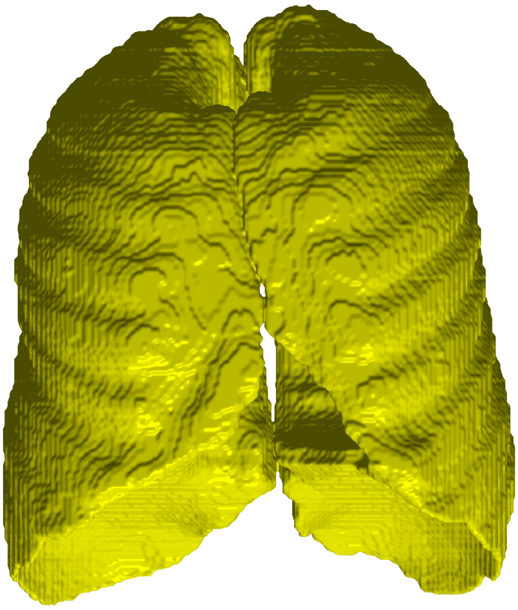


Fig. 4.2 3D visualization of the lungs from subject 2 (fixed) of the EMPIRE10 data. Visualization based on the automatic lung segmentation (binary mask) accompanying the EMPIRE10 data.

4.2 The Hospital of Vendsyssel data

Data comes from the Hospital of Vendsyssel, Denmark. Data consists of chest HRCT scans from 52 subjects, where 14 subjects were considered healthy and 38 subjects were diagnosed with having emphysema. Data was obtained at full inspiration and scans were performed on a Siemens SOMATOM Definition Flash CT scanner in the transverse (axial) plane. Of the 52 total scans, 10 scans from healthy subjects and 10 scans from subjects having emphysema were selected for segmentation by multi-atlas segmentation and graph cut segmentation. No selection criteria was used.

Manual delineation of the lungs from two subjects were available for select slices. One healthy subject (32 slices delineated) and one subject with emphysema (42 slices delineated). The delineation was performed by a medical student. These two subjects will be used for quantitative performance evaluation on slice-by-slice basis. The rest of the data will be qualitatively evaluated by visual inspection.

In Figure 4.3 transverse images of the scan of two healthy subjects are shown.

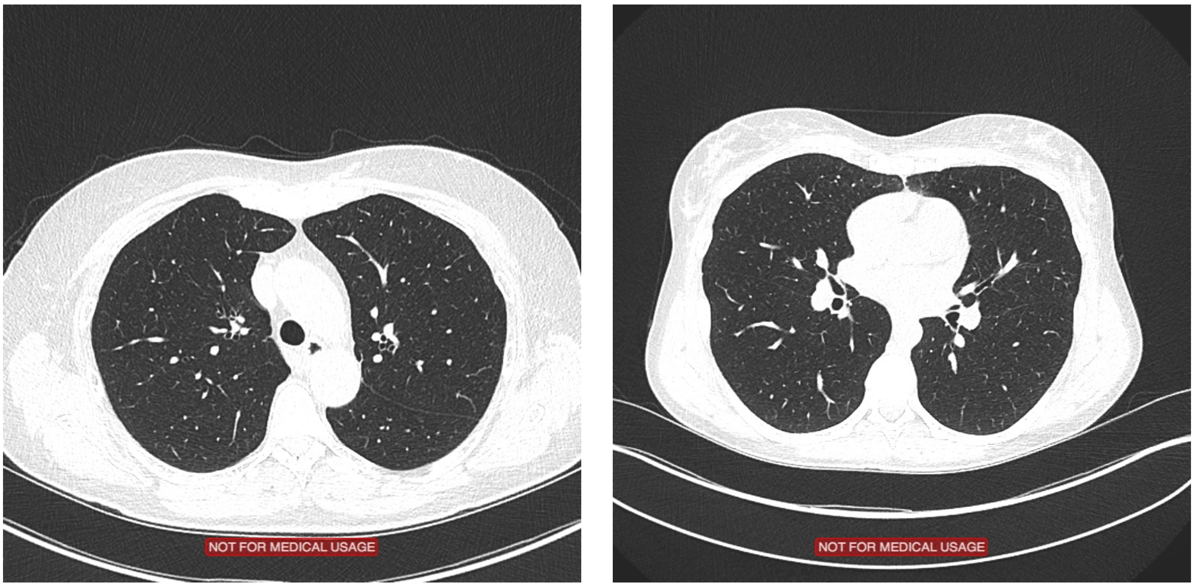


Fig. 4.3 Chest HRCT scan of healthy lungs from the Hospital of Vendsyssel data (transverse slice).

In Figure 4.3 transverse images of the scan of two subjects with emphysema are shown.

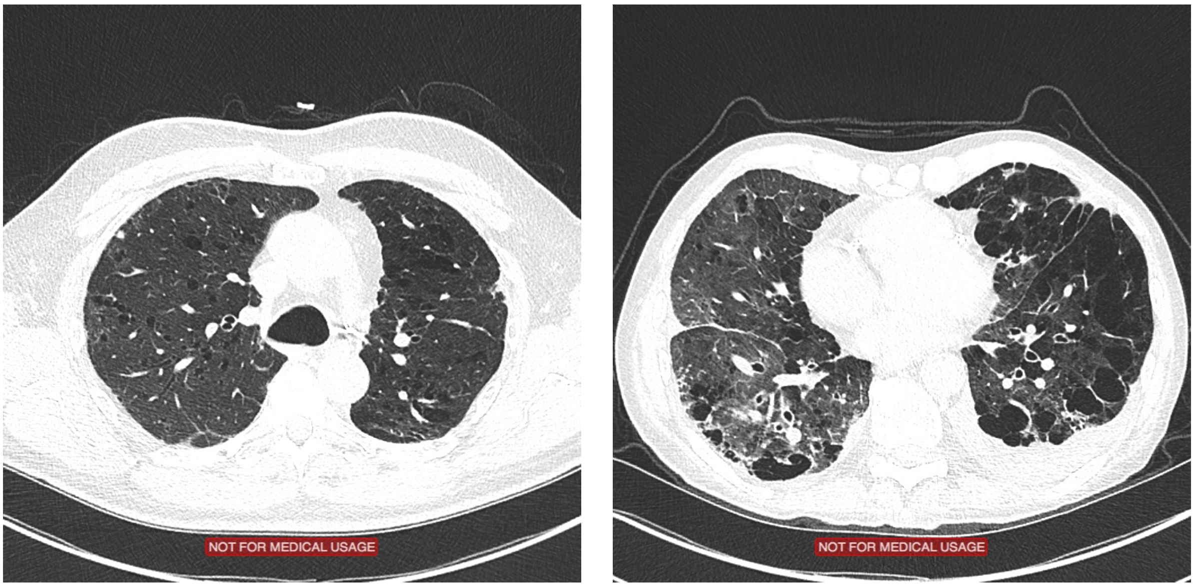


Fig. 4.4 Chest HRCT scan of emphysematous lungs from the Hospital of Vendsyssel data (transverse slice).

Chapter 5

Results

The multi-atlas segmentation method and the graph cut with multi-atlas registration method were used on the two different datasets described in Chapter 4, the EMPIRE10 dataset and the Hospital of Vendsyssel dataset. The results from the leave-one-out cross-validation experiment on the EMPIRE10 dataset will be presented first and the results from the Hospital of Vendsyssel dataset will follow.

5.1 Results from the EMPIRE10 dataset

The results in this section are based on the leave-one-out cross-validation experiment using the EMPIRE10 dataset. The EMPIRE10 dataset includes 19 individual and independent atlases, hence each leave-one-out segmentation includes 18 individual atlases for the segmentation and one atlas for evaluation.

In Figure 5.1 the evolution of DSC with respect to the number of atlases used for the multi-atlas segmentation is shown. The atlases are ranked and added according to their DSC with the highest ranking atlas as the first.

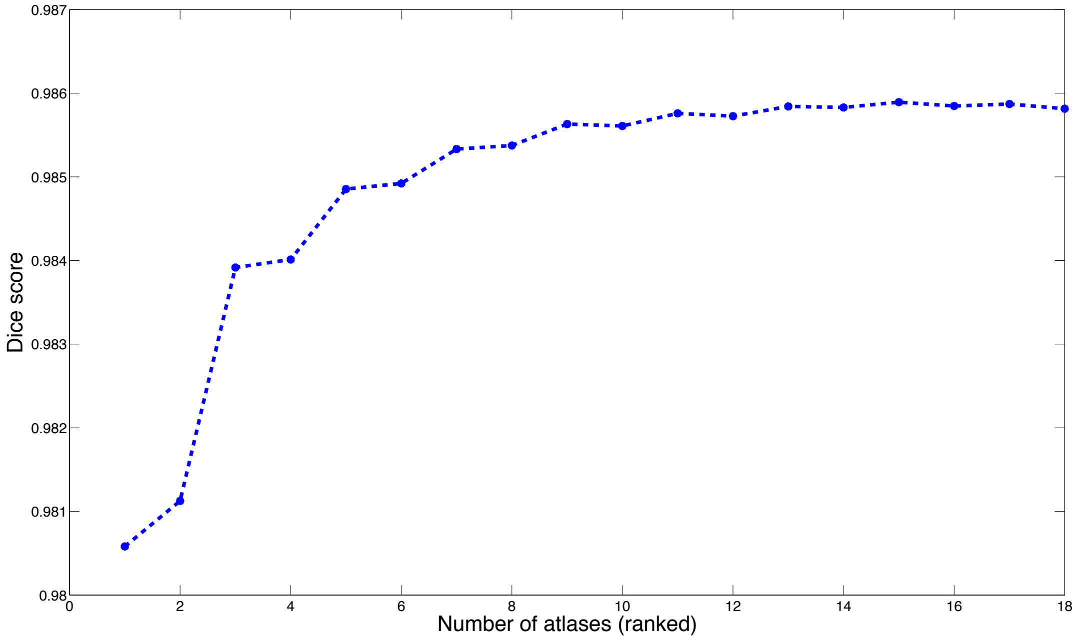


Fig. 5.1 Evolution of Dice score as more atlases are used in the multi-atlas segmentation method.

Figure 5.1 shows a flattening of the curve, reaching a plateau at around 14 to 18 number of atlases. This is in agreement with the results of [26] and shows a sufficient number of atlases for the segmentation.

In Figure 5.2 a bar plot of the DSC of the multi-atlas segmentation and graph cut with multi-atlas registration method are shown.

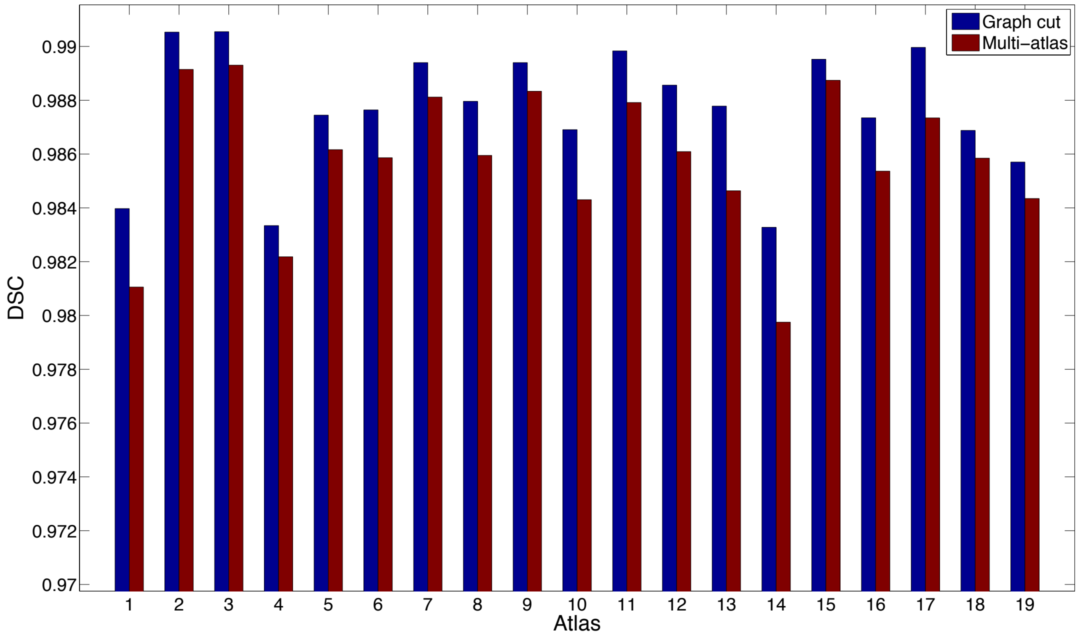


Fig. 5.2 Bar plot of DSC of the multi-atlas segmentation (red) and the graph cut segmentation (blue). The atlas numbers refers to the atlas from the EMPIRE10 dataset used for evaluation in the leave-one-out cross-validation experiment.

Figure 5.2 shows an increase in DSC for the multi-atlas registration and graph cut method in all cases compared to the multi-atlas segmentation.

In Figure 5.3 a box plot of the DSC of the multi-atlas segmentation and graph cut method are shown.

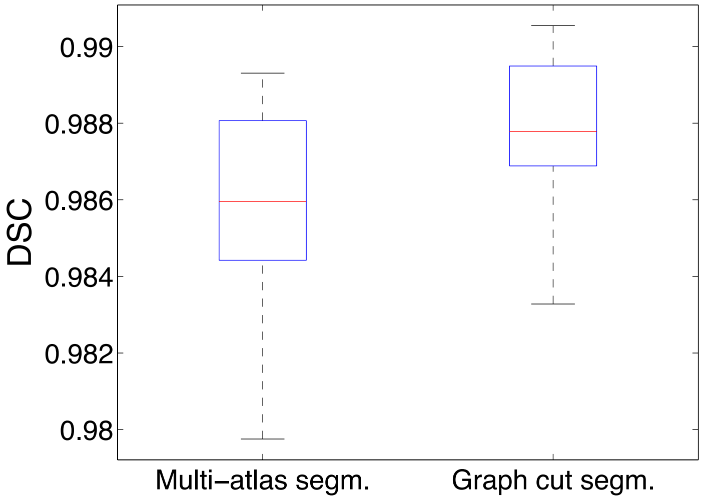


Fig. 5.3 Boxplot of multi-atlas segmentation (left) and graph cut segmentation (right).

Figure 5.3 shows a better performance in DSC of the graph cut segmentation compared to multi-atlas segmentation in terms of higher median value and higher values of the whiskers of the box plot. The graph cut segmentation has a statistical significant better performance ($p < 0.05$) than the multi-atlas segmentation confirmed by the Wilcoxon signed-rank test. The mean, median, maximum, minimum and standard deviation of the DSC for the multi-atlas segmentation and multi-atlas registration and graph cut segmentation are shown in Table 5.1.

Measure / Method	Mean DSC	Median DSC	Max. DSC	Min. DSC	Stdv. DSC
MAS	0.9858	0.9860	0.9893	0.9798	0.0027
GCS	0.9877	0.9878	0.9905	0.9833	0.0023

Tab. 5.1 The mean, median, maximum (max.), minimum (min.) and standard deviation values for the multi-atlas segmentation (MAS) and graph cut segmentation (GCS).

Results from the multi-atlas segmentation and graph cut segmentation are shown in Figure 5.4 along with the ground truth segmentation from the labeled images of the EMPIRE10 dataset. The images show an improvement of the graph cut segmentation compared to the multi-atlas segmentation especially at the border regions, where the graph cut segmentation smoothens the segmentation and corrects mis-registrations from the multi-atlas registration.

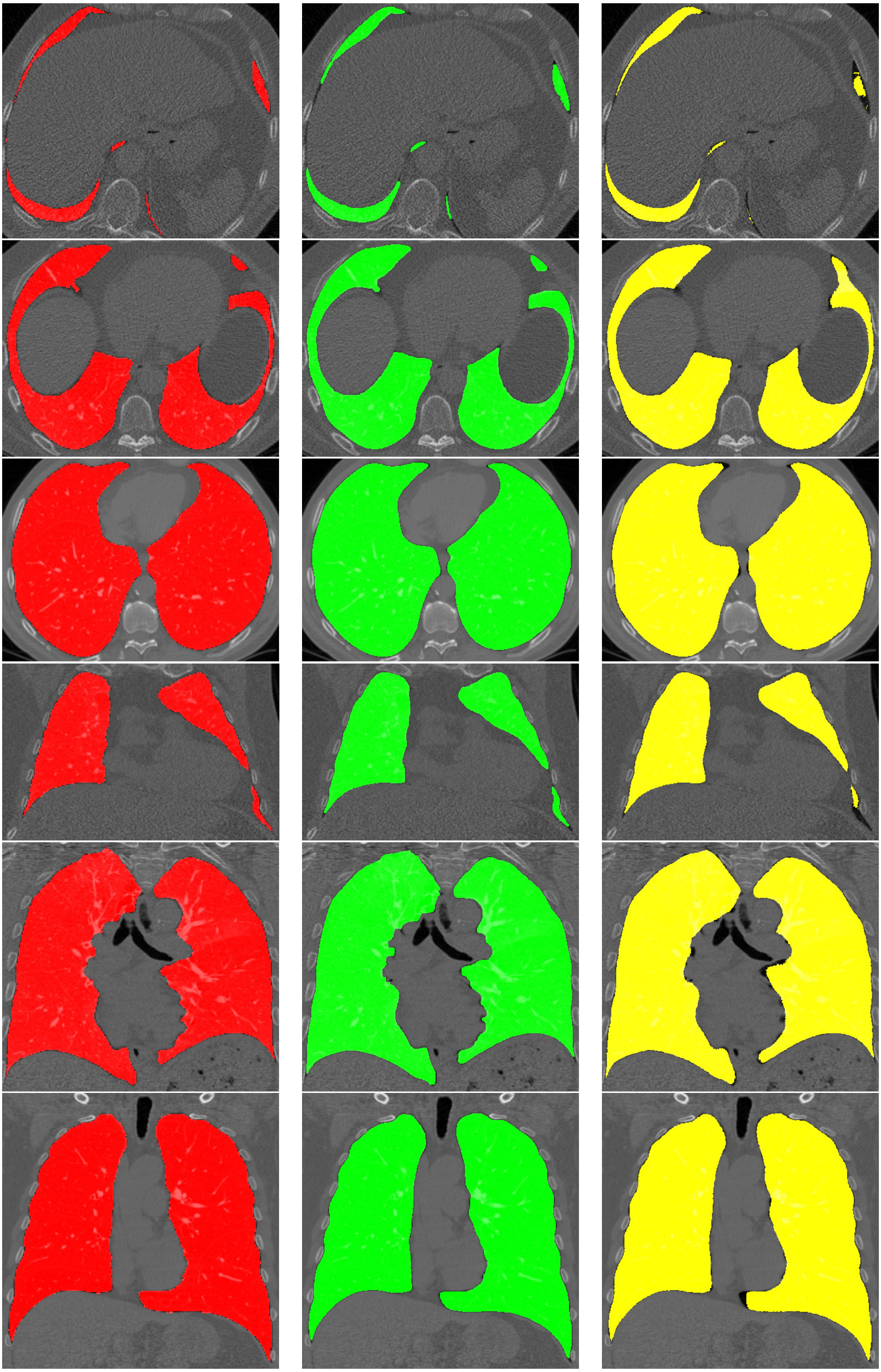
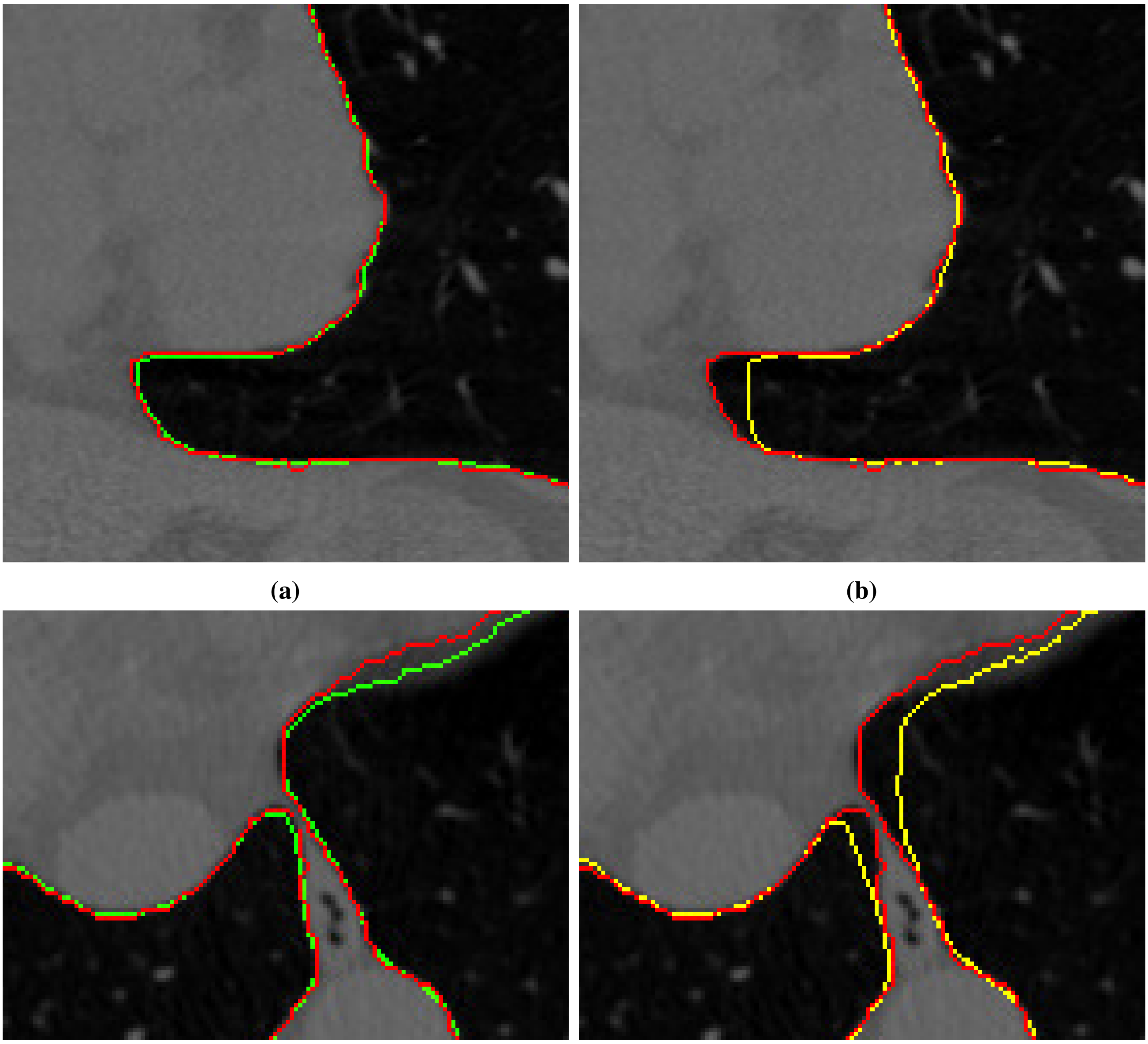


Fig. 5.4 Segmentation results. Each row shows a subject from the EMPIRE10 data. Ground truth = red, graph cut = green, multi-atlas = yellow.

By visual inspection, the results in Figure 5.4 show how the graph cut segmentation, which incorporates intensity modeling and neighboring modeling along with the spatial prior modeling from the multi-atlas registration, performs better than the multi-atlas segmentation. The intensity modeling of the graph cut segmentation corrects some of the mis-registration from the spatial model and in some cases expands the segmentation to darker areas part of the lungs, in other cases the segmentation is reduced as the mis-registration has led to segmentation of tissue not part of the lungs. It is also visible, that the graph cut segmentation smoothens the segmentation compared to the multi-atlas segmentation. This is due to the neighborhood modeling of the graph cut segmentation, which penalizes different labels for neighboring voxels.

Figure 5.5 shows some cropped and zoomed in examples of the improvements from the graph cut segmentation. The graph cut segmentation is closer to the ground truth segmentation from the EMPIRE10 dataset. It is evident that the intensity modeling and neighborhood modeling of the graph cut segmentation corrects for mis-registrations and draws the segmentation closer to the true lung boundary.

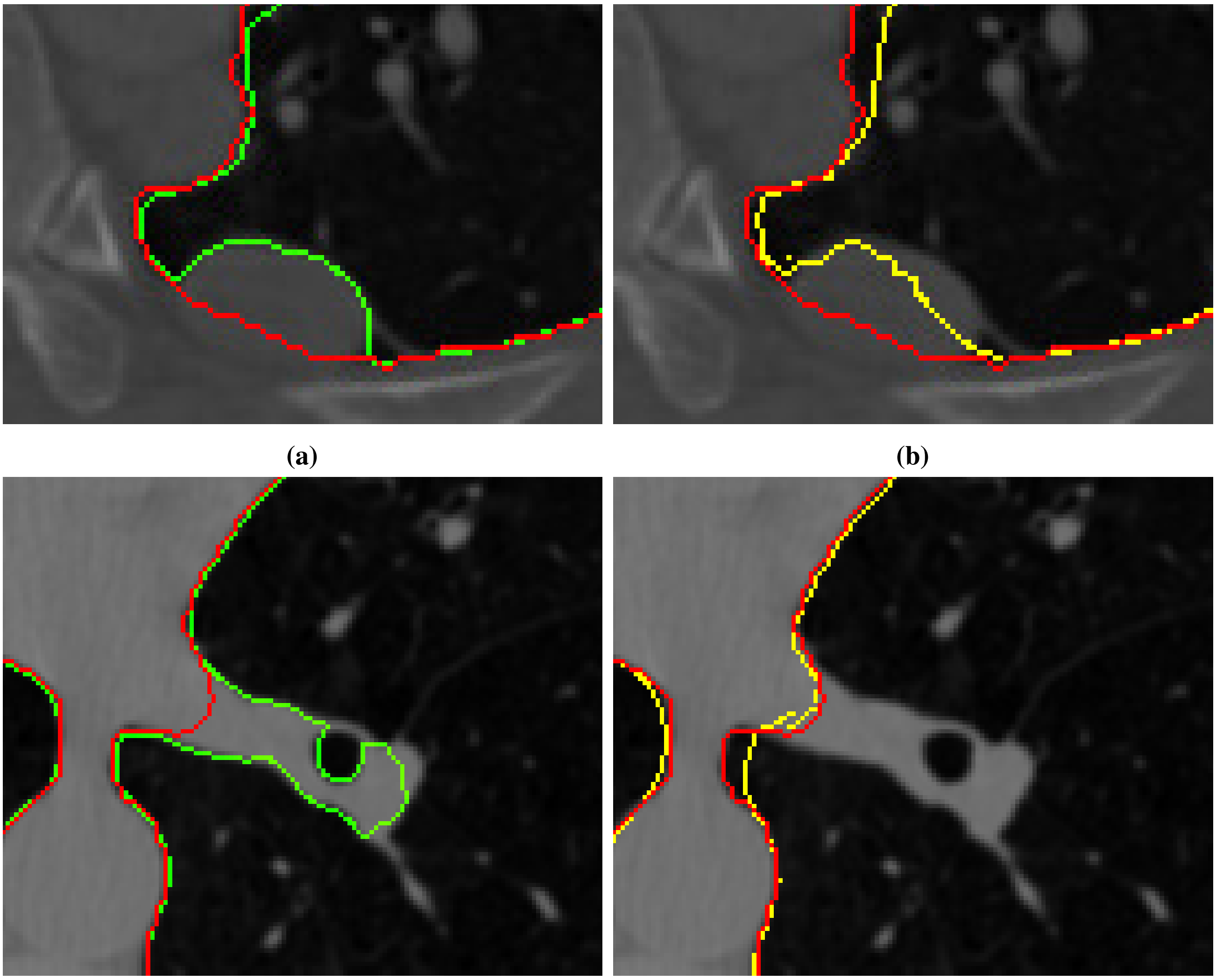


(a) (b)

(c) (d)

Fig. 5.5 Left: outline of result of graph cut segmentation (green) and ground truth (red). Right: outline of result of multi-atlas segmentation (yellow) and ground truth (red). Each row corresponds to a cropped image of one case in the EMPIRE10 dataset. The graph cut segmentation draws the segmentation closer to the ground truth segmentation and the true lung boundary.

As the above Figure 5.5 shows an improvement from the graph cut segmentation, Figure 5.6 below shows examples where by visual inspection it is evident that the graph cut segmentation deteriorates the segmentation compared to multi-atlas segmentation and performs worse than the multi-atlas segmentation compared to the ground truth segmentation.



(a) **(b)**

(c) **(d)**

Fig. 5.6 Left: outline of result of graph cut segmentation (green) and ground truth (red). Right: outline of result of multi-atlas segmentation (yellow) and ground truth (red). Each row corresponds to a cropped image of one case in the EMPIRE10 dataset. The graph cut segmentation deteriorates the segmentation compared to the multi-atlas segmentation as the segmentation is drawn away from the ground truth segmentation.

The first row of the above Figure 5.6 shows an example of a possible lung nodule or pleural wall thickening from a scan in the EMPIRE10 dataset. As the intensity of the nodule/thickening is very similar to the surrounding non-lung tissue, the intensity model of the graph cut segmentation deteriorates the segmentation by pulling it away from the true boundary. The bottom row shows another example of the graph cut segmentation pulling the segmentation away from the true boundary. Here it is not due to pathology, but an area close to the larger airways.

The hilar region of the lungs, where the bronchi, blood vessel, nerves and lymphs enter the lungs, is a challenging region for segmentation. An example of segmentation results from this region is shown in Figure 5.7.

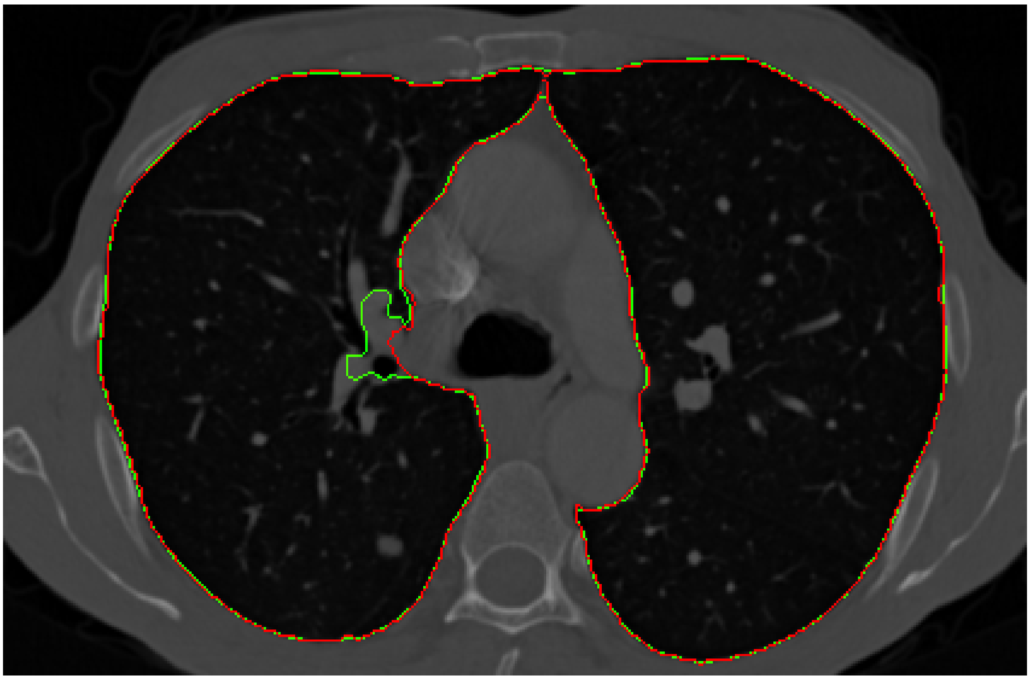


Fig. 5.7 Hilar region at the medial border of the lungs in an axial scan image. The ground truth segmentation is shown in red and the graph cut segmentation in green.

Segmentation results between the ground truth segmentation and the graph cut segmentation tend to differ most at the hilar region of the lungs, as in Figure 5.7.

5.1.1 Segmentation of airways

The automatic seed point selection and region growing algorithm for airway segmentation successfully segmented the main airways (trachea and main bronchi) in 18 of 19 cases of the EMPIRE10 dataset. In one of the cases the seed point was placed in area of the lung close to the trachea of the upper airways and therefore the segmentation was erroneous. This case needed manual placement of the seed point and afterwards the region growing segmentation proceeded as normal. The result of the airway segmentation for some of the cases in the EMPIRE10 dataset are shown in Figure 5.8 in a 3D rendering.

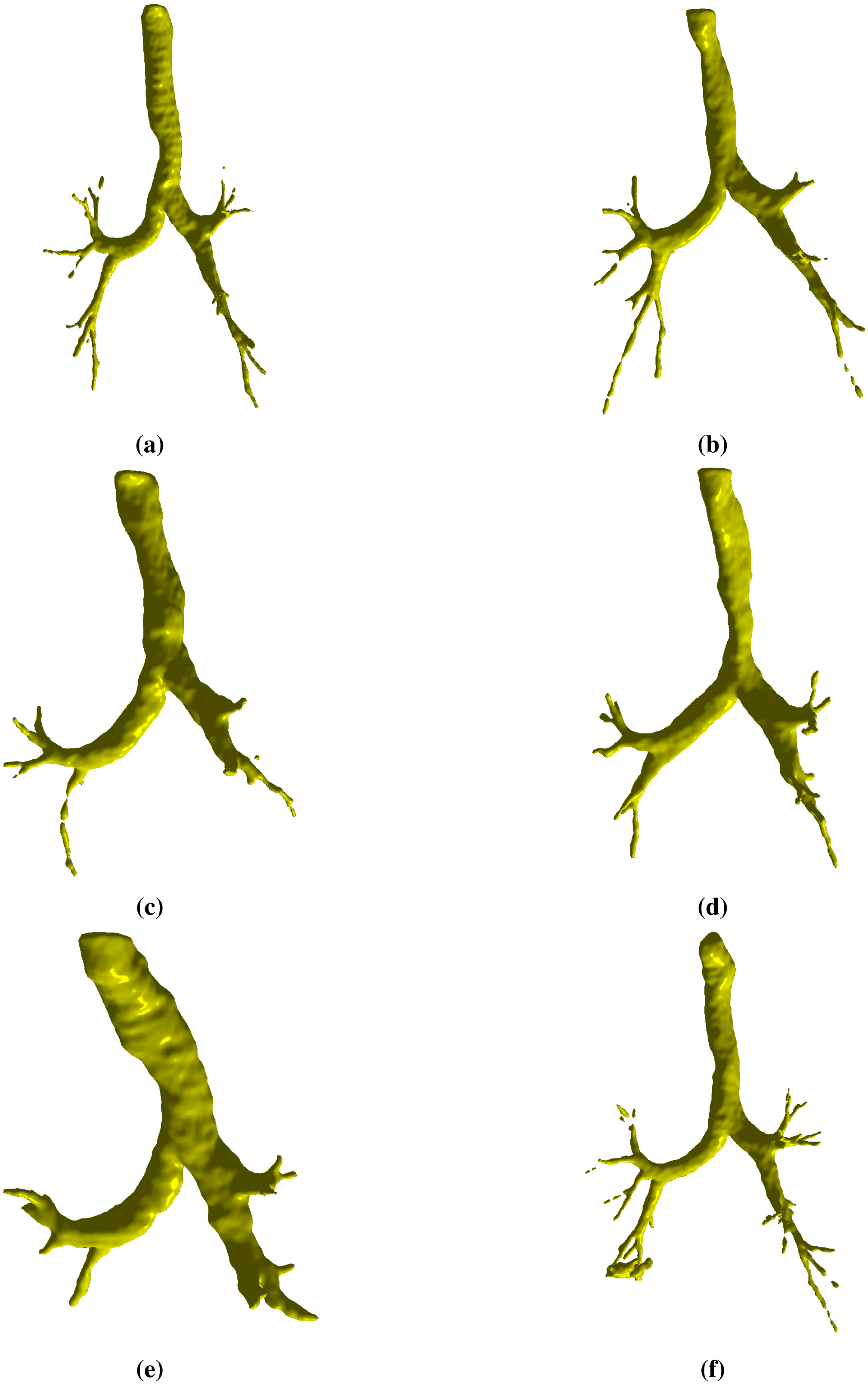
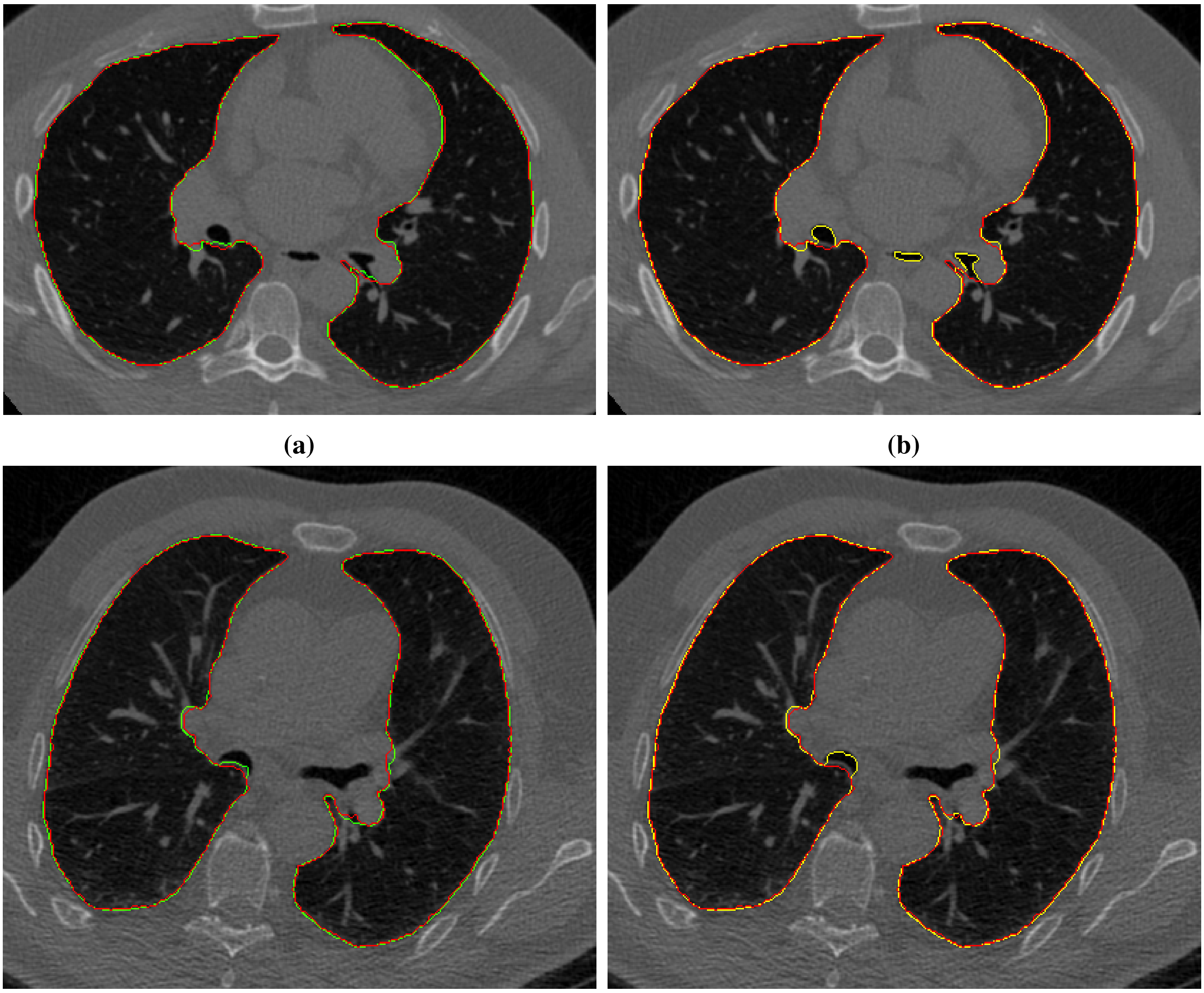


Fig. 5.8 Result of the airway segmentation from the EMPIRE10 dataset for 4 cases visualized in a 3D rendering.

As the above Figure 5.8 shows, the upper airways has successfully been segmented by the region growing algorithm. The branching depth of the segmentation varies as some of the airways descend into branches of the smaller airways, however, in all cases the upper airways consisting of the trachea and the primary bronchus are segmented, which are the ones to be excluded from the lung segmentation.

In Figure 5.9 some examples of the graph cut segmentation results before and after removal of the upper airways are shown.



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

Fig. 5.9 Left: Final graph cut segmentation with removed airways outlined in green and ground truth in red. Right: Graph cut segmentation without removed airways outlined in yellow and ground truth in red.

The above Figure 5.9 shows how the graph cut segmentation tends to segment the airways as the tissue border between lungs and airways are thin and intensities of the airways are similar to that of the lungs, which are incorporated in the intensity model of the graph cut segmentation.

5.2 Results from the Hospital of Vendsyssel dataset

The following results are for the data from the Hospital of Vendsyssel. Segmentation was successfully performed for 10 scans of healthy subjects and 10 scans of subjects with emphysema. Manual segmentations were available for select slices for 2 subjects: one healthy subject and one subject with emphysema. The results compared to these manual segmentations are made on a slice-by-slice basis and can therefore not be directly compared to the result of the EMPIRE10 dataset, as these were for a complete lung volume segmentation.

Figure 5.10 shows a bar plot of the DSC for each slice of the manually segmented healthy subject from the Hospital of Vendsyssel dataset.

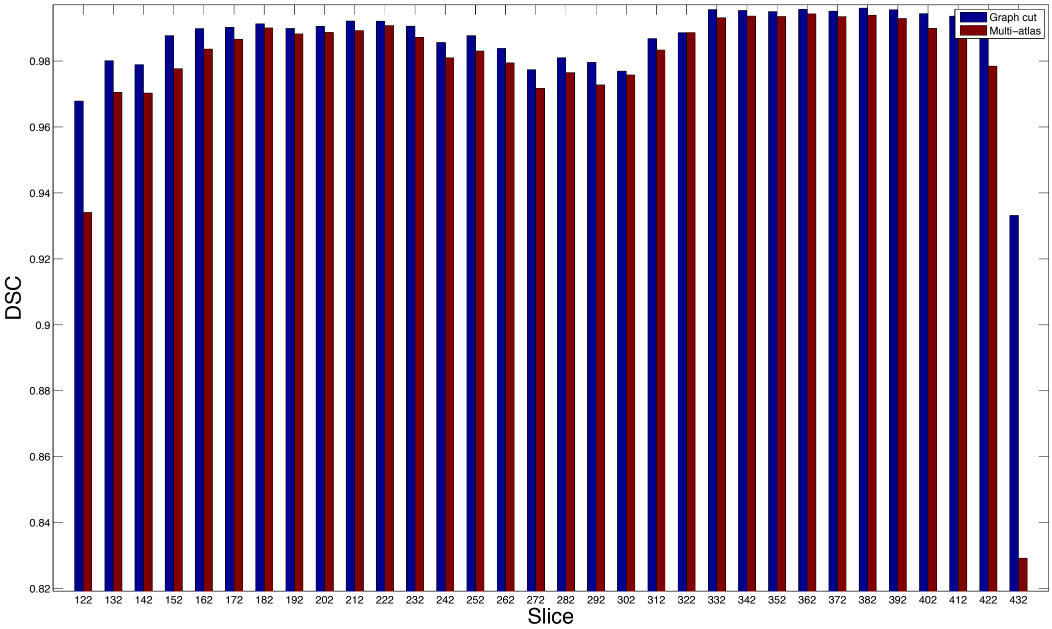


Fig. 5.10 Bar plot of the DSC of a healthy subject for multi-atlas segmentation and graph cut segmentation. The DSC is on a slice-by-slice basis.

The bar plot shows an increase in DSC for the graph cut segmentation for each slice. A box plot of the slice-by-slice results are shown in Figure 5.11.

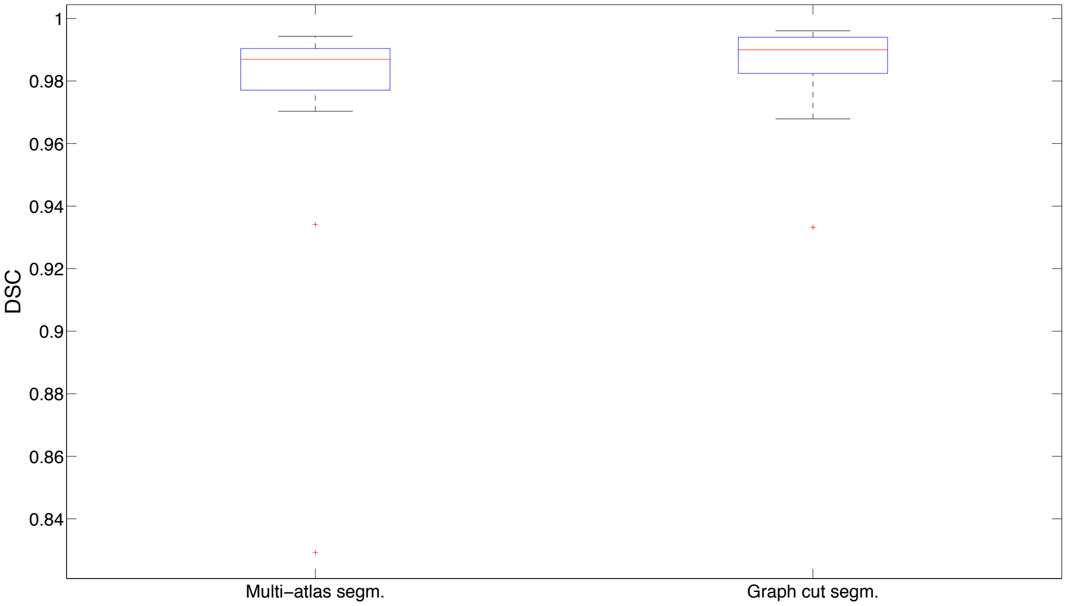


Fig. 5.11 Box plot of DSC of a healthy subject for multi-atlas segmentation and graph cut segmentation. The DSC is on a slice-by-slice basis.

The box plot of Figure 5.11 shows a slight overall increase in DSC for the graph cut segmentation with one less outlier. In Table 5.2 the slice-by-slice mean, median, maximum (max.), minimum (min.) and standard deviation values for the multi-atlas segmentation (MAS) and graph cut segmentation (GCS) are listed.

Method \ Measure	Mean DSC	Median DSC	Max. DSC	Min. DSC	Stdv. DSC
MAS	0.9785	0.9869	0.9943	0.8292	0.0296
GCS	0.9865	0.9900	0.9961	0.9332	0.0119

Tab. 5.2 The mean, median, maximum (max.), minimum (min.) and standard deviation values for the multi-atlas segmentation (MAS) and graph cut segmentation (GCS). The values are on a slice-by-slice basis.

Figure 5.12 shows coronal slices from the healthy subject with manual delineation and segmentation from multi-atlas segmentation and graph cut segmentation.

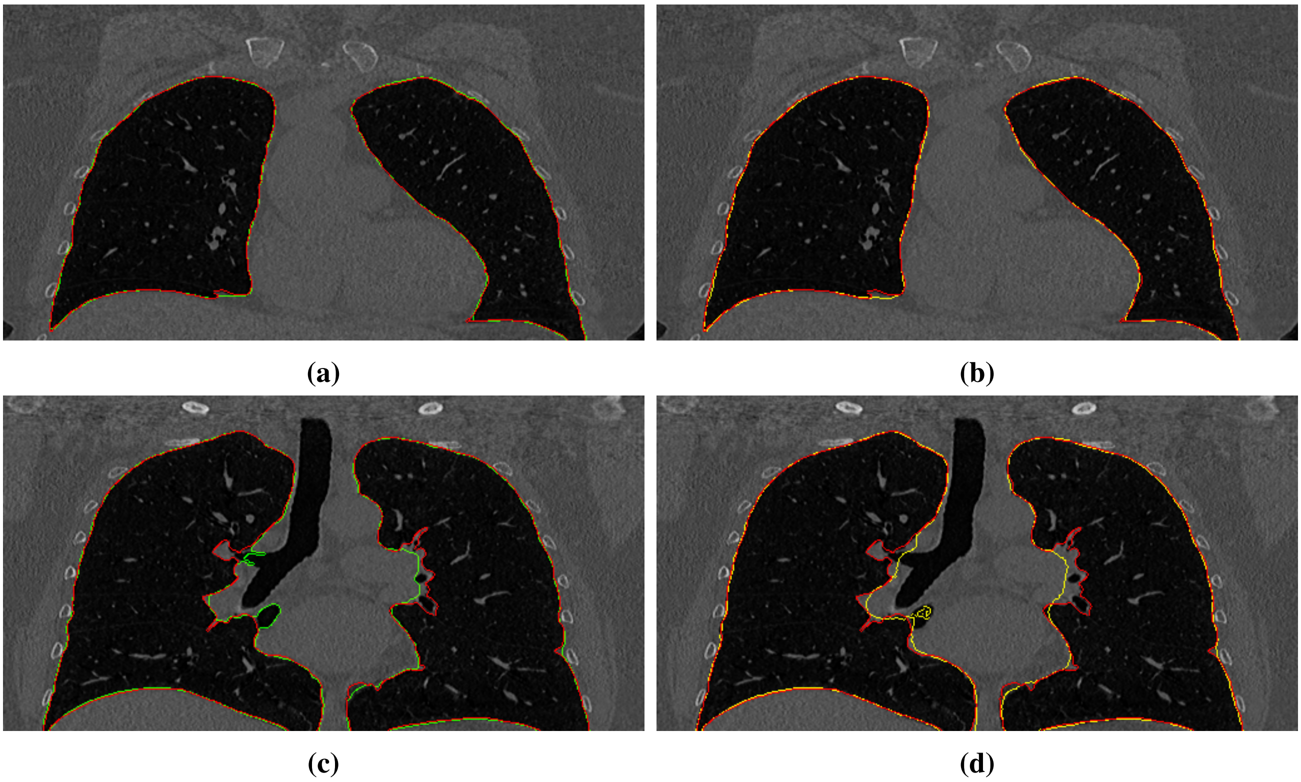

(a) (b)
(c) (d)

Fig. 5.12 Coronal slices of healthy subject with manual delineation in red, graph cut segmentation in green and atlas segmentation in yellow.

In Figure 5.13 the results of the segmentation of the subject with emphysema are shown. Figure 5.13 shows the bar plot of the DSC compared to the manual delineation on a slice-by-slice basis.

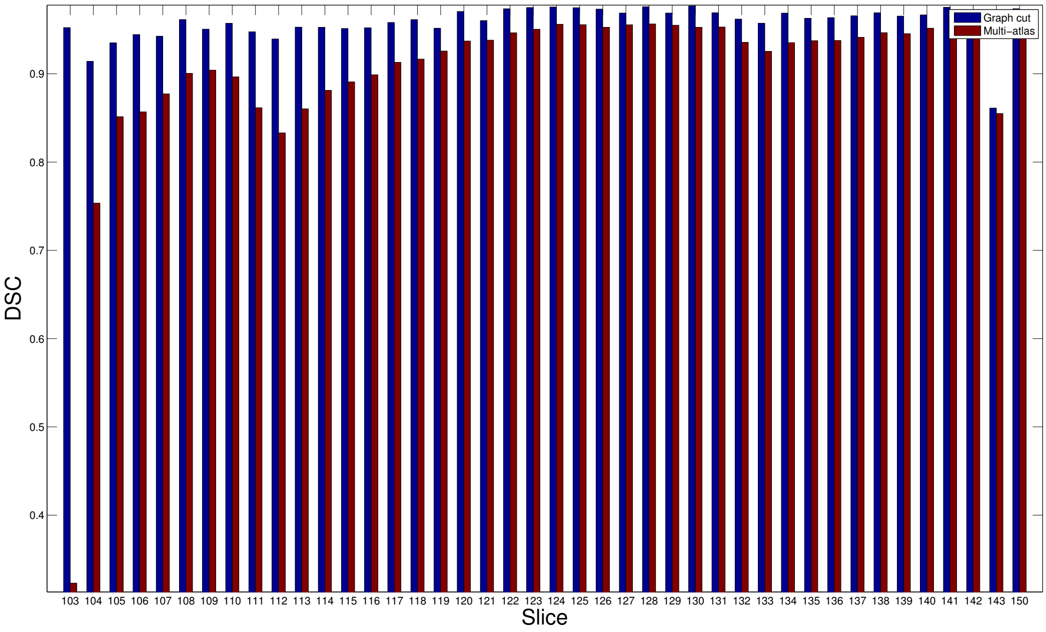


Fig. 5.13 Bar plot of the DSC of a subject emphysema for multi-atlas segmentation and graph cut segmentation. The DSC is on a slice-by-slice basis.

The bar plot shows an increase in DSC for the graph cut segmentation for each slice. A box plot of the slice-by-slice results are shown in Figure 5.14.

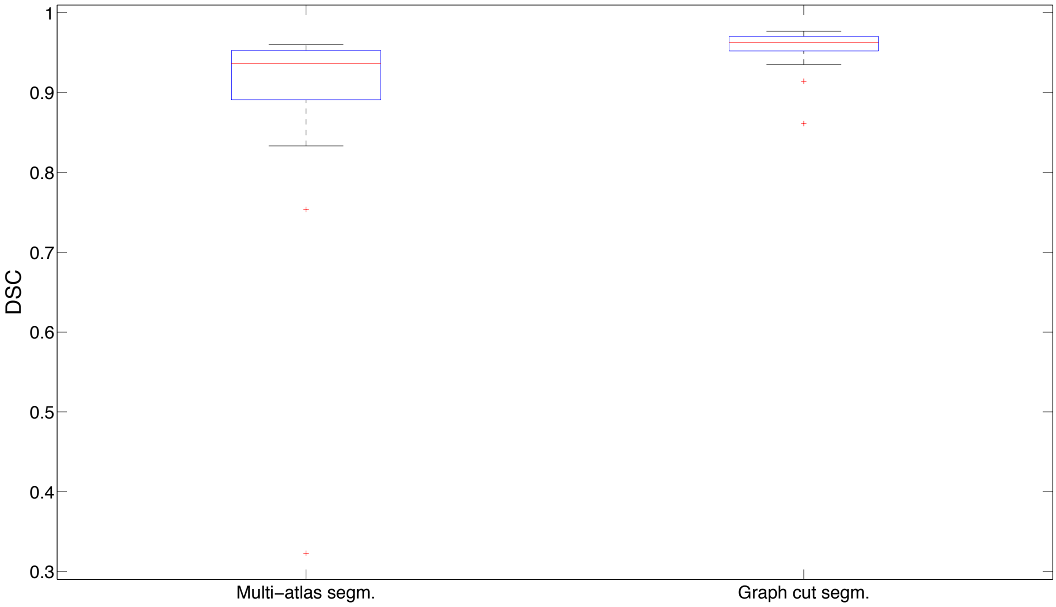


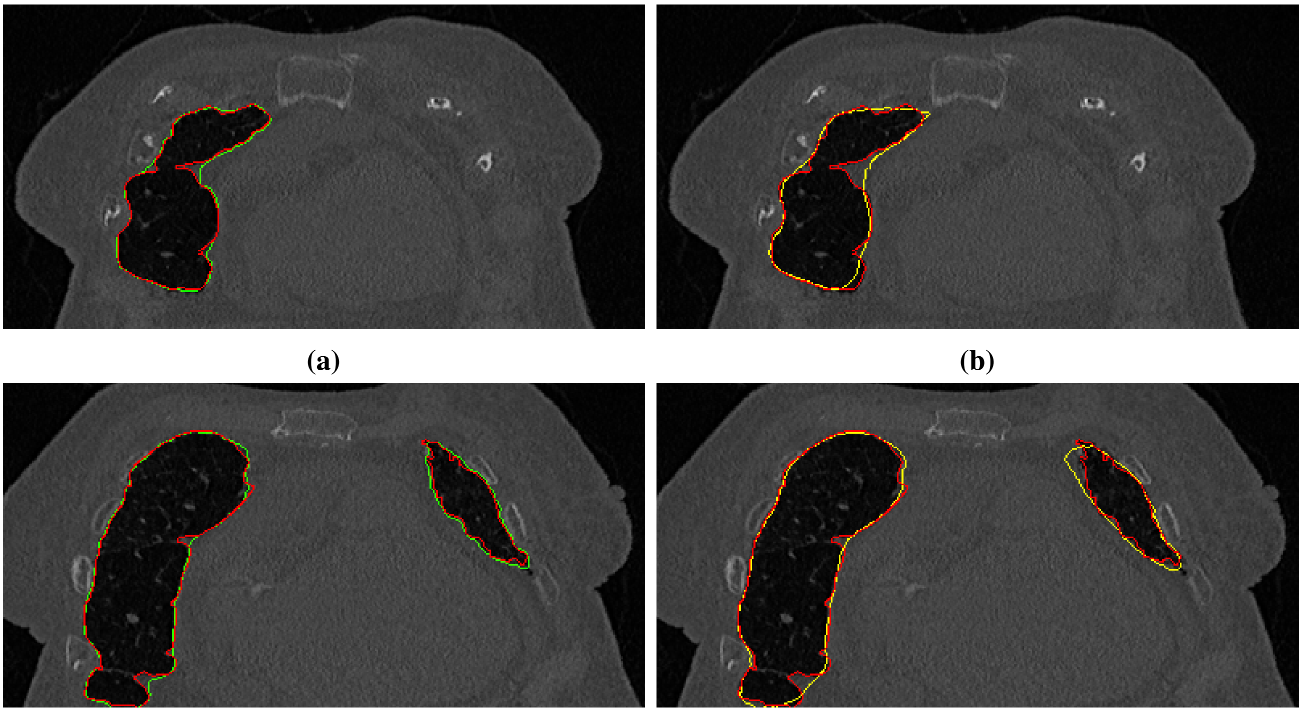
Fig. 5.14

The box plot of Figure 5.14 shows a larger overall increase in DSC for the graph cut segmentation compared to the multi-atlas segmentation than that of the healthy subject. Both the graph cut and multi-atlas segmentation contains two outliers, however, the multi-atlas segmentation contains outliers with DSC lower than the graph cut segmentation. In Table 5.3 the slice-by-slice mean, median, maximum (max.), minimum (min.) and standard deviation values for the multi-atlas segmentation (MAS) and graph cut segmentation (GCS) are listed.

Measure Method	Mean DSC	Median DSC	Max. DSC	Min. DSC	Stdv. DSC
MAS	0.9032	0.9365	0.9600	0.3230	0.1019
GCS	0.9584	0.9624	0.9769	0.8612	0.0202

Tab. 5.3 The mean, median, maximum (max.), minimum (min.) and standard deviation values for the multi-atlas segmentation (MAS) and graph cut segmentation (GCS). The values are on a slice-by-slice basis.

Table 5.3 shows lower DSC for both multi-atlas segmentation and graph cut segmentation compared to the DSC of the healthy subject, but the results show an increase in performance for the graph cut segmentation. Figure 5.15 shows coronal slices from the subject with emphysema with manual delineation and segmentation from multi-atlas segmentation and graph cut segmentation.



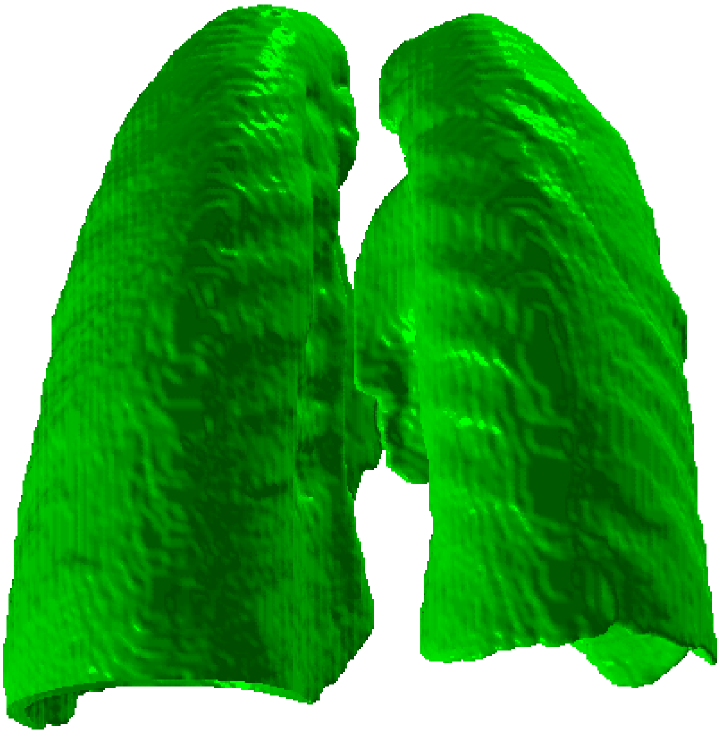
(a) (b)

(c) (d)

Fig. 5.15 Coronal slices of subject with emphysema with manual delineation in red, graph cut segmentation in green and atlas segmentation in yellow.

In Figure 5.16 a 3D rendered visualization of the result of the graph cut segmentation method of the lungs from one subject in the Hospital of Vendsyssel dataset is shown.

(a) Anterior view



(b) Posterior view

Fig. 5.16 3D rendered view of the result of the graph cut segmentation method of the lungs from one subject in the Hospital of Vendsyssel dataset.

Chapter 6

Discussion

In this thesis a method for lung segmentation from chest CT scans using multi-atlas registration and graph cuts has been presented. The performance of the method has been evaluated with respect to the Dice similarity coefficient and the results have been presented, which showed good segmentation results. The method was compared to a method based only on multi-atlas segmentation and showed favorable, statistically significant better results by an improved performance with respect to the Dice similarity coefficient. This chapter will present a discussion of the results and the evaluation of the method, improvements and limitations of the method as well as possibly future work.

6.1 Multi-atlas registration and segmentation

The multi-atlas registration served two purposes: to perform multi-atlas segmentation using majority voting and to build a probability map of a target image incorporated in the spatial prior model of the graph cut segmentation. The multi-atlas registration was based on individual and independent atlases from the EMPIRE10 dataset, which originated from a variety of different scanners, scanning protocols, slice thicknesses and breathing cycles. They were all registered to the target image using the same registration procedure. Although the atlas registration and multi-atlas segmentation showed good results, improvements could be made. Atlas selection for segmentation of a specific target image could improve multi-atlas registration results by selecting atlases from a pool of atlases based on the similarity with the target image. Atlas selection has been applied for segmentation of chest CT scans in [48], where the methods used, adaptive multi-atlas segmentation (AMAS) and adaptive local multi-atlas segmentation (ALMAS), showed statistically significant better results than traditional multi-atlas segmentation. Additionally, these methods can reduce computational resources used, while they stopping criteria for how many atlases are necessary to obtain a good performance for a specific segmentation. The method selects atlases based on the normalized mutual information (NMI) similarity measure between atlas and target image obtained by doing a computationally fast affine transformation, after which more refined registration are performed. Implementing the atlas selection method in the multi-atlas registration could improve results of both multi-atlas segmentation and graph cut segmentation and could reduce the number of atlases used for each target image by only selecting the most appropriate atlases based on the similarity measure.

For the multi-atlas segmentation, a majority voting decision fusion method was used to fuse the decisions of the multiple labels of the atlases registered to the target image. Majority voting simply thresholds the probability map of the target image build from the individual atlases at > 0.5 to select those labels which most atlases agree upon. Improved methods for decision fusion other than majority voting exist, which have proven to outperform majority voting. In the Selective and Iterative Method for Performance Level Estimation (SIMPLE) [32] an initial segmentation is made using all available atlases and traditional majority voting. All atlases are then evaluated with respect to this segmentation on the basis of the Dice similarity coefficient. Some atlases are rejected and some atlases are kept. The kept atlases are assigned weight based on Dice similarity coefficient according to their similarity with the estimated ground truth from the majority voting. This process is iterated and stopped when no atlases are rejected. This results in global weights assigned to the labels of individual atlases in contrast with majority weighting, where each atlas are weighted equally. The method has been further developed to include local weights, local SIMPLE [49], which has been applied to lung lobe segmentation from chest CT scans in patients with COPD and shown to produce statistically significant better results than majority voting and global SIMPLE. Therefore, using a local weighting for decision fusion in the multi-atlas segmentation could yield better results and the multi-atlas registration and graph cut segmentation could be compared to such a method as well.

6.2 Atlases

The atlases available for this study comes from the EMPIRE10 dataset [47]. They are obtained from a variety of different scanners, scanning protocols, slice thicknesses and breathing cycles. The atlases also comes from subjects both normal and with a lesser degree of pathology. The binary masks, the labels, of the atlases are based on an automatic lung segmentation method [5], which has been visually checked and manually corrected where necessary [47]. Other studies that perform similar multi-atlas registration as in this thesis use atlases with labels manually segmented by medical experts [5, 25, 26] and some have included multiple manual segmentation to test both inter- and intra-observer agreement between segmentations [25]. As the multi-atlas registration also depends on the quality of the atlas labels, it may improve on the results by having: a) a larger pool of atlases for atlas selection, b) a set of atlases from known sources, known scanning protocols and known patient history, and c) completely manually delineated labels from medical experts. This has not been available for this thesis, but could improve results and evaluation of performance. Furthermore, it could be of interest to do inter- and intra-observer agreement between different raters, as this could help to verify automatic lung segmentation as a more objective mean for lung segmentation, compared to possible subjective ratings between raters. This is true for evaluation of computer-aided diagnosis systems as well, where these systems might provide a more objective measure with better classification rates when making decisions compared to intra-observer agreement between raters.

6.3 Graph cut with shape term

The graph cut segmentation improved segmentation of the lungs in the leave-one-out cross-validation experiment using the EMPIRE10 dataset. This data primarily consists of healthy subjects or subjects with lesser pathology. The intensity modeling and neighborhood modeling of the graph cut segmentation improve the performance by correcting local mis-registrations from the multi-atlas registration. However, for certain pathologies in the lungs, the intensity and shape at the border of lung may be altered. This can lead to deteriorating segmentation results from the graph cut segmentation, as these abnormalities are not incorporated in the graph cut model. The multi-atlas registration provides a global shape prior for the graph cut segmentation in terms of the spatial prior model, but shape alterations with intensities different from that of the intensity model is not accounted for. Therefore, the graph cut model makes erroneous segmentations compared to the desired ground truth segmentation. In Figure 5.5 on page 36 and Figure 5.6 on page 37 of the results in Chapter 5, examples of improvements in segmentation using graph cut is shown and examples of deteriorations in segmentation using graph cut is shown.

As seen in the figures, the graph cut segmentation improves segmentation when intensities are similar to that of the lungs known by the intensity model (Figure 5.5) and deteriorates segmentation when intensities are more similar to that of the surrounding structures of the lungs (Figure 5.6). The pathology in the lungs of Figure 5.6 on page 37 alters both shape and intensity of the lungs and is therefore hard to segment. The multi-atlas segmentation performs the best as it contains global shape knowledge, which is not corrected for at the border by the intensity model of the graph cut segmentation. To segment such pathologies, a shape term as well as neighborhood prior constraints has been incorporated in a graph cut segmentation framework in [7]. They incorporate an energy term that penalizes at the segmentation boundary based on shape. This is done by comparing a vector connecting voxels p and q at the boundary of the lung to the gradient vector of a signed distance $\varphi(p)$ from the boundary of a given shape. The energy is low when the direction of the two vectors are similar, and therefore promotes a segmentation similar to "correct" lung shape. This incorporated into the graph cut segmentation framework as edge weights based on the shape term. The shapes are described by a level set distribution model (LSDM). In addition to the shape term, a neighbor constrained energy term is added, which assigns weights based on the distance of voxels at the boundary of the segmentation to neighboring structures, in this case the rib cage. The rib cage is segmented and it is argued that the lung segmentation will be drawn towards the correct lung border in case of pathology at the pleural wall, as the lung border must be "close" to the ribs. An example of a result of this method is shown in Figure 6.1 with images adapted from [7].

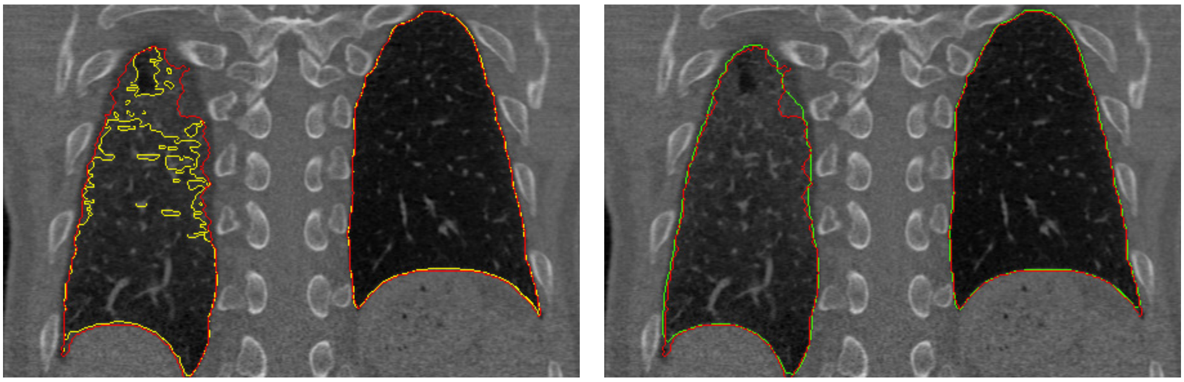
(a) True boundaries **(b)** Conventional method **(c)** Multi-shape graph cut

Fig. 6.1 The result of the multi-shape graph cut with neighbor prior constraints method. (a) shows the true boundaries, (b) the result of a conventional method and (c) the result of the multi-shape graph cut with neighbor prior constraints method. The lung is affected by pathology at the pleural wall (white arrow), which alters the shape and intensity of the lung and makes it difficult for correct lung border segmentation. Images adapted from [7]

As is evident in Figure 6.1, the multi-shape method is able to segment the lungs close to the true border even in case of pathology which alters shape and intensity and makes segmentation hard. It could be of interest in future work on the method presented in this thesis to investigate the use of a lung border proximity term to neighboring structures or to incorporate local shape knowledge in order to correct segmentation at pathologies, which alter shape and intensity. Also, validation and evaluation of the current method on more datasets with these pathologies including ground truth segmentation would be required.

6.4 Conventional segmentation methods and data quality

Conventional methods for lung segmentation have included thresholding methods with various morphological operations, such as the one in [24]. These methods are simple and requires less computational time than multi-atlas registration and graph cut segmentation. However, these methods are pure signal-based and relies only on the intensities. This generally works well for healthy subjects, but for patients with pathology, both shape and intensities are altered. In Figure 6.2 an example of one of the scans in the EMPIRE10 dataset is shown, along with a segmentation based on thresholding similar to [24] and a segmentation based on the graph cut segmentation presented in this thesis.

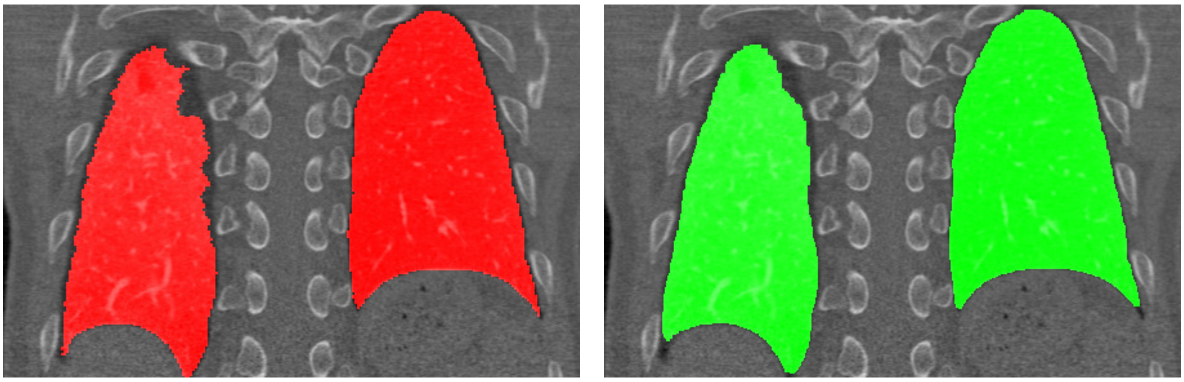


(a) Threshold segmentation **(b)** Graph cut segmentation

Fig. 6.2 Coronal view of scan with (a) ground truth segmentation outlined in red and threshold segmentation method outlined in yellow, and (b) ground truth segmentation outlined in red and graph cut segmentation method outlined in green.

It is evident in Figure 6.2 that the graph cut segmentation has superior performance compared to a conventional thresholding method. This is due to the diffuse changes in intensities caused by pathology. The graph cut segmentation method has global shape prior knowledge from the spatial prior model and smoothens the segmentation with the incorporation of intensity and neighborhood modeling.

In [5] a conventional method similar to [24] is used to segment the lungs from chest CT. However, this method has more advanced morphological operations applied to the segmentation. The method includes a convexity check of the lung boundaries, as the lung shapes are convex at the costal border. If errors are detected, the method tries to fix this by morphological closing, however, if further errors are detected the lungs are segmented using a multi-atlas segmentation method. The multi-atlas segmentation method is argued to be superior, however, requires more computational time and therefore the first attempt to segment is based on a conventional method to save time. This method is used to obtain the ground truth segmentation of the EMPIRE10 dataset. The method has been shown to perform well and is the basis for the evaluation of performance of the graph cut method presented in this thesis. In Figure 6.3 the ground truth segmentation of the EMPIRE10 dataset is shown on the left with the graph cut segmentation shown on the right.



(a) Ground truth segmentation **(b)** Graph cut segmentation

Fig. 6.3 Coronal view of scan with ground truth segmentation in red and graph cut segmentation in green.

Figure 6.3 shows irregularities at the border of the ground truth segmentation, while the boundaries of graph cut segmentation tends to be more smooth. Pathology present in lungs, such as in the figure above, makes segmentation more challenging than in healthy subjects, and it would therefore be of interest to perform evaluation and validation of the graph cut method on manual delineation of lungs with pathology to further validate the performance and investigate limitations.

An especially challenging region to segment using the graph cut segmentation method compared to ground truth is the hilar region of the lungs, where the bronchi, blood vessel, nerves and lymphs enter the lungs. This region contains intensities from the entering structures similar to the surrounding, non-lung structures and the correct "cut" of these structures at the hilar region is challenging. An example is shown in Figure 5.7 on page 38 of the results in Chapter 5, where it is visible how the graph cut segmentation does not follow the ground truth segmentation at the border of the hilar region of the lung. The ground truth segmentation is obtain by the method in [5], where the lungs are separated into the left and right lung and morphological closing with a sphere is applied to get a smooth border segmentation at the hilar region. A similar approach could be implemented as a post-processing step in the graph cut segmentation to obtain better segmentation at the hilar region of the lungs.

6.5 Evaluation of performance

The evaluation of the graph cut method and the multi-atlas segmentation method was based on a leave-one-out cross-validation of the EMPIRE10 dataset, as well as a comparison with manual segmentation of two patients from the Hospital of Vendsyssel dataset on select slices. The leave-one-out cross-validation experiment resulted in a mean Dice similarity coefficient (DSC) of 0.9877. The DSC is a commonly used measure of volume overlap for segmentation evaluation and in [5], which is the method used for segmentation of the ground truth segmentations in the EMPIRE10 dataset, a overall DSC of 0.95 was reported. The data used for evaluation of performance in [5] is based on scans from the NELSON lung cancer screening study, the Lung Image Database Consortium (LIDC) and from a database of patients with interstitial lung disease (ILD). The results are not directly comparable between the method of this thesis and the one in [5], as the data in [5] consists of 150 scans with subject containing various diseases such as COPD, lung cancer and ILD. The data available for evaluation of the graph cut segmentation method includes primarily healthy subjects and therefore it is desirable to be able to validate performance on scans including ground truth segmentation from patients with known lung pathology. The goal of the thesis was to segment healthy and emphysematous lungs robustly and accurately and while both healthy and emphysematous lungs have been successfully segmented based on visual inspection and with a DSC of 0.9584 on a subject with emphysema compared to ground truth, more validation data are needed for further evaluation performance.

6.6 Future work

To improve the graph cut segmentation method presented in this thesis, and to further evaluate its performance, future work may include the already mentioned implementation of atlas selection and weighted voting for the multi-atlas registration, a larger pool of atlases for the multi-atlas registration, an implementation of a shape term and neighborhood proximity term in the graph cut segmentation framework, and the separation of the lungs in order to apply morphological operations to the hilar region of the lungs. Also, the ability to divide the lungs into its distinct lung lobes might be of interest in order to separate quantification of the lungs into its anatomical regions. The fissures of the lung lobes visible on a CT scan may be included in a segmentation framework to divide the lungs into its lobes. Multi-atlas registration for lung lobe segmentation has been applied in [4] and [49].

Furthermore, airway segmentation might be improved by using wave front propagation as in [50] or appearance modeling as in [51]. An improvement of the airway segmentation would also be to detect the bifurcations of the airways to detect the primary bronchi bifurcation at the hilar region of the lungs.

As the parameters in the graph cut segmentation determine the weighting between the association potential and the interaction potential as well as the weighting between the intensity model and the spatial prior model, it would be of interest to do a thorough parameter tuning test. In this thesis the parameters have been chosen based on the work of other studies, however, an improvement might be seen by doing parameter tuning tests, although the gain in performance is expected to be modest.

The computational time and resources to perform the graph cut segmentation in the current work are high. This has not been prioritized in this thesis, but future work might include improvements in computational time and resources, as the time spent on multi-atlas registration ranges between 15 min. to 45 min, while the graph cut segmentation ranges between 30 min. and 8 hours in the current implementation. This could be reduced significantly by making use of parallel processing of the graph cut segmentation on the graphical processing unit (GPU) of a computer. The current implementation only runs on a single CPU core and is therefore slow for large volumes like the lungs. Multi-core optimization could also improve the computational performance significantly. Also, atlas selection might improve registration times by only selecting a subset of the atlases for registration.

Besides the improvements to the graph cut segmentation method, future work might include the application of the obtained segmentation from the graph cut method to computational analysis and quantification of pathology in the lungs, such as COPD. Computer-aided diagnosis (CAD) systems rely on a proper segmentation for quantification and classification. Therefore, it could be of interest to use the segmentation for feature extraction using texture analysis or density measures and feature classification to perform computerized analysis of e.g. emphysema, validated and correlated with clinical findings.

Chapter 7

Conclusion

In this thesis a method for automatic segmentation of the lungs from chest CT scans using multi-atlas registration and graph cut segmentation has been presented. The method incorporates spatial prior knowledge from multi-atlas registration, intensity modeling and neighborhood modeling into a graph cut segmentation framework. The spatial prior model from the multi-atlas registration provides global shape prior knowledge while the intensity and neighborhood modeling smoothens the segmentation and corrects local mis-registration errors. The method performed significantly better compared to a multi-atlas segmentation method in a leave-one-out cross-validation experiment and it successfully segmented lungs from both healthy subjects and patients with emphysema.

The method serves as an alternative to time-consuming and subjective manual delineations by providing an automated and objective segmentation. As lung segmentation is a prerequisite for further computational analysis of the lungs, the obtained segmentation may serve as a basis for quantitative analysis of the lungs.

Bibliography

[1] J. Vestbo, S. S. Hurd, A. G. Agustí, P. W. Jones, C. Vogelmeier, A. Anzueto, P. J. Barnes, L. M. Fabbri, F. J. Martinez, M. Nishimura, R. A. Stockley, D. D. Sin, and R. Rodriguez-Roisin, "Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: Gold executive summary.," *American journal of respiratory and critical care medicine*, vol. 187, pp. 347–365, Feb 2013.

[2] F. A. A. Mohamed Hoesein, B. de Hoop, P. Zanen, H. Gietema, C. L. J. J. Kruitwagen, B. van Ginneken, I. Isgum, C. Mol, R. J. van Klaveren, A. E. Dijkstra, H. J. M. Groen, H. M. Boezen, D. S. Postma, M. Prokop, and J.-W. J. Lammers, "Ct-quantified emphysema in male heavy smokers: association with lung function decline," *Thorax*, vol. 66, no. 9, pp. 782–787, 2011.

[3] L. Sorensen, S. Shaker, and M. de Bruijne, "Quantitative analysis of pulmonary emphysema using local binary patterns," *Medical Imaging, IEEE Transactions on*, vol. 29, pp. 559–569, Feb 2010.

[4] E. van Rikxoort, *Segmentation of anatomical structures in chest CT scans*. PhD thesis, Utrecht University, The Netherlands, 2009.

[5] E. M. van Rikxoort, B. de Hoop, M. A. Viergever, M. Prokop, and B. van Ginneken, "Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection," *Medical Physics*, vol. 36, no. 7, pp. 2934–2947, 2009.

[6] P. Lo, *Segmentation of Lung Structures in CT*. PhD thesis, University of Copenhagen, 2010.

[7] K. Nakagomi, A. Shimizu, H. Kobatake, M. Yakami, K. Fujimoto, and K. Togashi, "Multi-shape graph cuts with neighbor prior constraints and its application to lung segmentation from a chest CT volume," *Medical Image Analysis*, vol. 17, no. 1, pp. 62 – 77, 2013.

[8] T. Buzug, *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*. Springer, 2008.

[9] M. Takahashi, J. Fukuoka, N. Nitta, R. Takazakura, Y. Nagatani, Y. Murakami, H. Otani, and K. Murata, "Imaging of pulmonary emphysema: a pictorial review," *Int J Chron Obstruct Pulmon Dis*, vol. 3, no. 2, pp. 193–204, 2008.

[10] A. Faller, M. Schünke, G. Schünke, and E. Taub, *The Human Body: An Introduction to Structure and Function*. Basic sciences, Thieme, 2004.

[11] F. Martini, J. Nath, and E. Bartholomew, *Fundamentals of Anatomy and Physiology*. Pearson Education, Limited, 2011.

[12] A. Sirajuddin, "Normal thoracic anatomy and common variants," in *Clinically Oriented Pulmonary Imaging* (J. P. Kanne, ed.), Respiratory Medicine, pp. 1–17, Humana Press, 2012.

[13] Radiopaedia.org, "Lung parenchyma." `http://radiopaedia.org/articles/lung-parenchyma`, 2014.

[14] J. Shiraishi, Q. Li, D. Appelbaum, and K. Doi, "Computer-aided diagnosis and artificial intelligence in clinical imaging," *Seminars in Nuclear Medicine*, vol. 41, no. 6, pp. 449 – 462, 2011. Image Perception in Nuclear Medicine.

[15] L. Sorensen, M. Nielsen, P. Lo, H. Ashraf, J. Pedersen, and M. de Bruijne, "Texture-based analysis of copd: A data-driven approach," *Medical Imaging, IEEE Transactions on*, vol. 31, pp. 70–78, Jan 2012.

[16] F. A. M. Hoesein, M. Schmidt, O. M. Mets, H. A. Gietema, J.-W. J. Lammers, P. Zanen, H. J. de Koning, C. van der Aalst, M. Oudkerk, R. Vliegenthart, I. Isgum, M. Prokop, B. van Ginneken, E. M. van Rikxoort, and P. A. de Jong, "Discriminating dominant computed tomography phenotypes in smokers without or with mild {COPD}," *Respiratory Medicine*, vol. 108, no. 1, pp. 136 – 143, 2014.

[17] S. Bodduluri, J. D. N. Jr., E. A. Hoffman, and J. M. Reinhardt, "Registration-based lung mechanical analysis of chronic obstructive pulmonary disease (copd) using a supervised machine learning framework," *Academic Radiology*, vol. 20, no. 5, pp. 527 – 536, 2013.

[18] C. Dharmagunawardhana, S. Mahmoodi, M. Bennett, and M. Niranjan, "Quantitative analysis of pulmonary emphysema using isotropic gaussian markov random fields," in *9th International Conference on Computer Vision Theory and Applications*, pp. 44–53, January 2014.

[19] L. Sørensen, S. Shaker, and M. de Bruijne, "Texture based emphysema quantification in lung ct," in *The First International Workshop on Pulmonary Image Analysis*, 2008.

[20] T. Mcinerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," *Medical Image Analysis*, vol. 1, pp. 91–108, 1996.

[21] D. L. Pham, C. Xu, and J. L. Prince, "A Survey of Current Methods in Medical Image Segmentation," in *Annual Review of Biomedical Engineering*, vol. 2, pp. 315–338, Annual Reviews, 2000.

[22] Hu, Grossberg, and Mageras, "Survey of recent volumetric medical image segmentation techniques," in *Biomedical Engineering* (C. A. B. de Mello, ed.), InTech, 2009.

[23] M. Erdt, S. Steger, and G. Sakas, "Regmentation: A new view of image segmentation and registration," *Journal of Radiation Oncology Informatics*, vol. 4, no. 1, 2012.

[24] S. Hu, E. Hoffman, and J. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric x-ray ct images," *Medical Imaging, IEEE Transactions on*, vol. 20, pp. 490–498, June 2001.

[25] V. Fortunati, R. F. Verhaart, F. van der Lijn, W. J. Niessen, J. F. Veenland, M. M. Paulides, and T. van Walsum, "Tissue segmentation of head and neck ct images for treatment planning: A multiatlas approach combined with intensity modeling," *Medical Physics*, vol. 40, no. 7, pp. –, 2013.

[26] F. van der Lijn, T. den Heijer, M. M. Breteler, and W. J. Niessen, "Hippocampus segmentation in {MR} images using atlas registration, voxel classification, and graph cuts," *NeuroImage*, vol. 43, no. 4, pp. 708 – 720, 2008.

[27] F. van der Lijn, M. de Bruijne, S. Klein, T. den Heijer, Y. Hoogendam, A. van der Lugt, M. M. B. Breteler, and W. Niessen, "Automated brain structure segmentation based on atlas registration and appearance models," *Medical Imaging, IEEE Transactions on*, vol. 31, pp. 276–286, Feb 2012.

[28] M. Schaap, L. Neefjes, C. Metz, A. van der Giessen, A. Weustink, N. Mollet, J. Wentzel, T. W. van Walsum, and W. Niessen, "Coronary lumen segmentation using graph cuts and robust kernel regression.," *Information processing in medical imaging : proceedings of the ... conference*, vol. 21, pp. 528–539, 2009.

[29] D. Rueckert and J. Schnabel, "Registration and segmentation in medical imaging," in *Registration and Recognition in Images and Videos* (R. Cipolla, S. Battiato, and G. M. Farinella, eds.), vol. 532 of *Studies in Computational Intelligence*, pp. 137–156, Springer Berlin Heidelberg, 2014.

[30] L. A. Schwarz, *Non-rigid Registration Using Free-form Deformations*. PhD thesis, Technische Universität München, Germany, 2007.

[31] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. Miller, S. Pieper, and R. Kikinis, "3d slicer as an image computing platform for the quantitative imaging network," *Magnetic Resonance Imaging*, vol. 30, pp. 1323–1341, 11 2012.

[32] T. Langerak, U. Van der Heide, A. N. T. J. Kotte, M. Viergever, M. Van Vulpen, and J. P. W. Pluim, "Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (simple)," *Medical Imaging, IEEE Transactions on*, vol. 29, pp. 2000–2008, Dec 2010.

[33] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," *IEEE Trans Med Imaging*, vol. 23, pp. 903–921, Jul 2004.

[34] S. Kumar and M. Hebert, "Discriminative random fields," *International Journal of Computer Vision*, vol. 68, no. 2, pp. 179–201, 2006.

[35] Y. Boykov and O. Veksler, "Graph cuts in vision and graphics: Theories and applications," in *Handbook of Mathematical Models in Computer Vision* (N. Paragios, Y. Chen, and O. Faugeras, eds.), pp. 79–96, Springer US, 2006.

[36] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2Nd Edition)*. Wiley-Interscience, 2000.

[37] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient n-d image segmentation," *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109–131, 2006.

[38] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, pp. 1124–1137, Sept. 2004.

[39] K. Mori, J. Hasegawa, J. Toriwaki, H. Anno, and K. Katada, "Recognition of bronchus in three-dimensional x-ray ct images with applications to virtualized bronchoscopy system," in *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, vol. 3, pp. 528–532 vol.3, Aug 1996.

[40] Stefan Klein and Marius Staring and Keelin Murphy and Max A. Viergever and Josien P.W. Pluim, "elastix: a toolbox for intensity-based medical image registration," *IEEE Transactions on Medical Imaging*, vol. 29, pp. 196 – 205, January 2010.

[41] Denis P Shamonin and Esther E Bron and Boudewijn P.F. Lelieveldt and Marion Smits and Stefan Klein and Marius Staring, "Fast Parallel Image Registration on CPU and GPU for Diagnostic Classification of Alzheimer's Disease," *Frontiers in Neuroinformatics*, vol. 7, pp. 1–15, January 2014.

[42] Elastix, "References." `http://elastix.bigr.nl/wiki/index.php/References`, 2014.

[43] M. Staring, M. E. Bakker, J. Stolk, D. P. Shamonin, J. H. Reiber, and B. C. Stoel, "Towards local progression estimation of pulmonary emphysema using CT," *Med Phys*, vol. 41, p. 021905, Feb 2014.

[44] Elastix, "Parameter file database." `http://elastix.bigr.nl/wiki/index.php/Par0015`, 2014.

[45] MATLAB, *version 8.2.0.701 (R2013b).* Natick, Massachusetts: The MathWorks Inc., 2013.

[46] M. M. S. AG, "Mevislab." `http://www.mevislab.de`, 2014.

[47] K. Murphy, B. van Ginneken, J. M. Reinhardt, S. Kabus, K. Ding, X. Deng, K. Cao, K. Du, G. E. Christensen, V. Garcia, T. Vercauteren, N. Ayache, O. Commowick, G. Malandain, B. Glocker, N. Paragios, N. Navab, V. Gorbunova, J. Sporring, M. de Bruijne, X. Han, M. P. Heinrich, J. a. Schnabel, M. Jenkinson, C. Lorenz, M. Modat, J. R. McClelland, S. Ourselin, S. E. a. Muenzing, M. a. Viergever, D. De Nigris, D. L. Collins, T. Arbel, M. Peroni, R. Li, G. C. Sharp, A. Schmidt-Richberg, J. Ehrhardt, R. Werner, D. Smeets, D. Loeckx, G. Song, N. Tustison, B. Avants, J. C. Gee, M. Staring, S. Klein, B. C. Stoel, M. Urschler, M. Werlberger, J. Vandemeulebroucke, S. Rit, D. Sarrut, and J. P. W. Pluim, "Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge.," *IEEE transactions on medical imaging*, vol. 30, pp. 1901–20, Nov. 2011.

[48] E. M. van Rikxoort, I. Isgum, Y. Arzhaeva, M. Staring, S. Klein, M. A. Viergever, J. P. Pluim, and B. van Ginneken, "Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus," *Medical Image Analysis*, vol. 14, no. 1, pp. 39 – 49, 2010.

[49] M. Agarwal, E. A. Hendriks, B. C. Stoel, M. E. Bakker, J. H. C. Reiber, and M. Staring, "Local simple multi atlas-based segmentation applied to lung lobe detection on chest ct," 2012.

[50] U. L. Stephansen, R. W. Horup, M. Gram, J. T. Olesen, J. Carl, A. S. Korsager, and L. R. Østergaard, "Airway tree segmentation for optimal stent placement in image-guided radiotherapy," *International Workshop on Pulmonary Image Analysis*, pp. 135–146, 2011.

[51] P. Lo, J. Sporring, H. Ashraf, J. J. Pedersen, and M. de Bruijne, "Vessel-guided airway tree segmentation: A voxel classification approach," *Med Image Anal*, vol. 14, pp. 527–538, Aug 2010.