# Supervoxel-based Image Registration and Analysis with Application to Lung Cancer




**Adam Szmul**

Linacre College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Michaelmas 2019

# Acknowledgements

This DPhil was an incredible adventure and a wonderful journey for me, in so many ways. I am grateful to Cancer Research UK and EPSRC for making this journey possible by awarding me with a DPhil studentship.

I would like to thank my brilliant supervisors for their support: Prof. Julia Schnabel and Prof. Vicente Grau, who cared not only about the research but also about the person behind that: me as a human being. It was a great privilege to be one of your students. I am very grateful to Dr Bartłomiej Papież for your everyday help and meritorious advice, but also for being a dear friend. I am thankful to my clinical collaborators from the Churchill Hospital: Prof. Fergus Gleeson and Dr Tahreema Matin for your invaluable help.

It was a real pleasure to work with and along my team colleagues: Jola, Andre, Ben, Russ, Jesper, Monica and Thais, who created such a friendly work environment. I thank George, Andre and Dongli for being great friends and sharing so much time with me. I owe special thanks to George for all your help and for your patience with my writing struggles. Thank you Adam for all our fruitful discussions and intellectual challenges.

I am grateful to Prof. Tomasz Zielinski and Prof. Andrzej Skalski, who first showed me the research world and helped me to develop interest in uncovering more of it. I would like to thank Prof. Mark Jankinson, Prof. Marius Staring and Prof. Jens Rittscher for being my examiners.

Nothing of that could have ever happened without unconditional love and constant support from my parents. It is them who helped me to develop natural curiosity and appreciation of every little thing. I thank my sister and her family for their love, support and understanding.

Lastly, and most importantly, I thank with all my heart and soul the most important person in my life, my wife. Judyta, only your great sacrifices to follow my dream allowed me to complete my journey of the DPhil. I am extremely grateful for all your help and support throughout the journey.

# Statement of Originality

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at the University of Oxford or any other educational institution, except where due acknowledgement is made in the thesis.

Any contribution made to the research by others, with whom I have worked at the University of Oxford or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Adam Szmul
16/11/2019.

# Abstract

Lung diseases, including lung cancer, are amongst the largest burdens to healthcare systems worldwide. Improving the extraction of information from imaging data has the potential to provide more accurate diagnosis and more effective treatment. For example, assessing regional lung function could guide more effective radiotherapy treatment to spare well-functioning parts of the lungs. One meaningful regional representation is achieved by locally clustering similar image voxels together using the concept of supervoxels, which allows for bulk regional processing and analysis, whilst removing redundant information. This thesis aims to lay further foundations for supervoxel-based image analysis by presenting novel deformable lung image registration frameworks, as well as a method to estimate lung ventilation using supervoxels.

We demonstrate how a supervoxel-based image representation can be combined with graph cuts as a discrete optimisation-based approach to provide efficient and accurate 3D deformable image registration. Our work overcomes some of the previous limitations by posing the problem on a graph created by adjacent supervoxels. We further show that the application of a relaxed graph representation of the image, followed by guided image filtering over the estimated deformation field, allows us to model 'sliding motion'. Applying this method to lung image registration results in highly accurate image registration and anatomically plausible estimations of the deformations.

The resulting deformation fields might be further applied to estimate lung ventilation maps. We present a novel approach for estimating regional lung ventilation from dynamic lung CT imaging. Our method combines a supervoxel-based image representation with deformable image registration, performed between peak breathing phases, during which we track changes in intensity for a number of layers of previously extracted supervoxels. Such a region-based approach is expected to be physiologically

more consistent with lung anatomy than methods relying on voxel-wise relationships.

Our work also presents novel approaches for performing more accurate hyperpolarised $^{129}$Xenon MRI (XeMRI) analysis. We propose a multi-modal lung image registration enhanced with personalised motion model approach derived from lung 4DCT. The approach with *a prior* motion model is particularly important for regions where there is not enough information to reliably drive the registration process, as in the case of XeMRI and proton density MRI (pMRI) to CT registration. Subsequently we introduce a framework for breathing motion correction of the dynamic sequence of XeMRI using a lung atlas-based approach. Such a method for the breathing motion correction is expected to facilitate the analysis of temporal lung ventilation.

# Publications

## Journal Articles:

- Adam Szmul, Tahreema Matin, Fergus Gleeson, Julia A. Schnabel, Vicente Grau, and Bartłomiej W. Papież, **Patch-based lung ventilation estimation using multi-layer supervoxels**, Computerized Medical Imaging and Graphics, Volume 74, pages 49-60, 2019.

- Adam Szmul, Bartłomiej W. Papież, Andre Hallack, Vicente Grau, and Julia A. Schnabel, **Supervoxels for graph cuts based deformable image registration using guided image filtering**, Journal of Electronic Imaging, 26(6):26-26-13, 2017.

## Conference Publications:

- Adam Szmul, Tahreema Matin, Fergus Gleeson, Julia A. Schnabel, Vicente Grau, and Bartłomiej W. Papież, **XeMRI to CT Lung Image Registration Enhanced with Personalized 4DCT-derived Motion Model**, MICCAI Workshop: Thoracic Image Analysis, Springer, pages 260-271, 2018.

- Adam Szmul, Bartłomiej W. Papież, Daniel Bulte, Ozkan Doganay, Fergus Gleeson, Julia A. Schnabel, and Vicente Grau, **Atlas-based breathing motion correction for dynamic lung XeMRI**, ISMRM, 2018.

- Adam Szmul, Bartłomiej W. Papież, Tahreema Matin, Fergus V. Gleeson, Julia A. Schnabel, and Vicente Grau **Regional lung ventilation estimation based on supervoxel tracking**, SPIE Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling, 10576:10576-10576-7, 2018.

- Adam Szmul, Bartłomiej W. Papież, Russell Bates, Andre Hallack, Julia A. Schnabel, and Vicente Grau, **Graph Cuts-Based Registration Revisited: A Novel Approach for Lung Image Registration Using Supervoxels and Image-Guided Filtering**, IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW): WBIR 2016, pages 592-599, 2016.

- Bartłomiej W. Papież, Adam Szmul, Vicente Grau, Michael Brady, and Julia A. Schnabel, **Non-local graph-based regularization for deformable image registration**, MICCAI Workshop: Medical Computer Vision and Bayesian and Graphical Models for Biomedical Imaging, Springer, pages 199-207, 2016.

# Contents

**List of Figures** **iv**

**List of Tables** **vi**

**1 Introduction** **1**
1.1 Scientific Contributions . . . . . 2
1.2 Thesis Outline . . . . . 4

**2 Background on Lung Imaging** **7**
2.1 Lung Cancer . . . . . 7
2.1.1 Lung Cancer Treatment . . . . . 9
2.1.2 Lung Function Evaluation Techniques . . . . . 12
2.2 Lung Imaging Techniques . . . . . 13
2.2.1 X-ray and Computed Tomography . . . . . 13
2.2.2 Positron Emission Tomography . . . . . 15
2.2.3 Single Photon Emission Computed Tomography . . . . . 16
2.2.4 Magnetic Resonance Imaging . . . . . 17
2.2.5 Noble Gas Enhanced MRI . . . . . 18
2.2.6 Discussion of Lung Imaging Techniques . . . . . 21
2.3 Lung Image Registration . . . . . 21
2.3.1 Challenges in Lung Registration . . . . . 23
2.3.2 Similarity Measure . . . . . 25
2.3.3 Global Registration . . . . . 28
2.3.4 Non-linear Registration . . . . . 29
2.3.4.1 CT to CT lung registration . . . . . 29
2.3.4.2 Multi-modal lung image registration . . . . . 36
2.3.4.3 Motion models in application to pulmonary image registration . . . . . 39

2.3.5 Lung Image Registration Evaluation . . . 40
2.4 Application of Lung Image Registration . . . 42
2.4.1 Adaptive Radiotherapy . . . 42
2.4.2 Ventilation Estimation from 4DCT . . . 43
2.4.2.1 Intensity/density change ventilation estimation . . . 44
2.4.2.2 Jacobian-based . . . 45
2.4.2.3 Air/tissue fraction averaging . . . 45
2.5 Conclusions . . . 47

**3 Supervoxels for Deformable Image Registration 48**
3.1 Background . . . 50
3.1.1 Image Clustering . . . 50
3.1.2 MRFs and Graph Cuts . . . 51
3.2 Methods . . . 55
3.2.1 Energy Formulation . . . 55
3.2.2 Guided Image Filtering . . . 59
3.2.3 Symmetric Registration . . . 61
3.3 Experiments and Results . . . 62
3.3.1 Materials . . . 62
3.3.2 Parameter Setting . . . 63
3.3.3 Influence of Image Representation . . . 65
3.3.4 Impact of Filtering Method . . . 69
3.3.5 Graph Relaxation . . . 70
3.3.6 Sliding Motion . . . 74
3.3.7 Comparison with Other Methods . . . 76
3.4 Discussion . . . 78

**4 Patch-based Lung Ventilation Estimation 83**
4.1 Background . . . 84
4.1.1 Lung Ventilation Imaging . . . 86
4.1.2 Lung Image Registration . . . 87
4.2 Methods . . . 88
4.2.1 Supervoxel Tracking for Ventilation Estimation . . . 90
4.2.2 XeMRI to CT Alignment . . . 95
4.3 Experiments and Results . . . 96
4.3.1 Materials . . . 96
4.3.2 Influence of the Supervoxel Size . . . 97

4.3.3 Influence of the Number of Layers . . . 100
4.3.4 Global Correlation . . . 101
4.3.5 Spatial Overlap Comparison . . . 105
4.4 Discussion . . . 109

**5 Towards More Accurate Hyperpolarised $^{129}$Xenon MRI Analysis 116**
5.1 Multi-modal Lung Image Registration . . . 117
5.1.1 Background . . . 118
5.1.2 Methods . . . 120
5.1.3 Experiments and Results . . . 127
5.1.4 Discussion . . . 132
5.2 Atlas-based Breathing Motion Correction . . . 136
5.2.1 Introduction . . . 136
5.2.2 Methods . . . 139
5.2.3 Results . . . 141
5.2.4 Discussion . . . 145
5.3 Summary . . . 146

**6 Conclusions 148**
6.1 Summary . . . 148
6.2 Outlook . . . 151
6.2.1 Lobar Analysis Based on DXeV . . . 151
6.2.2 CNN-based Lung Image Registration . . . 152
6.2.3 Multi-modal Lung Descriptor Learning . . . 154

**References 157**

# List of Figures


2.1 An example of lung images of the same patient acquired with different imaging modalities . . . . . 14
2.2 Sample images of a lung cancer patient with reduced pulmonary function caused by tumour and enhanced by COPD . . . . . 17
2.3 Comparison of ventilation maps for a healthy subject and a COPD subject using different contrast agents . . . . . 19
2.4 An example of a CT checker-board of the same coronal slice image pair of a patient . . . . . 22
2.5 An example of a multi-modal CT and MRI image pair . . . . . 37

3.1 An example of superpixel image representation on a lung CT slice in coronal, axial and sagittal plane . . . . . 52
3.2 Diagram presenting the workflow of the proposed method . . . . . 54
3.3 An example of the SLIC algorithm clustering of the peak inhale 2D lung CT image . . . . . 57
3.4 Magnitude of the displacement field for different image representations and filtering methods . . . . . 66
3.5 Influence of the regularisation parameter $\alpha$ on the TRE . . . . . 69
3.6 Registration results after applying the proposed method . . . . . 71
3.7 Comparison of different variants of the method based on TRE . . . . 72
3.8 Influence of the relaxation parameter $v$ on TRE . . . . . 72
3.9 Visualisation of the landmark error distribution . . . . . 74
3.10 An example of maximum shear stretch . . . . . 75
3.11 Comparison of the proposed method with the closest implementations 77

4.1 The diagram shows the work-flow of the proposed approach . . . . . 89
4.2 Examples of different lung tissue patterns . . . . . 91
4.3 Comparison of the estimated ventilation maps for Patient 1 . . . . . 93

4.4 Comparison of the estimated ventilation maps for Patient 2 . . . . . 93
4.5 Comparison of the estimated ventilation maps for Patient 3 . . . . . 94
4.6 The plot shows the influence of the size of the supervoxels on the correlation of the estimated ventilation maps . . . . . 98
4.7 The plot shows the influence of the number of supervoxel layers on the correlation of the estimated ventilation maps . . . . . 101
4.8 The figure shows correlation and Spearman correlation coefficient between the HU averaging ventilation approach and XeMRI . . . . . 106
4.9 Ventilation percentiles . . . . . 107
4.10 Dice overlap coefficient calculated for masks created based on the ventilation percentile ranges . . . . . 109

5.1 To bring XeMRI into alignment with CT, two transformations are composed . . . . . 119
5.2 Diagram presenting the workflow of the proposed method . . . . . 121
5.3 The reference CT image in the coronal view and superpixels estimated for the lungs imposed on the image . . . . . 123
5.4 Coronal view of the CT scan of Patient 2 . . . . . 129
5.5 Personalised registration correlation . . . . . 131
5.6 Regularisation influence personalized motion model . . . . . 131
5.7 Distribution of the estimated labels . . . . . 133
5.8 A coronal view of DXeV sequence for Subject 4 before motion correction is shown, with tidal breathing motion of the lungs . . . . . 137
5.9 The peak breathing phases images . . . . . 138
5.10 Diagram presenting the work-flow of the proposed framework addressing breathing motion correction of DXeV . . . . . 140
5.11 Distribution of the intensity changes for Subject 3 . . . . . 143
5.12 Distribution of the intensity changes for Subject 4 . . . . . 143
5.13 An example of fitting quadratic curve to the intensity sequences . . . 144
5.14 Magnitude of the norm of residuals after quadratic interpolation . . . 145

6.1 An example of the intensity distribution of DXeV for left lung in time 153
6.2 An example of series of cross-section in coronal view of DXeV or left lung with an estimated separation between regions with different ventilation patterns . . . . . 153
6.3 An example of the application of the CNN-based super-resolution method to lung pMRI . . . . . 155

# List of Tables


3.1 Comparison of the proposed method against regular grid image representation . . . 67
3.2 Comparison of the proposed method with its variants based on the quantification of negative determinant of the Jacobian of deformations 68
3.3 Comparison of the proposed method applied over the full graph without the relaxation criterion . . . 68

4.1 The influence of the size of supervoxels on the Spearman correlation for deeds . . . 99
4.2 The influence of the size of supervoxels on the Spearman correlation for demons . . . 99
4.3 The influence of the size of supervoxels on the Spearman correlation for BLF demons . . . 99
4.4 The influence of the size of supervoxels on the Spearman correlation for supervoxel graph cuts . . . 100
4.5 The influence of the number of layers of supervoxels on the Spearman correlation for deeds . . . 102
4.6 The influence of the number of layers of supervoxels on the Spearman correlation for demons . . . 102
4.7 The influence of the number of layers of supervoxels on the Spearman correlation for BLF demons . . . 103
4.8 The influence of the number of layers of supervoxels on the Spearman correlation for supervoxel graph cuts . . . 103
4.9 Spearman's correlation coefficients calculated for all ventilation estimation methods . . . 104

5.1 Dice overlap calculated for masks of the lungs in the ventilation images with respect to the reference lung atlas . . . 141

# Chapter 1

# Introduction

It is difficult to imagine modern medicine without medical imaging. Medical imaging allows clinicians not only to examine patient's body that would otherwise be difficult without surgical intervention, but also see things that are impossible to detect with the naked eye. Medical imaging techniques have evolved from 2D images, through 3D volumetric imaging to dynamic 4D imaging, and have become of paramount importance in health care, being essential for screening, differential diagnosis, surgery planning, assisted surgical interventions and any follow-up that is required [1, 2].

In the current digital age, an increasing amount of medical data are acquired at every step of medical care. This not only increases the chance of successful patient diagnosis, therapy and follow-up, but also introduces new challenges that must be addressed by those in the field of medical analysis. With medical imaging being more widely available, complex medical treatment frequently involves the need to acquire medical images of a patient at different points in time, using different devices and complementary imaging modality techniques. There are several challenges here. For instance, assessment of differences between such images based on visual inspection alone, may not be of sufficient accuracy to perform reliable quantitative analysis of the data, while the analysis of misaligned images may lead to misdiagnosis or false

conclusions. To make the best diagnostic use of different images of the same patient, they need to be accurately aligned. This can be achieved in an image registration step, which aims to find the optimal transformation between the images that brings them into alignment. This thesis proposes several new approaches for enhancing lung image registration and lung image analysis with potential applications to lung cancer. Most of the proposed methods apply supervoxel-based image representation. Supervoxel clustering is a method to perform over-segmentation of an image, grouping voxels in spatial proximity and having similar intensity values. Therefore, voxels from the same supervoxel usually belong to the same organ, at the same time allowing for variation between the supervoxels within the organ. Although supervoxels have previously been applied to lung image registration, the technique remains relatively unexplored. This thesis aims to lay further foundations for their applications by presenting novel deformable lung image registration frameworks and a method to estimate lung ventilation using supervoxels. The next section presents the structure of the remaining part of the thesis.

## 1.1 Scientific Contributions

The research conducted throughout this thesis resulted in a number of scientific contributions to the medical image analysis field. The main scientific contributions covered in this thesis are:

- Proposing a novel method for 3D deformable image registration using graph cuts. Through the use of supervoxels, we made the graph cuts problem solvable in an efficient way. We further introduced a graph relaxation criterion based on a relation between intensities of adjacent supervoxels, which improves the 3D deformable image registration performance for application on lung scans. In this method the estimated displacement field is converted into a continuous domain

by applying image-guided filtering. We show that it is capable of preserving well the 'sliding motion' of the lungs.

- Presenting a novel method for estimating lung ventilation from dynamic lung Computed Tomography (CT) images that combines a supervoxel-based image representation with deformations estimated during deformable image registration, performed between peak breathing phases. For this we propose a method that tracks changes of the intensity of previously extracted supervoxels. Due to the patch-based approach of our method, it may be physiologically more consistent with lung anatomy than previous methods relying on voxel-wise relationships. Compared to other methods used in the field, in ours the ventilation is estimated for supervoxels, which tends to group voxels that have similar intensity values and are in close spatial proximity.

- An introduction of a personalised motion model-based multi-modal image registration method. This method is dedicated to facilitate image registration between hyperpolarised $^{129}$Xenon MRI (XeMRI) and CT, where pMRI to CT registration is used as an intermediate step. The work addresses the problem of insufficient information inside the lungs in proton density Magnetic Resonance Imaging (pMRI) by applying a personalised statistical lung motion model. The motion model is estimated using Principal Component Analysis applied to the results of intra-patient image registration of 4D Computed Tomography (4DCT) data. The framework compensates for local variations between images in the motion model parameters optimisation step, where for every supervoxel an optimal set of motion model parameters are estimated.

- A presentation of a framework for breathing motion correction in Dynamic Hyperpolarised $^{129}$Xenon MRI (DXeV). The analysis of such data without breathing motion correction might result in spatial misalignment of the subsequent

ventilation volumes. The proposed framework addresses this issue by constructing a CT-based lung atlas, which helps to bring all ventilation time frame volumes into alignment.

## 1.2 Thesis Outline

This thesis is structured as follows: this chapter presents an overview of this thesis and highlights the scientific contributions. Chapter 2 first introduces the clinical problem of lung cancer, giving a brief review of the types of lung cancer. We present the available diagnostic tools and treatment options, focusing on the applications of the imaging methods. This is followed by a review of the most common imaging techniques applied for detecting, classifying, planning and assisting therapy, as well as used for follow-up of lung cancer. Subsequently, a literature review on image registration methods applicable to these imaging techniques is given. Finally the chapter presents the existing methods used for estimating lung ventilation surrogates based on 4DCT.

Chapter 3 presents the first novel contribution of the thesis, introducing the supervoxel-based image registration technique that applies graph cuts during the optimisation step and guided image filtering to convert the estimated sparse displacement field into a continuous domain. The primary application of this method is for CT-CT registration, in particular for bringing lung 4DCT images into alignment. The proposed approach overcomes some of the limitations of previous approaches, which applied the graph cuts method to deformable registration. This is conducted by posing the problem on a graph created by adjacent supervoxels, reducing the quantity of nodes required to the number of supervoxels, instead of voxels, present in the image. We demonstrate how a supervoxel image representation combined with graph cuts-based optimisation can be applied to 3D data. We further show that the application

of a relaxed graph representation of the image, followed by guided image filtering over the estimated deformation field, allows us to model 'sliding motion'.

The deformation fields acquired in the process of lung 4DCT image registration can be further applied to estimate lung ventilation maps. In Chapter 4, we present a novel approach for regional lung ventilation estimation from dynamic lung CT imaging. Our method combines a supervoxel-based image representation with deformable image registration, performed between peak breathing phases, during which we track changes in intensity of previously extracted supervoxels for a number of layers. Such a region-oriented approach is expected to be more physiologically consistent with the anatomy of the lung than previous methods that rely on voxel-wise relationships, as this approach has the potential to mimic the lung physiology. Our results are compared with static ventilation images acquired from XeMRI. We bring the CT-based ventilation maps and XeMRI images into alignment using a dedicated image registration framework, where we combine a number of affine and deformable registration steps. In our study we use three patient datasets consisting of 4DCT pMRI and XeMRI.

Chapter 5 presents image registration applications dedicated to XeMRI. The first application is our initial work on a novel method for multi-modal lung image registration, constrained by a motion model derived from lung 4DCT. The motion model is estimated using Principal Component Analysis on the results of intra-patient image registration. The approach with *a prior* motion model is particularly important for regions where there is not enough information to reliably drive the registration process, as in the case of XeMRI and pMRI to CT registration. Simultaneously, the method addresses local variations between images during motion model parameters optimisation step performed in a supervoxel-based manner. Additionally, a method for breathing-related motion correction of dynamic XeMRI sequences is presented. This method applies an atlas-based approach, with the atlas defining the shape of

the lung, to align all ventilation images with a reference frame. Such a method for the breathing motion correction will facilitate the analysis of the temporal lung ventilation.

Chapter 6 summarises all our contributions and discusses future work. We propose a potential extension of the work and new directions which can be explored. We present how the approach from Sec. 5.2, where we introduce a supervoxel-based analysis of the temporal XeMRI signal, can be used to perform anatomic lung segmentation. Furthermore, we discuss the application of convolutional neural networks (CNNs) for deformable lung image registration and multi-modal self learned descriptors to enhance such procedures.

# Chapter 2

# Background on Lung Imaging

This chapter begins from a brief introduction of lung pathologies, focusing, in particular, on lung cancer in Sec. 2.1. This is followed by a discussion about the most commonly used diagnostic tools in the clinic, and available lung cancer treatment options. In Sec. 2.2 we present routinely used imaging techniques for this application, as well as some that are still in the research stage. Subsequently, methods for image registration with application to the lungs are discussed in Sec. 2.3 and their application to ventilation estimation are described in Sec. 2.4. This chapter is concluded in the final Sec. 2.5.

## 2.1 Lung Cancer

Pulmonary-related diseases are among the leading causes of death worldwide [3] with lung cancer being the most common cause of cancer death and the second most commonly diagnosed cancer type in both male and female populations [4]. Each year there are 46,000 new cases of lung cancer diagnosed in the UK with a mortality of over 35,000 per year [5]. There are two main types of lung cancer: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC), which represent, respectively, 12% and 87% of the lung cancer cases in the UK [5]. Even though NSCLC can be

further divided into subgroups, these are all treated in the same way and in general present better prognosis than SCLC [6]. This is primarily due to the higher growth fraction and earlier development of metastases observed in cases of SCLC. Cancer that originates from the lungs tissue is called primary lung cancer. A neoplasm that spreads to the lungs from a primary lesion, originating elsewhere, is called a secondary lung tumour.

Different treatment strategies can be offered to patients depending on the type of cancer they have, and their general health and fitness. The main treatment options are:

- surgery
- chemotherapy
- radiotherapy
- chemoradiotherapy
- symptom control care (palliative care)

The choice of the most appropriate treatment option for an individual patient is made by a multidisciplinary team of doctors and healthcare professionals of different specialities. We discuss the available treatment options in Sec. 2.1.1.

To establish a precise diagnosis and select the most appropriate treatment plan, it is important to examine a sample of the cancer tissue. Samples are collected using one of two ways: during bronchoscopy, if the tumour is located inside the airways, or by performing a separate biopsy. A needle biopsy through the skin is the preferable option for taking the tumour samples. When the latter is not possible, a surgical biopsy is performed, either as a keyhole procedure or during open surgery. The samples are assessed by a pathologist who classifies the cancer.

There are other important lung diseases. Among these, Chronic Obstructive Pulmonary Disease (COPD) commonly coexists with lung cancer. COPD is an umbrella term that encompasses multiple diseases, with chronic bronchitis and emphysema at both ends [7, 8]. COPD, as well as asthma, are lung obstructive diseases and can be characterized as disorders where the air flow in the airways is restricted or even fully closed, which subsequently affects the ability to breathe, thereby reducing lung function. This kind of disorder cannot be compensated by increasing the patient's breathing effort. While asthma is a temporary state of the reduction in the airways airflow, COPD is a permanent and irreversible condition. The main focus of this thesis, however, will be in addressing challenges in lung imaging with application to lung cancer. Following sections discuss lung cancer treatment options and the available lung function evaluation techniques.

### 2.1.1 Lung Cancer Treatment

The treatment option offered to a patient depends on a number of factors, such as cancer type, disease stage, and the patient's general condition [5, 9, 10].

In the case of NSCLC, the first line treatment option to be considered is surgery. The proximity of the cancer to the other organs, for example the location of the tumour near the heart, main airways and vessels or oesophagus may make such a strategy less preferable or even inadvisable. Typically, when SCLCs are diagnosed, the cancer has already spread elsewhere, making surgery an unsuitable option as complete tumour removal is no longer possible. When surgery is performed, however, the procedure can be classified according to how much lung tissue is removed. In lobectomy just one lobe is excised, whereas bilobectomy involves removal of two lobes. If there is a need to remove the whole lung, then the procedure is called pneumonectomy. This may be necessary if cancer that was originally located in the central parts of the lungs has spread to both lobes of the left lung or all three on

the right. If the cancer is diagnosed early and the tumour precisely localised, it is possible to perform a resection of an even smaller part of the lung than a lobe. Most of the lung surgeries are performed as open surgeries, through a full thickness chest wall incision - thoracotomy. Recent advances in the operation techniques have made keyhole surgery a more common alternative than before. However, this approach is still used only in a minority of procedures, in practice.

As complementary treatment the doctors may prescribe anti-cancer (cytotoxic) drugs. This is known as chemotherapy and works by disrupting the growth of cancer cells. The drugs are administered intravenously, i.e. directly to the blood stream, either through a peripheral cannula in a patient's arm or through a central line in a patient's chest. There are different drugs prescribed depending on the cancer type. SCLC usually responds well to chemotherapy, which not only works on the tumours inside the lungs, but also kills the cancer cells which have already spread through the blood stream to other organs. This treatment might also be applied before the surgery to shrink the tumour, making it easier to remove. Chemotherapy is systemic, which means that it affects the whole organism and is not free from side effects. The side effects depend on the medications received and individual response to the treatment. The chemotherapeutics tend to cause damage especially to the tissues that undergo intensive cell metabolism, such as the hair follicles, mucous membranes and bone marrow, resulting in hair loss, sickness, loss of appetite and lower resistance to infections. Due to these side effects, the use of chemotherapy may be limited in patients who are generally unwell and already suffer from other chronic medical conditions. Chemotherapy can be used on its own, before, after or alongside radiotherapy.

Radiotherapy, which involves using X-rays or proton beams to target the regions where cancer cells reside, is preceded by a planning stage. This requires imaging data acquisition and delineation of the regions to be treated and spared. The imaging data involve acquisition of high resolution Computed Tomography (CT), Positron Emission

Tomography (PET) /CT and dynamic lung CT, when it is available. The treatment depends on the type of cancer, its stage and localisation. After complete resection of early stage NSCLC with no cancer cells in the lymph nodes, radiotherapy usually is not needed. For SCLC, radiotherapy is typically used during the early stages of the disease as a complementary treatment to chemotherapy - this is known as concurrent chemoradiotherapy. Similarly to chemotherapy, radiotherapy can also be applied before the surgery to downsize the tumour. Since SCLC can metastatize to the brain, preventative radiation may be applied to the brain; this is called prophylactic cranial irradiation. Radiotherapy is not free from side effects. In addition to the effects observed after chemotherapy, in the irradiated areas patients may also suffer from skin problems, such as local redness, itchiness or change in skin colour. Other side-effects include a sore mouth or throat, leading to problems during swallowing, an intense cough, shortness of breath and hair loss on the neck or chest. Accurate therapy planning plays a crucial role in both successfully treating the tumour and minimising the side effects. Particularly in thoracic radiotherapy, one of the major risks is radiation-induced lung damage (RILD), which can be limited with accurate planning and delivery of the treatment [11, 12]. The longitudinal analysis of a recently conducted trial, IDEAL CRT [13], showed that all of 27 patients, which were not excluded from the study due to partial lung collapse and abscess, from a cohort of 33 participating in the study undergoing radiotherapy, suffered from RILD to some extend [14].

The discussed methods might not be applicable or effective for advanced lung cancer. In cases of disseminated disease, treatment aims to improve the patients' comfort and provide them with relief in their symptoms; this is known as palliative care. Pain control is then one of the main priorities. There are procedures that can help with breathlessness caused by fluid around lungs or by airway blockage. Radiotherapy can also be an element of palliative care. For example, it may be

delivered locally to reduce the size of tumour and unblock the airways or oesophagus.

### 2.1.2 Lung Function Evaluation Techniques

Quantifying lung dysfunction is important in disease diagnosis and therapy monitoring. The gold standard test for pulmonary function is the spirometry technique, which evaluates global lung function by measuring the volume and speed of air during inhalation and exhalation. Among the few parameters that can be measured, the most useful ones are the forced vital capacity (FVC) and forced expiratory volume in 1 second (FEV1). Although these methods are cheap, fast and widely used as a screening tool [15], they are inadequate for detecting local irregularities due to the following limitations: firstly, only global measurements of lung function are obtained, which do not give any spatial information about the lobes; and secondly, they fail to detect the mild and early stages of respiratory defects. Another limitation of the method is a problem with its reproducibility, especially in the case of inexperienced patients.

Another way to evaluate lung function is to perform the Cardiopulmonary Exercise Test (CPET), which is a highly sensitive, non-invasive stress test. The stress factor is the physical exercise performed during the test, such as cycling on an ergometer. Lung diseases as well as those related to the heart or muscles will limit the patient's performance. A CPET assesses how well the heart, lungs, and muscles are working both individually and in unison. The test is performed under the control of spirometry. The data collected during CPET provide essential information to tailor exercise prescriptions in pulmonary rehabilitation (percentage of maximal workload), to evaluate the responses to intervention, aid in the construction of prediction rules for defining risk, and for preoperative selection of potentially high-risk patients for cancer resection or lung volume resection surgery [16]. Both methods suffer from the same limitation: they are capable of providing information about the general func-

tion of the lungs, but lack meaningful local evaluation. Therefore, they are neither suitable for guiding more accurate radiotherapy for lung cancer patients, which could spare well functioning parts of the lungs, nor can they be used as a tool to evaluate the progress of the disease.

Because of these limitations, as well as the high death toll and lack of a meaningful change in lung cancer survival in the last 30 years, the high level of new lung cancer cases and expected increase of new COPD cases [17], a reliable tool in the form of local quantitative lung analysis is necessary for early diagnosis and therapy monitoring. Dedicated lung imaging techniques, capable of combining structural and functional information about the lungs, can play a crucial role here. We discuss lung imaging techniques in the following section.

## 2.2 Lung Imaging Techniques

There are a number of complementary imaging techniques that can be used for lung imaging, an example of which is shown in Figure 2.1. This subsection will briefly present the most relevant structural (anatomical) and functional imaging methods applicable for lung imaging.

### 2.2.1 X-ray and Computed Tomography

2D lung X-ray imaging is the most commonly used, inexpensive and widely available lung imaging technique. The technique, however, does not provide spatial depth information sufficient for an unequivocal cancer diagnosis or radiotherapy planning. An extension of this imaging technique is computed tomography (CT) [18], where volumetric 3D scans are obtained by composing attenuation measurements acquired from different angles during X-ray tube rotation around a patient lying on a table in a gantry. Changes in the attenuation of the emitted signal correspond to different

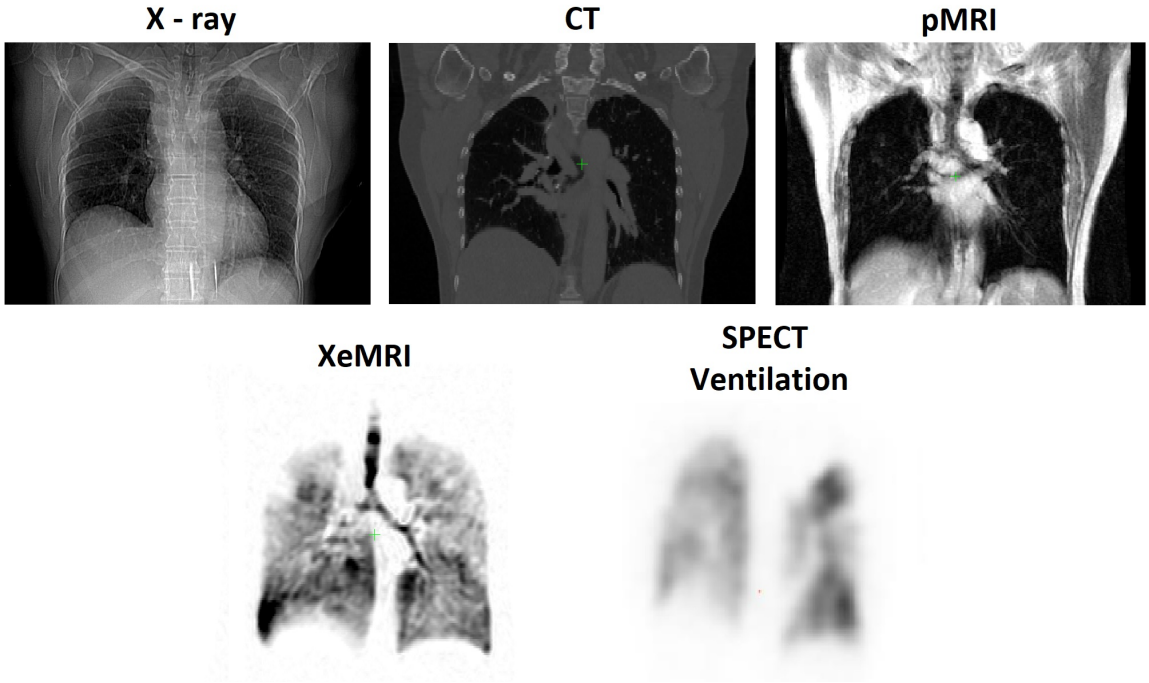



Figure 2.1: An example of lung images of the same patient acquired with different imaging modalities, shown in coronal view. In the upper row structural imaging methods are presented with X-Ray, Computed Tomography, and Magnetic Resonance Imaging (MRI). In the lower $^{129}$Xenon MRI (XeMRI) and $^{99m}$Tc-DTPA Single Photon Emission Tomography ventilation images, which show the distribution of the contrast agents ($^{129}$Xe and $^{99m}$Tc-DTPA) in the lungs. In XeMRI the distribution of hyperpolarised $^{129}$Xenon can be seen during a single breathold, whereas in SPECT it is over a longer period of time. Where applicable, similar location of the slice was presented. All the images are before bringing them into alignment, showing different location of the patient in the scanners as well as inconsistent levels of inhalations, challenges which are addressed by image registration.

types of tissue, due to the differences in their densities. The advantages of CT are its high resolution, wide availability, low scanning costs compared to other volumetric imaging methods and the speed of acquisition, which lead to low respiratory artefacts. Contrast-enhanced CT (CE-CT) is used for finding pathological changes by enhancing intensity differences between healthy structures and lesions, due to their differences in contrast uptake. Lesions usually differ in vessel distribution and blood supply compared to their surrounding tissue, which might make them more visible in CE-CT images. There might be different contrast agents used in CT imaging de-

pending on the application. Iodine-based CT contrast agents are the most commonly used. The agent provides extra attenuation which results in high intensity values in images. The main drawback is toxicity of iodine concentrations used for imaging, which limits its use in some clinical practices. The other group is gadolinium-based contrast agents. Due to their unique magnetic properties, gadolinium-based contrast agents are more commonly used in MRI. However, they can be successfully applied in CT imaging of the cardiovascular system or for pulmonary and aortic angiography. Due to its high density and a high atomic number, gold nanoparticles are an ideal radiopaque nanoparticulate contrast. In result it possess favorable X-ray attenuating properties and provides superior (about 2.7 times greater) contrast per unit weight compared to iodine [19]. The tumour microvasculature can be imaged using dynamic contrast-enhanced CT (DCE-CT) [20]. This results in a linear relationship between the concentration of contrast agent and signal attenuation. Another extension to CT is to add a second tube operating at a different voltage (or to use a single tube with the voltage being switched) to achieve Dual Energy CT (DE-CT) [21, 22]. In this application, aerosol inhaled Xenon can be used as a contrast agent to detect ventilation abnormalities through imaging of gas absorption. Dynamic lung CT, also known as 4DCT, provides a set of volumes acquired over a breathing cycle under respiratory monitoring. It is commonly used for radiotherapy planning, showing how the organs behave at different lung inflation levels. Although the structural information provided by CT is very useful, its limited functional imaging capabilities and ionising nature require the use of complementary imaging modalities for lung assessment.

### 2.2.2 Positron Emission Tomography

Positron emission tomography (PET) is a functional imaging technique. The images are based on the detection of photons originating from the annihilation of positrons emitted from a tracer agent with electrons from body tissues. The PET images are

acquired after injection of a radioactive tracer, such as flurodeoxyglucose (FDG) [23]. The lung cancer diagnosis and staging protocol include the acquisition of diagnostic CT volumes of the lung, followed by whole body PET. CT images are acquired with one deep breath hold, whereas PET images require shallow breathing during the acquisition time, which creates artefacts caused by breathing motion. Among the main limitations of the method are long acquisition times for PET (in the case of FDG, approximately 30 minutes), problems related to different respiratory patterns during imaging, along with the radioactive nature of the tracer and its high cost.

### 2.2.3 Single Photon Emission Computed Tomography

Another approach to functional imaging is Single Photon Emission Computed Tomography (SPECT), which is based on detecting photons emitted from a radioactive tracer in the human body. This method enables the acquisition of ventilation maps (most commonly after inhalation of aerosol $^{99m}Tc - DTPA$ tracer agent) and perfusion images (usually through injection of $^{99m}Tc - MMA$ tracer agent). Although lung ventilation maps are highly desirable, the long acquisition time in SPECT (as well as in PET) results in their low resolution. This, together with the radioactive nature of the tracers, are the main disadvantages of the method. An additional challenge is dealing with artefacts caused by the reconstruction methods of SPECT and PET images [24]. The majority of SPECT and PET scanners are integrated with CT, which enables combining functional and structural information. The CT scan is acquired under free breathing protocol and used for attenuation correction (AC-CT) in image reconstruction process. The AC-CT scan may be used as an intermediate step in PET to diagnostic CT registration, which can facilitate a diagnosis process [25]. An example of SPECT ventilation images and their overlay on CT is shown in Figure 2.2.

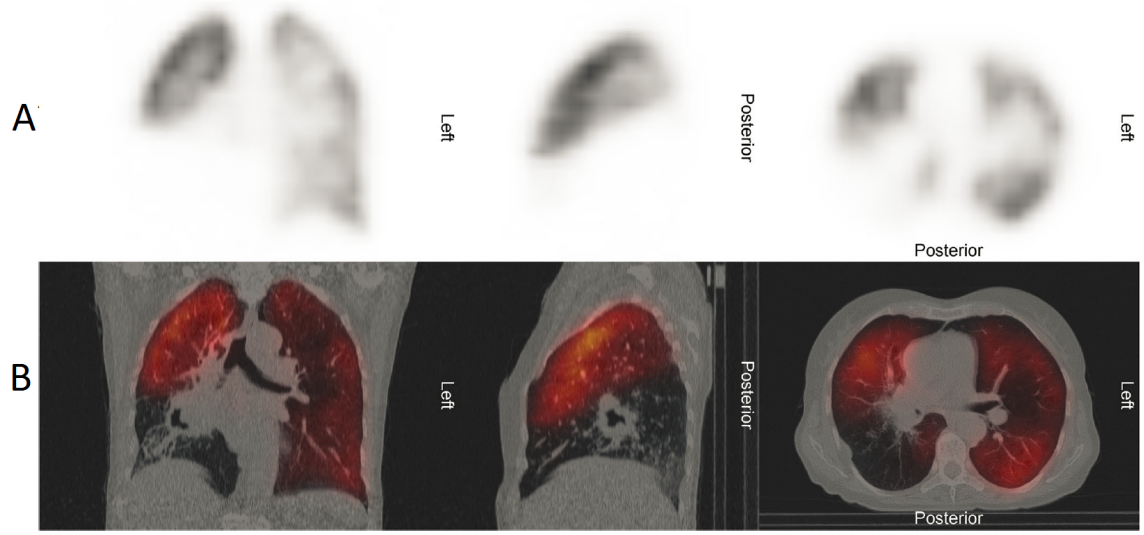



Figure 2.2: Sample images of a lung cancer patient with reduced pulmonary function caused by tumour and enhanced by COPD, in three views [26]. (A), perfusion SPECT images are shown, (B) fusion of SPECT and CT. The red colour in row B stays for the SPECT images. A perfusion defect is visible in the lower right lobe.

### 2.2.4 Magnetic Resonance Imaging

In contrast to the previously described imaging techniques, magnetic resonance imaging (MRI) is a non-ionising method [27]. The patient is placed in an external stationary magnetic field, which aligns the magnetic moment of protons (spin) in the same direction as the field. Then additional short radio frequency (RF) pulses are applied in a direction perpendicular to the field, which results in a flipping of the net alignment by a certain angle (flip angle). After switching off the RF pulse, the spinning protons return to the equilibrium state and realign with the external magnetic field. The body tissue can be characterised by two constant relaxation times: T1 and T2. During the imaging magnetisation in longitudinal (z-direction) and transverse (xy-direction) is measured, based on which T1 and T2 are calculated. The images are reconstructed as a result of calculations of the T1 and T2 times for protons in different tissues. When time to repetition (TR) is very long and time to echo (TE) very short, the MR signal is directly proportional to the unweighted number of spinning protons present in the scanned volume. Images acquired with such a technique

are minimally weighted by the relaxation times and are, therefore, termed proton density weighted MRI (pMRI). This imaging protocol will be of particular interest in this thesis. Previously built PET scanners were integrated with CT scanners. The latest hybrid PET scanners, however, provided by Siemens, and more recently by GE are available also with MRI scanners. Just as in CT, MRI can also be performed with contrast agents (usually based on gadolinium, due to its strong paramagnetic properties). Recently dynamic contrast-enhanced MRI (DCE-MRI), where a set of images is acquired after injection of a gadolinium-based contrast agent, has been used more frequently to extract information about the functional activity of tumours [20]. Compared to DCE-CT, the changes in the contrast intensity for DCE-MRI are not linearly related to its concentration in the tissue. A novel lung ventilation and perfusion imaging technique based on Fourier decomposition MRI has been introduced [28]. The method relies on a single acquisition of series of data and does not require any contrast agent. The main limitation of this method is its ability to provide only 2D images. Even after recently proposed extensions of the method [29], the approach still has not been widely applied.

### 2.2.5 Noble Gas Enhanced MRI

Due to low proton density in the lungs and artefacts at air-tissue interface, pMRI results in worse contrast, particularly inside the lungs compared to CT. As with CT lung imaging, pMRI does not directly provide functional information about the lungs. To overcome this limitation the noble gases $^{3}$Helium and $^{129}$Xenon were proposed as inhalable contrast agents [30], and their application to human lungs has been presented [31]. The inhalation of these noble gases allows one to directly visualise airspaces in the lungs in MRI. The lower density of the gas, compared to body tissue, is compensated for by increasing the polarisation level, for instance by an optical pumping and spin exchange hyperpolarisation process [32]. In a routine polarisation

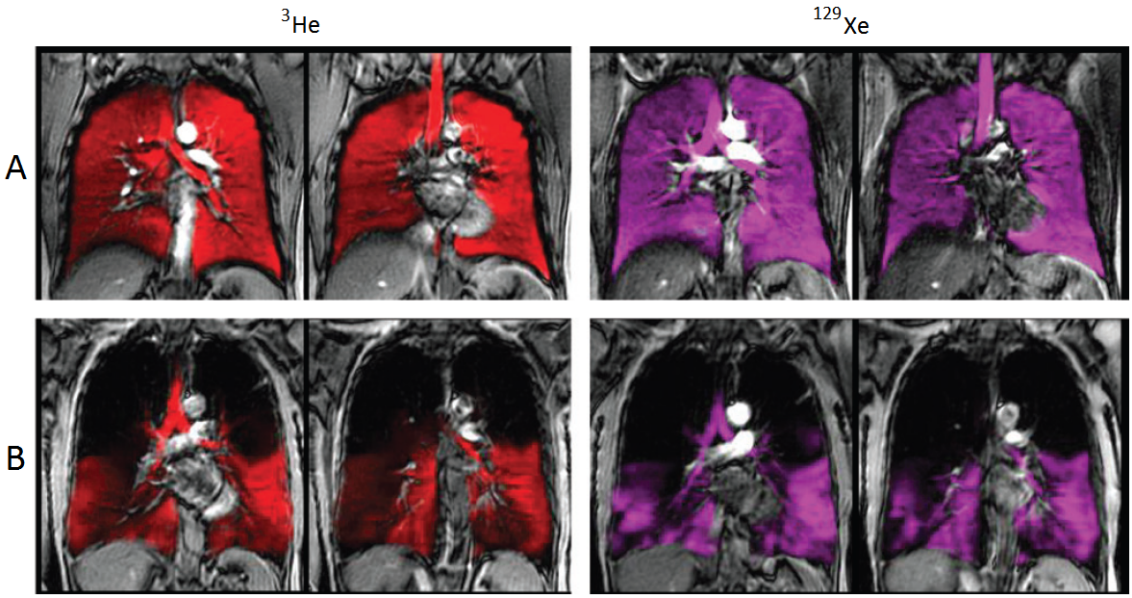



Figure 2.3: Comparison of ventilation maps for a healthy subject (A) and a COPD subject (B) using different contrast agents in coronal view [29]. The red colour represents $^{3}$He and purple $^{129}$Xe. For each subject two slices with HeMRI are shown and corresponding to them XeMRI images. Complementing the ventilation maps proton MRI images are displayed as a background. In the figure a uniform distribution of both hyperpolarised contrast agents can be seen for the healthy subject. In the case of the COPD subject, severe ventilation defects localised in the upper parts of the lungs can be observed for $^{3}$He as well as for $^{129}$Xe. Minor differences between the ventilation images for COPD subjects with different contrast agents are mainly due to differences in their physical properties.

procedure (polarisation of one litre of gas in one hour) the required increase is around 10% for $^{129}$Xe and 30% for $^{3}$He, which amounts to a 10000-fold increase in the nuclear spin polarization of a material in a magnetic field [33]. In noble gas-enhanced MRI images, only the gas is visible and lung surfaces or internal structures cannot be directly delineated. Thus, to establish a correspondence between the ventilation maps with physical structures, two scans are required - one acquired with a noble gas contrast agent, called either XeMRI or HeMRI, depending on the contrast agent, and the other using pMRI. Most of the initial studies in the noble gas-enhanced MRI field [34, 35, 36] were conducted using $^{3}$He, mainly due to its larger nuclear magnetic moment and higher level of polarisation compared with $^{129}$Xe. The large

increase of $^3$He costs [33], lack of natural resources and limited access to the existing sources of the gas, forced researchers to turn to $^{129}$Xe. Its abundant nature and high rate of solubility (10 times higher than $^3$He) make $^{129}$Xe a cheaper and more widely available alternative to $^3$He. An example of the comparison between HeMRI and XeMRI imaging is presented in Figure 2.3, from [37], where more details can be found. The study shows a slightly higher ventilation defect percentage rate for $^{129}$Xe, where this is calculated as the fraction of the gas signal intensity below a slice-specific threshold [38]. The reason for such results stems form the different physical properties of the two gases. $^{129}$Xe atoms are larger and heavier than $^3$He, which may make $^{129}$Xe more sensitive to the detection of mild obstruction and more helpful for the diagnosis of diseases in early stages. Because of the physical properties of $^{129}$Xe, breathing in one litre of the gas does not directly correspond to breathing one litre of air, as it is in pMRI imaging; instead, more effort is required during breathing. Unlike breathing air and subsequently holding one's breath, a similar procedure with $^{129}$Xe is hard to train without the gas and can affect the breath-hold manoeuvre. These variations can cause misalignments between XeMRI and pMRI, which may require compensation using image registration. As an extension to the static XeMRI lung ventilation imaging, dynamic XeMRI has been recently proposed [39]. In this method several volumetric images are acquired over a number of breathing cycles in a single imaging session. Hyperpolarised gas MRI is the safest functional lung imaging technique from all the reviewed methods, with no serious side effects being found so far [40], and thus is a very promising tool for accurate lung dysfunction diagnosis and therapy control. Currently, the main drawback of the technique is its accessibility, which is limited to a small number of imaging centers in the world.

### 2.2.6 Discussion of Lung Imaging Techniques

No imaging technique provides complete functional and structural information about the lungs. Fusion of different types of images can potentially facilitate diagnosis and treatment planning. There are a number of problems which need to be overcome in this task. The first issue is to establish correspondence between images acquired with different devices and at different points in time. When comparing CT images from the same patient taken at different time points, the intensity of corresponding lung structures vary because the volume of the lungs changes during breathing resulting in an altered density profile [41]. In the case of CT imaging, intensity attenuation is related to tissue density, whereas for MRI, the relationship is not obvious. CT scans are acquired during a single deep inhalation. During pMRI/XeMRI imaging, however, one litre of gas is usually inhaled, which results in the diaphragm being located at different positions than in CT. Moreover, the level of lung compression inhomogeneously increases from the apex downwards. During breathing, sliding motion occurs between the rib cage, lungs and diaphragm [42, 43, 44], and small features experience displacements that can be several times larger than their size. All this requires motion correction to achieve correspondence between images. The following section reviews the state-of-the-art components applied during this task.

## 2.3 Lung Image Registration

To take full advantage of the previously discussed imaging modalities, it is necessary to find spatial correspondence between the acquired images. This allows us to assess the same structures in two images that have been acquired from different imaging modalities or at different points in time. An example of such a case, where CT lung images where acquired at peak inhale and peak exhale is presented in Figure 2.4, showing high level of misalignment between the images. For radical lung cancer treat-

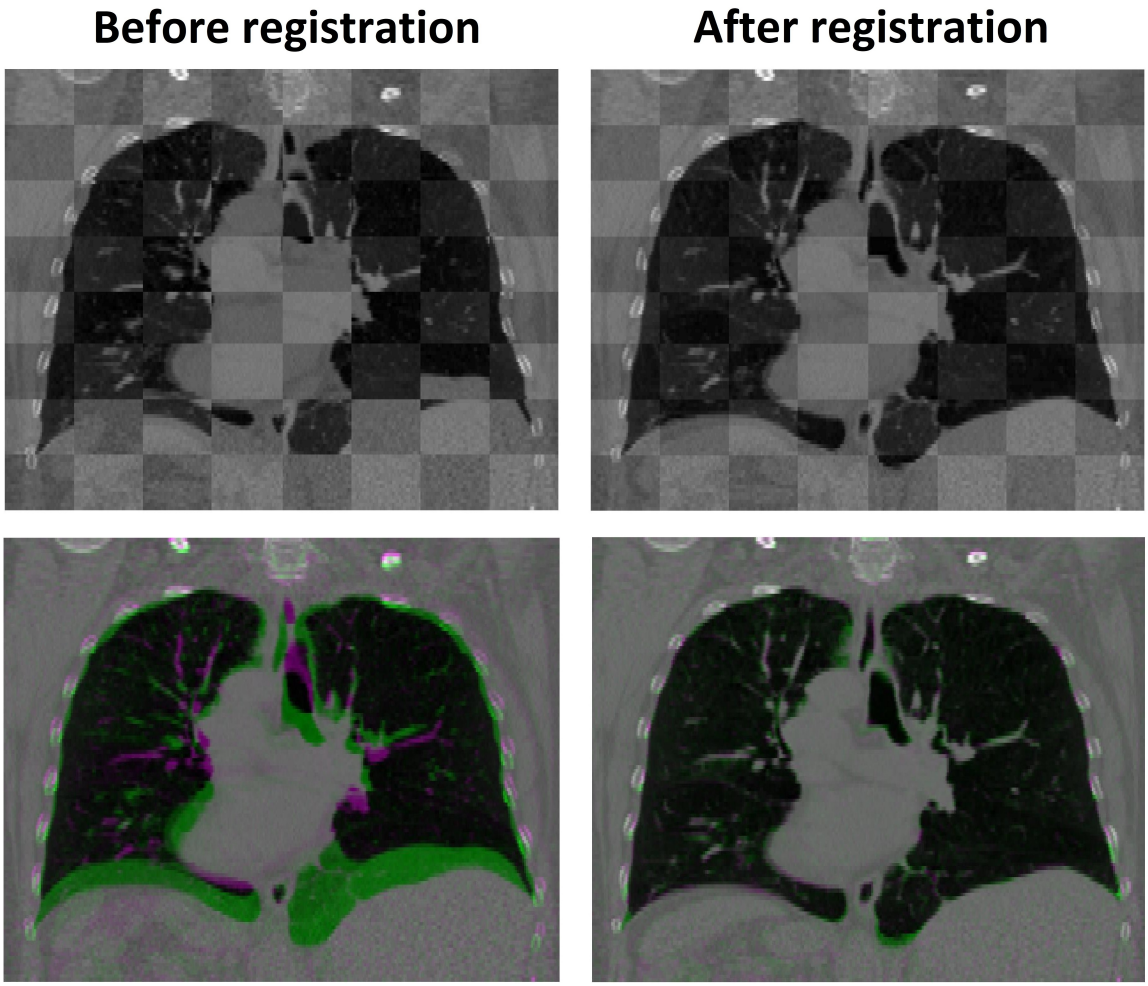



Figure 2.4: An example of a CT checker-board of the same coronal slice image pair of a patient. The images show lungs at peak inhale and peak exhale inhalation levels. In the lower row an overlay of peak inhale (magenta) and peak exhale (green) is presented. In the left, images before registration are presented, showing high level of misalignment between them. In the right, the results are presented after application of a deformable registration method, which compensated for most of the misalignments.

ment set-up errors have been reported between 1-5mm [45]. However, the advances in the therapy delivery, which can take into account patient's breathing motion over the course of treatment delivery, as well as the application of proton therapy to lung cancer treatment can potentially benefit from more accurate radiotherapy delivery.

The process of finding an optimal transformation between different images is called image registration and is one of the fundamental tasks in medical image processing. It remains an active and still challenging research field [46], especially for lung imaging

applications. In the general case of two images, one of them is considered to be fixed ($I_f$) and another moving - ($I_m$), which is transformed towards the first one. This process can be described as an energy, $E$, minimisation procedure of a cost function based on a similarity measure of the images, $\mathrm{Sim}(I_\mathrm{f}, I_\mathrm{m}(\mathbf{u}(\mathbf{x})))$, and $\alpha$-weighted regularisation, $(\mathrm{Reg}(\mathbf{u}))$, of the estimated displacement field,

$$\arg\min_{\mathbf{u}}(E(\mathbf{u})) = \mathrm{Sim}\Big(I_\mathrm{f}, I_\mathrm{m}\big(\mathbf{u}(\mathbf{x})\big)\Big) + \alpha\mathrm{Reg}\big(\mathbf{u}(\mathbf{x})\big) \tag{2.1}$$

where $I_\mathrm{m}(\mathbf{u}(\mathbf{x}))$ stands for $I_m$ warped toward $I_f$ by the displacement $\mathbf{u}$ corresponding to location $\mathbf{x}$. The similarity term measures the distance between images, whereas the regularisation term enforces estimation of plausible deformations. The registration process is usually divided into two steps: the first is a coarse linear alignment and the second is a non-linear registration. There are many design choices to be made before starting the registration process, such as which similarity measure to use, which transformation model to apply and how to perform the optimisation procedure.

### 2.3.1 Challenges in Lung Registration

There are a number of remaining challenges in lung registration, which make the process an extremely ambitious task. First is the difference in the volume of the lungs during imaging. In the case of 4DCT or full/peak inhale and full/peak exhale imaging, the difference in the lung volume and its position have an important clinical meaning. They might be, for instance, used for therapy planning. In other cases, however, it may be difficult to acquire two subsequent scans where the lungs are at exactly the same level of inflation. Inflated lungs increase their volumes, which results in a change in their density and therefore the observed intensity in CT. The change in the lung volume is not homogeneous across the whole volume in healthy subjects and is even more heterogeneous for patients suffering from lung/breathing disorders.

In a set of healthy lungs, the compression levels increase from the apex downwards towards the diaphragm, but when the lungs are damaged, this may not be the case. Therefore a simple global compensation for the change of the lung volume, based, for instance, on a calculation of the change in the segmented lung volume, is not applicable, as the changes are local in nature.

A second problem arises from the relative motion between the lung and surrounding tissues and is usually referred to as 'sliding motion'. This occurs, in particular, on the surface of the lung pleura between the lungs, rib cage and diaphragm. It is characterised by discontinuities in the tangential components of the displacement field along the direction normal to the lung surface, while the normal components remain smooth. From a mathematical point of view, this requires preservation of the physical properties of the tissue inside a lung lobe, where folding and overlapping of the structures should not be observed. At the same time, discontinuities in the transformation field due to the 'sliding motion' of the lungs, ribcage and diaphragm should be allowed. One of the approaches addressing this challenge is to perform lung segmentation before the registration and restrict the registration only to the lungs. Another aspect is intra-lobar 'sliding motion' inside a lung, which, for instance, causes segmentation methods involving the whole lung to lack sufficient specificity.

Lung image registration involving pMRI leads to a range of other difficulties. The density of lung tissue is low and therefore the amount of information provided by pMRI about its structure is limited. This makes segmenting airways or lobes from pMRI challenging tasks. Additional difficulties stem from the imaging time. Although efforts are made to make the pMRI faster without loss of accuracy, there is still a trade-off between breath-hold time and the resolution/size of acquired data.

Another challenge is in CT to CT registration and is related to the acquisition itself. AC-CT acquisition, used for the PET images reconstruction, was discussed when introducing PET imaging in the previous sections. The AC-CT might be used

as structural imaging for aligning PET with other modality images. However, its low dose acquisition protocol differs from the diagnostic high resolution CT, which is performed at full inhalation. Dynamic lung CT or 4DCT is a powerful tool which allows visualising the lung across the breathing cycle. It comes however at a price: potential emergence of artefacts. The 4DCT acquisition protocol is slightly different to those of other volumetric imaging techniques, where a whole volume is acquired at every imaging session. Instead, the acquisition is performed under control of spirometry and the images are acquired as sections of a volume. For each of the sections, a full breathing cycle is imaged. After the scan, the particular volumes corresponding to breathing phases are reconstructed based on the information from spirometry. Such an approach requires a regular breathing pattern by the patients during the imaging, otherwise the scanned section might not match well with adjacent ones. Such problems are called reconstruction artefacts and unfortunately are common in this examination. These challenges render most of the single-modality similarity measures along with conventional regularisers for deformable image registration unsuited to the task.

### 2.3.2 Similarity Measure

To perform image registration, some measure of distance between the images has to be calculated using an image similarity measure. This can be done, for instance, by a direct comparison of voxel intensities, stochastic appearance of intensities or by evaluating the relationship between neighbouring voxels. The most intuitive similarity measure is the sum of squared intensity differences (SSD):

$$\mathrm{SSD} = \sum_{\mathbf{x}\in\Omega} \Big( I_f(\mathbf{x}) - I_{\mathrm{m}}\big(\mathbf{u}(\mathbf{x})\big) \Big)^2 \tag{2.2}$$

where $\Omega$ is the domain of the overlap between images, $I_f(\mathbf{x}), I_{\mathrm{m}}(\mathbf{u}(\mathbf{x}))$ are the image intensities at the spatial position $\mathbf{x}$ and $\mathbf{u}$ is the transformation applied to the moving image, $I_m$. The main limitation of this approach is the assumption that the intensity variation of corresponding structures in both images is caused only by noise. However, this is invalid even in the case of single-modal lung images. In lung CT imaging, the intensity between images may differ due to tissue density changes during inspiration and expiration.

Cross-correlation (CC) or its normalised version, NCC [47], are common similarity measures for the CT to CT lung registration task. This is mainly due to the locally linear nature of the changes in intensity between two lung images acquired at different breathing phases. This kind of regional variations can be well captured by locally calculated CC. NCC is defined as:

$$\mathrm{NCC} = \frac{\sum_{\mathbf{x}\in\Omega}\Big(I_f(\mathbf{x}) - \overline{I_f(\mathbf{x})}\Big)\Big(I_m\big(\mathbf{u}(\mathbf{x})\big) - \overline{I_m\big(\mathbf{u}(\mathbf{x})\big)}\Big)}{\sqrt{\sum_{\mathbf{x}\in\Omega}\Big(I_f(\mathbf{x}) - \overline{I_f(\mathbf{x})}\Big)^2}\sqrt{\sum_{\mathbf{x}\in\Omega}\Big(I_m\big(\mathbf{u}(\mathbf{x})\big) - \overline{I_m\big(\mathbf{u}(\mathbf{x})\big)}\Big)^2}} \tag{2.3}$$

where $\overline{I_{\mathrm{f}}}$ and $\overline{I_{\mathrm{m}}}$ are the mean intensity values in $I_{\mathrm{f}}$ and $I_{\mathrm{m}}$, respectively. Its locally calculated variant [48] provided the best and second best results in the EMPIRE10 challenge [49].

When locally normalised version of cross correlation (LNCC) is calculated, the mean values of the images are calculated for patches instead of whole images, as in Eq. 2.3. In our work we followed the implementation provided in [47]:

$$\mathrm{LNCC}(\mathbf{x}) = \frac{\Big(I_f(\mathbf{x})I_m\big(\mathbf{u}(\mathbf{x})\big)\Big) * G_k - \Big(I_f(\mathbf{x}) * G_k\Big)\Big(I_m\big(\mathbf{u}(\mathbf{x})\big) * G_k\Big)}{\sqrt{\Big[I_f(\mathbf{x})^2 * G_k - \big(I_f(\mathbf{x}) * G_k\big)^2 + \epsilon\Big]\Big[I_m\big(\mathbf{u}(\mathbf{x})\big)^2 * G_k - \big(I_m\big(\mathbf{u}(\mathbf{x})\big) * G_k\big)^2 + \epsilon\Big]}} \tag{2.4}$$

The global mean values of the images, $\overline{I_f(\mathbf{x})}$, $\overline{I_m\big(\mathbf{u}(\mathbf{x})\big)}$ are substituted with their local counterparts calculated via convolutions ( $*$) with Gaussian kernels, $G_k$ which size specifies the size of the patch for which local correlation is calculated. A small constant $\epsilon$ is added to the local variances of the images in the denominator to prevent dividing by zero.

The most common approach to multi-modal image registration (such as CT to MRI image registration) is a measure of mutual information (MI) introduced by Viola and Wells III [50] and Maes *et al.* [51], which is based on Shannon's entropy. In [52], the authors introduced a normalised MI measure as a solution to the problem of changing image overlap. Another extension of the measure, $\alpha - \mathrm{MI}$, was proposed in [53] and is based on the calculation of graphs instead of histograms. A comprehensive review of MI-based methods can be found in [54].

In addition, similarity can be described as a measure not only of the relationship between two voxel intensities, but also as the likelihood of the voxel and its surrounding region being represented by other voxels/regions in the image. The modality independent neighbourhood descriptor (MIND) [55], which is used both in single-modal (CT to CT) and multi-modal (CT to MRI) registration, is one such example. The descriptor is defined, for each voxel position $\mathbf{x_p}$ with a spatial search region $R$, by a relation between a distance $D_p$ and a variance $V$:

$$\mathrm{MIND}(I, \mathbf{x}, \mathbf{r}) = \frac{1}{\eta} \exp\left(-\frac{D_p(I, \mathbf{x}, \mathbf{x}+\mathbf{r})}{V(I, \mathbf{x})}\right) \tag{2.5}$$

with $\eta$ normalising the values to 1. The distance $D_p$ measures the intensity difference between the voxels in a defined neighbourhood $\mathbf{r}$ of a control point and creates a vector of size $\mid R \mid$:

$$D_p(I, \mathbf{x}, \mathbf{x}+\mathbf{r}) = G_k * \big(I(\mathbf{x}) - I'(\mathbf{x}+\mathbf{r})\big)^2 \tag{2.6}$$

For each squared difference between the translated image $I'(\mathbf{x}+\mathbf{r})$ by a vector $\mathbf{r}$ and the original image, there is a subsequent convolution operation with Gaussian kernel $G_k$. The variance term $V(I,\mathbf{x})$ in the denominator represents an estimation of intensity distribution in the neighbourhood of a control point $\mathbf{x_p}$ due to the enhanced MIND response to patches that are similar to the control point and its attenuation of dissimilar ones. The variance is represented as a mean of voxel differences within a six-neighbourhood $N$:

$$V(I,\mathbf{x}) = \frac{1}{6}\sum_{n\in N} D_p(I,\mathbf{x},\mathbf{x}+\mathbf{n}) \tag{2.7}$$

The main advantage of MIND is that it provides a descriptor which is sensitive to similarities in gradients and texture and insensitive to direct differences in intensities between voxels.

### 2.3.3 Global Registration

One of the simplest forms of image registration is performed using a rigid transformation model [56]. The most common model in 3D medical applications has six degrees of freedom. The matching is performed by rotations and translations of one of the images, in order to find the position at which the similarity measure takes its lowest value. Rigid registration is usually an efficient way of matching structures which do not change significantly in a short period of time, like the human skull or adult bones. In the case of the lungs, this transformation is not suitable to represent all movements and, therefore, is typically used as a first step for coarse alignment of images. It provides a good global initialisation for the subsequent non-linear registration step. Without this step, the non-linear registration process may be too far from the desired local optimum and lead to erroneous solutions.

### 2.3.4 Non-linear Registration

The most challenging type of lung registration is non-linear (also known as deformable), as this involves many more degrees of freedom than its linear counterparts. Thus, functions must be minimised over a higher-dimensional space. Relying purely on similarity measures, which match spatial locations based on the chosen measure, might result in a physiologically implausible deformations [57]. The transformation for lung registration should preserve discontinuities, modelling 'sliding motion' by allowing for discontinuities at the lung boundaries and between lobes, while providing smooth deformations in other regions. Therefore, the transformation model may be inherently regularised, whereas in other cases, an additional regularisation term $\text{Reg}(\mathbf{u}(\mathbf{x}))$, which penalises implausible deformations, is introduced to further constrain the transformation model. The exact definition of the $\text{Reg}(\mathbf{u}(\mathbf{x}))$ term depends on the application, but usually is defined in terms of derivatives of the deformations [58]. A good general review of deformable image registration methods can be found in [59]. Non-linear image registration plays a crucial role in lung image alignment. In the following subsections we discuss methods addressing problems in CT-CT and multi-modal lung image registration.

#### 2.3.4.1 CT to CT lung registration

In CT an image is created by a reconstruction of the attenuation projections of X-ray beams caused by body tissue. Registration of such images may seem at first glance to be a classic single-modality problem, where intensities of the corresponding structures are consistent across images to be registered. Because of the problems discussed above, such as variation of the lung tissue density or 'sliding motion', it is a more challenging task. Below we briefly review the three most common CT to CT groups of lung image registration approaches, namely demons, free form deformations with B-splines and discrete optimisation-based methods.

**Demons:** The demons-based method is one of the most common and widely used approaches, which represents an optimisation as a force pushing the moving image towards the target image [60]. This method is based on a diffusion process and the deformations are calculated using intensity gradients. The optimisation is performed in an iterative manner, where finding the best displacement and regularisation steps are applied successively. Gaussian smoothing is used as a simplification of diffusion regularisation. It is possible to apply an additional Gaussian smoothing-based regularisation of the estimated updates, which works as a fluid-like regularisation.

As an extension to demons, Vercauteren *et al.* proposed calculating the displacement field in a diffeomorphic transformation space by using log-demons [61, 62], which prevents folding of the deformation field and preserves one-to-one mapping. They use the scaling-and-squaring approach [63] to calculate the final displacement field. The basic concept is to first attempt to find an unconstrained deformation field, $\mathbf{c}$, by minimising a similarity measure, but at the same time compute a smooth regularised version, $\mathbf{u}$, which introduces an additional constraint in the optimisation, minimising the distance between $\mathbf{c}$ and $\mathbf{u}$:

$$\arg\min_{\mathbf{u}} \left[\mathrm{Sim}\left(I_{\mathrm{f}}, I_{\mathrm{m}}(\mathbf{u}(\mathbf{x}))\right) + \mathrm{dist}(\mathbf{u}, \mathbf{c}) + \alpha \mathrm{Reg}(\mathbf{c})\right] \tag{2.8}$$

Although this method is widespread and relatively simple in implementation, it suffers from a few limitations. For example, setting the regularisation parameter (Gaussian kernels) requires a fair amount of calibration and has a significant influence on the obtained results. The method relies on gradient information, which can make it susceptible to local minima and dependent upon the initial position of the images. To address this issue the method is usually implemented in a multilevel variant.

In [64] the authors further extended the method, introducing bilateral filtering as regularisation of the deformation field to preserve 'sliding motion'. This method

addresses both challenges in lung image registration: 'sliding motion' and the variation of the tissue density. Bilateral filtering, which is an edge preserving filtering method, helped to model 'sliding motion', whereas the Normalized Gradient Field [65] overcame the issue of lung intensity changes during breathing.

In [42], the authors proposed a method which addresses the 'sliding motion' problem. The authors call the method slipping organ motion, with the 'sliding motion' capability embedded in the regularisation part of the diffusion-like image registration [66]. The energy function is formulated in a similar way as in the Thirion's demons force [60]. The method decouples diffusion regularisation into normal and tangential directions around the lung boundaries using automatically detected masks. The method has some limitations. It requires accurate lung and liver segmentation for calculating normal and tangential direction vectors; it neglects the 'sliding motion' which may occur inside the lungs. Modelling of that motion would depend on lung segmentation into individual lobes, which is not a trivial task in general, and is even more challenging in 4DCT due to the reconstruction artefacts

**Free-form deformations:** Another common approach for deformable image registration is the application of free-from deformations (FFDs) based on B-splines [67], where the deformation is applied by moving a mesh of control points placed on a regular grid. An extended B-spline method with hierarchical layers of different resolutions, where some of the control points were not optimised, achieving overall a non-uniform distribution was proposed in [68]. One of the most commonly used implementation of deformable image registration using FFDs based on B-splined is the Elastix toolbox [69], which also addresses computational requirements by applying the adaptive stochastic gradient decent method used as an optimiser [70]. A different approach was proposed in [71], presenting an implementation of the method for graphical processing units.

The combination of 'sliding motion' and FFDs was addressed in [72], where a

patient's body was segmented into static and moving parts separated by the pleural interface. The registration was performed independently for the separated regions. This approach was further developed in [73, 74] forcing the separated regions to match, preventing discontinuities between the regions. The preservation of the 'sliding motion' is achieved by decomposition of the basis functions of B-splines into orthogonal and tangential components with respect to the nearest lung interface. In [75] the authors proposed eXtended Free-Form Deformations, a discontinuity preserving method, introducing enriched basis functions instead of conventional basis functions. The new basis functions are created by a product operation between the conventional basis functions and an enrich function, which the authors proposed to be the sign function.

To overcome the challenge of lung tissue intensity changes, mass preserving lung image registration methods have been proposed. In [41] the authors proposed a new similarity measure: the sum of squared tissue volume difference. In a similar approach, the authors of [76] modified the cost function introducing a mass preserving correction step as the multiplication of the moving image by a locally calculated determinant of the Jacobian of the deformation on a voxel-wise basis. In both approaches the authors assumed that the total lung tissue mass remains nearly constant across the breathing, but it's distribution changes during breathing. Still, in [41] the authors performed a correction for the lung tissue changes, originating mainly from changes in the blood volume. At the same time, this characteristic of the lungs opens the possibility of investigating the ventilation of the lungs based on CT. As the change in the lung attenuation is mostly caused by the change in its air filling, it may be useful to quantify this change.

**Discrete methods:** Discrete optimisation-based registration, which uses a discretised space of allowed displacements, is another interesting approach for non-linear image registration. To use a discrete solver for deformable image registration, the problem is usually formulated in terms of Markov Random Fields (MRFs). The

optimisation is posed on an undirected graph that is built by a set of nodes, $P$. MRFs create a framework where long-range relations of distant nodes in the graph are modelled via a *knock-on effect* by short-range relations of usually just a pair of neighbouring nodes. In such a form, the optimisation is the process of finding a labelling $f$ for the set $P$ that minimises the energy $E(f)$, consisting of the sum of two terms:

$$E(f) = E_{data}(f) + \alpha E_{smooth}(f), \tag{2.9}$$

where $E_{data}(f)$ represents the disagreement between the labelling $f$ and the observed data, whereas an $\alpha$-weighted piecewise smoothness term $E_{smooth}(f)$ forces nodes connected by an edge to have similar labels. Depending on the application, the labels can represent different properties, for instance objects or background in segmentation [77] or a depth map in stereo-vision [78]. When such a formulation is applied to deformable image registration, the labels to be assigned represent predefined displacement vectors $[u_x, u_y, u_z]$. For this application, only a limited number of displacements is considered, thus this formulation requires a transformation model, which transfers the displacement field from a discrete domain to a continuous one. For a 3D application the number of labels can consist of thousands of allowed displacements, which might be a challenging optimisation task.

Initially, the lack of efficient optimisation algorithms limited discrete optimisation-based methods application for deformable image registration. An approach using dynamic programming as an optimisation method, **drop** [79], has been successfully introduced for brain image registration and then applied to the lungs. The main limitation of this method is that it estimates the best displacements only in the main directions, independently of each other. It made the algorithm faster at the cost of higher susceptibility to local minima. What is more, it does so in an iterative way, which renders the method similar to the continuous-optimisation based approaches and therefore becomes potentially vulnerable to be trapped in a local minimum.

These restrictions limit the advantages of the discrete optimisation-based approach, which allows for a dense displacement sampling and for some methods to find at least a strong local minimum. In [80], the authors proposed using the random walks algorithm for deformable image registration. The major drawback of random walker method for deformable image registration [80] is that it is computationally expensive. The main bottleneck is solving the system of combinatorial Laplace equation, which is time consuming since it must be solved for each of the displacement labels. This limits its application in the fields, especially for 3D images. In [81] the optimisation problem is posed on a minimum spanning tree (MST) extracted from a graph. The image is parametrised to regular cubes represented by nodes in the graph. The method, however, does not address the 'sliding motion' problem directly. More recently, a symmetric regularised correspondence fields method was introduced in [82] for images from patients with COPD and a supervoxel-based belief propagation algorithm in [83]. For the latter, the optimisation has also been posed on a minimum spanning tree, this time created based on a supervoxel image parametrisation. However, the tree representation, because of the hierarchical nature, misses many of the connections between regions either inside or outside of the lungs. This breaks up the regularisation components that would be expected between those regions. In contrast, a graph like the one proposed in Chapter 3 of this thesis, preserves all relevant connections. Both methods proposed in [81] and [83] use the belief propagation algorithm as the optimisation method to estimate displacement field. In [84] the authors, instead of finding the optimal displacement for every voxel or a patch, perform keypoints localisation. Then the optimal displacements are evaluated in a similar manner to [81, 83], also using belief propagation on a minimum spanning tree created from the key points. This step is followed by a dense displacement registration, applying B-splines. This method addresses the 'sliding motion' problem through lung segmentation and restricts the registration only to the lung masks.

Another example of the optimisation method applicable to lung image registration is the method of graph cuts. Although the method has been widely used in many fields of computer vision, such as image segmentation [77], de-noising [85], stereo vision [86] or motion detection [87], its application to image registration has been limited. A voxel-wise deformable registration method applying graph cuts has been proposed in [88] for both 2D and 3D brain images. An obvious limitation of the method is the size of the graph on which the problem is posed. While for 2D images the problem is still tractable, in 3D cases the size of the graph with every voxel being represented by a node, along with the high number of possible displacements, make the problem challenging in terms of computation. For instance, the method presented in [89], which is an extension of the approach from [88], reports computational times for 3D cases of up to 24h. A landmark-based graph cuts algorithm was proposed for medical image registration in [90]. In that work, a penalty cost is added to the data term for a node on which a landmark is localised, forcing its alignment with the corresponding landmark in the other image. Such an approach can be seen as putting a hard constraint on the deformation field in the presence of landmarks. A registration method using mutual information as a similarity measure with graph cuts as an optimisation method was later proposed in [91] and followed by a similar approach with a prior joint intensity distribution in [92]. Similar to the previously discussed graph cuts-based image registration approaches, the main limitation of these methods is their computation complexity. As they are applicable for 2D problems, where every pixel is represented by a node, they are not suitable for 3D problems, due to memory requirements and computational time.

Recent research has seen the first attempts to use the powerful tool of CNNs [93, 94] in the application of image registration [95, 96]. CNNs have established new standards in many fields of computer vision, such as segmentation, classification, tracking or single image super-resolution and are finding new applications in med-

ical image analysis. However, in the field of medical image registration, where the ground truth is usually unknown, the networks were trained using artificially applied deformations. Such an approach has obvious limitations, as the deformations taking place in real life situations usually involve complex motion, such as 'sliding motion' in the case of lung registration. Modelling such motion with artificial deformations seems to be a challenging task. Another problem which needs to be addressed is the application of augmentation to the training dataset, mainly due to limited datasets. Again, such an artificial manipulation of the dataset must be done very carefully, keeping in mind the lung physiology, to avoid implausible deformations. While the initial results may look promising, they are not widely applied in this field.

#### 2.3.4.2 Multi-modal lung image registration

The problem of aligning images which have not only been acquired at different points in time, but also originate from different scanners is called multi-modal image registration and is even more challenging than single-modality image registration. In this case an additional difficulty stems from the fact that anatomically corresponding regions in both images are represented by different ranges of intensity. Some structures may not even be present in the images as in the registration of functional imaging to structural. This section focuses on MRI to CT, PET to CT and XeMRI to CT registration problems. An example of multi-modal image pair of the same patient is presented in Figure 2.5, where initial misalignment between both images is shown and its gradual reduction after application of image registration methods.

**PET to CT Registration:** PET is one of the functional imaging methods, therefore finding its direct mapping to diagnostic CT is a particularly challenging task. To make the task more feasible, AC-CT, acquired along PET imaging and assumed to be in alignment, can be used to drive the registration. This has been presented in [25], where an elastic approach to image registration with 'sliding motion' modelling

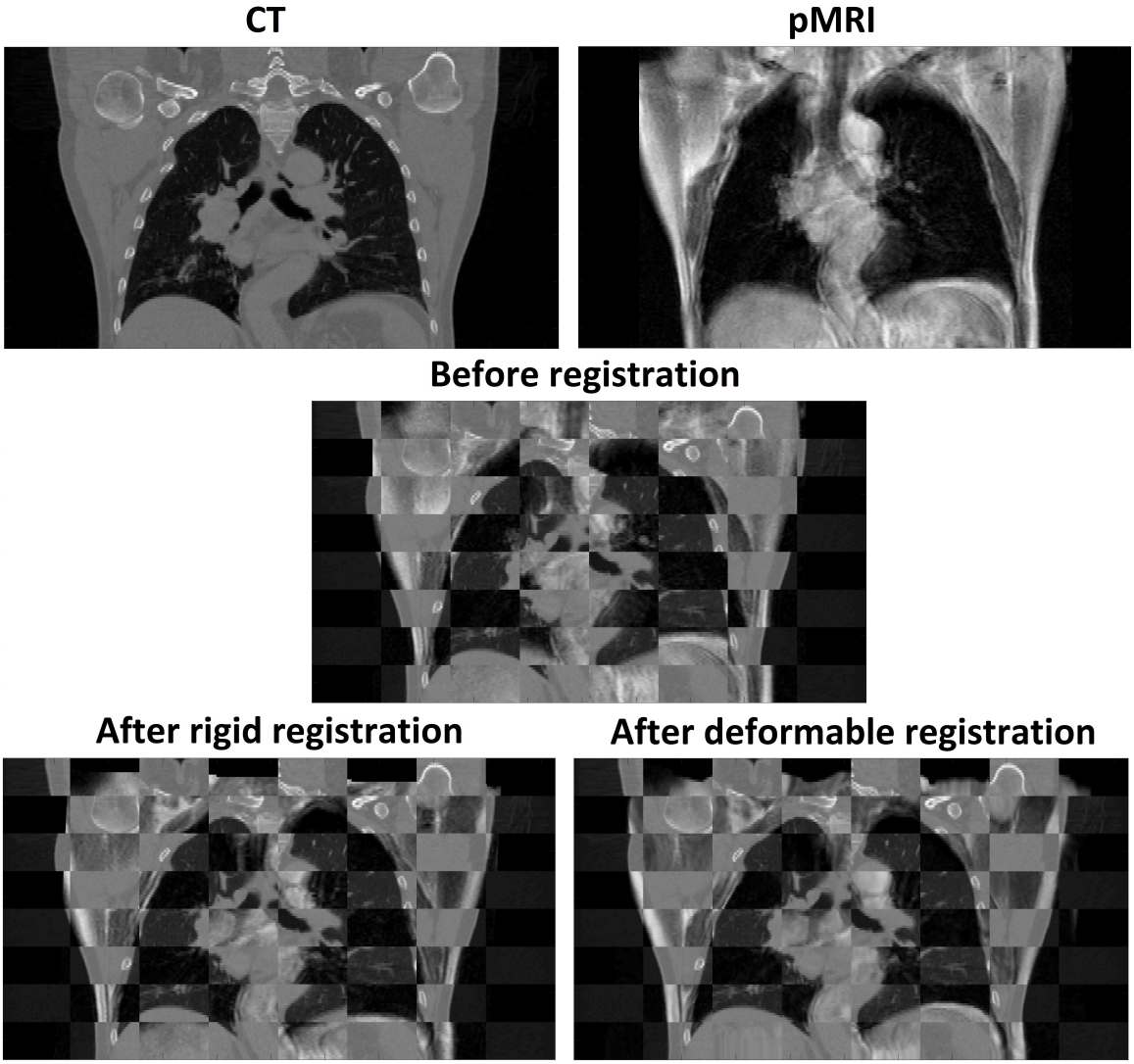



Figure 2.5: An example of a multi-modal CT and MRI pair of the same coronal slice of a patient image. In the upper row original images before image registration are presented. In the middle is a checker-board image showing the initial high level of misalignment between the images. Subsequently, in the lower row on the left the results of the initial rigid registration are shown, with the improved alignment between the images. In the lower row on the right, the image presents the results after application of a deformable registration method, which was capable of bringing both images into alignment. In the presented example deeds deformable image registration method [81] has been applied.

and the application of rigidity with topology preservation constraints [97] has been proposed. The rigidity is preserved by applying additional filtering by using the pre-segmented rigid structures (bones). In the soft tissue regions, the displacement field

after filtering remains unchanged, whereas for the rigid structures it is averaged over the neighbouring voxels belonging to the same region [98]. In [99] first the airway trees segmented from AC-CT and high resolution CT were rigidly aligned. This step was followed by deformable registration of segmented thoracic cavities of the images. This method was introduced with the bronchoscopy procedure planning and guidance in mind.

**MR to CT registration:** The problem of MR to CT image registration is a well established and deeply investigated area in brain imaging [100, 101]. In the case of lung imaging, however, there has not been as much work done. This application becomes more interesting as novel approaches to radiotherapy planning and guiding with the use of MRI are investigated [102]. Direct application of registration methods designed for different tasks may not be able to address the specific problems present in lung imaging. MR to CT lung registration has been addressed in [103], where mutual information as a similarity measure has been combined with the demons framework. The main contribution of that work is to introduce a computationally tractable technique to estimate the gradient of mutual information used as a similarity measure for deformable image registration, which is required to apply the demons framework. This was not the first approach to combine mutual information with demons framework. [104] proposed using a point-wise based approach for calculating mutual information, while in [105] NMI was applied with analytically estimated NMI gradients. One of the MRI registration problems which has recently required addressing is XeMRI to CT registration. Such a functional imaging to structural data registration is particularly difficult, due to the lack of direct correspondence between the images. The initial approaches relied on the application of global transformation methods such as in [106, 107], where rigid transformations were applied to HeMRI. In the latter, rigid registration was enhanced with semi-manual points-to-contour matching. Such an approach is likely to result in suboptimal results, mainly due to differences in the

lung volumes between CT ans MRI, as well as complex lung motion, which was also noticed by the authors. In the approach proposed in [108], the authors investigated two methods, which they called direct and indirect approaches. Both of them shared the same applied image registration framework: affine followed by deformable using the Ants method [109]. This is important that the XeMRI images have been acquired at the same breath-hold as proton MRI images. Such an acquisition makes both already spatially aligned, which prevents the problem of having different levels of lungs inflation between the both, even using a reference volume bag during the imaging. In [110] a comparison of the ventilation images acquired by static HeMRI and 4DCT-based ventilation maps has been presented. To bring the images into alignment an intermediate registration of pMRI to CT was performed. In this study, the authors applied affine registration for lung pairs from MRI and CT independently. That is, left lung from MRI was registered to the left lung from CT and an analogous procedure was performed for the right lung. This area, however, still lacks sufficient research and this thesis will address this problem presenting a novel approach for XeMRI to CT registration.

#### 2.3.4.3 Motion models in application to pulmonary image registration

Lung motion models may be applicable for lung image registration [111, 112]. For instance, a statistical motion model based on deformations estimated from 4DCT co-registration was proposed in [113]. The individual motion models estimated for each subject from the dataset were co-registered to an average shape and an intensity model was generated from reference frames from 4DCT. This resulted in the development of an average inter-subject model. In [114], after estimating the deformations from 4DCT, the surface point distribution model of the shape of the lungs was constructed. After applying Principal Component Analysis (PCA) to reduce the dimensionality, the statistical model between the estimated deformations and point-

based shape variations was calculated. A similar approach was presented in [115], with the diaphragm position used as a surrogate of the motion to control the model. To create a lung motion model, Finite Element Analysis (FEA) can also be used, as can be seen in [116], where a patient-specific bio-mechanical model has been proposed for lung CT registration. However, to achieve satisfying accuracy the FEA model-based method required an additional registration between the model-based estimated volume and the target volume. All of the afore mentioned methods have been applied to CT-to-CT registration problems. In the case of pMRI-to-CT registration, the task may be even more challenging due to the low out-of-plane spatial resolution of pMRI and lack of direct intensity correspondences.

### 2.3.5 Lung Image Registration Evaluation

A comparison of the performance of different methods is usually a problematic task, due to many reasons. First of all, there are often restrictions about releasing the data on which the research was conducted. This makes direct comparison between different methods more difficult or even impossible. Furthermore, even if other researchers have made their methods available to others, tuning the parameters for methods in which ones are not experts may be challenging and result in suboptimal performance of the method. Then the comparison unintentionally may not be totally fair [117].

To present a reliable comparison of different methods, the *Evaluation of Methods for Pulmonary Image Registration 2010* (EMPIRE10) challenge was proposed [118]. The participants used many different similarity measures. It is difficult to identify the differentiating characteristic of the best ranked methods. It can be observed that the methods using NCC, which is defined in the next section, had the best overall performance [109]. Some methods segmented the lungs as a pre-processing stage and used these masks to complement motion correction. In general, these methods obtained better results. However, such an approach becomes problematic in

CT lung cancer images, as well as in lung MRI images, where lung borders are not easy to segment. Nevertheless, it can be concluded that there is not one universal lung registration technique guaranteeing the best results. Different combinations of similarity measure, transformation models and optimisation can lead to similar results.

Another platform created for the development and comparison between different image registration methods is the Dir-Lab dataset [119, 120]. For each of the 10 4DCT cases, 300 manually placed landmarks were provided for extreme breathing phases. The landmarks are distributed not only on the surface of the lungs or on the main airways, but also on the smaller airways, blood vessels and other characteristic features matched between the volumes by experts. The evenly distributed landmarks in the lungs volume make the evaluation process more reliable. Since the release of the dataset it became a platform for comparing of the performance of many methods in terms of accuracy. It also made the comparison as fair as possible, as every author of the method, being an expert in its tuning, has a chance to optimise its performance and decide whether it uses the optimal set of parameters. One may also observe that some methods that have not been originally developed for lung registration have been tested on this dataset.

There are also other 4DCT datasets, such as POPI [121]. This platform, however, is less suitable for detailed evaluation, as there are only six cases with 100 landmarks provided. All discussed platforms are dedicated to single modality CT-CT image registration. The remaining challenge is the evaluation of multi-modal lung registration, especially for functional to structural image registration. Recently, first efforts were made to crate a platform, which would enable an evaluation of the 4DCT-based estimated ventilation maps with ventilation maps originating from SPECT and XeCT [122]. This seems to be a promising project, which, however, still needs some refinements, as the data are not consistent, and there are no landmarks for the evaluation

provided. Addressing these limitations would certainly attract researchers and help in the development of ventilation estimation methods, as well as multimodal image registration frameworks.

## 2.4 Application of Lung Image Registration

The previous section discussed the challenges of lung image registration and how they have been addressed to date. Image registration is often performed not only to bring images into alignment but also to obtain additional information, as will be explained in this section. One such application is adaptive radiotherapy, which can be used, for instance, to treat lung cancer patients. Another application is estimation of the lung ventilation. The following subsections discuss both these applications.

### 2.4.1 Adaptive Radiotherapy

Planning CT (pCT), which is used to prepare radiotherapy plans, is acquired in advance of the treatment delivery. Over that period, a plan detailing in which areas to deliver a prescribed dose, as well as the organs at risk along with the margins, are carefully delineated. The treatment is performed over several weeks following the planning stage. During that time, progressive anatomical changes may occur, which can result in differences between the planned dose distribution and that which is delivered in practice. Changes in the anatomy of cancer patients may influence the radiotherapy. They can be related, for instance, to the tumour growth or weight loss. Lower dose delivery to the tumour may not be enough to cure, whereas high radiation of the adjacent healthy tissue could result in RILD, severely worsening cancer survivors' quality of life [123]. In the recent study [14] of the IDEAL CRT dataset [13], the authors reported that all of the patients over a one year follow-up period developed RILD to some extent. Therefore, changes in the patients' anatomy

in some cases should not be neglected and the prepared radiation plans should be adjusted accordingly. To achieve this goal a registration between the pCT and CT acquired during radiotherapy, usually using cone beam CT (CBCT) is required [124].

### 2.4.2 Ventilation Estimation from 4DCT

The deformations resulting from peak inhale to peak exhale 4DCT registration may be used to estimate ventilation surrogates. This section discusses the existing approaches for this task, starting from those requiring estimation of the displacement field. It can be used for lung ventilation quantification [125], which has the potential to spare well-functioning parts of the lungs during radiotherapy. In the case of lung ventilation quantification based on 4DCT, an accurate registration is crucial. The most accepted approach tracks changes in lung tissue intensity, which directly corresponds to the tissue density. Most of the existing methods perform an initial Gaussian image smoothing to compensate for errors in registration. Having images accurately registered, such smoothing can be avoided, resulting in more detailed ventilation maps. The actual level of required accuracy of the estimated ventilation maps is not yet fully established. It is a field still under investigation, as results of different methods are not consistent. The clinical evaluation of this influence will be a subject of a clinical trail evaluation, which will help to establish the influence of the ventilation surrogates on radiotherapy plans and their clinical outcome. The estimated ventilation maps may be further used to correct the radiotherapy plan to avoid side effects due to irradiation of healthy tissue. These methods can be divided into intensity-based [126, 127] and Jacobian-based [128, 129] methods. Subsequently, a method which applies the relation between air and tissue fraction and does not require image registration is presented [122].

#### 2.4.2.1 Intensity/density change ventilation estimation

One of the approaches for ventilation estimation from 4DCT is to track changes in lung intensity throughout the breathing process. As CT intensity is directly related to the tissue density, it reflects the changes originating from filling the lungs with air. Following [126, 130], the ventilation estimate for the peak exhale phase is:

$$Vent^{exh}_{Dens} = \frac{F_{inh}(\mathbf{x}+\mathbf{u}_{inh}) - F_{exh}(\mathbf{x})}{1 - F_{inh}(\mathbf{x}+\mathbf{u}_{inh})}, \tag{2.10}$$

where $F_{exh}(\mathbf{x}), F_{inh}(\mathbf{x}+\mathbf{u}_{inh})$ are defined as fractions of air at voxel $\mathbf{x}$ in exhale or inhale CT volume respectively, and are defined as:

$$F_{exh}(\mathbf{x}) = -\frac{HU_{exh}(\mathbf{x})}{1000}, \tag{2.11}$$

$$F_{inh}(\mathbf{x}+\mathbf{u}_{inh}) = -\frac{HU_{inh}(\mathbf{x}+\mathbf{u}_{inh})}{1000}, \tag{2.12}$$

with $HU_{exh}(\mathbf{x})$ being the intensity value, expressed in Hounsfield units (HU), for a voxel at spatial position $\mathbf{x}$, of the peak exhale image. Similarly, $HU_{inh}(\mathbf{x}+\mathbf{u}_{inh})$ represents the intensity value of peak inhale image for the same voxel $\mathbf{x}$, being warped toward the peak inhale image by the displacement field $\mathbf{u}_{inh}$. The intensity values of the attenuation are divided by 1000HU, which is the value of the air. The result gives the fractions of air in every voxel $\mathbf{x}$. Substitution from Eq. 2.11, 2.12, in Eq. 2.10 results in:

$$Vent^{exh}_{Dens} = \left( \frac{HU_{exh}(\mathbf{x}) * G_{k1} - HU_{inh}(\mathbf{x}+\mathbf{u}_{inh}) * G_{k1}}{HU_{inh}(\mathbf{x}+\mathbf{u}_{inh}) * G_{k1} + 1000} \right) * G_{k2}, \tag{2.13}$$

where additionally $G_{k1}, G_{k2}$ are Gaussian kernels used for smoothing the images. Some researchers substituted Gaussian kernels with box filter kernels, [125], or with median filter kernels, [131, 132]. The study presented in this thesis followed the

original formulations of [126, 130], therefore, it will be subsequently referring to the Gaussian filtering. In [133] the authors investigated the reproducibility of the intensity-based approach for a number of different parameters and cases.

#### 2.4.2.2 Jacobian-based

Local lung volume change can be used as a measure of lung ventilation [128]. The determinant of the Jacobian of deformations estimated by deformable registration between different breathing phases is:

$$Vent_{JAC} = \begin{vmatrix} 1 + \frac{\partial u_x(\mathbf{x})}{\partial x} & \frac{\partial u_x \mathbf{x})}{\partial y} & \frac{\partial u_x(\mathbf{x})}{\partial z} \\ \frac{\partial u_y(\mathbf{x})}{\partial x} & 1 + \frac{\partial u_y(\mathbf{x})}{\partial y} & \frac{\partial u_y(\mathbf{x})}{\partial z} \\ \frac{\partial u_z(\mathbf{x})}{\partial x} & \frac{\partial u_z(\mathbf{x})}{\partial y} & 1 + \frac{\partial u_z(\mathbf{x})}{\partial z} \end{vmatrix} \tag{2.14}$$

with $u_x, u_y, u_z$ being the deformation fields estimated in the $x, y$ and $z$ directions respectively. The value of the determinant of the Jacobian of the deformations expresses how much the lung volume changes. A value of 1 indicates no change in volume, while values over 1 correspond to volume expansion, and values below 1 express shrinkage of the volume at the spatial position, $\mathbf{x}$. Such an approach was applied in [129], where XeCT-based ventilation volumes were compared with the ventilation estimated based on the determinant of the Jacobian. Recently a sensitivity analysis of the determinant of the Jacobian, which was used for lung cancer treatment planning, was presented [134]. The reproducibility of the Jacobian-based ventilation has been investigated in [135]. It has been shown that the method correlates well with lung function in emphysema patients [130].

#### 2.4.2.3 Air/tissue fraction averaging

In contrast to previously presented methods, the HU-based method [122] does not require image registration between breathing phases. Instead of performing image

registration between full inhale and full exhale phases, the authors propose to estimate lung ventilation in terms of a product of air and tissue fractions density calculated locally for all of the breathing phases and subsequently averaged across them:

$$Vent_{HU} = \frac{1}{N}\sum_{\phi=1}^{N} V_{\phi}(\mathbf{x}) \tag{2.15}$$

$$V_{\phi}(\mathbf{x}) = \begin{cases} \frac{HU_{\phi}(\mathbf{x})}{-1000} \times \frac{HU_{\phi}(\mathbf{x})+1000}{1000} & \text{for} \quad \mathbf{x} \in L(\phi) \\ 0 & \text{for} \quad \mathbf{x} \notin L(\phi) \end{cases} \tag{2.16}$$

where $\phi$ is a current number of the breathing phase in 4DCT with $N$ being the total number of breathing phases, and $L$ is the mask of the lungs delineated from the peak inhale phase. The method was reported to have good correlation with Galligas PET for a control group but it failed to achieve the same accuracy for individuals with severe ventilation distortions. Such behaviour could be expected due to the estimation of the ventilation based on a relation between air and tissue fraction for a single phase. The method represented by Eq. 2.15 assumes that physiological ventilation of blood-gas exchange is performed in the process of diffusion. The HU values provide information on the fractional of air/tissue content for each voxel location, therefore the HU-based model for physiological ventilation could compute the voxel-wise product of air and tissue densities. Such information is assumed to be a surrogate for the rate of blood-gas exchange in that voxel. The additional assumption of the model is that the static surrogate of the ventilation estimated for an individual breathing phase visible in CT reflects the dynamic ventilation when averaged over many breath phases acquired via 4DCT. A clearly problematic case is the detection and evaluation of the air trapped in the lungs for patients with severe COPD.

## 2.5 Conclusions

All the imaging methods presented in this chapter provide complementary information about the lungs. Fusion of multi-modal lung images has the potential to facilitate both diagnosis and therapy planning. However, combining information from XeMRI with CT, as well as exploring the information contained in dynamic XeMRI is a scarcely investigated field and requires further research. Most of the presented lung image registration approaches, as well as lung image analysis methods with application to lung ventilation estimation rely on voxel-wise or regular grid image representation. Such an approach does not take the anatomy and physiology directly into account. Methods performing image parametrisation with the consistency of the lung structure in mind, such as supervoxel clustering, which groups voxels spatially close and with similar intensity values into larger structures, have the potential to provide better results. One such example can be supervoxel-based deformable image registration designed for lung applications. In the following chapter, one such method will be described, which not only applies supervoxel-based image parametrisation, but also addresses 'sliding motion' problem in a form of a relaxed graph representation of the image, followed by guided image filtering over the estimated deformation field.

# Chapter 3

# Supervoxels for Deformable Image Registration

This chapter presents a novel method dedicated to solving the lung image registration problem with three main aims. The first of these is to propose a method applicable to lung CT image registration that reduces the size of the 3D deformable image registration problem. This is achieved by making the graph cuts method efficiently applicable to that task by using a supervoxel-based image representation. Based on the supervoxels representing groups of voxels of similar appearance and small spatial distance, a graph is constructed. The second is the introduction of a novel graph relaxation criterion based on the relation between intensities of adjacent supervoxels, which improves the 3D deformable image registration performance for lung applications. The third is showing the presented approach to be capable of preserving the sliding motion of the lungs. The conducted experiments also show that the proposed framework is robust to supervoxel extraction and the graph relaxation parameters. In terms of registration accuracy, the method provides reasonably good results for a range of parameters and modifications to the graph connectivities, as well as applied filtering methods. It is demonstrated that the proposed framework performs

favourably to most other similar deformable image registration methods and especially well compared to those which do not use a supervoxel image representation, when validated on a publicly available dataset. The methodology presented in this chapter has been published, initially as a conference work [136], and later extended to a journal article [137].

Among different imaging modalities applicable to lung imaging, CT remains the most widely used volumetric imaging method. This is the reference modality for the particular case of lung cancer, playing a key role in radiotherapy planning [138] and treatment monitoring [139]. In medical imaging, it is difficult for the clinical expert to directly compare image volumes acquired at different times. In order to accurately assess corresponding structures, an accurate image registration needs to be performed to allow for direct side-by-side comparison. Bringing CT images that were acquired at peak inhale and peak exhale breathing phases into alignment forms the first challenge addressed in this thesis. A practical application for this, in the context of lung imaging, is image-guided radiotherapy, where the dose plan is devised from a baseline scan and mapped, using image registration, to the patient for more accurate dose delivery (a concept known as 'dose painting'), taking the different breathing states into account. Another application is monitoring the patient's response to radiotherapy, where in addition to compensating for differences in patient breath hold, actual morphological or functional changes need to be assessed. For such sophisticated tasks, the image registration accuracy required is of voxel or even sub-voxel size. The estimated deformations can be further used for creating motion models applicable for pMRI/XeMRI to CT registration or to estimate lung ventilation maps. Common software that is available as part of existing radiotherapy planning tools does not always address all the aforementioned lung image registration challenges and may not be up-to-date with recently developed methods. In some cases, it is not possible to export deformation fields which would allow for direct comparison with novel methods

in the research domain. In the previous chapter, different image registration methods developed for lung application have been discussed. Regardless of the contributions in this field, challenges related to this task have yet to be fully answered.

## 3.1 Background

This section presents the chosen method for image clustering with the use of supervoxels. It is followed by the introduction of Markov Random Fields (MRFs) as an approach for discrete optimisation-based image registration. As other methods applying MRFs were discussed in the previous chapter, here only the proposed approach is introduced in detail.

### 3.1.1 Image Clustering

The simplest way to reduce the complexity of deformable image registration arising due to the large size of the images is to parametrise the problem using a regular grid structure. Down-sampling of an image can be considered as a simplification of this method. Even though such an approach efficiently reduces the complexity of the problem, it does not preserve coherent image regions or the high contrast edges that typically bound them. To overcome this limitation, superpixel algorithms have been proposed. These algorithms group pixels with similar characteristics within a local neighbourhood, performing pseudo-segmentation of images into regions that appear consistent visually. In the literature, many approaches have been proposed to perform image clustering, including mean-shift methods [78], gradient-based [140] and the watersheds method [141]. They have found a wide range of applications in imaging, including object localisation and class segmentation [142], human pose estimation [143] and depth estimation [144, 145]. While the traditional application of superpixels is to image (over)segmentation, here they are employed to remove

redundant image intensity information in order to reduce the number of nodes needed for the graph-based registration. At the same time, as superpixels tend to preserve consistent edges, this allows us to effectively preserve sliding boundaries between organs. For 3D applications, the clusters represent groups of voxels and are therefore called supervoxels.

To the best of our knowledge, the most popular image clustering method is Simple Linear Iterative Clustering (SLIC) [146]. This is mainly due to the speed of performance, direct control over the number of extracted supervoxels and supervoxel compactness. The SLIC method is designed to extract $k$ approximately equally-sized supervoxels. It starts by distributing seeds for each of the supervoxels, initially placed at intervals $S = \sqrt[3]{M/k}$ voxels apart, with $M$ being the total number of voxels in the image. Their initial positions are corrected by moving the seeds to positions of lowest gradient in a small neighbourhood (usually $3 \times 3 \times 3$). This step is performed to avoid placing a seed on an edge or a noisy voxel. Following that, every voxel in the image is assigned to the closest supervoxel, based on the distance measure:

$$D = \sqrt{(d_e)^2 + \left(d_I/S\right)^2 m^2}, \tag{3.1}$$

where $d_e = \|\mathbf{x} - \mathbf{c}\|$ is the Euclidean distance of a particular voxel $\mathbf{x}$ to the supervoxel centre $\mathbf{c}$, $d_I = \|(I(\mathbf{x}) - I(\mathbf{c}))\|$ is a voxel's intensity difference from the supervoxel's average intensity, and $m$ is a compactness parameter.

### 3.1.2 MRFs and Graph Cuts

To use a discrete solver for deformable image registration, the problem is usually formulated in terms of MRFs. These create a framework where long-range relations of distant nodes in the graph are modelled via a *knock-on effect* by short-range relations of usually just a pair of neighbouring nodes. The optimisation is posed on an undi-

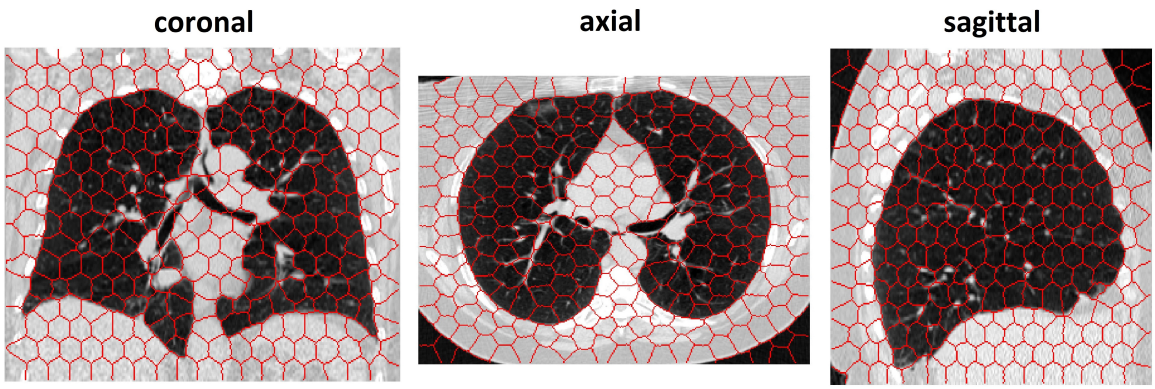



Figure 3.1: An example of superpixel image representation on a lung CT slice in coronal, axial and sagittal plane from Dir-Lab case 8. The superpixels correspond well with lung boundaries and airways inside lungs. For illustrative purposes, we use 2D superpixels and present them on 2D slices. However, in the framework we extract supervoxels from 3D volumes.

rected graph that is composed of a set of nodes $P$, connected by edges, as the process of finding a labelling $f$ for the set $P$ that minimises the energy $E(f)$, consisting of the sum of two terms:

$$E(f) = E_{data}(f) + \alpha E_{smooth}(f), \tag{3.2}$$

where $E_{data}(f)$ represents the disagreement between the labelling $f$ and the observed data, whereas an $\alpha$-weighted piecewise smoothness term, $E_{smooth}(f)$, forces nodes connected by an edge to have similar labels. Depending on the application, the labels can represent different properties, for instance, objects or background in segmentation [77] or a depth map in stereo-vision [78]. Such a formulation can also be applied to deformable image registration, where the labels to be assigned represent predefined displacement vectors $[u_x, u_y, u_z]$.

To address the computational cost issue inherent in pixel/voxel-wise formulations of the problem, the use of a supervoxel image representation is proposed, as shown in Figure 3.1. This approach allows us to reduce the complexity of the problem by

representing a group of voxels as a single supervoxel. The graph cuts algorithm has a worst-case computational complexity of $O(n^2)$ for a graph with $n$ nodes [147]. In our case, the number of nodes $n$ is reduced from the number of voxels to the number of supervoxels. When each of the supervoxels consists of $n_s$ voxels, then the number of supervoxels $N = \frac{n}{n_s}$, and the worst-case computational complexity of our formulation is expressed as $O(N^2)$. Subsequently, the number of displacements to be estimated is reduced from the number of voxels in the image to the number of extracted supervoxels, which is significantly lower. The use of supervoxels has certain limitations. Smaller supervoxels diminish the advantage of the reduction of the number of nodes in the graph (in the limit, the number of supervoxels can be equal to the number of voxels). On the other hand, when the supervoxels are large, they may not be able to accurately model the variations in displacement. In the proposed approach the graph cuts method is applied during the process of finding the optimum displacements, and later an edge-preserving filtering method, namely guided image filtering [148], is used to convert the displacement field derived from the sparse image representation into a piecewise continuous displacement field. To our knowledge, the closest published method to the one we present here is the one in [83]. Even though both methods share some common features, such as the supervoxel approach to the image parametrisation or discrete optimisation, the methods present completely different approaches to image registration. Our proposed method uses a multi-resolution approach to extract supervoxels, whereas in [83] the supervoxels are extracted from the original, unprocessed images. In our case, at each resolution level supervoxels with different relative size are extracted, starting with larger ones and gradually reducing their sizes. Another major difference lies in the graph representation, which will be presented in Sec. 3.2, along with the chosen method to convert the sparse displacement field into the locally continuous one.

The remainder of this chapter is structured as follows. Sec. 3.2 describes the

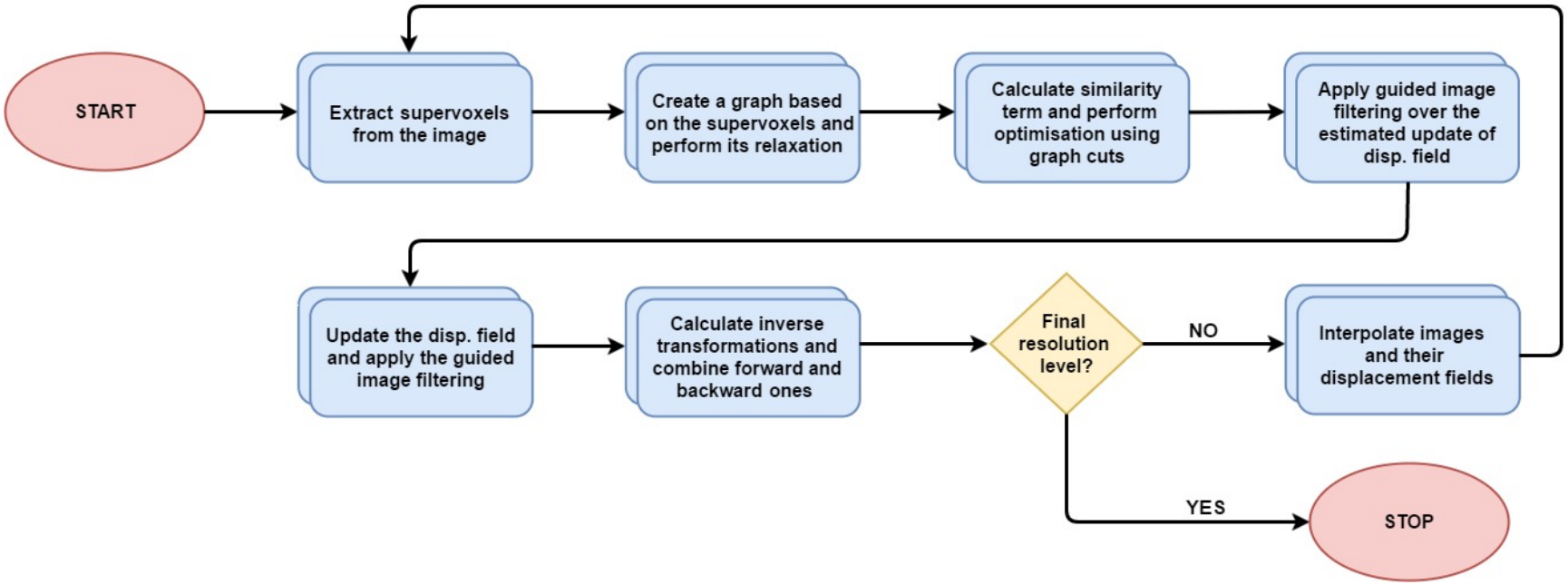



Figure 3.2: Diagram presenting the workflow of the proposed method, starting from extracting supervoxels and creating a graph based on them. This is followed by estimating a displacement for each supervoxel using the graph cuts optimisation method. The guided image filtering is first applied to the estimated update of the displacement field, similar to the fluid-like regularisation in demons, and then to the update combined with the previously estimated displacement field, similar to the diffusion-like regularisation in demons. Subsequently, inverse transformations are estimated and combined with corresponding forward ones. Double rectangles represent the symmetric nature of the method, where forward and backward transformation from the target to the moving image are estimated at the same time.

proposed method, introducing all of its components. In Sec. 3.3 we present the conducted experiments and obtained results. A discussion of the results is presented in Sec. 3.4.

## 3.2 Methods

This section presents the proposed method for deformable image registration, using supervoxel-based image representation and graph cuts as an optimisation routine. An overview of the proposed method is shown in Figure 3.2, with each component addressed individually in the following sections. Sec. 3.1.1 has already explained the way in which the images are over-segmented into supervoxels and described the selected image clustering method. The following Sec. 3.2.1 explains in detail the graph construction and energy function formulation used for the initial displacement field estimation. Sec. 3.2.2 describes the proposed transition from the discrete domain to the continuous with the application of the guided image filtering method, while Sec. 3.2.3 explains how to combine the estimated forward and backward transformations.

### 3.2.1 Energy Formulation

In this work, the optimisation problem is posed on an undirected graph, defined by the adjacency of previously extracted supervoxels. In the graph, each supervoxel is represented by a node and every pair of directly adjacent supervoxels is connected by an edge. The edges indicate a relation between the supervoxels while the edge values represent the strength of the relation. In CT imaging structures belonging to the same organ or consistent part of the organ, such as in our case lung lobe/section usually present similar intensity profile. There is a noticeable difference in mean intensity between supervoxels belonging to the lungs and those belonging to the neighbouring structures, such as liver, rib-cage or mediastinum. Based on this assumption we

calculate mean intensity value of all supervoxels and use the absolute difference between these intensities of adjacent supervoxel pairs to assign values of the edges in the graph connecting nodes representing the supervoxels. Such construction of the graph favours smooth changes between supervoxels having similar mean intensity values and thus probably representing the same structures. To allow for discontinuities in the deformation field between supervoxels clustering different organs, while at the same time preserving the smooth deformation field inside them, a relaxed form of the graph is introduced. The edges with lower values are assumed to connect different structures, and can thus be removed, by using a relaxation threshold parameter $\upsilon$:

$$V(p,q) = \begin{cases} 1 - \|\overline{I_{mov}(\mathbf{x}_p)} - \overline{I_{mov}(\mathbf{x}_q)}\| & \text{for} \quad \|\overline{I_{mov}(\mathbf{x}_p)} - \overline{I_{mov}(\mathbf{x}_q)}\| < \upsilon \\ 0 & \text{for} \quad \|\overline{I_{mov}(\mathbf{x}_p)} - \overline{I_{mov}(\mathbf{x}_q)}\| \geqslant \upsilon, \end{cases} \tag{3.3}$$

where $\overline{I_{mov}(\mathbf{x}_p)}$ and $\overline{I_{mov}(\mathbf{x}_q)}$ stand for mean intensity values of all voxels $\mathbf{x}_p, \mathbf{x}_q$ contained within the supervoxels represented by nodes $p$ and $q$, respectively, calculated for the moving image. The image intensities are initially normalised between 0 and 1, therefore the relaxation threshold parameter $\upsilon$ is a constant and lies within these ranges. The relaxed graph representation has an additional advantage: as there are fewer non-zero edges remaining in the graph, the overall complexity of the graph layout is reduced. An example of such clustering and a graph created over the image, as well as its relaxed form is shown in Figure 3.3. The proposed formulation is designed to result in physically plausible deformations inside the lungs and, in turn, a more realistic deformation field. The graph relaxation term that has been introduced can be considered as a form of pre-segmentation of the image, akin to lung masking, which is a common practice in the field of lung registration [43, 149, 150].

The data cost term is formulated as a mean error calculated for all voxels $\mathbf{x}$ in the fixed image $I_{fix}$ and moving image $I_{mov}$ clustered in a certain supervoxel represented

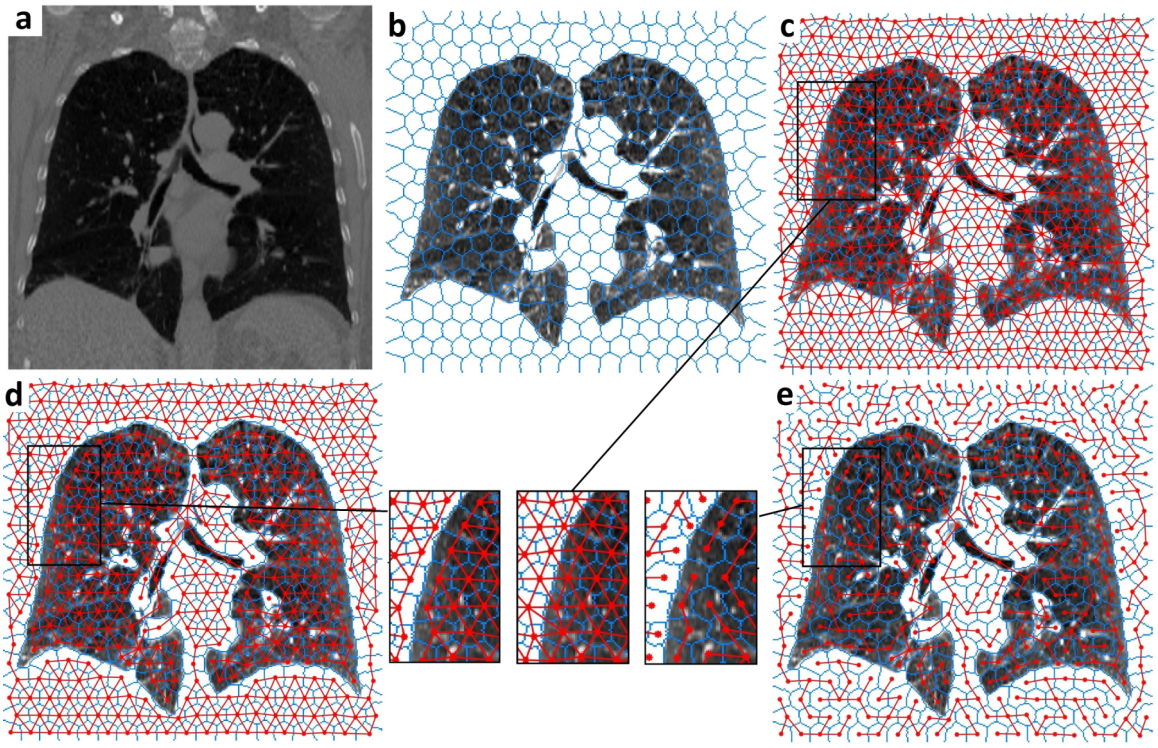



Figure 3.3: An example of the SLIC algorithm clustering of the peak inhale 2D lung CT image. In (a) the original CT image is shown, while (b) shows the corresponding supervoxel representation, with the borders of the supervoxels drawn in blue. The graph extracted based on the clustering is shown in figure (c), with the edges connecting adjacent supervoxels in red. The relaxed graph representation is shown in figure (d), whereas the minimum spanning tree-based simplification of the graph is shown in figure (e). Apart from Figure (a), the contrast in all images has been adjusted for better visualisation of supervoxels and graphs. For illustrative purposes, 2D superpixels are used and presented on a 2D slice. However, in our framework, supervoxels from 3D volumes are extracted.

by a node $p$, for the applied displacement $f_p$:

$$E_{data}(f) = \Sigma_p \overline{|\text{MIND}(I_{fix}, \mathbf{x}_p) - \text{MIND}(I_{mov}, \mathbf{x}_p + f_p)|}. \tag{3.4}$$

The proposed method does not aim to match supervoxels between two images. Instead, the intention is to find the optimal displacement for each of the supervoxels extracted only from the moving image based on the mean error of all their included voxels (3D image pixels) for each label, which represents a displacement. Such a

formulation is more robust to the potential differences in the shape of supervoxels if extracted in both images.

Many forms of the piecewise smoothness term for MRFs have been proposed, starting from the simplest Potts model, where the term takes the value 1 when labels are equal and 0 otherwise. In a linear model, the smoothness term changes linearly according to the distance between labels. These formulations of the piecewise smoothness term satisfy all metric requirements, including the triangle inequality. The most common way of formulating regularisation in deformable image registration is a quadratic model, which depends on a squared distance between labels $f_p$ and $f_q$ assigned to two neighbouring nodes $p$ and $q$, defined as:

$$\|f_p - f_q\|^2 = (u_{x_p} - u_{x_q})^2 + (u_{y_p} - u_{y_q})^2 + (u_{z_p} - u_{z_q})^2. \tag{3.5}$$

To recall, $f_p$ is a label representing a displacement vector $(x_p, y_p, z_p)$ assigned to the node $p$. Our proposed formulation uses the following piecewise smoothness term:

$$E_{smooth}(f) = \Sigma_{p,q \in N} \ (1 - \|I_{mov}(\mathbf{x}_p) - I_{mov}(\mathbf{x}_q)\|) \ \|f_p - f_q\|^2, \tag{3.6}$$

where the $(1 - \|I_{mov}(\mathbf{x}_p) - I_{mov}(\mathbf{x}_q)\|)$ factor stands for the edge values of the graph and $\|f_p - f_q\|^2$ represents the quadratic distance between labels.

We apply the graph cuts method [151] on this formulation. Due to the fact that we are using a quadratic regularisation term, which does not satisfy the triangle inequality condition for metrics, we use the $\alpha - \beta$ swap variant of the graph cuts method [151], which has relaxed requirements about piecewise smoothness term properties and can be used with a semi-metric piecewise smoothness term.

In the proposed formulation a multi-resolution approach is implemented. At each resolution level we extract a number of supervoxels containing a specified number of voxels. For each of the supervoxels we calculate an error vector with length equal to

the number of labels. Each value of the vector corresponds to the data term of the energy for a particular displacement represented by the label. In such a way, we create a data cost matrix, whose size is $(number\ of\ supervoxels) \times (number\ of\ labels)$. The optimisation complexity has been heavily reduced compared to the voxel-wise formulation, where the data cost matrix would have size $(number\ of\ voxels) \times (number\ of\ labels)$, bearing in mind that every supervoxel represents a group of voxels.

### 3.2.2 Guided Image Filtering

The labelling procedure that was described in the previous section results in the assignment of a displacement vector to each supervoxel, which estimates their optimal displacement and creates a displacement field. A direct application of the deformation estimated for a supervoxel would assign the exact same deformation to all voxels in each supervoxel which is generally not physiologically accurate and may result in foldings at borders of supervoxels that have different deformations, resulting in physiologically implausible deformations. To overcome this limitation, in [83] the authors proposed using multiple layers of supervoxels while in [79, 81] a free-form deformation (FFD) with B-splines [67] is used as an interpolation method.

In this work we follow an alternative solution proposed in [152]; namely, to smooth the displacement field, while at the same time preserving some of the discontinuities. Guided image filtering [148], where an output image intensity $I_{out}(\mathbf{x})$ at voxel position $\mathbf{x}$ is a linear combination of an input image $I_{in}$ and an image $I_g$ used for guidance, is calculated for all voxels $\mathbf{y}$ belonging to a certain neighbourhood $\mathcal{N}(\mathbf{x})$ of a voxel $\mathbf{x}$, and can be formulated in a general linear translation-variant filtering equation:

$$I_{out}(\mathbf{x}) = \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} \mathbf{W}_{\mathbf{x},\mathbf{y}}(I_g) I_{in}(\mathbf{y}), \tag{3.7}$$

where $\mathbf{W_{x,y}}$ is a filter kernel calculated for the guide image $I_g$ in a neighbourhood $\mathcal{N}(\mathbf{x})$ of a voxel $\mathbf{x}$. At the same time, following the key assumption of the guided image filtering, that there is a linear model between the guidance image and the output image in a windows $\omega_k$ centered at voxel $\mathbf{k}$, we can assume:

$$I_{out}(\mathbf{x}) = a_k I_g(\mathbf{x}) + b_k, \forall_{\mathbf{x}} \in \omega_k, \tag{3.8}$$

where $a_k$, $b_k$ are linear coefficients assumed to be constant in the window $\omega_k$. After formulating an optimisation problem to estimate the parameters $a_k$ and $b_k$ in a form that a difference between $I_{out}$ and $I_{in}$ is minimised, while linear relation between $I_{out}$ and $I_g$ is maintained in a form:

$$E(a_k, b_k) = \sum_{\mathbf{x}\in\omega_k} \left( (a_k I_g + b_k - I_{in})^2 + \epsilon a_k^2 \right). \tag{3.9}$$

As this equation takes form of the linear ridge regression, based on [153, 154] the solution for $a_k$, $b_k$ is given as:

$$a_k = \frac{\frac{1}{\omega}\sum_{\mathbf{x}\in\omega_k} I_g(\mathbf{x}) I_{in}(\mathbf{x}) - \mu_k \overline{I_{in}(\mathbf{x})}}{\sigma_k^2 + \epsilon}, \tag{3.10}$$

$$b_k = \overline{I_{in}(\mathbf{k})} - a_k \mu_k, \tag{3.11}$$

where $\mu_{I_g}$, $\sigma_k^2$ are the mean and variance of the guidance image $I_g$ calculated in a windows $\omega_k$ centered at voxel $\mathbf{k}$. An intuitive understanding for these equations is that for the regions of high variance, where $I_g$ changes a lot with in $\omega_k$, $\sigma_k^2 \gg \epsilon$, which gives $a_k \approx 1$ and $b_k \approx 0$. In contrast, for patches of low variance $\sigma_k^2 \ll \epsilon$, $a_k \approx 0$ and $b_k \approx \mu_{I_g}$. The kernel $\mathbf{W_{x,y}}$ can be defined as:

$$\mathbf{W}_{\mathbf{x},\,\mathbf{y}}(I_g) = \frac{1}{|\omega|^2} \sum_{k:(x,y)\in\omega_k} (1 + \frac{(I_g(\mathbf{x}) - \mu_{I_g})(I_g(\mathbf{y}) - \mu_{I_g})}{\sigma_k^2 + \epsilon}. \tag{3.12}$$

In the proposed method, this filtering method is applied over the dense displacement field $\mathbf{u}$ derived from the sparse representation using the moving image $I_{mov}$ for guidance. The discontinuities in the intensity, especially at the borders of organs such as the lungs, may result in discontinuities in the displacement field. On the other hand, small changes in the intensities between adjacent supervoxels may indicate that they are part of the same organ. This yields smooth deformations across anatomically consistent regions in the moving image, while at the same time preserves discontinuities in the deformation field at the region boundaries. The idea of guided image filtering is based on the edge-preserving approach, similar to bilateral filtering [155], which discourages smoothing across boundaries and has good edge-preserving smoothing properties. The main advantage of the guided image filtering method is that it is several times faster than bilateral filtering. Related approaches that were previously proposed include continuous optimisation-based registration for the lungs [64], where bilateral filtering was applied, and liver motion compensation [152] with guided image filtering.

### 3.2.3 Symmetric Registration

In order to enhance the registration process as well as more accurately model the complex nature of the lungs during the breathing process, we propose to apply symmetric registration to the proposed formulation [156]. At each resolution level, two deformation fields are estimated and these fields correspond to the registrations between, respectively, the moving image towards the target one and *vice versa*. Such a formulation not only compensates for a proportion of mis-registrations originating from the relaxed graph formulation, especially at the border of the lungs and diaphragm, where there is no direct link between the organs, and so a considerable amount of motion may be observed; it also removes the bias caused by the choice of the target and source. After performing an optimisation at each resolution level,

inverses of both previously estimated deformation fields are approximated based on a simple fixed-point approach originally proposed in [157] and subsequently combined together:

$$\mathbf{u}^{new}_{fix\rightarrow mov} = 0.5 \cdot (\mathbf{u}_{fix\rightarrow mov} + \mathbf{u}^{-1}_{mov\rightarrow fix}), \tag{3.13}$$

$$\mathbf{u}^{new}_{mov\rightarrow fix} = 0.5 \cdot (\mathbf{u}_{mov\rightarrow fix} + \mathbf{u}^{-1}_{fix\rightarrow mov}), \tag{3.14}$$

where $\mathbf{u}_{mov\rightarrow fix}$ and $\mathbf{u}_{fix\rightarrow mov}$ stand for the estimated deformation field transforming $I_{mov}$ to $I_{fix}$ and from $I_{fix}$ to $I_{mov}$ respectively.

## 3.3 Experiments and Results

This section evaluates the proposed method on a publicly available CT dataset. A number of experiments are carried out to support the proposed configuration of the method as well as the parameter settings. Sec. 3.3.1 describes the dataset, which was used for the evaluation of the proposed method. The parameter setting is described in Sec. 3.3.2, whereas Sec. 3.3.3 explores the influence of the image representation by performing experiments on regular grid image parameterisation. This is continued in Sec. 3.3.4 with an analysis of the effect on the results of the applied filtering method, by comparing results of Gaussian filtering and the proposed guided image filtering. Sec. 3.3.5 investigates the influence of the relaxation parameter and Sec. 3.3.6 shows how the method deals with sliding motions. Finally, Sec. 3.3.7 compares the proposed method with several others in the field that can be considered to be closely related.

### 3.3.1 Materials

The performance of the proposed image registration method is evaluated on the publicly available Dir-Lab dataset (www.dir-lab.com) [119], which consists of 10 cases of dynamic 3D CT volumes covering the full breathing cycle (also called 4D CT)

of patients suffering from oesophageal or lung cancer. The spatial resolution of the dataset varies between $0.97 \times 0.97 \times 2.5$ mm$^3$ and $1.16 \times 1.16 \times 2.5$ mm$^3$. The dataset provides 300 landmarks for peak inhale and peak exhale breathing phases, manually placed by an expert. Due to the well-distributed landmarks, this dataset creates a valuable validation framework for deformable lung image registration. We performed registrations between peak inhale and peak exhale breathing phases for the dataset to evaluate the performance of the method in terms of registration accuracy. The calculated target registration error (TRE) between the annotated landmarks is defined as the Euclidean distance over corresponding landmark locations before and after registration. The observer error for this dataset ranges from 0.75mm to 1.13mm [119]. To determine the observers error, manual registration of a subset of landmarks (200 for cases 1-5 and 150 for cases 6-10) was performed by independent observers (two independent observers were involved in the evaluation of cases 1-5, and three in the remaining cases 6-10). The observers performed the manual registration between unregistered extreme breathing phase CT volumes. The averaged inter-observer error for this dataset ranges from 0.75mm to 1.13mm [119]. The initial TRE of unregistered images was 8.46mm on average with 5.4 of std dev., with the highest TRE of 15mm and std dev. of 9 observed for case 8.

### 3.3.2 Parameter Setting

All the experiments were conducted with the same parameter settings for consistency. Any changes in the parameter settings are clearly stated in the remainder of this thesis. The experiments are performed at five resolution levels, at each of which the supervoxels were extracted with approximate sizes of $[50, 150, 400, 800, 500]$ voxels with maximum allowed displacement of $[3, 4.5, 4.5, 6, 4.5]$ voxels and quantification step of $[0.5, 0.75, 0.75, 1, 0.75]$ respectively. The approximate size of supervoxels at each resolution level was deliberately chosen not to be consistent across different

resolution levels. The intention was to allow supervoxels at subsequent resolution levels represent smaller structures. The initial size of the supervoxels was chosen based on the visual assessment of their overlay on CT images in terms of ability to represent gradually decreasing regions. The maximum allowed displacement was chosen to cover for the largest possible displacements to be estimated at each resolution level. Additionally, the number of supervoxels for which the maximum allowed displacement was assigned was monitored to make sure that larger ranges are not required. The quantification step was set to keep the number of allowed labels to be assigned at a constant level across all resolution levels. In the presented settings we chose to have 13 possible steps in each direction (x, y z), which gives in total 2197 possible different displacement vectors to be assigned. It was a reasonable choice between performance and the computational requirements. Further increase in step size was leading to higher demands in computational resources without gain in the accuracy. In the current implementation of the method, images are initially interpolated to isotropic voxel size for the first four resolution levels. For the final resolution level, which performs fine-tuning, the method was applied on the original, unprocessed images. The windows $w$ for guided image filtering were set to [5 5 5], apart from the final resolution level, which was set to [7 7 7] voxels and the smoothness parameter $\sigma$ to 0.001. The default value of $\sigma$ for the implementation of the guided image filtering used in our framework was 0.01. In our initial heuristic parameter search experiments, we have noticed more visually plausible results of the filter outcomes, as well as slight improvement in the TRE for some of the cases when $\sigma$ was set to 0.001. Such observations is most likely caused by smaller differences in intensities between supervoxels inside the lungs in CT scans, then can be observed between objects in natural images on which the implementation was initially tested. Therefore, we decided to fix $\sigma = 0.001$ in all our experiments, rather than leaving it at the default level of $\sigma = 0.01$. Such setting should be more sensitive to subtle changes in

intensities. The MIND descriptor was calculated using six neighbouring voxels, with range r=1. The parameters were chosen heuristically, based on the results from case 8, as this proved to be the most challenging case in the dataset with the highest initial misalignment. By optimising parameters only on one case, we expect the approach to be more robust to over-fitting.

### 3.3.3 Influence of Image Representation

An experiment was conducted using a regular grid representation, setting SLIC's compactness parameter $m$ to 4000, which forced the supervoxels to take a cubic shape and achieved TRE of 1.36mm, compared to 1.16mm for the proposed method. In all the other experiments $m$ was set to 40 to allow supervoxels to take an optimal non-cubic shape. Results of this comparison are presented in Table 3.1. Among all of the individual cases of the Dir-Lab dataset the TRE for the proposed method is the lowest. A higher difference was calculated in case 8, where the regular grid variant results in TRE of 2.42mm, whereas for the proposed method TRE was only 1.37mm. In Figure 3.4 the magnitude of the displacement field in the coronal view for both methods are presented.

One of the ways to measure the quality of the estimated deformation field may be through calculating the determinant of the Jacobian of the deformations, which can be seen as an indicator of complexity of the deformations. The Jacobian of the deformations is a matrix created by the first derivatives of the deformations calculated for each variable. Its determinant takes the form:

$$det(Jac(x,y,z)) = \begin{vmatrix} 1+\frac{\partial u_x(x,y,z)}{\partial x} & \frac{\partial u_x(x,y,z)}{\partial y} & \frac{\partial u_x(x,y,z)}{\partial z} \\ \frac{\partial u_y(x,y,z)}{\partial x} & 1+\frac{\partial u_y(x,y,z)}{\partial y} & \frac{\partial u_y(x,y,z)}{\partial z} \\ \frac{\partial u_z(x,y,z)}{\partial x} & \frac{\partial u_z(x,y,z)}{\partial y} & 1+\frac{\partial u_z(x,y,z)}{\partial z} \end{vmatrix}, \tag{3.15}$$

where $u_y(x,y,z), u_z(x,y,z), u_z(x,y,z)$ are the displacement fields estimated in $x, y, z$

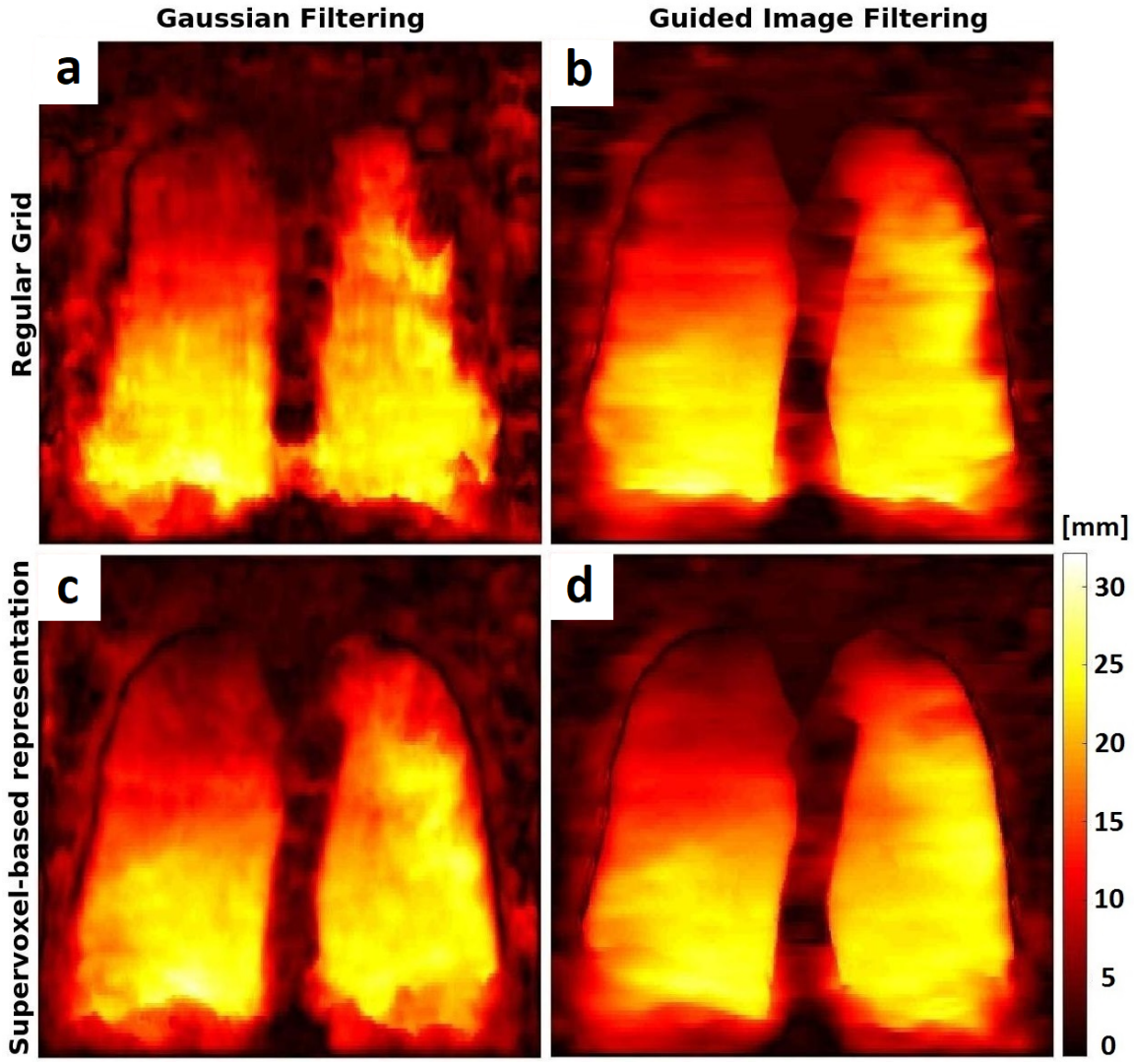



Figure 3.4: Magnitude of the displacement field for different image representations and filtering methods presented for case 8 from the Dir-Lab dataset. The images show a 2D coronal plane from the 3D magnitude volume. On the upper row (a, b) the results for regular grid image representation are shown, while the lower row (c, d) shows those obtained using supervoxel-based image representation. The left column (a ,c) presents the results for Gaussian filtering and the right (b, d) for guided image filtering. The best results were achieved with the proposed configuration shown in the lower right corner.

directions respectively. Table 3.2 presents the quantification of the negative determinant of the Jacobian of deformations estimated during the image registration, which shows the fraction of voxels for which a one-to-one mapping is violated. The mean of the determinant of the Jacobian of deformations is also calculated, showing the change in volume, and the standard deviation of the determinant of the Jacobian of deformations, all calculated inside the lungs for the Dir-Lab dataset. A trend can be observed that the variants using supervoxel image representation result in lower fractions of negative determinant of the Jacobian of deformations, as well as lower standard deviations of the determinant of Jacobian of deformations.

Table 3.1: Comparison of the proposed supervoxel-based image representation (Proposed) against regular grid (RG) image, and guided image filtering (GIF) against Gasussian filtering (Gauss), based on TRE in [mm] and standard deviation for the Dir-Lab dataset. The results shown in bold are those that are consistently best for the proposed method using supervoxels and guided image filtering

| | Target Registration Error in [mm] and standard deviation | | | | |
|---|---|---|---|---|---|
| Case | Initial | RG+Gauss | RG+GIF | Proposed+Gauss | Proposed+GIF |
| c1 | 3.89 ± 2.8 | 1.16 ± 0.7 | 0.98 ± 0.5 | 1.02 ± 0.5 | **0.95 ± 0.5** |
| c2 | 4.34 ± 3.9 | 1.16 ± 0.7 | 1.03 ± 0.7 | 1.03 ± 0.6 | **0.94 ± 0.5** |
| c3 | 6.94 ± 4.0 | 1.42 ± 1.0 | 1.16 ± 0.6 | 1.17 ± 0.6 | **1.10 ± 0.6** |
| c4 | 9.83 ± 4.8 | 1.62 ± 1.1 | 1.37 ± 0.9 | 1.42 ± 1.0 | **1.35 ± 0.9** |
| c5 | 7.48 ± 5.5 | 1.70 ± 1.8 | 1.38 ± 1.3 | 1.46 ± 1.4 | **1.33 ± 1.2** |
| c6 | 10.9 ± 6.9 | 1.63 ± 1.5 | 1.20 ± 0.7 | 1.25 ± 1.1 | **1.14 ± 0.7** |
| c7 | 11.0 ± 7.4 | 2.39 ± 3.3 | 1.58 ± 2.0 | 1.45 ± 1.6 | **1.13 ± 0.7** |
| c8 | 15.0 ± 9.0 | 4.54 ± 6.4 | 2.42 ± 3.9 | 2.37 ± 3.7 | **1.37 ± 1.9** |
| c9 | 7.92 ± 3.9 | 1.53 ± 1.2 | 1.27 ± 0.8 | 1.21 ± 0.7 | **1.14 ± 0.6** |
| c10 | 7.3 ± 6.3 | 1.52 ± 1.6 | 1.18 ± 1.0 | 1.25 ± 0.9 | **1.16 ± 0.8** |
| Mean | 8.46 ± 5.4 | 1.87 ± 1.9 | 1.36 ± 1.2 | 1.36 ± 1.2 | **1.16 ± 0.8** |

In Figure 3.5 the influence of the regularisation parameter $\alpha$ from Eq. 3.2 is presented, which weights data and regularisation terms. The best performance of our method in terms of TRE has been recorded for a value of $\alpha = 0.01$, which has been used in the presented experiments. An interesting observation is that even when $\alpha$ is set to a low value, strongly reducing the influence of the regularisation, the value

Table 3.2: Comparison of the proposed method with its variants based on the quantification of negative determinant of the Jacobian of deformations, which measures the fraction of voxels for which a one-to-one mapping is violated, as well as mean of the determinant of the Jacobian of deformation, showing the change in volume and standard deviation of the determinant of the Jacobian of deformations, which can be seen as a measure of complexity of the deformations, for the Dir-Lab dataset. Different variants of the method are described as: the proposed supervoxel-based image representation (Proposed), regular grid image representation (RG), guided image filtering (GIF) and Gasussian filtering (Gauss).

| Method | RG + Gauss | RG + GIF | Proposed + Gauss | Proposed + GIF |
|---|---|---|---|---|
| Fraction of neg J. | $8.5\times10^{-3}$ | $1.0\times10^{-3}$ | $1.8\times10^{-3}$ | $7.2\times10^{-4}$ |
| Mean Jacobian | 0.910 | 0.918 | 0.915 | 0.918 |
| Std of Jacobian | 0.33 | 0.19 | 0.22 | 0.16 |

Table 3.3: Comparison of the proposed method applied over the full graph without the relaxation criterion, after applying the proposed graph relaxation and after relaxing the graph to a minimum spanning tree (MST), based on TRE and standard deviation for Dir-Lab dataset. Best results are shown in bold.

| | Target Registration Error in [mm] and standard deviation | | | |
|---|---|---|---|---|
| Case | Initial | Full graph | Relaxed graph | MST |
| c1 | 3.89 ± 2.8 | **0.95 ± 0.5** | **0.95 ± 0.5** | 0.95 ± 0.5 |
| c2 | 4.34 ± 3.9 | **0.94 ± 0.5** | **0.94 ± 0.5** | 0.97 ± 0.6 |
| c3 | 6.94 ± 4.0 | 1.10 ± 0.6 | 1.10 ± 0.6 | **1.09 ± 0.6** |
| c4 | 9.83 ± 4.8 | 1.37 ± 0.9 | **1.35 ± 0.9** | 1.37 ± 0.9 |
| c5 | 7.48 ± 5.5 | 1.35 ± 1.2 | **1.30 ± 1.2** | 1.37 ± 1.3 |
| c6 | 10.9 ± 6.9 | 1.16 ± 0.7 | **1.14 ± 0.7** | 1.21 ± 0.8 |
| c7 | 11.0 ± 7.4 | 1.20 ± 0.9 | **1.14 ± 0.7** | 1.26 ± 1.0 |
| c8 | 15.0 ± 9.0 | 1.58 ± 2.4 | **1.37 ± 1.7** | 1.75 ± 2.7 |
| c9 | 7.92 ± 3.9 | 1.15 ± 0.6 | **1.14 ± 0.6** | 1.17 ± 0.7 |
| c10 | 7.3 ± 6.3 | **1.12 ± 0.8** | 1.16 ± 0.8 | 1.16 ± 0.8 |
| Mean | 8.46 ± 5.4 | 1.19 ± 0.7 | **1.16 ± 0.8** | 1.23 ± 1.0 |

of TRE is only slightly raised. This is most likely due to the applied guided image filtering in the transition to a locally continuous deformation field, as well as the multiresolution approach. Both of them can partially compensate for the limited influence of the regularisation term.

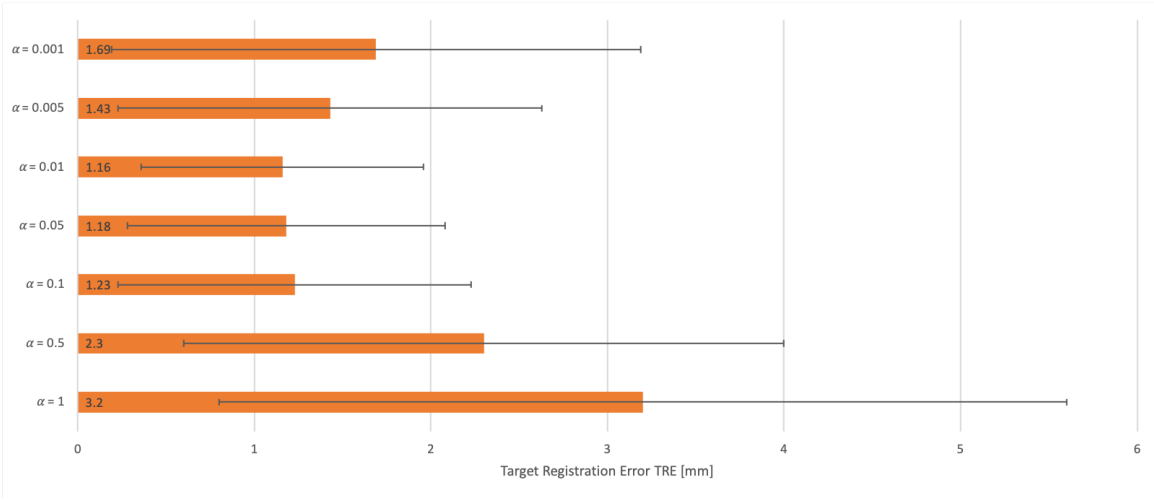



Figure 3.5: The influence of the regularisation parameter $\alpha$ on the TRE.

### 3.3.4 Impact of Filtering Method

Direct application of the estimated deformations from the graph cuts-based optimisation process is very likely to result in unwanted deformations, such as gaps and overlaps. To address this issue, in our approach we perform smoothing of the estimated deformation field with guided image filtering method. A conversion from a sparse displacement representation, where different displacement vectors are assigned to every supervoxel, to its smoothed version, was applied in our method to addresses this issue. Such a formulation eliminates potential problems with gaps in the displacement field, however the proposed workflow do not guarantee explicitly that overlaps will not be introduced to the deformation field. Table 3.2 presents the quantification of the negative determinant of the Jacobian of deformations estimated during the image registration, which shows the fraction of voxels for which a one-to-one mapping is violated. The mean of the determinant of the Jacobian of deformations is also calculated, showing the change in volume, and the standard deviation of the determinant of the Jacobian of deformations, all calculated inside the lungs for the Dir-Lab dataset. The measures have been calculated for all four variants of the proposed

method, namely using regular grid (RG) representation with Gaussian filtering, RG with guided image filtering (GIF), supervoxel based parametrisation with Gaussian filtering and finally the proposed supervoxel-based with GIF. A trend can be observed that the variants using supervoxel image representation result in lower fractions of negative determinant of the Jacobian of deformations, as well as lower standard deviations of the determinant of Jacobian of deformations. To show the influence of the filtering method on the results, experiments were conducted using both Gaussian filtering and guided image filtering. When supervoxel-based image representation was used, for Gaussian filtering the achieved TRE was 1.36mm, whereas for the proposed guided image filtering method it was 1.16mm. A similar trend was observed when both filtering methods for the regular grid image representation were applied, resulting in TRE of 1.87mm and 1.36mm for Gaussian and guided image filtering respectively. Guided image filtering has been performed with the fixed imaged as the guide. The results are summarised in Table 3.1. The visualisation of the influence of the filtering methods is shown in Figure 3.4, where the proposed method with guided image filtering is presented, as well as the proposed method with Gaussian filtering and, finally, regular grid with both filtering methods. The results presented in Table 3.2 show that the application of guided image filtering results in lower fractions of negative determinant of the Jacobian of deformations as well as lower standard deviations of the determinant of the Jacobian of deformations, regardless of image representation. The best overall results are achieved for the proposed method using supervoxel-based image representation and guided image filtering.

### 3.3.5 Graph Relaxation

We performed experiments on the graph created by all adjacent supervoxels (subsequently in the chapter called full graph), as well as after performing graph relaxation with threshold parameter $\upsilon$ set to 0.1, by removing edges according to Eq. 3.3. The

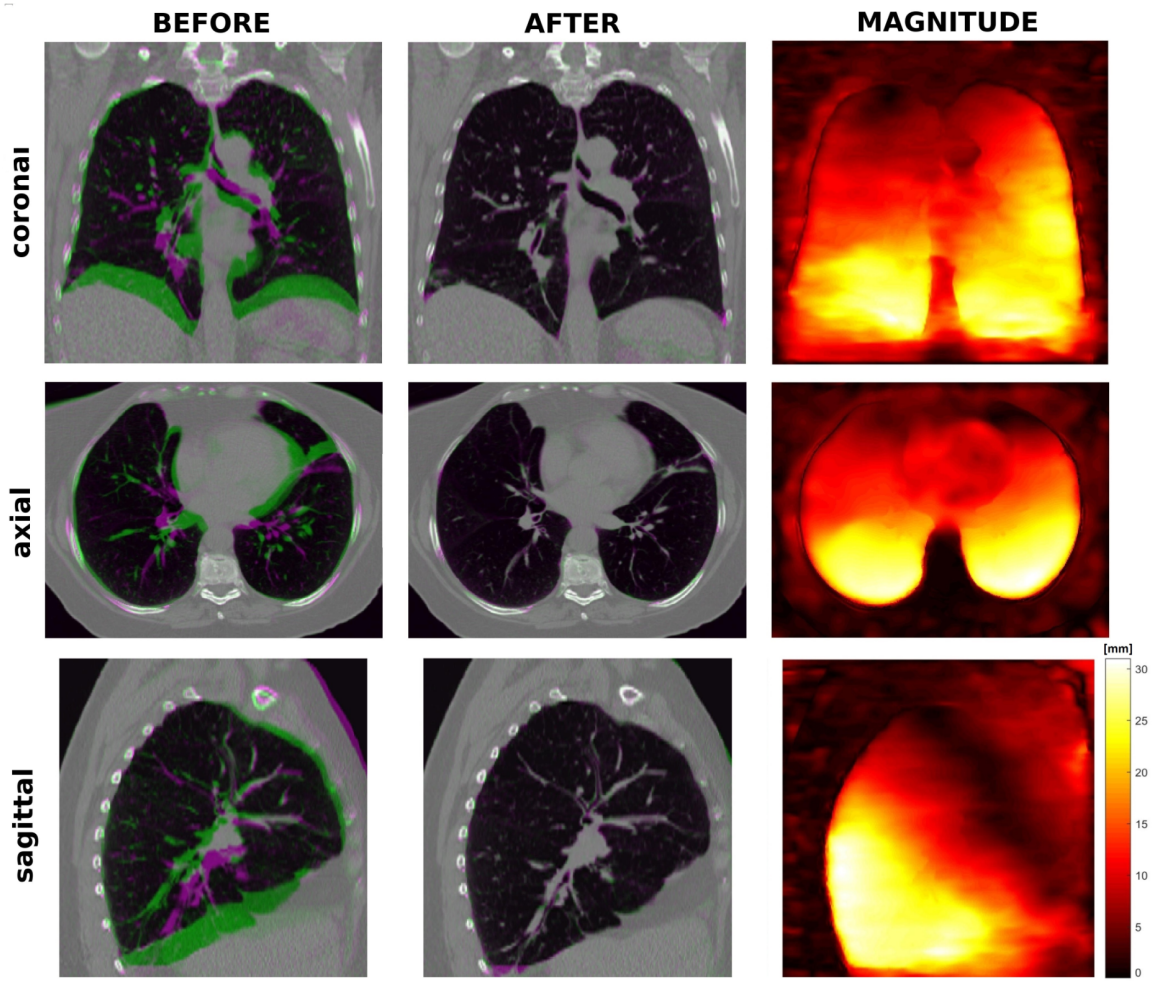



Figure 3.6: Registration results after applying the proposed method to the most challenging Dir-Lab case, (Case 8), showing the axial, coronal and sagittal views. In the left column an overlay of full inhale (green) and full exhale (magenta) before applying the registration is presented. In the middle column the same slices, after performing the proposed registration, are shown. The initial misalignment, visible in the left column, is almost completely eliminated by the proposed registration method. In the right column the magnitude of the displacement field in mm is presented.

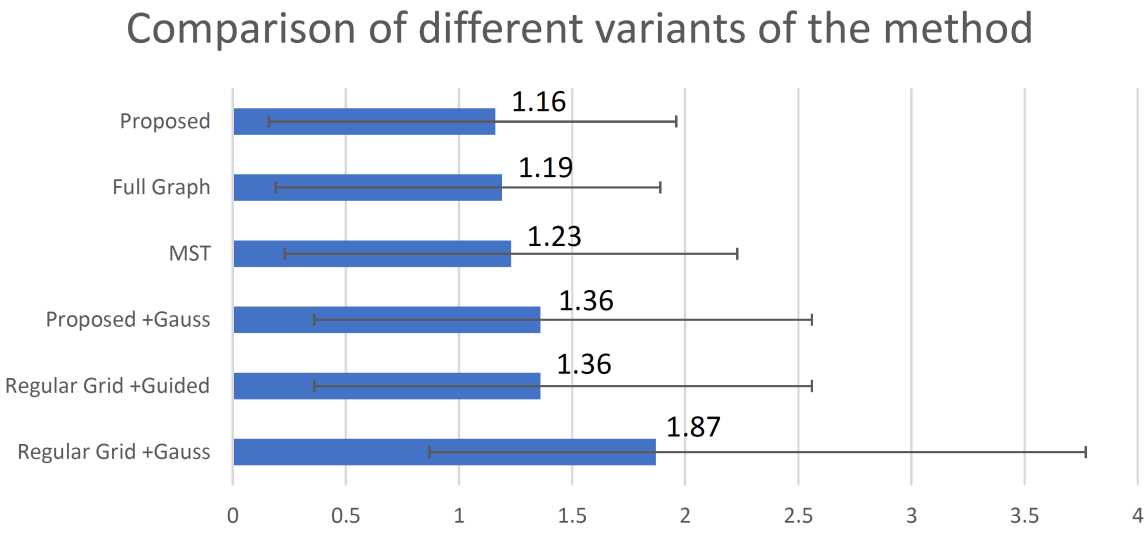



Figure 3.7: Comparison of different variants of the method based on TRE

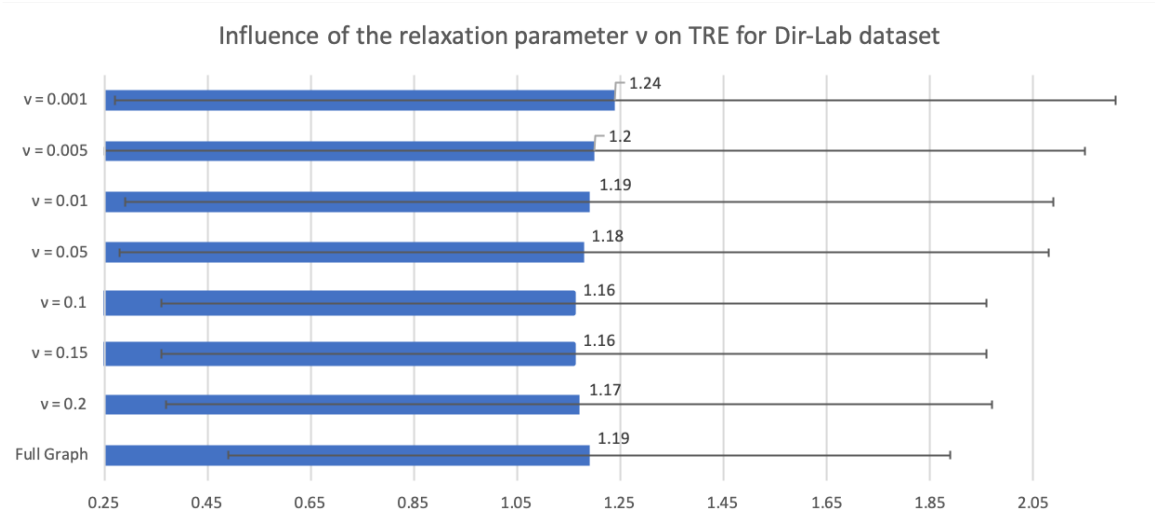



Figure 3.8: Influence of the relaxation parameter $v$ on TRE in [mm] and standard deviation for the DIR-Lab dataset.

percentage of edges removed by the $\upsilon$ threshold depends on its value, which influence on the final registration is evaluated in the following section. In general, if value of $\upsilon$ is close to 1 most of the edges are preserved, when it decreases more edges are removed. For instance for $\upsilon = 0.15$ on average approximately 50% of the total number of edges in the graph are removed. This value differs between the resolution levels, as for the initial, coarser levels differences between the neighbouring supervoxels are more pronounced and more edges are removed (approximately 60%). For the following finer levels the percentage of removed edges decreases reaching approximately 40%. The results for the three first cases were identical for both variants in terms of TREs. However, apart from case 10, where the full graph variant outperformed its relaxed counterpart, for the remaining six cases the TREs for the relaxed graph representation were lower. An additional experiment was conducted, where, akin to [81], an MST was extracted from the full graph. The optimisation was performed using the same energy formulation and parameter settings as for the full graph and its relaxed version before their performances were compared. The numerical results are presented in Table 3.3, where all three variants of the method are compared: the full graph approach, proposed relaxed graph representation, and minimum spanning tree relaxation - MST. The influence of the relaxation parameter $\upsilon$ on TRE was investigated by comparing TREs calculated using the full graph representation as $\upsilon$ is gradually changed. The results presented in Figure 3.8 show that the best performances of TRE of 1.16mm were achieved for $\upsilon = 0.15$ and $\upsilon = 0.1$. Further relaxation of the graph deteriorates TRE for the Dir-Lab dataset. The differences in the TRE values presented in Figure 3.8 for a range relaxation parameter $\upsilon$ settings clearly indicate for a minimum in the TRE for the parameter in range 0.15-0.1. These differences, however, do not show statistically significant improvement over the other settings ($p > 0.05$). Therefore, it is not possible to state that they will introduce meaningful changes in a clinical context. The visualisation of the registration re-

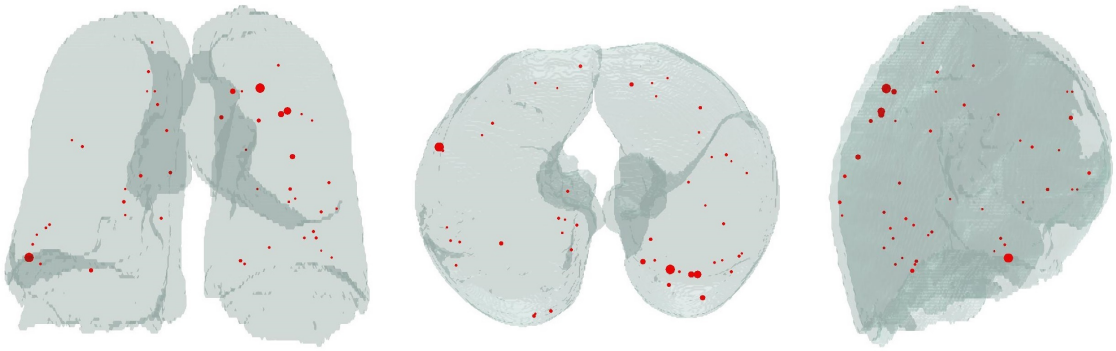


Figure 3.9: Visualisation of the landmark error distribution for the most challenging case 8 after registration using the proposed method in all three views with the lung mask in the background. Only landmarks with an error higher than 1.5mm are displayed. The size of the points corresponds to the landmark error.

sults after applying the proposed method, for the most challenging Dir-Lab case 8 is shown in Figure 3.6. The distribution of the landmark error for the same case after registration using the proposed method is presented in Figure 3.9.

### 3.3.6 Sliding Motion

To explicitly show how the proposed method deals with sliding motion, we calculated the maximum shear stretch of the deformation field, which can be used as a measure of sliding motion. Following the method presented in [158], we calculated the gradients of the deformation field with respect to the x, y and z directions, using a similar approach as that used to calculate the Jacobian of the deformations. Subsequently, the deformation field gradients were decomposed into eigenvalues, which were then used to find the maximum and minimum principal stretch components, $\gamma_1$ and $\gamma_3$, respectively. Then the maximum shear stretch is defined as: $\gamma_{max} = \frac{1}{2}(\gamma_1 - \gamma_3)$. The visualisation of the maximum shear stretch distribution on a logarithmic scale is shown for the example of case 8 in Figure 3.10. The highest value of shear stretch, namely $\gamma_{max} > 5$ appears almost exclusively at lungs borders, indicates a considerable sliding motion at the borders between the lungs and rib cage as well as the diaphragm

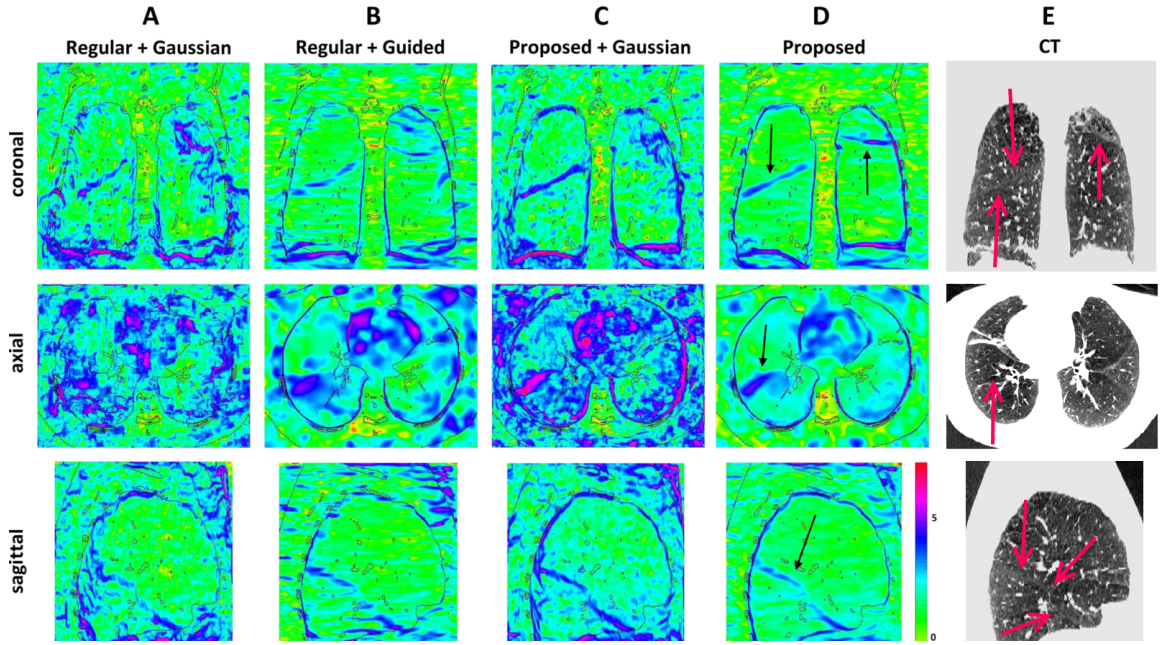



Figure 3.10: An example of maximum shear stretch calculated for the most challenging case 8 shown in the coronal, axial and sagittal views. The results for different variants of the proposed method were compared, presenting the regular grid with Gaussian filtering in the column A, the regular grid representation with guided image filtering in the column B, supervoxels-based image representation with Gaussian filtering in the column C and on the column D the proposed method, where supervoxel image representation with guided image filtering was used. Fissures, indicated by black arrows between the lung lobes, indicate a fair amount of sliding motion taking place inside the lungs. In column E the corresponding anatomical CT images are shown, from which the main contours were propagated in the previous images. The CT images are shown with adjusted intensity ranges to help visualise the internal structures of the lungs. Additionally, the fissures between the lung lobes are pointed with red arrows, which correspond to black arrows from column D.

and the liver. It is also clearly noticeable (as indicated by black arrows in Figure 3.10) that the value of shear stretch between the lung lobes is equal or slightly below 5, which implies interlobar sliding motion. This observed sliding motion inside the lungs corresponds with lung lobe boundaries visible in column E, where CT images with adjusted intensity value window to highlight them. Even though the clearly different intensity values for distinct lung lobes are more pronounced between left upper and lower lobes, the separation can be observed between most of the lobes in the lungs. One explanation for this might be that supervoxels when extracted followed the fissure separating the lobes, which followed by guided image filtering helped to emphasise differences in their motion during breathing.

### 3.3.7 Comparison with Other Methods

The proposed method is compared with other, closely related methods in the field, based on the TRE reported for the Dir-Lab data set. These are: continuous optimisation-based demons method enhanced with bilateral-filtering (BLF [64]), dense displacement sampling method using belief propagation on a minimum spanning tree (deeds [81]) and supervoxel-based with belief propagation method applied over a minimum spanning tree (SBP [83]), where the minimum spanning tree is extracted from a graph created by the adjacent supervoxels and the optimisation performed over a number of layers of supervoxels. The proposed method achieved a mean TRE of 1.16mm, compared to 1.96mm for BLF, 1.43mm for deeds and 1.23mm for SBP. For all of the cases the proposed method achieved the lowest TRE, except in case 8, where the best result was achieved with the SBP method. The numerical results are presented in Figure 3.11. The results achieved for the proposed method show statistically significant improvement compared to those using BLF and deeds (p-value $<0.05$). However, no statistical significance was found with the results for SBP.

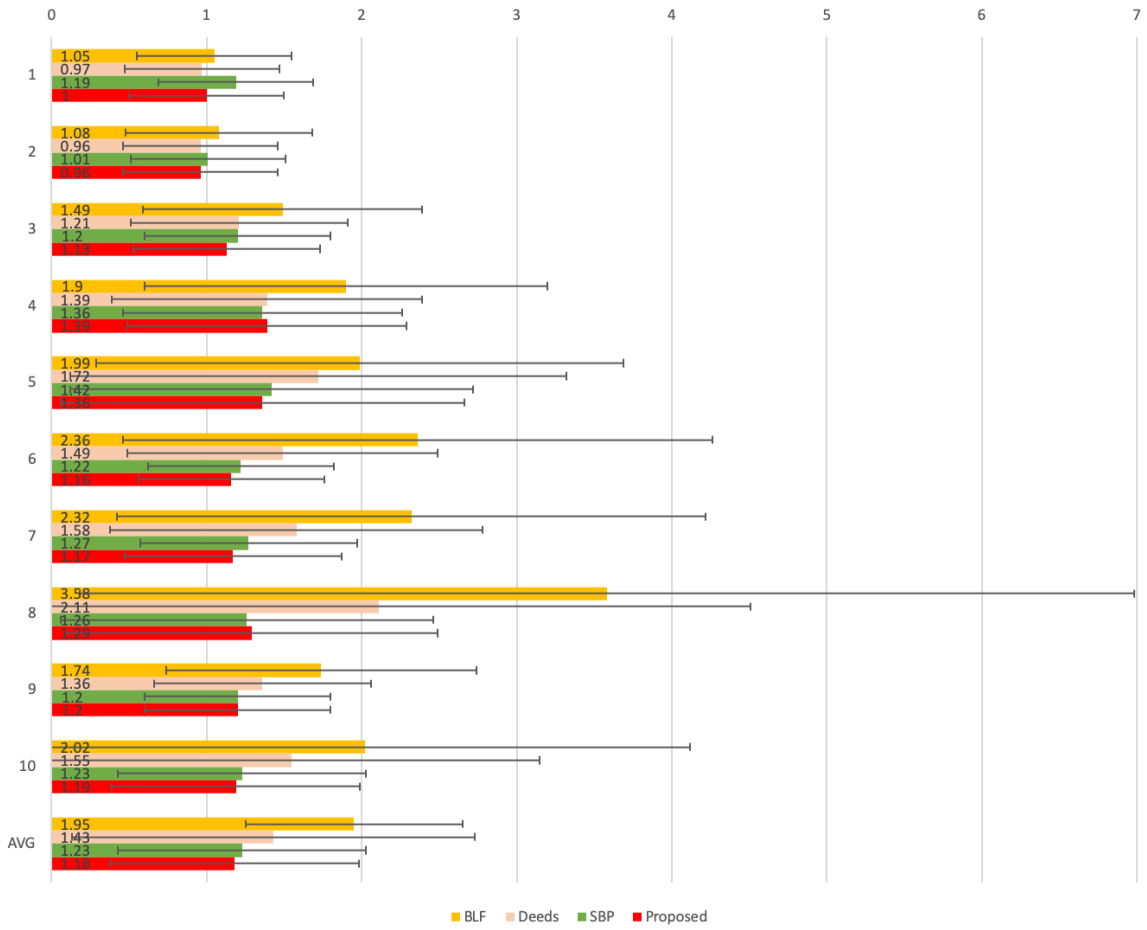



Figure 3.11: Comparison of the proposed method with the closest implementations based on the mean TRE for Dir-Lab data set. For each case, the TRE after registration are shown for: the proposed method, bilateral-filtering combined with demons (BLF [64]), the dense displacement sampling method based on belief propagation on a minimum spanning tree (deeds [81]) and supervoxel-based belief propagation (SBP [83]). Mean results show that the proposed method outperforms all of the compared methods on average and in all individual cases apart from case 8.

## 3.4 Discussion

The novel deformable image registration method proposed in this chapter successfully combines compact image representation via supervoxel clustering with an efficient graph cuts optimisation method and guided image filtering, providing a locally continuous displacement field. The justification for the chosen approach was shown by presenting at first the results for a regular grid image representation variant, in which cubic supervoxels were extracted. Even though the compactness parameter $m$ for SLIC was set to a relatively high value (4000), it is possible that some of the the supervoxels were imperfect cubes. Nevertheless, this only emphasises the fact that a perfectly regular cubic representation would further deteriorate the results. The variant of the method using regular grid image representation with Gaussian filtering achieved a TRE of 1.87mm, which is the worst result among all of the compared variants and seems to be the least suitable configuration for the application. After using supervoxels during the image representation stage, and maintaining Gaussian filtering, TRE is reduced to 1.36mm; this method, then, results in the largest change in TRE, manifesting the superiority of supervoxel-based image representation over a regular grid. While the relative difference between the regular grid and the supervoxel methods, compared to the initial TRE of 15mm, may not seem large, visual assessment of the magnitude of the displacement field shown in Figure 3.4 illustrates to what extent this is advantageous in practice. The displacement field for regular grid is heavily distorted at the borders of the lungs and so the lung shape is not recognisable, even though they would be expected to be locally consistent in their deformations. This clearly indicates that the supervoxel-based image representation has superior properties over its regular grid counterpart. Surprisingly, the supervoxel image representation with Gaussian filtering variant achieved the same TRE as the regular grid method with guided image filtering (1.36mm). This suggests that the guided image filtering method is more suitable than Gaussian filtering in this appli-

cation for image registration tasks and plays a role in additional structure oriented regularisation, which can even partially compensate for regular grid image representation. These observations strongly support the approach outlined in the proposed method. Good results for the minimum spanning tree-based version were expected, based on the previously reported results in the literature [81, 83]. The results position the MST-based method between the regular grid/Gaussian filtering versions and the proposed method. Coincidentally the same average TRE (1.23mm) is achieved in our test and the one reported in [83]. However, the specific cases of the dataset achieved different TREs. The second best result was achieved using the proposed version in a full graph representation (TRE of 1.19mm), and these results were further improved by using its relaxed variant (TRE 1.16mm). The relaxed graph representation provided improved or at least equally accurate results for all cases except for case 10. In the proposed approach the parameters were tuned using just one of the cases of the dataset - case 8, which is the most challenging and has the largest initial misalignment. Such an approach is more robust to over-fitting the parameters. As a drawback it is possible that such an approach will result in suboptimal parameters settings for one (or more) of the cases. A larger dataset could potentially result in more accurate parameter tuning. The visualisation of the results for different variants of the proposed method in Figure 3.7 shows the gradual improvement of the results. An additional experiment investigated the influence of the intensity term in the piecewise smoothness term of the energy formulation in Equation 3.6. The substitution of the proposed linear intensity term with a quadratic one resulted in TRE of 1.22mm, compared to 1.16mm for the linear term. This experiment supports the proposed formulation of the energy in Equation 3.6. In terms of the algorithm's run time, the performance of the proposed method in its current suboptimal and non-parallelised Matlab implementation varied from 63 min to 121 min per pairwise registration, depending on the Dir-Lab case.

Since the optimisation of some parameters for the presented method was performed on all data, there might be a possibility of over-fitting of the applied parameters to DirLab dataset. However, the point of the presented analysis was not to provide a perfect set of parameters, which would be suitable for every dataset. Medical image analysis is a particularly demanding field, where imaging datasets originating from different centres or acquired for different diseases might vary. For instance lung CT might be performed with different reconstruction kernels, at breath-hold or free breathing, which will have significant influence on the characteristics of the data. The common practice is to perform fine tuning of the parameters in order to achieve the best results of the particular dataset, with their specific characteristics. It is possible that for some datasets the optimal set of parameters of the proposed method might be slightly different, however the method was tested in a similar way that other methods evaluated on DirLab dataset and has shown good performance.

The methods we used for comparison were chosen to be closest to the proposed approach; in addition, these selected methods are considered to be state-of-the-art in lung image registration. The authors of the work [43], for instance, did not use the Dir-Lab dataset for their validation, making a direct comparison a difficult task. Instead, the results were compared with [64], which, like the work presented in [43], uses the log-demons optimisation strategy. The authors of [89], which can be seen as an extension of the work presented in [88], reported the computation time for a single 3D registration to be up to 24 hours. In comparison, our proposed method requires a run time of around 1-2 hours, depending on the volume size; specifically, for images of similar volume to those presented in [89], a run time of slightly over 1 hour was recorded. The methods proposed in [91, 92] have been presented for 2D cases only. While these potentially can be extended to 3D cases, this would have required significant code development, which was beyond the scope of this work. The best performing method on the Dir-Lab dataset to date is described in detail

in [159]; this method results in a TRE of 0.94mm on average. This method and other top performing methods on the Dir-Lab dataset use segmentations and perform registration only inside the masks, as all the landmarks are placed inside the lungs. Such an approach makes the methods unsuitable for direct comparison.

The proposed method follows the trend noticeable in image registration field where discrete optimisation-based methods perform very well and in some cases favourably compared to their continuous optimisation-based counterparts. The proposed method creates a promising framework for more advanced applications, such as multi-modal image fusion, where images originating from different modalities can be combined together. Another interesting application of the method could be a joint lung registration and lobe segmentation. The borders between the lobes appear to be well-distinguished, based on the deformation field, as show in Figure 3.10.

As the presented method does not require any input segmentations, it can potentially be generalised and applied on other organs. Related work, which applied supervoxels based regularisation within Demons framework, has been successfully applied to liver CT image registration [160]. The application to the liver image registration might be also relevant due to the sliding motion occurring at the surface of the organ. The method shows potential to be applicable to other medical image registration where sliding motion appears, such as multi-modal lung registration, as well as cardiac motion correction.

It is very important to establish the clinical relevance of the achieved results. In routine radiotherapy practice, average set-up errors are reported at 1-5mm level for photon therapy [45]. However, in the case of Stereotactic Ablative Radiotherapy (SABR), 0mm (understand as subvoxel) accuracy is required. This type of treatment is becoming increasingly popular in lung cancer and has established itself as a standard of care for early stage lung cancer in patients who are not medically operable (due to not being fit for the procedure). Another type of lung cancer treatment where

high level of accuracy plays a crucial role is proton beam therapy. Due to its Bragg peak shape of the dose-depth distribution, which allows for precise dose delivery to the target volume, at the same time limiting exposure of the tissue on its path, this type of treatment has become a very promising alternative to photon therapy. The rapid decline in the radiation of the proton beam after the target location is one of the key advantages of this method of radiotherapy. However, its drawback is that it requires precise planning and delivery, to make sure that the tumour will receive the prescribed dose and that healthy tissue will not be exposed. Therefore, the accuracy plays an even more important role than in the photon therapy and minimising set-up uncertainties is essential. Delivery of the doses (by photons or protons) to the healthy tissue is currently one of the limiting factors in radiotherapy. Approximately 2% of patients undergoing photon therapy die due to radiation induced lung damage. It has also been shown that increased dose delivery to the heart strongly correlates with higher mortality [161]. Recently it has been modelled that reducing the MR systematic tracking error to 0.2mm for MR-Linac might lead up to 14% gain in disease free survival for lung cancer patients [162]. Given this, the subvoxel accuracy achieved with the proposed methods are clearly relevant to the radiotherapy planning.

The application of the image registration results not only in the images getting aligned, but also in the estimation of the deformation fields. The estimated deformations may be further used, for instance, for tracking segmented organs in dynamic images, to create motion models or to estimate ventilation surrogates. The last application is particularly important for lung cancer radiotherapy planning. The ability to localise the well functioning regions of the lungs can be used to adapt radiotherapy plans and spare them from radiation. Such an application is presented in the following Chapter 4, where a novel approach for estimating ventilation surrogates based on 4DCT is proposed.

# Chapter 4

# Patch-based Lung Ventilation Estimation using Multi-layer Supervoxels

The deformation fields acquired from CT-CT image registration, as presented in the previous Chapter 3, might be further used to estimate lung ventilation surrogates. This chapter presents a novel method for lung ventilation estimation from dynamic lung CT imaging, combining a supervoxel-based image representation with deformations estimated during deformable image registration, performed between peak breathing phases. For this we propose a method that tracks the intensity of previously extracted supervoxels.

Patch-based approaches have received substantial attention over the recent years in medical imaging. They have been successfully applied to enhance image segmentation [163, 164], as well as to tumor segmentation [165, 166, 167] or analysis of its texture [168]. Other applications are patch-based descriptors for image similarity measure [55, 169] and patch-based image registration approaches [170, 83, 136]. Their potential applications is to provide more anatomically consistent ventilation

maps estimated on dynamic lung CT. An assessment of regional lung function may act as a guide for radiotherapy, ensuring a more accurate treatment plan. This, in turn, could spare well-functioning parts of the lungs. Correlation between the estimated ventilation maps with static ventilation images acquired from hyperpolarised $^{129}$Xenon MRI (XeMRI) is calculated for the evaluation of the method. The influence of different image registration methods used to estimate deformations between the peak breathing phases in the dynamic CT imaging is investigated. The presented method performs favourably to other ventilation estimation methods commonly used in the field, independently of the image registration method applied to dynamic CT. Due to the patch-based approach of our method, it is physiologically more consistent with lung anatomy than previous methods, which rely on voxel-wise relationships. In the proposed method the ventilation is estimated for supervoxels, which tend to group voxels that are in spatial proximity and have similar intensity values. The proposed method was evaluated on a dataset consisting of three lung cancer patients undergoing radiotherapy treatment, and this resulted in a correlation of 0.485 with XeMRI ventilation images, compared with 0.393 for the intensity-based approach, 0.231 for the Jacobian-based method and 0.386 for the Hounsfield units averaging method, on average. Within the limitation of the small number of cases analysed, results suggest that the presented technique may be advantageous for CT-based ventilation estimation. The results showing higher values of correlation of the proposed method demonstrate the potential of our method to more accurately mimic the lung physiology.

## 4.1 Background

The gold standard test for pulmonary function is the spirometry technique, which evaluates global lung function by measuring the volume and speed of air during in-

halation and exhalation. Although this method is cheap, fast, and widely used as a screening tool [15], it lacks regional lung evaluation and thus cannot be used for precision medicine such as individualised radiotherapy planning [126]. Assessment of the regional lung functionality has the potential to provide more accurate radiotherapy treatment [11], which could spare well-functioning parts of the lungs and be used for follow up to evaluate the treatment effectiveness. This chapter presents one potential application of patch-based methods: to estimate regional ventilation of the lungs based on dynamic lung CT (4DCT).

The estimation of ventilation surrogates based on CT images is an active field of research. In [110], for instance, the Dice overlap of percentage ventilation volumes between 4DCT originating ventilation estimates and $^3$Helium MRI (HeMRI) was investigated, and no significant difference was found. In that study, the ventilation images, both 4DCT originating ventilation surrogates and HeMRI, were divided into 10 percentile ranges (0-10% , 11-20%, and subsequently to 91-100%). Such preprocessing of the images allowed the authors to calculate Dice overlap between corresponding percentile ranges between the estimated ventilation surrogates and HeMRI. Their findings support the claim that well-ventilated regions in 4DCT-based ventilation surrogates and HeMRI are in agreement. A lobar oriented analysis of lung ventilation was presented in [171], comparing CT-based ventilation surrogates with HeMRI for asthma patients. The authors found good agreement between the investigated ventilation imaging methods. An interesting study was presented in [132], where, apart from investigating the correlation between different CT-based ventilation estimation techniques and hyperpolarised gas MRI ventilation images, the correlation between HeMRI and XeMRI was evaluated. They found significantly lower correlation between CT originating ventilation surrogates and hyperpolarised gas MRI ventilation maps than between HeMRI and XeMRI. In order to search for the best indicator of lung function, the authors of [172] investigated correlation between ventilation esti-

mated from XeMRI and the amount of gas transferred to red blood cells (RBCs). Their initial conclusion was that XeMRI ventilation may not be the optimal measure of the true regional lung function. The study was conducted on 17 patients; however only three of them were lung cancer patients. Despite the amount of work done in the field, there is still scope for further research in the methods dedicated to estimating ventilation from CT images.

The original idea of this chapter and preliminary results were presented in [173] and further extended to a journal publication [174]. The remaining part of this section is organised as follows: Sec. 4.1.1 presents a recapitulation of lung ventilation imaging methods and their application for ventilation imaging. Subsequently, in Sec. 4.1.2, lung image registration methods, chosen for the comparison when estimating lung ventilation surrogates, are discussed.

### 4.1.1 Lung Ventilation Imaging

Lung ventilation imaging allows for insight into local lung function. Most of the lung ventilation imaging techniques require the use of ionising imaging methods and radioactive tracers, such as SPECT [175] or PET [176]. A lung function imaging technique that has received substantial attention in the field is hyperpolarised noble gas MRI [30]. Its non-ionising nature makes it a promising technique for imaging ventilation, perfusion and gas transfer within the lungs [33]. In the static approach of this method, the patient breathes in a fixed volume of hyperpolarised noble gas and images are acquired during breath hold. Static hyperpolarised XeMRI provides images with direct visualisation of the distribution of the inhaled gas in the patient's lungs. Compared with the aforementioned lung functional imaging techniques, XeMRI is acquired during a single breath-hold. To allow for the localisation of ventilation images with respect to the patient anatomy, pMRI is acquired immediately before XeMRI acquisition. The aforementioned techniques have already been introduced in detail

in Chapter 2.

4DCT, on the other hand, is a structural imaging technique routinely used for radiotherapy planning, where images are acquired during free breathing under spirometry control. This can be used to estimate lung ventilation maps, based on the results of the registration between peak-exhale and peak-inhale breathing phase volumes. 4DCT based ventilation estimation methods can be roughly classified into those which rely on the value of the determinant of the Jacobian of the deformations to estimate the lung volume change [128, 129], methods tracking the changes in the intensity of the lungs, quantified in Hounsfield units (HU) [127, 126], and based on averaging a product of air and tissue fraction across all breathing phases [122]. Changes in tissue attenuation correspond to variations in tissue density and originate from filling the lungs with air. The Jacobian-based and intensity tracking methods require estimation of the deformation field between peak breathing phases from 4DCT, performed by deformable image registration. In contrast, the averaging method [122] does not need to estimate deformations. All the ventilation surrogates estimation techniques, mentioned above, have already been presented in detail in Chapter 2 of this thesis.

### 4.1.2 Lung Image Registration

Most of the ventilation estimation approaches require estimation of the deformation field between peak exhale and peak inhale lung volumes from 4DCT. To investigate the influence of the image registration method on the estimated ventilation maps in the proposed approach, a number of different methods were used: the deeds method [81] (referred to as deeds), classic demons method [60] (referred to as demons), its extended version applying bilateral filtering (BLF) [64] (referred to as demons BLF), and supervoxel-based graph cuts [137] (referred to as supervoxel graph cuts). The first three methods have been introduced and discussed in Chapter 2, whereas the supervoxel graph cuts were proposed in the previous Chapter 3.

All of the methods have shown good performance in the publicly available Dir-Lab dataset [119], with 300 manually annotated landmarks provided for evaluation purpose. Applying the deeds method resulted in a Target Registration Error (TRE) on the dataset of 1.43mm, for demons and their BLF extension TREs of 2.35mm and 1.95mm were measured, respectively, and for the supervoxel graph cuts TRE was 1.16mm. In this study, the methods were applied with the parameter settings provided when evaluated on Dir-Lab, as this dataset also consists of 4DCTs. These parameters were set completely independent of our testing dataset. Our choice of the methods covers the main registration categories in terms of optimisation, namely: discrete, continuous, and hybrid.

The remaining part of this chapter is organised in the following way: Sec. 4.2 introduces the proposed method for estimating ventilation in a patch-based manner. The results of the proposed experiments are presented is Sec. 4.3, which are finally discussed and concluded in Sec. 4.4.

## 4.2 Methods

This section presents the proposed supervoxel-based lung ventilation estimation method. The applied image clustering method, as well as how this method is incorporated to the proposed approach is explained in Sec. 4.2.1. Sec. 4.2.2 introduces the applied image registration framework, which brings the estimated ventilation maps from 4DCT to align with XeMRI ventilation images. A work-flow diagram presenting the method is shown in Figure 4.1. For illustrative purposes the superpixels are extracted from a 2D image, whereas in the actual method 3D supervoxels are applied.

Figure 4.1: The diagram shows the work-flow of the proposed approach for patch-based lung ventilation estimation. The first step is to perform image registration between peak breathing phase volumes. This is followed by clustering the peak exhale image into supervoxels. A number of layers of supervoxels is extracted, which is depicted by multiple boxes for the current and following steps. The previously estimated deformations are applied to the supervoxels. The original supervoxels extracted from the peak exhale image are interpolated and labelled using the same colour. The warped supervoxels are mapped over the peak inhale image to calculate changes in their mean intensity between the mapping over the peak exhale and peak inhale breathing volumes. Finally, the results obtained for all the supervoxel layers are averaged.

### 4.2.1 Supervoxel Tracking for Ventilation Estimation

One of the approaches for patch-based methods is supervoxel-based image representation. Supervoxel-based image clustering groups voxels that are in spatial proximity and are visually similar into larger structures. They have become the building blocks for many applications in computer vision and medical image analysis. The presented formulation applies Simple Linear Iterative Clustering (SLIC) [146] as a clustering method, due to its speed of performance and direct control over the number of extracted supervoxels and supervoxel compactness.

The existing methods [122, 126, 130, 128] for estimating ventilation from 4DCT were described and reviewed in Chapter 2 of this thesis. The presented methods take the structural relationship between voxels into account mainly at the image registration step or calculation of derivatives. Work presented in this chapter applies the proposed patch-based method, which additionally addresses this issue and estimates ventilation for larger, spatially and visually consistent regions. Such an approach may be more meaningful from an anatomical point of view. The lungs are divided into right and left lungs, which are further separated into lobes, while each lobe consists of a number of segments. All of them are supplied with separate airways which form a tree-like structure with no additional interconnections between the subtrees. Therefore, the ventilation of the neighbouring lobes or sections supplied by different branches of the airway tree may differ. Supervoxels have the potential to cluster anatomically consistent lung regions and, therefore, ventilation surrogates would not be estimated across anatomical borders of the lungs. This might not be a huge advantage in the case of patients with healthy lungs, where no pathological parenchyma changes are present. However, as this method is mostly dedicated for patients undergoing lung cancer radiotherapy treatment, it is a fair assumption that along with lung cancer, other pathological changes in the lungs might be present. For instance COPD is a common co-existing condition with lung cancer, where damaged lung tissue can

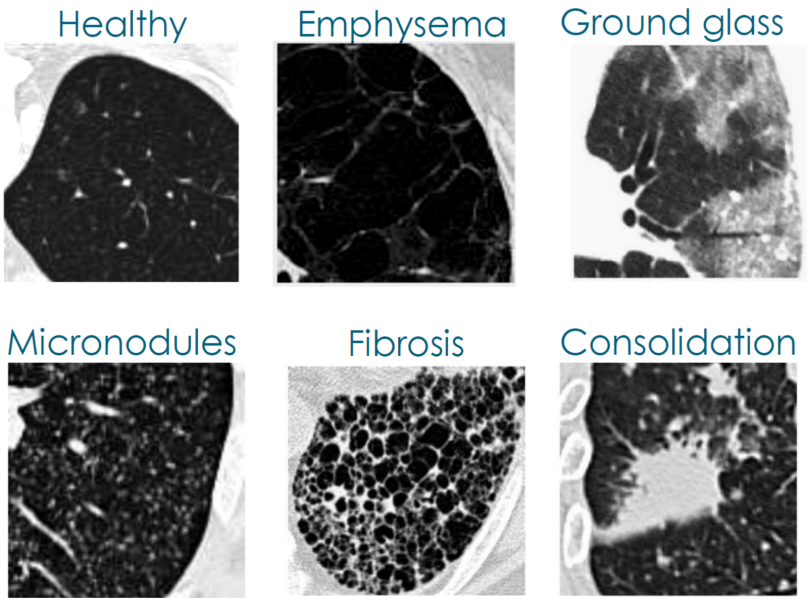



Figure 4.2: Examples of different lung tissue patterns. In the upper row first from the left a healthy lung tissue pattern is shown. The following images show emphysema, ground glass opacity, micronodules, fibrosis and consolidation [177].

be distinguished on lung CT scans. The damaged tissue has lower intensity compared to healthy lung tissue due to lower tissue density and air-like intensity profile. In contrast, fibrosis, ground glass opacity and consolidation will present higher intensity than healthy tissue [177]. The illustration of different types on lung tissue patterns is presented in Figure 4.2. In most of the previous studies ventilation images were estimated for peak exhale [130]. However, this work used the peak inhale breathing phase as a reference, due to the difference in the volume of the lungs between pMRI and CT then being the smallest. Such an approach should result in smaller deformations applied during pMRI to CT registration and, therefore more accurate results of the registration. A similar approach has been recently applied in [132] to address the same challenges. The estimate of the ventilation for peak inhale is formulated as follows:

$$Vent^{inh}_{Dens}(\mathbf{x}) = \left( \frac{HU_{exh}(\mathbf{x} + \mathbf{u}_{exh}) * G_{k1} - HU_{inh}(\mathbf{x}) * G_{k1}}{HU_{inh}(\mathbf{x}) * G_{k1} + 1000} \right) * G_{k2} \tag{4.1}$$

with $HU_{exh}(\mathbf{x})$ denoting the lung intensity value from peak exhale breathing phase image for voxel position $\mathbf{x}$ and $HU_{inh}(\mathbf{x}+\mathbf{u}_{inh})$ lung intensity value from peak inhale breathing phase image for the same voxel moved by displacement field $\mathbf{u}_{inh}$ and $G_{k1}, G_{k2}$ are Gaussian kernels used for smoothing and convolved (denoted by $*$ in the equation) with the images. All image intensities are given in HUs.

The first step is to extract supervoxels from the peak exhale breathing phase image and then calculate their mean intensities. Subsequently, the estimated deformations originating from image registration between peak exhale and peak inhale breathing volumes are applied to the extracted supervoxels. Such an approach helped mapping the same regions, represented by supervoxels, from the peak exhale to the peak inhale image (see Figure 4.1 ). Mean intensities of the transformed supervoxels are recalculated, this time over the peak inhale image. Then, this approach is applied to the voxel-wise intensity-based method, thus estimating ventilation for supervoxels rather than for each individual voxel:

$$Vent_{SLIC}(\mathbf{s}) = \frac{\overline{HU_{exh}(\mathbf{s}+\mathbf{u}_{exh})} - \overline{HU_{inh}(\mathbf{s})}}{\overline{HU_{inh}(\mathbf{s})} + 1000}, \tag{4.2}$$

where $\overline{HU_{inh}(\mathbf{s})}$ is the mean intensity of a supervoxel $\mathbf{s}$ extracted from peak inhale phase and $\overline{HU_{exh}(\mathbf{s}+\mathbf{u}_{exh})}$ stands for mean intensity of the same supervoxel $\mathbf{s}$ mapped to peak exhale breathing phase by deformation $\mathbf{u}_{exh}$. Then all voxels belonging to a particular supervoxels are assigned the estimated ventilation value from Eq. 4.2.

In the approach of [126, 130] (referenced to, henceforth, as 'classic approach'), images are preprocessed using Gaussian filtering to compensate for mis-registration or other image artefacts. Therefore the ventilation of each voxel is influenced by all of its neighbours within a certain spatial distance. The proposed approach estimates the ventilation surrogates for groups of voxels which are both in spatial proximity and have similar intensity values, such a formulation should more accurately correspond to anatomical structure of the lungs.

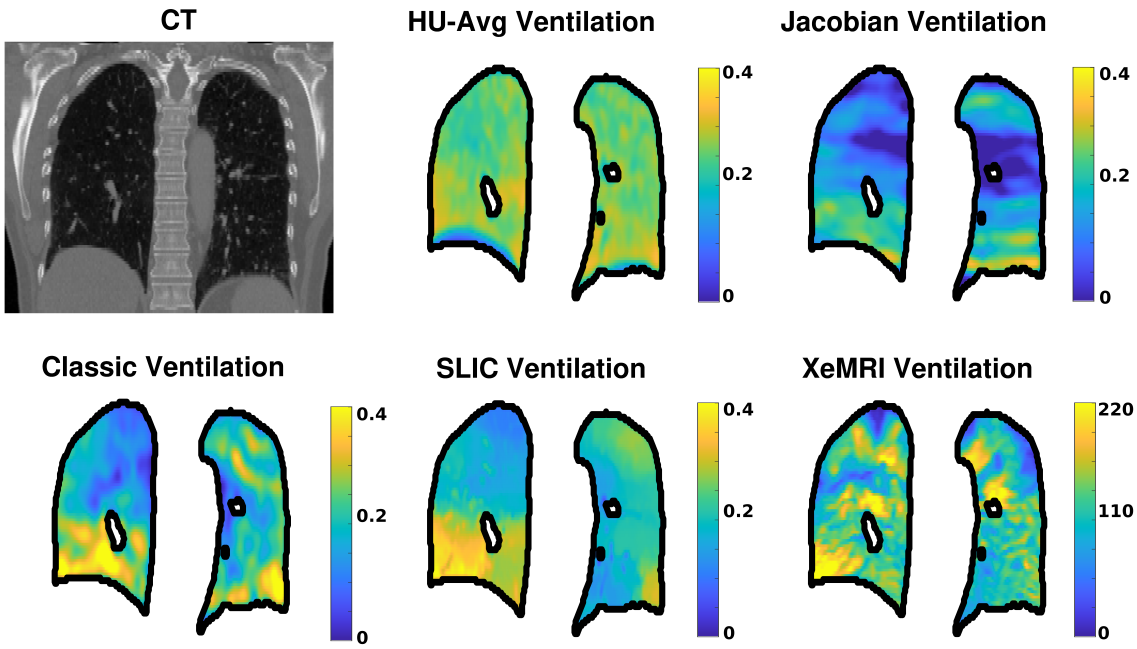



Figure 4.3: Comparison of the estimated ventilation maps using different methods for Patient 1 in coronal view. The results are presented for the supervoxel graph cuts image registration method. Lung contours extracted from the CT image are propagated onto ventilation images.

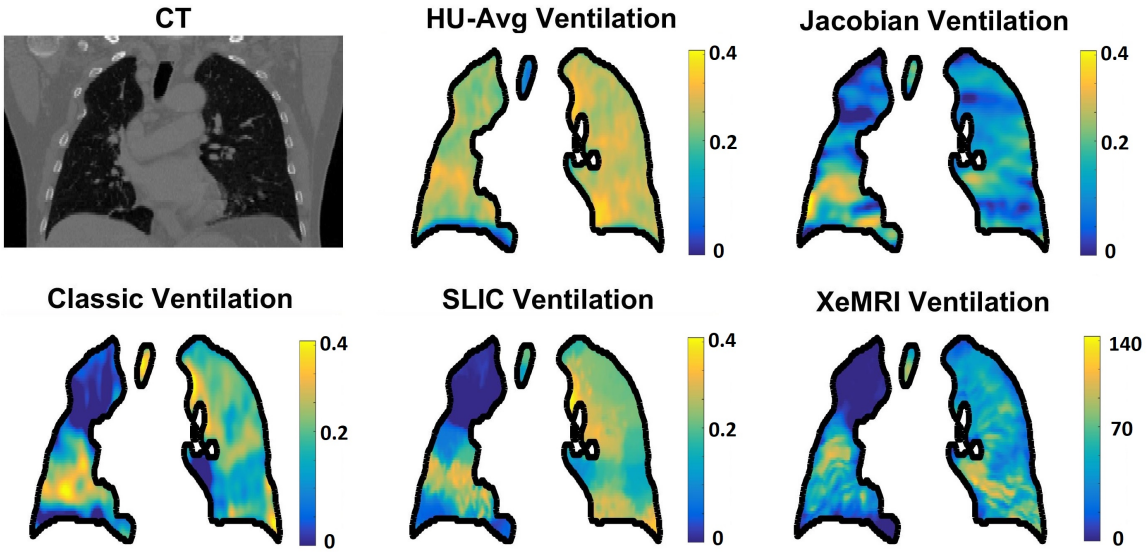



Figure 4.4: Comparison of the estimated ventilation maps using different methods for Patient 2 in coronal view. The results are presented for the supervoxel graph cuts image registration method. Lung contours extracted from the CT image are propagated onto ventilation images.

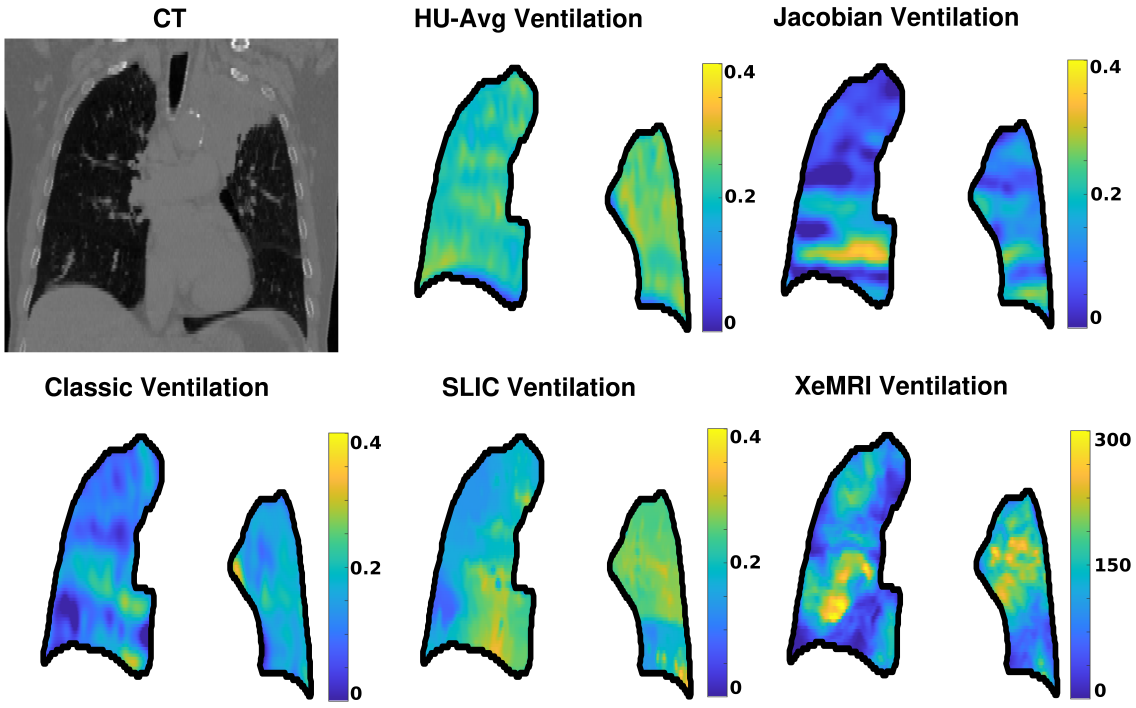



Figure 4.5: Comparison of the estimated ventilation maps using different methods for Patient 3 in coronal view. The results are presented for the supervoxel graph cuts image registration method. Lung contours extracted from the CT image are propagated onto ventilation images.

A particular clustering can have a strong influence on the results of the estimated ventilation maps. As there are no additional explicit constrains on the shape and size of the supervoxels, they may result in mis-segmentations. This issue is addressed by performing the extraction of several layers of supervoxels, akin to [83], where a similar approach has been applied to deformable image registration. Such an approach has been shown to improve the robustness of the method, as well as reduces the influence of the particular clustering and the need to add further constrains to the clustering. Therefore supervoxel clustering was performed several times, each time changing slightly the initial clustering parameters, which results in different clusterings. Then the final estimated ventilation for a voxel is calculated as the average over all the layers of supervoxels:

$$\overline{Vent}_{SLIC} = \frac{1}{n}\sum_{1}^{n} Vent^{n}_{SLIC}, \tag{4.3}$$

where $n$ is the number of layers of supervoxels. A visualisation of the ventilation maps from both methods and XeMRI for all patients is shown in Figure 4.3, 4.4, 4.5. To further compensate for the particular clusterings, Gaussian filtering with a small [7 7 7] kernel and $\sigma = 1$ is applied over the final $\overline{Vent}_{SLIC}$.

### 4.2.2 XeMRI to CT Alignment

In order to perform the evaluation of the correlation between 4DCT-based estimated ventilation maps and XeMRI ventilation images, these volumes need to be brought into alignment. As XeMRI ventilation images lack structural information, their direct registration with CT would be an extremely challenging task. This issue is addressed firstly by correcting for the misalignment between XeMRI and pMRI with a rigid registration method. Then pMRI to CT registration is performed. This task is divided into two steps: affine registration with mutual information as a similarity measure,

which is followed by deformable registration with the use of deeds method [81]. The reason why the method proposed in the previous Chapter 3 has not been applied here is that its performance is compromised when one of the volumes shows significantly lower spatial resolution. A further challenge is the extraction of supervoxels from such anisotropic data.

## 4.3 Experiments and Results

A number of experiments were conducted, in which different aspects of the estimated ventilation maps were investigated. The materials on which the experiments are performed are presented in Sec. 4.3.1. In the following Sec. 4.3.2 and Sec. 4.3.3 the influence of the supervoxels size and the number of layers of supervoxels are investigated, respectively. Then, in Sec. 4.3.4 the correlation coefficients between estimated ventilation maps and XeMRI is calculated. In Sec. 4.3.5 the spatial overlap based on ventilation percentile ranges is evaluated.

### 4.3.1 Materials

We performed the evaluation of the proposed method based on three cases (two male, one female, with ages ranging from 62 to 71 years old) consisting of 4DCT and XeMRI/pMRI data for patients undergoing radiotherapy treatment at Churchill Hospital in Oxford. The patients suffered from stage II lung cancer. The cancer staging for the patients was T2a-T3/4, N1-N2, and M0. TNM is a complex tumour staging description method, where T describes the size of the tumour and its location, N states whether the cancer has spread to the lymph nodes, and M tells about more distant metastasis of the cancer. In simplification, the patients included in our study had lung cancer tumours of size between 3-7cm (T2a-T4), with lung cancer cells being present in lymph nodes in the chest (N1-N2), however without any metastasis

to other lung lobes than the primary tumour has been localised or other parts of the patients body (M0). 4DCT was acquired on a GE Optima CT580 RT scanner with phase sorting. The patients lied supine on the couch with arms held above their head, and were coached on how to maintain regular tidal breathing. A mixture of $^{129}$Xe gas (80%) and air was polarised on-site to between 4% and 12%, by using a commercial polariser (model 1651; GE Healthcare, Milwaukee, Wisconsin) operating on the rubidium vapor spin-exchange optical pumping basis, and deposited in a Tedlar bag (Jensen Inert Products, Coral Springs, Florida). Hyperpolarised $^{129}$Xe was administered by first instructing the subjects, to exhale to FRC and then inhale the 1.0 litre contents of the Tedlar bag through 9.5mm inner diameter Tygon tubing (Cole-Palmer Instrument; Hanwell, London, England). Subjects were then instructed to hold their breath for up to 25 seconds for image acquisition [178]. The pMRI and XeMRI scans were obtained with a 1.5-T whole-body system (Signa HDx; GE Healthcare). The spatial resolution of the XeMRI/pMRI is 1.56x1.56x20 mm$^3$, whereas for CT it is 0.977x0.977x2.5 mm$^3$. All data were resampled to an isotropic resolution of 1x1x1 mm$^3$ size. The sequence of 4DCT consisted of 10 volumes of different breathing phases.

### 4.3.2 Influence of the Supervoxel Size

The size of the supervoxels is one of the factors that have an influence on the results of the estimated ventilation maps. On the one hand, supervoxels that are too large may not be capable of modeling the regional nature of the ventilation changes. On the other hand, if too small, the supervoxels may lose their structure-oriented properties and provide similar results to the voxel-wise approaches. The purpose of this experiment is to investigate how the size of the supervoxels influences the correlation of the estimated ventilation maps with XeMRI ventilation images, and if there is a local maximum indicating an optimal size of the supervoxels. To investigate

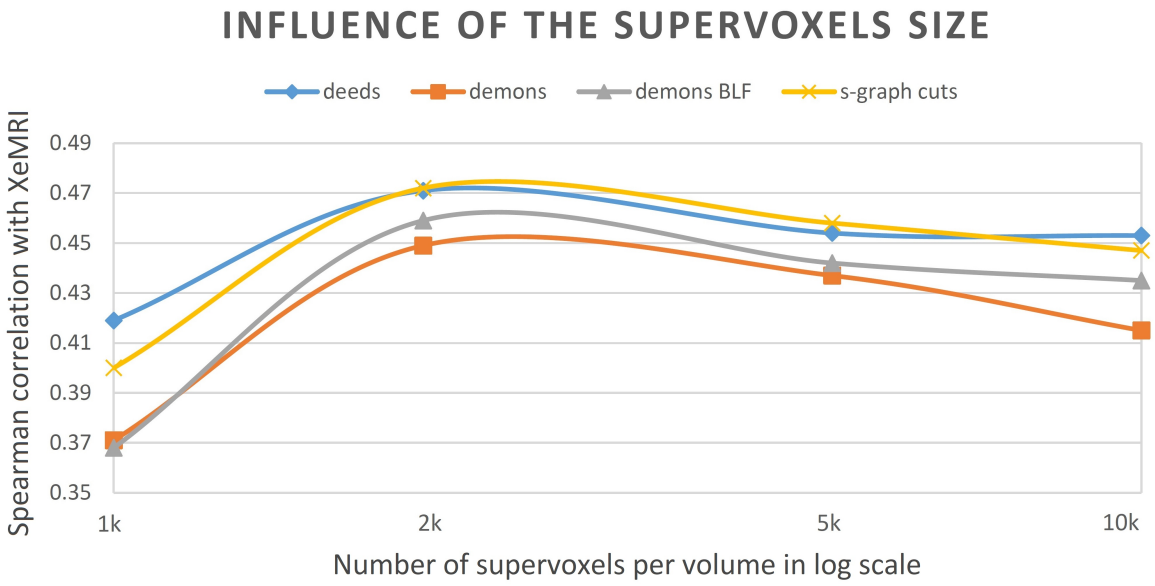



Figure 4.6: The plot shows the influence of the size of the supervoxels on the correlation of the estimated ventilation maps for all four registration methods with XeMRI ventilation images. The x-axis with supervoxel size is shown on a logarithmic scale. It is noticeable that for a low number of supervoxels the correlation is relatively low, then it reaches its maximum at approximately 2,000 of supervoxels.

these properties, the correlation of the SLIC-based estimated ventilation maps with XeMRI using different supervoxel sizes were measured. Approximately 1,000, 2,000, 5,000 and 10,000 supervoxels were extracted from every volume. The experiments were conducted with a fixed number of supervoxel layers, namely 15, which is a reasonable trade-off between performance and computational requirements. The same experiment was performed for all of the registration methods. A plot showing the trend for average results on a logarithmic scale is presented in Figure 4.6.

Here, detailed numerical results calculated for the influence of the supervoxels size are presented. The provided Tables 4.1, 4.2, 4.3, 4.4 were used to generate the plots presented in Figure 4.6, where the influence of the supervoxel size on the correlation between the results of the proposed estimated ventilation method and XeMRI ventilation images were investigated. Tables 4.1, 4.2, 4.3, 4.4 show the results calculated for the different image registration methods, the deeds [81], demons [60],

BLF demons [64] and supervoxel graph cuts [137], respectively.

Table 4.1: The influence of the size of supervoxels on the Spearman correlation between 4DCT-based estimated ventilation maps using the deeds [81] registration method and XeMRI ventilation images.

| Patient | $\sim$ 1k | $\sim$ 2k | $\sim$ 5k | $\sim$ 10k |
|---|---|---|---|---|
| 1 | 0.364 | 0.471 | 0.421 | 0.423 |
| 2 | 0.686 | 0.685 | 0.678 | 0.677 |
| 3 | 0.209 | 0.282 | 0.263 | 0.259 |
| Avg | 0.419 | 0.479 | 0.454 | 0.453 |

Table 4.2: The influence of the size of supervoxels on the Spearman correlation between 4DCT-based estimated ventilation maps using the demons [60] registration method and XeMRI ventilation images.

| Patient | $\sim$ 1k | $\sim$ 2k | $\sim$ 5k | $\sim$ 10k |
|---|---|---|---|---|
| 1 | 0.220 | 0.351 | 0.336 | 0.302 |
| 2 | 0.639 | 0.661 | 0.660 | 0.651 |
| 3 | 0.255 | 0.336 | 0.315 | 0.292 |
| Avg | 0.371 | 0.449 | 0.437 | 0.415 |

Table 4.3: The influence of the size of supervoxels on the Spearman correlation between 4DCT-based estimated ventilation maps using BLF demons [64] registration method with sliding motion preservation and XeMRI ventilation images.

| Patient | $\sim$ 1k | $\sim$ 2k | $\sim$ 5k | $\sim$ 10k |
|---|---|---|---|---|
| 1 | 0.236 | 0.417 | 0.380 | 0.372 |
| 2 | 0.651 | 0.653 | 0.650 | 0.641 |
| 3 | 0.218 | 0.308 | 0.295 | 0.294 |
| Avg | 0.368 | 0.459 | 0.442 | 0.435 |

Table 4.4: The influence of the size of supervoxels on the Spearman correlation between 4DCT-based estimated ventilation maps using supervoxel graph cuts [137] registration method with sliding motion preservation and XeMRI ventilation images.

| Patient | $\sim$ 1k | $\sim$ 2k | $\sim$ 5k | $\sim$ 10k |
|---|---|---|---|---|
| 1 | 0.385 | 0.450 | 0.428 | 0.417 |
| 2 | 0.603 | 0.675 | 0.667 | 0.659 |
| 3 | 0.213 | 0.298 | 0.278 | 0.264 |
| Avg | 0.400 | 0.474 | 0.458 | 0.447 |

For a sparse image representation with a low number of large supervoxels, the results show relatively poor correlation with XeMRI. However, by increasing the number of supervoxels, and hence, reducing their size, the correlation reaches its maximum for approximately 2,000 supervoxels extracted from a volume. Results gradually decay when increasing number of supervoxels of smaller size. All image registration methods exhibit a similar trend in their correlation with XeMRI.

### 4.3.3 Influence of the Number of Layers

The number of layers of supervoxels has an obvious influence on the estimated ventilation images. Using just a single layer of supervoxels, the results may be biased by the particular image clustering. By applying a larger number of layers, the results should be more robust and can be expected to result in the higher correlation with XeMRI. An experiment investigating the influence of the number of layers of supervoxels on the correlation between the estimated ventilation maps with XeMRI for all four image registration methods was conducted. Ventilation maps were estimated for 1, 5, 10, 15 and 25 layers of supervoxels. The size of the supervoxels was adjusted based on the previous experiment to approximately 2k supervoxels per volume. The obtained average results are shown in Figure 4.7. The proposed method achieves relatively low correlation when there is just one layer of supervoxels. The results improve noticeably after increasing the number of layers to 5 layers, while

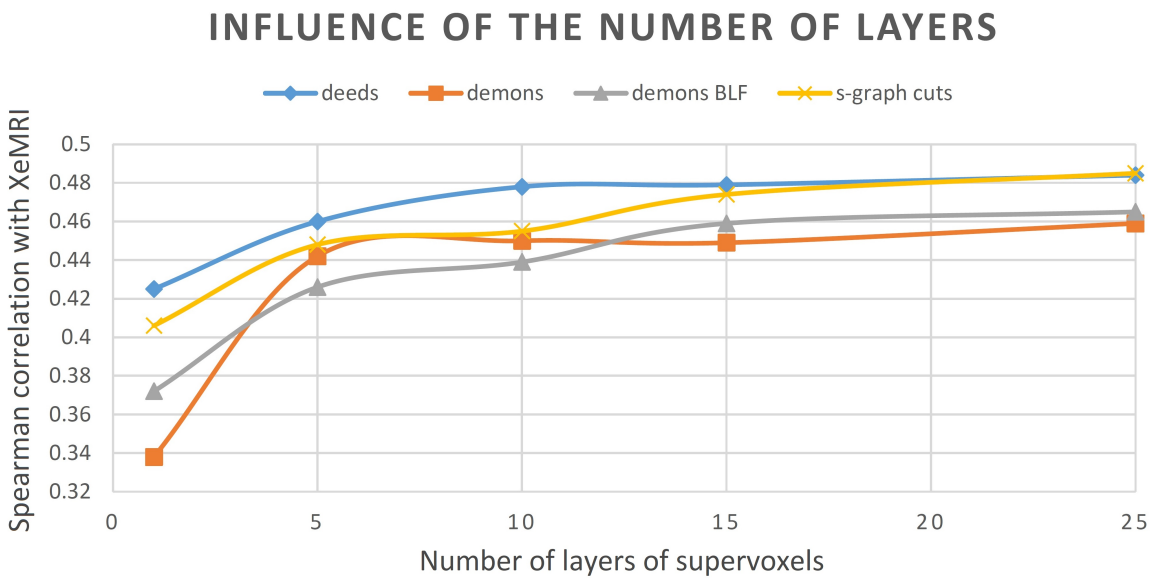



Figure 4.7: The plot shows the influence of the number of supervoxel layers on the correlation of the estimated ventilation maps with XeMRI ventilation images for all four registration methods. The correlation improves when the number of layers is increased, with noticeable change in the slope of the plot between just one layer of supervoxels and 10. Increasing the number of supervoxels from 15 to 25 resulted in only slight improvement in the overall correlation.

there is a more gradual improvement up to 15. With the use of more than 15 layers of supervoxels the correlation between XeMRI and the investigated ventilation surrogates estimated using different image registration methods only slightly improves for 25 layers. Similar trends can be observed independently of the applied registration method.

The detailed numerical results calculated for the influence of the number of supervoxels layers are presented in the following Tables 4.5, 4.6, 4.7 and 4.8. Based on these results, Figure 4.7 was prepared.

### 4.3.4 Global Correlation

To investigate the global relation of the estimated ventilation maps with XeMRI Spearman's correlation coefficient for all methods was calculated: the proposed SLIC-

Table 4.5: The influence of the number of layers of supervoxels on the Spearman correlation of XeMRI with deeds [81] registration method based ventilation surrogates.

| Patient | 1 layer | 5 layer | 10 layer | 15 layers | 25 layers |
|---|---|---|---|---|---|
| 1 | 0.374 | 0.411 | 0.469 | 0.471 | 0.476 |
| 2 | 0.644 | 0.689 | 0.686 | 0.685 | 0.690 |
| 3 | 0.257 | 0.281 | 0.280 | 0.282 | 0.287 |
| Avg | 0.425 | 0.460 | 0.478 | 0.479 | 0.484 |

Table 4.6: The influence of the number of layers of supervoxels on the Spearman correlation of XeMRI with demons [60] registration method without sliding motion preservation based ventilation surrogates.

| Patient | 1 layer | 5 layer | 10 layer | 15 layers | 25 layers |
|---|---|---|---|---|---|
| 1 | 0.200 | 0.326 | 0.354 | 0.351 | 0.369 |
| 2 | 0.500 | 0.662 | 0.663 | 0.661 | 0.667 |
| 3 | 0.315 | 0.338 | 0.335 | 0.336 | 0.332 |
| Avg | 0.338 | 0.442 | 0.450 | 0.449 | 0.459 |

based, the classic intensity based [126, 127], Jacobian based [128, 129] and HU averaging [122]. Despite the fact that the classic intensity based, Jacobian based and HU averaging methods were not initially developed for the evaluation on hyperpolarised gas MRI, they have been successfully used to study and compare ventilation surrogates with this ventilation imaging method [110, 132]. For all these methods $G_{k1} = G_{k2}$ with [15 15 15] voxel kernels size and $\sigma = 2$ voxels were used, as these settings have been previously reported in the literature [122, 132], and resulted in moderate correlation [132]. Only for the proposed SLIC-based method a smaller, [7 7 7] voxel kernel size with $\sigma = 1$ filter was applied. The reason for such setting was to introduce less smoothing compared to other methods. In general, it is shown that more smoothing results in higher correlation, however then local nature of ventilation images might be lost. From our observation, such setting was a reasonable trade-off between smoothing some voxel-wise artefacts and preserving the local nature of the estimated ventilation surrogates for the proposed method. A reader interested in the

Table 4.7: The influence of the number of layers of supervoxels on the Spearman correlation of XeMRI with BLF demons [64] registration method with sliding motion preservation based ventilation surrogates.

| Patient | 1 layer | 5 layers | 10 layers | 15 layers | 25 layers |
|---|---|---|---|---|---|
| 1 | 0.254 | 0.334 | 0.361 | 0.417 | 0.426 |
| 2 | 0.562 | 0.638 | 0.653 | 0.653 | 0.656 |
| 3 | 0.300 | 0.308 | 0.302 | 0.308 | 0.313 |
| Avg | 0.372 | 0.426 | 0.439 | 0.459 | 0.465 |

Table 4.8: The influence of the number of layers of supervoxels on the Spearman correlation of XeMRI with supervoxel graph cuts [137] registration method based ventilation surrogates.

| Patient | 1 layer | 5 layers | 10 layers | 15 layers | 25 layers |
|---|---|---|---|---|---|
| 1 | 0.291 | 0.377 | 0.398 | 0.450 | 0.477 |
| 2 | 0.652 | 0.677 | 0.671 | 0.675 | 0.678 |
| 3 | 0.275 | 0.289 | 0.297 | 0.298 | 0.299 |
| Avg | 0.406 | 0.448 | 0.455 | 0.474 | 0.485 |

influence of the applied filter sizes on the correlation of the ventilation surrogates with hyperpolarised gas MRI is referred to, for instance, [132].

In this experiment, the proposed method achieved a higher correlation on average for all cases compared with the voxel-wise method. The best results were obtained independently of the chosen image registration method. The classic approach achieved higher correlation compared with HU averaging method for all image registration methods, apart from the demons method. The lowest correlations with XeMRI were obtained for the Jacobian-based method for all registration methods. Only for this ventilation estimation approach, demons BLF achieved higher correlation than supervoxel graph cuts for one of the cases, yet still lower than the deeds method. The numerical results are presented in Table 4.9. Visualisation of the estimated ventilation maps for one of the cases is shown in Figure 4.4.

The correlation between the proposed ventilation estimation method and the clas-

Table 4.9: Spearman's correlation coefficients calculated for all ventilation estimation methods. For the classic methods, SLIC-based and Jacobian-based results achieved by investigated image registration methods were presented. In the case of HU-Avg no image registration method was applied. The highest correlations have been achieved for the proposed SLIC-based method.

| | Classic | | | | SLIC | | | | Jacobian | | | | HU-Avg |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Pat. no. | demons | demons BLF | deeds | graph cuts | demons | demons BLF | deeds | graph cuts | demons | demons BLF | deeds | graph cuts | no reg. |
| 1 | 0.205 | 0.306 | 0.356 | 0.348 | 0.369 | 0.426 | 0.476 | **0.477** | 0.274 | 0.273 | 0.280 | 0.278 | 0.399 |
| 2 | 0.605 | 0.598 | 0.645 | 0.633 | 0.667 | 0.656 | **0.690** | 0.678 | 0.335 | 0.349 | 0.414 | 0.339 | 0.444 |
| 3 | 0.279 | 0.256 | 0.186 | 0.198 | **0.332** | 0.313 | 0.287 | 0.299 | 0.096 | 0.105 | 0.109 | 0.101 | 0.315 |
| Avg | 0.363 | 0.386 | 0.396 | 0.393 | 0.459 | 0.465 | 0.484 | **0.48**5 | 0.235 | 0.242 | 0.268 | 0.231 | 0.386 |

sic method is calculated, showing that they correlate well (0.834), with the proposed method achieving the best correlation with XeMRI. The visualisation of the results is presented in Figure 4.8 for all cases. Worth noticing is the relatively high correlation achieved by the HU-Avg method. The method fails to provide local estimates of the ventilations, resulting in a blurred image of the shape of the lungs. This explanation can be confirmed by a closer investigation of the visual results in Figure 4.4 where the ventilation surrogates for one of the patients are presented. Similar behaviour was observed throughout the dataset. Local patterns of the ventilation changes can hardly be observed in the conducted experiments. It was not possible to see the trend visible for other methods where the results of the correlation with XeMRI for patient 2 were noticeably higher than for the other two. For Jacobian-based ventilation the results obtained by different image registration methods are more consistent than for the classic and SLIC-based, especially for Patients 1 and 3. Such an observation might indicate that Jacobian-based ventilation estimation method could be less susceptible to the image registration method applied to estimated deformations, based on which the determinants of the Jacobian are calculated. The correlation has been

evaluated only for the voxels inside the lung masks. Patients from our dataset suffered from severe ventilation defects, therefore, there were regions inside of the lungs where no ventilation in XeMRI images was acquired. When investigating Figure 4.8, the concentration of the values in the bottom line can be noticed for all methods, especially for Jacobian based ventilation estimation method. In SLIC and Classic method we can see that the concentration is mostly localised around zero. This indicates that both methods are correctly identifying regions with ventilation defects present in XeMRI. On the other hand in Jacobian based approach we can observe that the methods have classified these voxels as moderately ventilated (especially in Patients 2 and 3). Only for Hu-Avg method for Patient two hot spots in the bottom line can be observed. One localised close to zero indicates for correct identification of the ventilation defect. The other around 0.15 suggests that a region which shows no ventilation in XeMRI has been classified as moderately ventilated by the method. This corresponds to visualisation of the results from Figure 4.4, where ventilation defect in upper right lobe has not been detected by the method. In contrast, SLIC-based method the hot spot in the bottom line is well defined and shows that fewer voxels were classified incorrectly according to XeMRI. This might indicate that the method is well suitable for identifying regions suffering from ventilation defects, which is potentially important indicator in radiotherapy planning enhanced by ventilation information.

### 4.3.5 Spatial Overlap Comparison

A series of experiments was performed investigating the quantitative relation between the spatial distribution of the estimated ventilation maps and XeMRI ventilation images. Akin to [125], the estimated ventilation images were divided into five disjoint regions, based on calculated percentile distribution ranges (see Figure 4.9). These regions were further used to create masks corresponding to the percentile ranges.

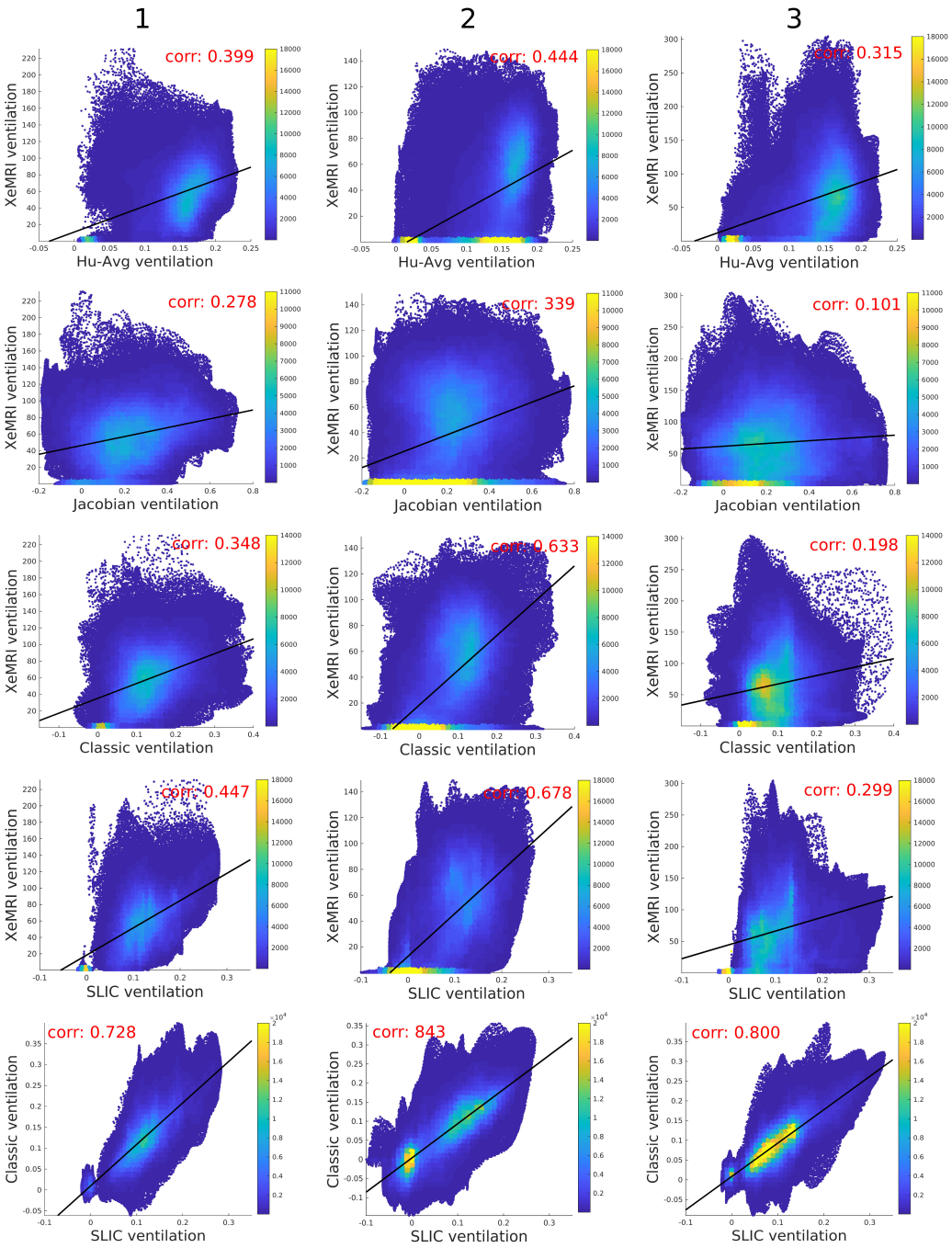



Figure 4.8: The figure shows correlation and Spearman correlation coefficient between the HU averaging ventilation approach and XeMRI in the first row. In the second row, correlation and Spearman correlation coefficient between the Jacobian-based ventilation approach and XeMRI are presented. In the following rows analogous correlations are presented for the classic approach and for the proposed SLIC-based method. In the last row, correlation between the classic and the proposed method are shown. Columns correspond to patient's number.

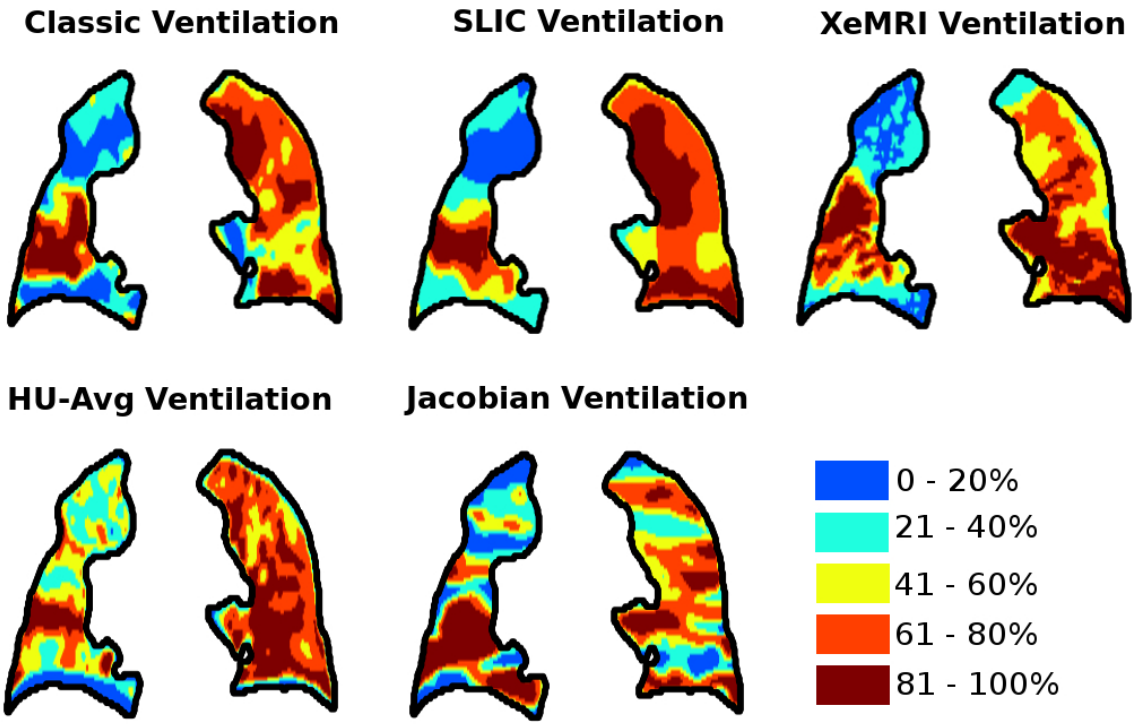



Figure 4.9: An example showing the partition of the ventilation images based on the percentiles of their intensities for patient 2 in a coronal slice view. The ventilation images were divided into five percentile ranges, which Dice overlap was subsequently calculated.

The Dice overlap coefficient with corresponding masks of XeMRI was calculated for each of them. This procedure was applied to all four lung ventilation estimation methods and the experiments were repeated for all investigated image registration approaches.

The visualisation of the results is presented in Figure 4.10. The general trend, where Dice overlap is the highest for the first percentile range (0-20%), for the intermediate ranges is lower, and then raises for the last percentile range (81-100%) is in agreement with similar experiments in the literature [125]. Similar patterns in the results can be observed for all of the tested image registration methods. For all of the evaluated percentile ranges, the highest correlation has been achieved for the proposed method (in red). The lowest Dice overlap has been obtained for the Jacobian-based ventilation method. The classic approach and HU averaging method achieved similar results. The results for the HU averaging method are independent of the image registration methods. However, the plots showing the results of the correlation for different image registration methods were included to enable the comparison with them.

A valuable piece of information from Figure 4.10 is that the highest correlations have been observed between the least and the most ventilated lung regions consistently for all investigated ventilation estimation methods and image registration methods. This means that all of the methods are most accurate in which parts of the lungs work well and which do not contribute to the gas exchange. This might have a direct clinical implication, as such information could be incorporated into radiotherapy planning to make sure that well functioning regions will be spared even at cost of poorly ventilated parts of the lungs.

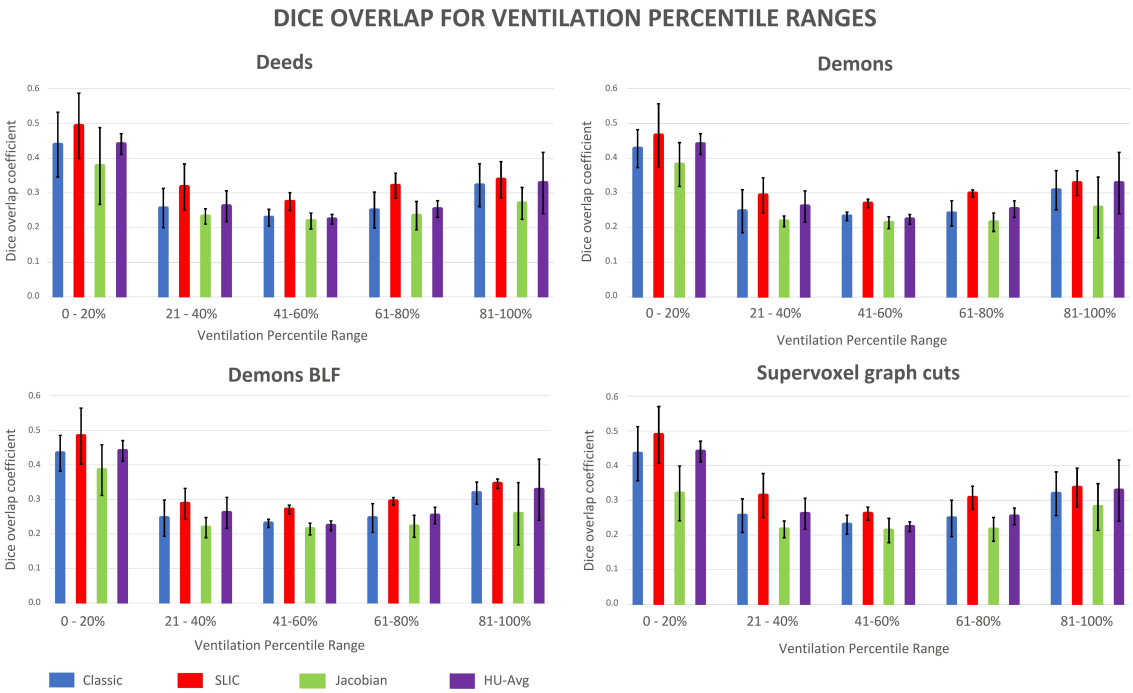



Figure 4.10: Dice overlap coefficient calculated for masks created based on the ventilation percentile ranges. The results for all four image registration methods: deeds, demons, demons with BLF and supervoxel graph cuts are investigated. For each of them, all four estimates of the ventilation were calculated. The presented results are averaged for all three patients from the dataset. The highest Dice overlap was achieved for the proposed method in red. It can be noticed that similar patterns in the results can be observed for all of the tested image registration methods.

## 4.4 Discussion

In this Chapter we introduced and investigated the patch-based approach for ventilation estimation. Several experiments were conducted, measuring the correlation of the estimated ventilation maps with XeMRI ventilation images. The evaluation was performed using a proposed registration framework bringing SLIC-based ventilation images into alignment with XeMRI. The presented results show that the proposed method results in higher correlation with XeMRI for a number of performed tests.

Following the discussion from Chapter 2, there are two main approaches for estimating lung ventilation surrogates based on the deformations resulting from 4DCT coregistration. The first applies the differences in intensity between corresponding

voxels of extreme breathing phases. The alternative approach relies on examining the value of the determinant of the Jacobian of the deformations estimated in the registration process. The proposed approach is based on the first one, where we track the changes in the intensity of the corresponding regions of the lungs between peak breathing phases, represented by supervoxels. The potential change in the size of the supervoxels between the peak breathing phases in such an approach does not need to be taken into account using the first approach for estimating lung ventilation surrogates. A method closer in approach to the Jacobian-based ventilation estimation method, would focus on addressing this behaviour. Then changes in supervoxels size would be estimates of the determinant of the Jacobian, measuring expansion or contraction of certain regions of the lungs. The choice of method was based on previous studies showing that methods relying on differences in lung tissue intensity results in ventilation maps better correlating with functionality of emphysema patients [127]. The other approach would be important for instance in MR-LINAC, where no CT is available, and therefore our method would not be suitable and only Jacobian based methods could be applied.

The influence of the supervoxel size was investigated in Sec. 4.3.2 for all of the image registration methods. The results presented in Figure 4.6 suggest that to achieve the best correlation, the size of the supervoxel needs to be chosen with care. In the conducted experiments, the highest correlation was achieved for approximately 2,000 supervoxels per volume. This means that the average supervoxel consists of 30,000 isotropic 1x1x1mm$^3$ voxels. This is related to their ability to model the anatomy of the lungs. This trend in the changes of correlation has been expected and supports the stated claim that structure-oriented patch-based approach is a better choice than a voxel-wise approach. The limited number of available cases did not allow us to create separate training and testing datasets. Apart from the image registration methods settings, all other parameters of the presented ventilation estimation method were

trained and tested with the use of the available dataset in a manner described in Sec. 4.3.1. The chosen approach might be seen as suboptimal, however due to the unique character of the dataset, creation of the separate training and testing datasets based on them was not feasible. Further evaluation on the larger dataset would be required to draw meaningful conclusions about superiority of the proposed method over other approaches.

The results from Sec. 4.3.3 show that a larger number of layers of supervoxels provides images more correlated with XeMRI. This is expected, as a larger number of layers of supervoxels should improve the robustness of the method for mis-segmentations originating from a particular supervoxel extraction. The investigation of the influence of the compactness of supervoxels was out of the scope of this work, because the intention is for the supervoxels to mimic the anatomy of the lungs. Therefore, this parameter was adjusted manually to make the supervoxels correspond in shape to lung structures.

When investigating the influence of the number of supervoxels on the correlation between XeMRI and ventilation surrogates, several different image registration methods was applied. Despite differences in approaches of the studied methods, a similar trend of the correlation for all of them could be observed. This observation suggests that the proposed method is not restricted to one of the evaluated methods, but is more universal, as the methods investigated in this study represent the main registration categories in terms of optimisation, namely: discrete, continuous, and hybrid. Given the small number of cases analysed, we cannot provide a definitive answer as to whether the registration methods are equivalent; however, we believe the similarity of the results provides some reassurance as to the robustness of the method in this respect.

The results presented in Sec. 4.3.4, where the demons image registration method was compared with its sliding motion preserving version, show that the incorporation

of the sliding motion improves the correlation of the estimated ventilation maps with XeMRI. This observation supports the claim that sliding motion preservation results in more anatomically plausible deformations of the registration. The deeds [81] and supervoxel graph cuts [137] methods showed slightly more accurate performance on the Dir-Lab dataset [119] than both demons methods, which is also reflected in the higher correlations calculated for these methods. The difference between the correlations calculated for the deeds [81] and supervoxel graph cuts [137] for all ventilation estimation methods is minimal (below 0.01). This may suggest that even though both methods represent different approaches to the image registration problem, both cope well with the task providing similar level of accuracy. This also suggests that more accurate image registration methods provide estimations of ventilation which result in higher levels of correlation with XeMRI. 4DCT data for all patients suffered from reconstruction artefacts, which made the registration a challenging task. Yet, it is noticeable that in our dataset, the highest correlations has been found for Pat. 2. One possible explanation for the higher correlations for Pat. 2 is that they suffered from the most severe ventilation defects, as can be seen in Figure 4.4, where upper lobe in right lung is poorly ventilated. As it has been shown in Figure 4.10, the highest scores of correlation has been achieved between poorly ventilated regions. Similar trend can be observed in Tables 4.5, 4.6, 4.7 and 4.8, which show that this observations is independent of the applied image registration method and even number of layers of supervoxels.

The analysis of the Dice overlap for the ventilation percentile ranges shows favourable performance of our method for all investigated ranges. The Jacobian-based ventilation maps resulted in lower correlation with XeMRI than the classic approach. This observation is in agreement with previous findings from [130]. An interesting observation is that the HU averaging methods achieves comparable results to the classic approach. Especially in the context of the results reported in [122], where the au-

thors found low correlation for the patients with severe ventilation defects, which is also the case in our study. One explanation for this can be that the results of this method are blurred and smoothed, as the volumes are averaged, which provides relatively high correlation but the visual inspection reveals lower regional specificity of the method. It has been previously showed that such behaviour could be observed for other ventilation estimation methods.

The presented work investigated the influence on the correlation of the estimated ventilation with XeMRI of using four different image registration methods. They represent distinct approaches to the task, covering most of the existing groups of methods. This work did not aim to evaluate which of the registration methods performed best, but rather wanted to show the general trend that the application of the proposed supervoxel-based ventilation estimation improves the correlation to XeMRI ventilation maps, irrespective of the underlying image registration.

An open point for a discussion may be whether the achieved level of correlation reaches the expectations and is applicable to enhance radiotherapy. The proposed SLIC-based method and the classic approach were able to identify severe ventilation defects, as shows an example in Figure 4.4. Precise localisation of the ventilation defects remains challenging to evaluate. In a recent study [179], the researchers have evaluated and compared ventilation maps estimated based on 4DCT and high quality exhale/inhale breath-hold CT (BHCT) with Galligas PET. They found that ventilation maps estimated using BHCT achieved higher correlation values. Such an observation may be mainly due to reconstruction artefacts present in 4DCT, and could suggest that the application of BHCT has the potential to improve the accuracy of the estimated ventilation maps.

This chapter presented a novel patch-based ventilation estimation method from 4DCT. It was shown that the proposed method achieves higher correlation coefficients compared to the classic approach when correlated with XeMRI. The results, although

already encouraging at the current stage, would gain on significance when evaluated on a larger cohort of patients. Additionally, an improved pMRI to CT registration, which brings XeMRI to the alignment with CT, may play an important role in this evaluation. The method applied in this study, as well as methods applied by other researchers, were not developed strictly for this task. Therefore, the pattern of the regularisation applied inside the lungs may not fully correspond to the anatomy of the lungs. In the presented study we applied the deeds [81] image registration method, which has been proposed for lung image registration purposes for application to lung CT and tested in multimodal lung image registration. However, the specific application of XeMRI/pMRI to CT registration, where MRI data are highly anisotropic, might be even more challenging than other applications. In the deeds method the similarity measure is evaluated with a descriptor-based MIND, which works well as an edge detector, and therefore the surfaces of the lungs as well as the main airways or vessels should in result be well aligned. However, in the case of the ventilation images the most important regions are within lung parenchyma, which is not well visible in MRI and, therefore, its accurate alignment with CT is a challenging task. This can mean that the registration error originating from pMRI to CT registration in our case can have non-negligible influence on the results.

We have presented a proof of concept for a patch-based ventilation estimation method from 4DCT. We have shown that our proposed method achieves higher correlation coefficients compared to the classic approach when correlated with XeMRI. The results, already encouraging at the current stage, would gain more significance when evaluated on a larger cohort of patients.

The presented results, where for the majority of the conducted experiments the proposed method for ventilation estimation from 4DCT achieves better correlations with XeMRI images, might indicate higher physiological consistency of the proposed approach with XeMRI ventilation images. Future work includes performing further

analysis on a larger patient group. An improvement in the correlation can be observed when more accurate image registration methods were applied to CT-CT registration. Further improvements may be expected when task dedicated registration between pMRI and CT was applied.

The following Chapter 5 in Sec. 5.1 presents one such approach in a form of a personalised motion model-based multi-modal image registration method. The presented approach addresses local variations between images in the supervoxel-based motion model parameters optimisation step. The second section (Sec. 5.2) introduces a framework for breathing motion correction in dynamic XeMRI. The analysis of such data without breathing motion correction might result in spatial misalignment of the subsequent ventilation volumes. The proposed framework addresses this issue by constructing a CT-based lung atlas, which helps to bring all ventilation time frames volumes into alignment.

# Chapter 5

# Towards More Accurate Hyperpolarised $^{129}$Xenon MRI Analysis

The work presented so far addressed CT-CT lung image registration and 4DCT-based lung ventilation estimation. Both of the presented methods are based on the application of supervoxel-based image representation. The next challenge, which directly arises from the previous chapter, is proton density MRI (pMRI)/hyperpolarised $^{129}$Xenon MRI (XeMRI) to CT registration. This is addressed in Sec. 5.1, which proposes a multi-modal lung image registration method enhanced with personalised motion model approach. In Sec. 5.2, we introduce an approach for breathing motion correction of the dynamic sequence of hyperpolarised gas XeMRI. Both methods are presented in full detail, but due to time constraints were subjected to limited testing and thus will require a more complete validation in future work.

## 5.1 Multi-modal Lung Image Registration Enhanced with a Personalised Motion Model

There are many methods dedicated to CT-CT lung registration. However, pMRI/ XeMRI to CT lung registration remains a hardly explored area. That is a non-trivial task, mainly because of the insufficient amount of information inside the lungs in pMRI, which is caused by the lower proton density of the lung tissue compared to other organs. Multi-modal lung image registration requires addressing more complex deformations and intensities that are not directly comparable between the acquired scans. An additional challenge discussed in Chapter 2 is the presence of 'sliding motion' between the surfaces of the lungs, the ribcage, and diaphragm. A difficulty also stems from the susceptibility to acquisition artifacts in pMRI caused by the interfaces between air and lung tissue. Lung pMRI is usually acquired with lower spatial resolution than CT, which makes it difficult to find corresponding anatomical features between the images inside the lungs, which could help to drive the registration in these regions.

This section presents our work on a novel method addressing the problem of insufficient amount of information inside the lungs in pMRI, applying a personalised 4DCT statistical motion model to a supervoxel-based graphical image registration [83, 137]. The motion model is estimated based on the results of intra-patient image registration using Principal Component Analysis (PCA). The approach with a prior motion model is particularly important for regions where there is not enough information to reliably drive the registration process, as in the case of XeMRI and pMRI to CT registration. Simultaneously, the method addresses local variations between images in the supervoxel-based motion model parameters optimisation step. The results are compared in terms of the plausibility of the estimated deformations and correlation coefficient with 4DCT-based estimated ventilation maps using state-of-the-art multi-

modal image registration methods. The method achieves higher average correlation scores, showing that the application of PCA-based motion model in the deformable registration helps to drive the registration for the regions of the lungs with insufficient amount of information. The main contribution of this work is a dedicated framework, which addresses the challenges of XeMRI to CT deformable registration in the form of supervoxel-based motion model enhanced method. The methodology has been previously described in [174].

### 5.1.1 Background

Even though a lot of research has recently been dedicated to lung MRI [180, 181, 182], it still is not as common as lung CT imaging. Therefore, MRI and CT registration for the lungs is not that well investigated as in other organs, such as the human brain. This registration, however, plays an important role in the analysis of XeMRI [30] with relation to 4DCT images. As XeMRI does not provide structural information, its correspondences to the patient anatomy rely on pMRI, which is acquired during the same imaging session but not within the same breath-hold. This can result in spatial misalignment between XeMRI and pMRI, which is corrected by initial rigid registration. 4DCT is acquired during free breathing, under spirometry control, when the patient is instructed to breath normally or can be couched before the session how to maintain regular tidal breathing. It is not, therefore, straightforward to directly map XeMRI to lung 4DCT, as the images are acquired at different breathing patterns. While 4DCT is performed at tidal breathing, XeMRI are acquired at a breathold after an inhalation of one litre of gas. This is likely to result in different levels of inflation and with no structural reference points present in both images. These issues can be addressed through an intermediate registration between pMRI and CT, as shown in Figure 5.1. This registration becomes particularly challenging for a number of reasons, including the lower spatial resolution of pMRI compared with CT, the

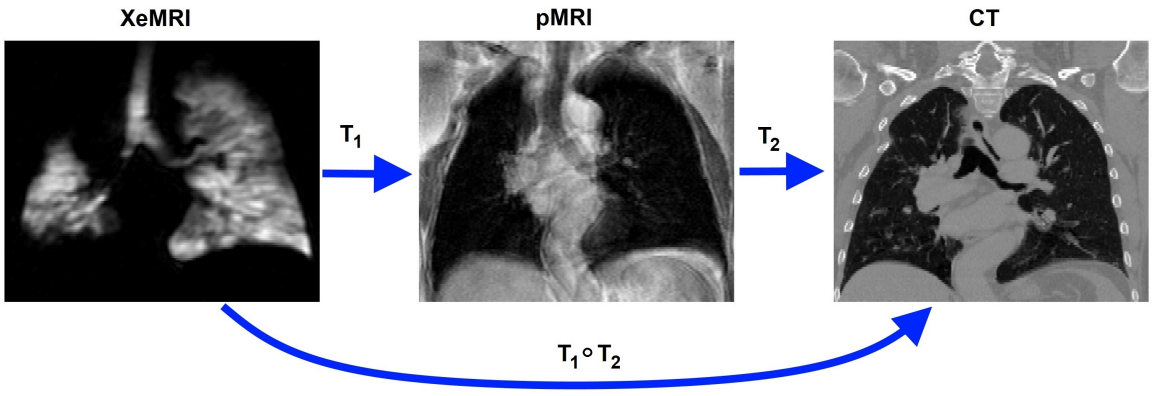



Figure 5.1: To bring XeMRI into alignment with CT, two transformations are composed: transformation $T_1$ that compensates for a possible initial misalignment between XeMRI and pMRI, and transformation $T_2$ estimated based on registration between pMRI and CT. The dedicated method addressing this problem is the main contribution of this work.

limited information from lung tissue in pMRI due to its low proton density, and the presence of susceptibility artifacts. Therefore, the registration can easily result in under- or over-estimation of deformations inside the lungs.

In the case of lung 4DCT the images are already rigidly aligned and only deformable image registration is required to bring all of them into alignment to a reference frame. In registering pMRI to CT images, an initial rigid body registration is required. The images not only have been acquired at different points in time, but also with different spatial orientation. While CT images are acquired in axial view, pMRIs in our dataset is performed in coronal direction, with (over 20 times) larger slice thickness. The multi-modal nature of the images, differences in the level of breath-hold and the strong anisotropic voxel sizes in pMRI are the main challenges to be addressed. Direct application of an off-the-shelf deformable image registration method to such data might not only result in suboptimal alignment, but also in physiologically implausible deformations. For these reasons, motion correction between CT and pMRI is usually performed as a two-step process, with coarse initial alignment covered by rigid registration, followed by deformable image registration. An

application of affine registration in this case might not be advisable, as the possible differences in the lung volumes in CT and pMRI could result in global scaling of the lungs, which does not correspond to equivalent levels of inhalation. Such differences in lung volumes, however, should be compensated for in the deformable registration step, which has the potential to more accurately mimic the complex motion of the lungs. Therefore, rigid registration, represented by translations and rotations only, is more suitable as an initial alignment method.

The following step is deformable registration, in which all the misalignments not covered by rigid registration should be corrected. pMRI to CT registration is not well investigated and not many dedicated methods have been proposed. One of the most established approaches for multi-modal image registration is FFDs on B-splines with mutual information as a similarity measure [67]. A method that has been originally proposed for lung CT registration, deeds [81], has shown good performance in multi-modal image registration application. This is mainly due to its image descriptor-based similarity measure, MIND [55], and an efficient optimisation posed on an MST. Both methods represent a voxel-wise registration approach, where estimated deformations are not restricted to the lung masks. In the case of XeMRI to CT registration a more suitable approach might be to focus the registration on the lung regions only. This is mainly because there is no relevant information about ventilation in XeMRI outside of the lungs. In this application, edge detecting descriptors, such as MIND may not be an optimal solution, as they would drive that registration mainly on contours of the lungs, which will not be present, and would fail to find match for the sparse features inside the lungs in pMRI.

### 5.1.2 Methods

The method proposed in this chapter consists of three main steps: (1) creation of a personalised lung motion model from 4DCT, (2) lung image clustering and (3) graph-

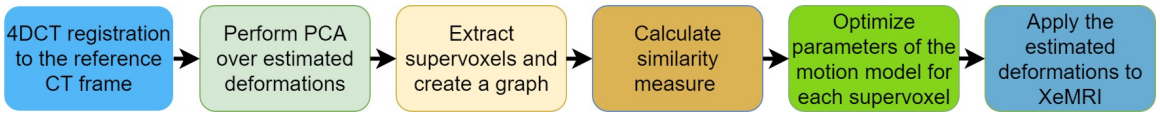



Figure 5.2: Diagram presenting the workflow of the proposed method. It starts from registering all the 4DCT volumes for each of the patients to the chosen reference frame. Principal Component Analysis (PCA) decomposition is applied over the estimated deformation fields to create a motion model. Subsequently, supervoxels are extracted from the lungs in the reference CT volume. A graph is created, where every supervoxel is represented by a node and all adjacent supervoxels are connected by an edge. For every supervoxel the best set of motion model parameters is found to bring pMRI into alignment with CT using graph cuts optimisation. Finally, as the ultimate goal of the registration framework, the estimated deformation field is applied to XeMRI.

based pMRI-to-CT registration. An overview of the proposed method with all the necessary operations within the framework is presented in Figure 5.2.

The proposed method creates a personalised motion model using displacements estimated during 4DCT registration to a reference volume. The supervoxel-based approach, presented in Chapter 2, is applied as the image registration method [137] to co-register different breathing phase volumes and estimate deformation fields, which will be used to create the personalised motion model. The method demonstrates potential to estimate abnormal lung motion more accurately, shows good performance in terms of TREs, plausibility of the resulting deformations for lung CT registration, and has the ability to address the sliding motion problem. In the proposed method, for each patient, all breathing phases from 4DCT are co-registered to a reference volume, which is chosen as the peak inhale breathing phase volume. The 4DCT data consists of 10 volumes; therefore as a result of the alignment 9 displacement fields are acquired. Following the registration, vectors comprising all the estimated deformation fields are created:

$$R^p(\mathbf{x}) = [V_1^p(\mathbf{x}), V_2^p(\mathbf{x}), ..., V_n^p(\mathbf{x})], \tag{5.1}$$

where $p$ is the direction of the deformations (anterior-posterior, left-to-right, superior-

to-inferior), $n$ is the number of volumes that are co-registered to the reference volume, and $\mathbf{x}$ is a voxel location.

Subsequently, Principal Component Analysis (PCA) is applied to the estimated deformations, $R^p(\mathbf{x})$, to obtain major motion patterns. The estimated major motion patterns for each of the 4DCT sequences create a personalised motion model. After applying PCA, Eq. 5.1 can be reformulated in terms of eigenvalues and eigenvectors:

$$R^p(\mathbf{x}) \simeq \mu_d^p + \sum_{i=1}^{n} \lambda_i^p \nu_i^p(\mathbf{x}), \tag{5.2}$$

where $\mu_d^p$ is the mean displacement, $\nu_i^p(\mathbf{x})$ is the i-th eigenvector and $\lambda_i^p$ is the corresponding eigenvalue for direction $p$ (anterior-posterior, left-to-right, superior-to-inferior) for each voxel's spatial location $\mathbf{x}$. The motion model is restricted to use the first eigenvector, as it covers the major motion pattern observed during the registration. The fraction of the variance explained by each of the principal components is a common quantitative measure in PCA of how well individual eigenvector contributes to the data. This is calculated as the ratio between the variance of that principal component and the total variance of the data. In the case of our data it was on average in anterior-posterior direction - 83%, in left-to-right direction - 82% and in superior-to-inferior direction - 95%. The restriction to the use of only that eigenvector makes the optimisation more efficient, while taking advantage of the personalised motion model application. Regional variations from the motion pattern are compensated for in the following step by applying supervoxel-based motion model parameters optimisation, similar to the image registration method described in Chapter 2.

The most common approach for using motion models is to perform a rough image alignment, followed by more detailed deformable image registration (see for example [116]). The method proposed in this chapter presents a different approach. Our approach is also different from a more standard global motion model PCA-based method, where the optimisation of the parameters of the motion model is performed

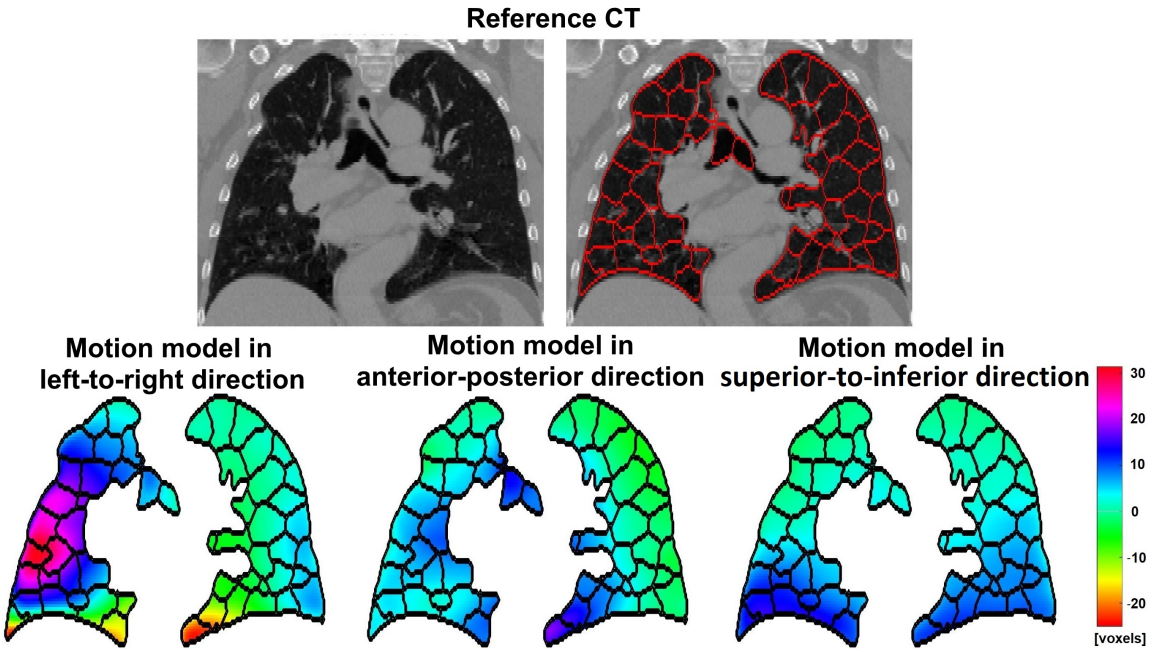



Figure 5.3: The reference CT image in the coronal view and superpixels estimated for the lungs imposed on the image are shown in the upper row. Below, the estimated motion model for the reference CT volume frame left-to-right, anterior-posterior and superior-to-inferior directions shown in coronal view with propagated superpixels from the CT image. For illustrative purposes, superpixels are extracted from a 2D image, whereas the proposed method uses supervoxels extracted from 3D volumes.

globally and a single set of parameters is found for the whole image. The proposed approach differs here, and the set of motion model parameters is found for every group of voxels represented as a supervoxel, which means that adjacent supervoxels might be assigned to different sets of motion model parameters. This potentially allows for more variation and accurate modelling of the lung motion, even if the estimated motion model does not fully represent possible lung motion. With its local optimisation approach, the proposed method is closer in assumptions to supervoxel-based deformable registration methods, such as [83, 136, 137]. First, our method combines motion model-based registration with deformable image registration into one framework. The global motion model is estimated for the segmented lungs from 4DCT, and then parameters of the motion model are optimised locally for groups of voxels. Second, the groups of voxels are represented by supervoxels. Supervoxel-based image representation has shown the potential to cluster voxels of similar appearance and spatial location, and that are therefore likely to undergo similar motion during breathing. We applied Simple Linear Iterative Clustering (SLIC) as clustering method. The method has been already presented in detail in Sec. 3.1.1 of Chapter 3. The same clustering method was applied in previous Chapters 3 and 4.

As the motion model has been created using one sequence of 4DCT per patient, such a limitation could make the model unable to fully mimic all the intra and inter breathing cycle variations. The proposed method, integrating deformable image registration with a motion model-based approach, is designed to compensate for such differences. Image registration, as a problem of finding the optimal transformation between two images, can be stated using a Markov Random Fields-based optimisation and posed on a graph, just as it was presented in Chapter 3. Following image clustering, a graph is created, where every supervoxel is represented by a node and all nodes corresponding to adjacent supervoxels are connected by an edge. The edge values are uniformly set to 1.

As an optimisation scheme a similar approach to [137] is applied, using graph cuts [151]. In the proposed method, labels to be assigned to each supervoxel are in the form of a set $l \in L$, where every label $l$ is a vector $[l_x l_y l_z]$ of parameters, where $l_x, l_y, l_z$ are parameters of the model in anterior-posterior, left-to-right and superior-to-inferior directions respectively. This is one of the main differences compared with the majority of other methods in the field, where labels usually directly represent displacements. The label is applied to the corresponding patch of the motion model, and therefore, even if the algorithm assigns the same labels to neighbouring supervoxels, they may potentially still have different displacements. The displacement inside the patch is not uniform and should mimic the motion of its tissue. Such an approach restricts the possible displacements of the patches to those which have been estimated for the particular regions of the lungs, and therefore results in more anatomically plausible estimated displacements. At the same time, the method still allows for local adjustments to the model by the estimated parameters. The estimation of the motion model parameters in $l_x$, $l_y$ and $l_z$ is one advantage of the proposed method, as it compensates for the residual differences, when ideal rigid alignment of pMRI and CT is difficult to achieve. This alignment is challenging mainly because of the multi-modal nature of the images, differences in position inside the scanner, as well as possible variations in the patient anatomy, for instance due to tumour presence. The model was created from 4DCT, based on co-registration of images acquired with the patient remaining at the same position in a scanner. Therefore the presented approach takes the advantage of a model based-approach to compensate for the misalignments and drive the registration for regions with not enough information, while at the same time allows for local derivations from the model to more accurately register the images.

As a similarity measure to find the optimal parameters of the motion model for every supervoxel the locally normalized correlation coefficient (LNCC) [47] was applied, which is a well established approach for measuring image similarity in medical

image registration. Our initial tests have shown that MIND [55], which was successfully used in the method presented in Chapter 3, resulted in suboptimal performance in this application. The main reason for that might be that MIND mostly relies on edges, whereas in this application we were interested in matching tissue inside the lungs, where no strong edges are present. The general formulation of the energy to be minimised during the optimisation process is:

$$E(l) = \underbrace{\sum_{p} \overline{LNCC(I_{fix}(\mathbf{x}_p), I_{mov}(\mathbf{x}_p + l_p \times R(\mathbf{x}_p))}}_{data\ term} + \alpha \underbrace{\sum_{p,q \in N} \|l_p - l_q\|^2}_{smoothness\ term}, \tag{5.3}$$

where the data term is formulated as a mean error calculated for all voxels $\mathbf{x}$ in the fixed image $I_{fix}$ and moving image $I_{mov}$ clustered in a certain supervoxel represented by a node $p$, for the applied motion model $R$ with the parameters represented by a label $l_p$. $R(\mathbf{x}_p)$ represents a 3D patch of motion model $R$ created by all voxels belonging to a supervoxel represented by a node $p$. The label $l_p$ is a vector which provides 'scaling' parameters for the $R(\mathbf{x}_p)$, and the operation $l_p \times R(\mathbf{x}_p)$ indicates multiplication of x, y, z- direction motion models from $R(\mathbf{x}_p)$ by a corresponding scaling factor from the estimated label vector $l_p$. The piecewise smoothness term represents quadratic distance between the labels. The influence of the piecewise smoothness term on the energy is controlled by a weighting parameter $\alpha$. Since no XeMRI ventilation signal is expected to be present outside of the lung, the presented registration framework is therefore restricted to estimating deformations inside the volume of the lungs. The lungs are segmented from CT and registration is done only inside the masks. Akin to [83], the method uses a single resolution level with multiple layers of supervoxels, slightly varying their size and initial location. The estimated deformations are averaged across all layers.

The displacements estimated for the pMRI-to-CT registration are propagated to

XeMRI, just as shown in Figure 5.1, resulting in their alignment with CT. Visual assessment of the framework's results are presented in Figure 5.4, showing the estimated displacement fields for the proposed method and those chosen to compare against.

### 5.1.3 Experiments and Results

Experiments have been performed on a dataset of the same three patients as in Chapter 4, where the data and their acquisition have been described in detail.

The first step in the proposed framework was to compensate for the initial misalignments between pMRI and XeMRI, as presented in Figure 5.1. The images were not acquired during a single breath-hold, therefore some misalignments might occur. pMRI and XeMRI were represented by volumes of lower spatial resolution than CTs, and therefore all the volumes were resampled to the CT resolution using cubic interpolation. In the presented application it is important to achieve good alignments at the apex and upper parts of the lungs. Following [36], pMRI volumes and reference volume from 4DCT were carefully initially aligned using rigid registration with normalised mutual information as a similarity measure. For consistency with the results presented in the previous Chapter 4, here we have applied the same parameters of the registration as were used previously. The resulting transformation was propagated to the corresponding XeMRI volumes, bringing them into rigid alignment with the reference CT volume.

Subsequently, deformable registration of pMRI-to-CT was performed and the results of the proposed registration method are compared with those obtained with the deeds deformable image registration method [81] and FFDs-based image registration approach using B-splines [67].

For the proposed method, we extracted 3D supervoxels consisting of approximately 500 voxels each, with the compactness parameter set to 0.1, and used 20

layers of supervoxels. The range of motion model parameters $l_x, l_y$ and $l_z$ was set between -0.6 and 0.6 as weighting parameters for the eigenvectors of the motion model, at the intervals of 0.1. This allowed us to cover the largest misalignments between pMRI and CT. The weighting parameter $\alpha$ was empirically set to 0.02 and locally normalised cross correlation was calculated for a 7x7x7 voxels patch size. The method has been implemented within the Matlab environment and its running time with the chosen parameters setting is approximately 45min on an i7 2.7GHz laptop machine. The running times for the deeds and B-splines were approximately 13min and 25min, respectively, with C++ implementations. The method is capable of further optimisation and parallelisation, which should result in significant running time reduction.

Visual inspection of the results from Figure 5.4 reveals that the displacements estimated by the proposed method are anatomically more consistent. The figure presents a coronal view of CT image in a) for Pat. 2. In b) XeMRI initially aligned with rigid registration is presented and in c) the estimated ventilation surrogates are shown. In the middle row in d), e) and f) the results of registration of XeMRI to CT are shown for three investigated methods: B-splines, deeds and the proposed, respectively. Blue arrows in d), e) and f) point to the corresponding regions and highlight implausible shapes of the results, especially for deeds, where 'wavy', not expected, patterns are introduced. The green arrows point to a region of a change between low and well ventilated part of the lungs from XeMRI. Here again a 'wavy' pattern in the images resulting from deeds registration can be observed. Additionally, the motion of low ventilation region seems to be underestimated by B-splines approach, as in image c) it is closer to the lung apex and seems to match with the border between the lung lobes. The best visual correspondence between them can be observed for the proposed approach in f).

To evaluate the obtained results, XeMRI ventilation images were compared with ventilation maps estimated from 4DCT, and this is obtained using image registration

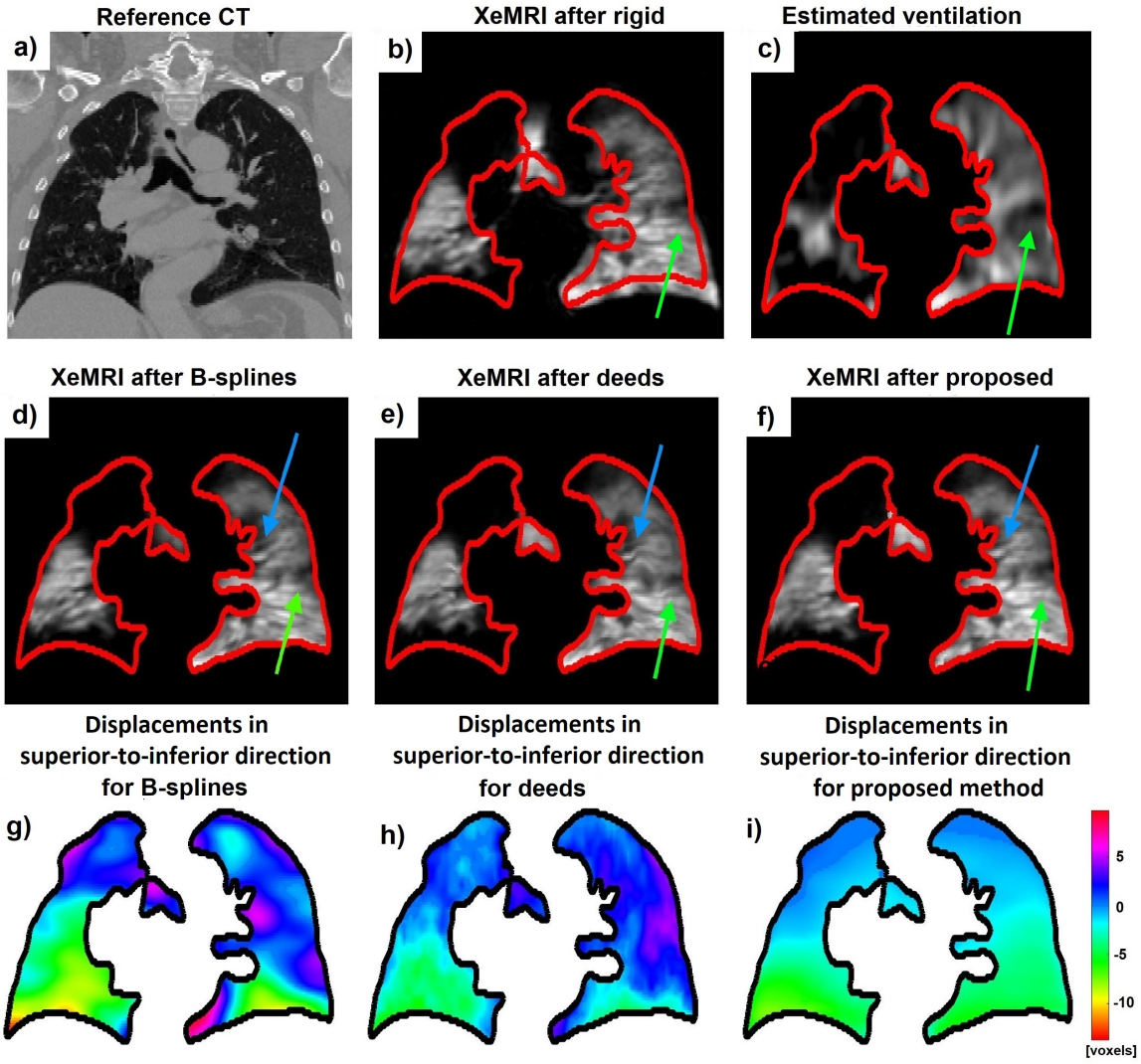



Figure 5.4: Coronal view of the CT scan of Patient 2 is shown in a). In b) XeMRI after applying rigid registration ($T_1$ from Figure 5.1) and in c) ventilation estimated from 4DCT are presented. The remaining figures show XeMRI ventilation images for the corresponding to CT slices. The lung border from CT is super-imposed on the ventilation images. The results of the XeMRI ventilation after applying deformable registration are shown only inside of the lung mask in the middle row. The possible under-estimation of the motion for B-splines [67] d) and deeds [81] e) are pointed to by green arrows, and implausible deformations by blue arrows. The results for the proposed method are shown in f). The bottom row shows displacements in superior-to-inferior direction for the corresponding slices for all the methods in g), h) and i).

of the dynamic CT sequence to a reference CT volume. Estimation of the ventilation maps was performed with the use of a method based on the changes of the lung intensity expressed in HUs between peak inhale and peak exhale breathing phases [127]. An alternative approach would be to estimate the ventilation from 4DCT based on determinant of Jacobian [126, 129] or with the use of supervoxel tracking [173]. Spearman's correlation coefficient was calculated between the registered XeMRI ventilation images and the estimated 4DCT-based ventilation maps. The method resulted in higher correlation coefficients for Patient 1 and Patient 2 (0.399 and 0.676) compared to other two image registration methods (0.289 and 0.586 for B-splines and 0.356 and 0.645 for deeds). For Patient 3 all methods achieved comparable results, with a slightly higher value for B-splines (0.265). On average the proposed method achieved the best score of 0.44, with deeds being the second highest-scoring method (0.4), while the lowest correlation was calculated for B-splines (0.38). Standard deviation of the determinant of the Jacobian of deformations, which can be seen as a measure of complexity of the deformations, for the presented method was on average 0.22, compared with 0.18 for B-splines and 0.23 for deeds. The results of the calculated correlations are shown in Figure 5.5. The correlations were calculated only for the regions being inside the presegmented lung masks, as the ventilation is not expected to be present outside of them.

In Figure 5.6 the influence of the regularisation parameter $\alpha$ from Eq. 5.3 is presented, which weights data and regularisation terms. The best performance of the proposed method in terms of Spearman's correlation coefficient has been observed for a value of $\alpha = 0.02$, which has been chosen for the presented experiments. The differences between the calculated correlations do not seem to be as high, which supports our point that the motion model based method in the application of pMRI/XeMRI to CT registration can potentially provide improved results in terms of correlation.

Finally, we decided to investigate the influence of the regularisation parameter $\alpha$

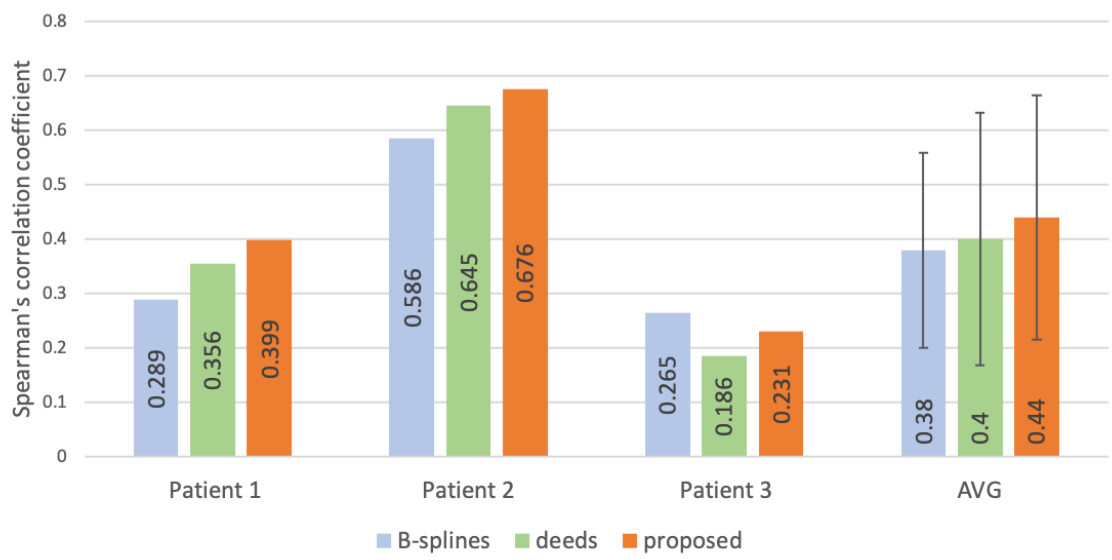



Figure 5.5: Spearmans correlation between CT-based estimated ventilation maps and XeMRI ventilation images for different pMRI-to-CT registration approaches, the global motion model approach and for the proposed method. The proposed method achieved the largest correlation. The correlations were calculated only for the regions being inside the presegmented lung masks, as the ventilation is not expected to be present outside of them.

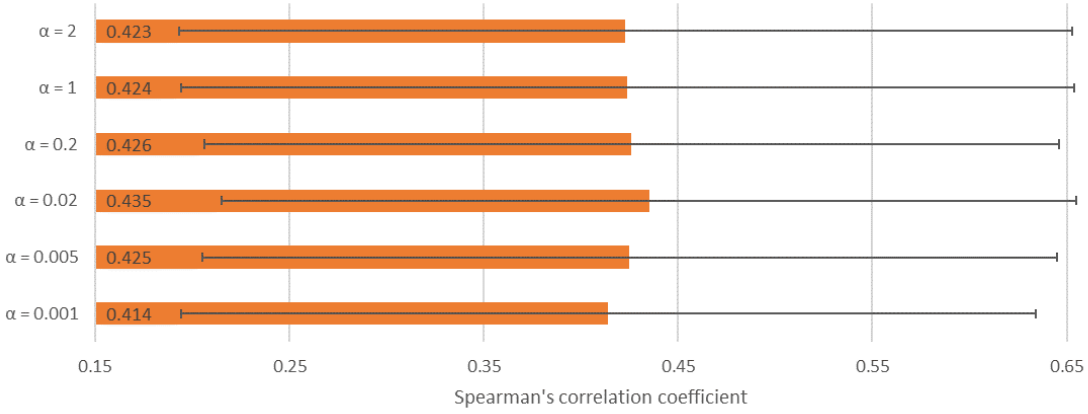



Figure 5.6: The influence of the regularisation parameter $\alpha$ on Spearman's correlation coefficient.

on the behaviour of the proposed method. The intuition would be that when regularisation parameter $\alpha$ increases, then the proposed method changes from a locally-oriented approach, with potentially different sets of motion model parameters assigned to every supervoxel, to a global method with a single set of parameters optimised for the whole image. To verify this hypothesis, we have looked at the distributions of estimated labels $l$, which is a vector $[l_x l_y l_z]$ of parameters, where $l_x, l_y, l_z$ are parameters of the model in the anterior-posterior, left-to-right and superior-to-inferior directions respectively. In Figure 5.7 we have presented their distribution for the optimal setting of $\alpha = 0.02$ and a high $\alpha = 2$, where almost all of the labels are equal. An interesting observation is that for the higher values of $\alpha$ two peaks in the histogram can be observed, when rather one was expected. This is because the left and right lungs in our dataset are not connected and a different set of motion model parameters was estimated for them. This point shows that a simple motion model for the whole image would not be an optimal solution in this case.

### 5.1.4 Discussion

We have presented a personalised model-driven method for pMRI-to-CT lung image registration. The proposed method was evaluated on three datasets of patients undergoing radiation treatment for lung cancer. The visual results showing estimated deformations presented in Figure 5.4, suggest that the proposed method better mimics the motion of the lungs. Changes in the direction of the motion estimated by B-splines and deeds, especially for the left lung, are unlikely to be present during breathing. Correlation was calculated between CT-based estimated ventilations and XeMRI brought into alignment with CT by the proposed method. On average, the proposed method outperformed other image registration approaches, as well as the global motion model, in terms of the correlation with ventilation maps estimated from 4DCT. The slightly lower score for Patient 3, compared to the registration

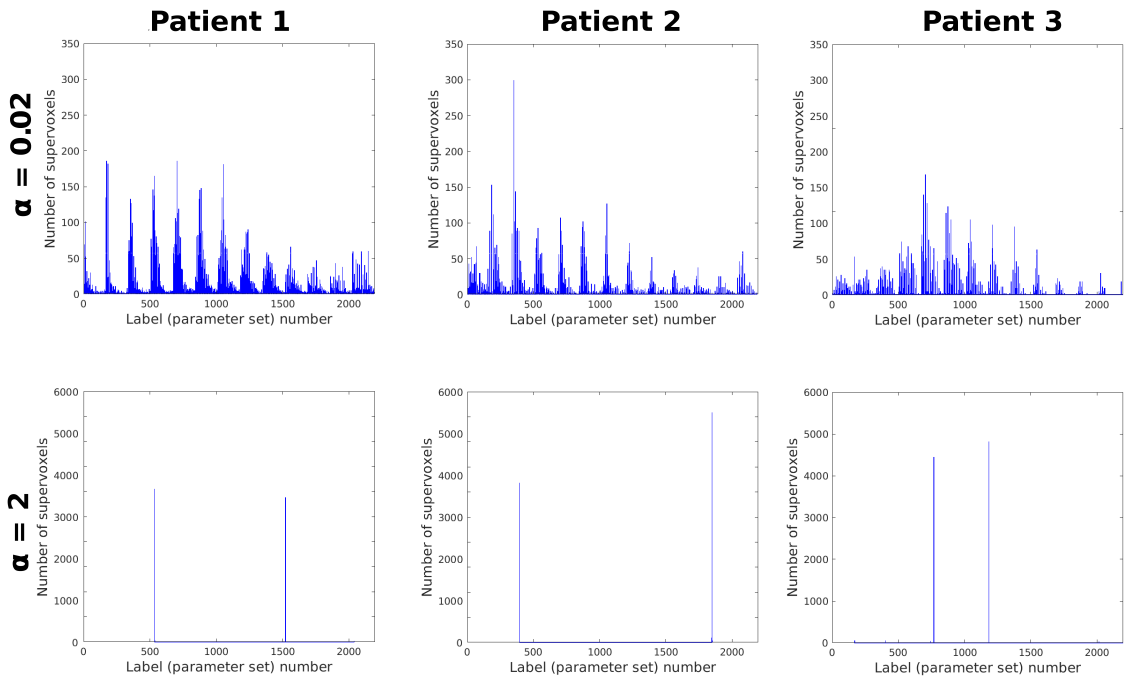



Figure 5.7: Distribution of the estimated labels representing motion model parameters for different settings of regularisation parameter $\alpha$. In the upper row the distribution of the estimated labels is shown for the optimal $\alpha = 0.02$ setting. When the $\alpha$ parameter is increased (lower row, $\alpha = 2$), the proposed method tends to show more global behaviour, where all supervoxels in a lung are assigned with the same label. Two peaks in the lower row show that different motion model parameters have been assigned to the left and right lung, as they were disjoint.

methods, was possibly caused by the fact that the difference in the lung volume between pMRI/XeMRI and CT was the lowest in this case. This observation seems to be supported by the fact that all the methods achieved comparable results for this patient.

One of the challenges in such methods is the lack of the ground truth or landmarks set in the corresponding images. Low out-of-plane resolution of pMRI and XeMRI (20mm), which is a few times larger than in CT (0.977mm), is another factor making the registration problem not a trivial one to address. The calculated correlation of XeMRI with 4DCT-based estimated ventilation maps resulted in the overall moderate correlation. One reason for that might be different breathing patterns in 4DCT compared to XeMRI/pMRI, related to physical properties of $^{129}$Xenon gas, which is much heavier than air. Ventilation maps based on 4DCT are estimated based on the changes of the tissue density, which should correspond to the lungs filling with air, however in practice they might provide complementary information [172]. Another limitation is the access to only one 4DCT scan of each patient. Therefore the proposed method might be prone to intra breathing cycle variations. This issue could potentially be eliminated by including more scans, such as diagnostic CT, of the same patient in the breathing motion creation step. The proposed motion model-based method also requires an accurate initial rigid registration. The upper parts of the lungs and apexes should be well aligned initially, or else the motion model-based registration might result in suboptimal performance. Such behaviour is imposed by the lung physiology and should not be considered as a limitation of the method.

As the proposed method has been evaluated on a limited number of cases, it should be primarily considered as a proof of concept. This work also starts a very important discussion about the quality of the registration between ventilation imaging and structural CT. This problem has not been fully developed and usually not much attention has been paid to this registration task, compared to for instance 4DCT used

to estimate ventilation maps. Our work shows that this task matters and we hope that the community becomes more aware of that. The presented method shows promising results for the challenging application of XeMRI to CT registration. The application of the PCA-based motion model in the deformable registration step of the framework appears to have the potential to help drive the registration for the regions of the lungs with insufficient amount of information. In the following Sec. 5.2 we propose a framework addressing the breathing motion problem in Dynamic Hyperpolarised $^{129}$Xenon Ventilation (DXeV) by applying an atlas-based image registration approach.

## 5.2 Atlas-based Breathing Motion Correction for Dynamic Lung XeMRI

This section presents a framework dedicated to compensate for the breathing motion in Dynamic Hyperpolarised $^{129}$Xenon Ventilation (DXeV), including preliminary results by aligning the images from different breathing phases, showing preliminary indications of improved accuracy for the ventilation analysis. This study has been previously presented as [183].

### 5.2.1 Introduction

In this thesis we presented so far methods for the analysis and applications of static hyperpolarised XeMRI. Such an imaging method enables the evaluation of lung ventilation in images acquired within a single breath-hold. The recently proposed DXeV [39] imaging technique extends these capabilities to capture ventilation volumes during breathing, thus creating a temporal sequence. The lung volume increases and decreases as the organ inflates and deflates. This breathing motion impacts on spatial correspondence between ventilation volumes of DXeV and therefore it impedes the temporal analysis of the localised ventilation. This work proposes a framework to compensate for the breathing motion for DXeV, aligning images from different breathing phases and, therefore, aiming to make the ventilation analysis more accurate. While motion compensation can be achieved through registration of lung masks, extracting such masks directly from DXeV is a challenge, since different lung regions receive inhaled gas at different times, or not at all if they are part of a ventilation defect. Alternatives include laborious manual annotation or an additional acquisition of structural data.

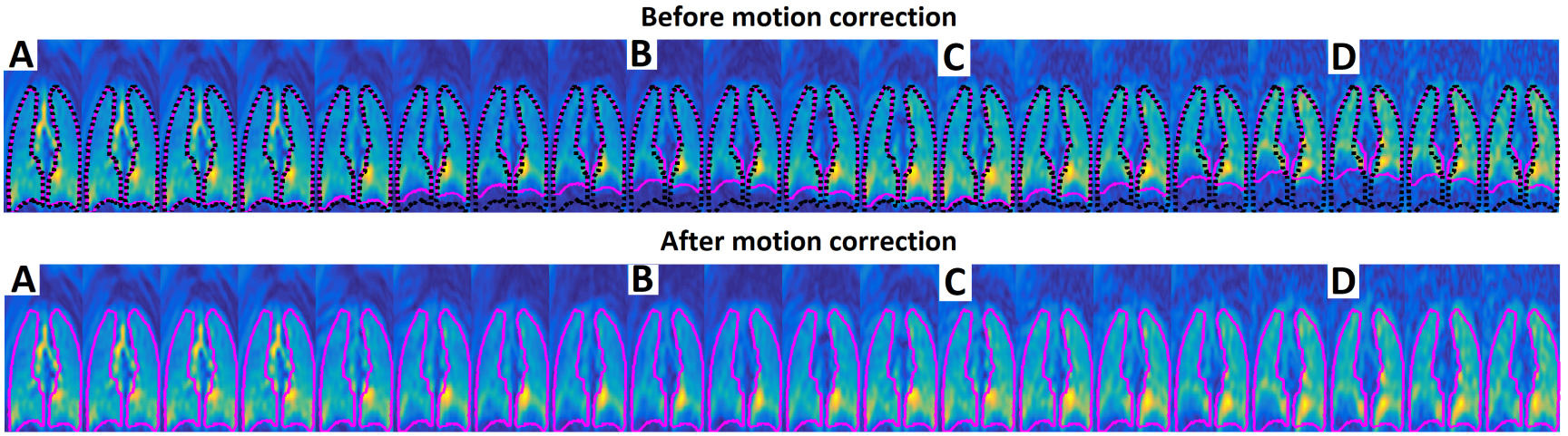



Figure 5.8: In upper row a coronal view of DXeV sequence for Subject 4 before motion correction is shown, with tidal breathing motion of the lungs, where the segmented lung mask is superimposed in black and the target lung atlas is shown in magenta. In lower row results after applying the proposed breathing motion correction for the same subject are shown, where the tidal breathing motion has been largely compensated and all masks correspond to each other spatially. The intensities are normalised for each of the time frames separately. For illustrative purpose, the last 20 ventilation volumes are presented, where breathing motion is most noticeable. From the sequence, the peak breathing phases are marked with letters A-D, and zoomed in the following figure. The presented results demonstrate consistency of our method, however, it was not possible to evaluate the accuracy of the registration.

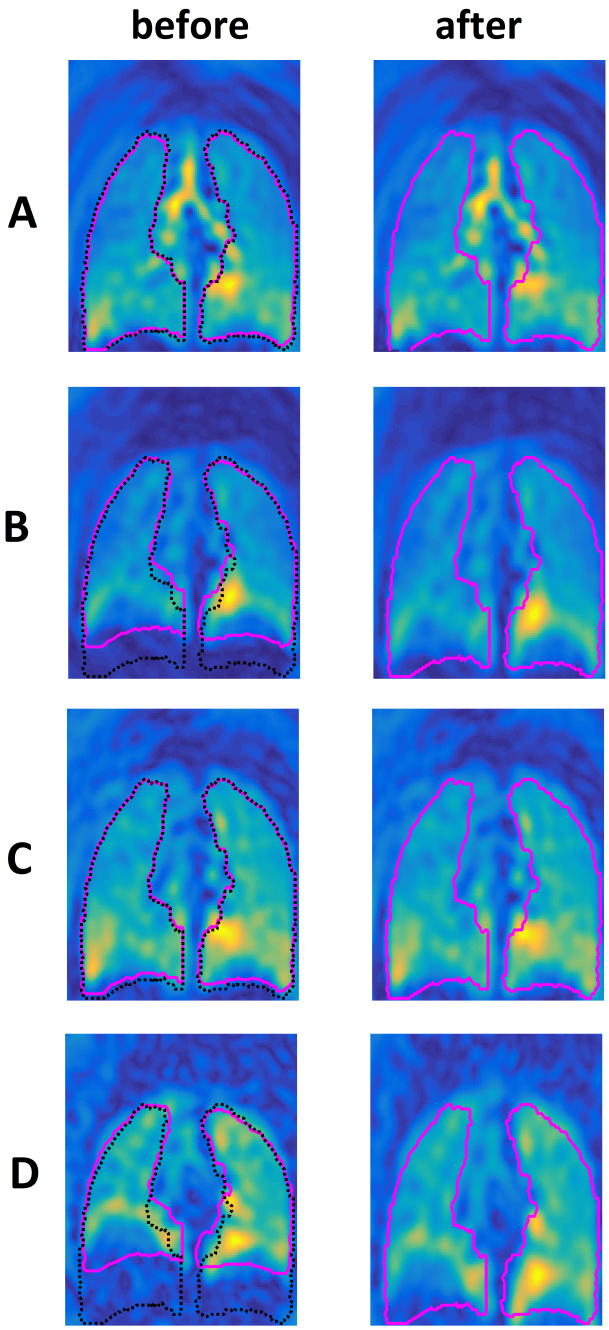



Figure 5.9: The peak breathing phases images (marked with letter A-D) from the previous figure are presented before and after applying the motion correction. The segmented lung mask is superimposed in black and the target lung atlas is shown in magenta before motion correction. The lung mask images after motion correction are shown in magenta in right column.

### 5.2.2 Methods

DXeV shows the temporal change of the hyperpolarised gas signal in the lungs, shown in Figure 5.8. In the following Figure 5.9 we zoom in the peak breathing phases images with superimposed lung masks. In the initial frames the patient breaths in the gas, therefore the gas starts to be visible in the trachea and in the following frames propagates through the airways to the remaining parts of the lungs. In the case of patients with ventilation defects this process might be interfered for instance by an airway blockage. Direct registration of such images between frames would be a suboptimal solution due to the aberrant gas distribution in the lungs. A work-flow diagram of the proposed method is presented in Figure 5.10. As no structural imaging data was available for the DXeV dataset, which would help to map all the frames into a common space, a lung atlas is built to delineate a plausible shape of the lungs in the ventilation images. The lung atlas was constructed from 10 CT volumes form the Dir-Lab dataset [119]. The atlas is further used to co-register all ventilation volumes in the sequence to the atlas. The lungs were initially segmented from peak inhale and peak exhale CT volumes using thresholding followed by morphological operations. The resulting 20 lung masks were co-registered using the Demons-based group-wise image registration method [184]. As a result an averaged shape of the lungs was created, which is subsequently used as a lung atlas.

The peak inhalation ventilation volume from DXeV was chosen as a reference volume in the dynamic sequence to which all other time frame volumes are aligned. This is because this volume provides the best definition of lung boundaries, and by representing the largest lung volume, it naturally includes the area of the lung in all other frames. On the chosen ventilation volume a thresholding was performed followed by morphological operations to segment an initial shape of the lungs. In this case, the application of affine image registration is more suitable than deformable registration. This is mainly due to difficulties in identification of the lung contours, so that when




Figure 5.10: Diagram presenting the work-flow of the proposed framework addressing breathing motion correction of DXeV. First a lung shape atlas is created based on 4DCT images. Then a reference frame from the temporal sequence is chosen and an initial segmentation of the lungs from this frame is performed. The atlas is registered with this segmentation. Subsequently, lungs are segmented from the remaining frames of the sequence, restricting the region of interest (ROI) only to the inside the lungs. After fitting the atlas to this segmentations, all of the images are aligned with the reference frame using affine registration, compensating for the breathing motion.

deformable image registration was applied, the results could be inconsistent with the actual lung shape. Subsequently, the atlas is registered to the reference ventilation volume of the DXeV using affine image registration.

A similar segmentation step to the previously described was performed on the remaining DXeV volumes from the sequence. This time the region of interest was restricted to the inside of the lung atlas. Since the atlas was registered with the peak inhalation volume, the remaining ventilation volumes of the lungs in the sequence should be localised inside of it. As a result, every time frame volume from the DXeV sequence has an initial shape of the lungs segmented. At this stage the difference between the segmented lungs for each ventilation volume from the sequence and the atlas were compensated for using an affine image registration method.

When lung image registration is driven by their mask in the surface matching manner, a few challenges need to be taken into account. The registration that aims to match surfaces of the lung masks requires carefully established regularisation, which would be responsible for estimating deformations of structures inside of the lungs, without their actual matching. This is a challenging task, as the lung motion is locally not smooth, and the magnitude of motion changes across the lungs. This would be especially important if deformable image registration was applied. Due to lack of structural information in the XeMRI images, we decided to apply affine

Table 5.1: Dice overlap calculated for masks of the lungs in the ventilation images with respect to the reference lung atlas for the full sequence, breath-hold phase and breathing phase, before and after motion correction. The temporal DXeV sequence was divided into two phases, when the subject holds their breath and when is breathing. The best results are highlighted in bold. For all cases the proposed motion correction method results in higher Dice overlap.

| Pat. no. | Full sequence before | Full sequence after | Breath-hold before | Breath-hold after | Breathing before | Breathing after |
|---|---|---|---|---|---|---|
| 1 | 0.930 | **0.990** | 0.905 | **0.978** | 0.958 | **0.996** |
| 2 | 0.940 | **0.999** | 0.951 | **0.999** | 0.933 | **0.999** |
| 3 | 0.960 | **0.998** | 0.984 | **0.998** | 0.941 | **0.997** |
| 4 | 0.958 | **0.997** | 0.962 | **0.997** | 0.956 | **0.997** |

registration, as a simplified approach. Having access to the corresponding pMRI data deformable approach, incorporating information about lung mask would be a more suitable choice.

### 5.2.3 Results

The proposed method was tested on a dataset of four DXeV sequences of healthy subjects, each consisting of 30 volumetric ventilation images. The original data, with clearly noticeable tidal breathing motion, is presented in Figure 5.8. The subjects initially take a breath in, hold their breath and then start to breath normally. This results in a peak inhalation observed in approximately the 10th frame from the sequence, and then in the two following peak inhalations in the later frames. The visualisation of the results of the proposed method is shown in the lower row of Figure 5.8, where initial misalignment is mostly compensated by the proposed method. It is noticeable that the registration to the atlas resulted in all volumes of the lungs

having visually similar appearance.

Subsequently, we divided the temporal DXeV sequence into two phases, when the subject holds their breath and when is breathing. After applying motion correction to the volumes we performed one more mask fitting to the volumes after motion correction. This was followed by calculation of Dice overlap between the mask of reference frame and the masks of the volume after motion correction. The results from Table 5.1 show that the proposed method results in close to 1 (average of 0.996) Dice overlap between the ventilation volumes and the reference volume.

The influence of the motion correction for landmarks placed inside the lungs is shown in Figures 5.11 and 5.12, where we present the results for Subject 3 (Figure 5.11) and Subject 4 (Figure 5.12), as for those the most tidal breathing motion has been observed. The reduction in the intensities of the ventilation images corresponding to the loss of the polarisation of the gas in the lungs is expected to be at an approximately constant rate. However, for the lower parts of the lungs (blue, red, pink, cyan) noticeable variations in the curve shapes can be observed, especially for the ventilation volumes pointed by arrows, which vary from the expected shape. The motion correction for the upper parts of the lungs (green, black, magenta, pink) results in almost no change in the intensity distribution, as little motion is observed there.

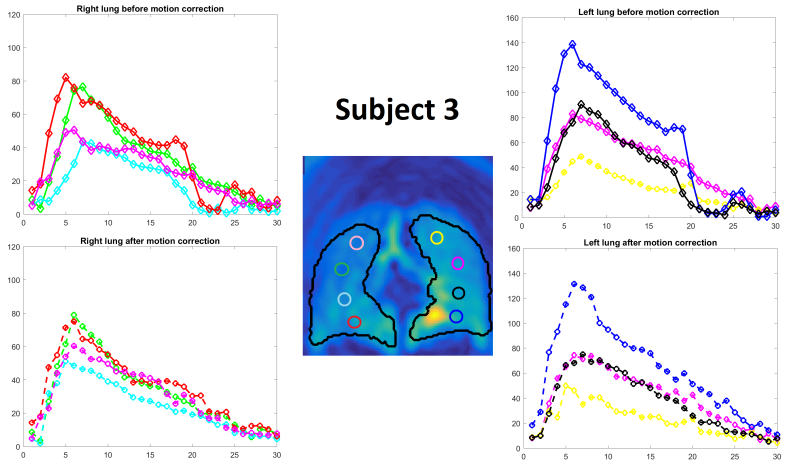



Figure 5.11: Distribution of the intensity changes for Subject 3. In the upper plots the distribution is plotted before applying motion correction for 8 landmarks. In the lower plots the distribution is plotted after applying motion correction for the same landmarks. Colours of the plots correspond to landmark locations

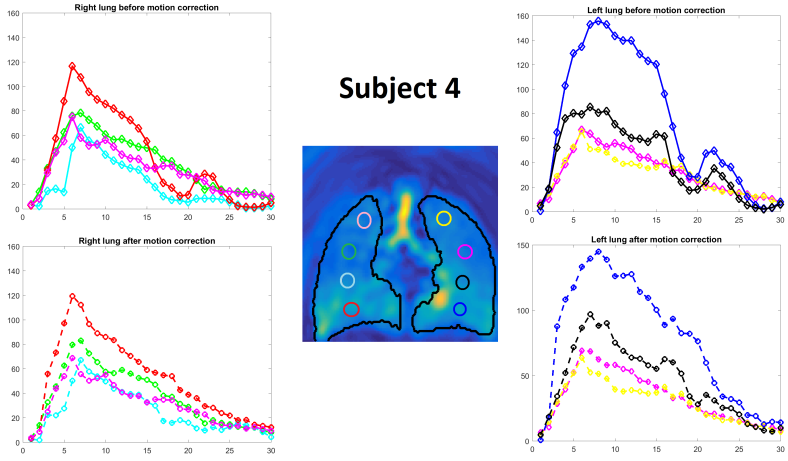



Figure 5.12: Distribution of the intensity changes for Subject 4. In the upper plots the distribution is plotted before applying motion correction for 8 landmarks. In the lower plots the distribution is plotted after applying motion correction for the same landmarks. Colours of the plots correspond to landmark locations

We have performed an additional experiment, which intended to measure the smoothness of the intensity changes in a voxel-wise manner across the time sequence. The norm of residuals after quadratic interpolation of the intensity sequence was

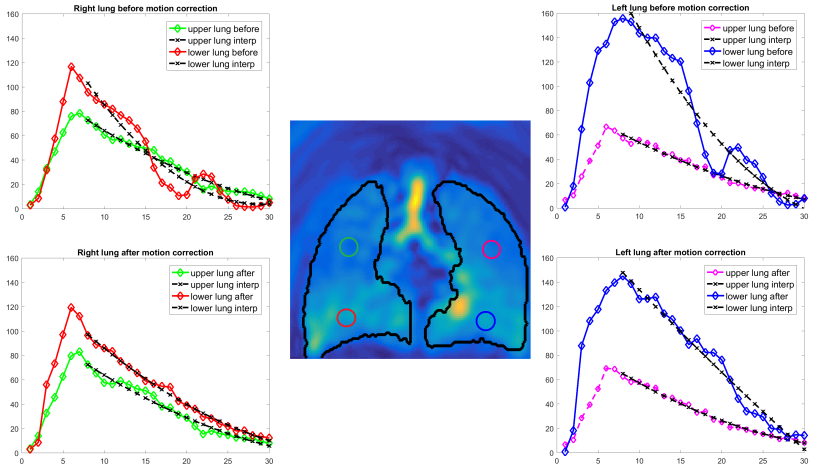



Figure 5.13: An example of fitting quadratic curve to the intensity sequences after the peak until the end. The results are presented for one of the subjects and for four points. The interpolated curves are plotted along the intensity sequence before motion correction (upper plots) and after motion correction was applied (lower plots). The landmarks for which the plots were made are pointed with corresponding colours on the XeMRI.

chosen as the approximate measure of the smoothness of the curves. This was performed on the decay part of the intensity plot, measured after the peak until the end. The visualisation of the interpolation results is presented in Figure 5.13, where the time intensity plots for four landmarks are shown in colours with corresponding fitted curves. It is noticeable that the curves are visually better fitted to the plots made for the landmarks after motion correction. We have performed the measurements for all four subjects and visualised the results in Figure 5.14. For all the subjects it is noticeable that images after motion corrections have lower magnitude of the residuals. The least pronounced difference is for Subject 2. The reason for that is that their breathing cycle was relatively shallow, and magnitude of the residuals was low even before the motion correction.

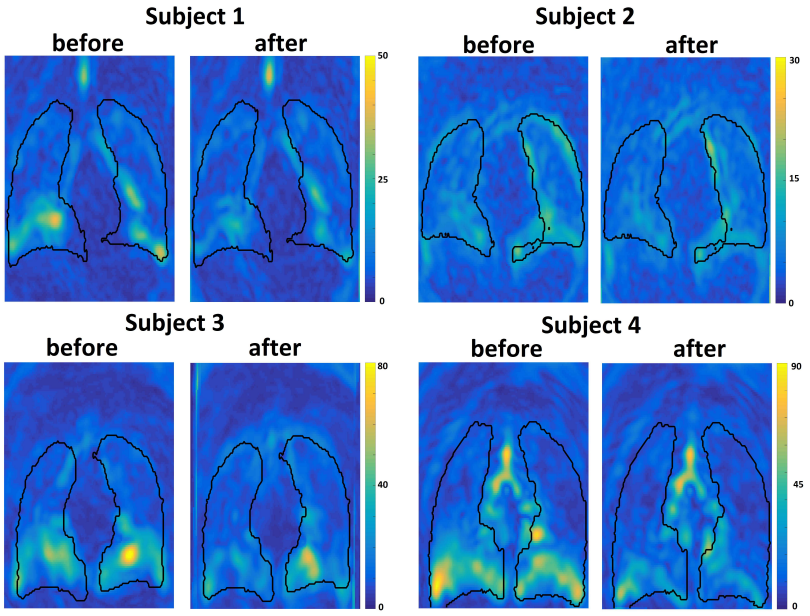



Figure 5.14: Magnitude of the norm of residuals after quadratic interpolation of the intensity sequences, measured after the peak until the end. For all of the subjects it is noticeable that image after performing the motion corrections has lower magnitude of the results. The least pronounced difference is for Subject 2.

### 5.2.4 Discussion

As shown in Figure 5.8 b) the proposed method compensates for the breathing motion in the dynamic sequence. The sudden changes in curves from Figure 5.11 and Figure 5.12, do not correspond to loss of the gas polarisation or any other physical property but are caused by misalignment of the subsequent ventilation volumes in the sequence. With the current evaluation approach, we were able to demonstrate that our registration framework provides consistent results. It was not possible to directly measure the accuracy of the registration of the proposed method. This work presented a framework for breathing motion correction in dynamic XeMRI. The presented results show that the analysis of the DXeV data without breathing motion correction might be prone to the spatial misalignment of the subsequent ventilation volumes and that the proposed breathing motion correction method shows the potential to compensate for such motion. An extension of this work could be to apply deformable

image registration. However, lack of structural images in the form of pMRI at all breathing stages limits this application. A different approach, using a motion model similar to what was presented in Sec. 5.1, appears to be more promising for this task. If additional 4DCT data were available for the patients, the method presented in the previous chapter would meet the requirements. The personalised motion model would be an interesting application for dynamic XeMRI. Otherwise, a motion model created on a separate dataset [111, 112] or even a FEM-based motion model, [116], could be used to substitute the applied affine image registration. The presented preliminary results show feasibility of the motion correction for the novel lung ventilation imaging modality. The dynamic XeMRI sequence after the breathing motion correction should allow for more accurate analysis of the temporal gas distribution in the lungs. Such an analysis could also be applied to to find different patterns in the gas distribution in time, which can potentially allow to localise different anatomical regions of the lungs.

## 5.3 Summary

In this chapter we have presented two methods addressing challenges related to motion correction for XeMRI. In the first section (Sec. 5.1) we have introduced a personalised motion model-based multi-modal image registration method. This method is dedicated to facilitate image registration between XeMRI and CT, where pMRI to CT registration is used as an intermediate step. This work addresses the problem of insufficient information inside the lungs in pMRI by applying a personalised 4D-CT statistical motion model. The motion model is estimated using PCA applied to the results of intra-patient image registration of 4DCT data. The framework compensates for local variations between images in the motion model parameters optimisation step. For every supervoxel an optimal set of motion model parameters are estimated. The

achieved results were compared with 4DCT-based ventilation maps estimated using state-of-the-art multi-modal image registration methods in terms of correlation with XeMRI. The method achieved on average higher correlation scores, showing that the application of PCA-based motion model in the deformable registration helps to drive the registration for the regions of the lungs with insufficient information, such as inside the lungs in pMRI.

In the second section (Sec. 5.2), we have proposed a framework for breathing motion correction in dynamic XeMRI. The analysis of such data without breathing motion correction might result in spatial misalignment of the subsequent ventilation volumes. The proposed framework addresses this issue by constructing a CT-based lung atlas, which helps to bring all ventilation time frame volumes into alignment. The atlas is initially fitted into a reference frame from the DXeV sequence and then all remaining volumes are co-registered to the reference frame. This results in compensating for most of the breathing motion and so should plausibly allow for a more accurate temporal analysis of the hyperpolarised $^{129}$Xenon gas intensity changes. All the registrations involved in this framework use affine transformations. The presented results show the necessity and demonstrate the feasibility of such motion correction for DXeV sequence.

# Chapter 6

# Conclusions

This chapter summarises the contributions of the thesis (in Sec. 6.1) and then provides an outlook (in Sec. 6.2) on how the project could be further developed.

## 6.1 Summary

This thesis has presented and discussed methods addressing the challenging problem of lung image registration and its further application to lung image analysis. The introduced methods take advantage of supervoxel-based image representation, which is particularly relevant when applied to lung imaging. Spatially and visually consistent regions represented by supervoxels have the potential to more accurately mimic the lung physiology during image analysis. The proposed approaches can be applied during the diagnosis and treatment of lung cancer. This thesis has discussed, in detail, several contributions to the medical imaging field, which can be applied to aid in the therapy of lung cancer; these are briefly summarised in this chapter.

The first contribution of this thesis is the introduction of a novel supervoxel-based image registration method, applying graph cuts as an optimisation technique and image guided filtering. In this method we have applied supervoxel-based image representation and posed the optimisation problem on a graph, where every supervoxel is

represented by a node, and all adjacent supervoxels are connected by an edge. Compared to a voxel-wise approach such a graph construction reduces the number of nodes required, specifically from the number of voxels to supervoxels, a significantly lower figure. We further introduced a graph relaxation criterion to preserve only the most relevant connections between supervoxels of similar mean intensity values, which, therefore, represent similar anatomical structures. The optimisation procedure results in the assignment of a displacement vector to every supervoxel, which estimates the optimal displacement of each individual supervoxel and, thus creates a displacement field. A direct application of the deformation estimated for a supervoxel would assign the exact same deformation to all voxels in each supervoxel. This is generally not physiologically accurate and may result in foldings at the borders of supervoxels that have different deformations, resulting in physiologically implausible deformations. We addressed this issue by applying guided image filtering, which smoothed the displacement field, while at the same time preserving some of the discontinuities. Such properties are especially important in lung image registration to help preserve 'sliding motion'. The method has been applied to CT-CT image registration and evaluated on a publicly available dataset of lung 4DCT images. We demonstrated how a supervoxel image representation combined with graph cuts-based optimisation can be applied to 3D data. We compared our method with other methods in the field and showed that ours performs favourably in terms of image registration accuracy. We showed that the application of a relaxed graph representation of the image, followed by guided image filtering over the estimated deformation field, allows us to model 'sliding motion'.

The second contribution is the introduction of an approach for estimating regional lung ventilation from dynamic lung CT imaging. The method combines a supervoxel-based image representation with deformable image registration, performed between peak breathing phases. After using clustering to over-segment the images represent-

ing peak exhale into supervoxels, we extract a number of layers of supervoxels, each with slightly different initial parameters. Subsequently, we apply the previously estimated deformations to the extracted supervoxels. The mean intensity values of each supervoxel is calculated to find its initial position. The mapping of the warped supervoxels to the peak inhale image allows us to calculate changes in their mean intensity between the mapping over the peak exhale and peak inhale breathing phase volumes. Finally, we average the results obtained for all the supervoxel layers. The ventilation is therefore estimated for spatially and visually consistent regions represented by supervoxels. Such a region-oriented approach is expected to be more physiologically consistent with lung anatomy than previous methods relying on voxel-wise relationships, as it has the potential to mimic the lung physiology. A dataset consisting of three lung cancer patients undergoing radiotherapy treatment enabled us to evaluate the method and compare the results with the most popular methods. We explored the influence of the size of the supervoxels as well as the number of supervoxel layers applied. Our results were compared with ventilation images acquired from XeMRI and achieved higher overall values of correlation than other methods for estimating ventilation surrogates used in the field.

This thesis has proposed further applications of lung image registration methods, presenting preliminary results of two methods addressing other challenges in XeMRI lung imaging and analysis. The first of these is pMRI/XeMRI to CT lung registration, which addresses challenges discussed in Chapter 2, such as sliding motion between the surfaces of the lungs, ribcage, and diaphragm, and most importantly the multi-modal nature of the images with directly incomparable intensities between them. We presented a novel method addressing the issue of insufficient information inside the lungs in pMRI, applying a personalised 4DCT statistical motion model for a supervoxel-based graphical image registration. The motion model is estimated based on the results of intra-patient image registration using PCA. The approach

using prior motion models is particularly important for regions where there is not enough information to reliably drive the registration process, as in the case of hyperpolarised $^{129}$Xenon MRI and proton density MRI to CT registration. Simultaneously, the method addresses local variations between images that can not be compensated globally by the motion model, in the supervoxel-based regional model parameters optimisation step. Furthermore, we presented a preliminary approach for motion correction of the dynamic sequence of hyperpolarised gas XeMRI. In that work a framework to compensate for the breathing motion for Dynamic Hyperpolarised $^{129}$Xenon Ventilation (DXeV) is proposed. The images from different breathing phases are aligned and this, therefore, increases the accuracy of the ventilation analysis process. A lung atlas is built to delineate a plausible shape of the lungs in the ventilation images. After applying the proposed breathing motion correction method the tidal breathing motion was found to have been largely compensated for and all masks of ventilation volumes corresponded to one another.

## 6.2 Outlook

### 6.2.1 Lobar Analysis Based on DXeV

In Sec. 5.2, we proposed a method to compensate for breathing motion in DXeV. The application of this method resulted in bringing all volumes from the temporal sequence into alignment. The aligned images may be further analysed using a supervoxel-oriented method, to identify distinct patterns in intensity changes. The hyperpolarised $^{129}$Xenon gas reaches different regions of the lungs at different time, due to the non-identical path of the airways. Our initial study has also identified the difference in intensity between the upper and lower parts of the lungs. Such variations may be useful for distinguishing anatomically consistent regions, such as the lobes of the lungs. We propose to perform the analysis after clustering of the temporal

sequence, to compensate for the potential suboptimal motion compensation and low spatial resolution of the data. The intensity of hyperpolarised gas found in the lungs, which can be estimated from the physical properties of the gas, allows us to elucidate the shape of intensity decay. Using this *a priori* knowledge, we can then perform parametrisation of the acquired measurements for every supervoxel. The lungs are anatomically divided into lobes: two in the left lung and three in the right. This is an additional piece of knowledge that should be incorporated into the framework. Based on this assumption, we can seek to identify a predefined number of different patterns. An example of the intensity distribution of DXeV for the left lung as a function of time is shown in Figure 6.1. Two main trends can be observed. In our pilot study we have applied a support vector machine (SVM) approach to separate the supervoxels into two groups; the visual results are presented in Figure 6.2. The resulting separation is anatomically reasonable and in the expected region. The challenging part of this project will be its application to patients with severe ventilation defects caused by lung cancer or COPD. As only subregions of the lungs will be visible in DXeV there will be a need to incorporate structural imaging.

### 6.2.2 CNN-based Lung Image Registration

The advent of convolution neural networks (CNNs) [93] in recent years has led to their applications in many computer vision and medical image analysis tasks, such as segmentation or classification. Coupled with the increasing interest in analysing large datasets, also known as the 'big data' approach, CNNs are creating new possibilities for medical imagine analysis. Image registration, however, remains a field where CNNs are still not frequently used and classic approaches are more popular. Most of the presented CNN-based approaches require training datasets with known ground truths of the deformations. Such an approach is difficult to achieve in many cases, particularly for lung registration, where the motion is complex and differs be-

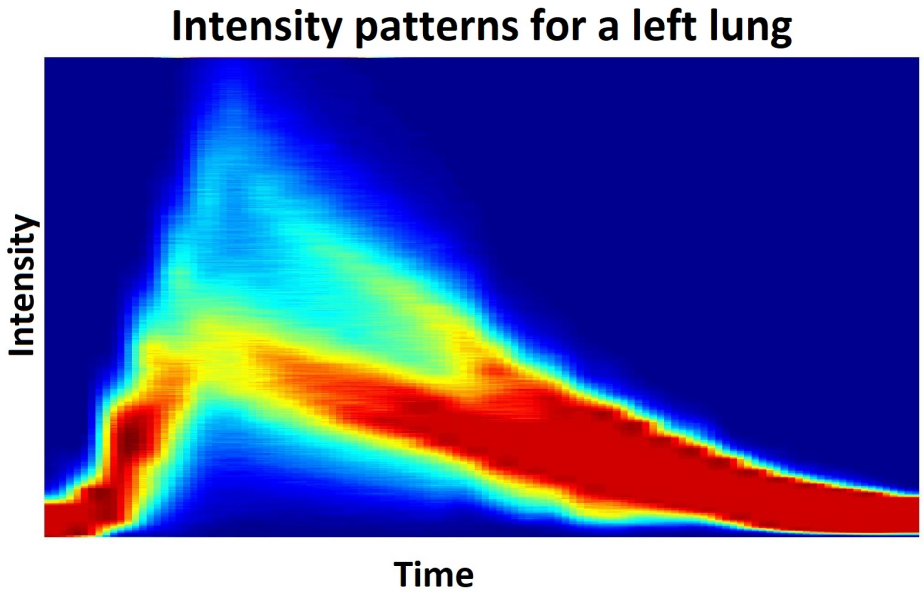



Figure 6.1: An example of the intensity distribution of DXeV for the left lung in time. Two main trend patterns can be observed. Colours correspond to the number of voxels having a particular value of intensity. Based on this distribution lung clustering into distinct regions can be performed.

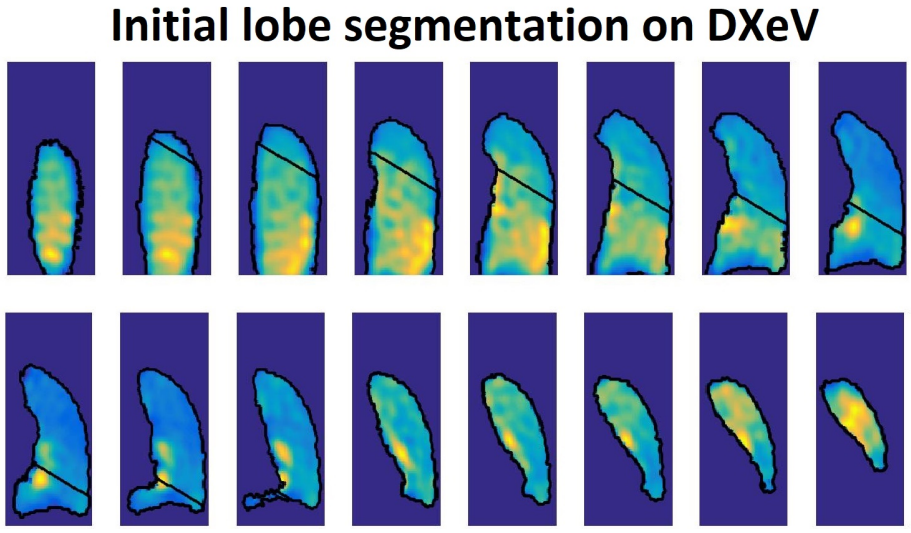



Figure 6.2: An example of series of cross-section in coronal view of DXeV or left lung with an estimated separation between regions with different ventilation patterns. Visual inspection suggests that the estimated separation is consistent with the anatomical border between the upper and lower lobes of the lung.

tween subjects, or even for the same person between different breathing cycles. Very recently, an unsupervised approach, where the ground truth deformations were not required, was presented in the form of a package called VoxelMorph [185], and showed comparable performance to the Ants registration method [48] when evaluated on brain image cases.

We would be interested in further exploring this area, particularly in its application to the lungs. This would involve the use of a task specific similarity measure, such as MIND [55]. The applied cost function would, just as in the VoxelMorph approach, penalise implausible deformations. The VoxelMorph approach provides deformations that are diffeomorphic. However, as in the lung application when modelling 'sliding motion', we expect to see discontinuities in the estimated displacement field and, hence, our method must address this issue.

One way of doing so would be to follow a well-established approach that involves masking the lungs, and imposing diffeomorphism inside and outside the lungs, thus allowing for discontinuities at the borders.

### 6.2.3 Multi-modal Lung Descriptor Learning

The approaches based on CNNs may be used not only to perform the alignment of the images, but also to find subject-specific multi-modal descriptors, which could be used in conventional image registration formulations. The existing methods for multi-modal image descriptors are mainly based on predefined formulations, such as MIND [55], scale-invariant feature descriptors (SIFT) [186] or DAISY [187] based on a predefined formulations. However, an approach to this task could also be based on learning the formulation from the data, using CNNs. Bringing images originating from different modalities into a common intensity space would make the registration problem easier to perform. The process would be in some sense similar to the encoding problem, this time without a limitation on the size of the output. The initial

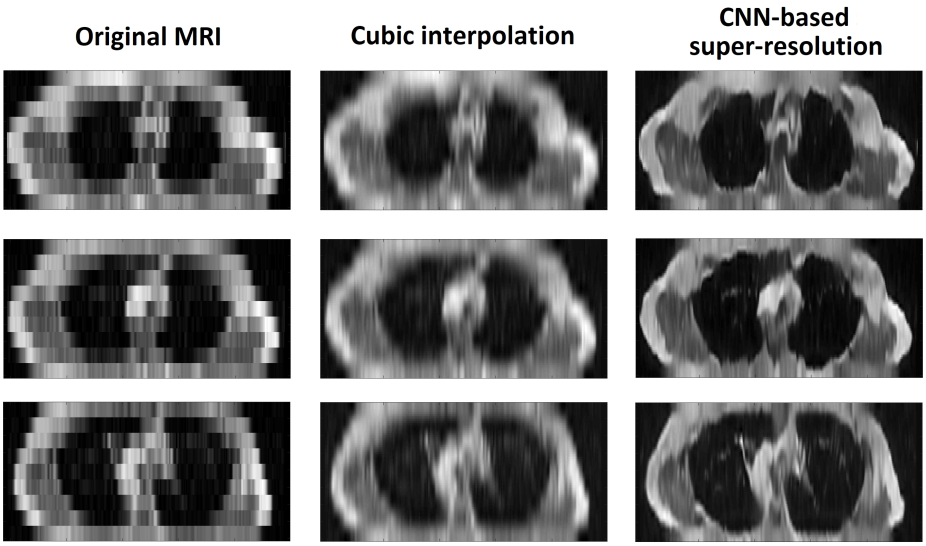



Figure 6.3: An example of the application of a CNN-based super-resolution method to lung pMRI. We show three different cross-sections in the axial direction. We compare the original pMRI in the left column with cubic interpolation images in the middle and the CNN-based super-resolution results on the right column. A noticeable improvement can be observed after the application of the CNN-based super-resolution method.

training would require patches of the same part of the body acquired with different imaging modalities. The best example of such medical data is The Visible Human Project [188], where images of male and female subjects acquired with a number of different modalities are available. The limited number of available training datasets would necessitate application of pre-trained networks and fine-tuning the parameters. An interesting path to explore would be to incorporate in the framework a content extraction approach, similar to the one presented in [189], where the content of an image was extracted along with different painting styles. The learned descriptors could be personalised by performing retraining during the registration on some parts of the images which are easier to register and considered to be already aligned. For instance, in the case of the lungs, less motion is observed in the upper parts of the lungs, but this gradually increases the further down the apex one observes. Therefore, after initial alignment upper parts ot the lungs could be used for fine-tuning of the

networks. In our work we have also observed higher registration errors in the lower parts of the lungs. Such a problem could be formulated in the form of the Generative Adversarial Network (GAN) framework [190], where additional networks would determine if the estimated descriptor is depicting the patch accurately enough.

In the case of pMRI to CT registration, this could be further enhanced by applying CNN-based super-resolution methods to pMRI. As pMRI is usually acquired with lower spatial resolution than CT scans, the images will require some pre-processing. One such method could be single image super-resolution [191], which we have tested on our pMRI dataset and achieved encouraging initial results. We present these in Figure 6.3, comparing them with the initial unprocessed data, as well as results obtained after applying cubic interpolation.

Medical image registration and image analysis in lung applications is a dynamic field, with new imaging methods and novel applications continually emerging. The methods presented in this thesis, along with the proposed directions of future work, have the potential to be applied in clinical practice, facilitating all steps of lung cancer treatment, where medical imaging is involved.

# References


[1] James H Scatliff and Peter J Morris. From röntgen to magnetic resonance imaging the history of medical imaging. *North Carolina medical journal*, 75(2):111–113, 2014.

[2] Klaus D Toennies. *Guide to medical image analysis*. Springer, 2017.

[3] Christopher J. L. Murray. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*, 385(9963):117–171, 2015.

[4] Jacques Ferlay, H.R. Shin, Freddie Bray, David Forman, C. Mathers, and D.M. Parkin. Cancer incidence and mortality worldwide: Iarc cancerbase no. *International Agency for Research on Cancer.*, 2, 01 2013.

[5] Cancer Research UK. Cancer incidence and mortality statistics for the UK. (May):1–8, 2015.

[6] Chee-Keong Toh, Fei Gao, Wan-Teck Lim, Swan-Swan Leong, Kam-Weng Fong, Swee-Peng Yap, Anne A.L. Hsu, Philip Eng, Heng-Nung Koong, Agasthian Thirugnanam, and et al. Differences between small-cell lung cancer and non-small-cell lung cancer among tobacco smokers. *Lung Cancer*, 56(2):161–166, 2007.

[7] Victor Kim and Gerard J. Criner. Chronic bronchitis and chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine*, 187(3):228–237, 2013.

[8] Tahreema N. Matin, X. Xu, Tom Doel, Vicente Grau, N. Rahman, A. Nickol, and Fergus V. Gleeson. COPD investigations. *Thorax*, 69(Suppl 2):2014–2016, 2014.

[9] London Cancer. Radiotherapy Guidelines for Treatment of Lung Cancer. 2014.

[10] National Institute for Health and Care Excellence. Lung cancer: diagnosis and management. (June), 2016.

[11] Rob H. Ireland, Bilal A Tahir, J. M. Wild, C. E. Lee, and M. Q. Hatton. Functional Image-guided Radiotherapy Planning for Normal Lung Avoidance. *Clinical Oncology*, 28(11):695–707, 2016.

[12] Yevgeniy Vinogradskiy, Austin Faught, Richard Castillo, Edward Castillo, Thomas Guerrero, Moyed Miften, and Arthur K. Liu. Using 4 DCT-ventilation to characterize lung function changes for pediatric patients getting thoracic radiotherapy. *Journal of Applied Clinical Medical Physics*, 2018.

[13] David B. Landau, Laura Hughes, Angela Baker, Andrew T. Bates, Michael C. Bayne, Nicholas Counsell, Angel Garcia-Alonso, Susan V. Harden, Jonathan D. Hicks, Simon R. Hughes, and et al. Ideal-crt: a phase 1/2 trial of isotoxic dose-escalated radiation therapy and concurrent chemotherapy in patients with stage ii/iii non-small cell lung cancer. *International Journal of Radiation Oncology* Biology* Physics*, 95(5):1367–1377, 2016.

[14] Catarina Veiga, David Landau, Jamie R. McClelland, Jonathan A. Ledermann, David Hawkes, Sam M. Janes, and Anand Devaraj. Long term radiologi-

cal features of radiation-induced lung damage. *Radiotherapy and Oncology*, 126(2):300–306, 2018.

[15] Erika J. Sims and David Price. Spirometry: an essential tool for screening, case-finding, and diagnosis of COPD. *Prim Care Respir J*, 21(2):128–130, 2012.

[16] Roberto P. Benzo, Sriram Paramesh, Sanjay A. Patel, William A. Slivka, and Frank C. Sciurba. Optimal protocol selection for cardiopulmonary exercise testing in severe COPD. *Chest*, 132(5):1500–1505, 2007.

[17] World Health Organization. World Health Statistics 2008. *World Health Organisation*, page 112, 2008.

[18] Willi A. Kalender. Computed tomography: fundamentals, system technology, image quality, applications. *Wiley VCH*, 2011.

[19] Hrvoje Lusic and Mark W Grinstaff. X-ray-computed tomography contrast agents. *Chemical reviews*, 113(3):1641–1666, 2012.

[20] James P. B. O'Connor, P. S. Tofts, K. A. Miles, L. M. Parkes, G. Thompson, and Alan Jackson. Dynamic contrast-enhanced imaging techniques: CT and MRI. *Br J Radiol*, 84(2), 2011.

[21] Brett A. Simon, Catherine Marcucci, Mansheung Fung, and Subhash R. Lele. Parameter estimation and confidence intervals for xe-ct ventilation studies: a monte carlo approach. *Journal of Applied Physiology*, 84(2):709–716, 1998.

[22] Catherine Marcucci, Daniel Nyhan, and Brett A. Simon. Distribution of pulmonary ventilation using Xe-enhanced computed tomography in prone and supine dogs. *Journal of Applied Physiology*, 90(2):421–430, 2001.

[23] Markus Nowak Lonsdale and Thomas Beyer. Dual-modality PET/CT instrumentation - today and tomorrow. *Eur J Radiol*, 73(3):452–460, 2010.

[24] Shan Tong, Adam M. Alessio, and Paul E. Kinahan. Image reconstruction for PET/CT scanners: past achievements and future challenges. *Imaging Med.*, 2(5):529–545, 2011.

[25] Habib Y. Baluwala, Laurent Risser, Julia A. Schnabel, and Kinda A. Saddi. Toward physiologically motivated registration of diagnostic CT and PET/CT of lung volumes. *Med Phys*, 40(2):021903, 2013.

[26] Katherina P. Farr, Stine Kramer, Azza A. Khalil, Anni Morsing, and Cai Grau. Role of perfusion SPECT in prediction and measurement of pulmonary complications after radiotherapy for lung cancer. *Eur J NuclMedMol Imaging*, 42:13151324, 2015.

[27] Ray H. Hashemi, William G. Bradley, and Christopher J. Lisanti. MRI: the basics. *Lippincott Williams and Wilkins*, 2010.

[28] Grzegorz Bauman, Michael Puderbach, Michael Deimling, Vladimir Jellus, Christophe Chefd'hotel, Julien Dinkel, Christian Hintze, Hans-Ulrich Kauczor, and Lothar R. Schad. Non-contrast-enhanced perfusion and ventilation assessment of the human lung by means of fourier decomposition in proton MRI. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 62(3):656–664, 2009.

[29] Grzegorz Bauman and Oliver Bieri. Matrix pencil decomposition of time-resolved proton MRI for robust and improved assessment of pulmonary ventilation and perfusion. *Magnetic resonance in medicine*, 77(1):336–342, 2017.

[30] Mitchell S. Albert, G. D. Cates, B. Driehuys, W. Happer, B. Saam, C. S. Springer, and Arnold Wishnia. Biological magnetic resonance imaging using laser-polarized $^{129}Xe$. *Nature*, 370(6486):199–201, 1994.

[31] John P. Mugler, Bastiaan Driehuys, James R. Brookeman, Gordon D. Cates, Stuart S. Berr, Robert G. Bryant, Thomas M. Daniel, Eduard E. De Lange, J. Hunter Downs, Christopher J. Erickson, William Happer, Denise P. Hinton, Neal F. Kassel, Therese Maier, C. Douglas Phillips, Brian T. Saam, Karen L. Sauer, and Mark E. Wagshul. MR imaging and spectroscopy using hyperpolarized 129Xe gas: Preliminary human results. *Magnetic Resonance in Medicine*, 37(6):809–815, 6 1997.

[32] Graham Norquay, Steven R. Parnell, Xiaojun Xu, Juan Parra-Robles, and Jim M. Wild. Optimized production of hyperpolarized $^{129}Xe$ at 2 bars for in vivo lung magnetic resonance imaging. *J Appl Phys*, 113(4), 2013.

[33] John P. Mugler and Talissa A. Altes. Hyperpolarized $^{129}Xe$ MRI of the human lung. *J Magn Reson Imaging*, 37(2):313–31, February 2013.

[34] Erika R. Ward, Laurence W. Hedlund, William C. Kurylo, Charles T. Wheeler, Gary P. Cofer, Mark W. Dewhirst, Lawrence B. Marks, and Zeljko Vujaskovic. Proton and hyperpolarized helium magnetic resonance imaging of radiation-induced lung injury in rats. *Int J Radiat Oncol*, 58(5):1562–1569, 2004.

[35] Edwin J. R. Van Beek, Jim M. Wild, Hans Ulrich Kauczor, Wolfgang Schreiber, John P. Mugler, and Eduard E. De Lange. Functional MRI of the lung using hyperpolarized 3-helium gas. *J Magn Reson Imaging*, 20(4):540–554, 2004.

[36] Jim M. Wild, Salma Ajraoui, Martin H. Deppe, Steven R. Parnell, Helen Marshall, Juan Parra-Robles, and Rob H. Ireland. Automatic image registration of lung CT and hyperpolarized helium-3 MRI via mutual information of proton MRI. *NMR Biomed*, 24(2):130–134, 2011.

[37] Miranda Kirby, Sarah Svenningsen, Alexei Ouriadov, Harvey O. Coxson, and David G. Mccormack. Hyperpolarized $^{3}He$ and $^{129}Xe$ MR imaging in healthy

volunteers and patients with chronic obstructive pulmonary disease. *Radiology*, 265(2), 2012.

[38] Neil Woodhouse, Jim M. Wild, Martyn N. J. Paley, Stanislao Fichele, Zead Said, Andrew J. Swift, and Edwin J. R. Van Beek. Combined helium-3/proton magnetic resonance imaging measurement of ventilated lung volumes in smokers compared to never-smokers. *J Magn Reson Imaging*, 21(4):365–369, 2005.

[39] Ozkan Doganay, Tahreema N. Matin, Anthony Mcintyre, Brian Burns, Rolf F. Schulte, Fergus V Gleeson, and Daniel Bulte. Fast dynamic ventilation MRI of hyperpolarized 129Xe using spiral imaging. *Magnetic resonance in medicine*, 79(5):2597–2606, 2018.

[40] Bastiaan Driehuys, Santiago Martinez-Jimenez, Zackary I. Cleveland, Gregory M. Metz, Denise M. Beaver, John C. Nouls, S. Sivaram Kaushik, Rafael Firszt, Christine Willis, Kevin T. Kelly, Jan Wolber, Monica Kraft, and H. Page McAdams. Chronic obstructive pulmonary disease: safety and tolerability of hyperpolarized $^{129}$Xe MR imaging in healthy volunteers and patients. *Radiology*, 262(1):279–89, 2012.

[41] Youbing Yin, Eric a Hoffman, and Ching-Long Lin. Mass preserving non-rigid registration of CT lung images using cubic B-spline. *Medical physics*, 36(9):4213–4222, 2009.

[42] Alexander Schmidt-Richberg, Rene Werner, Heinz Handels, and Jan Ehrhardt. Estimation of slipping organ motion by registration with direction-dependent regularization. *Med Image Anal*, 16:150–159, 2012.

[43] Laurent Risser, Francois Xavier Vialard, Habib Y. Baluwala, and Julia A. Schnabel. Piecewise-diffeomorphic image registration: application to the motion es-

timation between 3D CT lung images with sliding conditions. *Med Image Anal*, 17(2):182–193, 2013.

[44] Danielle Pace, Stephen Aylward, and Marc Niethammer. A locally adaptive regularization based on anisotropic diffusion for deformable image registration of sliding organs. *IEEE Trans. Med. Imaging*, 32:2114–2126, 2013.

[45] The Royal College of Radiologiests. On target : ensuring geometric accuracy in radiotherapy. 2008.

[46] Julia A. Schnabel, Mattias P. Heinrich, Bartłomiej W. Papież, and Sir Michael Brady. Advances and challenges in deformable image registration: From image fusion to complex motion modelling. *Med. Image Anal.*, 33(10):145–148, 2016.

[47] Gerardo Hermosillo, Christophe Chefd'Hotel, and Olivier Faugeras. Variational methods for multimodal image matching. *International Journal of Computer Vision*, 50(3):329–343, 2002.

[48] Brian Avants, C. Epstein, M. Grossman, and J. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal*, 12:26 41, 2008.

[49] Gang Song and Nicholas Tustison. Lung CT image registration using diffeomorphic transformation models. *Medical Image Analysis for the Clinic: A Grand Challenge*, pages 23–32, 2010.

[50] Paul Viola and William M. III Wells. Alignment by maximization of mutual information. *Int J Comput Vis*, 24(2):137–154, 1997.

[51] Frederik Maes, Andre Collignon, Dirk Vandermeulen, Guy Marchal, and Paul Suetens. Multi-modality image registration by maximization of Mutual Information. *IEEE Trans. Med. Imag*, vol. 16:pp. 187–198, 1997.

[52] Colin Studholme, Derek L. G. Hill, and David J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognit*, 32(1):71–86, jan 1999.

[53] Alfred O. Hero, Bing Ma, Olivier J. J. Michel, and John Gorman. Applications of entropic spanning graphs. *IEEE Signal Processing Mag*, 19(5):85–95, 2002.

[54] Josien P W Pluim, J. B. Antoine Maintz, and Max A. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Trans. Med. Imag*, 22(8):986–1004, aug 2003.

[55] Mattias P. Heinrich, Mark Jenkinson, Manav Bhushan, Tahreema Matin, Fergus V. Gleeson, Michael Brady, and Julia A. Schnabel. MIND: modality independent neighbourhood descriptor for multi-modal deformable registration. *Med Image Anal*, 16(7):1423–35, October 2012.

[56] John Ashburner and Karl J. Friston. Rigid body registration. *Statistical parametric mapping: The analysis of functional brain images*, pages 49–62, 2007.

[57] Torsten Rohlfing. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. *IEEE Trans. Med. Imag*, 31(2):153–163, 2012.

[58] Berthold K.P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[59] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. Deformable medical image registration: a survey. *IEEE Trans. Med. Imag*, 32(7):1153–1190, 2013.

[60] Jean-Philippe Thirion. Image matching as a diffusion process: an analogy with Maxwell's demons. *Med Image Anal*, 2(3):243–260, 1998.

[61] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Non-parametric diffeomorphic image registration with the demons algorithm. *MICCAI*, 10(Pt 2):319–326, 2007.

[62] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Symmetric log-domain diffeomorphic registration: a demons-based approach. *MICCAI*, 11(Pt 1):754–761, 2008.

[63] Vincent Arsigny, Olivier Commowick, Xavier Pennec, and Nicholas Ayache. A log-Euclidean framework for statistics on diffeomorphisms. *MICCAI*, 9(Pt 1):924–931, 2006.

[64] Bartłomiej W. Papież, Mattias P. Heinrich, Jerome Fehrenbach, Laurent Risser, and Julia A. Schnabel. An implicit sliding-motion preserving regularisation via bilateral filtering for deformable image registration. *Med Image Anal*, 18(8):1299–311, December 2014.

[65] Eldad Haber and Jan Modersitzki. Intensity gradient based registration and fusion of multi-modal images. *Methods Inf Med*, 46(3):292–9, 2007.

[66] Jan Modersitzki. Numerical Methods for Image Registration . 2004.

[67] Daniel Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and David J. Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imag*, 18(8):712–21, August 1999.

[68] Julia A. Schnabel, Daniel Rueckert, Marcel Quist, Jane M. Blackall, Andy D. Castellano-Smith, Thomas Hartkens, Graeme P. Penney, Walter A. Hall, Haiying Liu, Charles L. Truwit, Frans A. Gerritsen, Derek L. G. Hill, and David J. Hawkes. A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations. *MICCAI*, 2208:573–581, 2001.

[69] Stefan Klein, Marius Staring, Keelin Murphy, Max A. Viergever, and Josien P.W. Pluim. Elastix: a toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, 2010.

[70] Stefan Klein, Josien PW Pluim, Marius Staring, and Max A Viergever. Adaptive stochastic gradient descent optimisation for image registration. *International Journal of Computer Vision*, 81(3):227, 2009.

[71] Marc Modat, Gerard R. Ridgway, Zeike A. Taylor, Manja Lehmann, Josephine Barnes, David J. Hawkes, Nick C. Fox, and Sébastien Ourselin. Fast free-form deformation using graphics processing units. *Computer methods and programs in biomedicine*, 98(3):278–284, 2010.

[72] Ziji Wu, Eike Rietzel, Vlad Boldea, David Sarrut, and Gregory C. Sharp. Evaluation of deformable registration of patient lung 4DCT with subanatomical region segmentations. *Medical physics*, 35(2):775–781, 2008.

[73] Vivien Delmon, Simon Rit, Romulo Pinho, David Sarrut, and V. Delmon. Direction dependent b-splines decomposition for the registration of sliding objects. In *Proceedings of the Fourth International Workshop on Pulmonary Image Analysis*, pages 45–55, 2011.

[74] Vivien Delmon, Simon Rit, Romulo Pinho, and David Sarrut. Registration of sliding objects using direction dependent b-splines decomposition. *Physics in Medicine & Biology*, 58(5):1303, 2013.

[75] Rui Hua, J.M. Pozo, Z.A. Taylor, and A.F. Frangi. Multiresolution eXtended Free-Form Deformations (XFFD) for non-rigid registration with discontinuous transforms. *Medical Image Analysis*, 36:113–122, November 2016.

[76] Vladlena Gorbunova, Jon Sporring, Pechin Lo, Martine Loeve, Harm A. Tiddens, Mads Nielsen, Asger Dirksen, and Marleen de Bruijne. Mass preserving image registration for lung CT. *Med Image Anal*, 16(4):786–795, 2012.

[77] Yuri Boykov and Gareth Funka-Lea. Graph cuts and efficient N-D image segmentation. *Int. J. Comput. Vis.*, 70(2):109–131, 2006.

[78] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient graph-based image segmentation. *Int. J. Comput. Vis.*, 59(2):167–181, 2004.

[79] Ben Glocker, Nikos Komodakis, Georgios Tziritas, Nassir Navab, and Nikos Paragios. Dense image registration through MRFs and efficient linear programming. *Med Image Anal*, 12(6):731–741, 2008.

[80] Dana Cobzas and Abhishek Sen. Random walks for deformable image registration. *Med. Image Comput. Comput.-Assist. Intervent. (MICCAI 2011)*, pages 557–565, 2011.

[81] Mattias P. Heinrich, Mark Jenkinson, Michael Brady, and Julia A. Schnabel. MRF-based deformable registration and ventilation estimation of lung CT. *IEEE Trans. Med. Imag*, 32(7):1239–48, July 2013.

[82] Mattias P. Heinrich, Heinz Handels, and Ivor J. A. Simpson. Estimating large lung motion in COPD patients by symmetric regularised correspondence fields. *Med. Image Comput. Comput.-Assist. Intervent. (MICCAI 2015)*, 9350:338–345, 2015.

[83] Mattias P. Heinrich, Ivor J.A. Simpson, Bartłomiej W. Papież, Sir Michael Brady, and Julia A. Schnabel. Deformable image registration by combining uncertainty estimates from supervoxel belief propagation. *Med. Image Anal.*, 27:57–71, 2016.

[84] Jan Rühaak, Thomas Polzin, Stefan Heldmann, Ivor J. A. Simpson, Heinz Handels, Jan Modersitzki, and Mattias P. Heinrich. Estimation of Large Motion in Lung CT by Integrating Regularized Keypoint Correspondences into Dense Deformable Registration. *IEEE Trans. Med. Imag.*, 36(8):1746 – 1757, 2017.

[85] Ashish Raj and Ramin Zabih. A graph cut algorithm for generalized image deconvolution. *IEEE Int. Conf. Comput. Vis.*, 2:1048–1054, 2005.

[86] Yuri Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26:1124 – 1137, 2004.

[87] Junhwan Kim, Vladimir Kolmogorov, and Ramin Zabih. Visual correspondence using energy minimization and mutual information. *IEEE Int. Conf. Comput. Vis.*, pages 1033–1040 vol.2, Oct 2003.

[88] Tommy W. H. Tang and Albert C. S. Chung. Non-rigid image registration using graph-cuts. *Med. Image Comput. Comput.-Assist. Intervent. (MICCAI 2007)*, 10(Pt 1):916–924, 2007.

[89] Ronald W. K. So, Tommy W. H. Tang, and Albert C. S. Chung. Non-rigid image registration of brain magnetic resonance images using graph-cuts. *Pattern Recognit.*, 44(10-11):2450–2467, 2011.

[90] Herve Lombaert, Yiyong Sun, and Farida Cheriet. Landmark-based Non-rigid Registration via Graph Cuts. *Lec. Notes Comput. Sc.*, 2007.

[91] Ronald W. K. So and Albert C. S. Chung. Non-rigid image registration by using graph-cuts with mutual information. *IEEE Image Proc.*, Sept(1522-4880):4429–4432, 2010.

[92] Ronald W. K. So and Albert C. S. Chung. Learning-based non-rigid image registration using prior joint intensity distributions with graph-cuts. *IEEE Image Proc.*, 2:717–720, 2011.

[93] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, pages 1097–1105, 2012.

[94] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[95] Shun Miao, Z. Jane Wang, and Rui Liao. A CNN regression approach for real-time 2D/3D registration. *IEEE Transactions on Medical Imaging*, 35(5):1352–1363, 2016.

[96] Koen A.J. Eppenhof, Maxime W. Lafarge, Pim Moeskops, Mitko Veta, and Josien P.W. Pluim. Deformable image registration using convolutional neural networks. In *Medical Imaging 2018: Image Processing*, volume 10574, page 105740S. International Society for Optics and Photonics, 2018.

[97] Marius Staring, Stefan Klein, and Josien P.W. Pluim. A rigidity penalty term for nonrigid registration. *Medical physics*, 34(11):4098–4108, 2007.

[98] Habib Y Baluwala, Kinda A. Saddi, and Julia A. Schnabel. Elastic registration of chest CT images with log un-biased deformations and rigidity constraint. *Proceedings - International Symposium on Biomedical Imaging*, 5:1235–1238, 2011.

[99] William E. Higgins, Ronnarit Cheirsilp, Xiaonan Zang, Patrick Byrnes, Trevor Kuhlengel, Thomas Allen, Jennifer Toth, and Rebecca Bascom. Pet/ct image registration for multimodal bronchoscopic procedure planning and guidance.

In *B103. ADVANCES IN DIAGNOSTIC BRONCHOSCOPY*, pages A4383–A4383. American Thoracic Society, 2016.

[100] André Collignon, Frederik Maes, Dominique Delaere, Dirk Vandermeulen, Paul Suetens, and Guy Marchal. Automated multi-modality image registration based on information theory. In *Information processing in medical imaging*, volume 3, pages 263–274, 1995.

[101] John N.H. Brunt. Computed tomography-magnetic resonance image registration in radiotherapy treatment planning. *Clinical Oncology*, 22(8):688–697, 2010.

[102] Martin J. Menten, Andreas Wetscherek, and Martin F. Fast. MRI-guided lung SBRT: present and future developments. *Physica Medica*, 2017.

[103] Laurent Risser, Mattias P. Heinrich, Daniel Rueckert, and Julia A. Schnabel. Multi-modal diffeomorphic registration using mutual information: application to the registration of CT and MR pulmonary images. *4th International Workshop on Pulmonary Image Analysis*, pages 35–44, 2011.

[104] Huanxiang Lu, M. Reyes, A. Serifovic-Trbalic, S. Weber, Y. Sakurai, H. Yamagata, and P. Cattin. Multimodal diffeomorphic demons registration based on point-wise mutual information. *IEEE International Symposium on Biomedical Imaging*, pages 372–375, 2010.

[105] Marc Modat, Jamie R. McClelland, and Sébastien Ourselin. Lung registration using the niftyreg package. *Medical image analysis for the clinic-a grand Challenge*, 2010:33–42, 2010.

[106] Rob H. Ireland, Chris M. Bragg, Mark McJury, Neil Woodhouse, Stan Fichele, Edwin J.R. van Beek, Jim M. Wild, and Matthew Q. Hatton. Feasibility of Im-

age Registration and Intensity-Modulated Radiotherapy Planning With Hyperpolarized Helium-3 Magnetic Resonance Imaging for NonSmall-Cell Lung Cancer. *International Journal of Radiation Oncology*Biology*Physics*, 68(1):273 – 281, 2007.

[107] Rob H. Ireland, Salma Ajraoui, Martin H. Deppe, Steven R. Parnell, Helen Marshall, Juan Parra-Robles, and Jim M. Wild. Automatic image registration of lung CT and hyperpolarized helium-3 MRI via mutual information of proton MRI. *NMR Biomed*, 24(2):130–134, 2011.

[108] Bilal A. Tahir, Andrew J. Swift, Helen Marshall, Juan Parra-Robles, Matthew Q. Hatton, Ruth Hartley, Richard Kay, Christopher E. Brightling, Wim Vos, Jim M. Wild, and Rob H. Ireland. A method for quantitative analysis of regional lung ventilation using deformable image registration of CT and hybrid hyperpolarized gas/(1)H MRI. *Physics in medicine and biology*, 59(23):7267–77, dec 2014.

[109] Brian Avants, N. J. Tustison, G. Song, P. Cook, A. A. Klein, and J. C. Gee. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage*, 54:20332044, 2011.

[110] Lindsay Mathew, Andrew Wheatley, Richard Castillo, Edward Castillo, George Rodrigues, Thomas Guerrero, and Grace Parraga. Hyperpolarized $^{3}$He magnetic resonance imaging. comparison with four-dimensional X-ray computed tomography imaging in lung cancer. *Acad Radiol*, 19(12):1546–1553, 2012.

[111] Jamie R. McClelland, Jane M. Blackall, Ségolene Tarte, Adam C Chandler, Simon Hughes, Shahreen Ahmad, David B. Landau, and David J. Hawkes. A continuous 4d motion model from multiple respiratory cycles for use in lung radiotherapy. *Medical Physics*, 33(9):3348–3358, 2006.

[112] Jamie R. McClelland, D. J. Hawkes, T. Schaeffter, and A. P. King. Respiratory motion models: A review. *Med. Image Anal.*, 17(1):19–42, 2013.

[113] Jan Ehrhardt, Ren Werner, Alexander Schmidt-Richberg, and Heinz Handels. Statistical Modeling of 4D Respiratory Lung Motion Using Diffeomorphic Image Registration. *IEEE Trans. Med. Imaging*, 30:251–265, 02 2011.

[114] Xiaoxiao Liu, Ipek Oguz, Stephen M. Pizer, and Gig S. Mageras. Shape-correlated deformation statistics for respiratory motion prediction in 4D lung. *Proc.SPIE*, 7625:7625 – 7625 – 10, 2010.

[115] Qinghui Zhang, Alex Pevsner, Agung Hertanto, YuChi Hu, Kenneth E. Rosenzweig, C. Clifton Ling, and Gig S. Mageras. A patient specific respiratory model of anatomical motion for radiation treatment planning. *Med. Phys.*, 34(12):4772–4781, 2007.

[116] Lianghao Han, Hua Dong, Jamie R. McClelland, Liangxiu Han, David J. Hawkes, and Dean C. Barratt. A hybrid patient-specific biomechanical model based image registration method for the motion estimation of lungs. *Med. Image Anal.*, 39:87–100, 2017.

[117] Sven Kabus, Tobias Klinder, Keelin Murphy, Bram van Ginneken, Cristian Lorenz, and Josien P. W. Pluim. Evaluation of 4D-CT lung registration. *MICCAI*, pages 747–754, 2009.

[118] Keelin Murphy, Bram Van Ginneken, Joseph M. Reinhardt, Sven Kabus, Kai Ding, Xiang Deng, Kunlin Cao, Kaifang Du, Gary E. Christensen, Vincent Garcia, Tom Vercauteren, Nicholas Ayache, Olivier Commowick, Grgoire Malandain, Ben Glocker, Nikos Paragios, Nassir Navab, Vladlena Gorbunova, Jon Sporring, Marleen De Bruijne, Xiao Han, Mattias P. Heinrich, Julia A. Schnabel, Mark Jenkinson, Cristian Lorenz, Marc Modat, Jamie R. McClelland,

Sebastien Ourselin, Sascha E. A. Muenzing, Max A. Viergever, Dante De Nigris, D. Louis Collins, Tal Arbel, Marta Peroni, Rui Li, Gregory C. Sharp, Alexander Schmidt-Richberg, Jan Ehrhardt, Rene Werner, Dirk Smeets, Dirk Loeckx, Gang Song, Nicholas Tustison, Brian Avants, James C. Gee, Marius Staring, Stefan Klein, Berend C. Stoel, Martin Urschler, Manuel Werlberger, Jef Vandemeulebroucke, Simon Rit, David Sarrut, and Josien P W Pluim. Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. *IEEE Trans. Med. Imag*, 30(11):1901–1920, 2011.

[119] Richard Castillo, Edward Castillo, Rudy Guerra, Valen E. Johnson, Travis McPhail, Amit K. Garg, and Thomas Guerrero. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Phys Med Biol*, 54(7):1849–1870, 2009.

[120] Edward Castillo, Richard Castillo, Josue Martinez, Maithili Shenoy, and Thomas Guerrero. Four-dimensional deformable image registration using trajectory modeling. *Physics in Medicine and Biology*, 55(1):305–327, 2010.

[121] Jef Vandemeulebroucke, Simon Rit, Jan Kybic, Patrick Clarysse, and David Sarrut. Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. *Medical physics*, 38(1):166–178, 2011.

[122] John Kipritidis, Michael Hofman, Shankar Siva, Jason Callahan, Pierre-Yves Le Roux, Henry Woodruff, William Counter, and Paul Keall. Estimating lung ventilation directly from 4D CT Hounsfield unit values. *Medical Physics*, 43:33–43, 01 2016.

[123] Gilles Defraene, Wouter van Elmpt, Wouter Crijns, Pieter Slagmolen, and Dirk De Ruysscher. CT characteristics allow identification of patient-specific sus-

ceptibility for radiation-induced lung damage. *Radiotherapy and Oncology*, 117(1):29–35, 2015.

[124] Alison J. Cole, Catarina Veiga, Ursula Johnson, Derek D'Souza, Narinder Lalli, and Jamie R. McClelland. Toward adaptive radiotherapy for lung patients: Feasibility study on deforming planning CT to CBCT to assess the impact of anatomical changes on dosimetry. *Physics in medicine and biology*, 2018.

[125] Richard Castillo, Edward Castillo, J. Martinez, and T. Guerrero. Ventilation from Four Dimensional Computed Tomography: Density versus Jacobian Methods. *Medical Physics*, 55:4661–4685, 2010.

[126] Thomas Guerrero, Kevin Sanders, Josue Noyola-Martinez, Edward Castillo, Yin Zhang, Richard Tapia, Rudy Guerra, Yerko Borghero, and Ritsuko Komaki. Quantification of regional ventilation from treatment planning CT. *International Journal of Radiation Oncology*Biology*Physics*, 62(3):630 – 634, 2005.

[127] Tokihiro Yamamoto, Sven Kabus, Tobias Klinder, Cristian Lorenz, Jens von Berg, Thomas Blaffert, Billy W. Loo, and Paul J. Keall. Investigation of four-dimensional computed tomography-based pulmonary ventilation imaging in patients with emphysematous lung regions. *Phys. Med. Biol.*, 56(7):2279–98, 2011.

[128] Joseph M. Reinhardt, Gary E. Christensen, Eric A. Hoffman, Kai Ding, and Kunlin Cao. Registration-derived estimates of local lung expansion as surrogates for regional ventilation. In *Information Processing in Medical Imaging*, pages 763–774, 2007.

[129] Joseph M. Reinhardt, Kai Ding, Kunlin Cao, Gary E. Christensen, Eric A. Hoffman, and Shalmali V. Bodas. Registration-based estimates of local lung

tissue expansion compared to xenon CT measures of specific ventilation. *Medical Image Analysis*, 12(6):752 – 763, 2008.

[130] Tokihiro Yamamoto, Sven Kabus, Tobias Klinder, Jens von Berg, Cristian Lorenz, Billy W. Loo, and Paul J. Keall. Four-dimensional computed tomography pulmonary ventilation images vary with deformable image registration algorithms and metrics. *Med. Phys.*, 38(2011):1348–1358, 2011.

[131] John Kipritidis, Shankar Siva, Michael S. Hofman, Jason Callahan, Rodney J. Hicks, and Paul J. Keall. Validating and improving CT ventilation imaging by correlating with ventilation 4D-PET/CT using 68Ga-labeled nanoparticles. *Medical physics*, 41(1), 2014.

[132] Bilal A. Tahir, P.J.C. Hughes, S.D. Robinson, H. Marshall, N.J. Stewart, G. Norquay, A. Biancardi, H-F Chan, G.J. Collier, K.A. Hart, and et al. Spatial comparison of CT-based surrogates of lung ventilation with hyperpolarized Helium-3 and Xenon-129 gas MRI in patients undergoing radiation therapy. *International Journal of Radiation Oncology Biology Physics*, 2018.

[133] K. Du, J. E. Bayouth, K. Ding, G. E. Christensen, K. Cao, and J. M. Reinhardt. Reproducibility of intensity-based estimates of lung ventilation. *Medical Physics*, 40(6):063504, 2013.

[134] Wei Shao, E. Sarah, Yue Gerard, Taylor J. Pan, Patton, Joseph M. Reinhardt, Oguz C., John E. Bayouth Durumeric, and Gary E. Christensen. Sensitivity analysis of jacobian determinant used in treatment planning for lung cancer. *Proc.SPIE Medical imaging*, 10574:10574 – 10574 – 9, 2018.

[135] Kaifang Du, J. E. Bayouth, K. Cao, G. E. Christensen, K. Ding, and J. M. Reinhardt. Reproducibility of registration-based measures of lung tissue expansion. *Medical Physics*, 39(3):1595–1608, 2012.

[136] Adam. Szmul, Bartłomiej W. Papież, Russell Bates, Andre Hallack, Julia A. Schnabel, and Vicente Grau. Graph Cuts-Based Registration Revisited: A Novel Approach for Lung Image Registration Using Supervoxels and Image-Guided Filtering. *Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, pages 592–599, 2016.

[137] Adam Szmul, Bartłomiej W. Papież, Andre Hallack, Vicente Grau, and Julia A. Schnabel. Supervoxels for graph cuts-based deformable image registration using guided image filtering. *J Electron Imaging*, 26(6):061607, 2017.

[138] Michael R. Kaus, Thomas Netsch, Sven Kabus, Vladimir Pekar, Todd McNutt, and Bernd Fischer. Estimation of Organ Motion from 4D CT for 4D Radiation Therapy Planning of Lung Cancer. *Med. Image Comput. Comput.-Assist. Intervent. (MICCAI 2004)*, 3217:1017–1024, 2004.

[139] Elisabeth Weiss, Krishni Wijesooriya, S. Vaughn Dill, and Paul J. Keall. Tumor and normal tissue motion in the thorax during respiration: Analysis of volumetric and positional variations using 4D CT. *Med. Image Comput. Comput.-Assist. Intervent. (MICCAI 2004)*, 67:296307, 2007.

[140] Andrea Vedaldi and Stefano Soatto. Quick shift and kernel methods for mode seeking. *ECCV Lec. Notes Comput. Sc.*, 5305(4):705–718, 2008.

[141] Vaïa Machairas, Matthieu Faessel, David Cárdenas-Peña, Théodore Chabardes, Thomas Walter, and Etienne Decencière. Waterpixels. *IEEE Trans. Image Proc.*, 24(11):3707–3716, 2015.

[142] Brian Fulkerson, Aandrea Vedaldi, and S. Soatto. Class segmentation and object localization with superpixel neighborhoods. *IEEE Int. Conf. Comput. Vis.*, pages 670–677, Sept 2009.

[143] Hanguen Kim, Sangwon Lee, Dongsung Lee, Soonmin Choi, Jinsun Ju, and Hyun Myung. Real-Time Human Pose Estimation and Gesture Recognition from Depth Images Using Superpixels and SVM Classifier. *Sensors*, 15(6):12410–12427, 2015.

[144] Lawrence C. Zitnick and Sing Bing Kang. Stereo for image-based rendering using image over-segmentation. *Int. J. Comput. Vision*, 75(1):49–65, October 2007.

[145] Beyang Liu, S. Gould, and D. Koller. Single image depth estimation from predicted semantic labels. *Comput. Vis. Pattern Recognit.*, pages 1253–1260, June 2010.

[146] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012.

[147] Robert Sedgewick. Algorithms in c, part 5: Graph algorithms. *Addison-Wesley, Reading, Massachusetts*, 2001.

[148] Kaiming He, Jian Sun, and Xiaoou Tang. Guided Image Filtering. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(6):1397–1409, 2013.

[149] Jan Rühaak, S. Heldmann, T. Kipshagen, and B. Fischer. Highly accurate fast lung CT registration. *SPIE Med. Imag.*, 86690Y, 2013.

[150] Jef Vandemeulebroucke, Olivier Bernard, Simon Rit, Jan Kybic, Patrick Clarysse, and David Sarrut. Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT. *Med Phys*, 39(2):1006, 2012.

[151] Yuri. Boykov and M.P. Jolly. Interactive graph cuts for optimal boundary region segmentation of objects in N-D images. *ICCV 2001*, 1(July):105–112, 2001.

[152] Bartłomiej W. Papież, Jamie Franklin, Mattias P. Heinrich, Fergus V. Gleeson, and Julia A. Schnabel. Liver motion estimation via locally adaptive over-segmentation regularization. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 427–434. Springer International Publishing, 2015.

[153] Norman R Draper and Harry Smith. *Applied regression analysis*, volume 326. John Wiley & Sons, 1998.

[154] Trevor Hastie, Robert Tibshirani, Jerome Friedman, and James Franklin. The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, 27(2):83–85, 2005.

[155] Carlo Tomasi and Roberto Manduchi. Bilateral filtering for gray and color images. *Proceedings of the Sixth International Conference on Computer Vision (ICCV- 98)*, pages 839–846, 1998.

[156] Gary E. Christensen and Hans J. Johnson. Consistent image registration. *IEEE transactions on medical imaging*, 20(7):568–582, 2001.

[157] Mingli Chen, Weiguo Lu, Quan Chen, Kenneth J Ruchala, and Gustavo H Olivera. A simple fixed-point approach to invert a deformation field. *Med. Phys.*, 35(1):81–88, 2008.

[158] Ryan E. Amelon, Kunlin Cao, Joseph M. Reinhardt, Gary E. Christensen, and Madhavan L. Raghavan. A measure for characterizing sliding on lung boundaries. *Annals of Biomedical Engineering*, 42(3):642–650, 2014.

[159] Lars Knig and Jan Rühaak. A fast and accurate parallel algorithm for non-linear image registration using normalized gradient fields. *IEEE International Symposium on Biomedical Imaging*, April 2014.

[160] Bartłomiej W Papież, James M Franklin, Mattias P Heinrich, Fergus V Gleeson, Michael Brady, and Julia A Schnabel. GIFTed Demons: deformable image registration with local structure-preserving regularization using supervoxels for liver applications. *Journal of Medical Imaging*, 5(2):024001, 2018.

[161] S Vivekanandan, DB Landau, N Counsell, DR Warren, A Khwanda, SD Rosen, E Parsons, Y Ngai, L Farrelly, L Hughes, et al. The impact of cardiac radiation dosimetry on survival after radiation therapy for non-small cell lung cancer. *International Journal of Radiation Oncology* Biology* Physics*, 99(1):51–60, 2017.

[162] L Bendall, BW Papiez, MA Hawkins, and JD Fenwick. Isotoxic dose escalation with real-time imaging on an mr-linac in lung radiation therapy. *International Journal of Radiation Oncology Biology Physics*, 102(3):S207, 2018.

[163] Aulien Lucchi, Kevin Smith, Radhakrishna Achanta, Graham Knott, and Pascal Fua. Supervoxel-based segmentation of mitochondria in em image stacks with learned shape features. *IEEE Transactions on Medical Imaging*, 31(2):474–486, 2012.

[164] D. Mahapatra, P. J. Schffler, J. A. W. Tielbeek, J. C. Makanyanga, J. Stoker, S. A. Taylor, F. M. Vos, and J. M. Buhmann. Automatic Detection and Segmentation of Crohn's Disease Tissues From Abdominal MRI. *IEEE Transactions on Medical Imaging*, 32(12):2332–2347, Dec 2013.

[165] Weijie Chen, Maryellen Giger, and Ulrich Bick. A Fuzzy C-Means (FCM)-Based Approach for Computerized Segmentation of Breast Lesions in Dynamic Contrast-Enhanced MR Images. *Academic radiology*, 13:63–72, 02 2006.

[166] Benjamin Irving, James Franklin, Bartłomiej W. Papież, Ewan Anderson, Ricky Sharma, Fergus V. Gleeson, Michael Brady, and Julia A. Schnabel. Pieces-of-parts for supervoxel segmentation with global context: Application to dce-mri tumour delineation. *Medical Image Analysis*, 32:69–83, 03 2016.

[167] Po Su, Jianhua Yang, Hai Li, Linda Chi, Zhong Xue, and Stephen T. Wong. Superpixel-Based Segmentation of Glioblastoma Multiforme from Multimodal MR Images. *Multimodal Brain Image Analysis*, pages 74–83, 2013.

[168] Helen Winter, Jolanta Mirecka, Bartłomiej Papież, Benjamin Irving, Mark Jenkinson, Michael Chappell, Ricky A. Sharma, and Julia A. Schnabel. Clinical feasibility of texture analysis in DCE MRI of patients receiving selective internal radiation therapy. *European Journal of Surgical Oncology*, 44:S30 – S31, 2018.

[169] Andre Hallack, Bartłomiej W. Papież, James Wilson, Lai Mun Wang, Tim Maughan, Mark J. Gooding, and Julia A. Schnabel. Correlating Tumour Histology and ex vivo MRI Using Dense Modality-Independent Patch-Based Descriptors. *Patch-Based Techniques in Medical Imaging*, pages 137–145, 2015.

[170] Minjeong Kim, Guorong Wu, Qian Wang, Seong-Whan Lee, and Dinggang Shen. Improved image registration by sparse patch-based deformation estimation. *NeuroImage*, 105:257 – 268, 2015.

[171] Bilal A. Tahir, Cedric Van Holsbeke, Rob H. Ireland, Andrew J. Swift, Felix C. Horn, Helen Marshall, John C. Kenworthy, Juan Parra-Robles, Ruth Hartley,

Richard Kay, and et al. Comparison of CT-based lobar ventilation with 3He MR imaging ventilation measurements. *Radiology*, 278(2):585–592, 2015.

[172] Leith J. Rankine, Ziyi Wang, Bastiaan Driehuys, Lawrence B. Marks, Chris R. Kelsey, and Shiva K. Das. Correlation of Regional Lung Ventilation and Gas Transfer to Red Blood Cells: Implications for Functional-Avoidance Radiation Therapy Planning. *International Journal of Radiation Oncology Biology Physics*, 2018.

[173] Adam Szmul, Bartłomiej W. Papież, Tahreema Matin, Fergus Gleeson, Julia A. Schnabel, and Vicente Grau. Regional lung ventilation estimation based on supervoxel tracking. *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*, 10576:10576 – 10576 – 7, 2018.

[174] Adam Szmul, Tahreema Matin, Fergus Gleeson, Julia A. Schnabel, Vicente Grau, and Bartłomiej W. Papież. XeMRI to CT Lung Image Registration Enhanced with Personalized 4DCT-derived Motion Model. *MICCAI Workshops: Thoracic Image Analysis*, 2018.

[175] Paul J. Roach, Denis J. Gradinscak, Geoffrey P. Schembri, Elizabeth A. Bailey, Kathy P. Willowson, and Dale L. Bailey. SPECT/CT in V/Q Scanning. *Seminars in Nuclear Medicine*, 40(6):455 – 466, 2010.

[176] Shankar Siva, Roshini Thomas, Jason Callahan, Nicholas Hardcastle, Daniel Pham, Tomas Kron, Rodney J. Hicks, Michael P. MacManus, David L. Ball, and Michael S. Hofman. High-resolution pulmonary ventilation and perfusion PET/CT allows for functionally adapted intensity modulated radiotherapy in lung cancer. *Radiotherapy and Oncology*, 115(2):157–162, 2015.

[177] Adrien Depeursinge, Alejandro Vargas, Alexandra Platon, Antoine Geissbuhler, Pierre-Alexandre Poletti, and Henning Müller. Building a reference multime-

dia database for interstitial lung diseases. *Computerized medical imaging and graphics*, 36(3):227–238, 2012.

[178] Tahreema N. Matin, Najib Rahman, Annabel H. Nickol, Mitchell Chen, Xiaojun Xu, Neil Stewart, Tom Doel, Vicente Grau, James M. Wild, and Fergus V. Gleeson. Chronic Obstructive Pulmonary Disease: Lobar Analysis with Hyperpolarized 129 Xe MR Imaging. *Radiology*, 282:857–868, 2016.

[179] Enid M. Eslick, John Kipritidis, Denis Gradinscak, Mark J. Stevens, Dale L. Bailey, Benjamin Harris, Jeremy T. Booth, and Paul J. Keall. Ct ventilation imaging derived from breath hold ct exhibits good regional accuracy with galligas pet. *Radiotherapy and Oncology*, 127(2):267–273, 2018.

[180] James M. Wild, H. Marshall, M. Bock, L.R. Schad, P.M. Jakob, M. Puderbach, F. Molinari, E.J.R. Van Beek, and Jürgen Biederer. MRI of the lung (1/3): methods. *Insights into imaging*, 3(4):345–353, 2012.

[181] Jürgen Biederer, M. Beer, W. Hirsch, James M. Wild, M. Fabel, M. Puderbach, and E.J.R. Van Beek. MRI of the lung (2/3). Why when how? *Insights into imaging*, 3(4):355–371, 2012.

[182] Jürgen Biederer, S. Mirsadraee, M. Beer, F. Molinari, C. Hintze, G. Bauman, M. Both, E.J.R. Van Beek, James M. Wild, and M. Puderbach. MRI of the lung (3/3)-current applications and future perspectives. *Insights into imaging*, 3(4):373–386, 2012.

[183] Adam Szmul, Bartłomiej W. Papież, Ozkan Doganay, Fergus Gleeson, Daniel Bulte, Julia A. Schnabel, and Vicente Grau. Atlas-based breathing motion correction for dynamic lung XeMRI. *ISMRM*, 2018.

[184] Bartłomiej W. Papież, Bogdan J. Matuszewski, Lik-Kwan Shark, and Wei Quan. Facial expression recognition using diffeomorphic image registration

framework. *Mathematical Methodologies in Pattern Recognition and Machine Learning*, pages 179–194, 2013.

[185] Guha Balakrishnan, Amy Zhao, Mert R. Sabuncu, John Guttag, and Adrian V. Dalca. An Unsupervised Learning Model for Deformable Medical Image Registration. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 9252–9260, 2018.

[186] David G Lowe. Object recognition from local scale-invariant features. *IEEE International Conference on Computer vision*, 2:1150–1157, 1999.

[187] Engin Tola, Vincent Lepetit, and Pascal Fua. Daisy: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5):815–830, 2010.

[188] U.S. National Library of Medicine (NLM). The Visible Human Project. *https* : *//www.nlm.nih.gov/research/visible/visible_human.html*, 1995.

[189] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2414–2423, 2016.

[190] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, pages 2672–2680, 2014.

[191] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1646–1654, 2016.