1-1-2013

Quantitative Analysis of Thoracic Computed Tomography Images

Gang Song
University of Pennsylvania, songgang97@gmail.com

Follow this and additional works at: http://repository.upenn.edu/edissertations

Part of the Computer Sciences Commons

Recommended Citation
Song, Gang, "Quantitative Analysis of Thoracic Computed Tomography Images" (2013). Publicly Accessible Penn Dissertations. 803. http://repository.upenn.edu/edissertations/803

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/edissertations/803
For more information, please contact libraryrepository@pobox.upenn.edu.
Quantitative Analysis of Thoracic Computed Tomography Images

Abstract
Lung disease has risen to the third leading cause of chronic morbidity and mortality in the United States. The diagnosis, differentiation, and classification of the severity of various lung diseases rely on clinical assessment, thoracic imaging using computed tomography (CT), and pulmonary function testing (PFT). While being the reference standard for assessment of the lung's mechanical function, PFT strictly permits a global measurement of lung physiology.

In contrast, high-resolution image analysis is a powerful tool with the potential for regional as well as global quantification of diseases. Imaging plays an increasingly important role in lung disease diagnosis. Most current pulmonary imaging techniques are used clinically to assess anatomic changes and to provide qualitative or semi-quantitative estimates of disease severity. Although generally effective, radiologic interpretation of CT images is time-consuming, requiring considerable expertise. It is also largely qualitative and prone to intra-observer diagnostic variability. It is therefore desirable to have automated quantitative analysis from imaging modalities.

This dissertation investigates the problem of using thoracic computed tomography images for automatically and quantitatively analyzing and diagnosing lung diseases. We discuss how to complete a holistic automatic pipeline for clinical studies and study its various components. The final goal is to illustrate how to build a complete pipeline using CT images as input and yielding the desired clinical results as output. In the case when a single lung image is provided, we propose an algorithm for segmenting small airways from CT images. When the images at both inspiration and expiration phases are available, we discuss the use of image registration algorithms to compute lung kinematics, especially focusing on various diffeomorphic transform models. Finally, these image-derived quantitative metrics are analyzed for clinical studies of specific diseases, including differentiation of interstitial lung diseases and chronic obstructive pulmonary diseases, and quantification of small airway air trapping and emphysema. We show that quantitative CT imaging, integrating segmentation, registration, feature computation, feature selection and pattern recognition, can provide better biomarkers for diagnosis and prognosis.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Computer and Information Science

First Advisor
James Gee

Second Advisor
Jianbo Shi

This dissertation is available at ScholarlyCommons: http://repository.upenn.edu/edissertations/803
QUANTITATIVE ANALYSIS OF THORACIC COMPUTED
TOMOGRAPHY IMAGES

Gang Song

A DISSERTATION
in
Computer and Information Science

Presented to the Faculties of the University of Pennsylvania
in
Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2013

Supervisor of Dissertation
James Gee
Associate Professor of Radiologic Science

Co-Supervisor of Dissertation
Jianbo Shi
Associate Professor of Computer and Information Science

Graduate Group Chairperson
Val Tannen, Professor of Computer and Information Science

Dissertation Committee
Jean Gallier, Professor of Computer and Information Science
Warren Gefter, Professor of Radiology
Camillo J. Taylor, Associate Professor of Computer and Information Science
Grace Parraga, Associate Professor of Medical Biophysics
ACKNOWLEDGEMENT

I would like to express my deepest appreciation and gratitude to my two advisors, James Gee and Jianbo Shi, for their insightful guidance and patient mentorship, all the way through my graduate studies at the University of Pennsylvania. I am truly fortunate to have had the opportunity to work with both of them.

I would like to thank my committee members, Jean Gallier, Warren Gefter, Grace Parraga, and Camillo J. Taylor, for their invaluable feedback and thought-provoking suggestions to the completion of this dissertation.

I would like to thank all my collaborators for generously sharing their time and ideas. Special thanks are extended to Eduardo Mortani Barbosa Jr., Drew Torigian, and Warren, for providing enormous genuine intellectual advice during our long-time collaboration.

I was lucky to work with a group of wonderful people at PICSL and the CIS Department. I am sincerely grateful to Brian Avants, Nicholas Tustison, and Paul Yushkevich. They have inspired my research since the beginning. I would also like to thank Daniel Adler, Elena Barnardis, Caroline Brun, Phillip Cook, Tessa Sundaram Cook, Sandhitsu Das, Jeffrey Duda, Yang Alex Liu, Hui Sun, Anh Truong, Hongzhi Wang, Yan Wang, Jue John Wu, Yuanjie Zheng, and literally every current and past member of PICSL for their tremendous help and friendship.

I would like to thank my friend, Tingting Sha, for her constant moral support and endless encouragement over the past years. I have owned much to her, which goes beyond words. I would like to thank my amazing parents, Laozhi Song and Chaxiang Wei, for their unconditional love and dedication. I could not have done this without them.

Finally, I would like to thank and dedicate this dissertation to my grandmother, Yuhua Wang. She was the kindest person to me.
ABSTRACT

QUANTITATIVE ANALYSIS OF THORACIC COMPUTED TOMOGRAPHY IMAGES

Gang Song

Jianbo Shi

James Gee

Lung disease has risen to the third leading cause of chronic morbidity and mortality in the United States. The diagnosis, differentiation, and classification of the severity of various lung diseases rely on clinical assessment, thoracic imaging using computed tomography (CT), and pulmonary function testing (PFT). While being the reference standard for assessment of the lung’s mechanical function, PFT strictly permits a global measurement of lung physiology.

In contrast, high-resolution image analysis is a powerful tool with the potential for regional as well as global quantification of diseases. Imaging plays an increasingly important role in lung disease diagnosis. Most current pulmonary imaging techniques are used clinically to assess anatomic changes and to provide qualitative or semi-quantitative estimates of disease severity. Although generally effective, radiologic interpretation of CT images is time-consuming, requiring considerable expertise. It is also largely qualitative and prone to inner-observer diagnostic variability. It is therefore desirable to have automated quantitative analysis from imaging modalities.

This dissertation investigates the problem of using thoracic computed tomography images for automatically and quantitatively analyzing and diagnosing lung diseases. We discuss how to complete a holistic automatic pipeline for clinical studies and study its various components. The final goal is to illustrate how to build a complete pipeline using CT images as input and yielding the desired clinical results as output. In the case when a single lung
image is provided, we propose an algorithm for segmenting small airways from CT images. When the images at both inspiration and expiration phases are available, we discuss the use of image registration algorithms to compute lung kinematics, especially focusing on various diffeomorphic transform models. Finally, these image-derived quantitative metrics are analyzed for clinical studies of specific diseases, including differentiation of interstitial lung diseases and chronic obstructive pulmonary diseases, and quantification of small airway air trapping and emphysema. We show that quantitative CT imaging, integrating segmentation, registration, feature computation, feature selection and pattern recognition, can provide better biomarkers for diagnosis and prognosis.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACKNOWLEDGEMENT</strong></td>
<td>ii</td>
</tr>
<tr>
<td><strong>ABSTRACT</strong></td>
<td>iii</td>
</tr>
<tr>
<td><strong>LIST OF TABLES</strong></td>
<td>x</td>
</tr>
<tr>
<td><strong>LIST OF ILLUSTRATIONS</strong></td>
<td>xviii</td>
</tr>
<tr>
<td><strong>CHAPTER 1 Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 High Resolution Computed Tomography</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Quantitative Information Within HRCT Thoracic Images</td>
<td>3</td>
</tr>
<tr>
<td>1.2.1 Within a Single CT image</td>
<td>4</td>
</tr>
<tr>
<td>1.2.2 Between a Pair of CT Images</td>
<td>5</td>
</tr>
<tr>
<td>1.2.3 Groups of Images</td>
<td>7</td>
</tr>
<tr>
<td>1.2.4 Between Imaging and Other Diagnosis Techniques</td>
<td>7</td>
</tr>
<tr>
<td>1.3 Quantitative Analysis of HRCT Pulmonary Images</td>
<td>7</td>
</tr>
<tr>
<td>1.3.1 A Holistic Pipeline</td>
<td>8</td>
</tr>
<tr>
<td>1.3.2 Automated Image Feature Computation</td>
<td>8</td>
</tr>
<tr>
<td>1.3.3 Utilize Image Features for Disease Analysis</td>
<td>10</td>
</tr>
<tr>
<td>1.4 Overview of the Dissertation</td>
<td>11</td>
</tr>
<tr>
<td><strong>CHAPTER 2 Segmentation of Airway by Removing Paths of Leakage</strong></td>
<td>13</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>13</td>
</tr>
<tr>
<td>2.2 Methods</td>
<td>14</td>
</tr>
<tr>
<td>2.2.1 Initial Segmentation Using Speed Function on Adjusted Gradients</td>
<td>15</td>
</tr>
<tr>
<td>2.2.2 Path Candidate Generation from Initial Segmentation</td>
<td>17</td>
</tr>
<tr>
<td>2.2.3 Leakage Removal Using Cost Function on Path Nodes</td>
<td>18</td>
</tr>
</tbody>
</table>
CHAPTER 3 Review of Diffeomorphic Deformation Models in Image Registration

3.1 Introduction

3.2 From Small Deformation to Large Deformation in Image Registration

3.2.1 A First Example: Horn-Schunck Optical Flow

3.2.2 Demons algorithm

3.2.3 Compositive Demons and Inverse Transform

3.2.4 Diffeomorphism Group

3.2.5 Connection Between Small deformation and Diffeomorphism

3.3 Smooth Vector Fields and Sobolev Gradient

3.3.1 Examples of Regularization Terms and Their optimization Operators

3.3.2 Sobolev Norm and Sobolev Space

3.3.3 Registration in Sobolev Space

3.4 Diffeomorphism with Time-varying Velocity Fields

3.4.1 Beg’s LDDMM algorithm

3.4.2 Numerical Algorithm of LDDMM

3.4.3 Comparison with Horn-Schunck Optical Flow

3.4.4 Comparison with Compositive Demons

3.5 Diffeomorphism with Stationary Velocity Fields

3.5.1 Computation of Exponential Mapping

3.5.2 Ashburner’s DARTEL algorithm

3.5.3 Vercauteren’s Diffeomorphic Demons Algorithm

3.5.4 Vercauteren’s LogDemons Algorithm

3.5.5 Comparison with DARTEL and Composite Demons
## Chapter 6.2.1

### PFT Parameters and Image Metrics

6.2.1 PFT Parameters and Image Metrics ........................................... 93

6.2.2 Feature Selection ........................................................................... 94

6.2.3 Classification with SVMs ................................................................. 96

6.2.4 Experiment setup ........................................................................... 99

6.3 Results .................................................................................................. 99

6.4 Discussion ........................................................................................... 102

### Chapter 7

#### Quantification of Small Airway Air Trapping and Emphysema in Obstructive Pulmonary Disease

7.1 Introduction ........................................................................................ 106

7.2 Materials and Methods ........................................................................ 108

7.2.1 Patient selection ............................................................................. 110

7.2.2 Image acquisition ........................................................................... 110

7.2.3 Image analysis ................................................................................ 110

7.2.4 Statistical analysis .......................................................................... 112

7.3 Results ................................................................................................ 112

7.4 Discussion ........................................................................................... 122

7.5 Conclusion ........................................................................................... 126

### Chapter 8

#### Conclusion

8.1 Conclusion .......................................................................................... 127

### Bibliography

BIBLIOGRAPHY ......................................................................................... 128
LIST OF TABLES

Table 3.1  Comparison of the parameterization and the regularization terms. For
the operator $L$, the regularization term is $||L \cdot ||^2$. Tikhonov regularization

corresponds to Gaussian smoothing. ........................................... 59

Table 5.1  Results using greedy SyN for each scan pair, per category and overall. Rankings and final placement are from a total of 34 competing
algorithms. The total running time for each pair is shown on the last


column. ............................................................................. 83

Table 5.2  Results using exponential mapping for each scan pair, per category and overall. Rankings and final placement are from a total of 34 competing
algorithms. The total running time for each pair is shown on the last


column. ............................................................................. 84

Table 6.1  Statistics of ILD patients. .................................................. 91

Table 6.2  Statistics of COPD patients. ................................................ 92

Table 6.3  All 21 PFT values (1st and 3rd rows) and corresponding relevance
values (2nd and 4th rows) to the disease types. ................................. 93

Table 6.4  Computed image metrics and their relevance to the disease types. $G_1$
denotes the metrics computed from the expiratory images; $G_2$ from
the inspiratory images. Metrics of $G_3$ are generated by subtracting $G_1$
from $G_2$. Within each of $G_1/G_2/G_3$, the metrics are indexed from 1
to 31. The numbers in last three columns are the relevance values to
the disease types of ILD and COPD. ............................................. 95
Table 6.5 The first 20 selected features from mRMR on both PFT values and image metrics. Note these features include both PFT values and image metrics. The image metrics includes both the first order statistical measurements obtainable from the attenuation histogram and other more sophisticated metrics of texture descriptions. $G1$: from the expiratory volumes, $G2$: from the inspiratory volumes, $G3$: subtracting $G1$ from $G2$ (same as in Table 6.4).

Table 7.1 Calculated volumes (in ml) and sample standard deviation for all patients (n=17), subgroup of negligible emphysema (n=7), and subgroup with moderate/severe emphysema (n=10).

Table 7.2 Pearson linear correlation coefficients and respective p-values between key PFT parameters and calculated emphysema and small airway air trapping volumes.
| Figure 1.1 | A representative thoracic HRCT image volume, displayed in 2D slices of (a) axial, (b) sagittal, and (c) coronal planes. The resolution is $1 \times 1 \times 1\text{mm}^3$. |
| Figure 1.2 | A visual comparison of CT images of (a) interstitial lung disease and (b) chronic obstructive pulmonary disease. Note the visually different texture patterns for the two diseases. |
| Figure 1.3 | An example of airway segmentation using CT image. (a) The input CT image. (b) The segmented airway tree using our method in Chapter 3. |
| Figure 1.4 | Deformation between inspiration and expiration. The deformation field (b) is computed from the CT image at inspiration (a) and the CT image at expiration (c). A 2D slice of the 3D deformation field is shown as black-white warped grids. |
| Figure 1.5 | An example of using run length feature to separate ILD from COPD on a dataset of 25 patients. The Y axis is the computed value of run length texture feature. The blue stars on the left of the X axis are from ILD group and the red stars on the right are from COPD group. Using this feature computed from CT images, ILD and COPD can roughly separated. |
Figure 2.1  Example of the proposed adjusted image gradients.  (a) A region containing a thin airway (b) Gradients computed as $|\nabla I|$. Note that it has high gradients inside the airway, which prevent fast marching.  (c) Airway segmentation using (b) to compute speed image and MSFM.  (d) Adjusted gradients computed using $g(x)$. Only the gradients in the bright airway wall are preserved.  (e) Airway segmentation using adjusted gradients in (d) to compute speed image and MSFM.

Figure 2.2  Path graph construction. Only the airway tree in the right lung is shown.  (a) Initial segmentation $S$ using fast marching with adjusted gradients.  (b) Distance transform $D$ inside $S$ with color bar shown on the right.  (c) Graph $G$ using discretization on $D$. Nodes $\{n\}$ are colored by red, green and blue. Each node $n$ has the same discretization value in $D_0$ and belong to the same connected component.  (d) Final segmentation of airway tree using our path segment removal approach.

Figure 2.3  Example of the proposed path leakage cost function on a node of leakage.  Left: shape of the node is similar to airway, but is not part of airway.  First and second rows in the middle: on left) one 2D slice with the path $n$ shown in brown; on right) $k = 8$ sections of the neighborhood of $n$ in different directions, each section $N_i(n)$ shown in different colors.  Last row in the middle: bar plots of $\frac{1}{|N_i(n)|} \sum_{x \in N_i(n)} g(x)$ with each $i$ shown in corresponding colors.  Right: surface rendering of the $k = 8$ sections. Location $a$ has an leakage in the upward direction; however location $b$ is difficult to segment locally. Our proposed cost $S(n)$ is defined as the minimum of the $k = 8$ values (the red bar here), which is significantly lower than other bars.
Figure 2.4  Surface rendering of an example result using the proposed method. (a) Segmented airway using our approach. (b) Comparison with ground truth. The blue regions are labeled in ground truth but missing in our results. The red regions are labeled in our results but missing in ground truth. The brown regions exist in both ground truth and our results. (c) Example of the trachea and bronchi region, which is excluded when computing the rates.

Figure 3.1  A classical example showing the progress of deforming a half C to a full C along a diffeomorphism. The deforming grid accompanies each deformed image.

Figure 3.2  Fig.1 from [7]. Inversion and composition in a small deformation setting. The compositions are not identity transform.

Figure 3.3  Adapted from of Fig.11 from [15]. Each column is a pair of $\phi_1$ (first row) and $v_t$ (second row). Different time points of $v_t$ are superimposed on the same figure.

Figure 4.1  Computed tomography images of two lung lobes moving in different directions. Top row: images before/after applying the transform. Bottom row: the two local affine transforms and the computed polyaffine transform using the proposed approach.
Figure 4.2 Polyaffine transform using the previous approach \[5\]. (a) Two input affine transforms are shown in red and green. The starting and ending locations are indicated by the boxes in solid and dashed lines respectively. (b) The input local regions (shown as red and green areas) for each transform with their velocity plotted inside. (c) Computed velocity field using Eqn. 4.4 interpolated from the input local regions. (d) The resulting transform. The ending locations of each local region are plotted as boxes in solid lines at the arrow heads. The solid arrows (the result) deviate from the dashed arrows (the input), showing that local transforms are not preserved precisely.

Figure 4.3 Polyaffine transform using our proposed method. (a) Two input affine transforms (same as in Fig. 4.2(a)) are shown in red and green. The starting and ending locations are indicated by the boxes in solid and dashed lines respectively. (b) The trajectories of the input local regions (shown as red and green areas) for each transform with their velocity plotted inside. (c) Computed velocity field using Eqn. 4.8 interpolated from the input region trajectories. (d) The resulting transform. The ending locations of each local region are plotted as boxes in solid lines at the arrow heads. The solid arrows (the result) are overlapped with the dashed arrows (the input), showing that local transforms are preserved precisely.

Figure 4.4 The trajectories of two local regions overlap. Each arrow represents an input local transform, plotted in black and red. The boxes at the arrow tails are the starting locations of each input transform; the boxes at the heads are the ending locations.
Figure 4.5  Comparison of the transforms computed using (a,c,e) dist$(x, M^*_i)$, and
(b,d,f) dist$(x, \partial M^*_i)$. (a,b) The transforms are constructed from the
same input, one horizontal translation (green arrow) and one vertical
translation (red arrow). (b) is visually more smooth than (a). (c,d)
The weights $w(x)$ for the local vertical translation, range in $[0, 1]$. (d)
is more polarized except along boundary $\partial M^*_i$. (e,f) The resulting
velocity below the green trajectory has an upward vertical component
in (e) but not in (f). Similar effects are seen on the right of the red
trajectory. $M^*_i$ are shown as green and red dotted boundaries in (c,d)
and solid areas in (e,f).

Figure 4.6  Comparison of our proposed method and the baseline method [5],
using (a) two local translation transforms and (b) two local rotation
transforms. The x-axis corresponds to (a) the translation offset, or
(b) the rotation angle. The y-axis corresponds to the average error
of the resulting transforms in the local affine regions, with standard
deviations plotted as error bars.

Figure 5.1  The four categories evaluated in the EMPIRE10 challenge: the lung
boundaries, fissures, labeled landmarks and singularities in the defor-
mation field, from [72].

Figure 5.2  Examples of good registration results, from two subjects: (a) the fix
subject, (b) the moving subject, (c) the warped moving subject.

Figure 5.3  Example of bad registration shown in enlarged image region. The
vessels and the fissure in the upper region do not align well. The
remaining lower region, however, has a good registration.

Figure 6.1  Pipeline of the proposed feature selection analysis.
Figure 6.2  Examples of SVMs on a synthetic 2D data set. Each data point has its $x$ and $y$ coordinates as two features. Type I points are marked with red dots; type II with blue crosses. SVM hyperplanes are shown in green. Note that Type I points cannot be separated from type II points by any straight lines. (a) The linear kernel SVM misclassifies several points, marked with circles. (b) The RBF kernel SVM classifies all the points with a curve as classification hyperplane.

Figure 6.3  Relevance (mutual information) of image metrics to different disease types. The $x$ axis has the image metrics index listed in Table 6.4. Red bars are for metrics from $G_1$, the expiratory images; blue bars for $G_2$, the inspiratory images; and green bars for $G_3$, difference of $G_1$ to $G_2$.

Figure 6.4  Relevance (mutual information) of PFT values to different disease types. The $x$ axis lists all the 21 PFT values in Table 6.3. Note that the range of the $y$ axis is similar to the range from those image metrics in Fig. 6.3.

Figure 6.5  Classifying ILD and COPD using the first $n = 1$ to 20 selected features with (a) linear SVM and (b) RBF SVM, using MaxRel or mRMR on PFT values, image metrics or both PFT and image metrics. For each curve, the $x$ axis is the number of selected features for training and testing with SVMs; the $y$ axis is the average error rate from leave-one-out tests. One subject is excluded when learning SVM parameters and that subject is used to test the accuracy of the trained SVM.

Figure 6.6  Receiver operating characteristic curves for using different values to separate ILD and COPD. The green curve uses FEV1/FVC value. The red curve uses the RBF SVM scores computed from the first 10 features selected by mRMR. SVMs are trained and tested using the leave-one-out strategy.
Figure 7.1 Illustration of the effectiveness of the registration algorithm. (a) Representative coronal thoracic HRCT image. (b) Superimposed 3D lung and central airways segmentation. (c) Coronal subtraction image (inspiration-expiration) prior to registration. Vessel and airway misalignment is noted. (d) The same subtraction after registration. Vascular and airway alignment are substantially improved, indicating the correctness of the registration mapping. Brighter areas indicate greater difference in attenuation values.

Figure 7.2 Representative thoracic HRCT images in (a) axial, (b) sagittal, and (c) coronal planes in subgroup 1 patient with obstructive lung disease (negligible emphysema). Segmented small airway air trapping shown in red and segmented emphysema is shown in green in (d) axial, (e) sagittal, and (f) coronal planes.

Figure 7.3 Representative thoracic HRCT images in (a) axial, (b) sagittal, and (c) coronal planes in subgroup 1 patient with obstructive lung disease (moderate/severe emphysema). Segmented small airway air trapping shown in red and segmented emphysema is shown in green in (d) axial, (e) sagittal, and (f) coronal planes.

Figure 7.4 Computed volumes for each change threshold for the negligible emphysema subgroup (n = 7 patients). Error bars denote 1 standard deviation. ATV, small-airway air-trapping volume; EV, emphysema volume; HU, Hounsfield units; TSLVe, total segmented lung volume in expiration; TSLVi, total segmented lung volume in inspiration.
Figure 7.5  Computed volumes for each change threshold for the moderate or severe emphysema subgroup (n = 10 patients). Error bars denote 1 standard deviation. ATV, small-airway air-trapping volume; EV, emphysema volume; HU, Hounsfield units; TSLVe, total segmented lung volume in expiration; TSLVi, total segmented lung volume in inspiration.

Figure 7.6  Pearsons correlations between forced expiratory volume in 1 second (FEV1)/ forced vital capacity (FVC) and emphysema volume (EV), small airway air trapping (AT) volume, and total AT volume at each specific change threshold (5300 Hounsfield units [HU], in 25-HU increments) for subgroup 1 (n = 7 patients).

Figure 7.7  Pearsons correlations between residual volume (RV) and emphysema volume (EV), small airway air trapping (AT) volume, and total AT volume at each specific change threshold (5300 Hounsfield units [HU], in 25-HU increments) for subgroup 1 (n = 7 patients).

Figure 7.8  Pearsons correlations between residual volume (RV) and emphysema volume (EV), small airway air trapping (AT) volume, and total AT volume at each specific change threshold (5300 Hounsfield units [HU], in 25-HU increments) for subgroup 2 (n = 10 patients).

Figure 7.9  Pearsons correlations between forced expiratory volume in 1 second (FEV1)/ forced vital capacity (FVC) and emphysema volume (EV), small airway air trapping (AT) volume, and total AT volume at each specific change threshold (5300 Hounsfield units [HU], in 25-HU increments) for subgroup 2 (n = 10 patients).
CHAPTER 1

Introduction

The lung is the essential respiration organ for human bodies. Its principal function is to transport oxygen into the bloodstream and release carbon dioxide into the atmosphere. In general, the lung can be described as an elastic organ and an elaborate network of connecting airways, vasculature and pulmonary interstitium. Normal lung tissue has a homogeneous appearance delineated by lobular boundaries, airways and vascular trees. Numerous pulmonary diseases affect normal lung mechanics by disrupting the material composition of the lung, impairing the integrity of the lung’s ability to oxygenate the blood effectively.

Lung disease has risen from to one of the major leading causes of chronic morbidity and mortality in the United States in recent years [117]. According to the American Lung Association, one in six deaths in the United States can be attributed to lung decease, and more than 35 million Americans live with chronic lung disease [2]. Identifying and diagnosing different lung diseases is challenging. The classical ways include pulmonary function tests (PFT), lung tissue biopsy analysis and imaging tests.

Pulmonary function tests enable clinicians to evaluate an individual’s respiratory status [40]. It is a noninvasive method of assessing the integrated mechanical function of the lung, chest wall, and respiratory muscles. For example, using spirometry, one can measure how
much air the lung can hold and how quickly the air can move out when a patient exhales. It currently comprises the reference standard for pulmonary functional assessments.

Lung tissue biopsy analysis is useful when diagnosing interstitial lung disease. A small amount of lung tissue sample is obtained from a patient's lungs for a detailed laboratory analysis. The amount of tissue sample varies from the head of pin to a larger one using an invasive surgical biopsy.

Imaging test is another clinical diagnosis method. There are various types of imaging approaches, which include chest X-ray, computed tomography (CT) and gas diffusion imaging. The objective of this dissertation is to study how to use lung CT images for a quantitative analysis of various pulmonary diseases.

1.1 High Resolution Computed Tomography

CT imaging provides a noninvasive way to study pulmonary morphology and lung anatomy. Tissues of different radiation attenuation show distinct values in their CT images. The values are measured using Hounsfield Units (HU), which ranges from -1024 to +3071. Air, for example, does not absorb x-rays and has the least radiation attenuation with the value as -1000HU. Bones, containing calcium, are normally above 200HU. Fatty tissues and blood are in the range from -100HU to 200HU. Today’s high-resolution CT (HRCT) is able to achieve near-isotropic 3D resolution on the order of 0.5 mm and to scan the whole lung in 10-25 seconds in expiration or inspiration breath-hold (Fig. 1.1).

Thoracic HRCT of lung imaging is particularly useful for examination of lung parenchyma in high resolution. It is used to assess a variety of lung pathologies. HRCT image analysis is a powerful tool with the potential for regional as well as global quantification of pulmonary diseases.

Imaging has its unique advantages compared with the other diagnosis methods. PFT is non-
invasive, but it can only provide an overall measurement of pulmonary mechanics. Tissue biopsy analysis provides accurate information but only around the sampled regions.

On the contrary, imaging provides both a non-invasive and accurate way to analyze the lung regions. First, CT images provide information on all locations of the lung, not just the sampled regions in tissue biopsy. Important anatomies like airways and vascular trees are visible in 3D HRCT images. Additionally, measurements from PFTs are not typically sensitive during early disease development, because affecting a portion of the lung’s functional ability may not become evident if the rest of the lung is compensating for the dysfunction. However, imaging can provide direct visualization for clinicians and radiologists to make a diagnosis on early disease development.

1.2 Quantitative Information Within HRCT Thoracic Images

There is a wealth of information within HRCT thoracic images for lung disease analysis and diagnosis. This section discusses the information that is available to extract from different image sources.
1.2.1 Within a Single CT image

To diagnose lung diseases, patients usually undertake scanning at a particular phase of breathing, for example, at the end of inspiration, or at the end of expiration. There are various types of information that one can interpret from a single thoracic CT image, e.g. intensity values, texture information, etc.

The intensity value of each individual voxel is useful in CT images. Different tissues have different voxel intensities using CT modalities. This fact is useful when differentiating normal tissue from pathological one. For example, the region with emphysema usually contains darker voxels. Thus, image thresholding is useful to identify emphysema \[79\]. By computing the volume of voxels of intensities below a certain threshold, one can interpret the severity level of the emphysema.

Texture information also provides distinct visual clues to diagnose lung disease. Although each individual alveolar is unrecognizable using current CT imaging resolution, the overall tissue exhibits different visual texture patterns for different lung diseases. For example, interstitial lung disease (ILD) has certain characteristic texture patterns, when compared with the CT image of a patient with chronic obstructive pulmonary disease (COPD). Fig. 1.2 shows one image from ILD and one from COPD, where ILD demonstrates a fibrosis texture pattern, which does not exist in the image of the COPD patient. Such texture features are crucial to determine the diagnosis and to assess severity of the disease.

Furthermore, important anatomies, like the airway trees, the vascular trees and individual lobes, are visible from the CT image. An example of the airway tree is shown in Fig. 1.3. Segmenting the anatomy from CT images helps to identify diseases specifically related to certain regions, like airway wall attenuation for airway diseases \[114\]. Using automatic image segmentation techniques, it is able to have region-specific study by segmenting the anatomy of interest from the CT image. In other cases, certain anatomies may also be desired to be excluded for statistical study for a more accurate analysis.
Figure 1.2: A visual comparison of CT images of (a) interstitial lung disease and (b) chronic obstructive pulmonary disease. Note the visually different texture patterns for the two diseases.

1.2.2 Between a Pair of CT Images

The lung is an elastic body, and the respiratory motion is a dynamic procedure. When multiple CT images scanned at different phases in a breathing cycle are available, the pulmonary motion can be determined by studying the deformation between images taken at different phases. A common practice is to take one CT scanning when the patient holds the breath at the end of inspiration and take the other at the end of expiration. One can analyze the pulmonary mechanism by computing the deformation field between these two images. Fig. 1.4 shows an example of the deformation field, computed from the image at inspiration to the image at expiration.

Once the deformation field is available, certain physical mechanism properties, like Jacobian or strains, can be further computed to analyze the pulmonary motion. When more than two phases are scanned, a longitudinal 4D analysis is also possible.
Figure 1.3: An example of airway segmentation using CT image. (a) The input CT image. (b) The segmented airway tree using our method in Chapter 2.

Figure 1.4: Deformation between inspiration and expiration. The deformation field (b) is computed from the CT image at inspiration (a) and the CT image at expiration (c). A 2D slice of the 3D deformation field is shown as black-white warped grids.
1.2.3 Groups of Images

When studying one type of disease, we normally need statistics from a set of images taken from a group of patients with the same diseases. Another case is longitudinal studies, where we collect the images of the same patient over multiple time points. These cases lead to groupwise statistical analysis.

1.2.4 Between Imaging and Other Diagnosis Techniques

Information within images is valuable and it becomes even more valuable when integrating together with other clinical diagnoses. For example, studying correlations between image features and the numbers measured from pulmonary function tests are crucial to research different diseases. Also, one can investigate the redundancy between imaging diagnosis and other diagnosis tests by studying whether the information computed from images is statistically necessary when other diagnosis methods are also available.

1.3 Quantitative Analysis of HRCT Pulmonary Images

While much information exists within thoracic CT images, it requires tremendous effort for radiologists to observe images, extract information and make diagnosis. Sometimes certain analysis is very time-consuming or even impossible for humans to complete. For example, it can take hours or days to segment anatomy structures from a 3D image. This makes it infeasible to analyze a large-scale dataset. Also, certain tasks are difficult for human vision, like creating a voxel-wise pulmonary deformation field by aligning two 3D images manually. Furthermore, with the exponential increase of available imaging data, the required effort and time will increase explosively.

Therefore, it is desirable to have an automatic computational approach to utilize the information from HRCT thoracic images for quantitative analysis and diagnosis. The advantages using an automatic quantitative approach include:
• Reduced human work and errors. This is essential for a reproducible research.

• Quantitative evaluation over qualitative interpretation. For example, rather than giving an assessment as mild or severe, a continuous number is preferred, which is more suitable for progressive evaluation.

• Retrospective study on available imaging data. It would be economical to utilize existing images when new computational methods become available, without the need of recruiting new patients and taking similar images.

1.3.1 A Holistic Pipeline

In some aspect, the automatic quantitative approach can be viewed as translating the visual interpretation and reasoning of radiologists to a computational pipeline, or even enhancing human interpretation by fusing other information in a systematic framework. This is a challenging task and the main topic of this dissertation.

Quantitative analysis of CT images is not a problem of an isolated module. Instead, it is more like an integrated pipeline linking different modules using a holistic approach to integrate various information. In this dissertation, we separate it into two sub-problems: how to compute various information from images, and how to utilize such information for disease analysis.

1.3.2 Automated Image Feature Computation

The first sub-problem we are going to address in this dissertation is how to compute various features from thoracic CT images automatically. As previously discussed, there are multiple types of information which are available from images. Each kind requires a different computation method to be extracted. Some of them, like voxel-wise intensity, can be straightforward as a direct query of CT image arrays. Then the intensity information can be used together with image thresholding operation, or histogram analysis.
Other types of information are more challenging to extract. In these cases, more elaborated algorithms are needed. For example, in the example of differentiating ILD from COPD using HRCT images (see Fig. 1.2), one possible way to define the texture is to use run length feature ([118]), based on computing the voxels of the same intensity level along different directions. Fig. 1.5 shows that ILD and COPD have separate distinct values of this feature. Thus, this is an automatic way to replace the visual interpretation from radiologists to diagnose ILD from COPD.

![Figure 1.5: An example of using run length feature to separate ILD from COPD on a dataset of 25 patients. The Y axis is the computed value of run length texture feature. The blue stars on the left of the X axis are from ILD group and the red stars on the right are from COPD group. Using this feature computed from CT images, ILD and COPD can roughly separated.](image)

In the following chapters, we are also going to discuss other types of image features. For segmenting anatomical structures from a single image, this dissertation will focus on how to segment the airways in Chapter 2. For computing the deformation field from a pair of images, we will discuss how to use image registration methods to recover an invertible and smooth pulmonary deformation in Chapter 3, 4 and 5.
1.3.3 Utilize Image Features for Disease Analysis

The second sub-problem in this dissertation is how to utilize the features computed from CT images in the application of disease analysis. Some image features have clinical or physical interpretation, which can be easily translated using radiologists’ knowledge. For example, using automatic segmentation of the whole lung region, the volume of segmentation is highly correlated with the Total Lung Capacity (TLC) measured in the pulmonary function test, as is discussed in Chapter 7. Also, after aligning a pair of images at inspiration and expiration phases, the intensity change for each voxel corresponds to the local density change around the voxel during respiration. These features can be used directly for correlation study or further analysis.

In other cases, it is not clear whether the computed features are useful for further analysis. For example, there exist many possible definitions to compute image texture features. It is desirable to find out which definitions can describe some specific type of disease. In this case, feature selection can be used in the pipeline to select the most relevant features for accuracy, or to remove redundant features for simplicity. Furthermore, feature selection can be applied not only to the image features, but also to other measurements (like pulmonary function tests) together with image features, to improve diagnosis. We are going to discuss in detail about how to utilize feature selection to find optimal texture features in Chapter 6.

Besides feature selection, a common practice is to classify different diseases using image features. This is similar to the case when radiologists make a diagnosis decision based on the information they observe. We are going to discuss using machine learning methods like support vector machines as pattern classifiers to separate diseases in Chapter 6.
1.4 Overview of the Dissertation

The topic of this dissertation is to study how to use lung CT images for quantitative analysis of various pulmonary diseases. The goal is to illustrate how to build a complete pipeline using CT images as input and the desired clinical results as the output. In the remaining chapters, we are going to discuss in detail how to compute different types of features from images and how to utilize these features for disease analysis.

In Chapter 2, we discuss how to segment airway trees from one single thoracic CT image. We propose a new approach specifically aimed for segmenting thin airway trees.

Different deformation models used in registering two images are reviewed in Chapter 3. We unify several image registration algorithms in the same framework and compare different ways to parameterize the deformation model. We specifically discuss the diffeomorphic deformation models, which are suitable for the physical kinematics of lung respiration.

A new deformation model is proposed in Chapter 4 to balance the parameter complexity and the model flexibility. Such a new model is suitable as an initial representation of lung movement for further image registration.

In Chapter 5, we discuss how to adapt the diffeomorphic deformation model in pairwise lung CT image registration. We demonstrate the accuracy of our registration algorithm by evaluating them on a public benchmark.

We provide a systematic pipeline in Chapter 6 to identify an optimal set of texture-based image features in classifying interstitial lung disease and chronic obstructive pulmonary disease. We demonstrate that image features provide extra information to the measurements of pulmonary function test.

In Chapter 7, we discuss how to build a pipeline to quantify of small airway trapping and emphysema using image registration and airway segmentation. We show that small airway
trapping can be quantitatively evaluated and automatically located from the CT images.

We conclude with summary and highlight the contributions of this dissertation in Chapter 8.
CHAPTER 2

Segmentation of Airway by
Removing Paths of Leakage

2.1 Introduction

Segmentation of airway trees from computed tomography images is critical for various clinical applications involving pulmonary diseases. Diameters of the fourth generation airway in a typical CT image are about two or three voxels wide. The limitation of imaging resolution and noise lead to the inhomogeneity of image intensities inside airway walls and also the blurring effects around airway walls. For example, low-dose CT imaging is sometimes preferred in order to reduce radiation exposure (105). These factors make the balance between detecting over-segmentation and extending thin airway very critical in airway segmentation.

Various algorithms have been proposed in the literature. Schlathölter et al used level set methods for a simultaneous segmentation and tree reconstruction framework (86). The authors proposed several heuristic rules to detect leakage in the growing regions. Tschirren et al proposed to keep an active region of cylinder shape (105). By tracking the orientation of the active cylinder, the active region was extended to next possible airway location. A
multi-threshold approach was adopted in [106] to increase robustness in growing airway
trees. Recent work from Christian et al [14] used gradient vector flow to guide growing
direction. The work of [63] focused on extending thin airway by computing the shortest
paths inside a search sphere from end points in the initial segmentation.

Many of these algorithms have shown successful segmentation of the bronchi and trachea. 
However, for the third and higher generations in the airway tree, current segmentation
results still have room for improvement, indicated by a recent evaluation on 15 airway
segmentation algorithms [62].

In this chapter we explore a new approach for airway segmentation. Instead of mixing
airway segmentation and leakage detection at every iteration as in [14, 63], we divide this
problem into a hypothesis generation of thin airway paths and a post processing procedure
of removing leakage path candidates. For the purpose of generating as many hypotheses as
possible, we propose a novel speed function for thin airway. To exclude leakage regions, we
propose a novel cost function defined on the whole path candidate. Such a scheme is more
flexible when evaluating the whole path and can be viewed as complementary to current
region growing methods.

2.2 Methods

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

We use fast marching as the first step to generate the initial airway segmentation. The arrival time is modeled by the Eikonal equation

\[
|\nabla T| F = 1,
\]

(2.1)

where \( F \) is the speed function of the propagation front and \( T \) is the arrival time. Given the 3D image volume \( I \), one common form of the gradient based speed function is

\[
F = e^{-\beta \|
abla I\|}.
\]

(2.2)

Here \( \beta \) is a scalar weighting coefficient.

By definition \( \nabla I \) usually has high values at the locations close to the boundary of airway, both inside and outside the airway. Thus fast marching will have a low speed along the boundary inside airway. For the trachea and bronchi, this is not a practical problem since the airway boundary takes a small portion in the whole volume. However, for thin airway (like fourth order airway), where the diameter of the airway is about two voxels wide, the speed \( F \) would be low along whole thin airway segment and this would prevent fast marching from successfully extending further (See Fig. 2.1b).

Using \( \|
abla I\| \) directly is not suitable for front propagation along thin airway with partial volume effect. One way to deal with this problem is to interpolate image to a higher resolution, or to assign a special label for airway boundaries in fast marching. Here we propose a different solution by adjusting the definition of image gradients. The ideal gradient for airway segmentation should be defined only on the wall of airway (brighter voxels) rather than inside airway (darker voxels), such that the speed inside airway are high while low
Figure 2.1: Example of the proposed adjusted image gradients. (a) A region containing a thin airway (b) Gradients computed as $|\nabla I|$. Note that it has high gradients inside the airway, which prevent fast marching. (c) Airway segmentation using (b) to compute speed image and MSFM. (d) Adjusted gradients computed using $g(x)$. Only the gradients in the bright airway wall are preserved. (e) Airway segmentation using adjusted gradients in (d) to compute speed image and MSFM.

outside. To formalize this notion, the following modified gradient is proposed:

$$
g(x) = \begin{cases} 
|\nabla I(x)| & \text{if } |I(x + \frac{1}{2}\vec{n}_x) - I(x)| < |I(x) - I(x - \frac{1}{2}\vec{n}_x)|; \\
0 & \text{otherwise.}
\end{cases} \tag{2.3}
$$

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

With this gradient definition, the corresponding front propagation equation becomes $F = e^{-\beta g(x)}$. A seed point $x_0$ in the trachea region is selected as the initial condition $T(x_0) = 0$. A comparison of results using $g(x)$ and $\nabla I(x)$ is shown in Fig.2.1. The fast marching gives the correct segmentation using the proposed $g(x)$ on this thin airway segment.

Another issue in solving front propagation for thin airway is the choice of connectivities. Many airway segments are extended along diagonal directions. However, Sethian’s original fast marching method (FMM, [90]) is not accurate along these directions as it computes the derivative using 6-connectivity. In order to adopt the 26-connectivity in fast marching,
we applied the multi-stencil fast marching method (MSFM, [46]) to solve the equation in practice. Compared to FMM, MSFM solves the equation along several predefined stencils to cover the entire 26 neighboring locations. For those stencils that are not aligned with the natural coordinate system, the equation is derived using directional derivatives. Details of MSFM can be found in [46]. Note that a 26-connectivity, which contains connectivity along diagonal directions in 3D, is especially important for fast marching in thin airway when one location has no 6-connectivity neighbors in the airway.

2.2.2 Path Candidate Generation from Initial Segmentation

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

Figure 2.2: Path graph construction. Only the airway tree in the right lung is shown. (a) Initial segmentation $S$ using fast marching with adjusted gradients. (b) Distance transform $D$ inside $S$ with color bar shown on the right. (c) Graph $G$ using discretization on $D$. Nodes $\{n\}$ are colored by red, green and blue. Each node $n$ has the same discretization value in $D_0$ and belong to the same connected component. (d) Final segmentation of airway tree using our path segment removal approach.
More formally, a graph $G$ is built to describe the structure of the binary mask $S$. Each node in $G$ is a cluster of voxels in $S$, corresponding to part of the airway or to part of the leakage. We apply an approach similar to [86] and [47] to get the graph $G$.

First a distance field, $D(x)$, to the initial seed point $x_0$ is computed in the domain $S$. Then each continuous value $D(x)$ is discretized into an integer $j$ such that

$$D_0(x) = j, \text{ if } jh \leq D(x) < (j + 1)h,$$

in which $h$ is the bin width for discretization. By assigning a node $n$ to one connected component of the same integer value in $D_0(x)$, a graph $G$ is constructed from $S$. Neighboring connected components are connected in the graph. The bin width $h$ controls the shape of each node in the airway so that a node is roughly a tube-like structure. An example of the graph construction using discretization is illustrated in Fig.2.2.

The propagation from initial seed point $x_0$ to a voxel $x_1$ in $S$, denoted as $C^{x_1}_{x_0}$, can be traced by solving ordinary differential equation

$$\frac{dC(t)}{dt} = -\nabla \frac{D}{|\nabla D|}, \text{ with } C(t_1) = x_1 \text{ and } C(0) = x_0. \quad (2.5)$$

We define the path from $x_0$ to $x_1$ on the graph $G$ as the series of nodes that intersect with $C(t)$:

$$P_{x_0}^{x_1} = \{n | n \in G, n \cap C_{x_0}^{x_1} \neq \emptyset \} \quad (2.6)$$

### 2.2.3 Leakage Removal Using Cost Function on Path Nodes

Each path segment node $n$ in $G$ is a candidate for leakage removal. We consider three properties of the path segment: its volume, its vesselness measurement and how it is separated from the background in the image. The first two properties have been investigated in the literature [86, 14, 63]. Here we use Frangi’s vesselness definition [36] on dark tubes.
to measure each voxel in $n$. Those voxels of vesselness smaller than $t_{vessel}$ are removed from node $n$. The nodes of a volume greater than a threshold $t_{vol}$ are also removed. These two properties make sure that path segment $n$ is of a tube-like shape, however, it may still contain leakage areas (see Fig. 2.3 for an example).

We propose a novel leakage cost function to estimate how each path node can be separated from background. For each node $n$, its medial axis is computed as the segment direction. The local coordinate system use this direction as $z$ axis of the segment. The surrounding region of $n$ (obtained from image dilation operation) is then divided into $k$ sections by different angles uniformly distributed in the $x$-$y$ plane. Fig. 2.3 shows an example of $k = 8$ sections in different colors. If the node has leaked into the background from some direction, the average gradient in that section would be low.

![Figure 2.3: Example of the proposed path leakage cost function on a node of leakage. Left: shape of the node is similar to airway, but is not part of airway. First and second rows in the middle: on left) one 2D slice with the path $n$ shown in brown; on right) $k = 8$ sections of the neighborhood of $n$ in different directions, each section $N_i(n)$ shown in different colors. Last row in the middle: bar plots of $\frac{1}{|N_i(n)|} \sum_{x \in N_i(n)} g(x)$ with each $i$ shown in corresponding colors. Right: surface rendering of the $k = 8$ sections. Location $a$ has an leakage in the upward direction; however location $b$ is difficult to segment locally. Our proposed cost $S(n)$ is defined as the minimum of the $k = 8$ values (the red bar here), which is significantly lower than other bars.](image)

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

\[
S(n) = \min_i \frac{1}{|N_i(n)|} \sum_{x \in N_i(n)} g(x)
\]  

(2.7)

where \(g(x)\) is the adjusted gradient in section 2.2.1, \(N_i(n)\) is the neighborhood of \(n\) in the \(i\)-th angle. The nodes of \(S(n)\) lower than a threshold \(t_{\text{path}}\) are considered as leakage and are removed from \(G\). The final airway segmentation is the connected component in \(G\) that contains most remaining nodes.

2.3 Results

The proposed method was applied to extract the airway trees in four canine subjects at 15 months after pulmonary lobe surgery (part of lobes were removed). The images were acquired under forced inflationary pressure of 30 cm H\(_2\)O. All images have a slice thickness of 1.25 mm and in-plane resolution of 0.46\(\times\)0.46 mm\(^2\). The airway trees in these volumes were semi-supervised extracted and manually modified as the ground truth for evaluation.

In all our experiments, the parameter \(\beta\) used in computing speed function was set to 0.05. The initial segmentation threshold \(t\) was 200. The distance field \(D(x)\) was discretized by bin width \(h = 2\). The value of \(h\) is related to the length of the node in practice. We tested \(h\) from 2 to 4 and got similar results. We applied the multi-scale strategy when computing Frangi’s vesselness, using four scales \(s = 0.5, 1, 2, 3\) (see the discussion of scale \(s\) in \[36\]). For each path node \(n\), its surrounding neighborhood was obtained by computing the dilation of radius= 3 and was further divided to \(k = 8\) sections along the segment direction. The vesselness threshold \(t_{\text{vessel}}\) was 30 and the volume threshold \(t_{\text{vol}}\) was 1000. The final segmentation was given by removing all path nodes whose path leakage cost \(S(n)\) was lower than 10.

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

Figure 2.4: Surface rendering of an example result using the proposed method. (a) Segmented airway using our approach. (b) Comparison with ground truth. The blue regions are labeled in ground truth but missing in our results. The red regions are labeled in our results but missing in ground truth. The brown regions exist in both ground truth and our results. (c) Example of the trachea and bronchi region, which is excluded when computing the rates.

All final segmentation results $S$ were evaluated against the manually labeled airway trees $S_g$. Using $|\cdot|$ to denote the number of voxels in a set, we computed the recall rate as $\frac{|S\cap S_g|}{|S|}$, the false alarm rate as $\frac{|S-S_g|}{|S|}$ and the missing rate as $\frac{|S_g-S|}{|S_g|}$. As trachea and bronchi are relatively easy to segment using region growing approaches and our main interest was to extract thin airway, these regions were excluded in computing both $S$ and $S_g$. Another reason is that the trachea and bronchi might take up to 95% volume in the whole airway tree; thus it should be excluded as a huge bias in evaluating thin airway segmentation. An example of our final segmentation is illustrated in Fig.2.4. Missing regions $S_g - S$ are colored in blue and false alarm regions $S - S_g$ in red (in Fig.2.4b).

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

2.4 Discussion

In this chapter we proposed a new method for extracting airway trees from 3D computed tomography images. We focused our method on removing leakage regions while still segmenting most thin airway. Compared to existing region growing methods, our method had a unique path removal procedure to exclude potential leakage. A graph of path candidates was constructed from the arrival time using the fast marching method. The leakage regions were identified from all path segment candidates using our proposed leakage cost function. A similar way of checking different directions was proposed in [115]. They used gradients in different radial directions from one voxel to track vessels. In comparison, our cost function was defined on the node of a whole path segment for leakage removal. Each node was a higher level structure, which was more robust to noise and had a more flexible definition. Furthermore, this was not a linear measurement, which means it could not be represented as an integration along the path. Thus we did not use Dijkstra’s algorithm, which was used in [63].

The second contribution in this chapter was that we proposed a form of adjusted gradient in computing speed image. We also applied multi-stencil fast marching for a 26 connectivity neighborhood in narrow airway segmentation. We showed that thin airway regions, as narrow as 1 to 3 voxels wide, could be extracted by our adjusted image gradients and utilizing the connectivities along diagonal directions in fast marching.

It should be noted that our approach is not contradictory to existing region growing methods. By separating the leakage detection as a post processing step, our method can take the advantage of current work on region growing while reducing leakage. The result from
one approach can be used as input to the other. For example, the missing rate in our experiment may be further improved using the method in [63] in the stage of generating path candidates.

The number of datasets used in the evaluation was limited. A larger evaluation data set is needed in our future work. Also, we used manual segmentation as our ground truth. An alternative for constructing the reference is combining results from different algorithms as in [62]. It would also be interesting to test our algorithm on the output of other algorithms to verify the effectiveness of post-processing.
CHAPTER 3

Review of Diffeomorphic Deformation Models in Image Registration

3.1 Introduction

Image registration is an important medical image analysis technique for computational research. It generates a transform or a mapping between two given images. Such a mapping is essential for many medical quantitative research problems. One can construct an anatomy template by registering a group of images. Various statistics can be defined if images are registered to the same spatial space. Also, the mapping is useful in building the motion model for pulmonary and cardiac images. It also helps to discover the longitudinal changes and quantify the effects of surgical treatments.

Image registration is normally formulated as an optimization problem. An image similarity function defines how two images are called “similar” after registration. However, this problem is ill-posed. There are infinite ways to warp one image into the other. We normally
assume the optimal transform should be smooth, which is controlled by the regularization term in the energy function.

Different parameterization of transforms leads to different regularizations. Historically, transforms with small deformations were first studied as an optical flow problem [53]. Thirion proposed his demons algorithm [101], which was a heuristic approach inspired by the diffusion process. Due to its practical success, there were theoretical efforts [19] to justify the demons algorithm later. Small deformations parameterize the transform using a displacement field, which is analogous to the elasticity of a spring. The performance of these models usually degrades rapidly for images with large deformation. One important reason is that the small deformation model does not preserve the topology in the image. In other words, there is no guarantee regarding the one-to-one mapping property for the existence of the inverse transform.

Large deformation methods aim to solve this problem by adopting the notion of diffeomorphism in image registration. An example was given in Fig. 3.1. A diffeomorphism is a differentiable map with a differentiable inverse, which forms a group structure. It has a solid mathematical foundation. Unlike small deformation methods, large deformation methods solve a ”deforming” path of one image transforming into the other. The transform model was no longer defined on the end point of the transform, but on the entire spatial-temporal path. This is analogous to tracking motion of the fluid.

Diffeomorphism has been shown as an effective model in computing large deformation image registration. Trouvé gave some theoretical analysis about the diffeomorphism group in [104]. Beg et al derived a rigorous model [15] using the velocity fields as parameterization. To simplify the computation, Ashburner [7] proposed to use a stationary velocity field to approximate the time-varying velocity field. Especially he introduced a fast numerical method to compute the derivative in the optimization step. Building on Arsigny’s Log-Euclidean framework [6], Vercauteren et al adopted the similar idea and extended the classical demons into the diffeomorphism transform [111], which was efficient in computation. Later, this
approach was formulated more rigorously [110], which worked completely in the log-domain. Recent improvement by Mansi et al [69] extended the deformation with incompressibility constraints. Also, interesting work [22] from Chen et al independently proposed another more flexible way to control both the divergence and the curl of the deformation field.

Figure 3.1: A classical example showing the progress of deforming a half C to a full C along a diffeomorphism. The deforming grid accompanies each deformed image.

One way to understand these diffeomorphic approaches is to study the different forms of parameterization used in the diffeomorphism model. In this chapter, we focus on four popular diffeomorphic models: LDDMM from Beg et al [15], DARTEL from Ashburner [7], diffeomorphic demons [111] and LogDemons [110] from Vercauteren et al.

Our discussion proceeds in two steps. We first unify the general image registration in the same framework, using the small deformation as examples. We discuss various regularization approaches applied in these methods and show how they can be unified using the Sobolev gradient. We explain how the gradient of the image similarity term is mapped to the transformation space through the regularization term.

After exploiting the shared framework, we focus on discovering the different variants by contrasting the parameterization of the diffeomorphic transformation models. We compare the relationship between these algorithms and discuss how they can be converted to each other by simplifying assumptions or changing optimization schemes, etc. We show that these large deformation algorithms may be viewed as extensions of the common small deformation image registration methods.

For a clear comparison, the variant notations used by different authors are unified in this chapter. We especially fill in the mathematical details, assuming a common engineering
background for readers, to aid in a complete understanding of the derivation of the most important equations. Many techniques commonly used in diffeomorphism can find roots in the derivation of the small deformation models and the regularization terms.

There are other important problems in diffeomorphic image registration but we could not cover them in this chapter. A few of them include the image similarity terms, the B-spline transform model, the second-order optimization schemes, the symmetric scheme and inverse consistency problem, the numerical methods with Eulerian and Lagrangian coordinates, and other interesting constraints based on the velocity parameterization.

3.2 From Small Deformation to Large Deformation in Image Registration

The goal of image registration is to find a viable transform from one image to the other. Such a transform defines pixel-wise correspondences between two images. To simplify notation, we assume that any image \( I \) discussed here is a scalar function defined on one spatial domain: \( I : \Omega \subseteq \mathbb{R}^d \rightarrow \mathbb{R} \). For example, a 3D image has \( d = 3 \). A transform \( \phi \) is a mapping between two spatial domains: \( \Omega \rightarrow \Omega \). If one transform \( \phi \) is applied to an image \( I \), a new warped image is generated as the composition of two functions \( I \circ \phi \), which is also a function \( \Omega \rightarrow \mathbb{R} \).

Given an template image \( I_0 \) (i.e. the fixed image) and an individual image to be warped \( I_1 \) (i.e. the moving image), the image registration problem optimizes a transform \( \phi \) such that the warped moving image \( I_1 \circ \phi \) is close to the fixed image \( I_0 \) and \( \phi \) is smooth. This is formulated as minimization of an energy function \( E(\phi) \), which usually comprises two parts:

\[
E(\phi) = \lambda_S S(I_0, I_1, \phi) + \lambda_R R(\phi),
\]

in which \( \lambda_S \) and \( \lambda_R \) are two scalar weights. The first term in the right is an image similarity function \( S(I_0, I_1, \phi) \), evaluating the similarity between two images using the transform \( \phi \).
How to choose an appropriate similarity function is a fundamental problem in image registration, which, however, is not a topic in this chapter. Here we adopt the sum of square difference (SSD) of two images as the similarity function throughout this chapter, for both its simplicity and familiarity. One can write SSD as:

$$S(I_0, I_1, \phi) = \|I_1 \circ \phi - I_0\|^2.$$  \hspace{1cm} (3.2)

The second part is a transform regularization function $R(\phi)$, which is usually solely defined on the transform. This term is a prior on the transform smoothness. For some specific parameterized transformation models, for example, the affine transformation or the B-spline transformation, the smoothness is implicitly encoded through corresponding parameterization forms. However, for general non-rigid registration when no other information is available, the transform $\phi$ is defined using a deformation vector field $u$: $\Omega \rightarrow \mathbb{R}$,

$$\phi(x) = x + u(x), \ x \in \Omega.$$ \hspace{1cm} (3.3)

Such a parameterization is simply adding an identity transform with a vector field. Regularization functions are normally defined on $u$ instead of directly on $\phi$. $u(x)$ is a zero vector field when $\phi(x) = x$. This form is strongly connected to Hooke’s law about the equilibrium position of elastic material.

In the rest of this section we particularly elaborate on the mathematical details for the easy understanding of the derivation details, which is important on the discussion of large deformation models.

### 3.2.1 A First Example: Horn-Schunck Optical Flow

There are many choices of regularization terms. One simple model is to constrain first order variation of the displacement field $u$. This is denoted as the membrane energy model [7]
and provides one form of Tikhonov regularization \[53\]:

\[
R(u) = \frac{1}{2} \int_\Omega \sum_{i=1}^d \sum_{j=1}^d \left( \frac{\partial u_i(x)}{\partial x_j} \right)^2 \, dx = \frac{1}{2} \int_\Omega \sum_{i=1}^d \| \nabla u_i(x) \|^2 \, dx = \frac{1}{2} \int_\Omega \| \nabla u(x) \|^2 \, dx.
\] (3.4)

Solving the Euler-Lagrangian equation of this functional gives the negative Laplacian function as its derivative:

\[
\frac{\partial R(u)}{\partial u_i} = -\sum_{j=1}^d \frac{\partial^2 u_i(x)}{\partial x_j^2} = -\Delta u_i(x), \quad i = 1 \cdots d.
\] (3.5)

When \( u \) is close to zero, the SSD function of Eq.3.2 can be approximated using the first order Taylor expansion:

\[
I_1 \circ (x + u) - I_0 \approx (I_1 - I_0) + \langle u, \nabla I_1(x) \rangle
\] (3.6)

Combining it with the Tikhonov regularization leads to the overall energy function of Horn-Schunck optical flow method:

\[
E(u) = \lambda_S \int_\Omega ((I_1 - I_0) + \langle u, \nabla I_1 \rangle)^2 \, dx + \lambda_R \int_\Omega \sum_{i=1}^d \| \nabla u_i(x) \|^2 \, dx
\] (3.7)

The Euler-Lagrangian equations of \( E(u) \) at each pixel \( x \) gives the linear system of the optimal point:

\[
\lambda_S (I_1 - I_0 + \langle u, \nabla I_1(x) \rangle) \frac{\partial I_1(x)}{\partial x} - \lambda_R \Delta u_i(x) = 0, \quad i = 1 \cdots d
\] (3.8)

An iterative optimization scheme was given in \[53\], which is essentially the Jacobi method of solving the large sparse linear system. Such a method only works for small \( u \). However, we are going to reformulate it and explore some important ideas adopted in large deformation methods.
Notice that the first term in the left is the derivative of the image similarity function and the second term is the derivative of the regularization function, we actually have a general form of the functional derivative as:

$$ \nabla_u E(u) = \nabla_u S(u) + \nabla_u R(u). \quad (3.9) $$

Various optimization schemes can be derived from this representation. For example, the steepest gradient descent is an iterative scheme as $u \leftarrow u - \epsilon \nabla E(u)$. By introducing the second order approximation of $E(u)$, one can also apply Gauss-Newton method, Levenberg-Marquardt algorithm, or the Newton method when the second order derivative of $E$ is feasible. In this chapter we mostly focused on the derivation of the first order gradient.

The optimal point can also be computed by iteratively solving the linear system $\nabla_u E = 0$. Specifically if we denote $Q$ as an operator on $u$: $Q(u) = \nabla_u R(u)$, we have $\nabla_u S(u) + Q(u) = 0$. $u$ is solved as:

$$ u = Q^{-1}(-\nabla_u S(u)) \quad (3.10) $$

As the regularization $R$ is solely defined on $u$, $Q^{-1}$ is independent of choice of $S$. This formulation now can be viewed as projecting of $-\nabla_u S(u)$ to a new vector space. Thus each iteration consists of two steps. The first gives the initial estimation of the displacement from the image observation. The second regularizes the estimation by projecting it to a smooth vector space. Such a two step approach is practically adopted in the optimization of many image registration algorithms and we will elaborate on this later.

For Horn-Schunck optical flow, $Q^{-1}$ is indeed the inverse operator of Laplacian and thus filters out high-frequency components of the deformation field $u$ in the Fourier domain. Later we will also analyze other transformation models through such a perspective of projecting the similarity function gradient to the smooth vector space.
3.2.2 Demons algorithm

Before introducing image registration of large deformation, we briefly go over another important small deformation registration method: Thirion’s demons algorithm [101]. The demons algorithm is one efficient method applied in medical image registration and has many variants. It also serves as a baseline algorithm of other large deformation models.

In Thirion’s original paper [101], the demons algorithm considered the image registration as a diffusion process. The pixels of images were pushed by demons with local forces inspired by optical flow equations. However, the diffusion scheme was described more or less like a heuristic procedure without a clear theoretical explanation. Cachier et al. [19] justified the algorithm as an alternating minimization of an energy function. The idea is to introduce an auxiliary variable of $c$ as the raw correspondences between pixels, which can be viewed as the observation of the underlying transform $\phi$:

$$E(\phi, c) = \frac{1}{\sigma_i^2} \| I_1 \circ c - I_0 \|^2 + \frac{1}{\sigma_x^2} \| c - \phi \|^2 + \frac{1}{\sigma_T^2} \int_\Omega \| \nabla u(x) \|^2 dx, \quad (3.11)$$

where $\sigma_i$ accounts for the noise on the image intensity, $\sigma_x$ accounts for a spatial uncertainty on the correspondences and $\sigma_T$ controls the amount of regularization. Now the demons algorithm is explained as an alternating optimization over $c$ and $\phi$. The first step solves for $c$ by optimizing $S(c) = \frac{1}{\sigma_i^2} \| I_1 \circ c - I_0 \|^2 + \frac{1}{\sigma_x^2} \| c - \phi \|^2$ given current $\phi$. This is similar to the optical flow equation with the constraint that $c$ should be close to $\phi$. The second step solves for $\phi$ by optimizing $R(\phi) = \frac{1}{\sigma_i^2} \| c - \phi \|^2 + \frac{1}{\sigma_T^2} \int_\Omega \| \nabla u(x) \|^2 dx$ with given $c$. This is actually the projection step we discussed in the last section. $c$ and $\phi$ are typically initialized as the identity transform.

Since $c$ and $\phi$ have to be close, let $c = \phi + w$. The minimization of $S(c)$ here used the
same first order expansion of $I_1 \circ c$ as in the optical flow algorithm:

$$S(c) = S(\phi + w) \approx \frac{1}{\sigma^2_t} (I_1 \circ \phi - I_0 + \langle w, \nabla_\phi I_1 \rangle)^2 + \frac{1}{\sigma^2_x} \|w\|^2,$$  \hspace{1cm} (3.12)

where $\nabla_\phi I_1$ means $(\nabla I_1)(\phi)$. The optimal $w$ is now given by

$$\left( \nabla_\phi I_1 \nabla_\phi^T I_1 + \frac{\sigma^2_x}{\sigma^2_t} \right) w + (I_1 \circ \phi - I_0) \nabla_\phi I_1 = 0$$  \hspace{1cm} (3.13)

Using the Sherman-Morrison formula, the solution yields one choice of demon displacement $w$ as:

$$w = -\frac{I_1 \circ \phi - I_0}{\|\nabla_\phi I_1\|^2 + \sigma^2_x/\sigma^2_t} \nabla_\phi I_1$$  \hspace{1cm} (3.14)

Optimizing $\phi$ (or $u = \phi - x$) given the auxiliary variable $c$ is a quadratic form:

$$R(\phi) = \frac{1}{\sigma^2_x} \|c - \phi\|^2 + \frac{1}{\sigma^2_T} \int_\Omega \|\nabla u(x)\|^2 dx,$$  \hspace{1cm} (3.15)

which has a closed form solution. Using the Euler-Lagrangian equation yields the optimal $u$ given $c$ as:

$$u = \left( \text{Id} - \frac{\sigma^2_x}{\sigma^2_T} \Delta \right)^{-1} (c - x),$$  \hspace{1cm} (3.16)

where $\text{Id}$ is the identity operator. Comparing this with Eq(3.10), now it is clear that this regularization step is also projecting the auxiliary variable $c - x$ using a smoothing operator $Q^{-1} = (\text{Id} - \sigma^2_x/\sigma^2_T \Delta)^{-1}$. Thirion’s demons algorithm applies Gaussian smoothing for efficient computation: $Q^{-1} u = G * u$, where $G$ is a Gaussian kernel and $*$ represents image convolution. For now let us simply assume that $(\text{Id} - \sigma^2_x/\sigma^2_T \Delta)^{-1} = G$. We will come back to the exact regularization terms corresponding to Gaussian smoothing later.

Now we summarize the demons algorithm in Alg. 1. It is called the additive demons since
the update step is additive.

**Algorithm 1 Additive Demons Algorithm**

1. Given current transformation $\phi$, compute the correspondence update field $w$ using Eq. 3.14.
2. Additive update: $c \leftarrow \phi + w$.
3. Regularization: $\phi \leftarrow \text{Id} + G \ast (c - \text{Id})$. Go back to 1.

### 3.2.3 Compositive Demons and Inverse Transform

The additive update performs a gradient descent on $w$ in the vector space $\mathbb{R}^d$. Another way to combine to $\phi$ and $w$ is through composition:

$$ c = \phi \circ (\text{Id} + w) \quad (3.17) $$

Its first order approximation at $w \rightarrow 0$ is similar to the additive one: $c = \phi + \nabla \phi \cdot w$. The corresponding similarity term becomes:

$$ S(c) = S(\phi \circ (\text{Id} + w)) \approx \frac{1}{\sigma_i} (I_1 \circ \phi - I_0 + \langle w, \nabla (I_1 \circ \phi) \rangle)^2 + \frac{1}{\sigma_x^2} \|w\|^2, \quad (3.18) $$

and the corresponding demons force $w$ becomes:

$$ w = -\frac{I_1 \circ \phi - I_0}{\|\nabla (I_1 \circ \phi)\|^2 + \sigma_i^2/\sigma_x^2} \nabla (I_1 \circ \phi) \quad (3.19) $$

We list the compositive demons algorithm in Alg. 2. If we set $\sigma_i$ as a pixel-wise function: $\sigma_i(x) = I_1 \circ \phi - I_0$, Eq. 3.19 is close to the form of the original demons force in [101]. There are other variants of demons forces by choosing different forms of $S$ and various second-order optimization schemes.

The compositive demons requires to compute the transform composition $\phi \circ (\text{Id} + w)$ instead of simply adding $\phi + w$. It also computes the gradient of the warped image $I_1$ as $\nabla (I_1 \circ \phi)$ (Eq. 3.19). In contrast, the additive update needs to warp the gradient image: $\nabla \phi I_1$ (Eq. 3.19).
Algorithm 2 Compositional Demons Algorithm
1. Given current transformation $\phi$, compute the correspondence update field $w$ using Eq. 3.19.
2. Additive update: $c \leftarrow \phi \circ (\text{Id} + w)$.
3. Regularization: $\phi \leftarrow \text{Id} + G * (c - \text{Id})$. Go back to 1.

Eq. 3.14, which may have higher numerical error since the gradient image is usually a sparse image.

A more important difference between the compositive and the additive update schemes is related to the invertibility of the transforms. When $w$ is small, it is approximated that $\text{Id} + w$ is invertible with $(\text{Id} + w)^{-1} = \text{Id} - w$. But this is generally false for large deformation $w$.

An example was illustrated in [7] (see Fig. 3.2). In the additive update scheme $c \leftarrow \phi + w$, $c$ usually quickly becomes non-invertible after several iterations. The adding operation in the vector space does not preserve invertibility.

On the other side, the composition of two invertible transforms is still invertible. This implies that the compositive scheme is able to yield an invertible transform when $w$ is small in each iteration, which was validated in [111]. This means that a large and invertible transform can be generated from the composition of many small deformation fields. In the following sections we are going to formalize this intuition and reveal this as the background for many large deformation transformation models.

3.2.4 Diffeomorphism Group

The small deformation algorithms we discussed so far parameterize the transform as $\phi(x) = x + u(x)$. Unless $u(x)$ is close to zero, such a form is not necessarily invertible and in general cannot preserve topology. In many applications, the nonrigid transform between two images are very large, where “large” is defined in a historical context, that is, relative to transformations typically captured by elastic or demons type algorithms. Traditionally, these small deformation algorithms could not both explicitly enforce a one-to-one mapping and provide a high-quality registration solution. It is therefore necessary to exploit the
Figure 3.2: Fig.1 from [7]. Inversion and composition in a small deformation setting. The compositions are not identity transform.
large deformation framework which can preserve topology while capturing a wide range of transformations.

Large deformation is conveniently represented as a diffeomorphism in mathematics. A diffeomorphism \( \phi \) is a globally one-to-one continuous and smooth mapping with a continuous and smooth inverse. In other words, \( \phi^{-1} \) exists and both \( \phi \) and \( \phi^{-1} \) are invertible. The diffeomorphism forms a group \( \text{Diff} \) under the composition operation, i.e. \( \phi_1 \circ \phi_2 \in \text{Diff} \) when \( \phi_1, \phi_2 \in \text{Diff} \):

\[
\text{Diff} = \{ \phi : \Omega \to \Omega | \phi \text{ and } \phi^{-1} \text{ are differentiable} \}.
\] (3.20)

Suppose there are a series of small deformations \( (\psi_k \in \text{Diff})_{0 \leq k \leq n} \), we recursively define \( \phi_0 = \text{Id} \) and \( \phi_{k+1} = \phi_k \circ \psi_k \). (\( \phi_k \)) forms a polygonal line in \( \text{Diff} \). Note that the solution of compositive demons falls into this representation. In a continuous setting this polygonal line become a curve: \( \phi(x, t) \), \( 0 \leq t \leq 1 \). Here \( t \) is a time variable and \( \phi(x, t) \) is the position of one particle whose position was \( x \) when \( t = 0 \).

When \( \phi(x, t) \) is differentiable in \( t \), we have the important o.d.e that generates a diffeomorphism:

\[
\frac{d}{dt} \phi(x, t) = v(\phi(x, t), t),
\] (3.21)

where \( v \) satisfies continuity conditions in order to guarantee the existence of the solution. But it is important to point out that \( v(x, t) \) does not need to be invertible w.r.t \( x \).

By introducing the time variable \( t \), we can define a curve of transformations between two images \( I_0 \) and \( I_1 \). Define \( \phi_t \) as the transform at time \( t \) and similarly for \( v_t \):

\[
\phi_t(x) = \phi(x, t)
\] (3.22)

\[
v_t(x) = v(x, t)
\] (3.23)
At time $t = 0$, $\phi_0 = \text{Id}$; at time $t = 1$, we get the final transform $\phi = \phi_1$. Since $\phi_t$ has an inverse transform, we can also define the transform between two time points $\phi_{s,t}$ as:

$$\phi_{s,t} = \phi_t \circ (\phi_s)^{-1}. \quad (3.24)$$

We immediately have the following properties:

$$\phi_{t,s} = \phi_{s,t}^{-1} \quad (3.25)$$
$$\phi_{t,s} = \phi_{r,t} \circ \phi_{s,r} \quad (3.26)$$
$$\phi_1 = \phi_{0,1} \quad (3.27)$$
$$\phi_1^{-1} = \phi_{1,0} \quad (3.28)$$

The o.d.e Eq. 5.1 gives an important way to parameterize $\phi$ with the time-variant velocity fields $v_t$. Instead of only modeling the target transform $\phi_{0,1}$, such a parameterization encodes the whole temporal path that determines how one image $I_1$ is deformed into another image $I_0$. Since $\phi_t$ is a diffeomorphism, this path is an invertible one-to-one mapping. An invertible transform $\phi_1$ can be constructed from a time-variant velocity vector fields $v_t$, which does not need to be invertible. Such a parameterization of $\phi$ avoids handling the existence of $\phi^{-1}$ directly, which is a desirable property. In the following sections we discuss in detail how different approaches compute $v$ in the setting of image registration. We show the relationship between different diffeomorphism approaches and small deformation algorithms by deriving them from each other.

### 3.2.5 Connection Between Small deformation and Diffeomorphism

We already saw that the composite demons is close to large deformation image registration. It is helpful to view Eq. 5.1 by discretization of $t$. Suppose we discretize $\phi_t$, $0 \leq t \leq 1$ into
a series of \((\phi_i)_{0 \leq i \leq N}\), at the \(i\)-th time point, we have:

\[
\begin{align*}
\frac{d}{dt}\phi_t &= \mathbf{v}_t(\phi_t),
\frac{\phi_{i+1} - \phi_i}{\delta t} &\approx \mathbf{v}_t(\phi_t),
\phi_{i+1} &= \phi_i + \delta t\mathbf{v}_t \circ \phi_t,
\phi_{i+1} &= (\text{Id} + \delta t\mathbf{v}) \circ \phi_t.
\end{align*}
\]

(3.29)

Let \(\mathbf{w} = \delta t\mathbf{v}\) and we have:

\[
\phi_{i+1} = (\text{Id} + \mathbf{w}) \circ \phi_i
\]

(3.30)

Comparing with the composition used in Eq. 3.17 we see the difference is whether the small deformation approximation occurs before or after the last transform \(\phi_i\). The advantage of Eq. 3.17 is that it is easier to compute the derivative of \(\nabla_w(\phi \circ (\text{Id} + \mathbf{w}))\); but physically it is more natural to compose the update \(\mathbf{w}\) after the current transform \(\mathbf{w}\). We will discuss the derivation from Beg. et al [15] in the Sec. 3.4.

Another important fact about this discretization is that although \(\text{Id} + \mathbf{v}\) is not necessarily invertible, \(\text{Id} + \epsilon \mathbf{v}\) is usually invertible for a small scalar \(\epsilon\) with \((\text{Id} + \epsilon \mathbf{v})^{-1} \approx \text{Id} - \epsilon \mathbf{v}\). Thus \(\phi_N\) is also invertible. We will revisit this in Ashburner’s method [7] and Vercauteren’s methods [111, 110].

Theoretically we only require \(\mathbf{v}\) to be continuous to satisfy the existence of \(\phi\). It is more desirable that \(\mathbf{v}\) is also smooth, both for numerical issues and to provide an additional smoothness control on \(\phi\). The smoothness constraint we discussed before about the small deformation \(\mathbf{w}\) is now equivalent to the smoothness of the velocity field \(\mathbf{v}\). The regularization term \(R(\mathbf{u})\) where \(\mathbf{u} = \phi - \text{Id}\) naturally becomes a function on \(\mathbf{v}\): \(R(\mathbf{v})\).
3.3 Smooth Vector Fields and Sobolev Gradient

Before we discuss various forms of parameterization $v$ of diffeomorphism $\phi$, it is necessary to discuss the regularization on the vector fields $v$. A general prior about the final solution of the optimal transform is that it has to be a smooth one. The main aim of regularization in diffeomorphic image registration is to guarantee smoothness of $v$ under noisy imaging conditions and thus implicitly of the resulting transform $\phi$. Velocity field regularity also improves stability of the numerical integration scheme for both the forward and the inverse transform.

The regularization term is an important factor not only in formulating the registration energy function itself, but also in choosing the corresponding optimization schemes. A vector field $v$ is usually defined in an Euclidean space $L^2$ of square integrable functions. However, this vector space does not have any smoothness constraints. Hilbert space was introduced in the image registration domain by many researchers [104, 15, 48, 121]. Hilbert space can be viewed as a space of smooth functions. This section summarizes the theoretical analysis of regularization as minimizing the image similarity directly in Hilbert space. Using the notion of optimization in Hilbert space we show that various diffeomorphism optimization schemes are unified in the same framework in the later sections.

3.3.1 Examples of Regularization Terms and Their optimization Operators

The smoothness of a vector field $v$ is usually defined using its $k$-th order spatial derivatives. Before the discussion of Hilbert space, it is helpful to revisit the formulation of several common regularization terms used in the literature [53, 15, 7, 111] $R(v)$ and their derivatives $\partial R/\partial v$, which all can be derived using Euler-Lagrangian equations. Note that it requires the $k$-th order derivatives of $v$ vanish at the boundary of the domain $\Omega$. 

39
Membrane Model

In Sec. 3.2.1 we already went over the membrane model:

\[
R(v) = \frac{1}{2} \int_{\Omega} \sum_{i=1}^{d} \| \nabla v_i(x) \|^2 \, dx ,
\]

(3.31)

\[
\frac{\partial R(v)}{\partial v_i} = -\Delta v_i(x), \ i = 1 \cdots d .
\]

(3.32)

Bending Energy Model

\[
R(v) = \frac{1}{2} \int_{\Omega} \sum_{i,j,k=1}^{d} \left( \frac{\partial^2 v_i(x)}{\partial x_j \partial x_k} \right)^2 \, dx
\]

(3.33)

\[
\frac{\partial R(v)}{\partial v_i} = \sum_{j,k=1}^{d} \frac{\partial^4 v_i(x)}{\partial x_j^2 \partial x_k^2} = \Delta^2 v_i(x), \ i = 1 \cdots d .
\]

(3.34)

Laplacian model

\[
R(v) = \frac{1}{2} \int_{\Omega} \sum_{i=1}^{d} (\Delta v_i(x))^2 \, dx = \frac{1}{2} \int_{\Omega} \sum_{i=1}^{d} \left( \sum_{j=1}^{d} \frac{\partial^2 v_i(x)}{\partial x_j^2} \right)^2 \, dx
\]

(3.35)

\[
\frac{\partial R(v)}{\partial v_i} = \sum_{j,k=1}^{d} \frac{\partial^4 v_i(x)}{\partial x_j^2 \partial x_k^2} = \Delta^2 v_i(x), \ i = 1 \cdots d .
\]

(3.36)

Note that the Laplacian model and Bending model [7] actually share the same derivation (comparing Eq. 3.34 and Eq. 3.36).
Cauchy-Navier Model

This is used in Beg’s paper [15] and it is a variant of the Laplacian model. With a little abuse of notation, define the operator \( \mathcal{L} \) as

\[
\mathcal{L} = \gamma \text{Id} - \alpha \Delta
\]

Thus,

\[
R(v) = \frac{1}{2} \int_{\Omega} \| \mathcal{L} v \|^2 \, dx = \frac{1}{2} \int_{\Omega} \sum_{i=1}^{d} \left( \gamma v_i(x) - \alpha \sum_{j=1}^{d} \frac{\partial^2 v_i(x)}{\partial x_j^2} \right)^2 \, dx
\]

(3.37)

(3.38)

\[
\frac{\partial R(v)}{\partial v_i} = \gamma^2 v_i^2 - 2 \gamma \alpha \sum_{j=1}^{d} \frac{\partial^2 v_i(x)}{\partial x_j^2} + \alpha^2 \sum_{j,k=1}^{d} \frac{\partial^4 v_i(x)}{\partial x_j \partial x_k^2}
\]

(3.39)

\[
= (\gamma \text{Id} - \alpha \Delta)^2 v_i
\]

\[= L^2 v_i, \ i = 1 \cdots d.\]

Tikhonov Regularization and Gaussian Smoothing

In Sec. 3.2.2 we conveniently explain how the Gaussian smoother \( G \) acts as an inverse operator: \( G = (\text{Id} - \sigma_x^2 / \sigma_y^2 \Delta)^{-1} \) for some \( \Delta \). Strictly speaking, \( G \) relates to the general Tikhonov Regularization [76, 69]¹:

\[
R(v) = \frac{1}{2} \int_{\Omega} \sum_{i=1}^{d} v_i^2(x) + \sum_{k=1}^{\infty} \left( \sum_{1 \leq j_1, \cdots, j_k \leq d} \frac{1}{\sigma_{2k}^2 k!} \left( \frac{\partial^k v_i(x)}{\partial x_{j_1} \cdots \partial x_{j_k}} \right)^2 \right) \, dx
\]

(3.40)

¹Note: A typo is in [69]: the subscript of the equation in Sec 2.2 should be \( 1 \leq i_1, \cdots, i_k \leq d \), not \( i_1 + \cdots + i_k = k \).
As a practice on solving regularization, we provide the full derivation detail in this section.

By applying Euler-Lagrangian, we have:

\[
\frac{\partial R}{\partial v_i} = v_i + \sum_{k=1}^{\infty} \sum_{1 \leq j_1 \leq \cdots \leq j_k \leq d} \frac{(-1)^k}{\sigma^{2k} k!} \frac{\partial^{2k}}{\partial x_{j_1}^2 \cdots \partial x_{j_k}^2} v_i
\]

\[
= v_i + \sum_{k=1}^{\infty} \frac{(-1)^k}{\sigma^{2k} k!} \left( \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2} \right)^k v_i
\]

\[
= \left( \text{Id} + \sum_{k=1}^{\infty} \frac{(-1)^k}{\sigma^{2k} k!} \Delta^k \right) v_i.
\]

When \( k \) only goes to 1, not \( \infty \), \( \partial R/\partial v_i \) becomes \( (\text{Id} - 1/\sigma^2 \Delta) \), which was used in Sec. 3.2.2.

Using the notation of operator \( Q \) in Eq. 3.10, we can now compute its inverse operator \( Q^{-1} \) from the Fourier transform of the vector field \( v_i(x) \). Let \( \mathcal{F} \) denote the Fourier transform, \( \omega \) be the frequency variable, and \( \mathcal{F}(v_i(x)) = \hat{v}_i(\omega) \). We have

\[
\mathcal{F} \left( \Delta^k v_i(x) \right) = (-1)^k \omega^{2k} \hat{v}_i(\omega)
\]

\[
\mathcal{F} (Qv_i(x)) = \mathcal{F} \left( \left( \text{Id} + \sum_{k=1}^{\infty} \frac{(-1)^k}{\sigma^{2k} k!} \Delta^k \right) v_i(x) \right)
\]

\[
= \sum_{k=0}^{\infty} \frac{\omega^{2k}}{\sigma^{2k} k!} \hat{v}_i(\omega)
\]

\[
= \exp \left( \frac{\omega^2}{\sigma^2} \right) \hat{v}_i(\omega)
\]

To get the Fourier Transform of operator \( Q^{-1} \), we simply compute the inverse of the coefficient at each \( \xi \):

\[
\mathcal{F} (Q^{-1} v_i(x)) = \exp \left( -\frac{\omega^2}{\sigma^2} \right) \hat{v}_i(\omega)
\]

Now it is clear that the inverse Fourier transform leads to the operator as Gaussian convolution smoother \( G \).
### 3.3.2 Sobolev Norm and Sobolev Space

The various examples of regularization terms $R(v)$ showed in the previous Sec. 3.3.1 all share the same quadratic form of $k$-th order derivatives.

$$R(v) = \sum_{i=0}^{K} \lambda_i \langle v^{(i)}, v^{(i)} \rangle,$$  \hfill (3.44)

where $v^{(i)}$ is the $i$-th order derivative of $v$. For simplicity, we omit the scalar coefficient $\lambda_i$ in this section. Following the discussion in [121], using the differential operator $L = (D^0, \cdots, D^k)$ and the $L^2$ norm $\| \cdot \|_{L^2} = \langle \cdot, \cdot \rangle$, the Sobolev norm $\langle \cdot, \cdot \rangle_{H^k}$ is defined as:

$$\langle v, v \rangle_{H^k} = \sum_{i=0}^{k} \langle v^{(i)}, v^{(i)} \rangle_{L^2} = \langle L^* v, L v \rangle_{L^2}.$$  \hfill (3.45)

The quadratic form requires that the square of up to the $k$-th order derivative of smooth function $v$ is integrable. Thus we formalize it as $v$ belongs to the Sobolev space $H^k$: $H^k = \{ f : \| f \|_{H^k} < \infty \}$. For simplicity, we omit $k$ in $H^k$. The inner product in Sobolev space is defined as:

$$\langle u, v \rangle_H = \langle L u, L v \rangle_{L^2} = \langle L^+ L u, v \rangle_{L^2},$$  \hfill (3.46)

where $L^+$ is called as the adjoint operator of $L$.

Now we can simply write the regularization term $R(v)$ as: $R(v) = \| v \|_{H^k}$. Moreover the derivative of $R(v)$ has the form:

$$\nabla v R = L^+ L v = \left( \sum_{i=0}^{k} (-1)^i \Delta^i \right) v.$$  \hfill (3.47)
3.3.3 Registration in Sobolev Space

Now we reexamine the image registration energy function Eq. 3.1

\[ E(v) = \lambda_S S(I_0, I_1, v) + \lambda_R R(v), \]  

(3.1)

Using the notion of the smooth function space \( H \), the regularization term \( R(v) \) can be replaced by the constraint that \( v \in H \):

\[ E(v) = S(I_0, I_1, v), \text{ with } v \in H. \]  

(3.48)

Now the energy is the image similarity term only. Accordingly the optimization needs to be computed in Sobolev space to keep \( v \) remaining in \( H \) at each iteration step. Sobolev gradient \( (\nabla_v E)_H \) is used to replace the previous \( L^2 \) gradient \( (\nabla_v E)_{L^2} \). We derive their relationship using the first order Taylor expansion with Sobolev inner product:

\[ E(v + h) \approx E(v) + \langle (\nabla_v E)_{L^2}, h \rangle_{L^2} \]
\[ = E(v) + \langle (\nabla_v E)_H, h \rangle_H \]
\[ = E(v) + \langle \mathcal{L}^+ \mathcal{L} (\nabla_v E)_H, h \rangle_{L^2} \]

(3.49)

Thus we can express the Sobolev gradient in terms of the \( L^2 \) gradient as:

\[ (\nabla_v E)_H = (\mathcal{L}^+ \mathcal{L})^{-1}(\nabla_v E)_{L^2} \]  

(3.50)

The projecting idea we discussed in deriving Horn-Schunck optical flow Sec. 3.2.1 is now also validated using the Sobolev gradient by comparing Eq. 3.50 with Eq. 3.10. Also from Sec. 3.3.1 we know that the Gaussian smoothing is the inverse operator for the gradient of the first-order Tikhonov regularization. This validates the Gaussian smoothing step used in
the demons algorithm Alg.1 and 2. Thus we conclude that various registration algorithms we have discussed can be unified as minimizing the image similarity using the Sobolev gradient on the transform parameter $v$, which is also true for the diffeomorphic approaches from Beg et al [15], Vercauteren et al [109, 110].

Sobolev Gradient and Preconditioning

Another interesting property about the Sobolev gradient was discussed in [121]. If the operator $L$ is written as $L = (\text{Id}, L_1)$, we get:

$$L^+ L = \text{Id} + L_1^+ L_1$$

Eq. 3.50 now can also be explained using $(\text{Id} + L_1^+ L_1)^{-1}$ as a preconditioner for the $L^2$ gradient of $(\nabla_v E)_{L^2}$. This connects the Sobolev gradient with various Gauss methods with preconditioner like the Gauss-Newton method and the Levenberg-Marquardt algorithm. Thus we can also unify the Levenberg-Marquardt algorithm applied in Ashburner’s method [7].

3.4 Diffeomorphism with Time-varying Velocity Fields

We have formalized the unifying registration framework using the Sobolev gradient. From this section, we study several forms of parameterization $v$ applied in large deformation algorithms and discuss the connections among them and with small deformation algorithms.

Small deformation transforms normally use a small displacement field $u$ to model the mapping $\phi = x + u$. In contrast, a large deformation transform introduces an extra time variable $t$ to encode the warping path $\phi(x, t)$ between two images. The first parameterization is using
the o.d.e Eq. 5.1

$$\frac{d}{dt} \phi_t(x) = v_t(\phi_t(x)), \quad (5.1)$$

Optimizing the diffeomorphism transform $\phi$ is equivalent to optimizing the time-varying velocity field $v_t$. An example of $v_t$ was illustrated in [15] (see Fig. 3.3).

3.4.1 Beg’s LDDMM algorithm

Beg et al. proposed Large Deformation Diffeomorphic Metric Mapping (LDDMM) algorithm, which used this o.d.e as parameterization to optimize the following energy function:

$$E(v) = \frac{1}{2} \| I_0 \circ \phi_{1,0} - I_1 \|^2 + \frac{\sigma^2}{2} \int_0^1 \|Lv_t\|^2 dt, \quad (3.52)$$
where $\phi_{1,0}$ is defined by Eq. 3.28 in Sec. 3.2.4, the operator $L$ represents the Cauchy-Navier regularizer $R(v)$: $L = \gamma \text{Id} - \alpha \Delta$ in Sec. 3.3.1 with the Hilbert space $H$. Note that we omit the integral over the spatial domain $\int_x \cdot dx$ for simplicity since $v$ can be reinterpreted as a vector with an element for each pixel. The image similarity term is then $S(v) = \frac{1}{2} \|I_0 \circ \phi_{1,0} - I_1\|^2$.

Beg et al gave a rigorous way in computing the derivative of $\nabla_v E$ in [15]. The key in the proof is to introduce the Gateaux variation of $\phi_{s,t}$ w.r.t $v$ (Lemma 2.1 from [15]). A small perturbation of $v$ at time $r$ (i.e. $h_r$) affects all the transform $\phi_t$ for $t > r$ cumulatively. Since the derivation in [15] is detailed, we cite this lemma without proof here:

$$\partial_h \phi_{s,t} = D\phi_{s,t} \int_s^t (D\phi_{s,r})^{-1} h_r \circ \phi_{s,r} dr \quad (3.53)$$

Although it is hard to define the Fréchet derivative for $\phi$, which means that the simple chain rule could not be applied directly to compute $\nabla_v S$, Beg however showed that $\partial_v S$ still has a nice structure related to the Fréchet derivative (see the proof of Theorem 2.1 in [15]):

$$\partial_h S(v) = - \int_0^1 \langle |D\phi_{t,1}|(J^0_t - J^1_t)D(J^0_t), h_t \rangle dt , \quad (3.54)$$

where $J^0_t = I_0 \circ \phi_{t,0}$, $J^1_t = I_1 \circ \phi_{t,1}$, $D$ is the Jacobian matrix and $\| \cdot \|$ is the determinant value of the matrix. For the regularization term, the Gateaux variation is easy to compute:

$$\partial_h R(v) = \sigma^2 \int_0^1 \langle L^+ L v, h_t \rangle dt \quad (3.55)$$

Now the overall Gateaux variation $\partial_h E(v)$ can be represented as: $\partial_h E(v) = \int_0^1 \langle \nabla_v E_t, h_t \rangle dt$ and $\nabla_v E_t$ is defined as:

$$\nabla_v E_t = L^+ L(\sigma^2 v_t) - |D\phi_{t,1}|(J^0_t - J^1_t)D(J^0_t) \quad (3.56)$$
Beg et al proposed to use the Sobolev gradient instead of this $L^2$ gradient. From Eq. 3.50, we get the gradient:

\[(\nabla_{v_t} E_{t})_H = \sigma^2 v_t - (L^+ \mathcal{L})^{-1}(|D\phi_{t,1}|(J^0_t - J^1_t)D(J^0_t))\] (3.57)

### 3.4.2 Numerical Algorithm of LDDMM

In its numerical implementation, the time-varying velocity fields are discretized into $N$ time points $(v_t)_0 \leq i \leq N-1$. For each time point $i$, the steepest descent scheme in the Sobolev space is applied:

\[v_{ti} \leftarrow v_{ti} - \epsilon(\nabla_{v_{ti}} E_{ti})_H.\] (3.58)

The inverse of $L^+ \mathcal{L}$ in the Sobolev gradient can be computed from the Fourier domain. One has to compute $\phi_{t,1}$ and $\phi_{t,0}$ for $(\nabla_{v_{ti}} E_{ti})_H$ from Eq. 3.54. $\phi_{t,1}$ can be computed directly by composing $(1/N)\phi_{tN-1} \circ \cdots \circ (1/N)\phi_{ti}$. For $\phi_{t,0}$, one needs to compute the inverse of $\phi_{0,t}$. Beg proposed to use a Semi-Lagrangian scheme in Sec. 3.3 of [15] for stable numerical computation, which is not covered in this chapter. But in general to get the inverse transform is computationally intensive using a series of $v_t$.

LDDMM method requires that one store $N$ velocity fields. In each iteration, it needs to compute also $N$ gradient fields, $N$ compositions for $\phi_{t,1}$ and solve $N$ inverse problems. Overall, this is a very expensive algorithm.
3.4.3 Comparison with Horn-Schunck Optical Flow

Revisiting the Horn-Schunck optical flow update Eq. 3.8 with the approximation Eq. 3.6, one can rewrite it as:

\[ \nabla u E = (I_0 \circ \phi - I_1)DI_1 - \lambda \Delta u . \] (3.59)

Comparing with the gradient at time \( t \) of LDDMM (Eq. 3.56), \( \nabla v_t E_t \) is conveniently explained as the optical flow between \( J^0_t \) and \( J^1_t \) at time \( t \), weighted by the Jacobian determinant \( |D\phi_{t,1}| \) introduced by change of variable of \( y = \phi_{t,1}(x) \).\(^2\)

Thus the LDDMM algorithm can be summarized as computing \( v_{t_i} \) in the way of the optical flow for all time \( t_i \) and composing all \( (v_{t_i}) \) to get \( \phi_{t,1} \) and \( \phi_{t,0} \).

3.4.4 Comparison with Compositive Demons

Compositive demons described in 3.2.3 can also be viewed as a greedy optimization of LDDMM. In compositive demons, when the iterative optimization procedure uses iterations, the update \( w_i, 1 \leq i \leq N \) from each iteration can be viewed as the velocity fields. At iteration \( i \), the greedy optimization fixes all the velocity fields obtained before \( i \) and only optimize \( w_i \).

The advantage of such a greedy way is that this does not require to store all the velocity fields before \( i \); only their composition is needed. The disadvantage is that such a greedy method does not put any constraint on the overall quality of \( w_i \). Each \( w_i \) is a local optimal but the transform path following all \( w_i \) may be very curvy. In fact, Beg et al also showed that the regularization used in LDDMM has the same optimal transform as the shortest path to Id.

\(^2\)One small difference is that the image similarity is defined as \( \|I_0 \circ \phi_{1,0} - I_1\|^2 \), not \( \|I_1 \circ \phi_{0,1} - I_0\|^2 \) (Eq 3.2).
3.5 Diffeomorphism with Stationary Velocity Fields

The LDDMM method [15] uses the time varying velocity fields $v_t$ to parameterize the diffeomorphism $\phi$. Although the derivation is rigorous, the algorithm has a high complexity and also requires stable numerical methods in solving the inverse transform. One way to simplify the problem is to use stationary velocity fields $v$, which does not change across time, to generate the diffeomorphism [6, 7, 109, 110, 49]. The o.d.e that generates the diffeomorphism is now written as:

$$\frac{d}{dt} \phi_t(x) = v(\phi_t(x)), \quad (3.60)$$

The stationary velocity field $v$ is referred as the one-parameter generator of a subgroup $\Phi = (\exp(tv))_{t \in \mathbb{R}}$, where $v$ belongs to the tangent space of $\Phi$. The final transform is parameterized as the exponential at unit time $\exp(v)$. The theory about such an infinite-dimensional space with a Lie group structure is still under research [6, 110, 97, 68].

Using a stationary velocity field $v$ to replace the time varying velocity fields $v_t$ greatly simplify the computation. Especially we have

$$\phi_{s,t} = \phi_{t-s} = \exp((t-s)v) \quad (3.61)$$

This shows that using stationary velocity fields, $\phi_{s,t}$ only depends on the length between two time points and not on the starting point.

Although not every diffeomorphism can be parameterized by stationary velocity fields [7], such a model performs well empirically for medical image registration [49, 111]. Parameterization of stationary velocity fields not only simplifies the computation; it also provides the possibility of statistics study using one velocity field, which is desirable in many applications.
3.5.1 Computation of Exponential Mapping

In Sec. 3.4.2 we discussed that computing $\phi$ requiring $N$ times composition of vector fields and solve $N$ steps of the p.d.e of the inverse transform $\phi^{-1}$. With the stationary field, such complexity can be reduced to $\log N$ times composition for both $\phi$ and $\phi^{-1}$.

Arsigny et al [6] introduced an efficient method to compute the exponential mapping $\phi = \exp(u)$ by exploiting the following property:

$$
\exp(v) = \exp(N^{-1}v)^N \text{ for integer } N.
$$

(3.62)

Using the first order approximation $\exp(N^{-1}v) \approx \text{Id} + N^{-1}v$ for sufficiently large $N$, the following ”scaling and squaring” algorithm (Alg.3) computes recursively $K$ times for $N = 2^K$:

Algorithm 3 Scaling-and-squaring for Exponential Mapping

1. Scaling: choose $K$ s.t $2^{-K}v$ is close to 0, e.g, $\max_x \|2^{-K}v(x)\| \leq 0.5$.
2. First-order approximation: $\phi^0 = \text{Id} + N^{-1}v \approx \exp(2^{-N}v)$.
3. Squaring: recursive $K$ steps $\phi^{k+1} \leftarrow \phi^k \circ \phi^k$.

Another important advantage is that one can also compute $\phi^{-1}$, the inverse of $\phi$, using the same scaling-and-squaring method with $\exp(-N^{-1}v) \approx \text{Id} - N^{-1}v$.

3.5.2 Ashburner’s DARTEL algorithm

Ashburner proposed to use stationary velocity field for diffeomorphic image registration. The algorithm was named DARTEL: Diffeomorphic Anatomical Registration using Exponentiated Lie algebra. DARTEL can be viewed as a stationary version of LDDMM in many ways. DARTEL uses the image registration energy function as:

---

3In fact, this is the minimal amount of computation required. Increased integration accuracy requires a dense sampling of the o.d.e. in time and can require many more integration points than discretization points, in particular if an accurate/consistent forward and inverse transformation is desired.
\[ E(v) = \frac{1}{2} \| I_0 - I_1 \circ \phi \|^2 + \frac{\sigma^2}{2} \| \mathcal{L}v \|^2 , \quad (3.63) \]

Compared with Eq. 3.56, the \( L^2 \) first order derivative of \( S \) becomes:

\[ \nabla_v S(v) = \int_0^1 |D\phi_{-t}|(I_0 \circ \phi_{-t} - I_1 \circ \phi_{1-t})D(I_1 \circ \phi_{1-t}) \, dt , \quad (3.64) \]

Previously we discussed that it is expensive to compute this gradient in LDDMM. But luckily, Ashburner [7] also derived a recursive scheme for \( \nabla_v S(v) \) using \( \log N \) times operation. Using together the scaling-and-squaring algorithm to compute \( \phi \), each gradient update altogether only needs \( \log N \) steps of operation.

Ashburner used the Levenberg-Marquardt algorithm in \( L^2 \) space as the optimization scheme, for which the second order derivatives of the energy also was computed. From the discussion in Sec. 3.3.3, one could also choose to use the gradient descent in the Sobolev space acting as the preconditioner, which was used in Beg et al’s LDDMM algorithm [15].

**Comparison with LDDMM**

DARTEL and LDDMM both models the transform from the o.d.e generator. Both algorithms compute the integration over time using \( N \) discretized time points. Both have a rigorous derivation of the energy gradient. However, the DARTEL algorithm used the stationary velocity field \( v \) as the parameterization of the transform \( \phi \). DARTEL is a simplified way to approximate LDDMM by approximating the velocity fields at all time points with their average.

The DARTEL update of the velocity field can be viewed as the average of the optical flow at every time point. Instead of having to store the velocity field at every time point, DARTEL only stores one velocity field. LDDMM needs to compose (minimum) \( N \) different velocity
fields to get the transform and compute the gradient at each time point. DARTEL only needs $\log N$ times composition to compute both transform and the gradient.

Another difference between DARTEL and LDDMM is whether the optimization occurs in $L^2$ space or the Sobolev space. DARTEL computes sparse matrix inverse used in each Levenberg-Marquardt iteration. Although Ashburner proposed to use the multi-grid method with Gauss-Seidel iteration, such a computation still has high complexity compared to the convolution approach in the Fourier domain used in LDDMM. The higher frequency in the operator $L^+L$ used in $L^2$ gradient also leads to numerical instability.

3.5.3 Vercauteren’s Diffeomorphic Demons Algorithm

Vercauteren extended the composite demons algorithm in Sec. 3.2.3 with diffeomorphism updates [111]. Following notations used in Sec. 3.2.2 in each iterative step, the diffeomorphic transformation can be represented as:

$$\phi \leftarrow \phi \circ \exp(w)$$  \hspace{1cm} (3.65)

If $\phi$ was initialized as a diffeomorphism (i.e. $\text{Id}$), the composition of the fields $w$ from all the iterations are $\phi = \exp(w_N) \circ \cdots \circ \exp(w_1)$. As we know that $\exp(w) = \text{Id} + w + o(\|w\|)$, the composite demons is a first order approximation of such a parameterization. Note that the term $\|w\|^2 = \|c - \phi\|^2$ used in Eq. 3.18 can be now reinterpreted as $\|w\|^2 = \|\log(\phi^{-1} \circ c)\|^2$.

Unlike LDDMM and DARTEL, Vercauteren et al did not compute the exact analytical form of $\nabla_w \exp(w)$ for all $w$. In contrast they explored from the perspective of Lie group. The velocity field $v$ belongs to the the tangent space of the diffeomorphism transformation group. The exponential mapping $\exp(w)$ maps from a neighborhood of 0 in the vector tangent space to a neighborhood of $\text{Id}$ in the diffeomorphism transform group.

When $w$ is close to 0, one only checks the derivative $\partial_w \exp(w)$ at 0, similar to the small
The update step used in diffeomorphic demons actually is an approximation of DARTEL algorithm. Instead of averaging the optical flow at every time points, diffeomorphic demons only used the optical flow (in other words, the demons forces) at time 0 as an approximation. Note that this is validated only when $\mathbf{w}$ is small and not true in general. Unlike DARTEL, the final transform $\phi$ cannot be represented as the exponential mapping any more.

Similarly to compositive demons, diffeomorphic demons did not track all the velocity fields generated in each iteration. It is more efficient and more greedy than LDDMM. The difference with compositive demons is that the transform is updated by composing with $\exp(\mathbf{w})$, instead of $(\text{Id} + \mathbf{w})$ directly. This can be roughly reinterpreted as going forward one unit time in the diffeomorphism group space, instead of in the additive vector space.

Diffeomorphic demons thus can be viewed as a middle method between DARTEL and compositive demons. In practice, these three method have comparable performance \cite{50,111} in image matching results but may result in different transforms.

### 3.5.4 Vercauteren’s LogDemons Algorithm

In order to keep the final transform in the exponential form $\exp(\mathbf{v})$, Vercauteren et al proposed LogDemons \cite{110} to update the velocity field directly in the log domain. The aim
is to combine update $\mathbf{w}$ to current velocity $\mathbf{v}$ in a new exponential transform such that:

$$\exp(Z(\mathbf{v}, \epsilon \mathbf{w})) = \exp(\mathbf{v}) \circ \exp(\epsilon \mathbf{w}) ,$$  
\hspace{1cm} (3.67)

where $\epsilon$ is a small scalar to emphasize that $\epsilon \mathbf{w}$ is small.

Although mathematically it is still not clear about the Lie group structure of the infinite-dimensional space, Bossa et al has shown the Baker-Campbell-Hausdorff formula for Lie group could be applied successfully. In the settings of LogDemons, Vercauteren used the following first order approximation:

$$Z(\mathbf{v}, \epsilon \mathbf{w}) = \mathbf{v} + \epsilon \mathbf{w} + \frac{1}{2}[\mathbf{v}, \epsilon \mathbf{w}] + \frac{1}{12}[\mathbf{v}, [\mathbf{v}, \epsilon \mathbf{w}]] + O(\|\epsilon \mathbf{w}\|^2) ,$$  
\hspace{1cm} (3.68)

with the Lie bracket $[\cdot, \cdot]$ defined as: $[\mathbf{v}, \mathbf{w}] = |D(\mathbf{v})|\mathbf{w} - |D(\mathbf{u})|\mathbf{w}$ and $|D \cdot|$ represents the Jacobian determinant.

LogDemons also defines the regularization of $\phi$ as $\|D \log(\phi)\|^2 = \|\mathbf{v}\|^2$. Thus LogDemons has the update step and regularization step as:

$$\mathbf{c} = \exp(\mathbf{v}) \text{ and } \mathbf{v} \leftarrow Z(\mathbf{v}, \mathbf{w})$$  
\hspace{1cm} (3.69)

$$\mathbf{v} \leftarrow G * \mathbf{v}$$  
\hspace{1cm} (3.70)

**Comparison with DARTEL**

LogDemons completely works in the log domain of the exponential mappings. It has the same parameterization as DARTEL, which makes them share some similar advantages, like efficient computation of exponential mapping and providing the velocity fields for further statistical analysis.

Unlike DARTEL, LogDemons is extended from diffeomorphic demons approach. It needs
to satisfy the assumption that each update of velocity field has to be small. In contrast, DARTEL has a more rigorous way to compute the gradient, which can be explained as the averaging of optical flows at all time points. Another difference is that LogDemons only used the optical flow at time 0, as in the diffeomorphic demons. But LogDemons utilizes BCH formula to integrate the update back into the velocity field in the log domain.

3.6 Conclusion

Diffeomorphism transform is a nice mathematical model for the image registration problem of large deformation. We mainly discussed four types of parameterization models used in the diffeomorphism registration: Beg’s LDDMM [15], Ashburner’s DARTEL [7], Vercauteren’s diffeomorphic demons [111] and LogDemons [110].

A unified framework was first established from the discussion of two classic small deformation methods: Horn-Schunck optical flow [53] and Thirion’s demons [101]. We explained in detail the important derivation of the small deformation model and discussed their intuition. These small deformation methods serve as the base algorithm for their large deformation counterparts.

For all the small and large deformation methods we discussed in this chapter, they shared the same optimization procedure. First, the derivatives of image similarity term w.r.t the transform parameters were computed. Second, the parameters was updated using a smoothing operation according to different types of regularization. Third, the transform was recomputed using the updated parameters.

The smoothing operation on the transform parameters were derived from the regularization terms. We discussed several regularization terms and their derivatives using Euler-Lagrange equation. These regularizations can all be unified as the quadratic terms of first and higher order of derivatives. The smooth step was generalized as the optimization in the Sobolev space. The gradient of the image similarity term and the gradient of the transform reg-
ularization term were connected through Sobolev gradient. Specifically we justified the Gaussian smoothing adopted in the demons approach as solving the general Tikhonov regularization. And we also explained its connection to the preconditioners used in Newton’s optimization methods.

After proposing this generalized common framework for all models, we focused on large deformation models. We introduced the diffeomorphism as a group structure. By introducing a time variable, one diffeomorphism was generated from velocity fields through the o.d.e. The optimization variable is not a transform at one time point, but rather the whole transform path from one image to the other.

Starting with LDDMM, we discussed the first form of parameterization as time-varying velocity fields. We showed that it could be viewed as solving optical flow at every time points. Also the composite demons algorithm could be viewed as a greedy version of optimizing the time-varying velocity fields.

LDDMM was derived rigorously from the general o.d.e form and the Sobolev gradient was explicitly introduced in the optimization. Beg also showed that its connection with metric on the diffeomorphism forms. Thus from many aspects LDDMM can be served an exemplar for large deformation methods.

The computation complexity of LDDMM is, however, rather high. It needs to solve $N$ times inverse transforms and also compute $N$ gradient fields for $N$ discretized time points. Ashburner proposed DARTEL to use a stationary velocity field to replace the time-varying ones. The DARTEL algorithm updated the one velocity field using the average of optical flows at all times points, which could be viewed as a simplified version of LDDMM.

The stationary velocity fields forms a one parameter subgroup of transformation. The transform is the exponential mapping of the velocity field. Ashburner used the efficient scaling-and-squaring method proposed by Arsigny to compute such an exponential mapping. He also discovered a recursive structure to compute the gradient as well. This reduced the
complexity from $N$ to $\log(N)$. But Ashburner also proposed to solve the optimization with Newton’s method in $L^2$ space, which was still expensive for computing the inverse of a large sparse matrix.

Vercauteren’s diffeomorphic demons and LogDemons further reduced the computation overhead by using only the demons force (or equivalently, the optical flow) at time point 0 as the gradient update. While this was only valid when the velocity is small, he demonstrated its effectiveness in practice. Diffeomorphic demons simply replaced the small deformation update in the compositive demons by its exponential mapping in each iteration. However, the overall transform did not have the form of the exponential mapping. This was improved in LogDemons algorithm using the BCH formula from the Lie group theory to recompute a new velocity field in every composition. Such a property made LogDemons closer to DARTEL algorithm. One problem for Vercauteren’s two methods is that they exploited certain properties of Lie group, which is still not fully understood mathematically.

In summarization we concluded that LDDMM used the most general parameterization but its optimization was the most expensive. It can be explained as computing optical flows at all time points. Diffeomorphic demons adopted a very heuristic parameterization but its optimization was very simple, just like the demons algorithm. DARTEL and LogDemons were two methods in the between. DARTEL was a direct reinterpretation of LDDMM using a stationary velocity field. LogDemons was the modified diffeomorphic demons with a stationary velocity field. A summary of comparison of parameterization and regularization for all the methods is listed in Table 3.1.
Table 3.1: Comparison of the parameterization and the regularization terms. For the operator $\mathcal{L}$, the regularization term is $\|\mathcal{L} \cdot \|^2$. Tikhonov regularization corresponds to Gaussian smoothing.

<table>
<thead>
<tr>
<th>Small Deformation Methods</th>
<th>Parameterization</th>
<th>Regularization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horn-Schunck optical flow [53]</td>
<td>$\phi = \text{Id } + u$</td>
<td>$\mathcal{L} = \nabla \phi$, $\mathcal{L} u$</td>
</tr>
<tr>
<td>Additive demons [101][111]</td>
<td>$\phi \leftarrow \phi + w$</td>
<td>Tikhonov</td>
</tr>
<tr>
<td>Compositive demons [101][111]</td>
<td>$\phi \leftarrow \phi \circ (\text{Id } + w)$</td>
<td>Tikhonov</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diffeomorphism Methods</th>
<th>Parameterization</th>
<th>Regularization</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDDMM [15]</td>
<td>$\frac{d}{dt} \phi = v(\phi, t)$</td>
<td>$\mathcal{L} = \text{Id } - \alpha \Delta$</td>
</tr>
<tr>
<td>DARTEL [7]</td>
<td>$\frac{d}{dt} \phi = v(\phi) \leftrightarrow \phi = \exp(v)$</td>
<td>$\mathcal{L} = \nabla \phi$ and others</td>
</tr>
<tr>
<td>Diffeomorphic demons [111]</td>
<td>$\phi \leftarrow \phi \circ \exp(w)$</td>
<td>Tikhonov</td>
</tr>
<tr>
<td>LogDemons [110]</td>
<td>$\phi = \exp(v)$</td>
<td>Tikhonov</td>
</tr>
</tbody>
</table>
CHAPTER 4

From Affine to Polyaffine Transformation Model

4.1 Introduction

The transform models applied in image registration have a wide span of degrees of freedom. The transform can be as simple as an affine transform [55], which is a linear function defined on the whole image domain and only requires 12 scalar parameters for an image of three dimensions. In contrast, one can also use a deformation field as the transform in a non-rigid image registration [102, 17], which has an arbitrary number of degrees of freedom at the cost of expensive computation and difficult optimization. Between these extremes of parameterization, many other transforms have been studied, such as B-Splines used in free-form deformation [84], Geodesic Interpolating Splines [20] and finite element method [35]. These transforms are capable of describing a wide range of non-rigid transforms while using fewer parameters than a dense displacement field.

The polyaffine transform is a parameterization for deformable maps that fills the gap between a global affine transform and a deformation field transform. It exploits the prior
knowledge that for many applications of medical image registration, the underlying anatomical structure is comprised of multiple local regions. An example of two lobes moving in different directions is illustrated in Fig. 4.1. Each local region is roughly rigid and can be approximated by a different local affine transform. The Polyaffine transform is a mathematical framework that fuses the local affine transforms through one non-rigid transform. Among different approaches like the piecewise affine transform [61, 83], a Log-Euclidean framework was first proposed in [5] especially to construct a smooth and invertible diffeomorphism, which can also be computed efficiently. This framework has been adopted in multiple image registration applications [5, 25, 99, 88, 89].

The key concept in [5] is to construct a stationary velocity field by fusing multiple regions with different affine velocity. The stationary velocity field is further integrated over time to generate an invertible transform. The construction method used in [5], however, as we are going to discuss in Section 4.2 could not guarantee that the resulting final transform gives the exact same value of the input transforms in each local region (Fig. 4.2). Our work reformulates the polyaffine transform as finding a feasible solution to a constrained problem. This new approach guarantees the local affine transformations are preserved. To achieve this, we demonstrate that the weight function used to fuse affine velocities has to be defined using a time-varying function in the framework of diffeomorphisms. The trajectory of each local affine region is proposed for computing the weight function in modeling a time-varying diffeomorphism with a series of stationary diffeomorphic transforms. With this new concept of region trajectory, our approach preserves each local transform and remains efficient in its implementation.

4.2 Methods

Consider a set of $K$ affine transforms, $T_i(x) = A_i x + b_i$, each defined on a local region $M_i$. We want to integrate these local transforms into one transform $\phi(x)$ defined on the whole image domain $\Omega$. For simplicity, these $K$ local regions are subregions of $\Omega$ and do
not overlap with each other. The polyaffine problem can be formulated as finding a feasible solution \( \phi(x) : \Omega \rightarrow \Omega \) satisfying the constraints

\[
\phi(x) = A_i x + b_i , \text{ when } x \in M_i
\]  

A direct way to construct such a \( \phi \) is by simply averaging each local affine transform using a weight function \( w_i \) as in [91]:

\[
\phi(x) = \sum_{i=1}^{K} w_i(x) T_i(x).
\]

A popular choice of weighting function is defined by the distance from the point \( x \) to each mask \( M_i \):

\[
w_i(x) \propto \exp\left(-\frac{\text{dist}(x,M_i)}{\sigma^2}\right).
\]

Then the \( K \) weights \( w_i(x) \) computed on the location \( x \) are further normalized such that \( \sum_i w_i(x) = 1 \).

Figure 4.1: Computed tomography images of two lung lobes moving in different directions. Top row: images before/after applying the transform. Bottom row: the two local affine transforms and the computed polyaffine transform using the proposed approach.

This transform is smooth in the sense of its deformation gradient. However, one significant drawback is that it is not invertible in general [5]. To obtain the invertibility, i.e. making \( \phi \) a diffeomorphism, the velocity field \( v(x,t) \) was introduced in constructing \( \phi \) in [5].
4.2.1 Diffeomorphism in Polyaffine Model

A diffeomorphism $\phi$ can be obtained by solving the ordinary differential equation over a time variable $t$:

$$\frac{d\phi(x,t)}{dt} = v(\phi(x,t), t)$$

(4.2)

At time $t = 1$, the diffeomorphism $\phi(x, 1)$ can be obtained by integrating the velocity field $v(x, t)$ over time from $t = 0$ to $1$. The affine transform $T(x) = Ax + b$ is an example of such a diffeomorphism. Using homogenous coordinates, an affine transform can be generated by a velocity field $v$ [5]:

$$v(x,t) = Lx + u$$

with

$$\begin{bmatrix} L & u \\ 0 & 0 \end{bmatrix} = \log \left( \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \right)$$

(4.3)

Arsigny et al. proposed in [5] to first construct a velocity field $v$ by fusing the $K$ local affine velocity fields:

$$v(x) = \sum_i w_i(x)v_i(x), \text{ with } v_i(x) = L_i x + u_i, w_i(x) \propto \exp \left( -\frac{\text{dist}(x, M_i)}{\sigma^2} \right)$$

(4.4)

It should be noted that here $v(x)$ is independent of $t$, i.e. $v(x) = v(x, t)$. Thus it is a stationary velocity field. To integrate $\phi$ from the velocity field $v$, a general approach is to discretize time from $t = 0$ to $1$ into $N$ time points. In the special case when $v$ is stationary, an efficient recursive scaling-and-squaring method was proposed in [5] to compute the total field in only $\log N$ steps.

4.2.2 Polyaffine Transform Preserving Trajectories

Introducing the velocity field $v$ as the generator for the transform guarantees the invertibility of the final transform. This, however, adds a new constraint in finding a feasible solution to Eqn. [4.1]. Instead of requiring the resulting transform $\phi(x)$ to match each local affine
Figure 4.2: Polyaffine transform using the previous approach [5]. (a) Two input affine transforms are shown in red and green. The starting and ending locations are indicated by the boxes in solid and dashed lines respectively. (b) The input local regions (shown as red and green areas) for each transform with their velocity plotted inside. (c) Computed velocity field using Eqn. 4.4 interpolated from the input local regions. (d) The resulting transform. The ending locations of each local region are plotted as boxes in solid lines at the arrow heads. The solid arrows (the result) deviate from the dashed arrows (the input), showing that local transforms are not preserved precisely.
transform $T_i(x)$ within $M_i$ at $t = 1$, it requires the velocity $v(x)$ to match each transform at all times from $t = 0$ to 1. Thus we can specialize the problem of Eqn.\[4.1\] into a more restricted one:

$$v(\phi(x,t),t) = v_i(\phi(x,t),t) \text{ for } x \in M_i$$ (4.5)

The solution of the new problem Eqn.\[4.5\] ensures that the whole temporal trajectory (from $t = 0$ to 1) for all $x$ in $M_i$ matches $T_i$. Indeed, given the uniqueness of the solution to the O.D.E of Eqn.\[4.2\], we have $\phi(x,t) = T_i(x,t)$ for $x \in M_i$ since $T_i(x) = A_i x + b_i$ is the solution when $v(\phi(x,t),t) = v_i(\phi(x,t),t)$. A natural solution of $\phi(x,t)$ in Eqn.\[4.5\] is to use a time-varying weight function to define the time-varying velocity field at any time $t$:

$$v(y,t) = \sum w_i(y,t)v_i(y) \text{, with } y = \phi(x,t) \text{ and}$$

$$w_i(y,t) \propto \exp \left( -\frac{\text{dist}(y,\phi(M_i,t))}{\sigma^2} \right)$$ (4.6)

The weight $w_i(y,t)$ needs to be defined over $t$ by tracking each $M_i$ at time $t$, $\phi(M_i,t) = \{\phi(x,t), x \in M_i\}$. If a point $y = \phi(x,t)$ belongs to the $i$-th transform, the weight $w_i$ should have:

$$w_i(y,t) = 1 \text{ and } w_{j \neq i}(y,t) = 0 \text{ for } x \in M_i.$$ (4.7)

This new definition of velocity field is different from the one used in [5], i.e. Eqn.\[4.4\] where it only matches $v_i$ at time $t = 0$. The new definition of the weight $w_i$ depends on $t$ and needs to track the trajectory of the local mask $M_i$.

The definition in Eqn.\[4.4\] [5] could not guarantee that the velocity $v(\phi(x,t),t)$ is still dominated by $T_i$ when a point $x \in M_i$ at $t = 0$ moves to a new location $\phi(x,t)$ at time $t$. Thus it could not preserve the trajectory of each local affine region. This makes it an inviable solution to Eqn.\[4.5\] (illustrated in Fig.\[4.2\]).
4.2.3 Extend Local Region to Trajectory of Local Region

Although Eqn. 4.6 gives a feasible solution to Eqn. 4.5, it also eliminates the nice property of stationary velocity and is inefficient in computation. We propose a novel way to define a stationary velocity field \( v \) which still satisfies Eqn. 4.5. Define the trajectory of the region \( M_i \) from time \( t_1 \) to \( t_2 \) as notation \( M^*_i|_{t_1}^{t_2} = \cup_{\tau=t_1}^{t_2} \phi(M_i, \tau) \).

Without loss of generality, we first assume that these region trajectories do not overlap in the spatial domain \( \Omega \). Define a new stationary weighting function \( w(x) \) independent of \( t \):

\[
v(y) = \sum w_i(y)v_i(y) \\
\text{with} \ w_i(y) \propto \text{dist}(y, M^*_i|_{0}^{1})
\]

Since for \( i \neq j \), \( M^*_i|_{0}^{1} \cap M^*_j|_{0}^{1} = \emptyset \), for \( x \in M_i \) we have \( \text{dist}(\phi(x, t), M^*_i|_{0}^{1}) = 0 \) and \( \text{dist}(\phi(x, t), M^*_j|_{i \neq j,0}^{1}) \gg 0 \). Thus the new stationary \( w(x) \) still satisfies the same property of the time-varying version of \( w_i(x, t) \) in Eqn. 4.7: \( w_i(T_i(x, t)) = 1 \) and \( w_{i \neq j}(T_i(x, t)) = 0 \) for \( x \in M_i \). By introducing the region trajectory \( M^*_i|_{0}^{1} \), the proposed stationary weight \( w_i \) gives the same \( \phi \) as the time-varying version in Eqn. 4.6. The only difference between the two resulting \( \phi \) is in the intermediate areas outside of \( \cup_{i=1}^{K} M_i \).

When comparing the proposed stationary weight in Eqn. 4.8 and its generalized time-varying version in Eqn. 4.6, one should notice that \( w(y, t) \) in Eqn. 4.6 is defined on the spatial-temporal domain \( \Omega \times [0,1] \), while \( w(y) \) in Eqn. 4.8 is a "squeezed version" that collapses the region trajectory \( M^*_i|_{0}^{1} \) along the temporal axis into the spatial domain. The result using the proposed weight for the case in Fig. 4.2 is shown in Fig. 4.3 where our solution clearly preserves the input affine transform.

4.2.4 Series of Stationary Velocity Field for Trajectory Collision

One critical assumption in eliminating \( t \) from Eqn. 4.6 is that all trajectories \( M^*_i|_{0}^{1} \) do not overlap in the spatial domain, \( M^*_i|_{0}^{1} \cap M^*_j|_{0}^{1} = \emptyset \). Otherwise it will be ambiguous to
Figure 4.3: Polyaffine transform using our proposed method. (a) Two input affine transforms (same as in Fig. 4.2(a)) are shown in red and green. The starting and ending locations are indicated by the boxes in solid and dashed lines respectively. (b) The trajectories of the input local regions (shown as red and green areas) for each transform with their velocity plotted inside. (c) Computed velocity field using Eqn. 4.8 interpolated from the input region trajectories. (d) The resulting transform. The ending locations of each local region are plotted as boxes in solid lines at the arrow heads. The solid arrows (the result) are overlapped with the dashed arrows (the input), showing that local transforms are preserved precisely.
determine which local affine transform should be used when a point $y$ belongs to multiple trajectories.

This non-collision assumption is nevertheless not true in general. Consider two local affine transforms $T_1$ and $T_2$ in Fig. 4.4 defined on region $M_1$ and $M_2$ respectively. The end of trajectory of $M_1$, $T_1(M_1)$, is overlapped with the starting position of $M_2$. In this case, for point $y$ inside $T_1(M_1) \cup M_2$, it is ambiguous to define its weight $w_i(y)$ using Eqn. 4.8.

![Figure 4.4: The trajectories of two local regions overlap. Each arrow represents an input local transform, plotted in black and red. The boxes at the arrow tails are the starting locations of each input transform; the boxes at the heads are the ending locations.](image)

Our solution to this dilemma is to find a period from time $t_1$ to $t_2$ so that the trajectories within this period are not overlapped. By the time $t = 1$ when $M_1$ moves to $T_1(M_1)$, $M_2$ also moves to $T_2(M_2)$ and $T_1(M_1) \cap T_2(M_2) = \emptyset$. In general we need at any time $\tau_1, \tau_2 \in [t_1, t_2]$, no local regions are overlapped, $\phi(M_i, \tau_1) \cap \phi(M_j, \tau_2) = \emptyset$. When these trajectories are disjoint in the spatial-temporal domain, such $[t_1, t_2]$ is feasible. Thus it is possible to break the time from 0 to 1 into a sequence of $C + 1$ time points $[t_0, \ldots, t_C]$, such that

$$t_0 = 0 \text{ and } t_{k+1} = \max \{\tau | M_i^\tau |_{t_k} \cap M_j^\tau |_{t_k} = \emptyset, \forall i \neq j\} \quad (4.9)$$

For each non-collision period $[t_{k-1}, t_k]$, a stationary velocity $v^k$ and its associated weight
function are defined in the same way as in Eqn. 4.8. For each stationary velocity \( v^k \), its transform \( \phi^k \) is computed using the efficient scaling-and-squaring method in [5]. The final transform \( \phi \) is the concatenation of these \( C \) diffeomorphism transforms and is also a diffeomorphism: \( \phi = \phi^C \circ \cdots \circ \phi^2 \circ \phi^1 \).

### 4.3 Implementation

We have proposed the notion of region trajectory to construct the polyaffine model. Unlike previous models, our polyaffine transform guarantees that the final transform is the exact value of \( T_i(x) \) in each individual local region \( M_i \). Before we discuss some of our implementation details, here is the summary of the construction steps:

1. Compute the collision time points \( \{ t_0 = 0, t_1, \ldots, t_C = 1 \} \) using Eqn. 4.9.
2. Construct the stationary velocity \( v^k \) for each collision period \( t = t_{k-1} \) to \( t_k \).
3. Compute the corresponding diffeomorphism \( \phi^k \) from \( v^k \).
4. Concatenate all \( \phi^k \) to get the final diffeomorphism \( \phi \).

#### 4.3.1 Choice of Weight Function

A common practice to define the weight function \( w_i(x) \) is to use the distance between the point \( x \) to the region \( M_i \) [5], or its trajectory \( M_i^* \) in our case. The weight is computed using a Gaussian function [5], or other decreasing functions [25]. The weights are further normalized by scale so that \( \sum_i w_i(x) = 1 \).

There are two drawbacks of using directly \( \text{dist}(x, M_i^*) \). The first is a practical issue. For \( x \) in \( M_i^* \), its distance to the \( j \)-th transform region trajectory cannot be infinite, and \( w_{j \neq i}(x) \) is in general a small, nonzero scalar after normalization. Some arbitrary thresholding needs to be performed to approximate \( w_{j \neq i}(x) = 0 \). Finding a suitable \( \sigma \) in the Gaussian function for different regions is also arbitrary.
Figure 4.5: Comparison of the transforms computed using (a,c,e) dist\((x, M_i^*)\), and (b,d,f) dist\((x, \partial M_i^*)\). (a,b) The transforms are constructed from the same input, one horizontal translation (green arrow) and one vertical translation (red arrow). (b) is visually more smooth than (a). (c,d) The weights \(w(x)\) for the local vertical translation, range in \([0, 1]\). (d) is more polarized except along boundary \(\partial M_i^*\). (e,f) The resulting velocity below the green trajectory has an upward vertical component in (e) but not in (f). Similar effects are seen on the right of the red trajectory. \(M_i^*\) are shown as green and red dotted boundaries in (c,d) and solid areas in (e,f).
The second drawback is that such a distance does not consider the relative spatial configuration of different region trajectories. For example, in Fig. 4.5(b), there are two local transforms. One is a horizontal translation (on left), and the other is a vertical translation (on right). By examining the vertical component of the velocity field below the horizontal trajectory, it can be seen that there is some upward vertical velocity component, which comes from the small weights on the right vertical translation. Ideally, there should be no such vertical velocity around this region since the vertical velocity should be blocked by the horizontal trajectory on its left.

To eliminate this drawback, we propose a new weighting scheme by dividing the image into subdomains. The subdomains \( \{ M^*_i \} \) are a partition of the whole image domain \( \Omega \). The subdomain \( M^*_i \) can be viewed as the region of influence of the trajectory \( M^*_i \) in the center. Based on this intuition, we use the distance to each trajectory to define the subdomain \( M^*_i \) as the area containing all the points closer to \( M^*_i \) than other trajectories:

\[
M^*_i = \{ x \in \Omega | \text{dist}(x, M^*_i) = \min_{j=1...K} \text{dist}(x, M^*_j) \} \tag{4.10}
\]

Such a definition partitions the whole image domain \( \Omega \) into \( K \) disjoint subdomains. The proposed weight is correspondingly defined by the distance from the point \( x \) to the boundary of the subdomain \( M^*_i \). If \( x \) is close to the boundary, the weights should be nonzero for all adjacent local transforms. If \( x \) is in the center and far away from the boundary, its weight should be polarized. Note that \( d(x) = \text{dist}(x, \partial M^*_i) \), and \( \sigma \) is a constant. We can define \( w_i \) before normalization using a simple piecewise linear function.

\[
w_i(x) \propto \begin{cases} 
1 & \text{if } x \in M^*_i \text{ and } d(x) > \sigma \\
0.5 + d(x)/2\sigma & \text{if } x \in M^*_i \text{ and } 0 \leq d(x) < \sigma \\
0.5 - d(x)/2\sigma & \text{if } x \notin M^*_i \text{ and } 0 \leq d(x) < \sigma \\
0 & \text{if } x \notin M^*_i \text{ and } d(x) > \sigma 
\end{cases} \tag{4.11}
\]
An example of the computed weight function and the resulting transforms on two local affine transforms is illustrated in Fig. 4.5. Such a weighting scheme takes consideration of the relative position of each trajectory mask. Along the boundaries of the subdomains are the locations where the weighting matters most, while inside each subdomain it is dominated by one local transform. This can be viewed as an efficient approximate solution to the following equation:

\[ w_i^* = \min \int \| \nabla w_i \|^2 dx, \text{where } w_i(x) = 1 \text{ for } x \in M_i \text{ and } w_i(x) = 0 \text{ for } x \in M_{j \neq i} \quad (4.12) \]

In one extreme case when each local region is one point and uniformly distributed, such a partition becomes the Voronoi diagram on the image domain if time \( t = 0 \) is only considered.

### 4.3.2 Computing Mask Trajectories and Their Collision

Another important implementation detail is how to efficiently compute each region trajectory \( M_i^* \) given its predefined velocity \( v_i(x) = L_i x + u_i \). This is trivial when \( M_i \) is just a single point. However, when \( M_i \) is a region, there could be multiple trajectories. When a region can be efficiently parameterized as a polyhedron, one can track the trajectory of each vertex and compute the collision of any two polyhedra. Here we instead choose a simpler approximation by using a random point set to represent an arbitrary region.

For a region \( M \), a point set \( \{p_a\} \) is uniformly sampled inside using a sampling diameter \( d \). The region can be then approximated by the union of dilations from the point set \( \{p_a\} \) with distance \( d \). The region \( M \) transformed at time \( t \), \( T(M,t) \), is also approximated by the dilation of the point set \( \{T(p_a,t)\} \). The condition of collision detection in Eqn. 4.9 is implemented using:

\[ M_i^*|_{t_1}^{t_2} \cap M_j^*|_{t_1}^{t_2} = \emptyset \iff \min_{\tau_i = t_1}^{t_2} \min_{\tau_j = t_1}^{t_2} \min_{a,b} \text{dist}(T_i(p_{i,a}, \tau_i), T_j(p_{j,b}, \tau_j)) < d \quad (4.13) \]

One interesting fact about the distance between the two point trajectories of a local affine
region is that it is well bounded by a convex function of time variable $t$ when the region deforms over time. For simplicity without using extra notations in the homogenous coordinates, suppose $L$ is the logarithm of the affine matrix $A$ (see Eqn. 4.3), we have $T(x, t) = \exp(tL)x$. We have the following theorem.

**Theorem 1.** $f(t) = \|\exp(tL)(p - q)\|$ is bounded by a convex function for a given matrix $L$ and any two points $p$ and $q$.

**Proof 1.**

$$\|\exp(tL)(p - q)\| \leq \| \sum_{n=0}^{\infty} \frac{(tL)^n}{n!} \|p - q\| \leq \sum_{n=0}^{\infty} \frac{(t\|L\|)^n}{n!} \|p - q\| = \exp(t\|L\|)\|p - q\|$$

This bound ensures that if a sampling diameter $d$ is small enough, the dilation of the point set $\{T(p, t)\}$ at time $t$ should cover the mask $T(M, t)$ and thus is a good approximation.

### 4.4 Results

We evaluated the accuracy of our proposed approach using synthetic experiments. Two translation transforms were defined on two rectangular regions, similar to the inputs in Fig. 4.5. A scalar $r \in [0, 1]$ was used to control the relative range of the translation offset, $t_1 = r \times [100, 0]$ and $t_2 = r \times [0, 50]$. When $r = 0$, they were identity transforms; when $r$ increased, the two affine trajectories overlap. The accuracy was evaluated by the difference between the resulting transform and the corresponding affine transform for each pixel in all regions, $e(x) = \|\phi(x) - (A_ix + t_i)\|$. The mean and variance of all $e(x)$ for $x \in M_1 \cup M_2$ were computed versus $r$.

We compared our proposed polyaffine construction using the region trajectory with the approach in [5] as the baseline, where the weight was computed based on the region only. When the translation was relatively small ($r < 0.3$), both methods had a good accuracy. However, when the translation increased, the error in the baseline method also increased quickly since the region trajectories became significantly different from the regions. In
contrast, our approach still maintained a high accuracy with near-zero error. Similar results could be observed when comparing the case of two rotation transforms by changing the rotation angles (Fig. ??(b)).

4.5 Discussion

In this work we presented a novel approach for constructing a polyaffine transform which can precisely preserve each affine transform using one diffeomorphism. The polyaffine problem is formulated as finding a feasible solution with a new constraint that preserves the affine trajectories of each local region. The natural solution uses a time-varying weight function, which is time and memory consuming in implementation. Our approach instead uses a composition of one or more diffeomorphisms of stationary velocities and is thus both efficient in computation and accurate in preserving the new constraints.

The key in constructing our weight function is to use the trajectory of each local region, instead of the region itself, to define the fusion weight for each local transform. In the case when only the region is used in defining the weight, the affine velocity is only preserved at time $t = 0$. This may lead to inaccuracy in meeting the constraints. In contrast, our
new approach preserves the affine velocity for any time $t$, and therefore preserves the affine transform.

Depending on the values of input affine transform and their regions, our approach requires one or more stationary velocity fields to preserve different affine velocities in each region over time. The number of stationary velocity fields are determined by the collision detection of region trajectories when collapsing them along the temporal direction. One real-life analogy to such a scheme is the traffic lights in an intersection. When two trajectories are overlapped at the intersection, a traffic light will indicate when the points along one direction should stop to allow the points along the other to pass the intersection, which solves the ambiguity in defining the velocity function at the intersection. Note that depending on the input affine transforms, there might be no collision and just one stationary velocity is enough.
CHAPTER 5

Diffeomorphic Registration of a Pair of Lung CT Images

5.1 Introduction

To accurately quantify both global and regional assessments of lung biomechanics from image data, an automatic non-rigid image registration algorithm plays an important role. The image registration computes a transformation mapping between two configurations of the lung at inspiration and expiration, from which measures of local deformation can be calculated and compared within and across subjects.

Various lung image registration algorithms have been proposed in the literature. A small sample includes optical flow ([13, 29]), compressed optical flow ([21]), Laplacian-filtered image ([31, 32]), free-form deformation ([85, 116]). A more detailed review was presented in Chapter 3. In this chapter we are interested in diffeomorphic transformations which, by definition, preserve topology. Topology preservation is fundamental to making comparisons between objects in the natural world as such transformations permit comparisons to be made across time points in an individual’s disease process or to study development patterns.
across a large population. This fits naturally for the study of pulmonary kinematics. Some recent work on applying diffeomorphic transformation models on lung image registration includes [26, 37].

In this chapter we apply two different diffeomorphic models discussed before to the data sets provided for the EMPIRE10 lung registration evaluation [56], using our open-source toolbox ANTS (Advanced Normalization Tools, [12]). Built on the ITK (Insight Segmentation and Registration Toolkit [75]) framework, this software package comprises a suite of tools for image normalization and template building based on previously published research. The results indicate that ANTS provides a competitive set of registration tools for intra-subject lung CT image registration.

5.2 Methods

In general, image registration is the process of finding the optimal transformation, $\phi$, within a specified transformation space which maps each $x$ of image $I(x)$ to a location in image $J(z)$ such that a specified cost function, $C$, defined on $I$, $J$ and $\phi$, is minimized.

One typical setting of lung CT image registration for pulmonary kinematics consists of two chest CT images at different phases of a breathing cycle of the same subject (i.e. the scenario in EMPIRE10 [72]). Following a commonly adopted two-step strategy, our registration pipeline begins with an affine transformation for initial global alignment, which precedes a deformable diffeomorphic transformation with increased degrees of freedom.

5.2.1 Affine Transformation

The affine transformation is optimized with respect to translation, rotation, scaling and shearing. The successive optimization of each component allows for individual control over increasing degrees of freedom.

The affine transform is represented as a projection matrix $A$ and a translation vector $t$: 

\[ A \cdot x + t = z \]
\[ T(x) = Ax + t. \] The projection matrix is decomposed into the product of the rotation matrix \( R \), the scaling matrix \( S \) and the shearing matrix \( K \), \( A = RSK \).

Given the rotation axis of \((u, v, w)\) and the rotation angle \( \theta \), the rotation is parameterized by the quaternion of four-parameters \( R_T = (a, b, c, d) \in \mathbb{R}^4 \) with \((a, b, c, d) = (\cos \frac{\theta}{2}, u \sin \frac{\theta}{2}, v \sin \frac{\theta}{2}, w \sin \frac{\theta}{2})\). The unitary constraint \( \|R_T\|_2 = 1 \) keeps the scale unchanged under rotation. The rotation matrix \( R \) is given by:

\[
\begin{pmatrix}
 a^2 + b^2 - c^2 - d^2 & 2bc - 2ad & 2ac + 2bd \\
 2ad + 2bc & a^2 - b^2 + c^2 - d^2 & 2cd - 2ab \\
 2bd - 2ac & 2ab + 2cd & a^2 - b^2 - c^2 + d^2
\end{pmatrix}.
\]

The given lung masks of each image pair were the input to the affine registration. For these binary masks, mutual information\(^{[112, 67]}\) was used as the similarity function. More implementation details about the affine registration can be found in \(^{[93]}\).

### 5.2.2 Diffeomorphic Transformations

In contrast to many other transformation models which reside in the domain of vector spaces, a diffeomorphism is a differentiable mapping with a differentiable inverse. Modeling transformations with diffeomorphisms ensures certain unique and desirable topological properties.

We assume that the diffeomorphism, \( \phi \), is defined on the image domain, \( \Omega \), and maintains an affine transform at the boundary such that \( \phi(\partial\Omega) = A(\text{Id}) \) where \( A(\text{Id}) \) is an affine mapping applied to the identity transformation. \( \phi \), over time, parameterizes a family of diffeomorphisms, \( \phi(x, t) : \Omega \times t \to \Omega \), which can be generated by integrating a time-dependent, smooth velocity field, \( \mathbf{v} : \Omega \times t \to \mathbb{R}^d \), through the ordinary differential equation
The deformation field yielded by $\phi$ is $u(x) = \phi(x,1) - x$.

The following minimizing variational form was proposed for optimization in diffeomorphic normalization for inexact image matching in [34, 71, 16]:

$$v^* = \arg\min_v \left\{ \int_0^1 ||Lv||^2 dt + \lambda \int_\Omega ||I \circ \phi(x,1) - J||^2 d\Omega \right\}. \tag{5.2}$$

The first term on the right represents a mathematical metric between $I$ and $J$ given an appropriate norm, $L$, on the velocity field, $v$. The second term is the image similarity metric of square intensity difference with weight $\lambda$ accounting for the inexact matching.

To accommodate a variety of medical image normalization tasks, one typically encounters more complex intensity transfers between one anatomical instance $J$ and another instance $I$. This leads to the generalization of Equation (5.2):

$$v^* = \arg\min_v \left\{ \int_0^1 ||Lv||^2 dt + \lambda \int_\Omega \Pi_\sim(I, \phi(x,1), J)d\Omega \right\}. \tag{5.3}$$

where $\Pi_\sim$ is a similarity metric depending on the images and the mapping and $\lambda$ controls the degree of exactness in the matching.

**Symmetric Normalization**

Exploiting the fact that the diffeomorphism, $\phi$, can be decomposed into two components $\phi_1$ and $\phi_2$, [10] constructs a symmetric alternative to Equation (5.3). This leads to the

(o.d.e.)

$$\frac{d\phi(x,t)}{dt} = v(\phi(x,t), t), \quad \phi(x,0) = x. \tag{5.1}$$
symmetric variant of Equation (5.3)

\[ \{v_1^*, v_2^*\} = \arg\min_{v_1, v_2} \left\{ \int_0^{0.5} (|L v_1(x, t)|^2 + |L v_2(x, t)|^2) \, dt + \lambda \int_\Omega (I \circ \phi_1(x, 0.5), J \circ \phi_2(x, 0.5)) \, d\Omega \right\}. \tag{5.4} \]

The corresponding symmetric Euler-Lagrange equations are similar to [71]. Finding \(v_1^*\) minimizes the variational energy from \(t = 0\) whereas \(v_2^*\) minimizes from \(t = 1\). Thus, gradient-based iterative convergence deforms \(I\) and \(J\) along the geodesic diffeomorphism, \(\phi\), to a fixed point midway between \(I\) and \(J\) thus motivating the denotation of the solution strategy as Symmetric Normalization (SyN).

In practice to reduce the significant computational and memory cost of the dense-in-time gradient calculations and requisite reintegration of the diffeomorphisms after each iterative update, we offer a greedy variant as a lower-cost alternative:

\[ \nabla \Pi = \frac{\partial}{\partial \phi_i} \Pi_\sim(I(\phi_1^{-1}(x, 0.5)), J(\phi_2^{-1}(x, 0.5))) \tag{5.5} \]

for \(i \in \{1, 2\}\). \(\phi_1(x, 0.5)\) and \(\phi_2(x, 0.5)\) are then updated from the previous iteration according to

\[ \phi_i(x, 0.5) = \phi_i(x, 0.5) + \delta(K \ast \nabla \Pi(\phi_i(x, 0.5))). \tag{5.6} \]

where \(\delta\) is a user-specified step parameter. \(K\) is the Green kernel of \(L\), which is approximated by a Gaussian kernel ([18]). The gradient is then mapped back to the origin of each diffeomorphism to update the full mapping by enforcing \(\phi^{-1}(\phi(x, 1)) = x\).

**Exponential Mapping**

Ashburner introduced DARTEL (Diffeomorphic Anatomical Registration using Exponentiated Lie algebra) as a rapidly computed alternative to time parameterized diffeomorphic
schemes ([7]). The key difference between a time-varying diffeomorphism and a diffeomorphism generated by an exponential mapping ([7]) is that the exponential mapping maintains only a single vector field that is constant in time.

By exponentiation of a constant velocity field, a diffeomorphism can be generated from the following o.d.e (cf Equation (5.1))

\[
\frac{d\phi(x,t)}{dt} = v(\phi(x,t)), \quad \phi(x,0) = x.
\]  

(5.7)

Note that there is no explicit time parameter in the velocity field. Theoretically, restricting the velocity field to be constant in time reduces the size of the space that may be generated ([4]).

5.2.3 Cross Correlation Similarity Metric

Previous research uses optical flow ([43, 29]) or its variant ([21]), which implicitly assumes the intensity consistency between two image. However, for the lung imaging, the image pair are usually from two breathing phases, in which the local density changes are linearly reflected in the intensity changes. In this case, the invariance of cross correlation to the linear intensity change makes it a suitable similarity function. Furthermore we compute cross correlation in a neighborhood around each voxel to accommodate the inhomogeneity of the density changes throughout the whole lung. The local cross correlation is integrated over the lung volume as the overall similarity in the diffeomorphic transformation. One may write the (squared) cross-correlation for the diffeomorphic image registration as:

\[
\Pi_\sim(\mathcal{I}, \mathcal{J}) = \text{CC}(x) = \frac{\sum_i (\mathcal{I}(x_i) - \mu_I(x)) (\mathcal{J}(x_i) - \mu_J(x))^2}{\sum_i (\mathcal{I}(x_i) - \mu_I(x))^2 \sum_i (\mathcal{J}(x_i) - \mu_J(x))^2},
\]

(5.8)

where \(x\) is at the center of \(N^3\) square window, \(\mu\) is the mean value within the window centered at \(x\) and \(x_i\) iterates through that window.
5.2.4 Optimization Strategy

The affine registration of the binary lung masks was used to initialize the transform. We composed the affine transformation with the deformable transformation field before performing any interpolation or downsampling. In this way, the image warping never required more than a single image interpolation step and was able to always refer back to the original full-resolution image sources.

The greedy Symmetric Normalization and the exponential mapping were then used separately for the following diffeomorphic deformable registrations. Both the affine and the two diffeomorphic image registrations were computed using our open source toolbox ANTS ([12]). For the diffeomorphic registration, the intensity values outside the lung masks were set to zeroes. The intensity values inside the lung masks were normalized to the range of $[0, 1]$ by the linear adjustment: $\frac{\max(I) - I}{\max(I) - \min(I)}$.

For both the affine and diffeomorphic registration, the gradient descent was used in the optimization. Convergence during the optimization is achieved if the slope of linear regression of the cost values of last 12 iterations is close to zero. A multi-resolution approach was applied in both steps to accelerate computation speed and avoid trapped in local minimum.

5.3 Results

The dataset of 20 pairs of chest CT scans was provided in the EMPIRE10 competition ([56]), as part of MICCAI 2010 Grand Challenges. Each pair of scans was acquired from a single subject, in other words no inter-subject registrations were included. The scans came from a variety of sources and are provided by several different institutions. Scans might be acquired at various phases in the breathing cycle (full inspiration, full expiration, phase from 4D breathing data). Subjects might either have lung diseases or be healthy. Data from a variety of scanners were included and a variety of voxel sizes occur. In addition to the CT data, binary lung masks were provided for each scan. The lung masks were
Table 5.1: Results using greedy SyN for each scan pair, per category and overall. Rankings and final placement are from a total of 34 competing algorithms. The total running time for each pair is shown on the last column.

<table>
<thead>
<tr>
<th>Scan Pair</th>
<th>Lung Boundaries Score</th>
<th>Rank</th>
<th>Fissures Score</th>
<th>Rank</th>
<th>Landmarks Score</th>
<th>Rank</th>
<th>Singularities Score</th>
<th>Rank</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>0.00</td>
<td>6.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.18</td>
<td>2.00</td>
<td>0.00</td>
<td>11.50</td>
<td>6h 28m</td>
</tr>
<tr>
<td>02</td>
<td>0.00</td>
<td>11.00</td>
<td>0.00</td>
<td>15.00</td>
<td>0.34</td>
<td>3.00</td>
<td>0.00</td>
<td>12.50</td>
<td>8h 13m</td>
</tr>
<tr>
<td>03</td>
<td>0.00</td>
<td>5.50</td>
<td>0.00</td>
<td>12.50</td>
<td>0.29</td>
<td>1.00</td>
<td>0.00</td>
<td>12.00</td>
<td>7h 1m</td>
</tr>
<tr>
<td>04</td>
<td>0.00</td>
<td>2.50</td>
<td>0.00</td>
<td>16.50</td>
<td>0.84</td>
<td>4.00</td>
<td>0.00</td>
<td>14.00</td>
<td>7h 52m</td>
</tr>
<tr>
<td>05</td>
<td>0.00</td>
<td>13.00</td>
<td>0.00</td>
<td>16.00</td>
<td>0.00</td>
<td>5.50</td>
<td>0.00</td>
<td>13.50</td>
<td>3h 44m</td>
</tr>
<tr>
<td>06</td>
<td>0.00</td>
<td>16.00</td>
<td>0.00</td>
<td>15.00</td>
<td>0.30</td>
<td>7.00</td>
<td>0.00</td>
<td>14.00</td>
<td>2h 40m</td>
</tr>
<tr>
<td>07</td>
<td>0.00</td>
<td>1.50</td>
<td>0.08</td>
<td>2.00</td>
<td>1.14</td>
<td>1.00</td>
<td>0.00</td>
<td>10.00</td>
<td>7h 47m</td>
</tr>
<tr>
<td>08</td>
<td>0.00</td>
<td>4.00</td>
<td>0.00</td>
<td>3.50</td>
<td>0.59</td>
<td>3.00</td>
<td>0.00</td>
<td>12.50</td>
<td>7h 40m</td>
</tr>
<tr>
<td>09</td>
<td>0.00</td>
<td>2.00</td>
<td>0.00</td>
<td>6.50</td>
<td>0.50</td>
<td>2.00</td>
<td>0.00</td>
<td>13.00</td>
<td>7h 7m</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>1.50</td>
<td>0.00</td>
<td>15.00</td>
<td>1.55</td>
<td>12.00</td>
<td>0.00</td>
<td>13.50</td>
<td>8h 26m</td>
</tr>
<tr>
<td>11</td>
<td>0.01</td>
<td>13.00</td>
<td>0.00</td>
<td>4.00</td>
<td>0.65</td>
<td>4.00</td>
<td>0.00</td>
<td>11.50</td>
<td>6h 1m</td>
</tr>
<tr>
<td>12</td>
<td>0.00</td>
<td>10.00</td>
<td>0.00</td>
<td>13.50</td>
<td>0.00</td>
<td>1.50</td>
<td>0.00</td>
<td>14.50</td>
<td>6h 32m</td>
</tr>
<tr>
<td>13</td>
<td>0.00</td>
<td>3.00</td>
<td>0.07</td>
<td>7.00</td>
<td>0.76</td>
<td>3.00</td>
<td>0.00</td>
<td>13.00</td>
<td>2h 6m</td>
</tr>
<tr>
<td>14</td>
<td>0.00</td>
<td>2.00</td>
<td>0.21</td>
<td>2.00</td>
<td>0.80</td>
<td>1.00</td>
<td>0.00</td>
<td>9.50</td>
<td>8h 11m</td>
</tr>
<tr>
<td>15</td>
<td>0.00</td>
<td>8.00</td>
<td>0.00</td>
<td>7.00</td>
<td>0.58</td>
<td>1.00</td>
<td>0.00</td>
<td>12.50</td>
<td>5h 9m</td>
</tr>
<tr>
<td>16</td>
<td>0.00</td>
<td>3.50</td>
<td>0.01</td>
<td>10.00</td>
<td>0.93</td>
<td>5.00</td>
<td>0.00</td>
<td>13.50</td>
<td>55m</td>
</tr>
<tr>
<td>17</td>
<td>0.00</td>
<td>6.50</td>
<td>0.07</td>
<td>30.00</td>
<td>0.64</td>
<td>3.00</td>
<td>0.00</td>
<td>14.00</td>
<td>1h 22m</td>
</tr>
<tr>
<td>18</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.94</td>
<td>1.00</td>
<td>0.00</td>
<td>21.00</td>
<td>12h 41m</td>
</tr>
<tr>
<td>19</td>
<td>0.00</td>
<td>14.00</td>
<td>0.00</td>
<td>12.00</td>
<td>0.45</td>
<td>3.00</td>
<td>0.00</td>
<td>14.50</td>
<td>7h 9m</td>
</tr>
<tr>
<td>20</td>
<td>0.00</td>
<td>3.50</td>
<td>0.14</td>
<td>1.00</td>
<td>0.83</td>
<td>1.00</td>
<td>0.00</td>
<td>10.50</td>
<td>7h 42m</td>
</tr>
</tbody>
</table>

| Avg       | 0.00                  | 6.37 | 0.03           | 9.52 | 0.66           | 3.20 | 0.00                | 13.05 |       |

Average Ranking Overall: 8.03
Final Placement: 1
automatically generated using an algorithm by van Rikxoort et al. [107]. They visually checked and manually corrected the masks if necessary.

Table 5.2: Results using exponential mapping for each scan pair, per category and overall. Rankings and final placement are from a total of 34 competing algorithms. The total running time for each pair is shown on the last column.

<table>
<thead>
<tr>
<th>Scan Pair</th>
<th>Lung Boundaries</th>
<th></th>
<th>Fissures</th>
<th></th>
<th>Landmarks</th>
<th></th>
<th>Singularities</th>
<th></th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Score</td>
<td>Rank</td>
<td>Score</td>
<td>Rank</td>
<td>Score</td>
<td>Rank</td>
<td>Score</td>
<td>Rank</td>
<td>Time</td>
</tr>
<tr>
<td>01</td>
<td>0.00</td>
<td>2.50</td>
<td>0.21</td>
<td>14.00</td>
<td>2.46</td>
<td>13.00</td>
<td>0.01</td>
<td>27.00</td>
<td>8h 27m</td>
</tr>
<tr>
<td>02</td>
<td>0.00</td>
<td>11.00</td>
<td>0.00</td>
<td>15.00</td>
<td>0.44</td>
<td>13.00</td>
<td>0.00</td>
<td>12.50</td>
<td>10h 3m</td>
</tr>
<tr>
<td>03</td>
<td>0.00</td>
<td>5.50</td>
<td>0.00</td>
<td>12.50</td>
<td>0.41</td>
<td>14.00</td>
<td>0.00</td>
<td>12.00</td>
<td>8h 23m</td>
</tr>
<tr>
<td>04</td>
<td>0.00</td>
<td>2.50</td>
<td>0.00</td>
<td>16.50</td>
<td>1.59</td>
<td>20.00</td>
<td>0.00</td>
<td>28.00</td>
<td>10h 11m</td>
</tr>
<tr>
<td>05</td>
<td>0.00</td>
<td>13.00</td>
<td>0.00</td>
<td>16.00</td>
<td>0.00</td>
<td>5.50</td>
<td>0.00</td>
<td>13.50</td>
<td>6h 54m</td>
</tr>
<tr>
<td>06</td>
<td>0.00</td>
<td>16.00</td>
<td>0.00</td>
<td>2.00</td>
<td>0.37</td>
<td>16.00</td>
<td>0.00</td>
<td>14.00</td>
<td>4h 7m</td>
</tr>
<tr>
<td>07</td>
<td>0.00</td>
<td>6.00</td>
<td>1.44</td>
<td>18.00</td>
<td>2.82</td>
<td>15.00</td>
<td>0.00</td>
<td>25.00</td>
<td>8h 53m</td>
</tr>
<tr>
<td>08</td>
<td>0.00</td>
<td>4.00</td>
<td>0.04</td>
<td>15.00</td>
<td>1.16</td>
<td>17.00</td>
<td>0.00</td>
<td>26.00</td>
<td>8h 31m</td>
</tr>
<tr>
<td>09</td>
<td>0.00</td>
<td>6.00</td>
<td>0.00</td>
<td>20.00</td>
<td>0.59</td>
<td>14.00</td>
<td>0.00</td>
<td>13.00</td>
<td>8h 5m</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>1.50</td>
<td>0.00</td>
<td>15.00</td>
<td>3.00</td>
<td>20.00</td>
<td>0.03</td>
<td>28.00</td>
<td>9h 16m</td>
</tr>
<tr>
<td>11</td>
<td>0.04</td>
<td>16.00</td>
<td>0.09</td>
<td>16.00</td>
<td>1.15</td>
<td>13.00</td>
<td>0.00</td>
<td>25.00</td>
<td>8h 14m</td>
</tr>
<tr>
<td>12</td>
<td>0.00</td>
<td>10.00</td>
<td>0.02</td>
<td>28.00</td>
<td>0.27</td>
<td>20.00</td>
<td>0.00</td>
<td>14.50</td>
<td>5h 14m</td>
</tr>
<tr>
<td>13</td>
<td>0.00</td>
<td>3.00</td>
<td>0.06</td>
<td>2.50</td>
<td>1.08</td>
<td>18.00</td>
<td>0.00</td>
<td>13.00</td>
<td>2h 41m</td>
</tr>
<tr>
<td>14</td>
<td>0.00</td>
<td>3.00</td>
<td>5.78</td>
<td>20.00</td>
<td>2.35</td>
<td>10.00</td>
<td>0.01</td>
<td>26.00</td>
<td>9h 13m</td>
</tr>
<tr>
<td>15</td>
<td>0.00</td>
<td>8.00</td>
<td>0.00</td>
<td>7.00</td>
<td>0.64</td>
<td>13.00</td>
<td>0.00</td>
<td>12.50</td>
<td>7h 56m</td>
</tr>
<tr>
<td>16</td>
<td>0.00</td>
<td>3.50</td>
<td>0.01</td>
<td>8.50</td>
<td>1.38</td>
<td>20.00</td>
<td>0.00</td>
<td>13.50</td>
<td>1h 8m</td>
</tr>
<tr>
<td>17</td>
<td>0.00</td>
<td>6.50</td>
<td>0.05</td>
<td>16.00</td>
<td>0.98</td>
<td>16.00</td>
<td>0.00</td>
<td>14.00</td>
<td>2h 2m</td>
</tr>
<tr>
<td>18</td>
<td>0.00</td>
<td>4.00</td>
<td>0.90</td>
<td>9.00</td>
<td>2.93</td>
<td>16.00</td>
<td>0.00</td>
<td>24.00</td>
<td>7h 45m</td>
</tr>
<tr>
<td>19</td>
<td>0.00</td>
<td>14.00</td>
<td>0.00</td>
<td>26.00</td>
<td>0.50</td>
<td>13.00</td>
<td>0.00</td>
<td>29.00</td>
<td>9h 32m</td>
</tr>
<tr>
<td>20</td>
<td>0.00</td>
<td>3.50</td>
<td>1.99</td>
<td>10.00</td>
<td>1.70</td>
<td>10.00</td>
<td>0.00</td>
<td>25.00</td>
<td>8h 52m</td>
</tr>
<tr>
<td>Avg</td>
<td>0.00</td>
<td>6.97</td>
<td>0.53</td>
<td>14.35</td>
<td>1.29</td>
<td>14.82</td>
<td>0.00</td>
<td>19.77</td>
<td></td>
</tr>
</tbody>
</table>

Average Ranking Overall 13.98
Final Ranking 11

We reported results of two methods in this study. Both used the same affine registration on the lung masks to initialize the deformation. In the step of diffeomorphic registration, the first method used greedy SyN (noted as \textit{gsyn}) and the second used exponential mapping (noted as \textit{exp}). The registration process was fully automatic. The same set of parameters was used across all image pairs. The affine registration applied a four-level image pyramid and \(10^4\) locations were sampled in computing the mutual information of the lung binary masks. A five-level image pyramid was applied in both \textit{gsyn} and \textit{exp}. The neighborhood radius in computing cross correlation was 2 voxels. The Gaussian regularization kernel
The deformation fields were evaluated in four categories: lung boundaries, fissure alignments, labeled landmarks and singularities in the deformation (Fig. 5.1). More details about the evaluation and ranking protocol can be found in [72]. Table 5.1 shows the evaluation results of our greedy Symmetric Normalization model $gsyn$. Table 5.2 shows the corresponding results for the exponential mapping model. Computation was performed on a Linux workstation of Intel Xeon QuadCore CPU at 3GHz with 16 GB main memory.

![Figure 5.1: The four categories evaluated in the EMPIRE10 challenge: the lung boundaries, fissures, labeled landmarks and singularities in the deformation field, from [72].](image)

Out of the 34 submitted algorithms, our $gsyn$ got the first placement in the overall ranking and $exp$ got the 11th. Greedy SyN outperformed exponential mapping, which shows the benefit using time-dependent velocity field over constant-in-time velocity field in this application. In general, greedy SyN gave a satisfactory registration result. Most fissures and lung boundaries were aligned well with error close to zero. There are almost no singularities in the deformation fields due to the theoretical properties of diffeomorphic transformation models.

The average displacement of the manually labeled landmarks was within one voxel for most pairs after registration. See Fig. 5.2 for the examples of the registration. An example of bad result is illustrated in Fig.5.3 where the vessels in the upper half of the image were not aligned.
Figure 5.2: Examples of good registration results, from two subjects: (a) the fix subject, (b) the moving subject, (c) the warped moving subject.

Figure 5.3: Example of bad registration shown in enlarged image region. The vessels and the fissure in the upper region do not align well. The remaining lower region, however, has a good registration.
5.4 Discussion

In this study we applied two diffeomorphic transformation models to the data provided by the EMPIRE10 challenge. Cross correlation was used as the similarity function to compensate the intensity change in the lung volumes at different breathing phases. By carefully choosing the parameters, e.g. the cross correlation window size and the Gaussian regularization kernel width, for the specific application, our general image registration algorithm achieved a competitive result among other state-of-the-art methods, with the greedy Symmetric Normalization ranked No.1 and the exponential mapping ranked No.11 in the final outcome involving a total of 34 algorithms.

The whole registration pipeline was built on using our open-source toolbox ANTS. We should note that as a general purpose image registration software, ANTS has demonstrated success in other applications [8, 57]. Besides the image preprocessing, the programming of registration pipeline involved simple scripting, which makes the results reproducible. We believe that such an open source toolbox could benefit the research community and promote the idea of reproducible research, as described by Dr. Kovacevic ([58]), which “refers to the idea that, in "'computational’ sciences, the ultimate product is not a published paper but, rather, the entire environment used to produce the results in the study (data, software,etc.).”

The accuracy of current registration results can be further improved. As shown in Fig.5.3, some blood vessels and fissures were not aligned precisely. More domain specific knowledge, like the positions of the vessels and fissures from lung anatomy segmentation, should be utilized to improve upon current general registration techniques. The running time cost for our diffeomorphic approaches is very high (see last column in Table 5.1 and 5.2). The running time is linear to the volume size multiplied by the number of iterations. This can be improved by exploring possibilities of multi-threading computation and GPU technology in the future implementation. Another worthy direction is to use quantized displacement to form the registration problem as a discrete energy minimization problem, which can
be solved efficiently by minimizing the Markov Random Field ([39] [94]). The discretized deformation field can then be further optimized by other diffeomorphic registration method like ANTS.
CHAPTER 6

Feature Selection for
Characterization of ILD and COPD

6.1 Introduction

Based on clinical, imaging, and pathological characteristics, most types of chronic lung disease can be grouped into two basic categories: interstitial lung disease (ILD) and chronic obstructive pulmonary disease (COPD). ILD is a heterogeneous group of diseases in which the hallmark is chronic, progressive, predominantly interstitial inflammation with varying degrees of fibrosis of the lung parenchyma, often leading to reduced lung volume, decreased lung compliance, and restrictive physiology. COPD is characterized by chronic airflow limitation due to airway inflammation and lung parenchymal destruction which is not fully reversible and is usually progressive and obstructive.

The diagnosis, differentiation, and classification of the severity of ILD and COPD rely on clinical assessment, thoracic imaging (using CT and chest radiography), and pulmonary
function testing (PFT). PFT is a noninvasive method of assessing the integrated mechanical function of the lung, chest wall, and respiratory muscles. It currently comprises the reference standard for pulmonary functional assessment. Using PFT, the heterogeneous group of ILD typically exhibits a restrictive physiology pattern whereas COPD typically manifests an obstructive physiology pattern. PFT strictly permits a global assessment of lung physiology. In contrast, HRCT image analysis is a powerful tool with the potential for regional as well as global quantification of pulmonary diseases.

With the increasing amount of proposed image metrics, research inquiry concerns the effectiveness of these metrics. Previous research ([120, 27, 95, 66, 100, 113, 80, 92]) has been conducted to investigate the correlation between various quantitative image metrics and different PFT values.

Instead of performing a correlation study, we address the question of the effectiveness of image metrics primarily from the view of feature selection. We put more emphasis on what image metrics and what PFT values can characterize ILD and COPD in a quantitative framework. Image metrics and PFT values are viewed as candidates for selecting which of these best characterize the corresponding diagnosis (i.e., ILD or COPD). We are interested in whether and what features from image metrics provide additional information for differential diagnosis when PFT values are provided. This also differs from the classification work of [118] in the sense that we do not tend to train any classifiers directly. The selected features can be used as inputs for any available classifier. We use Support Vector Machine (SVM) [108] as an example to test the efficiency of the feature selection results. One advantage of SVM is that the training is not affected by the order of input features, which is not considered in some previous work like [92].

A minimal-redundancy-maximal-relevance (mRMR) information framework was introduced in [82] for such a feature selection task. The ideal selected features satisfy two constraints: maximal relevance and minimal redundancy. The relevance of both image and PFT features concerns the ability of such features in matching an existing classification (in our case, from
clinical and radiologic diagnosis). It is usually computed in terms of mutual information, correlation, or statistical tests. However, in order to get a compact subset of features to classify different types of disease, it is not enough to consider only the features with highest relevance. The selected features need to be as independent from each other as possible. This is known as the criterion of minimal redundancy, which makes the features more compact for certain diseases, in comparison with previous studies ([120, 27, 95, 66, 100, 113, 80, 92]) focusing on the relevance (more specifically, correlation) between all the available image features and PFT features.

In this study, we provide a systematic relevancy/redundancy analysis comparing 31 various statistical image metrics and 21 PFT values obtained in patients with diagnosed ILD and COPD. The analysis framework is described in Section 2. In Section 3, we provide the results of our analysis in comparing the characterization performance of both image and PFT features. This is followed by discussion of the results and possible clinical applicability of our approach.

### 6.2 Materials and Methods

Table 6.1: Statistics of ILD patients.

<table>
<thead>
<tr>
<th>ILD Patient: Total 14</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>8</td>
</tr>
<tr>
<td>Male</td>
<td>6</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Mean Age</td>
<td>56.1</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>12.0</td>
</tr>
<tr>
<td>Median Age</td>
<td>56</td>
</tr>
<tr>
<td>Degree of PFT</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>3</td>
</tr>
<tr>
<td>Mild</td>
<td>3</td>
</tr>
<tr>
<td>Moderate</td>
<td>4</td>
</tr>
<tr>
<td>Severe</td>
<td>2</td>
</tr>
<tr>
<td>Moderately Severe</td>
<td>1</td>
</tr>
<tr>
<td>Very Severe</td>
<td>1</td>
</tr>
</tbody>
</table>

Fourteen patients with ILD and 11 patients with COPD were retrospectively identified. The demographics of these selected patients are provided in Table 6.1. Every pa-
Table 6.2: Statistics of COPD patients.

<table>
<thead>
<tr>
<th>COPD Patient: Total 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Degree of PFT Obstruction</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

![Image](image_url)

Figure 6.1: Pipeline of the proposed feature selection analysis.

Patient underwent both thoracic HRCT image acquisition and PFT within 3 days of each other. HRCT was performed for both inspiration and expiration on a 64 multi-detector row CT scanner (Siemens Medical Solutions, Malvern, PA) with reconstruction of contiguous 1 mm axial images with a 1mm interval and a Siemens B41f kernel. Inspiratory and expiratory image datasets were then analyzed through computational software developed in our laboratory which is capable of generating several hundred distinct metrics encompassing various aspects of lung physiology (e.g., which include pulmonary volumetric and gross tissue indices, attenuation histogram statistics, deformation indices, co-occurrence and run-length matrix texture indices, and attenuation mask indices) gleaned from the relevant literature. For this study, we only focused on a subset of these metrics.

Our whole image-processing pipeline (Fig. 6.1) is fully automatic. First, a segmentation algorithm is applied to extract lung regions and segment the trachea and central bronchi from CT images. Next, we compute 31 different types of image metrics (Table 6.4) on the
Table 6.3: All 21 PFT values (1st and 3rd rows) and corresponding relevance values (2nd and 4th rows) to the disease types.

<table>
<thead>
<tr>
<th>PFT</th>
<th>FVC</th>
<th>FEV1</th>
<th>FEV1/FVC</th>
<th>FEF max</th>
<th>FEF 25-75%</th>
<th>FEV50</th>
<th>FIF50</th>
<th>MVV</th>
<th>SVC</th>
<th>IC</th>
<th>ERV</th>
<th>Rel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.088</td>
<td>0.026</td>
<td>0.37</td>
<td>0.357</td>
<td>0.218</td>
<td>0.307</td>
<td>0.021</td>
<td>0.251</td>
<td>0.197</td>
<td>0.083</td>
<td>0.146</td>
<td>0.083</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PFT</th>
<th>TGV</th>
<th>RV (Pleth)</th>
<th>TLC (Pleth)</th>
<th>RV/TLC</th>
<th>DLCO unc</th>
<th>DLCO cor</th>
<th>DL/VA</th>
<th>VA</th>
<th>Raw</th>
<th>sGaw</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.268</td>
<td>0.268</td>
<td>0.35</td>
<td>0.218</td>
<td>0.041</td>
<td>0.15</td>
<td>0.066</td>
<td>0.269</td>
<td>0.101</td>
<td>0.074</td>
</tr>
</tbody>
</table>

DLCOcor, corrected carbon monoxide diffusion; DLCOunc, uncorrected carbon monoxide diffusion; DL/VA, carbon monoxide diffusion corrected or alveolar volume; ERV, expiratory reserve volume; FEF25-75%, forced expiratory flow at 25% and 75% of forced vital capacity; FEV50, forced expiratory flow at 50% of forced vital capacity; FIF50, forced expiratory flow at 50% of forced vital capacity; IC, inspiratory capacity; FVC, forced vital capacity; max, maximum; MVV, maximum voluntary ventilation; PFT, pulmonary function testing; Raw, airway resistance; RV (Pleth), residual volume on plethysmography; sGaw, specific airway conductance; SVC, slow vital capacity; TGV, thoracic gas volume; TLC (Pleth), total lung capacity on plethysmography; VA, alveolar volume.

6.2.1 PFT Parameters and Image Metrics

Table 6.3 lists all 21 PFT values used in the study. These parameters were either directly obtained or calculated for each patient, using standard protocols following guidelines of the American Thoracic Society. As previously mentioned, 93 image metrics (Table 6.4) were generated for each patient. We use G1 to denote the 31 features from the expiratory images, G2 for those from the inspiratory images, and G3 for the difference of the expiratory/inspiratory metric pairs. This subset of image metrics includes those first-order statistical measures generated from the attenuation histogram, as well as more sophisticated expiratory and inspiratory image datasets separately. We also subtract expiratory metrics from their inspiratory counterparts to obtain additional 31 metrics. This translates into a total of 93 image metrics for each patient. We then compute the relevance of each PFT value and each image metric to the corresponding disease type. The minimal-redundancy-maximal-relevance (mRMR) criterion [82] is applied to select a mixed subset of PFT values and image metrics to show the degree of redundancy between them. Finally, the selected features are further fed into classifiers to evaluate the relationship between the number of selected features and the accuracy of ILD/COPD classification, which reflects the efficiency of the feature selection procedure.
cated second-order statistical quantities related to the texture of lung parenchyma, i.e., the co-occurrence [44] and run-length [28] matrix texture indices. Putting the 21 PFT values and the 93 computed image metrics together, there are 114 different numerical scalars for each patient, which will be cited as the 114 features in the study.

### 6.2.2 Feature Selection

The primary motivation for this study is to determine how image metrics perform relative to PFT values in matching to the clinical diagnosis of ILD and COPD. We also select an optimal subset of mixed image metrics and PFT values under the principle of minimal-redundancy-maximal-relevance.

In accordance with the specifications of the algorithm in [82], each of the $N = 25$ patients is given a label $c$ based on the clinical diagnosis, either ILD or COPD. Such a disease type $c$ is a random variable. Also, each patient has a set of $J = 93$ computed image metrics $X = \{x_1, \ldots, x_J\}$ and a set of $K = 21$ measured PFT values $Y = \{y_1, \ldots, y_K\}$. Let the mixed feature set $Z = X \cup Y$. Each element $z_i \in Z$ is also viewed as a random variable.

Peng et al [82] proposed to use mutual information to measure the relevance between a feature $z_i$ and the disease type $c$: $I(z_i, c) = \int \int p(z_i, c) \log \frac{p(z_i, c)}{p(z_i)p(c)} dz_i dc$.

The relevance measure $I(z_i, c)$ shows the amount of information that knowing the feature $z_i$ provides to classify the disease type as $c$. This relevance value is always nonnegative. If two variables $z_i$ and $c$ are independent, which means that $z_i$ provides no information in predicting $c$, their relevance is 0. If the value of $z_i$ is totally determined by $c$, $z_i$ and $c$ are not independent and their relevance becomes the entropy of $c$, which is a measure of the uncertainty associated with $c$. Also, the maximal relevance is bounded by the greater of the entropies of $z_i$ and $c$. Note that since the uniform distribution gives the maximum entropy as $\log(2)$ for all possible binary random variables $c$, the upper bound for $I(z_i, c)$ is $\log(2) = 0.69$. 
Table 6.4: Computed image metrics and their relevance to the disease types. $G_1$ denotes the metrics computed from the expiratory images; $G_2$ from the inspiratory images. Metrics of $G_3$ are generated by subtracting $G_1$ from $G_2$. Within each of $G_1/G_2/G_3$, the metrics are indexed from 1 to 31. The numbers in last three columns are the relevance values to the disease types of ILD and COPD.

<table>
<thead>
<tr>
<th>No.</th>
<th>Metrics Type</th>
<th>$G_1$</th>
<th>$G_2$</th>
<th>$G_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>lung region volume</td>
<td>0.325</td>
<td>0.276</td>
<td>0.073</td>
</tr>
<tr>
<td>2</td>
<td>relative volume $&lt; -910$</td>
<td>0.439</td>
<td>0.454</td>
<td>0.071</td>
</tr>
<tr>
<td>3</td>
<td>attenuation mean $\leq 38$</td>
<td>0.416</td>
<td>0.407</td>
<td>0.181</td>
</tr>
<tr>
<td>4</td>
<td>attenuation variance $\geq 98$</td>
<td>0.171</td>
<td>0.197</td>
<td>0.181</td>
</tr>
<tr>
<td>5</td>
<td>sum $\leq 38$</td>
<td>0.295</td>
<td>0.353</td>
<td>0.005</td>
</tr>
<tr>
<td>6</td>
<td>attenuation skewness $\leq 118$</td>
<td>0.276</td>
<td>0.159</td>
<td>0.026</td>
</tr>
<tr>
<td>7</td>
<td>attenuation kurtosis $\leq 118$</td>
<td>0.392</td>
<td>0.201</td>
<td>0.079</td>
</tr>
<tr>
<td>8</td>
<td>attenuation grey level entropy $\leq 118$</td>
<td>0.021</td>
<td>0.121</td>
<td>0.005</td>
</tr>
<tr>
<td>9</td>
<td>5% attenuation value $\leq 74$</td>
<td>0.47</td>
<td>0.527</td>
<td>0.353</td>
</tr>
<tr>
<td>10</td>
<td>95% attenuation value $\leq 74$</td>
<td>0.463</td>
<td>0.22</td>
<td>0.006</td>
</tr>
<tr>
<td>11</td>
<td>5% attenuation mean $\leq 74$</td>
<td>0.488</td>
<td>0.416</td>
<td>0.036</td>
</tr>
<tr>
<td>12</td>
<td>95% attenuation mean $\leq 74$</td>
<td>0.055</td>
<td>0.065</td>
<td>0.071</td>
</tr>
<tr>
<td>13</td>
<td>energy $\leq 118$ $\leq 44$</td>
<td>0.019</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>14</td>
<td>entropy $\leq 118$ $\leq 44$</td>
<td>0.159</td>
<td>0.215</td>
<td>0.159</td>
</tr>
<tr>
<td>15</td>
<td>correlation $\leq 118$ $\leq 44$</td>
<td>0.003</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>16</td>
<td>inverse difference moment $\leq 118$ $\leq 44$</td>
<td>0.021</td>
<td>0.121</td>
<td>0.025</td>
</tr>
<tr>
<td>17</td>
<td>inertia $\leq 44$</td>
<td>0.003</td>
<td>0.022</td>
<td>0.005</td>
</tr>
<tr>
<td>18</td>
<td>cluster shade $\leq 44$</td>
<td>0.083</td>
<td>0.021</td>
<td>0.074</td>
</tr>
<tr>
<td>19</td>
<td>cluster prominence $\leq 44$</td>
<td>0.113</td>
<td>0.074</td>
<td>0.104</td>
</tr>
<tr>
<td>20</td>
<td>Haralick’s correlation $\leq 44$</td>
<td>0.034</td>
<td>0.034</td>
<td>0.001</td>
</tr>
<tr>
<td>21</td>
<td>short run emphasis $\leq 118$</td>
<td>0.049</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>22</td>
<td>long run emphasis $\leq 118$</td>
<td>0.074</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>23</td>
<td>grey level non-uniformity $\leq 118$</td>
<td>0.243</td>
<td>0.316</td>
<td>0.007</td>
</tr>
<tr>
<td>24</td>
<td>run-length non-uniformity $\leq 118$</td>
<td>0.325</td>
<td>0.353</td>
<td>0.006</td>
</tr>
<tr>
<td>25</td>
<td>run percentage $\leq 118$</td>
<td>0.157</td>
<td>0.101</td>
<td>0.101</td>
</tr>
<tr>
<td>26</td>
<td>low grey level run emphasis $\leq 23$</td>
<td>0.285</td>
<td>0.157</td>
<td>0.113</td>
</tr>
<tr>
<td>27</td>
<td>high grey level run emphasis $\leq 23$</td>
<td>0.005</td>
<td>0.114</td>
<td>0.005</td>
</tr>
<tr>
<td>28</td>
<td>short run low grey level emphasis $\leq 28$</td>
<td>0.285</td>
<td>0.157</td>
<td>0.058</td>
</tr>
<tr>
<td>29</td>
<td>short run high grey level emphasis $\leq 28$</td>
<td>0.005</td>
<td>0.114</td>
<td>0.005</td>
</tr>
<tr>
<td>30</td>
<td>long run low grey level emphasis $\leq 28$</td>
<td>0.285</td>
<td>0.157</td>
<td>0.113</td>
</tr>
<tr>
<td>31</td>
<td>long run high grey level emphasis $\leq 28$</td>
<td>0.005</td>
<td>0.114</td>
<td>0.005</td>
</tr>
</tbody>
</table>
Ranking all the features in \( Z \) by their relevance values \( I(z_i, c) \) gives an order of all features. The first ranked \( k = 20 \) features are used as inputs to a classifier to distinguish ILD from COPD. This strategy is denoted as maximum relevance (MaxRel) and is the base feature selection strategy.

MaxRel strategy does not consider the redundancy between the selected features. For a fixed number of selected features, such redundancy prevents the selected feature subset from covering the distributions of all the feature candidates. To remove duplicated features while preserving the maximal relevance, the mRMR (minimal-redundancy-maximal-relevance) framework proposed in \cite{82} selects a subset of \( S \) from \( Z \) such that:

\[
S = \arg \max_{S \subseteq Z} \left\{ \sum_{z_i \in S} I(z_i, c) - \frac{1}{\|S\|} \sum_{z_i, z_j \in S} I(z_i, z_j) \right\}
\]  

(6.1)

The first term in Eq. 6.1 maximizes the total relevance of the selected features with the corresponding disease types; the second term minimizes the total redundancy of all pairs of the selected features. The framework in \cite{82} gives a heuristic way to optimize Eq.6.1. We use the online toolbox \cite{81} to compute the relevance weights and select features.

Given the diagnosis \( c \), we analyze the relevance of each feature, which indicates its individual characterization power. Then we use mRMR to select a subset of the first \( k = 20 \) features from the total of 114 features.

If the subset would include only PFT values, this indicates that image metrics would be redundant for differentiating ILD and COPD; otherwise, this indicates that image metrics provide extra information that PFT values do not possess for differentiating ILD and COPD.

### 6.2.3 Classification with SVMs

To evaluate quantitatively their effectiveness in classifying ILD and COPD, the selected features are used to train a binary pattern classifier. An efficient feature selection scheme
should achieve low classification error rate with minimal number of features. We studied the relationship between classification error and the number of selected features. Also, we compared classification using only PFT values to using image metrics as additional features.

In this study we use Support Vector Machines (SVM) as classifiers. SVMs have been used widely for feature classification tasks. Suppose that each patient has \( k \) types of selected features, mapping to a data point in a \( k \)-dimensional space. SVM takes a data point as input and the class label \( c \) is determined by comparing its output as either positive or negative.

Two different types of SVM are investigated in this study. The first is the linear SVM using a linear plane to separate two classes of points in the \( k \)-dimensional space. More formally, if the features for one patient is a vector \( \mathbf{x} = (x_1, \ldots, x_k) \), the linear SVM classifier is

\[
\mathbf{w} \cdot \mathbf{x} - b \begin{cases} 
\geq 0 & \text{for the positive class;} \\
< 0 & \text{for the negative class.}
\end{cases}
\]

The parameters \( \mathbf{w} \) and \( b \) are learned from the training samples to maximize the margins between hyperplanes that separates data points of two classes. An example of a linear SVM using 2 features is illustrated in 2D as in Fig. 6.2a. In this 2D case, the two features for each data point are its x and y coordinates and the separating plane is a straight line. Note that the linear SVM is unable to classify all the data points correctly using a straight line.

Instead of using the dot product in the above equation, SVM can apply other kernels and data points are implicitly mapped to higher dimensional spaces so that better classification can be achieved. More theoretical discussion about SVMs can be found in [87]. In this study we use the radial basis function (RBF) SVM, which is a popular choice of one nonlinear SVM for many other classification tasks.

Fig. 6.2b shows an example of using RBF SVM on the same data points, the hyperplane of which is a curve instead of a straight line to classify all the data points. By using a nonlinear kernel, RBF SVM usually can classify data of more complicated distributions than linear
Figure 6.2: Examples of SVMs on a synthetic 2D data set. Each data point has its x and y coordinates as two features. Type I points are marked with red dots; type II with blue crosses. SVM hyperplanes are shown in green. Note that Type I points cannot be separated from type II points by any straight lines. (a) The linear kernel SVM misclassifies several points, marked with circles. (b) The RBF kernel SVM classifies all the points with a curve as classification hyperplane.

SVMs. We compare these two SVMs to understand whether the performance improvement is from the choice of nonlinear kernel or from the more efficient features.

We use SVMs as classifiers to evaluate the efficiency of the selected feature subset whose size increases from $k = 1$ to 20 ranked by $mRMR$ or $MaxRel$. Both linear SVM and RBF SVMs are tested. The linear SVM is the basic format of SVM and is helpful in investigating the efficiency of the selected features and isolating the influence of choice of kernels. The RBF SVM is a popular choice for general classification tasks [87]. All features are normalized by dividing them by their maximum absolute values.

We apply the leave-one-out strategy as cross-validation in training SVM for classification. In each test, one subject is excluded when training SVMs and that subject is used as the test set. The error rates are computed as the average ratio of wrongly classified subject by repeating using each of the patients as the test subject.
6.2.4 Experiment setup

First, the relevance values were computed between each PFT value and image metric to the disease type. Next, to investigate if the image metrics are redundant to the PFT values in classifying ILD and COPD, three groups of feature selection were tested. The first selected only PFT measurements. The second selected only image metrics. The third selected from both groups. We evaluated two feature selection strategies, \textit{MaxRel} and \textit{mRMR}, with both the linear SVM and the RBF SVM. By varying different groups of features, different feature selection strategies, and different SVM models, there were \(3 \times 2 \times 2 = 12\) experiments in total.

In each feature selection experiment, \(k = 1\) to 20 different features were selected from the corresponding feature group (PFT values, image metrics, or both). The \(k\) features were then fed to the SVM. Leave-one-out strategy was used in training and testing SVM. The classification rates were collected against different \(k\).

6.3 Results

![Figure 6.3: Relevance (mutual information) of image metrics to different disease types. The x axis has the image metrics index listed in Table 6.4. Red bars are for metrics from \(G1\), the expiratory images; blue bars for \(G2\), the inspiratory images; and green bars for \(G3\), difference of \(G1\) to \(G2\).](image)

Figure 6.3: Relevance (mutual information) of image metrics to different disease types. The x axis has the image metrics index listed in Table 6.4. Red bars are for metrics from \(G1\), the expiratory images; blue bars for \(G2\), the inspiratory images; and green bars for \(G3\), difference of \(G1\) to \(G2\).

We first compared the relevance of each feature to the diagnosis of disease. Fig. 6.3
graphically depicted the relevance weights of all the image metrics. The relevance weights of PFT values were plotted in Fig. 6.4 for comparison. For example, the ratio of the forced expiratory volume in 1 second to the forced vital capacity (\( \text{FEV1/FVC} \)) had a high relevance value (0.47), concordant with its clinical application as most important distinguishing PFT feature.

The relevance values from image metrics and PFT values were in a similar range. The top PFT parameter, \( \text{FEV1/FVC} \), was 0.47; the top image metric, 5% attenuation value from the inspiratory dataset, was 0.52. This showed that some image metrics were as good as or even better discriminators than PFT values for characterization of ILD and COPD.

![Figure 6.4: Relevance (mutual information) of PFT values to different disease types. The x axis lists all the 21 PFT values in Table 6.3. Note that the range of the y axis is similar to the range from those image metrics in Fig. 6.3](image)

The feature selection results were grouped by SVM models in Fig. 6.5a and 6.5b. The blue, cyan, and red curves corresponded to the three different feature groups: PFT values, image metrics, and both. The red curve (\( mRMR \) on both) was below the blue one (\( mRMR \) on PFT), indicating the reduced error rate for the combination of PFT values and image metrics. For example, in case of the linear SVM using 6 features, \( mRMR \) using both PFT and image metrics achieved 2.08% error rate, whereas \( mRMR \) using only PFT values achieved 4.32%.

In comparison of different SVMs, the RBF SVM generally achieved a lower error rate than
Figure 6.5: Classifying ILD and COPD using the first $n = 1$ to 20 selected features with (a) linear SVM and (b) RBF SVM, using MaxRel or mRMR on PFT values, image metrics or both PFT and image metrics. For each curve, the $x$ axis is the number of selected features for training and testing with SVMs; the $y$ axis is the average error rate from leave-one-out tests. One subject is excluded when learning SVM parameters and that subject is used to test the accuracy of the trained SVM.
the linear SVM using the same number of features. For example, using four features, the RBF SVM achieved an error rate of 1.92% compared to the error rate of the linear SVM as 5.12%, using mRMR to select both image metrics and PFT values. The RBF SVM needed six features from mRMR on both to get the lowest error rate whereas the linear SVM needed eight features.

For the feature selection strategy, mRMR demonstrated consistently better performance than MaxRel in both the linear SVM and the RBF SVM. For example, using first 7 features with linear SVM, mRMR achieved a 1.44% error rate using the first 7 features, whereas MaxRel achieved an error rate of 4.16%. When the number of selected features increased beyond 12, the error rates of all strategies converged. This was true for all three feature groups (PFT, image, both) with linear SVM and two groups (image, both) with RBF SVM.

A subset of the first 20 features using mRMR on both PFT and image metrics were selected by averaging the ranks of each feature from all leave-one-out experiments, as listed in table 6.5. Both PFT values and image metrics were selected, 14 image metrics and 6 PFT measurements. Also, image features computed both from the first-order statistical measurements (such as attenuation value) and from the texture descriptions (such as grey level run emphasis) were selected in the same subset.

6.4 Discussion

Various image metrics have been proposed and used in the literature to differentiate ILD and COPD. In this study, we studied their relevance values to the corresponding diagnoses in comparison with those of PFT values. One notices that the two most relevant (i.e., the most powerful in characterizing and distinguishing the two diagnoses) PFT values are FEV1/FVC and TGV (thoracic gas volume), which reflect the degree of obstruction and total lung volume, concordant with clinical diagnostic criteria. A more interesting result is that these PFT values rank behind the best image metrics (as shown in Table 6.5). FEV1/FVC was
Table 6.5: The first 20 selected features from mRMR on both PFT values and image metrics. Note these features include both PFT values and image metrics. The image metrics includes both the first order statistical measurements obtainable from the attenuation histogram and other more sophisticated metrics of texture descriptions. $G_1$: from the expiratory volumes, $G_2$: from the inspiratory volumes, $G_3$: subtracting $G_1$ from $G_2$ (same as in Table 6.4).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Feature</th>
<th>Rank</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$G_1$: attenuation mean</td>
<td>11</td>
<td>$G_3$: mean</td>
</tr>
<tr>
<td>2</td>
<td>$G_1$: lung region volume</td>
<td>12</td>
<td>FEF 25-75%</td>
</tr>
<tr>
<td>3</td>
<td>$G_1$: sum</td>
<td>13</td>
<td>RV (pleth)</td>
</tr>
<tr>
<td>4</td>
<td>FEV1/FVC</td>
<td>14</td>
<td>$G_2$: mean</td>
</tr>
<tr>
<td>5</td>
<td>$G_1$: run length nonuniformity</td>
<td>15</td>
<td>$G_2$: run length nonuniformity</td>
</tr>
<tr>
<td>6</td>
<td>TGV</td>
<td>16</td>
<td>FEF1</td>
</tr>
<tr>
<td>7</td>
<td>$G_1$: 95% attenuation value</td>
<td>17</td>
<td>$G_1$: 5% attenuation mean</td>
</tr>
<tr>
<td>8</td>
<td>FEF max</td>
<td>18</td>
<td>$G_2$: 95% attenuation mean</td>
</tr>
<tr>
<td>9</td>
<td>FEF50</td>
<td>19</td>
<td>$G_2$: sum</td>
</tr>
<tr>
<td>10</td>
<td>$G_1$: relative volume $&lt; -950$</td>
<td>20</td>
<td>$G_2$: relative volume $&lt; -950$</td>
</tr>
</tbody>
</table>

ranked fourth and thoracic gas volume ranked sixth. Image metrics such as attenuation mean and lung region volume (ranked first and second, respectively) demonstrated higher relevance, indicating that image metrics can be even more powerful than the reference standard PFT values for characterization of ILD versus COPD.

Moreover, the information derived from image metrics cannot be obtained solely from PFT values, as indicated by the low degree of redundancy of the selected 20 features and also from the observation that classification error rate decreased by adding image metrics to the PFT values to distinguish ILD and COPD. Fig. 6.6 shows the comparison of the receiver operating characteristic (ROC) curves using different feature sets. The green curve is the ROC using only FEV1/FVC value as threshold value to separate ILD and COPD. The red curve shows using the RBF SVM score computed from the first 10 selected features (from mMRM). Clearly, more features can get a lower false positive rate while maintaining a higher true positive rate. This suggests that image features provide non-redundant information to which was provided by PFT values in classifying ILD and COPD.

Two types of SVMs were applied in this study to classify ILD from COPD. The RBF SVM generally demonstrated better classification than the linear SVM, which suggests that adopting the nonlinear RBF kernel will generally result in improved classification.
Figure 6.6: Receiver operating characteristic curves for using different values to separate ILD and COPD. The green curve uses FEV1/FVC value. The red curve uses the RBF SVM scores computed from the first 10 features selected by mRMR. SVMs are trained and tested using the leave-one-out strategy.

performance. To get a clear assessment of how image metrics improved classification, we trained the simpler linear SVM, which only used a linear hyperplane to separate data, such that the classification rate would be mainly affected by the efficiency of selected features. Comparing the error rates when only image metrics were used (the cyan curves in Fig.6.5) to the ones when only PFT values were used (the blue curves in Fig.6.5), the former clearly showed better performance in the case of linear SVM (Fig.6.5a).

There are limitations of the current study. First, the sample size is relatively small compared to the number of existing features. It would be beneficial to apply the same strategy to a large-scale database. However, our methodology of feature selection and the use of SVM as a classifier are invariant to the sample size and feature number.

Second, there was no spirometric control of patient expiratory effort, although patients were provided with breathing instructions at the time of imaging. It is conceivable that the PFT values and the computed image metrics may differ in the sense that respiratory maneuvers differ between CT and PFT.
Third, one side effect of removing the redundancy of features is that this is an exclusive selection. If two features are very similar, one of them may get excluded; meanwhile, both features will still be interesting for a further study if they have very high relevancy values to the disease type.

It should be noted that we did not devise an automated diagnosis system for ILD and COPD in this study. Although feasible and worthy of future study, such an automated diagnosis system approach was beyond the scope of this current study. Rather, the focus of this study was the demonstration that various image metrics are capable of providing useful information to better characterize the clinical diagnosis compared to the information provided by reference standard PFT values.

Although the distinction of COPD and ILD is usually evident from both clinical and imaging standpoints, the real power of this approach will be unleashed if these methods can be extended to subclassify different ILD and COPD phenotypes in the future, which is a far more challenging clinical task, one that is not easily addressed by standard imaging modalities or PFT, and will be the focus of future studies.
CHAPTER 7

Quantification of Small Airway Air Trapping and Emphysema in Obstructive Pulmonary Disease

7.1 Introduction

Obstructive pulmonary disease is most often due to chronic obstructive pulmonary disease (COPD), which is a major global public health problem. It is the fourth leading cause of chronic morbidity and mortality in the United States, and is anticipated to rank fifth in 2020 in burden of disease caused worldwide [1 51]. Moreover, among the four major causes of mortality, namely cardiovascular disease, malignant neoplasm, cerebrovascular disease, and COPD, the last is the only one that has been steadily rising in prevalence [51].

COPD is defined as chronic, progressive airflow limitation that is not fully reversible, associated with a range of pathological changes in the lungs with significant extra-pulmonary effects, caused by chronic inflammation and structural changes [1]. The chronic airflow limitation is caused by a mixture of small airway disease (obstructive bronchiolitis) and
parenchymal destruction (emphysema). The relative contributions of these two components vary substantially from patient to patient. The presence and extent of each component has the potential to affect clinical presentation, disease severity, prognosis, and therapeutic response [1].

The concept of COPD phenotyping addresses this variability in pathophysiologic presentation, by attempting to separate emphysema (which we designate as static air trapping) from small airway disease (which we designate as dynamic air trapping) and to quantify their relative contributions to each individual patient’s clinical presentation [78, 77]. Currently, the diagnosis, classification of severity, and clinical follow up of COPD patients rely on pulmonary function testing (PFT), which includes spirometry, gas diffusion testing, and lung volumes by body plethysmography [1]. Based on spirometric parameters, patients are then categorized in normal, restrictive, obstructive, or combined patterns of disease, where the obstructive pattern is typically associated with COPD [1]. Furthermore, the degree of obstruction can be assessed quantitatively. Nevertheless, PFT parameters have a fundamental limitation, as they provide a global assessment of pulmonary function but cannot provide information regarding regional heterogeneity of disease nor quantification of the relative contributions of emphysema and small airway disease.

Imaging plays a secondary but increasingly important role in COPD phenotyping, particularly in the evaluation of emphysema [42]. Most pulmonary imaging used clinically currently assesses anatomic changes and provides qualitative or gross semi-quantitative estimates of disease severity. However, functional assessment is limited and quantitative assessment is rarely used outside the research setting. Small airway disease is often difficult to diagnose directly. Nonetheless, through detection of regions of pulmonary air trapping as a surrogate marker for small airway disease, paired inspiratory and expiratory thoracic high resolution computed tomography (HRCT) may provide a means for quantitative COPD phenotyping.

On HRCT, the volume of emphysema is typically measured through application of an attenuation threshold mask (which may vary from -890 to -980 HU) upon inspiratory images
This is equivalent to static air trapping, in the sense that emphysematous areas of the lung lose elastic recoil and therefore do not change appreciably in volume between inspiration and expiration. In contrast, small airway air trapping is characterized by lung parenchymal areas that are not emphysematous but which fail to increase in attenuation (or to decrease in volume) beyond a certain level between full inspiration and end expiration. In clinical practice, the presence of small airway air trapping is typically determined qualitatively through visual detection of persistent areas of low attenuation on expiratory HRCT images relative to inspiratory HRCT images, and the degree of small airway air trapping is assessed semi-quantitatively (none, mild, moderate, severe). Inability to objectively quantify the amount of small airway air trapping severely limits one’s ability to phenotype COPD and to accurately assess response following therapeutic intervention, particularly in the setting of new drug development.

The goals of this preliminary study are to propose a computer-assisted methodology to extract quantitative data about emphysema and small airway air trapping via a deformable registration model between paired inspiratory and expiratory HRCT image datasets, automatic segmentation of pulmonary volumes, and to study their correlations with PFT parameters. We hypothesize that small airway air trapping quantification in addition to emphysema quantification will provide better functional correlation between HRCT and PFT when compared to emphysema quantification alone.

### 7.2 Materials and Methods

Approval from our Institutional Review Board (IRB) and a Health Insurance Portability and Accountability Act (HIPAA) waiver were obtained prior to study initiation for retrospective analysis of patient image datasets and PFT results.
Figure 7.1: Illustration of the effectiveness of the registration algorithm. (a) Representative coronal thoracic HRCT image. (b) Superimposed 3D lung and central airways segmentation. (c) Coronal subtraction image (inspiration-expiration) prior to registration. Vessel and airway misalignment is noted. (d) The same subtraction after registration. Vascular and airway alignment are substantially improved, indicating the correctness of the registration mapping. Brighter areas indicate greater difference in attenuation values.
7.2.1 Patient selection

The selection criteria for inclusion in the study were an obstructive pattern on PFT, as measured by FEV1/FVC<0.7 (indicating an obstructive pulmonary disease) and an available paired inspiratory and expiratory thoracic HRCT within 3 days of PFT. 17 patients were chosen with the following characteristics: 11 females and 6 males; mean age of 56.76 years (range of 39 to 86 years). The severity of obstruction on PFT was classified following the 2005 ATS/ERS guidelines as mild (3 patients), moderate (3 patients), severe (5 patients), and very severe (6 patients). 11 patients had a primary clinical or radiologic diagnosis of COPD, whereas 3 patients had nontuberculous mycobacterial infection, 1 patient had cystic fibrosis, 1 patient had lymphangioleiomyomatosis (LAM), and 1 patient had granulomatous-lymphocytic interstitial lung disease (GLILD).

7.2.2 Image acquisition

All patients were scanned using 64-slice multi-detector row computed tomography (MDCT) (Siemens medical systems, Malvern, PA). Images were acquired with 120 kVp, variable mAs due to automatic tube current modulation, 64 x 0.75 mm collimation, pitch of 0.8, and gantry rotation time of 0.5 seconds first during full inspiration and then during end expiration. Images were reconstructed with 1 mm slice thickness and 1 mm interval (no interslice gap) using a B41f kernel and a 512 x 512 matrix.

7.2.3 Image analysis

HRCT inspiratory and expiratory datasets were then anonymized and sent to the registration pipeline that automatically co-registered both datasets and generated a voxel by voxel correspondence between them.

We used our diffeomorphic registration algorithm to perform the deformable image registration. It generates a symmetric diffeomorphic deformation field [11, 9] between the
inspiratory and expiratory datasets and guarantees the smoothness and invertibility of the resultant deformation field, which is essential for analyzing pulmonary kinematics.

Aerated lung in inspiration (ALi) and expiration (ALE) were segmented using an attenuation mask that excluded any voxel above -50 Hounsfield units (HU). The central airways from the trachea to the 5th or 6th generation were removed from the lung volume using the airway segmentation algorithm in Chapter 2. A seed point in the top trachea was first localized by detecting circles in each slice from top to bottom. From the seed point, the airway tree was segmented by applying a multi-stencil level set method [45]; regions with segmentation leakage were further automatically refined using multiple geometric constraints. The difference between ALi or ALe and central airways volume (CAV) was calculated as the total segmented lung volume in inspiration (TSLVi) and expiration (TSLVe), respectively.

Emphysema volume (EV), corresponding to static air trapping, was calculated in inspiration using an attenuation mask that segmented voxels less than -950 HU, excluding central airways. The difference between TSLVi and EV was calculated as non-emphysema volume (NEV).

The deformable registration provided mappings of voxels between inspiration and expiration image volumes. For each voxel in the expiration volume and its correspondent in the inspiration volume, the difference of attenuation values was computed as the attenuation change from inspiration to expiration (Fig. 7.1). Small airway air trapping volume (ATV) was defined and calculated as the volume of all voxels in which the attenuation increase from inspiration to expiration was lower than a specified change threshold, excluding emphysematous voxels. The values of the change threshold were tested from 5 to 300 HU in 25 HU increments. Total air trapping volume (tATV) was also computed as EV+ATV.

Different change thresholds were applied to obtain ATV and corresponding tATV (EV+AT). We evaluated which change threshold was optimal to assess the severity of AT by choosing the specific threshold that provided strongest correlations with the forced expiratory volume
in 1 second to forced vital capacity ratio (FEV1/FVC) and also with residual volume (RV) as shown in the results section below.

7.2.4 Statistical analysis

These volumes (TSLVi, TSLVe, EV, NEV, ATV, tATV) computed from HRCT image analysis were compared with PFT parameters through correlation analysis. Statistical Analysis was performed to compute Pearson’s linear correlation coefficient (r), respective p-values (p), and also the sample standard deviations for the volumes computed.

7.3 Results

Mean computed volumes (+ standard deviation) were calculated for each air trapping change threshold and plotted. Patients were further subdivided into two subgroups (Fig. 7.2 and 7.3). Subgroup 1 was characterized by negligible EV (less than 2%, mean 0.88%, range 0.07-1.88%), with n=7 patients (Fig. 7.4). Subgroup 2 was characterized by moderate or severe EV (greater than 12%, mean 23.36%, range 12.02-34.38%), with n=10 patients (Fig. 7.5). Mean volumes and standard deviation for all patients and for both subgroups are shown on Table 7.1. The computed difference between mean TSLVi mean TSLVe between both subgroups is statistically significant (p = 0.024).

<table>
<thead>
<tr>
<th>TSLVi (ml)</th>
<th>TSLVe (ml)</th>
<th>EV (ml)</th>
<th>EV (%)</th>
<th>TSLVi-TSLVe (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vol</td>
<td>Std Dev</td>
<td>Vol</td>
<td>Std Dev</td>
<td>Vol</td>
</tr>
<tr>
<td>All Patients</td>
<td>5355.43</td>
<td>1315.38</td>
<td>4240.41</td>
<td>1496.48</td>
</tr>
<tr>
<td>Negligible EV</td>
<td>4492.36</td>
<td>706.9</td>
<td>3196.84</td>
<td>619.85</td>
</tr>
<tr>
<td>Mod/Severe EV</td>
<td>5959.58</td>
<td>1325.86</td>
<td>4970.9</td>
<td>1511.68</td>
</tr>
</tbody>
</table>

TSLVi = total segmented lung volume in inspiration; TSLVe = total segmented lung volume in expiration; EV = emphysema volume; EV % denotes percentage of emphysema (compared to TSLVi); and TSLVi-TSLVe % denotes percentage of volume change between inspiration and expiration. *Statistically significant difference, p=0.024.

Pearson linear correlation coefficients (r) and respective p-values (p) were calculated between...
Figure 7.2: Representative thoracic HRCT images in (a) axial, (b) sagittal, and (c) coronal planes in subgroup 1 patient with obstructive lung disease (negligible emphysema). Segmented small airway air trapping shown in red and segmented emphysema is shown in green in (d) axial, (e) sagittal, and (f) coronal planes.
Figure 7.3: Representative thoracic HRCT images in (a) axial, (b) sagittal, and (c) coronal planes in subgroup 1 patient with obstructive lung disease (moderate/severe emphysema). Segmented small airway air trapping shown in red and segmented emphysema is shown in green in (d) axial, (e) sagittal, and (f) coronal planes.
Figure 7.4: Computed volumes for each change threshold for the negligible emphysema subgroup (n = 7 patients). Error bars denote 1 standard deviation. ATV, small-airway air-trapping volume; EV, emphysema volume; HU, Hounsfield units; TSLVe, total segmented lung volume in expiration; TSLVi, total segmented lung volume in inspiration.
Figure 7.5: Computed volumes for each change threshold for the moderate or severe emphysema subgroup (n = 10 patients). Error bars denote 1 standard deviation. ATV, small-airway air-trapping volume; EV, emphysema volume; HU, Hounsfield units; TSLVe, total segmented lung volume in expiration; TSLVi, total segmented lung volume in inspiration.
computed volumes and key PFT parameters (Table 7.2). Strong and clinically meaningful correlations are described specifically below.

Table 7.2: Pearson linear correlation coefficients and respective p-values between key PFT parameters and calculated emphysema and small airway air trapping volumes.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EV</td>
<td>-0.034</td>
<td>0.889</td>
<td>-0.55</td>
<td>0.022</td>
<td>-0.862</td>
<td>0.001</td>
<td>0.46</td>
</tr>
<tr>
<td>AT at 5 HU</td>
<td>-0.14</td>
<td>0.591</td>
<td>-0.528</td>
<td>0.029</td>
<td>-0.709</td>
<td>0.001</td>
<td>0.423</td>
</tr>
<tr>
<td>AT at 25 HU</td>
<td>-0.027</td>
<td>0.917</td>
<td>-0.548</td>
<td>0.023</td>
<td>-0.77</td>
<td>0.001</td>
<td>0.511</td>
</tr>
<tr>
<td>AT at 50 HU</td>
<td>0.033</td>
<td>0.901</td>
<td>-0.521</td>
<td>0.032</td>
<td>-0.701</td>
<td>0.002</td>
<td>0.604</td>
</tr>
<tr>
<td>AT at 75 HU</td>
<td>0.139</td>
<td>0.595</td>
<td>-0.886</td>
<td>0.126</td>
<td>-0.574</td>
<td>0.016</td>
<td>0.557</td>
</tr>
<tr>
<td>AT at 100 HU</td>
<td>0.271</td>
<td>0.292</td>
<td>-0.193</td>
<td>0.457</td>
<td>-0.422</td>
<td>0.092</td>
<td>0.395</td>
</tr>
<tr>
<td>AT at 125 HU</td>
<td>0.366</td>
<td>0.149</td>
<td>-0.05</td>
<td>0.848</td>
<td>-0.313</td>
<td>0.221</td>
<td>-0.262</td>
</tr>
<tr>
<td>AT at 150 HU</td>
<td>0.435</td>
<td>0.082</td>
<td>0.05</td>
<td>0.848</td>
<td>-0.242</td>
<td>0.349</td>
<td>-0.17</td>
</tr>
<tr>
<td>AT at 175 HU</td>
<td>0.474</td>
<td>0.055</td>
<td>0.113</td>
<td>0.666</td>
<td>-0.198</td>
<td>0.446</td>
<td>-0.112</td>
</tr>
<tr>
<td>AT at 200 HU</td>
<td>0.504</td>
<td>0.041</td>
<td>0.155</td>
<td>0.553</td>
<td>-0.168</td>
<td>0.52</td>
<td>-0.073</td>
</tr>
<tr>
<td>AT at 225 HU</td>
<td>0.518</td>
<td>0.033</td>
<td>0.184</td>
<td>0.479</td>
<td>-0.146</td>
<td>0.575</td>
<td>-0.045</td>
</tr>
<tr>
<td>AT at 250 HU</td>
<td>0.530</td>
<td>0.029</td>
<td>0.206</td>
<td>0.428</td>
<td>-0.131</td>
<td>0.617</td>
<td>-0.024</td>
</tr>
<tr>
<td>AT at 275 HU</td>
<td>0.539</td>
<td>0.026</td>
<td>0.221</td>
<td>0.393</td>
<td>-0.119</td>
<td>0.649</td>
<td>-0.009</td>
</tr>
<tr>
<td>AT at 300 HU</td>
<td>0.556</td>
<td>0.023</td>
<td>0.233</td>
<td>0.368</td>
<td>-0.11</td>
<td>0.674</td>
<td>0.003</td>
</tr>
<tr>
<td>EV+AT at 5 HU</td>
<td>-0.061</td>
<td>0.816</td>
<td>-0.565</td>
<td>0.018</td>
<td>-0.81</td>
<td>0.001</td>
<td>-0.469</td>
</tr>
<tr>
<td>EV+AT at 25 HU</td>
<td>-0.034</td>
<td>0.897</td>
<td>-0.593</td>
<td>0.012</td>
<td>-0.848</td>
<td>0.001</td>
<td>-0.529</td>
</tr>
<tr>
<td>EV+AT at 50 HU</td>
<td>0.007</td>
<td>0.979</td>
<td>-0.613</td>
<td>0.009</td>
<td>-0.852</td>
<td>0.001</td>
<td>-0.632</td>
</tr>
<tr>
<td>EV+AT at 75 HU</td>
<td>0.085</td>
<td>0.747</td>
<td>-0.561</td>
<td>0.019</td>
<td>-0.826</td>
<td>0.001</td>
<td>-0.643</td>
</tr>
<tr>
<td>EV+AT at 100 HU</td>
<td>0.185</td>
<td>0.476</td>
<td>-0.429</td>
<td>0.064</td>
<td>-0.776</td>
<td>0.001</td>
<td>-0.561</td>
</tr>
<tr>
<td>EV+AT at 125 HU</td>
<td>0.262</td>
<td>0.311</td>
<td>-0.376</td>
<td>0.137</td>
<td>-0.734</td>
<td>0.001</td>
<td>-0.485</td>
</tr>
<tr>
<td>EV+AT at 150 HU</td>
<td>0.318</td>
<td>0.213</td>
<td>-0.313</td>
<td>0.222</td>
<td>-0.705</td>
<td>0.002</td>
<td>-0.429</td>
</tr>
<tr>
<td>EV+AT at 175 HU</td>
<td>0.355</td>
<td>0.163</td>
<td>-0.271</td>
<td>0.292</td>
<td>-0.685</td>
<td>0.002</td>
<td>-0.392</td>
</tr>
<tr>
<td>EV+AT at 200 HU</td>
<td>0.379</td>
<td>0.134</td>
<td>-0.242</td>
<td>0.349</td>
<td>-0.671</td>
<td>0.003</td>
<td>-0.366</td>
</tr>
<tr>
<td>EV+AT at 225 HU</td>
<td>0.396</td>
<td>0.115</td>
<td>-0.221</td>
<td>0.394</td>
<td>-0.659</td>
<td>0.004</td>
<td>-0.347</td>
</tr>
<tr>
<td>EV+AT at 250 HU</td>
<td>0.409</td>
<td>0.103</td>
<td>-0.205</td>
<td>0.43</td>
<td>-0.651</td>
<td>0.005</td>
<td>-0.332</td>
</tr>
<tr>
<td>EV+AT at 275 HU</td>
<td>0.418</td>
<td>0.095</td>
<td>-0.193</td>
<td>0.458</td>
<td>-0.644</td>
<td>0.005</td>
<td>-0.321</td>
</tr>
<tr>
<td>EV+AT at 300 HU</td>
<td>0.425</td>
<td>0.089</td>
<td>-0.183</td>
<td>0.481</td>
<td>-0.638</td>
<td>0.006</td>
<td>-0.312</td>
</tr>
</tbody>
</table>

EV = emphysema volume; NEV = non-emphysema volume; FVC = functional vital capacity; FEV1 = forced expiratory volume in the 1st second; FEF25-75% is the forced expiratory volume between 25 and 75% of FVC; ERV = end respiratory volume; RV = residual volume; TLC = total lung capacity. All PFT parameters are percentage of predicted (%P). RV and TLC are obtained by plethysmography.

A strong positive correlation was observed between TSLVi and total lung capacity (TLC) as measured by body plethysmography (r= 0.83). A strong negative correlation was demonstrated between EV and FEV1/FVC (r=-0.80).

Subgroup analysis revealed distinct trends within each subgroup. For subgroup 1 (negligible emphysema/static air trapping), there was no significant correlation of EV and FEV1/FVC or RV (respectively, r=0.10 and r=-0.11). On the other hand, there was a strong negative correlation between ATV and FEV1/FVC (minimum r=-0.85 at a change threshold of 75 HU) and a strong positive correlation between ATV and RV (maximum r=0.95 at a change threshold of 75 HU) (Fig. 7.6 and 7.9). For subgroup 2 (moderate or
Figure 7.6: Pearsons correlations between forced expiratory volume in 1 second (FEV1)/forced vital capacity (FVC) and emphysema volume (EV), small airway air trapping (AT) volume, and total AT volume at each specific change threshold (5300 Hounsfield units [HU], in 25-HU increments) for subgroup 1 (n = 7 patients).
Figure 7.7: Pearsons correlations between residual volume (RV) and emphysema volume (EV), small airway air trapping (AT) volume, and total AT volume at each specific change threshold (5300 Hounsfield units [HU], in 25- HU increments) for subgroup 1 (n = 7 patients).
severe emphysema/static air trapping), there was a weaker negative correlation between EV and FEV1/FVC (r=-0.52) and slightly improved negative correlation between ATV and FEV1/FVC (minimum r=-0.64 at a change threshold of 25 HU). Adding ATV to EV (tATV) yielded a small improvement in correlation strength (minimum r=-0.68 at a change threshold of 25 HU). There was no significant correlation between EV and RV (r=0.26), whereas there was moderately strong correlation between ATV and RV (r=0.68 at a change threshold of 5 HU) (Fig. 7.8 and 7.9).

Figure 7.8: Pearsons correlations between residual volume (RV) and emphysema volume (RV), small airway air trapping (AT) volume, and total AT volume at each specific change threshold (5300 Hounsfield units [HU], in 25-HU increments) for subgroup 2 (n = 10 patients).
Figure 7.9: Pearsons correlations between forced expiratory volume in 1 second (FEV1)/forced vital capacity (FVC) and emphysema volume (EV), small airway air trapping (AT) volume, and total AT volume at each specific change threshold (5300 Hounsfield units [HU], in 25-HU increments) for subgroup 2 (n = 10 patients).
7.4 Discussion

Emphysema quantification on thoracic HRCT utilizing attenuation masks in inspiration has long been performed [73]. However, small airway air trapping quantification on thoracic HRCT has only been implemented relatively recently through various approaches. The simplest involves qualitative and semi-quantitative assessment of low attenuation areas in the lungs on expiratory HRCT images [3, 65]. More sophisticated quantitative approaches of small airway air trapping (in chronological order of appearance in the literature) involve measurements of lung cross-sectional areas of air-trapping on expiratory CT images and lung cross-sectional area changes between inspiratory and expiratory CT images [64], use of optical flow techniques to register and subtract paired inspiratory and expiratory multi-detector row HRCT image datasets to generate attenuation difference color maps and histograms [103, 33], application of a -860 HU attenuation mask on expiratory HRCT images [70], and measurement of attenuation changes [70] or volumetric changes [119] of the lungs between inspiratory and expiratory HRCT image datasets.

Our current approach utilizes symmetric diffeomorphic deformable registration to generate voxel-to-voxel mapping, which allows automatic computation of attenuation differences between inspiration and expiration. By varying the threshold of the attenuation difference, we obtained an adequate match between the quantitatively segmented ATV and the qualitatively assessed ATV. The computed volume can potentially be used to predict some PFT parameters and also as a quantitative metric to assess true ATV.

PFT only provides global information about respiratory mechanics and physiology, which makes it relatively insensitive to early/mild disease conditions and also causes difficulties in the interpretation of results when there are co-existing disease states that cause opposite pathophysiologic effects, such as when emphysema and smoking related interstitial lung diseases coexist. These fundamental limitations of PFT could potentially be overcome by an imaging modality that not only globally but also regionally quantifies disease patterns, pro-
viding for quantification of their individual contributions to the overall clinical presentation, and therefore allowing for early diagnosis and treatment. Moreover, obtaining functional information from HRCT would be highly desirable, given that this method is the most widespread advanced lung imaging modality, such that no new imaging equipment or imaging protocols would be required. From a clinical perspective, obtaining such information may also provide new diagnostic and therapeutic insights on prevalent conditions beyond COPD such as asthma [30, 24] and cystic fibrosis [41].

The optimal threshold for quantifying small airway air trapping from HRCT varies among patients, as defined by the best match between automatically segmented air trapping areas and visual assessment of these regions by experienced thoracic radiologists, noting that PFT cannot quantify small airway air trapping directly. We observed that the best agreement was within the change threshold range of 25-75 HU. This range coincided with the strongest correlation with key PFT parameters as described above. Although there is currently no established way to quantify small airway air trapping due to absence of a reference standard, we hypothesize that the optimal change threshold for quantification of small airway air trapping lies within the 25-75 HU range. In our patient sample, there was no benefit of increasing the change threshold beyond 100 HU, as can be seen in the volume curves and correlation tables. Any value above 100 HU essentially included most of the segmented lung volume, noting substantially weaker correlations at higher change thresholds with key PFT parameters (Table 7.2).

Analysis of the severity of emphysema/static air trapping clearly delineated two strikingly different patient subgroups within our patient sample. Subgroup 1 (with negligible emphysema/static air trapping) had less than 2% EV/TSLVi (mean 0.88%, range 0.07-1.88%). Conversely, subgroup 2 (with moderate or severe emphysema) had greater than 12% EV/TSLVi (mean 23.36%, range 12.02-34.38%). Phenotypically, these subgroups would correspond to small airway predominant COPD and emphysema predominant COPD, respectively. This distinction may be clinically useful, as prognosis, treatment strategies, and
outcomes vary between the groups [1, 78, 77]. Subgroup 1 had a mean volume change between inspiration and expiration \(((TSLVi-TSLVe)/TSLVi) \times 100\%\) of 28.59\% + 10.06\% compared to 17.98\% + 9.35\% for subgroup 2, noting that the difference is statistically significant \((p = 0.024)\). This difference is likely secondary to increased loss of elastic recoil of the moderate/severe emphysema group in comparison to the negligible emphysema group. Additionally, there was a strikingly different behavior of the correlation patterns between the two subgroups. For instance, as subgroup 1 had negligible emphysema, we hypothesize that most of the obstructive physiology in this group is caused by small airway air trapping. Accordingly, there was no significant correlation between EV and FEV1/FVC or RV, although there was a fairly strong negative correlation between ATV and FEV1/FVC \((r=-0.85\) at a change threshold of 75 HU) and stronger positive correlation between ATV and RV \((\text{maximum } r=0.95\) at a change threshold of 75 HU\), indicating that quantification of small airway air trapping has potential to assess the clinical degree of obstruction in this group whereas quantification of emphysema/static air trapping has not. On the other hand, subgroup 2 had moderate or severe emphysema/static air trapping and smaller contributions of small airway air trapping. For this subgroup, EV correlated moderately with FEV1/FVC \((r=-0.52\) and weakly with RV \((r=0.26)\). ATV correlated similarly but slightly more strongly with FEV1/FVC \((r=-0.64\) at a change threshold of 25 HU\), but much more strongly with RV \((r=0.68\) at a change threshold of 5 HU\). There was no significant change in correlations when adding ATV to EV in this group. Nonetheless, due to the stronger correlations with ATV in comparison to EV, our data suggests that even in the moderate/severe emphysema subgroup, there is still benefit of adding ATV to more standardized EV to predict the severity of obstruction based on PFT.

Another interesting observation is that the optimal change thresholds for calculation of ATV varied between the subgroups. In subgroup 1, the optimal change threshold was 75 HU (maximizing strength of correlation with FEV1/FVC and RV), whereas in subgroup 2, the optimal change thresholds were respectively 25 HU (maximizing strength of correlation with FEV1/FVC) and 5 HU (maximizing strength of correlation with RV). We hypothe-
size that this difference is mostly secondary to the greater difference between inspiration (TSLVi) and expiration (TSLVe) volumes within subgroup 1 compared to subgroup 2. From a pathophysiologic perspective, subgroup 2 had more severe parenchymal destruction leading to increased static air trapping and greater loss of elastic recoil, with consequently smaller mean voxel attenuation differences between inspiration and expiration. Conversely, subgroup 1 had relatively preserved lung parenchyma with air trapping caused mostly by small airway disease. Therefore, this subgroup had a greater difference in lung volumes between inspiration and expiration, with consequently greater mean voxel attenuation differences.

There are limitations of the current study. Firstly, the sample size was small and the study was conducted retrospectively. Further inquiries to determine the validity and clinical applicability of our approach will necessitate larger scale studies and prospective study designs. Secondly, there was no spirometric control of patient expiratory effort during HRCT. Therefore, even though patients were instructed at the time of imaging and most expiratory HRCT images were deemed acceptable by the interpreting radiologists, it is conceivable that the spirometric volumes would differ from HRCT measured volumes given the technical differences between the studies. This could potentially be overcome in future studies by use of respiratory bellows to control the timing of expiratory image acquisition. Thirdly, HRCT volumes were measured discretely at two distinct time points with the patient in the supine position whereas spirometric volume and flow parameters are measured continuously with the patient in the sitting position, using standardized maneuvers to provide volume-time curves. Nonetheless, it is interesting to note that in spite of these limitations, our results still demonstrate strong correlations between TSLVi and TLC and also between FEV1/FVC and EV + ATV, indicating that the volumetric segmentation of the aerated lungs and central airways was reasonably accurate and that HRCT derived volumes can still reasonably predict key PFT parameters. Lastly, there is no current reference standard for quantification of small airway air trapping, as PFT can only provide surrogate markers of global air trapping which do not distinguish between the specific contributions of emphysema (static
air trapping) and small airway disease.

7.5 Conclusion

This study demonstrates the feasibility and potential of computer-assisted inspiratory and expiratory multi-detector row HRCT image analysis for the detection and quantification of small airway air trapping (dynamic air trapping) and emphysema (static air trapping) in obstructive lung disease using pre-existing image datasets. Through automated image registration and image analysis of inspiratory and expiratory HRCT images, we were able to separately segment and quantify volumes of emphysema and small airway air trapping, which demonstrated strong correlations with key PFT parameters. We were also able to distinguish two phenotypes of obstructive pulmonary disease: primarily due to emphysema/static air trapping and primarily due to small airway air trapping. These observations suggest that quantification of different components of pulmonary air trapping from multi-detector row HRCT in patients with obstructive lung disease is feasible, and may have great potential to assess phenotypic differences in patients with COPD.
CHAPTER 8

Conclusion

This dissertation investigated the problem of using thoracic computed tomography images for automatically and quantitatively analyzing and diagnosing lung diseases. We studied various components and discussed how to complete a holistic automatic pipeline for clinical studies by integrating multiple components together. The technical and clinical contributions are highlighted in the following:

1. Proposed a novel airway segmentation algorithm which specifically targeted at small airway segmentation. We proposed a form of adjusted image gradients and applied it with multi-stencil fast marching to generate multiple airway hypotheses. The oversegmentation areas were further removed using a novel cost function.

2. Unified parameterization forms of different diffeomorphic deformation models. A common framework was generalized as the optimization in the Sobolev space by deriving the smoothing operation from regularization terms.

3. Adapted the symmetric diffeomorphic registration algorithm for lung registration. We evaluated our registration algorithm in the public EMPIRE10 challenge, which had the top state-of-the-art performance.
4. Proposed a new novel way to construct an accurate poly-affine model. By introducing the trajectories of local regions instead of using regions themselves, the approach precisely encoded each local affine transform using a diffeomorphism with one or more stationary velocity fields.

5. Used feature selection to find optimal image features to differentiate ILD and COPD. Some image metrics were not only as good discriminators as PFT for the characterization of ILD and COPD, but were also not redundant when PFT values were available. Image metrics of attenuation histogram statistics and texture descriptions were valuable for further investigation in computer-assisted diagnosis.

6. Distinguished two phenotypes of obstructive pulmonary diseases using image registration in the pipeline. Quantification of emphysema related and small airway related air trapping had great potential to elucidate phenotypic differences in patients with COPD using thoracic HRCT images.

We showed that quantitative CT imaging analysis pipelines, integrating whole lung registration, anatomy segmentation, texture computation, feature selection and machine learning methods, provided better biomarkers for diagnosis and prognosis. Image-based quantitative metrics allow reproducible correlation with relevant anatomical and physiological parameters, and thus enable superior analysis of treatment response and ultimately contributing to the betterment of patient care.
BIBLIOGRAPHY


