

COMPUTER ASSISTED DETECTION
AND MODELLING OF PAEDIATRIC
AIRWAY PATHOLOGY FROM MEDICAL
IMAGES

A DISSERTATION PRESENTED
BY
BENJAMIN JOHN IRVING
TO
THE CENTRE FOR HEALTH INFORMATICS AND MULTIPROFESSIONAL EDUCATION
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
MEDICAL IMAGE COMPUTING

SUPERVISORS:
PAUL TAYLOR
AND
ANDREW TODD-POKROPEK

UCL
LONDON
SEPTEMBER 2012

I, Benjamin John Irving, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ABSTRACT

THIS THESIS PROPOSES a novel computer assisted detection framework to analyse airway shape change and to detect signs of paediatric pulmonary tuberculosis (TB). The method can accurately distinguish TB patients with airway involvement, from non-TB patients using CT scans. This model is also applied to X-ray radiographs to segment the airways.

As a first step, a CT airway segmentation algorithm is proposed, and then evaluated as part of the EXACT'09 airway segmentation challenge. Algorithms are then implemented to extract the airway centreline, detect branch points and label each airway branch. A number of cases had the appearance of complete obstruction in some bronchi, and a method is presented to identify and segment beyond these obstructions.

A statistical shape model is developed using airway shape variation, which requires correspondence between airways. Thus, a method to register regions of the airway tree is proposed. This method generates landmarks on the airway surface. Using these landmarks, a template mesh is then aligned to each airway by thin-plate-spline warp and local vertex alignment. This develops a corresponding surface mesh representation for each airway that can be used for statistical analysis.

A statistical model of the variation of local regions of the airway is constructed and features derived from this are used to train a classifier to detect abnormal airway variation. The method is able to accurately detect TB in unseen cases. It is also compared to second method based on automated bronchi measurements.

Finally, a method is developed that uses the previous 3D model to segment the airways in 2D radiographs. This uses an optimisation algorithm to fit the 3D model to image features in the radiograph.

These methods create novel tools for airway analysis and have wider applications in medical imaging.

DEDICATED TO MY FATHER, DAVID, WHO FIRST GOT ME INTERESTED IN ENGINEERING AND PROBLEM-SOLVING. HE WOULD HAVE BEEN PROUD THAT I'VE FOLLOWED A SIMILAR DIRECTION.

ACKNOWLEDGEMENTS

I WOULD LIKE TO THANK my supervisor Paul Taylor for guiding me through this PhD, giving invaluable advice along the way, and finding funding for conferences around the world. Andrew Todd-Pokropek, my co-supervisor, provided important insights and ideas for this multi-disciplinary project. Tania Douglas for proposing the initial idea for this research and providing an opportunity for me to be based at MIRU, UCT for two months. Thanks to everyone who has given me advice over the last three years at CHIME, CMIC and Medical Physics at UCL, especially Henry Potts, Dionisio Acosta, Kath Woolf, Elizabeth-Ashley Durham, Chris Martin, Tryphon Lambrou, Jamie McClelland, John Hipwell, Yipeng Hu and Dave Hawkes. During the write-up of my thesis a number of people have provided helpful feedback and proof reading. Thanks to Tania Douglas, Kit Vaughan and Michael Evans. My internship at MedicSight during my PhD rounded my knowledge, so thanks to Greg Slabaugh and the rest of the MedicSight R&D team. The experience of experts in TB and radiology have been essential to this project and I would like to thank Savvas Andronikou, Pierre Goussard and Rob Gie for providing clinical input. Thanks to Stef Stiener for providing access to data from Lodox Statscan. I would like to acknowledge the Commonwealth Scholarship Commission for providing funding that made this PhD possible, and organising fantastic international student events, from sustainable development conferences at Cumberland lodge to cocktails at Buckingham palace. I have had a great experience as a PhD student at UCL and will miss all the people, and especially, the unusual conversations at CHIME lunch times. Thanks to Prof Dipak Kalra, Espy Rodrigues and Tanya Harbinson for looking after me at CHIME. Finally, thanks to family and friends for all the support they given me over these last 3 years, and particularly to my mother and sister, Jenne and Michaela, for all the support and encouragement during the writeup of my thesis.

Contents

1	Introduction	22
1.1	Tuberculosis	23
1.2	Paediatric Pulmonary Tuberculosis	24
1.3	Lymphadenopathy and airway deformation	26
1.3.1	Airway anatomy	27
1.3.2	Lymphadenopathy	27
1.3.3	Airway involvement	29
1.3.4	Other radiographic signs of paediatric TB	30
1.3.5	Other diseases affecting the trachea and bronchi	30
1.4	Airway imaging	33
1.5	Computer-assisted detection of chest pathology	37
1.6	Thesis outline	40
1.7	Contributions	42
2	Automated Airway Segmentation	44
2.1	Background	46
2.2	Segmentation requirements	48
2.3	Method	49
2.3.1	Initialisation	51
2.3.2	Segmentation of Trachea and Main Bronchi	53
2.3.3	Morphological Filtering	54
2.3.4	Region Growing and Leak Removal	58
2.4	CT Datasets	63
2.4.1	EXACT'09 Adult dataset	63
2.4.2	GOSH 2010 paediatric nonTB CTs	63
2.4.3	Tygerberg 2010 paediatric TB CTs	64
2.4.4	Tygerberg 2012 paediatric TB and non-TB CTs	64
2.5	Evaluation and Results	64
2.5.1	EXACT'09 challenge (adult patients)	65
2.5.2	EXACT'09 Segmentation evaluation	66

2.5.3	EXACT'09 challenge general results	70
2.5.4	Improvements to method after EXACT'09	74
2.5.5	Paediatric airway segmentation	74
2.6	Discussion	80
2.7	Conclusion	81
2.8	Chapter summary	83
3	Airway Structure	84
3.1	Background	85
3.1.1	Skeletonisation	85
3.1.2	Anatomical branch labelling	88
3.2	Method	91
3.2.1	Airway Skeletonisation	91
3.2.2	Branch Point Detection	92
3.2.3	Branch connectivity labelling	93
3.2.4	Pruning	94
3.2.5	Anatomical branch labelling	94
3.2.6	Airway surface labelling	96
3.3	Results and discussion	97
3.4	Conclusions	99
3.5	Chapter Summary	101
4	Airway registration and surface correspondence	102
4.1	Background	103
4.2	Method	107
4.2.1	Surface point projection	108
4.2.2	Thin-plate-spline alignment	114
4.2.3	Local mesh alignment	115
4.2.4	Local shape models	119
4.3	Evaluation	120
4.3.1	Mesh registration	120
4.3.2	Mesh quality	122
4.4	Results	124
4.4.1	Algorithm parameters	124
4.4.2	Parameters and accuracy of mesh alignment	126
4.5	Discussion and Conclusions	128
4.6	Chapter summary	133
5	Obstructed Branch Detection	134
5.1	Background	135
5.2	Method	136

5.2.1	Identification of obstruction	138
5.2.2	Position of obstruction	139
5.2.3	Segmentation of disconnected airway regions	139
5.3	Results and discussion	141
5.3.1	Dataset	141
5.3.2	Detection of missing and incomplete branches	142
5.3.3	Segmentation of disconnected airway regions	144
5.4	Conclusion	148
5.5	Chapter summary	148
6	Automated detection of paediatric tuberculosis from airway shape	149
6.1	Background	151
6.1.1	Introduction to feature extraction and classification	151
6.1.2	Statistical shape models	152
6.1.3	Airway Analysis	153
6.2	Method	155
6.2.1	Feature Vectors	155
6.2.2	Classification	159
6.3	Evaluation Methods and Data	161
6.3.1	Datasets	161
6.3.2	Shape Model Evaluation	163
6.3.3	Evaluation of Classification	164
6.4	Results	166
6.4.1	Shape model derived features	167
6.4.2	Radius derived features	180
6.4.3	Run time	181
6.5	Discussion	181
6.6	Conclusion	184
6.7	Chapter summary	186
7	Model projection and airway segmentation in radiographs	187
7.1	Background	188
7.2	Method	190
7.2.1	Silhouette Detection	191
7.2.2	Model Projection and alignment	193
7.2.3	Image Features	195
7.2.4	Optimisation	198
7.3	Evaluation	199
7.3.1	Dataset	201
7.3.2	Manual annotations	201
7.3.3	Comparison	201

7.4	Results	203
7.4.1	Airway segmentation	203
7.5	Discussion	210
7.6	Conclusion	211
7.6.1	Future work	212
8	Conclusions	213
8.1	Summary	213
8.2	General remarks and future directions	216
A	Statistical and classification methods	219
A.1	Generalised procrustes analysis	219
A.2	Principal component analysis	221
A.3	Classification methods	222
B	Structure analysis: additional examples and evaluation	224
C	Data from registration analysis	228
D	Shape model features	233
	References	258

List of Figures

1.1	An airway segmentation with labelled branches	28
1.2	Coronal minimum intensity projection (MinIP) image with arrows indicating lymphadenopathy. Image courtesy of Prof S. Andronikou, Radiology Department, University of Witwatersrand.	31
1.3	Coronal minimum intensity projection (MinIP) image with arrows indicating lymphadenopathy. Stenosis of the left main bronchus (LMB) is visible. Image courtesy of Prof S. Andronikou, Radiology Department, University of Witwatersrand.	31
1.4	Coronal minimum intensity projection (MinIP) image with arrows indicating lymphadenopathy. Stenosis of the left main bronchus (LMB) and bronchus intermedius (BI) is visible. Image courtesy of Prof S. Andronikou, Radiology Department, University of Witwatersrand	31
1.5	Axial, sagittal and coronal slices showing the airways in a chest CT scan	34
1.6	Visualisation of the airways using Statscan with labels showing the visible airway branches and the subcarinal angle (SCA). Image courtesy of Virginia Saunders, MIRU, University of Cape Town.	36
2.1	General steps in an airway segmentation algorithm	47
2.2	Outline of the airway segmentation algorithm. Steps include: identification of the trachea, segmentation of the main bronchi, morphological enhancement and segmentation of the remaining bronchi, and finally a region growing and leak removal step.	50
2.3	Automatic identification of start and end slice positions for the airway segmentation algorithm	52
2.4	Manual selection of start and end slice positions	53
2.5	Trachea and potential misclassifications in the initial slice	54
2.6	Excluding leaks in trachea and main bronchi segmentation by restricting the region growing step based on compactness of the previous slice. r_0 is the maximum radius, r_i and cm_i are the radius and centre-of-mass at the i th iteration	55

2.7	Airway enhancement using morphological reconstruction to enhance circular regions in CT slices	57
2.8	Greyscale closing and reconstruction applied to a CT slice a) Original image b) closing the image with a single structuring element c) morphological reconstruction d) difference image e) threshold	59
2.9	Multidimensional morphological filtering	60
2.10	Leak removal step: identification, labelling and analysis of the changes in the growth front of each branch. a) shows detection of the bifurcation from the connected regions of the growth front b) new growth fronts are generated c) growth fronts are assigned labels based on connectivity to previous growth fronts unless bifurcation occurs	61
2.11	Airway segmentation overlayed on a coronal CT slice. In this case there is at the bifurcation between the trachea and RMB.	61
2.12	Airway segmentation of a CT scan from a 20 month patient with rendering of surrounding boney tissue	62
2.13	Adult airway segmentation with bone rendering (from the EXACT'09 training set)	67
2.14	Segmentation of test cases for the EXACT'09 airway segmentation competition (<i>green</i> is correctly segmented and <i>blue</i> is false positive)	68
2.14	Segmentation of test cases for the EXACT'09 airway segmentation competition (<i>green</i> is correctly segmented and <i>blue</i> is false positive)	69
2.15	Branches detected vs false positive rate for the algorithms in the EXACT'09 airway segmentation competition. Red shows our method.	73
2.16	Tree length detected vs false positive rate for the algorithms in the EXACT'09 airway segmentation competition. Red shows our method.	73
2.17	Examples from the Tygerberg 2012 non-TB paediatric evaluation set (age 2.4 ± 2.8 years) (S: Segmentation, LR: Segmentation with leak removal step)	76
2.17	Examples from the Tygerberg 2012 non-TB paediatric evaluation set (age 2.4 ± 2.8 years)(S: Segmentation, LR: Segmentation with leak removal step)	77
2.18	Examples from the Tygerberg 2012 TB paediatric evaluation set (age 3.1 ± 3.8 years)(S: Segmentation, LR: Segmentation with leak removal step). Some cases have complete obstruction of a bronchus	78
2.18	Examples from the Tygerberg 2012 TB paediatric evaluation set (age 3.1 ± 3.8 years)(S: Segmentation, LR: Segmentation with leak removal step). Some cases have complete obstruction of a bronchus	79
3.1	Bifurcation point detection from the wavefront a) the branches are disconnected in the bifurcation region b) a minimisation algorithm can be used to find the best point of intersection.	88

3.2	Graph representation of the airway tree	89
3.3	Representation a branch point in a one-voxel thick skeleton	92
3.4	Simplified pseudocode for the branch labelling algorithm	93
3.5	Centreline identification from a segmented airway tree (red, green and blue are used to illustrate distinct airway branches)	95
3.6	Airway branch labelling from centreline (red, green and blue are used to illustrate distinct airway branches)	98
3.7	Smoothed centrelines with segmentation and coronal CT slice. The centrelines may not connect at bifurcation points because each branch is individually smoothed. The smoothing algorithm is discussed in Section 4.2.1 of the next chapter.	100
4.1	m-rep representation of a tubular object	106
4.2	The voxel representation of the centreline, the interpolated centreline (blue), and the smoothed and resampled centreline (red). The tangent is calculated for each resampled point (p_i)	109
4.3	Mapping points to the airway surface by ray/triangle intersection of orthogonal vectors. Each set of vectors is translated to p_i and euler angles are used to align the central vector with the centreline tangent. The intersection between the vectors and the airway surface is then found.	110
4.4	Example case 1, mapping corresponding points onto the surface mesh of an airway. Red points represent the surface landmark points obtained from the intersection of the vectors (blue) with the airway surface. Red vectors are tangents to the centreline.	112
4.5	Example case 2, mapping corresponding points onto the surface mesh of an airway	113
4.6	Registration of a template mesh to an airway mesh for the RMB-RUL-BI region	116
4.7	Registration of a template mesh to an airway mesh for the Trachea-RMB-LMB region	117
4.8	Comparison between the mesh representation of a single airway and the registered template mesh. The closest point method shows greater error when branches are narrowed from stenosis	119
4.9	Relabelling the surface mesh where m is the face to be relabelled. m is labelled according to the face's position in relation to the branch end points e_{j1} and e_{j2}	121
4.10	Generating a volume from an airway surface mesh. The volume is represented by voxels that each have a volume $1/125$ of the original voxels in the airway segmentation	122

4.11	A triangle (sides a , b and c) with incircle (radius r) and circumcircle (radius R)	123
4.12	Section of the airway centreline showing the voxel coordinates (squares) and the smoothed and interpolated centreline (derived from the voxels)	125
4.13	Contour plots showing the proportion of error when comparing the accuracy of the registered template mesh to the original mesh. The error is shown as a function of the weighting parameters α and γ of F_1 and F_3 , respectively ($\beta = 1$).	127
4.14	Mean histogram plots of the mesh quality measure q_1 (ratio of shortest and longest edge of each face) for the entire training set. a) shows the mean template before vertex alignment b) shows the mean template using the closest point method c) shows the mean template using the closest point with mesh preservation term d) shows the proposed method	129
4.15	Mean histogram plots of the mesh quality measure q_2 (ratio of face incircle and circumcircle radii) for the entire training set. a) shows the mean template before vertex alignment b) shows the mean template using the closest point method c) shows the mean template using the closest point with mesh preservation term d) shows the proposed method	130
4.16	Registration of the template mesh to the trachea-RMB-LMB section for the entire test set. The first 36 cases are patients diagnosed with TB and the remainder are controls. Obstructed regions are automatically excluded (discussed in Chapter 6).	131
4.17	Registration to the template mesh to the RMB-RUL-BI section for the entire test set. The first 31 cases are patients diagnosed with TB and the remainder are controls. Obstructed regions are automatically excluded (discussed in Chapter 6).	132
5.1	Missing and incomplete categories of obstructions. Branches are defined as incomplete if a centreline has been detected and missing if a centreline has not been detected	137
5.2	Diagram showing the procedure to identify missing child branches . .	138
5.3	Airway shape and branch diameter profiles for obstructed branches for two patients. The top figures show a silhouette of the two patient airways and the bottom figures show branch radius for the RMB bronchus as a function of position along the centreline. The airways are normalised for size (more details in Equation A.2 of Appendix A.1) before branch diameter measurements so the diameters are relative measurements.	140
5.4	Variation along the first PCA mode of RMB	143
5.5	MDS reduction to 2 dimensions of 4 dimensional feature space.	144
5.6	Point model with identified obstruction	145

5.7	Branch labelling and identification of obstructed branch	146
5.8	Segmenting past complete obstruction	147
6.1	Shape models are generated from 3-branch regions (RMB-RUL-BI region in this example)	157
6.2	Branch radius measures using ray-surface intersection. 4 intersections are shown for each equidistant centreline point.	158
6.3	Modes of variation 1-5 for the trachea-LMB-RMB ($b_i = 0$ is the mean model and $b_i = \pm 3\sqrt{\lambda_i}$ represents the shape 3 standard deviations along each mode)	168
6.4	Modes of variation 6-10 for the trachea-LMB-RMB ($b_i = 0$ is the mean model and $b_i = \pm 3\sqrt{\lambda_i}$ represents the shape 3 standard deviations along each mode)	169
6.5	Modes of variation 1-5 of the RMB-RUL-BI ($b_i = 0$ is the mean model and $b_i = \pm 3\sqrt{\lambda_i}$ represents the shape 3 standard deviations along each mode)	170
6.6	Modes of variation 6-10 of the RMB-RUL-BI ($b_i = 0$ is the mean model and $b_i = \pm 3\sqrt{\lambda_i}$ represents the shape 3 standard deviations along each mode)	171
6.7	Example airways from the non-TB and TB datasets with overlaid 11 component representation of the first 3 branches. Small deviations of the model's shape from the airway shape is explained by higher order modes that have been removed from the representation.	172
6.8	The proportion of the total variance in the dataset that is represented by n principal modes. Principal modes are ordered by variance and, therefore, higher modes will make smaller contributions. In this dataset, the first 11 modes represent $\approx 90\%$ of the variation.	173
6.9	Generalisation $G(M)$ and specificity $S(M)$ of the model with confidence intervals, as a function of the number of modes (M), for the Trachea-LMB-RMB and RMB-RUL-BI regions.	174
6.10	Fisher mapping of first 10 modes plotted against 11th mode for the trachea-RMB-LMB region	174
6.11	Fisher mapping of first 10 modes plotted against 11th mode for the RMB-RUL-BI region	176
6.12	ROC curves for the two regions of interest using the training and test sets (the error bars show the 95% confidence interval for the sensitivity)	176
6.13	Classification based on 3-branch combinations (the error bars show the 95% confidence interval for the sensitivity)	177
6.14	Including obstructed branches in airway region classification (the error bars show the 95% confidence interval for the sensitivity)	178
6.15	AUC as a function of number of modes used in the classification. . .	179

6.16	ROC curves for the two regions of interest using the training and test sets (the error bars show the 95% confidence interval for the sensitivity). a) and b) Average diameter based features (described in Section 6.2.1.3). c) and d) Features based on observed stenosis (described in Section 6.2.1.3). e) and f) PDM derived features (repeat of Figure 6.12).	182
7.1	The proposed projection method to segment and model airway variation in chest radiographs. The 3D model is projected onto processed chest radiographs. The model parameters are then optimised with an energy function to obtain a locally optimal fit of the 3D model. The optimisation procedure updates the 3D shape, projected shape and energy function	191
7.2	The viewing direction (E) and outward facing normal (N) of a triangle, which are used to define the direction of the face	192
7.3	Airway silhouette detection and projection. Black points show the detected silhouette vertices on the trachea, LMB and RMB and the projected points onto a surface. The locations of the skeleton (blue points) and the branch start and end points (green points) are also projected.	194
7.4	Procrustes analysis of the 4 manual landmarks of the image (defining the branch start and end positions) with the projected branch landmark points. Green dots show the manual landmarks on the airway of the image and red squares show the aligned landmarks from the 3D model	195
7.5	Improvement in airway visibility after the local normalisation filter is applied to the image	196
7.6	Comparing greyscale of kernels inside and outside the airway region for each silhouette point	197
7.7	The optimisation method calculates a new \mathbf{b} at each step. A new 3D model is generated for this \mathbf{b} and projected onto the surface to recalculate \mathbf{E} . These steps are repeated until a local minimum is reached.	200
7.8	An example of the optimised 3D shape and projection	200
7.9	The manually drawn outline of an example airway	202
7.10	Case 11 showing fitting of the aligned shape model to the airway. a) and b) show the silhouette points before and after the shape model fit overlaid on the gradient image. c) and d) show the same fit but with the manual annotation overlaid. Arrows indicate the direction to the closest point on the manual annotation to measure the error.	206

7.11	Case 33 showing fitting of the aligned shape model to the airway. a) and b) show the silhouette points before and after the shape model fit overlayed on the gradient image. c) and d) show the same fit but with the manual annotation overlayed. Arrows indicate the direction to the closest point on the manual annotation to measure the error.	207
7.12	Airway region for all test images after local normalisation	208
7.13	Airway region with segmentation outline (green) and manual annotations (red) for all test images	209
B.1	Example of multithreshold pruning steps	225
B.2	Example T1	226
B.3	Example T8	227
C.1	Example quality assessment for a single airway	231

List of Tables

2.1	Evaluation measures for the twenty cases in the test set. The table columns are as follows: <i>branch count</i> is the number of branches detected correctly, <i>branch detected</i> is the proportion of branches detected compared to the ground truth, <i>tree length</i> is the sum of the length of all correctly detected branches, <i>tree length detected</i> is the tree length compared to the ground truth, <i>leakage count</i> is the number of correct regions bordering incorrect regions, <i>leakage volume</i> is the volume of regions wrongly detected, and <i>false positive rate</i> is the fraction of wrongly detected regions out of all detected regions.	71
2.2	Segmentation methods evaluated in EXACT'09 (Lo et al., 2009c) . . .	72
3.1	Threshold values for pruning	97
5.1	Leave-one-out cross validation of SVM for the RMB. The class labels are as follows: no missing branches (class 1), missing RUL (class 2) and missing BI (class 3)	144
5.2	Manually and automated labelling of obstructions and disconnected branches (no. of airways)	146
6.1	Mean age and std deviation of each dataset	162
7.1	Selecting the optimisation parameters α_e , α_g and r using the training set. Lowest mean error was found at $\alpha_e=0.4$, $\alpha_g=0.6$ and $r=0.003$. . .	203
7.2	The mean error after alignment (<i>initial error</i>) and SSM optimisation (<i>final error</i>). <i>Percentage change</i> shows the change after SSM optimisation	204
7.3	The mean error after alignment (<i>initial error</i>) and SSM optimisation (<i>final error</i>). <i>Percentage change</i> shows the change after SSM optimisation	205

B.1	Manual interaction required for 21 of the 90 test cases using <i>keep</i> , <i>remove</i> or <i>insert</i> to edit skeleton. Each number is the unique label of the branch that requires interaction.	226
C.1	V_{dif} for closest point and mesh preserving terms (F1, F2)	229
C.2	V_{open} for closest point and mesh preserving terms (F1, F2)	229
C.3	V_{dif} for entire method (F1, F2, F3) ($\beta = 0.5$)	229
C.4	V_{open} for entire method (F1, F2, F3) ($\beta = 0.5$)	230
C.5	V_{dif} for entire method (F1, F2, F3) ($\beta = 1.0$)	230
C.6	V_{open} for entire method (F1, F2, F3) ($\beta = 1.0$)	231
C.7	Accuracy of the registered template mesh compared to the original mesh for each training case in the trachea-LMB-RMB region. Cases with obstruction are automatically excluded (discussed in more detail in Chapter 6).	232
D.1	Principal components used for classification of the Trachea-RMB-LMB region. Rows are cases, columns are the 11 components used for classification. Gosh non-TB training set: C1-C37, Tygerberg TB training set: C38-C78, Tygerberg TB test set: C79-C114, and Tygerberg non-TB test set: C115-C160.	234
D.2	Principal components used for classification of the RMB-RUL-BI region. Rows are cases, columns are the 11 components used for classification. Gosh non-TB training set: C1-C35, Tygerberg TB training set: C36-C64, Tygerberg TB test set: C65-C96, and Tygerberg non-TB test set: C97-C137.	239

Notation

<i>TB</i>	Tuberculosis	22
<i>WHO</i>	World Health Organisation	23
<i>LMB</i>	Left Main Bronchus	27
<i>RMB</i>	Right Main Bronchus	27
<i>RUL</i>	Right Upper Lobe Bronchus	27
<i>BI</i>	Bronchus Intermedius	27
<i>RLL</i>	Right Lower Lobe Bronchus	27
<i>LUL</i>	Left Upper Lobe Bronchus	27
<i>LLL</i>	Left Lower Lobe Bronchus	27
<i>FTB</i>	fibre-optic trachea-bronchoscopy	33
<i>MPR</i>	Multi-planar reconstruction	33
<i>MinIP</i>	Minimum intensity projection	33
<i>SSD</i>	Shaded surface display	33
<i>VR</i>	Volume rendering	35
<i>CAD</i>	Computer-assisted detection	37
<i>HU</i>	Hounsfield Units	45
<i>SE</i>	Structuring Element	47
<i>DICOM</i>	Digital Imaging and Communications in Medicine standard	51
<i>CM</i>	Centre of mass	54
<i>FPR</i>	False positive rate	66
<i>ICP</i>	Iterative Closest Point	105
<i>m - rep</i>	Medial representation	105
<i>TPS</i>	Thin-plate Spline	114
<i>ICA</i>	Independent Component Analysis	152
<i>FDA</i>	Fisher Discriminant Analysis	152
<i>ASM</i>	Active Shape Model	152
<i>PDM</i>	Point Distribution Model	152
<i>COPD</i>	Chronic Obstructive Pulmonary Disease	154
<i>CV</i>	Cross Validation	164
<i>LOOCV</i>	Leave-one-out Cross Validation	164
<i>Se</i>	Sensitivity	165
<i>Sp</i>	Specificity	165

<i>TP</i>	True Positives	165
<i>TN</i>	True Negatives	165
<i>FP</i>	False Positives	165
<i>FN</i>	False Negatives	165
<i>ROC</i>	Receiver Operating Characteristic	165
<i>AUC</i>	Area-under-the-curve	165
<i>CI</i>	Confidence Interval	165
<i>S_e</i>	Standard Error	166
<i>BC_a</i>	Bias corrected and accelerated	166
<i>AAM</i>	Active Appearance Model	189
<i>PCA</i>	Principal Component Analysis	221
<i>GPA</i>	Generalised Procrustes Analysis	219
<i>SVD</i>	Singular Value Decomposition	220
<i>LDA</i>	Linear Discriminant Analysis	222
<i>SVM</i>	Support Vector Machines	222

Related publications

Below is a list of publications related to the thesis and authored or co-authored by the author of this thesis:

- B.J. Irving, P. Taylor, and A. Todd-Pokropek. 3D segmentation of the airway tree using a morphology based method. In *Proc. of Second International Workshop on Pulmonary Image Analysis*, pages 297–307, 2009
- B. Irving, A. Todd-Pokropek, and P. Taylor. Airway segmentation and analysis from computed tomography. In A. El-Baz and J.S. Suri, editors, *Lung Imaging and Computer Aided Diagnosis*. CRC press, 2011c
- B. Irving, P. Goussard, R. Gie, A. Todd-Pokropek, and P. Taylor. Identification of paediatric tuberculosis from airway shape features. In *Proc. Medical Image Computing and Computer Assisted Interventions*, volume 14, pages 133–140, 2011a
- B. Irving, P. Goussard, R. Gie, A. Todd-Pokropek, and P. Taylor. Segmentation of obstructed airway branches in ct using airway topology and statistical shape analysis. In *Proc. IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 447–451, 2011b
- P. Lo, B. van Ginneken, J.M. Reinhardt, T. Yavarna, P.A. de Jong, B. Irving, C. Fetita, M. Ortner, R. Pinho, J. Sijbers, M. Feuerstein, A. Fabijanska, C. Bauer, R. Beichel, C.S. Mendoza, S. Zayed, R. Wiemker, J.Lee, A/P. Reeves, S. Born, O. Weinheimer, E.M. van Rikxoort, J. Tschirren, K. Mori, B. Odry, D.P. Naidich, L. Hartmann, E.A. Hoffman, M. Prokop, J.H. Pedersen, and M. de Bruijne. Extraction of Airways from CT (EXACT’09). *IEEE Transactions on Medical Imaging*, 2012. (In Press)
- S. Andronikou, B. Irving, T. Hlabangana, T. Pillay, P.Taylor, P. Goussard, and R. Gie. Technical developments of post processing of paediatric airway imaging. *Pediatric Radiology*, 2012. (In Press)

A paper on later sections of this thesis is also in preparation.

1

Introduction

Detection of pulmonary tuberculosis (TB) in children is challenging, and a large number of cases are missed. Due to this difficulty, several tests are combined for diagnosis – including chest imaging. Lymphadenopathy (pathology of the lymph nodes) is a key sign of paediatric TB, but due to poor visibility in medical images, lymphadenopathy is commonly identified by the deformation of surrounding structures, such as the airways.

Automated analysis of the airway has the potential to assist in the detection of abnormal airway shape, the detection of TB and provide additional information to the clinician when viewing multi-slice CT. There is also potential to improve the visualisation and analysis of the airways in 2D radiographs.

This thesis develops methods to analyse both airway structure and shape variation in CT scans, which can be used to assist in the detection of signs of TB. This research includes segmentation of the airways, detection of the structure, registration of multiple airway shapes, detection of obstructions and classification of airway shape. These methods also have potential to be applied to other areas of airway research

and detection of other airway pathology, as well as 2D radiographs to assist in the segmentation of the airways

In this chapter, paediatric TB, the signs of paediatric TB in airway imaging, and the development of an airway analysis method are introduced. An outline of the thesis is presented in Section 1.6 and the contributions of this research are discussed in Section 1.7.

1.1 Tuberculosis

TB has a considerable effect on the quality of life in developing countries – particularly in Africa, and is the second biggest cause of death from a single infectious agent (WHO, 2012). The World Health Organisation (WHO) estimates the incidence of TB to be 276 (new cases/100 000 pop/yr) in Africa, higher than any other region in the world (WHO, 2012). The second highest is in South-East Asia with an incidence of 193. South Africa, a country with one of the highest burdens of TB, has an estimated incidence of 981 and an estimated prevalence (the number of cases at a given time) of 795 per 100 000, as estimated by the WHO for 2010.

TB is spread from an infected person by the bacterium *Mycobacterium Tuberculosis* in aerosol droplets, for example by coughing, and a person is infected by inhaling the bacterium. There are a number of stages to TB infection and these depend on whether the patient has been previously exposed to TB. Primary tuberculosis occurs when a patient has not been previously exposed. Post-primary TB is a recurrence of TB after a patient has been previously infected and then undergone an asymptomatic period. Progressive TB is the progression of the disease from primary to post-primary TB without a break (Collins and Stern, 2007; Grainger and Allison, 1986; Marais, 2007).

Primary TB is the first infection of tuberculosis and is, therefore, the most common form affecting children. The *Mycobacterium* is inhaled, travels through the airways and enters the alveoli. There is then development of regions of consolidation –that is, a mass that is a combination of bacteria and inflammation from the bodies immune response. The consolidation is known as the primary focus (or Ghon focus). At this point, the spread of TB is generally contained and the patient does not develop the disease, but the patient will have developed antigens in response to the bacterium and will test positive to an antigen skin test (Gie, 2003).

If the infection is not contained at this point, it will spread via the lymphatic system and infect lymph nodes, typically the hilar, paratracheal or mediastinal nodes. The combination of the Ghon focus and the infected lymph node is known as the Ghon complex. Lymphadenopathy – the pathology of the lymph nodes – causes the lymph nodes to enlarge which can lead to deformation of the airways (Gie, 2003; Grainger and Allison, 1986).

Children have more malleable airways than adults and lymphadenopathy is a common feature of pulmonary TB in children. Therefore, airway compression from the lymph nodes is common, such as displacement of the trachea and bronchi, as well as complete or partial obstruction of the bronchi. Obstruction of the bronchi can, in turn, lead to a hyper-inflated or collapsed lobe (Gie, 2003; Goussard and Gie, 2007; Grainger and Allison, 1986). The lymph nodes can also erode through the pleura causing pleural effusion or erode into the bronchi (Gie, 2003; Goussard and Gie, 2007; Grainger and Allison, 1986). However, pleural effusion is rare in primary TB (Andronikou and Wieselthaler, 2004). Miliary TB, characterised by small nodules scattered throughout the lungs, can also sometimes occur in primary TB (Grainger and Allison, 1986).

1.2 Paediatric Pulmonary Tuberculosis

While there has been considerable focus and recent advancements in the detection of TB in adults, paediatric TB is poorly detected and under reported (Nelson and Wells, 2004; Sandgren et al., 2012; Swaminathan and Rekha, 2010). However, TB is a major cause of childhood morbidity and mortality, particularly in developing countries. There is, therefore, a need to improve existing diagnostic techniques and develop new methods of detection to supplement the existing ones.

A considerable proportion of the TB burden is in children. Nelson and Wells (2004) estimate that approximately 11% of the 8.3 million TB cases worldwide in 2000 were children. Cases are typically under reported; the WHO reports only smear-positive cases and most children with TB under 5 years of age are smear negative (Nelson and Wells, 2004).

Current methods of detecting tuberculosis in adults have considerably poorer accuracy in children. Diagnosis of TB is substantiated in less than 50% of paediatric cases

and this can be even poorer in developing countries (Andronikou and Wieselthaler, 2004; Coulter, 2008). The limitations of the current methods for diagnosing paediatric TB are as follows:

1. **Symptoms** are vague and shared by other diseases (Andronikou et al., 2004; Coulter, 2008; du Toit et al., 2002; Gie, 2003; Salazar et al., 2001). Symptoms include cough, weight loss and fever; the patient can also be asymptomatic. Symptoms have been used successfully to screen for TB in children older than three years (sensitivity 82.3%, specificity 90.2%), however the results are much poorer for younger children (sensitivity 51.8 %, specificity 92.5 %) and children with HIV (sensitivity 56.2 %, specificity 61.8 %) (Marais et al., 2005). This is made more difficult by the similarity in symptoms between patients with and without TB if they are infected with HIV (Andronikou et al., 2004; Coulter, 2008; du Toit et al., 2002; Gie, 2003; Salazar et al., 2001).
2. **Contact with an infected adult** is used as an indicator, but this is often difficult to establish if the infected adult is not in the child's home (Gie, 2003).
3. **Direct detection of the bacterium** is the "gold standard" but there are difficulties acquiring the samples from children. A culture of *M. Tuberculosis* from sputum i.e. material such as phlegm brought up from the lungs, is the most common method of detecting TB in adults. However, this method is not effective in children (Andronikou and Wieselthaler, 2004; Gie, 2003; Marais et al., 2006; Salazar et al., 2001). In a study by Salazar et al. (2001), only 19% of infants and 52% of 2 to 4 year olds that were diagnosed with definite or probable pulmonary tuberculosis had a positive sputum sample or gastric aspirates. Similarly, Marais et al. (2006) report a positive test of between 30% and 40% for cultured samples and less than 15% for sputum smear microscopy for children with TB. This is because a nonproductive cough makes it difficult to obtain sputum, and childhood TB is paucibacillary so it yields a low bacterium count (Gie, 2003). Gastric washing can be used to test for bacilli in swallowed sputum. However, this is difficult to obtain and also has a low yield of bacterium (Andronikou et al., 2004; Salazar et al., 2001).
4. **Antigen skin tests** are common but, as discussed earlier, if a person has

been exposed to TB, and even though the infection is contained, the patient will still develop sensitivity to Myobacterium Tuberculosis and test positive in an antigen skin test. Antigen tests will also give false negatives in severely ill patients (Coulter, 2008; Gie, 2003).

5. **Interferon- γ testing** is also used, but cannot distinguish between latent or active TB (Coulter, 2008).
6. **Chest radiography** stills plays a key role in the detection of pulmonary TB. However, a major problem in developing countries is poor quality paediatric chest X-rays and inter-observer differences in the analysis. This can be due to clinicians who are not trained in the task (Coulter, 2008).
7. New diagnostic tests such as direct fluorescence sputum microscopy and faster liquid culture methods have similar problems detecting TB in children because of the low sputum yield and paucibacillary nature of the disease in children (Sandgren et al., 2012).

Due to the difficulty of confirming TB, a combination of tests is generally used: contact with an adult case, a positive tuberculin test, and signs in a chest X-ray. This is a useful procedure but is less effective in endemic areas because contact with an infected adult is difficult to establish and tuberculin only indicates exposure (Marais, 2007; Marais et al., 2006). Sandgren et al. (2012) highlight the need for further research into the diagnosis and treatment of paediatric TB, and highlight the increasing clinical interest in paediatric TB.

1.3 Lymphadenopathy and airway deformation

Paediatric TB could benefit from additional tests, and better analysis of airway deformation has the potential to assist in the detection of TB. This section discusses signs of lymphadenopathy caused by tuberculosis and its effect on the airways.

1.3.1 Airway anatomy

The thoracic cavity is enclosed by the ribs, sternum and spine and is divided into two pleural cavities and a mediastinum separating the two. The pleural cavities contain the lungs and the mediastinum includes the heart, trachea and oesophagus. The lungs are divided into separate lobes by fissures. The right lung is divided into 3 lobes and the left into 2 lobes. The heart occupies more space on the left side and so the right lung is larger than the left. The left lung does not have a middle lobe but instead has an extra attachment onto the upper lobe, the lingula (Moses et al., 2005).

The airway tree supplies oxygen to the lungs and begins at the trachea. The trachea bifurcates to form the left main bronchus (LMB) and right main bronchus (RMB), which enter the lungs. The RMB bifurcates into the right upper lobe bronchus (RUL), supplying that lobe, and the bronchus intermedius (BI), which further divides into the right middle lobe bronchus (RML) and the right lower lobe bronchus (RLL). As there is no middle lobe on the left side, the LMB bifurcates into the left upper lobe bronchus (LUL) and the left lower lobe bronchus (LLL). The airways continue to divide into smaller and smaller bronchi and bronchioles until they terminate at the alveoli where gaseous exchange takes place. The main divisions of the airway are shown in Figure 1.1 where names of the bronchi describes the lung region that is supplied (Moses et al., 2005).

The lungs also have a large blood supply. The pulmonary artery carries deoxygenated blood from the heart to the lungs where it is oxygenated and leaves via the upper and lower pulmonary veins. Bronchial arteries arising from the aorta supply the lung tissue with oxygenated blood. Lymphatic vessels pass through lymph nodes that are present in the lungs and then drain into the mediastinum (Moses et al., 2005).

1.3.2 Lymphadenopathy

Mediastinal and hilar lymphadenopathy are strong indicators of childhood TB. However, there are inherent difficulties detecting lymphadenopathy from chest radiographs because of its radiotranslucence, and CT remains the “gold standard” for detecting lymphadenopathy (Andronikou et al., 2004; Swingler et al., 2005). Even in CT it is often difficult to distinguish the node from surrounding tissue other than from

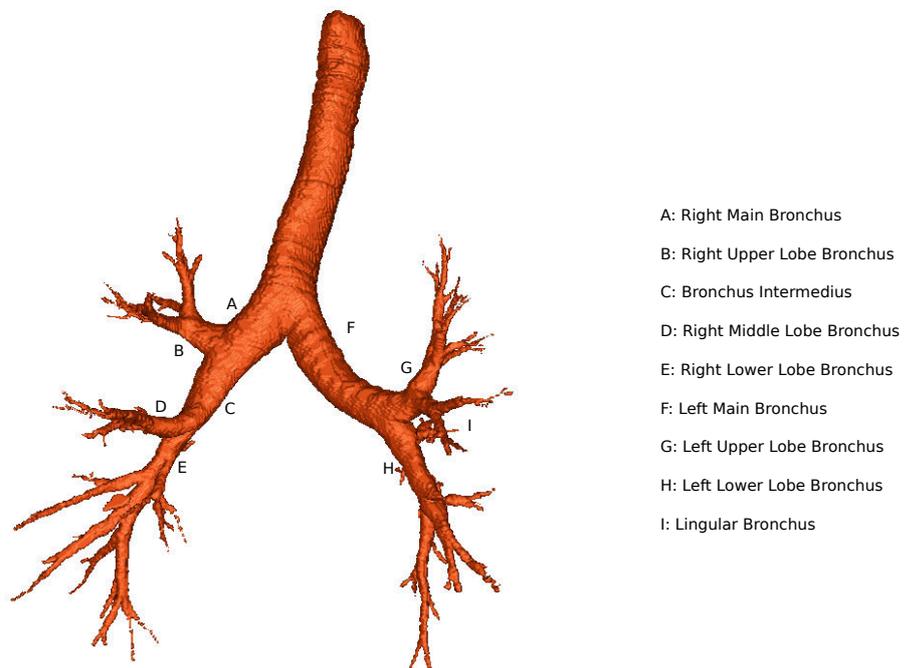


Figure 1.1: An airway segmentation with labelled branches

deformation of surrounding structures.

Leung et al. (1992) performed a study involving 142 paediatric patients, with a mean age of 5.9 years, and definite or probable TB. They noted that hilar or mediastinal lymph node enlargement was present in 92% of cases. Of particular interest is that they noted that 100% of patients under 3 years of age had enlarged lymph nodes. Andronikou et al. (2004) performed a study involving 100 children with a mean age of 22 months who were diagnosed with TB using the WHO criteria. Of these children, 92 had visible lymph nodes on CT and of these 46 patients had nodes greater than 1 cm and, therefore, considered definitely pathological. These studies suggest that lymphadenopathy is present in the majority of paediatric TB patients and identification of the lymph nodes can be considered a key sign of paediatric TB. In fact, according to du Toit et al. (2002) lymphadenopathy is one of the few distinguishing signs that can be used to diagnose paediatric TB.

Some previous studies have used chest radiographs to identify lymphadenopathy (Leung et al., 1992) but X-ray images have been found to be a poor method of detecting lymphadenopathy (Andronikou and Wieselthaler, 2004; Andronikou et al., 2004; du Toit et al., 2002; Swingler et al., 2005). It is important to note that lymph nodes are not clearly visible in X-ray or CT, but deformation and displacement of other structures allow detection and measurement of lymph nodes in CT.

1.3.3 Airway involvement

Airway involvement due to lymphadenopathy is relatively common. Studies indicating 35% (Andronikou et al., 2004) and 38% (Theart et al., 2005) of paediatric TB cases have some airway involvement. Deformation can occur due to one enlarged lymph node, or compression can be due to an enlarged lymph node on both sides or a lymph node and a vessel, and can be used as an indicator of TB (Andronikou et al., 2004; Goussard and Gie, 2007).

The trachea, LMB, RMB and BI are generally affected by lymphadenopathy, which can compress the airways externally or erode into the airways (Gie, 2003; Goussard and Gie, 2007). This generally results in partial compression of the airways. If the airway is partially obstructed then a “ball-valve” effect may develop where the airway becomes blocked only on exhalation, leading to hyperinflation of the lung.

If a bronchus is completely obstructed then lobar collapse can occur. According to Goussard and Gie (2007), deviation of the trachea can be an indication of paratracheal lymph gland compression and if both main bronchi are narrowed this can indicate enlarged subcarinal lymph nodes.

Figures 1.2, 1.3 and 1.4 show minimum intensity projections (MinIP) of a coronal view of the airways with lymphadenopathy and airway stenosis. Figure 1.4 shows the clearest airway stenosis and low density ring-enhancing nodules are also visible.

1.3.4 Other radiographic signs of paediatric TB

There are a number of other radiographic signs besides airway deformation that are associated with paediatric TB. In radiographs the following signs were observed: air space consolidation was seen in 80% of patients; hilar lymphadenopathy was seen in 72% of patients – however these are generally vague and unclear; hyperinflation of the lung was seen in 32% of patients and nodular lesions were seen in 28% of patients (Kim et al., 2006). In CT, signs were as follows: consolidation was seen in 100% of patients (and cavitation within the consolidation was seen in 30% of patients); hilar lymphadenopathy was seen in 100% of patients; hyperinflation of the lung was seen in 47% of patients and nodular lesions were seen in 41% of patients.

Consolidation appears in most cases and is visible in radiographic and CT examinations. This can be identified by regions of greater attenuation in the lungs. However, there are a number of causes for consolidation and, therefore, the specificity of consolidation for detecting paediatric pulmonary TB is low.

1.3.5 Other diseases affecting the trachea and bronchi

Congenital cardiac disease can cause airway obstruction by compression from enlarged cardiovascular structures. Lee et al. (2002) assessed 52 paediatric patients using bronchoscopy. Each patient in the study had both airway and congenital cardiac disorders. They found that 67 % of patients had external compression of the lower airways. Of these cases 11 % was compression of the trachea, 15 % was compression of the right main bronchus (RMB) and 67 % was compression of the left main bronchus (LMB). Where there was RMB involvement, 75 % of these cases also had LMB

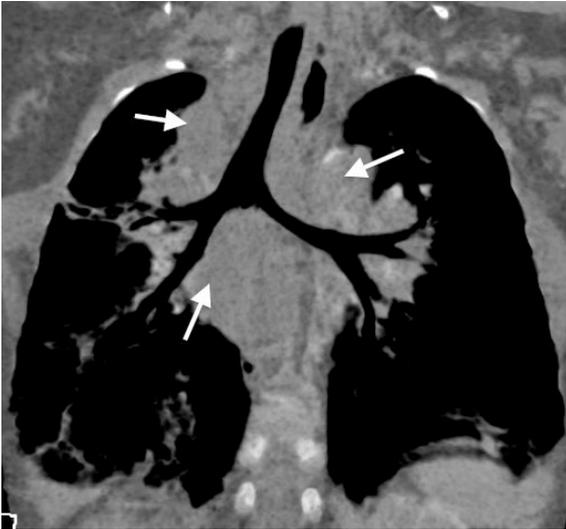


Figure 1.2: Coronal minimum intensity projection (MinIP) image with arrows indicating lymphadenopathy. Image courtesy of Prof S. Andronikou, Radiology Department, University of Witwatersrand.

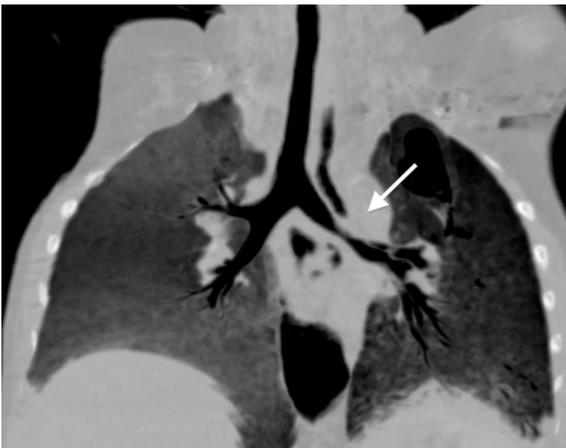


Figure 1.3: Coronal minimum intensity projection (MinIP) image with arrows indicating lymphadenopathy. Stenosis of the left main bronchus (LMB) is visible. Image courtesy of Prof S. Andronikou, Radiology Department, University of Witwatersrand.

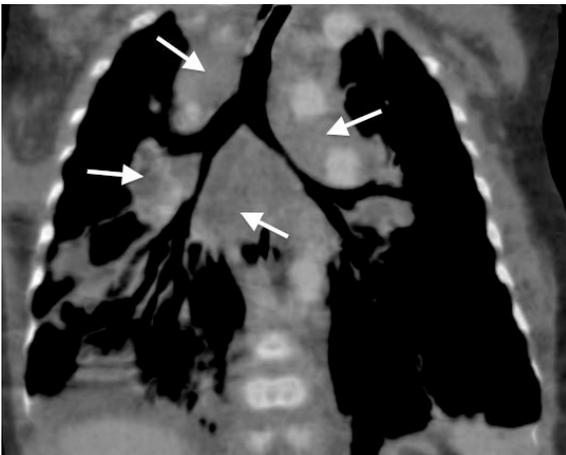


Figure 1.4: Coronal minimum intensity projection (MinIP) image with arrows indicating lymphadenopathy. Stenosis of the left main bronchus (LMB) and bronchus intermedius (BI) is visible. Image courtesy of Prof S. Andronikou, Radiology Department, University of Witwatersrand

involvement. The cause of the compression included dilation of the pulmonary arteries or aorta, or abnormalities of the aortic arch.

Lung cancer causes the largest number of cancer related deaths in the world (Bourke, 2007). Tumours can invade a number of lung structures and can result in airway obstruction. Tracheal tumours are rare but cause increasing airway obstruction and can be inoperable if detected too late. Bronchial tumours are much more common and lead to obstruction of the airways. As the disease progresses this can lead to lobar collapse and consolidation in the lung lobe. Tumours can also affect the hilar, paratracheal and mediastinal lymph nodes, causing enlargement and leading to airway obstruction and displacement (Bourke, 2007; Brewis et al., 1995).

Bronchiectasis is a pathological dilation of the bronchi caused by a chronic bacterial infection of the epithelium in the lung. Chronic infection causes tissue damage which leads to the bronchial dilation (Bourke, 2007).

A large number of other diseases affect the airways and Boiselle et al. (2008) provide an overview, including signs associated with each disease. These include post-intubation stenosis, amyloidosis, inflammatory bowel disease, relapsing polychondritis, rhinoscleroma, sarcoidosis and a number of other diseases. Post-intubation stenosis is a common cause of stenosis of the trachea. Amyloidosis can lead to tracheal and bronchial wall thickening. On rare occasions, inflammatory bowl disease cause tracheobronchial complications including sometimes irreversible stenosis. Relapsing polychondiritis can affect both central and distal airways through inflammation and tracheomalacia. Rhinosclerma can lead to bronchial wall thickening and luminal narrowing as well as nodularity of the mucosa. Sarcoidosis can lead to smooth, irregular and nodular stenosis of the airways. This illustrates the vast array of disease affecting the airways and the potential of the airways to be used as a sign for computer assisted disease detection.

As shown above, tuberculosis is not the only cause of lymphadenopathy or airway deformation. However, in patients under 5 years, other causes of lymphadenopathy are relatively uncommon in endemic TB areas, such as South Africa, and therefore lymphadenopathy is a good indicator of paediatric TB (Gie, 2003). Also, due to the difficulty of diagnosing paediatric TB (confirmed in <50% of cases), clinicians use a combination of several tests including imaging, sputum and tuberculin. The diagnosis of paediatric pulmonary TB partly relies on the identification of lymphadenopathy in

the images (Andronikou et al., 2004). Therefore, our aim is to provide an automated approach to this detection, which would be used in conjunction with other clinical tests.

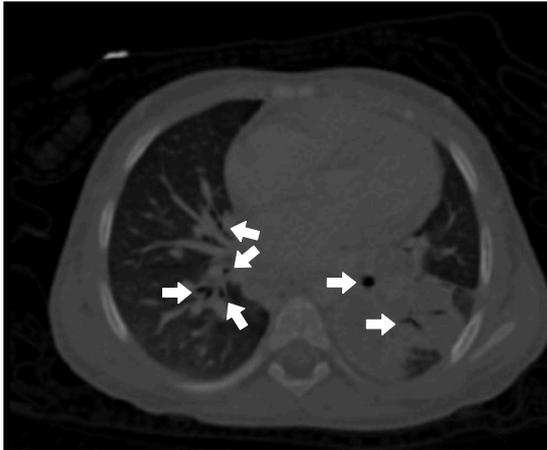
1.4 Airway imaging

CT is established as a key method of evaluating paediatric airway involvement (Andronikou et al., 2012). While fibre-optic trachea-bronchoscopy (FTB) is considered the “gold standard” for evaluating airway compression, the procedure is invasive and is performed under general anaesthesia. Modern CT scanners are becoming a viable alternative for identifying stenosis because of the high resolution images and no need for sedation – due to the speed of the acquisition. An additional advantage of CT is the ability to view external causes of airway pathology at the same time as viewing a rendering of the airway.

CT produces a series of axial slices over a patient volume. There are a number of common methods that are used by clinicians to visualise a CT volume – including for visualisation of the airways. Multi-planar reformation (MPR) is a common technique that reconstructs 2D images, in a number of orientations, from the CT axial slices including the coronal, sagittal and oblique views. Figure 1.5 shows example axial, sagittal and coronal slices with airway regions visible.

Minimum intensity projection (MinIP) is a useful method of visualising the airways, in particular. MinIP constructs an image by using the minimum intensity encountered by a ray passing through a volume from a defined viewpoint (Andronikou et al., 2012; Rubin, 2003). This technique highlights air filled structures and allows visualisation of multiple airway bronchi that lie on different planes, but it only produces a two dimensional visualisation and, therefore, does not visualise the bronchial relationships. Figures 1.2, 1.3 and 1.4 show MinIP images of the airways and lymphadenopathy. A more complete visualisation of the airway is available in these figures compared to a single coronal slice in Figure 1.5c.

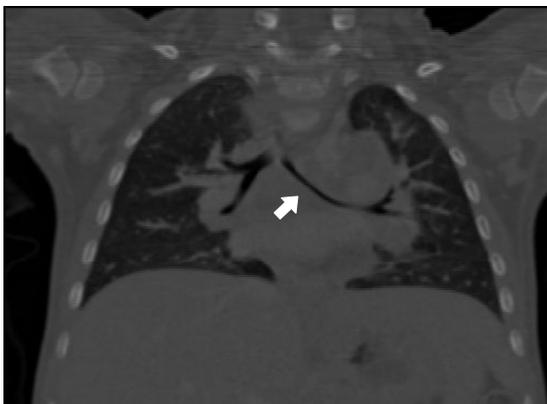
A number of common methods are available for 3D visualisation of the airway. Shaded surface display (SSD) provides a surface rendering using manually defined thresholds. This can be adjusted to visualise the airways. Partial volume effects mean that smaller bronchi have a similar greyscale to the lung region, making it



(a) Axial slice: arrows indicate examples of smaller bronchi.



(b) Sagittal slice: arrows indicate the right upper lobe bronchus and bronchus intermedius.



(c) Coronal slice: arrow indicates the stenosed left main bronchus.

Figure 1.5: Axial, sagittal and coronal slices showing the airways in a chest CT scan

impossible to visualise only the airways with this method. Volume rendering (VR) is similar to SSD but transparency is assigned to greyscale values allowing more detailed visualisations. VR has been used by clinicians to examine stenosed airways from TB (du Plessis et al., 2009). Both VR and SSD can be used to examine the inside of the bronchial lumen and is known as virtual bronchoscopy.

Radiation dose and limited access – particularly in rural areas – makes CT currently infeasible for routine detection of TB (Andronikou and Wieselthaler, 2004). Radiographs do not have these limitations but generally show poor airway visibility. Digital radiography (DR) may offer improved visualisations of the airways – especially Lodox Statscan. Lodox Statscan differs from conventional DR because it is a linear slit scanning X-ray machine. A rotating anode X-ray tube is mounted on a C-arm. A fan beam is emitted via a variable collimator of width 0.4 mm to 1.0 mm. The detector is fixed to the other end of the C-arm. During an examination the C-arm moves across the patient to acquire the image. A scan is completed in less than 13 seconds (Lodox, 2006).

Lodox linear slit scanning radiographic offers improved visualisation of the trachea and main bronchi (Pitcher et al., 2008). In a study evaluating the airway visibility of the airways in 102 patients between 6 months and 13 years, the LMB and RMB could be seen and measured in over 90% of the cases by trained clinicians (Daya et al., 2010; Sanders et al., 2009). Sanders et al. (2009) also suggest that airway shape in X-rays could be used to detect abnormalities caused by lymphadenopathy. The improvements in airway visibility could be due to higher kV used for the standard examinations and lower scatter of the slit scanning technology (Irving et al., 2008). An additional benefit of the linear slit scanning technology is the low radiation dose compared to conventional radiography (Irving et al., 2008; Maree et al., 2007). Figure 1.6 shows a typical Chest AP image acquired using Statscan for a paediatric patient.

While CT is generally not feasible for routine detection of TB, it remains a key technology for detection of lymph nodes and airway visualisation (a key sign of paediatric TB). It is expected that low-dose CT will play an increasingly larger role in the diagnosis and analysis of paediatric TB. Low-dose CT can reduce the radiation to the equivalent of two chest radiographs, and even with more noise introduced in these scans, can still provide the same diagnostic information as conventional CT (Donnelly et al., 2001; Peng et al., 2011). As CT becomes more common in TB

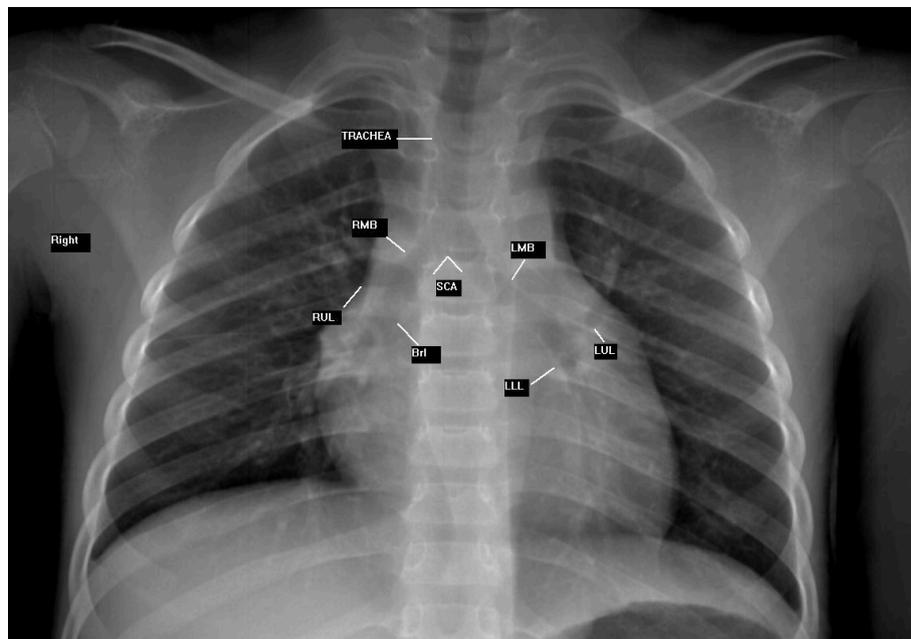


Figure 1.6: Visualisation of the airways using Statscan with labels showing the visible airway branches and the subcarinal angle (SCA). Image courtesy of Virginia Saunders, MIRU, University of Cape Town.

endemic areas, there is potential for radiographs to be replaced for both screening and follow up examinations. This will have a considerable impact on the detection of paediatric TB (Peng et al., 2011). However, with the number of slices produced during each examination, the workload of clinicians will also increase, and thus, along with increased use of CT, automated airway and lymph node analysis will become increasingly important.

1.5 Computer-assisted detection of chest pathology

Computer assisted detection (CAD) has the potential to make a large contribution to the detection of chest pathology. Typically, analysis of radiographs and CT relies on visual inspection of pathology, and CAD can add quantitative analysis to aid diagnosis (Bagci et al., 2012).

Typical CAD methods for detecting infectious disease in the chest region, from CT scans and radiographs, use texture and shape features to enhance regions of pathology, such as nodules, thickening and opacities, and to derive features for classification. Examples of shape features include thickness, area, volume, curvature, irregularity, and examples of texture features include greyscale contrast, variance, entropy and correlation (Bagci et al., 2012; Katsuragawa and Doi, 2007).

CAD for TB detection in CT and radiographs has focussed on post-primary disease commonly found in adults. Radiographic signs of infectious disease, also associated with TB, include: nodules, ground-glass opacities, and consolidation. Nodules have varied appearance such as round, irregular, solid, non-solid, soft tissue or ground glass opacities. Shape image features such as opacity, shape index, curvature and irregularity are often used to extract candidates and provide features for classification. Consolidation and cavitation are commonly detected using texture based features for regions in the lung fields (Bagci et al., 2012; Katsuragawa and Doi, 2007).

van Ginneken et al. (2002b) developed a CAD system to detect TB using local texture analysis of the lung region. The lung fields are segmented in the radiograph and divided into a total of 42 regions for analysis. Multiscale filters using Gaussian derivatives are then applied to the regions. Texture features are computed from the

filtered regions, based on standard deviation, skewness and kurtosis, and used with a k-nearest-neighbour classifier to detect signs of pathology. The method achieved an area-under-the-curve of 0.82 for the receiver operating characteristic (ROC) curve on a dataset of radiographs (147 images with textural abnormalities and 241 normal images).

As discussed earlier, due to the nature of primary pulmonary tuberculosis (common in young children), signs detected in previous CAD methods are uncommon in paediatric TB, and instead, lymphadenopathy and airway deformation are important for detection in chest radiographs and CT. Consolidation is also common in paediatric TB but has a low specificity because there are a number of other common causes.

There is potential to use automated airway shape analysis in the assessment of paediatric TB. An airway shape derived method would only be effective for the approximately 40% of paediatric TB patients showing airway involvement. However, due to poor detection of existing methods, this could be used alongside other diagnostic procedures to target cases with airway deformation and improve the detection of TB. An important part of the diagnostic routine for TB patients with suspected airway involvement is to undergo CT and bronchoscopy to quantify the severity and cause of airway involvement (du Plessis et al., 2009). Therefore, an airway analysis tool would also be useful to assess the severity of the airway involvement.

As discussed earlier, lymphadenopathy has poor visibility in chest radiographs but can be seen in a chest CT. Of the 100 paediatric TB cases evaluated, Andronikou et al. (2004) detected the presence of lymph nodes in 92 cases, but only 46 had nodes greater than 1 cm, which were considered definitely pathological, and 35 cases had airway involvement. Studies that detected lymphadenopathy established a consensus from a number of clinicians on the presence of lymph nodes based on features such as the lower attenuation of the centre of a node, peripheral ring-enhancement, presence of calcification and deformation of surrounding structures (Andronikou et al., 2004; Kim et al., 2006). These features also have potential to be used for detecting signs of TB in 46% of patients but are only applicable to CT and not X-ray. Airway shape is a key sign of paediatric TB that can be used in both CT and radiographs to aid detection of TB as well as general airway analysis. Inclusion of additional texture features based on lymphadenopathy would further improve detection in CT but this is unlikely in radiographs.

Therefore, in this thesis a novel method that uses airway shape change to detect signs of paediatric TB is introduced. However, no previous CAD algorithm has used airway shape to detect TB. Most variation is seen in the larger bronchi including the LMB, RMB, BI and Trachea, and variation can include deformation and stenosis, varying from mild to severe. Because of the complexity of the shape changes, branch cross sections may not be enough and a statistical model is built for regions of the airway to model pathological shape change.

The purpose of this model is to distinguish normal airway shape variation from pathological shape variation. The trachea and bronchi change shape during breathing and, because it is not possible to perform breath-hold CT scans for young patients, breathing will affect the appearance of the airways in CT. However, variation of the airways due to breathing is expected to be distinguishable from pathological variation. Holden and Ardran (1957) observed that the trachea and bronchi lengthen and dilate slightly during quiet inspiration, which can be distinguished from local narrowing, narrowing of certain branches relative to others and pathological deformation.

This method focusses on detection of abnormal variation of the airway in CT examinations to detect TB. Applying this method to radiographs will extend the application to routine screening. However, airway visibility in radiographs is poor. As part of this work, a 3D model is used to assist in the segmentation of 2D radiographs. A 3D model derived from CT images with far better airway visibility can be used to constrain the 2D airway segmentation in radiographs. There is potential to extend this analysis to detecting pathological airway shape change in radiographs.

This thesis describes a framework for automatically segmenting, extracting the structure of the airway and analysing the shape variation in a dataset of normal CT scans compared to CT scans with airway pathology. This method has the potential to provide additional information to the clinician regarding airway deformation and stenosis, for patients with potential airway involvement, undergoing CT examinations. A classification tool based on a dataset of CTs may be used to identify variation that is difficult or time consuming to see in axial slices. The 3D model of normal and abnormal variation includes more detail than can be identified in 2D radiographs and has the potential to assist in the segmentation and analysis of radiographs.

1.6 Thesis outline

This thesis presents research on airway segmentation, airway registration and classification of airway pathology, and is applied to the problem of detecting airway stenosis and deformation caused by lymphadenopathy from paediatric pulmonary tuberculosis in CT images. Chapters 2 and 3 segment and extract the structure of the airway to allow more advanced analysis in Chapters 4 - 7. An outline of each chapter is presented here and a more detailed summary of each chapter is presented in Chapter 8.

Chapter 2 presents a method to segment the airway tree and is used as the basis of further airway analysis. A start point in the trachea is identified and used to initialise the segmentation. Simple thresholding and region growing – with a leak restriction step – is used to segment the trachea, the LMB and the RMB. Then, a morphological filtering method is used to enhance airway regions in the axial, coronal and sagittal slices. A threshold is applied to identify airway locations, and a region growing method which is initialised at the trachea seed point, is used with a leak removal step to extract the final airway. The method was evaluated on an adult dataset as part of the EXACT’09 airway segmentation competition and achieved a fair result compared to other state-of-the-art airway segmentation algorithms. This segmentation algorithm was used to segment paediatric airways for the remainder of this study.

Next, the structure of the airways is identified, as discussed in **Chapter 3**. An existing airway skeletonisation algorithm is used to extract the branching centreline of the airway tree. A pruning method is used to extract the true branches, and the branching structure is used to identify anatomical branches and label the surface mesh of the airways.

Once the structure of each airway has been identified from airway segmentation and centreline detection, it is necessary to register each airway surface mesh, to compare variation between airway trees. **Chapter 4** introduces a method to register regions of the airway tree in a dataset of chest CT scans. Landmarks are projected onto the airway surface, and used to guide the thin-plate-spline warp of a template mesh to each airway. Individual vertices of the template mesh are then aligned to each airway, using a method adapted for stenosed tubular shapes. The final alignment

was evaluated in terms of the accuracy and mesh quality.

Due to severe pathology, a number of cases have main bronchi that have the appearance of being completely obstructed in the CT scan. These cases produce incomplete airway segmentations, and, for visualisation and analysis, it is important to segment beyond obstructions. **Chapter 5** uses the airway structure defined by the labelled skeleton and the statistical shape model of each branch of interest, to detect missing bronchi. The direction of the branch and point of obstruction are then used to search for additional airway regions. The method was evaluated on a limited number of cases with obstruction but showed promising initial results.

Chapter 6 builds on the previous chapters to evaluate paediatric airway stenosis and deformation caused by tuberculosis. A statistical shape model (SSM) is created for regions of the airway tree from a dataset of CT scans showing normal and pathological airway variation. The modes of variation from principal component analysis are used as features for classification. This method is compared to a second method using features derived from measurements of the branch diameter. These different methods were evaluated on a test set of patients with and without tuberculosis from Tygerberg Hospital, South Africa. The SSM derived features could accurately distinguish TB from non-TB paediatric patients using local airway shape changes and illustrate the potential of including automatic assessment of airway shape in the detection of paediatric pulmonary tuberculosis.

This airway assessment method could be applied to 2D radiographs. However, segmentation of the airways in radiographs is a challenging task because of poor visibility and overlapping bone and heart structures. In **Chapter 7**, the projection of a 3D statistical shape model of the airways to 2D is investigated, in order to assist in the segmentation of the 2D region. Silhouette vertices on the mean airway model are detected and projected onto the 2D radiograph. This projection is initially aligned to the airway using four manual annotations. An energy function is created from the filtered image gradient and greyscale-derived features, and is used along with a regularisation term to fit the 3D model the 2D radiograph. The results show that this method can be used to segment the trachea and main bronchi but could be improved by incorporating additional radiographic views.

As a research student at UCL, I was given this project to initiate and develop. Thus, I have developed all aspects of this detection system from scratch with the

support of my supervisors.

1.7 Contributions

The main contributions of this thesis to the fields of medical image computing and computer assisted detection are as follows:

1. **Airway segmentation and evaluation**

An airway segmentation method is developed that extends current morphological based methods, by applying a filter in three dimensions to overcome problems at branch bifurcation. Novel segmentation leak removal steps were also introduced for the main bronchi and the peripheral branches. This method was independently evaluated as part of of the EXACT'09 airway segmentation challenge.

2. **Registration of multiple airway trees to assist in analysis**

A method to register the surface mesh of multiple airway trees is presented, which has not been previously attempted. The method generates landmarks in the airway, warps a template mesh to each airway and performs surface match. Novel methods to consistently select regions of the airway and match the template to stenosed airways are also presented. This method is evaluated in terms of mesh quality and registration accuracy, and yields good results.

3. **Detection of airway tree obstructions and segmentation beyond the obstruction**

When severe pathology is present, some bronchi may appear completely obstructed. A novel approach is presented that builds a statistical shape model of each bronchi and uses the shape to detect obstructions. The CT volume is then searched for bronchi beyond the obstruction. No previous segmentation method has the ability to detect the point of obstruction and segment beyond the obstruction. This method was evaluated on a limited number of cases with obstructions but shows potential to improve general airway segmentation methods, and has other applications beyond airway analysis.

4. Extraction of airway features to represent normal and pathological shape variation

Clinical studies indicate that there is potential to use airway shape changes to assist in the detection of tuberculosis. No studies have attempted to detect TB using airway shape changes or use statistical shape models to model pathology in the airways. Features based on variation of a statistical shape model built on regions of the airway are extracted. These features are compared to bronchial radius derived features.

5. Detection of paediatric tuberculosis in regions of the airway using a statistical shape model of airway shape changes

A classifier was built using the airway shape features representing 90% of the variation in a TB and non-TB dataset. The classifier was applied to regions of the airway tree and performed well on an validation set. This showed that a statistical shape model of normal and abnormal variation of the airway tree can be used as an additional test to assist in the detection of TB and has the potential to be extended to other airway pathology. From a clinical perspective, it confirms previous observations that lymphadenopathy has a high specificity for TB in paediatric patients.

6. Projection of the 3D airway model to assist in the segmentation and analysis of the airways in chest radiographs

Airway segmentation of radiographs is a challenging problem with limited success. In this thesis, the CT derived 3D statistical shape model is used along with features on the radiograph to optimise the fit of a 3D model to the 2D radiograph. This novel approach to 2D airway segmentation shows promising results and there is potential to improve the method by using multiple radiographic views.

2

Automated Airway Segmentation

Abstract. Airway segmentation from CT is a challenging topic given partial volume effects, image variability and difficulty in establishing a gold standard segmentation. This chapter extends existing airway segmentation methods and is used to segment the airways in CT examinations from paediatric patients. The method initially segments the larger bronchi. A 3D morphological filtering approach is then used to enhance and segment the smaller bronchi, and finally a leak removal step is used to remove false branches. The method was evaluated on an adult dataset as part of the EXACT'09 airway segmentation challenge, and used as the basis for further airway analysis discussed in this thesis. The algorithm performed fairly well in the challenge. The method achieved an FPR (lower is better) of 1.27%, out of the 15 teams, 4 teams performed better (lower FPR) and 10 teams performed worse. At this FPR, the branch detected was 43.5% (9 teams performed better and 5 teams performed worse).

Computed tomography (CT) plays an important role in the evaluation of suspected airway involvement in disease, and clinicians have a number of tools at their disposal to visualise that airway. The tools currently available at clinical work stations include: multiplanar reconstruction, maximum and minimum intensity projections,

shaded surface display and volume rendering (VR) (Rubin, 2003). These tools are used for general visualisation of CT volumes and are not specific to airway analysis. VR provides a 3D visualisation of the volume by assigning different colour, intensity and opacity values to the different Hounsfield unit (HU) values of the CT volume. This visualisation can be adjusted to enhance tissues of interest. An internal view with flythrough can also be simulated with virtual bronchoscopy (an example workflow is discussed by (Siegel, 2003)). These methods essentially label the HU range corresponding to the airways. The large difference between the HU value of tissue regions and the HU value of air filled regions means that larger bronchi can be extracted using a simple threshold (Sonka et al., 1996). However, non-airway air filled regions may also be rendered, and smaller bronchi may not be rendered due to partial volume effects. These methods also require considerable manual interaction. Airway segmentation methods have been developed to overcome these challenges and automate airway extraction.

Airway segmentation from CT images is used to determine which voxels in the volume belong to the airway anatomy and which voxels belong to other anatomical structures. This can be a binary classification or probabilities can be assigned to each voxel in the CT volume. Thus, while VR can visualise the airways, by identifying the regions belonging to the airway, airway segmentation can be used to form the basis of more advanced airway analysis methods and be used to improve visualisation of the airway for the clinician. This is particularly important for smaller bronchi that may have similar intensities to the lung regions due to partial volume effects, and require advanced identification methods.

This chapter proposes a morphology based airway segmentation algorithm developed to detect the airways in paediatric CT scans. Current airway segmentation algorithms are outlined in Section 2.1. Requirements for the segmentation method are listed in Section 2.2. The method is outlined in section 2.3, which includes detection of a seed point in the trachea, segmentation of the main bronchi until bifurcation, filtering and segmentation of the smaller bronchi, and false branch detection. The segmentation method is applied to adult and paediatric CT examinations; the datasets are presented in Section 2.4. Section 2.5 presents an evaluation of the segmentation method. The method was evaluated as part of the EXACT'09 airway segmentation competition on an adult dataset (Section 2.5.2) and then applied to a paediatric air-

way dataset (Section 2.5.5). A discussion of the advantages, limitations and future development of this algorithm is presented in Section 2.6.

2.1 Background

Airway segmentation is an active research field and a number of algorithms have been developed. However, this is a challenging field and there is still plenty of room for improvement; one evaluation showed that no algorithm (even with considerable false positives) was able to segment more than 77% of the reference airway branches (Lo et al., 2009c).

Segmentation algorithms tend to have a number of steps in common but often use very different mathematical approaches at each step. An outline of the steps in a typical airway segmentation algorithm is shown in Figure 2.1. The first step defines a seed point. This point is usually in the trachea and is either manually selected or found by identifying a cross-section of the trachea. The second step is often a filter used to enhance the airways before a threshold is applied. Once possible airways have been identified, then a region-growing method is used to select the connected airway region. Rules are often applied to restrict incorrect growth into non-airway regions or remove leaks, either as part of the region-growing procedure or in the following step. Algorithms often focus on one of these steps; some algorithms may ignore the airway enhancement stage and focus on rule-based methods to select the correct airways, whereas others may not apply any rule-based restrictions on the region-growing but use sophisticated airway enhancement filters.

Airway segmentation is a challenging task because the resolution of the CT scan limits the visibility or results in discontinuities of the smaller bronchi. Challenges facing correct airway segmentation include lack of a “gold standard”, partial volume effects and image artefacts. A “gold standard” segmentation is difficult to establish and the most commonly used methods of evaluation include expert annotations, although these are time consuming, phantom studies or comparisons to other algorithms. Partial volume effects have a large influence on the smaller bronchi because a voxel can contain a combination of airway wall and lumen (Sonka et al., 1996). The volume can also be affected by movement artefacts, such as breathing and heart motion, and reconstruction artefacts.

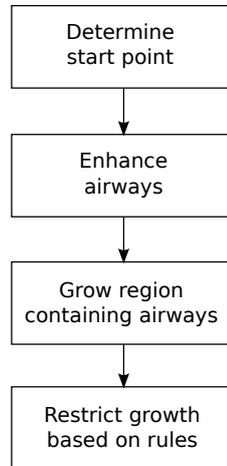


Figure 2.1: General steps in an airway segmentation algorithm

The main airways are easier to segment than smaller bronchi because of a well-defined wall and, therefore, greater contrast between wall and the lumen. Simple thresholding methods are often used to extract them while more sophisticated methods are used for the smaller bronchi. This saves processing time and allows the algorithm to be tailored more specifically to the smaller airways (as demonstrated by Sonka et al. (1996))

A selection of important airway segmentation methods can be loosely grouped into the following categories: adaptive thresholding (Mori et al., 1996), adaptive thresholding with branch analysis (Schlathölter et al., 2002; Swift et al., 2002; van Ginneken et al., 2008), fuzzy connectivity (Tschirren et al., 2005a), rule base segmentation (Park et al., 1998; Sonka et al., 1996), morphology based connection cost (Fetita et al., 2004) and morphological filtering (Aykac et al., 2003; Kiraly et al., 2002; Pisupati et al., 1996). A thorough review of the subject can be found in a chapter on airway segmentation that was written by the author of this thesis and published in *Lung Imaging and Computer Aided Diagnosis* (El-Baz and Suri, 2011).

Aykac et al. (2003) and Pisupati et al. (1996) perform greyscale morphological reconstruction on each CT slice to extract the regional minima and, therefore, enhance the airways. The greyscale morphological reconstruction is applied to each slice using a range of marker images that have been created from greyscale closing the original image with a range of structuring elements (SE) . The SEs are chosen to be a similar

size to the cross section of the airways present in the slice in order to enhance them. The reconstructed image is subtracted from the original image and a threshold is applied. 2D seeded region growing is then applied slice by slice to extract the airways from the binary volume. Pisupati et al. (1996) applied their method to CT images of canine lungs, while Aykac et al. (2003) applied their method to CT images of human lungs.

Morphology based methods have also been applied by Prêteux et al. (1999) and Fetita et al. (2004). Fetita et al. (2004) use an algorithm that calculates the connection cost between points in an image – that is, the smallest threshold that will provide a binary path between two points – to enhance the airways in CT slices. Once enhanced, the images are not thresholded and 3D reconstruction is performed on the greyscale images using a local energy minimisation algorithm. This energy function takes into account the tree topology of the airways as well as the greyscale intensity. There are a number of other non-morphology based methods have been used to segment the airways.

2.2 Segmentation requirements

The airway segmentation procedure, developed in this chapter, is used to form the basis of an airway shape analysis algorithm. Therefore, the key requirements of the airway segmentation are as follows:

- Accurate segmentation of 3 generations of bronchi
 - Requirement to include bronchi of interest for determining lymphadenopathy involvement (Trachea, LMB, RMB, RUL and BI) (see Chapter 6)
 - Child branches of bronchi of interest are also required to determine branch completeness (see Chapters 3 and 5)
- No false branches in the initial generations
- Adaptable to variation in scan parameters
 - Resolution
 - Scanner characteristics

- Adaptable to patient variation
 - Small airway cross sections in patients ages 0 - 5 years
 - Severe stenosis and deformation in some cases
 - Variation in lung opacity due to consolidation, cavitation and hyperinflation of lung lobes

Therefore, an algorithm was developed specifically to segment the airways in paediatric CT examinations with a low false positive rate. A morphological filtering approach was selected because of the ability to tailor the filter to airway size (i.e. smaller paediatric airways) and distinguish airways circular airways from lung regions surrounded by consolidation that may have a similar size a greyscale value.

The underlying morphological filter used in this method is similar to that of Aykac et al. (2003) and Pisupati et al. (1996). However, this method is extends the analysis to 3-dimensions (a requirement for accurately extracting bifurcation points), extends the region growing approach, introduces a separate approach for the main bronchi and develops a false branch removal algorithm. This method is adapted to segment airways in lower resolution paediatric CT examinations. The authors are not aware of other airway segmentation methods that have been applied to paediatric airways.

2.3 Method

An overview of the segmentation algorithm is shown in Figure 2.2. The first step involves: detection of the start slice, detection of a slice containing the trachea, and then detection of the trachea itself. Next, the trachea and main bronchi are segmented using a threshold and region growing approach while restricting the segmentation to the main bronchi to avoid leaking into false regions. Morphological filtering is applied in the axial, coronal and sagittal planes to enhance the smaller bronchi, and finally region growing with leak removal is performed.

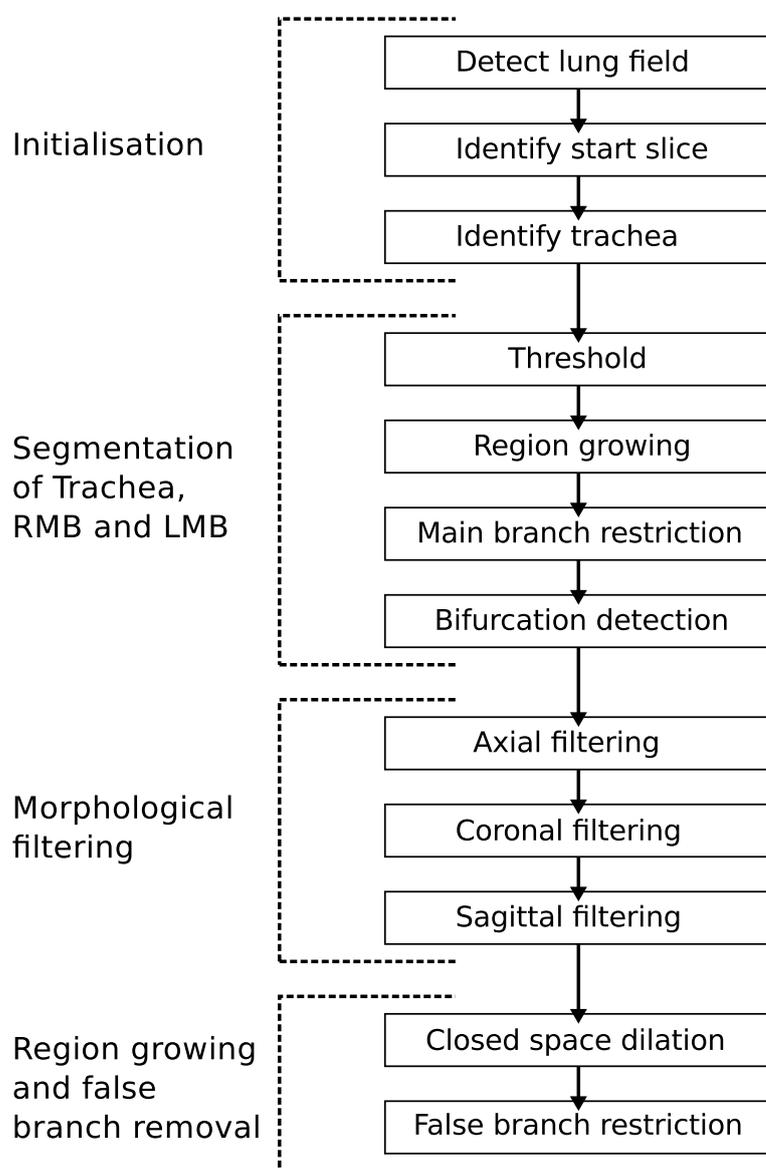


Figure 2.2: Outline of the airway segmentation algorithm. Steps include: identification of the trachea, segmentation of the main bronchi, morphological enhancement and segmentation of the remaining bronchi, and finally a region growing and leak removal step.

2.3.1 Initialisation

2.3.1.1 Lung region detection

This airway segmentation method requires an initialisation point at the start of the trachea. To select this point automatically, the slices are orientated using the DICOM header information and the lung regions in the coronal cross section are roughly segmented by thresholding for air regions in the coronal plane through the centre of the volume (air filled regions were classified as regions below -500 HU)¹. Connected regions in the slice were identified and the largest connected region not connected to the edge of the volume was identified as 1 or both lungs. More sophisticated lung detection methods exist (van Rikxoort et al., 2009) but this method was satisfactory for identifying a start point in the trachea.

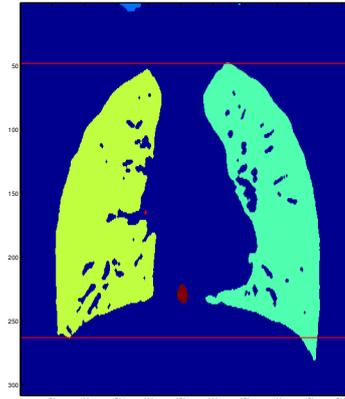
This step is included because the large variation in the region covered by a CT scan, makes it difficult to select an appropriate starting slice without knowing the location of the lung fields. In this algorithm 10 slices above the upper most point of the segmented lung region was chosen as the starting axial slice for the segmentation. Figure 2.3a) shows the connected components in the centre slice that were used to determine the start and end slice. Figure 2.3b) and c) show the start and end slices identified from the lung volume, and the axial view of the start slice.

This method uses the centre coronal section to initialise the scan. Therefore, the start position on the trachea may vary depending on the position of the patient. If improved stability is required for the anatomical location of the start slice then entire lung field could be extracted for improved accuracy. However, the start slice only defines the point from which the segmentation is performed. The system also includes a manual option for selecting the start and end slice, if a different start point is required for another study, as shown in Figure 2.4.

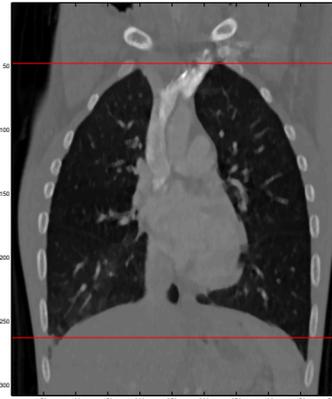
2.3.1.2 Trachea detection

After an axial slice containing a cross section of the trachea has been selected, the trachea is found in that slice by applying a threshold to identify air filled regions.

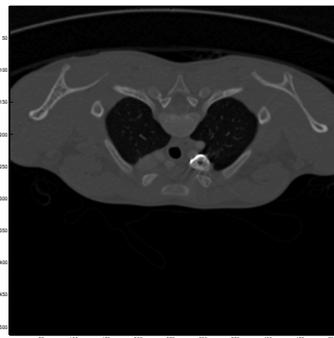
¹Certain cases did not conform to the HU scale and, therefore, if the slice contained only values above -500 HU, the threshold was set to *MinVal+500*



(a) Lung regions are identified from the connected regions in the coronal slice



(b) CT cross section with identified start and end positions



(c) Identified start slice containing a cross section of the trachea

Figure 2.3: Automatic identification of start and end slice positions for the airway segmentation algorithm

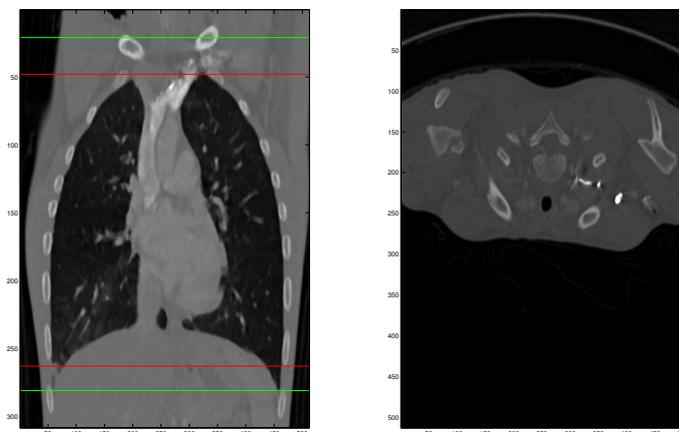


Figure 2.4: Manual selection of start and end slice positions

There are sometimes a number of regions that may have the appearance of a trachea in the slice (see Figure 2.5). Therefore, extracted regions are compared using position, size and compactness in a similar way to that of existing methods (Aykan et al., 2003; Mori et al., 1996)² Compactness (C) was calculated from perimeter (P) and area (A) as $C = \frac{P^2}{A}$. The segmentation of the starting slice is used as a seed for the rest of the segmentation.

2.3.2 Segmentation of Trachea and Main Bronchi

Once the trachea is identified in the initial slice, then thresholding and region growing are applied to the axial slices using the segmentation from the previous slice as seed points. This is applied progressively to each slice from the start of the trachea.

Segmentation using this greyscale thresholding approach is limited to the larger branches. If smaller branches are included in the segmentation, the similar greyscale intensities of the small bronchi and lungs will cause leaking into the lungs. Therefore, for each step, the newly segmented region is evaluated and steps are taken to stop the segmentation of smaller vessels. If the segmented region is more than double the area of the previous slice then a leak is considered to have occurred through

²The area parameter was calculated as the area of the object once the optimal area – chosen to be 400 pixels – was subtracted. The position of the object was defined as the euclidean distance from the ideal position – which was chosen to be (200, 256). These parameters were found from the training set.

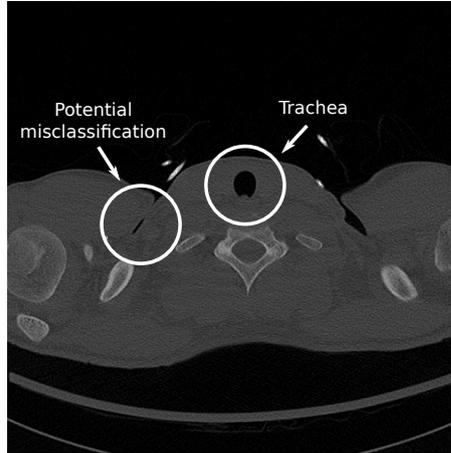


Figure 2.5: Trachea and potential misclassifications in the initial slice

smaller bronchi into the lungs. This leak is excluded by finding the centre-of-mass (CM) and the maximum radius (r_0) of the segmented region in the previous slice, and iteratively reducing the radius and recalculating the CM until a specified compactness ($C = \frac{P^2}{A}$) is reached; the circular region specified by r and CM are used to restrict the segmentation of the current slice (see Figure 2.6).

The position of the bifurcation of the trachea and main bronchi is estimated by labelling the number of connected regions in each slice that have been seeded by one connected region in the previous slice. If two separate regions are seeded by one connected region in the previous slice then bifurcation is considered to have occurred. When the trachea bifurcates, each of the main bronchi are followed and when the main bronchi bifurcate then the procedure is stopped. Due to the smaller branches, for paediatric airway segmentation, the thresholding approach is only applied until the first bifurcation (the start of the RMB and LMB).

2.3.3 Morphological Filtering

To detect bronchi beyond the trachea, LMB and RMB, morphological greyscale reconstruction is applied to all CT slices to extract local minima and, therefore, enhance the airways.

Greyscale reconstruction is an extension of binary reconstruction. Binary reconstruction is the application of successive dilations within objects of a binary image as

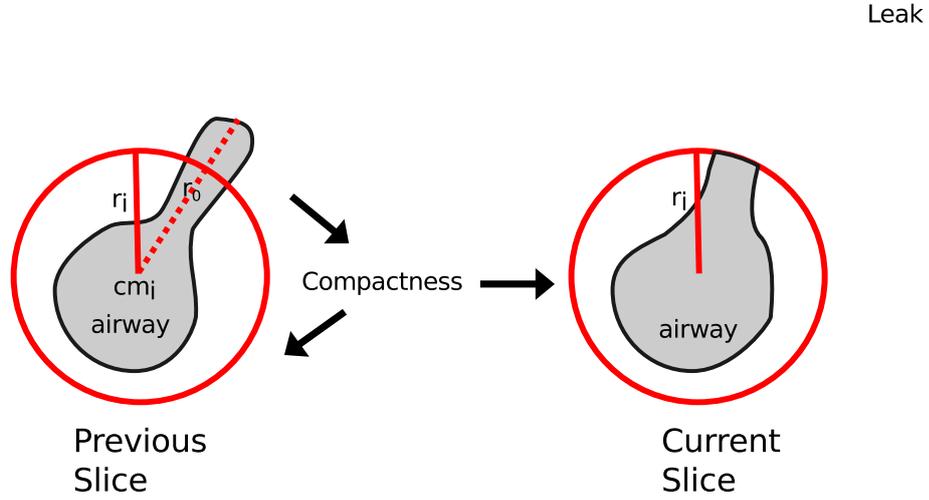


Figure 2.6: Excluding leaks in trachea and main bronchi segmentation by restricting the region growing step based on compactness of the previous slice. r_0 is the maximum radius, r_i and cm_i are the radius and centre-of-mass at the i th iteration

shown below (Vincent, 1992, 1993):

$$\rho_B(X) = \lim_{n \rightarrow +\infty} \delta_B^{(n)}(X) \quad (2.1)$$

where

$$\delta_B^{(1)}(X) = (X \oplus K) \cap B \quad (2.2)$$

and $\delta_B^{(n)}$ is δ_B applied n times:

$$\delta_B^{(n)}(X) = \delta_B \circ \delta_B \circ \dots \delta_B(X) \quad (2.3)$$

The marker (X) is a subset of the mask (B) and is made up of seed points – that is, where the dilation starts in the object (Vincent, 1992, 1993). B is generally the image that is being operated on. K is a SE used in the dilation and \oplus is the dilation operator.

For greyscale reconstruction, a marker and mask image are used where every pixel in the marker image has a pixel intensity less than or equal to the intensity of the corresponding pixel in the mask image (Vincent, 1992, 1993). The marker and mask

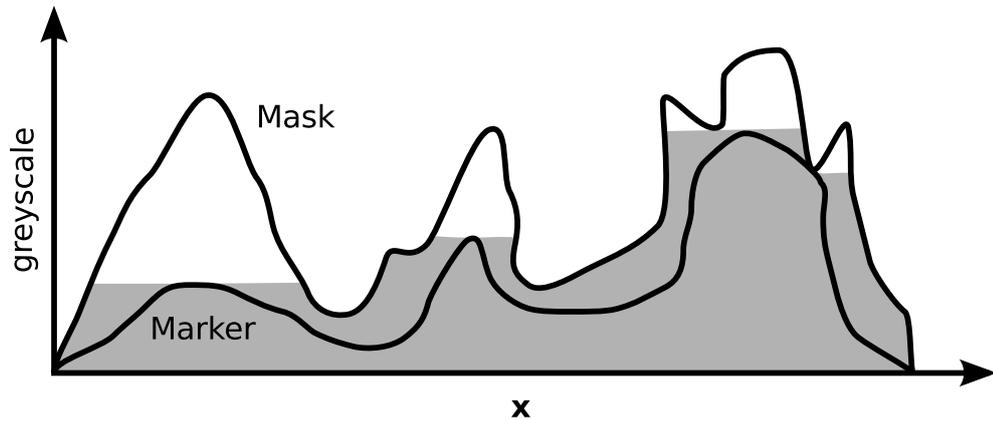
images are then thresholded, over a range of thresholds, and binary reconstruction is performed on each threshold. The maximum pixel intensities of the binary reconstructed images form the greyscale reconstructed image. With a good choice of a marker image this method can be used to create a reconstructed image with the intensity peaks (local maxima) removed. This can then be subtracted from the original image to enhance the peaks. Figure 2.7 is an illustration of airway enhancement by greyscale reconstruction in 1 dimension. Figure 2.7a shows the mask and marker greyscale images, with the reconstructed image represented by the shaded area. Figure 2.7b shows the difference image created by subtracting the reconstructed image from the original mask image.

In this study, the greyscale closing of the image of interest is used as the marker image:

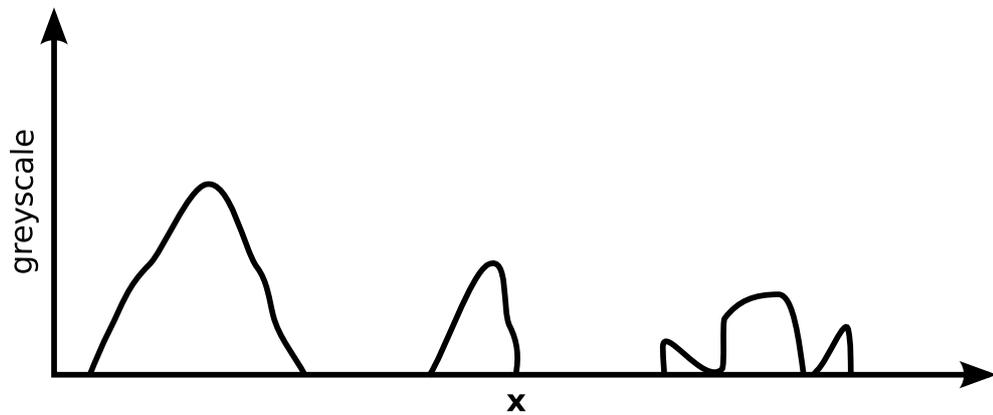
$$X = B \bullet D = (B \oplus D) \ominus D \quad (2.4)$$

where X is the marker image, B is the original image (used in reconstruction as the mask image) and D is the SE. D will control the shape of the marker image and, therefore, which airways are enhanced in the reconstruction. Applying morphological reconstruction a number of times to a slice using different SE sizes to produce marker images will enhance airways of different cross sections. Therefore, morphological reconstruction was applied using marker images produced using a range of SEs. The smallest SE chosen in this study is a 4-connected binary SE. Larger SEs are created by applying successive dilations to the smallest SE i.e. $D_n = D \oplus D \oplus \dots \oplus D$ n times.

As described earlier, reconstruction is applied using the range of marker images for each slice. As a closed image will have greater or equal intensity values compared to the original image, the reconstruction is applied using the inverted marker and mask images. The reconstructed image is the inverted again. Each reconstructed image is subtracted from the original image to enhance airways. These subtracted images are thresholded and the union of the thresholded images provides a binary slice with all possible airway locations. The threshold value (T) is obtained from this threshold fraction (T_{frac}) in terms of the minimum (P_{min}) and maximum (P_{max})



(a) Morphological greyscale reconstruction in one dimension. The marker image has a greyscale value less than or equal to the mask image and is chosen to extract local maxima of a circular appearance



(b) The difference image is found by subtracting the reconstructed image from the original image.

Figure 2.7: Airway enhancement using morphological reconstruction to enhance circular regions in CT slices

greyscale values in the reconstructed slice:

$$T = T_{frac} * (P_{max} - P_{min}) + P_{min} \quad (2.5)$$

Morphological filtering was applied slice by slice to the volume. Figure 2.8 shows each step of the morphological filter.

Aykac et al. (2003) and Pisupati et al. (1996) applied a similar filtering to each axial slice. We apply 3D filtering i.e. to each slice in the axial, coronal and sagittal plane. This is because of poor detection of branches parallel to the slice if just one direction is used. The axial plane is filtered last in order to enhance segmentation of bifurcation areas that can appear large and non circular from the axial plane if branches are parallel to the slice. Smaller branches parallel to the axial plane are, therefore, segmented first leaving more circular bifurcation areas that are detected with the axial filter (see Figure 2.9).

2.3.4 Region Growing and Leak Removal

Once possible regions have been detected using morphological filtering and reconstruction, bounded space dilation is applied to the binary volume as a region growing technique from the initial seed point (Masutani et al., 1998).

Closed space dilation is particularly useful because airway branching can be monitored while the region growing is taking place. Closed space dilation is described as follows:

$$X \oplus_B K = X_N = (X_{N-1} \oplus K) \cap B = \dots \quad (2.6)$$

where $X = X_0$ is the initial seed, B is the region being segmented, K is the kernel and \oplus is the morphological dilation operation (Masutani et al., 1998).

This method is seeded by the initial slice in the trachea, and 3D closed space dilation is applied until the segmentation volume does not increase using a $3 \times 3 \times 3$ SE. 2D closed space dilation, which is used previously (Aykac et al., 2003), is applied to each slice individually and requires a number of forward and backward passes through the whole volume to capture the 3D data. While 3D dilation allows any complexity of topology to be followed, we specify a constraint that the object remains 26-connected.

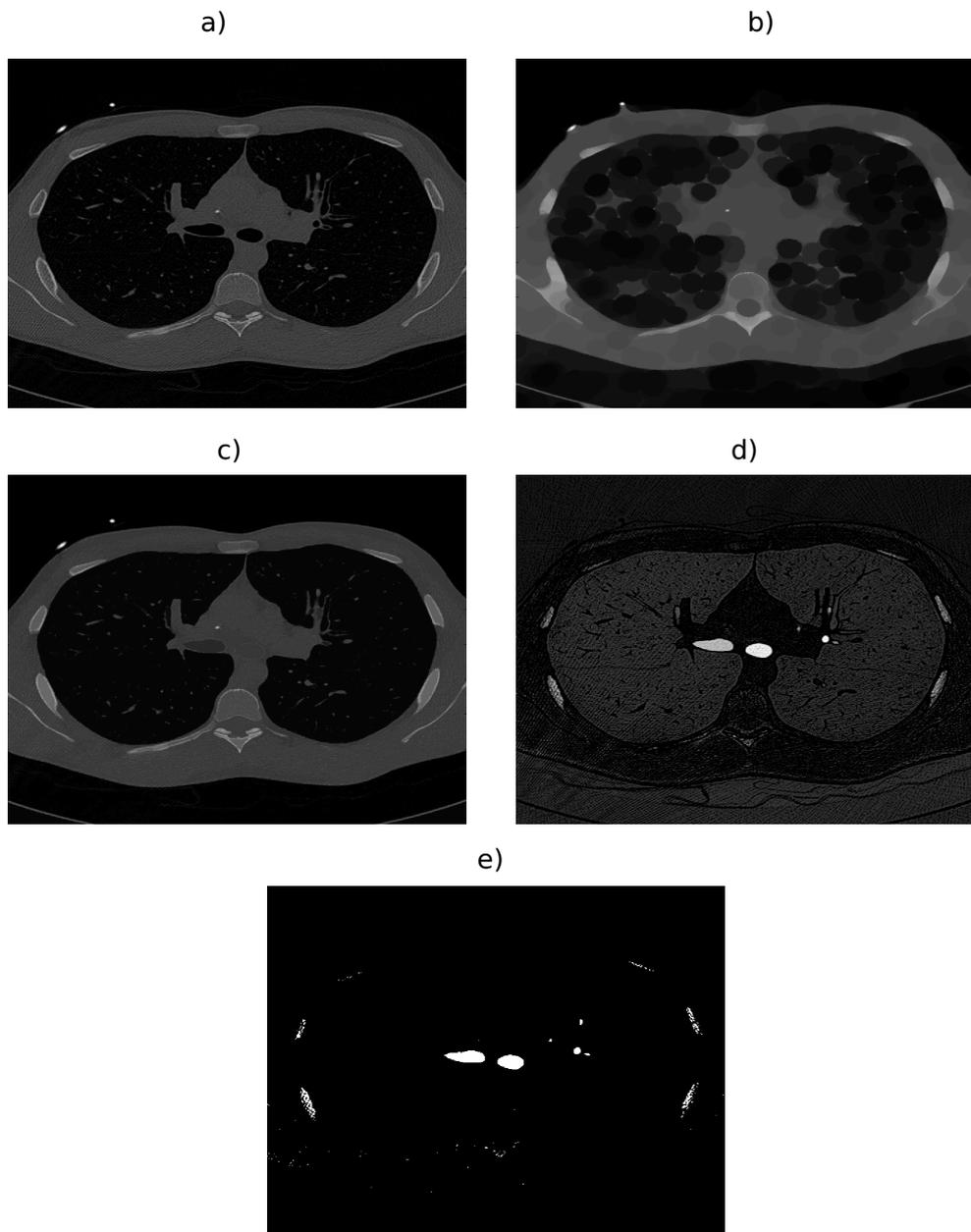


Figure 2.8: Greyscale closing and reconstruction applied to a CT slice a) Original image b) closing the image with a single structuring element c) morphological reconstruction d) difference image e) threshold

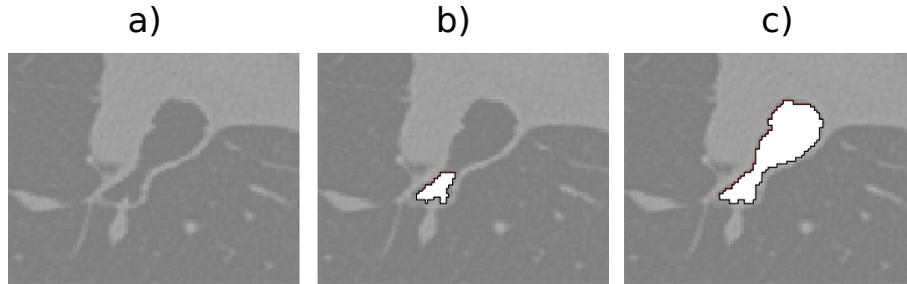


Figure 2.9: Multidimensional morphological filtering is used to enhance non-circular bifurcation points by first identifying vessels of a circular appearance in the coronal and sagittal directions. a) cropped CT slice with a non-circular airway region b) coronal and sagittal filtering c) axial filtering

In some cases other objects that are of a similar size and shape to that of the airways are enhanced by the morphological reconstruction and remain when thresholded. If after thresholding, the object is 26-connected to the airway region it will be segmented; this causes leaks. Leaking is prevented by applying a restriction on the dilation. The change in cross section of each branch can be monitored by calculating the area of the “growth front” of each branch, where the growth front is the area added to the connected region for each iteration of the closed space dilation (Masutani et al., 1998). The growth front of each branch is calculated by labelling each 26-connected from the difference between the bounded dilation and the previous iteration (Figure 2.10b). Each connected region is assigned the label of the connected region from the previous step (Figure 2.10c). If two 26-connected region are assigned the same label then a bifurcation has occurred and new labels are created. To detect leakage where the volume increase substantially, the volume of the last 3 dilations for each branch are compared to the 3 dilations before that. If the ratio of the volumes is above a specified threshold – chosen to be double the size – then closed space dilation along this branch is stopped.

This segmentation method results in a binary labelled volume identifying the connected airway tree. Figure 2.11 shows an example of the airway segmentation in a coronal slice and Figure 2.12 shows an example airway segmentation with bone rendering of a 20 month patient.

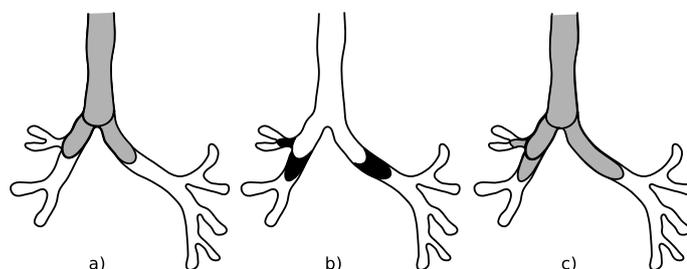


Figure 2.10: Leak removal step: identification, labelling and analysis of the changes in the growth front of each branch. a) shows detection of the bifurcation from the connected regions of the growth front b) new growth fronts are generated c) growth fronts are assigned labels based on connectivity to previous growth fronts unless bifurcation occurs

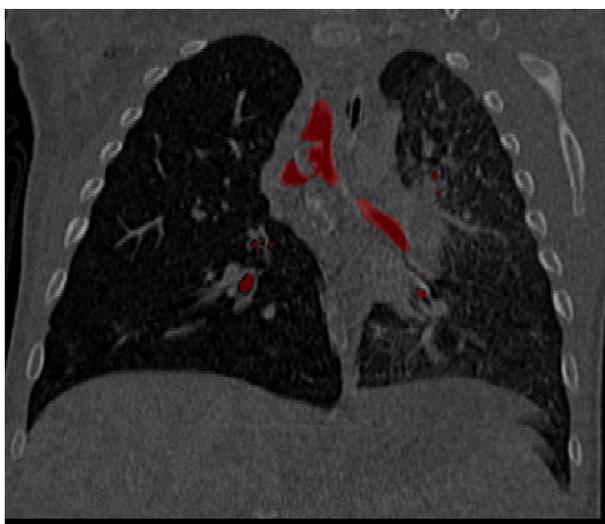


Figure 2.11: Airway segmentation overlaid on a coronal CT slice. In this case there is at the bifurcation between the trachea and RMB.

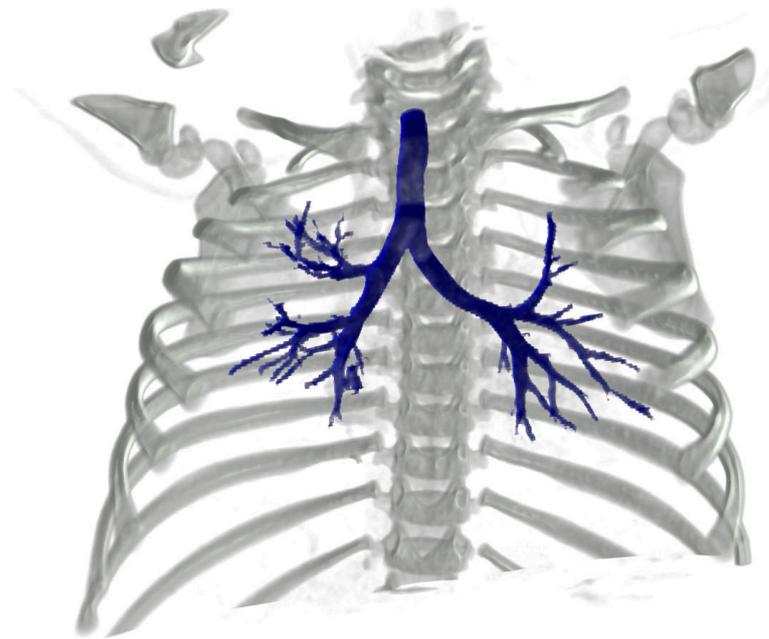


Figure 2.12: Airway segmentation of a CT scan from a 20 month patient with rendering of surrounding boney tissue

2.4 CT Datasets

A number of CT datasets were used throughout this study. Adult CT scans were used for training and evaluation of the segmentation algorithm as part of the EXACT'09 challenge. A number of datasets containing paediatric CT scans of patients with and without tuberculosis (TB) were collected and segmented. The *GOSH 2010 paediatric nonTB CT dataset* and *TYG 2010 paediatric TB CT dataset* were used as a training set, and the *Tygerberg 2012 paediatric TB and non-TB CT dataset* were used as a test set in Chapters 3, 4 and 6.

2.4.1 EXACT'09 Adult dataset

This dataset was used for evaluation of airway segmentation algorithms as part of the EXACT'09 airway segmentation challenge. The set consisted of 20 training CT scans and 20 test CT scans with various pathology. The dataset ranged from healthy volunteers to patients severe airway and lung pathology (Lo et al., 2009c).

CT scanners used in this dataset included the Siemens Sensation 64, the Siemens Sensation 16, Siemens Volume Zoom, Philips Mx8000 IDT 16, Philips Brilliance 16P, Toshiba Aquilion and the GE LightSpeed 16 (Lo et al., 2009c). Slice thickness ranged from 0.6 to 1.25 mm. In the test set Case 21 and Case 38 were taken at inspiration while the remainder of the test set were taken at expiration.

2.4.2 GOSH 2010 paediatric nonTB CTs

A paediatric Chest CT dataset was acquired in 2010 from Great Ormond Street Hospital, London, UK and consisted of 37 CT scans of patients without TB. The mean age (with standard deviation) was 3.8 ± 1.8 years.

These images were acquired using the Siemens SOMOTOM Definition and the pixel size in each slice varied from 0.28 mm to 0.51 mm and the slice thickness varied from 1 mm to 1.5 mm.

2.4.3 Tygerberg 2010 paediatric TB CTs

A paediatric Chest CT dataset was acquired in 2010 from Tygerberg Hospital, South Africa and consisted of 52 scans of patients with TB. The mean age was 1.8 ± 2.2 years.

These images were acquired using the Siemens SOMATOM Sensation 40, Siemens SOMATOM Emotion 6 and the Toshiba Aquilion. The pixel size varied from 0.24 mm to 0.54 mm and the slice thickness varied from 0.6 to 1.5 mm.

2.4.4 Tygerberg 2012 paediatric TB and non-TB CTs

Another paediatric Chest CT dataset was acquired in 2012 from Tygerberg Hospital, South Africa and was chosen to contain both TB and non-TB patients. There were 42 non-TB patients with a mean age of 3.1 ± 3.8 years and 48 TB patients with a mean age of 2.4 ± 2.8 years.

These images were acquired using the Siemens SOMATOM Sensation 40 and the Siemens SOMATOM Emotion 6. The pixel size in each slice varied from 0.21 - 0.54 mm and the slice thickness varied from 0.6 to 1.5 mm.

Non-TB patients in the dataset include children with the following conditions: parenchymal lung disease such as interstitial lung disease, congenital lung malformations which affect the parenchymal tissue and not the airways; suspected lung metastases; and infective conditions other than TB. This will include cases of bronchiectasis, cystic fibrosis and pleural disease. Patients with suspected TB undergo a CT scan when there are signs or symptoms of airway involvement and therefore, all TB cases in this dataset have suspected airway involvement.

2.5 Evaluation and Results

Building a “gold standard” validation set is difficult because of the time required to manually segment a CT scan, and smaller bronchi are easily missed. However, expert manual segmentations remain the best available form of validation. For example, Aykac et al. (2003) use 8 manually segmented scans for evaluation. Other methods of evaluation include using a phantom (Tschirren et al., 2002), evaluation of a subset

of axial slices containing the segmentation by an expert (Fetita et al., 2004) or comparison of manual airway labels to automated airway labelling from the segmentation (van Ginneken et al., 2008). Many segmentation algorithms are also evaluated qualitatively by comparing the 3D segmentation and the number of bronchi segmented to the CT volume (Kiraly et al., 2002).

This section evaluates the airway segmentation algorithm that has been outlined in this Chapter. This algorithm was developed in order to form the basis of a paediatric airway analysis. Due to the lack of paediatric gold standard data, the method was evaluated on an adult dataset as part of the EXACT’09 airway segmentation challenge. This section discusses the EXACT’09 airway segmentation challenge, the evaluation of this method using the challenge and general results of the challenge. Paediatric airway tree segmentations are also presented.

2.5.1 EXACT’09 challenge (adult patients)

Evaluation of airways segmentation methods is a difficult task because different CT datasets and varying performance measures are used by each group; the CT datasets often vary considerably in resolution and image quality. The Pulmonary Image Analysis workshop ³ at MICCAI 2009 held the EXACT’09 airway segmentation challenge in an attempt to evaluate various airway segmentation methods. 22 teams registered for the challenge and 15 submitted results for evaluation (Lo et al., 2009c). Lo et al. (2009c) evaluated each team’s performance based on the accuracy of segmentations of a set of 20 CT volumes. A gold standard was developed for each volume and compared to the segmentation from each algorithm.

40 CT scans were selected as the EXACT’09 dataset, and divided into 20 training images and 20 testing images. The CT images contained a selection of different acquisition conditions and pathologies. These CTs were provided to the teams and the segmentations from each team were submitted to EXACT’09 for evaluation. A reference segmentation was found for each test CT scan by combining the segmentations from each algorithm and using trained observers to remove false branches. Each segmentation from each team was evaluated by comparing the segmentation to the appropriate reference segmentation.

³www.lungworkshop.org

The segmentations were divided into individual branches for evaluation. Branch points were found using a fast marching algorithm as a region growing method. A number of slices, from each branch, containing the image and segmentation information, were then viewed by trained observers and evaluated. More than one observer viewed each branch and labels of “correct”, “partly wrong”, “wrong” or “unknown” were assigned. The labelled branches from all segmentation algorithms were combined and correct voxels were used to form the reference segmentation for each test CT volume. Each segmentation was then evaluated by comparing each segmented branch to the reference segmentation (Lo et al., 2009c).

2.5.1.1 Limitations of the evaluation

This evaluation achieved a complex task of using a common method of evaluation for a number of airway segmentation methods. However, there are some limitations that have to be considered when viewing the results:

- Ground truth images for the training set were not available to the teams – making it difficult to optimise parameters of the segmentation algorithm
- A binary segmentation was submitted. Different methods had varying sensitivities and false positive rates (FPR) . Without probabilities assigned to each voxel or segmentations at multiple FPRs, a ROC curve can’t be plotted, making it difficult to compare the sensitivity of two segmentation algorithms with different false positive values.

2.5.2 EXACT’09 Segmentation evaluation

The algorithm parameters were adjusted using the 20 training images and these parameters were fixed for the 20 test images. The parameters were qualitatively adjusted by comparing the segmented airway tree to the CT volume. The number of dilations used to create the largest SE for filtering the axial, coronal and sagittal slices were set to 12, 6 and 6 respectively. Therefore, for each slice in each orientation SEs from a 4-connected binary SE to an SE after 6 or 12 dilations are used to generate marker images. The threshold fraction used for the axial, coronal and sagittal slices were set to 0.3, 0.4 and 0.4 respectively.

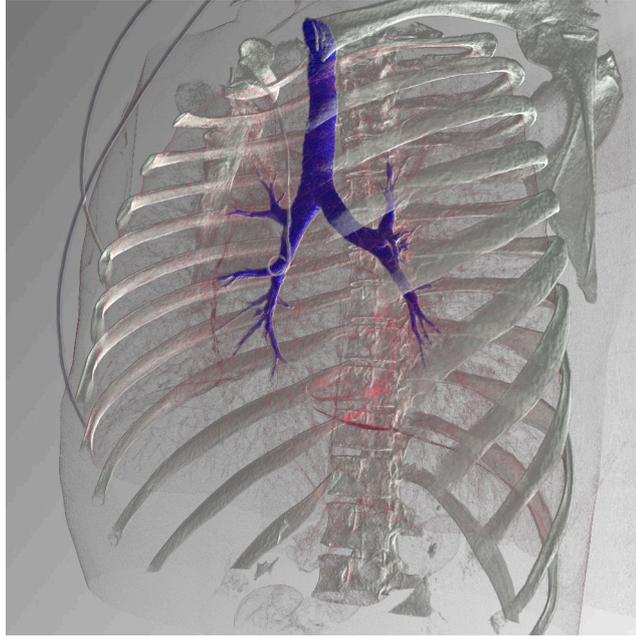


Figure 2.13: Adult airway segmentation with bone rendering (from the EXACT’09 training set)

The segmentations of the test images were submitted to EXACT’09 for analysis. Figure 2.13 shows an example airway segmentation with bone rendering. Figure 2.14 shows the segmentations from the 20 test images. Green (lighter regions) shows the correct segmentation and blue (darker regions) shows falsely classified regions.

Table 2.1 shows the accuracy of the algorithm for the 20 test cases provided. The meaning of each heading is as follows: *branch count* is the number of branches detected correctly, *branch detected* is the proportion of branches detected compared to the ground truth, *tree length* is the sum of all correctly detected branches, *tree length detected* is the tree length compared to the ground truth, *leakage count* is the number of correct regions bordering incorrect regions, *leakage volume* is the volume of regions wrongly detected, and *false positive rate* is the fraction of wrongly detected regions out of all detected regions. *branch detected* and *tree length detected* are similar measures – with the difference that detected branches of any length will make the same contribution to the former while correctly detected longer branches will contribute more to the latter. Ideally an algorithm would achieve a high proportion of *branch detected* and *tree length detected* with a low *false positive rate*. The ground

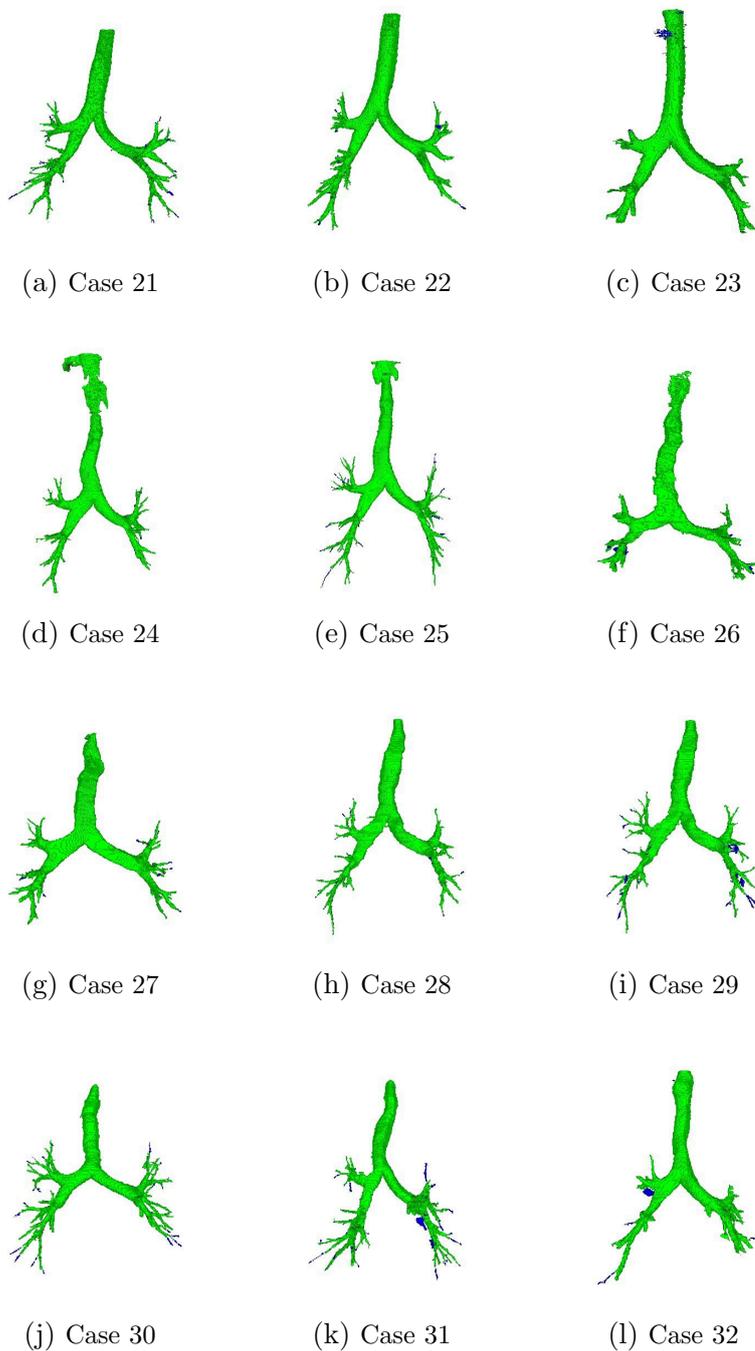
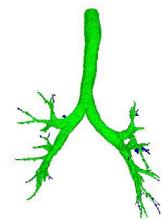
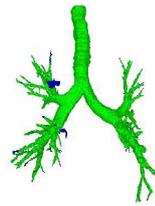


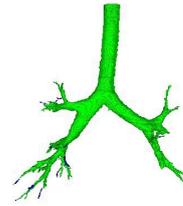
Figure 2.14: Segmentation of test cases for the EXACT'09 airway segmentation competition (*green* is correctly segmented and *blue* is false positive)



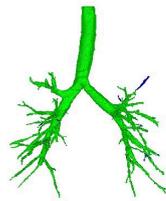
(m) Case 33



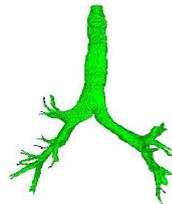
(n) Case 34



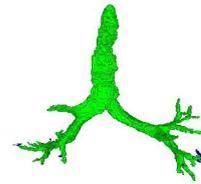
(o) Case 35



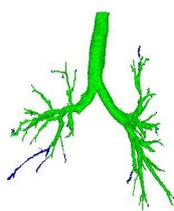
(p) Case 36



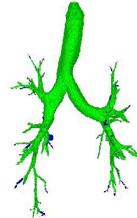
(q) Case 37



(r) Case 38



(s) Case 39



(t) Case 40

Figure 2.14: Segmentation of test cases for the EXACT'09 airway segmentation competition (*green* is correctly segmented and *blue* is false positive)

truth is determined from union of segmentations from all teams, where false branches are removed by clinical experts. Thus, the proportion of correctly detected branches was low for all teams. This algorithm produced a fair result; the mean number of branches detected was 43.5%, the tree length detected was 36.4% and the number of false positives was 1.27%.

The time taken to process a case varied according to computer specifications and the number of slices containing airways. The mean and standard deviation of the time taken to complete each of the 20 cases was 71 ± 18 minutes using a single core of a quad core 2.83GHz system. The algorithm evaluated in this competition was an initial version. Considerable improvements were made to the speed of the algorithm (discussed in Section 2.5.4).

2.5.3 EXACT'09 challenge general results

Figure 2.15 and Figure 2.16 show the *Branch Detected* and *Tree Length Detected* plotted in terms of the *False Positive Rate* for the 15 teams. The number in each figure refers to the team number and each algorithm is outlined in Table 2.2. Our algorithm is highlighted in red. The plots show that at the FPR (lower is better) of 1.27, out of the 15 teams, 4 teams had a lower FPR (better) and 10 teams had a higher FPR (worse). At this FPR point, 9 teams had a higher sensitivity (better) and 5 teams had a lower sensitivity (worse).

The lack of an ROC makes it difficult to compare performance of algorithms at different specificities. However, teams 5, 13 and 15 achieved a very good performance in the competition. Team 13 use a gradient vector flow method that is applied after an initial segmentation and centreline extraction step (Bauer et al., 2009). Team 15 develop an interactive airway segmentation method (Tschirren et al., 2009). Initially a fully automated airway segmentation is performed and then the user is provided with a framework to add or remove branches. The organisers of the airway segmentation competition (Team 5) also presented a method (Lo et al., 2009a). This method achieved a very low FPR and used an airway appearance model trained on previously segmented airways. The low FPR means that this method would be useful as a basis for the paediatric airway analysis but a trained model might not be accurate for severe forms of pathology that are not found in the training set. It is also worth

	Branch count	Branch detected (%)	Tree length (cm)	Tree length detected (%)	Leakage count	Leakage volume (mm ³)	False positive rate (%)
CASE21	101	50.8	55.7	50.4	0	0.0	0.00
CASE22	76	19.6	48.4	14.7	0	0.0	0.00
CASE23	33	11.6	23.9	9.2	0	0.0	0.00
CASE24	86	46.2	63.4	39.0	0	0.0	0.00
CASE25	123	52.6	91.4	36.3	3	11.6	0.06
CASE26	55	68.8	37.8	57.5	9	16.4	0.23
CASE27	83	82.2	55.6	68.7	2	95.2	1.01
CASE28	79	64.2	58.4	53.3	7	121.3	1.34
CASE29	92	50.0	62.6	45.4	3	17.9	0.19
CASE30	94	48.2	66.6	43.6	2	48.2	0.51
CASE31	91	42.5	59.0	33.6	4	1487.4	11.71
CASE32	45	19.3	35.8	16.5	5	803.2	6.94
CASE33	85	50.6	59.6	40.5	0	0.0	0.00
CASE34	140	30.6	90.3	25.2	3	32.4	0.17
CASE35	82	23.8	53.8	17.4	0	0.0	0.00
CASE36	148	40.7	147.0	35.7	1	1.9	0.01
CASE37	65	35.1	46.9	26.4	1	94.5	0.76
CASE38	73	74.5	44.1	66.3	4	57.5	0.65
CASE39	155	29.8	111.0	27.1	1	1.4	0.01
CASE40	116	29.8	80.7	20.9	5	256.1	1.71
Mean	91.1	43.5	64.6	36.4	2.5	152.3	1.27
Std. dev.	32.3	19.1	28.2	17.1	2.6	362.7	2.91
Min	33	11.6	23.9	9.2	0	0.0	0.00
1st quartile	73	29.8	46.9	20.9	0	0.0	0.00
Median	86	44.4	58.7	36.0	2	17.1	0.18
3rd quartile	123	64.2	90.3	53.3	5	121.3	1.34
Max	155	82.2	147.0	68.7	9	1487.4	11.71

Table 2.1: Evaluation measures for the twenty cases in the test set. The table columns are as follows: *branch count* is the number of branches detected correctly, *branch detected* is the proportion of branches detected compared to the ground truth, *tree length* is the sum of the length of all correctly detected branches, *tree length detected* is the tree length compared to the ground truth, *leakage count* is the number of correct regions bordering incorrect regions, *leakage volume* is the volume of regions wrongly detected, and *false positive rate* is the fraction of wrongly detected regions out of all detected regions.

noting that a considerable proportion of the airway tree was not segmented by any of the algorithms, which means there is room to improve airway segmentation methods. The algorithm with the highest sensitivity detected 76.5 % of the airway with a FPR of 15.6 %.

Method	Summary
1	Morphological filtering and region growing Irving et al. (2009)
2	Morphological filtering, Connection cost and Region growing Related to Fetita et al. (2004)
3	Restricted region growing
4	Branch edge detection and region growing
5	Airway appearance model and airway/vessel guide Method was submitted by EXACT'09 challenge evaluator(Lo et al., 2009c)
6	Region growing with revaluation of ROI
7	Tube detection and knowledge based region growing
8	Adaptive region growing
9	Local centricity region growing
10	Region growing with leak detection
11	Wave propagation and template matching
12	Adaptive region growing
13	Gradient vector flow
14	Multi threshold with topological restrictions Method based on van Ginneken et al. (2008)
15	Region growing User manually corrects branches Based on Mori et al. (1996)

Table 2.2: Segmentation methods evaluated in EXACT'09 (Lo et al., 2009c)

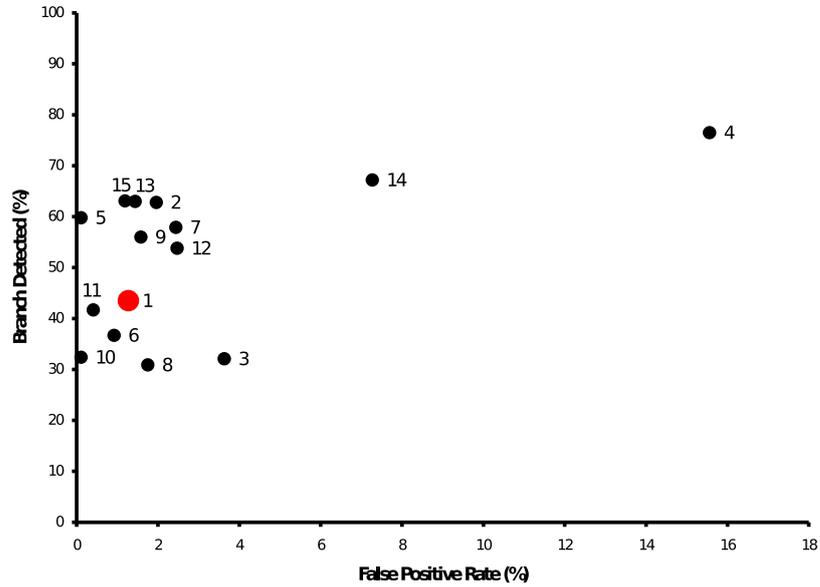


Figure 2.15: Branches detected vs false positive rate for the algorithms in the EXACT'09 airway segmentation competition. Red shows our method.

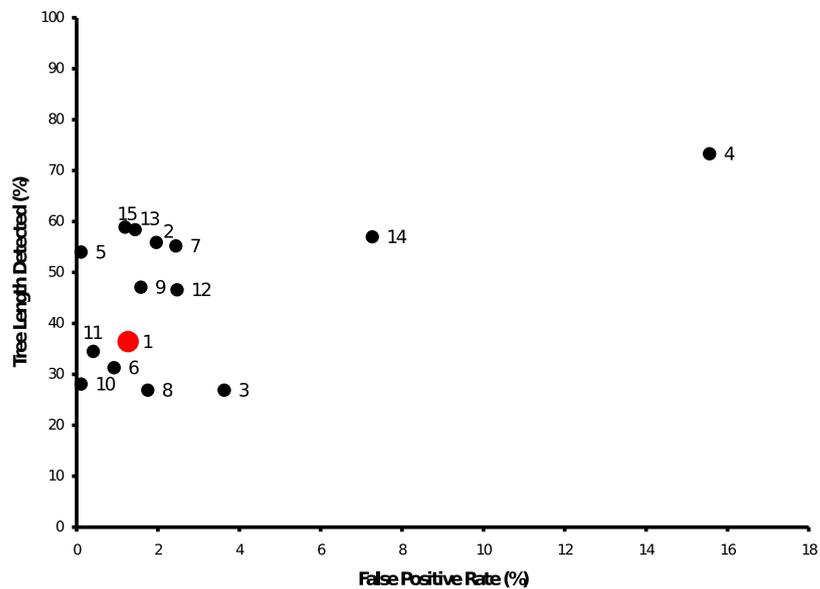


Figure 2.16: Tree length detected vs false positive rate for the algorithms in the EXACT'09 airway segmentation competition. Red shows our method.

2.5.4 Improvements to method after EXACT'09

The EXACT'09 challenge evaluated an initial version of the airway segmentation algorithm. Since the evaluation, a number of modifications were implemented to mainly improve the running time (but also accuracy).

Underlying algorithms were reimplemented more efficiently in Matlab and underlying Matlab function were chosen that took advantage of multi-core processing. The morphological filter – the bottleneck in the code – was limited to axial slices containing the lung region.

The structuring element used in the method is initialised as a 4-connected binary image and successive structuring elements are created by dilating the SE with the original SE. This results in the SE having a diamond shaped appearance. The accuracy of extracting circular airways was reduced by this shape and the structuring elements were replaced by pixel approximations of circular SEs. This improved the accuracy of the algorithm.

2.5.5 Paediatric airway segmentation

The number of dilations used to create the largest SE for filtering the axial, coronal and sagittal slices were set to 9, 5 and 5 respectively (from [12, 6, 6] in the adult set). The threshold fraction used for the axial, coronal and sagittal slices were set to 0.3, 0.45 and 0.45 respectively (from 0.3, 0.4, 0.4 for the adult cases). These parameters were chosen using a subset of the training set.

The time taken to segment each airway in the paediatric test set was $168 \pm 57s$. A considerable improvement over previous run times due to the modifications discussed in Section 2.5.4. However, some of the improvement is due to the smaller volumes used for paediatric patients.

Paediatric airway segmentations were evaluated qualitatively by comparing the segmentation to the CT volumes. Segmentations generally performed well up to the fourth generation of bronchi and in some cases many more generations. Obstructed branches could not be segmented using this method.

Figures 2.17 and 2.18 show example segmentations from paediatric CTs of patients with and without TB. As shown, the leak removal step removes false branches but can also reduce the number of generations that were segmented. The vast major-

ity of leaking in the segmentation algorithm occurs in the peripheral branches (See Figure 2.17h) and this method is effect for these cases. Some branches have almost completely obstructed branches due to severe airway pathology. The False branch detection algorithm can remove bronchi beyond a severe stenosis due to the nature of the branch radius monitoring algorithm (See Figure 2.18h).

Some cases have completely obstructed bronchi (as shown in the Figures). This method extracts a completely connected airway tree and is not able to segment obstructed branches.

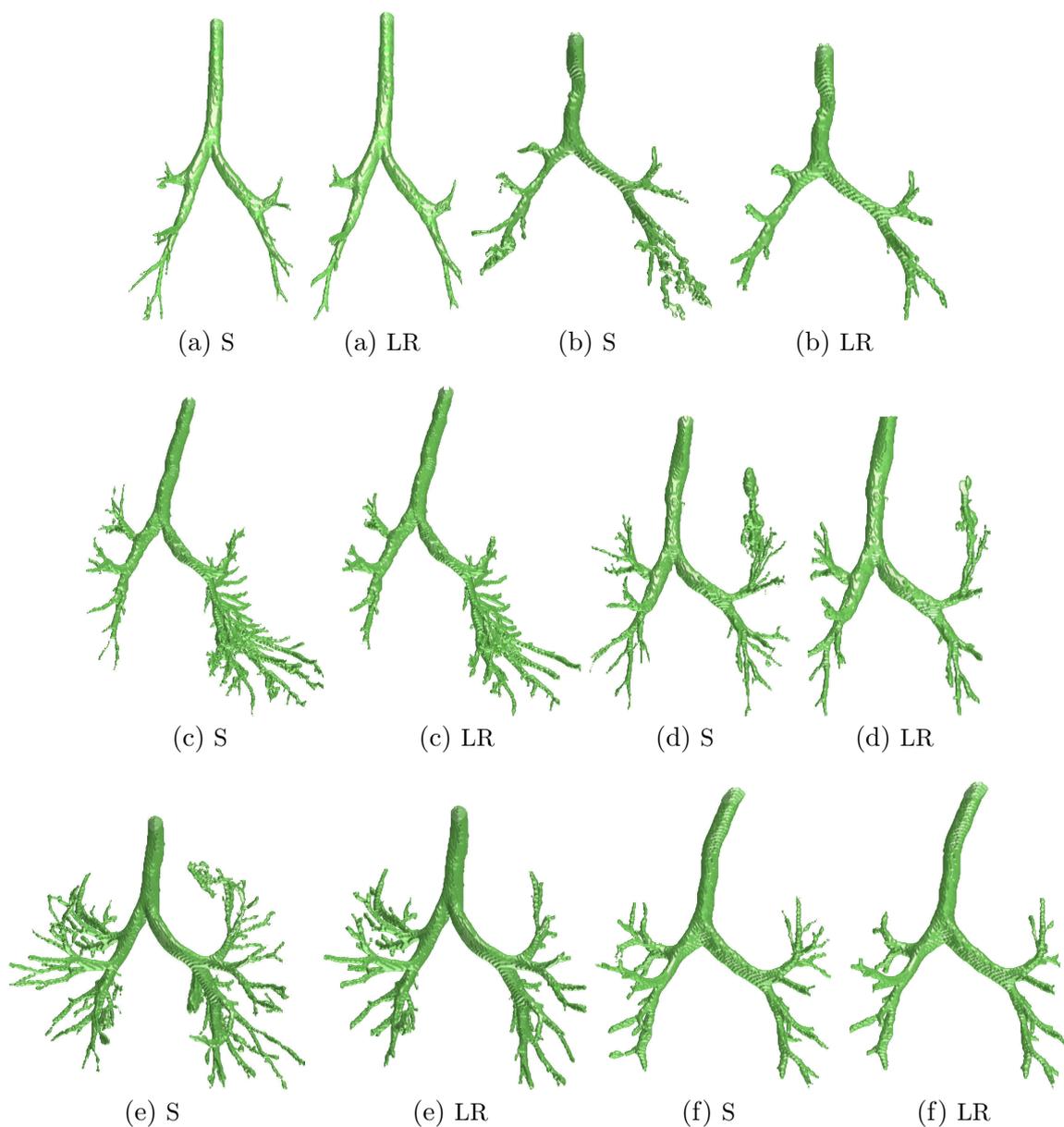


Figure 2.17: Examples from the Tygerberg 2012 non-TB paediatric evaluation set (age 2.4 ± 2.8 years) (S: Segmentation, LR: Segmentation with leak removal step)

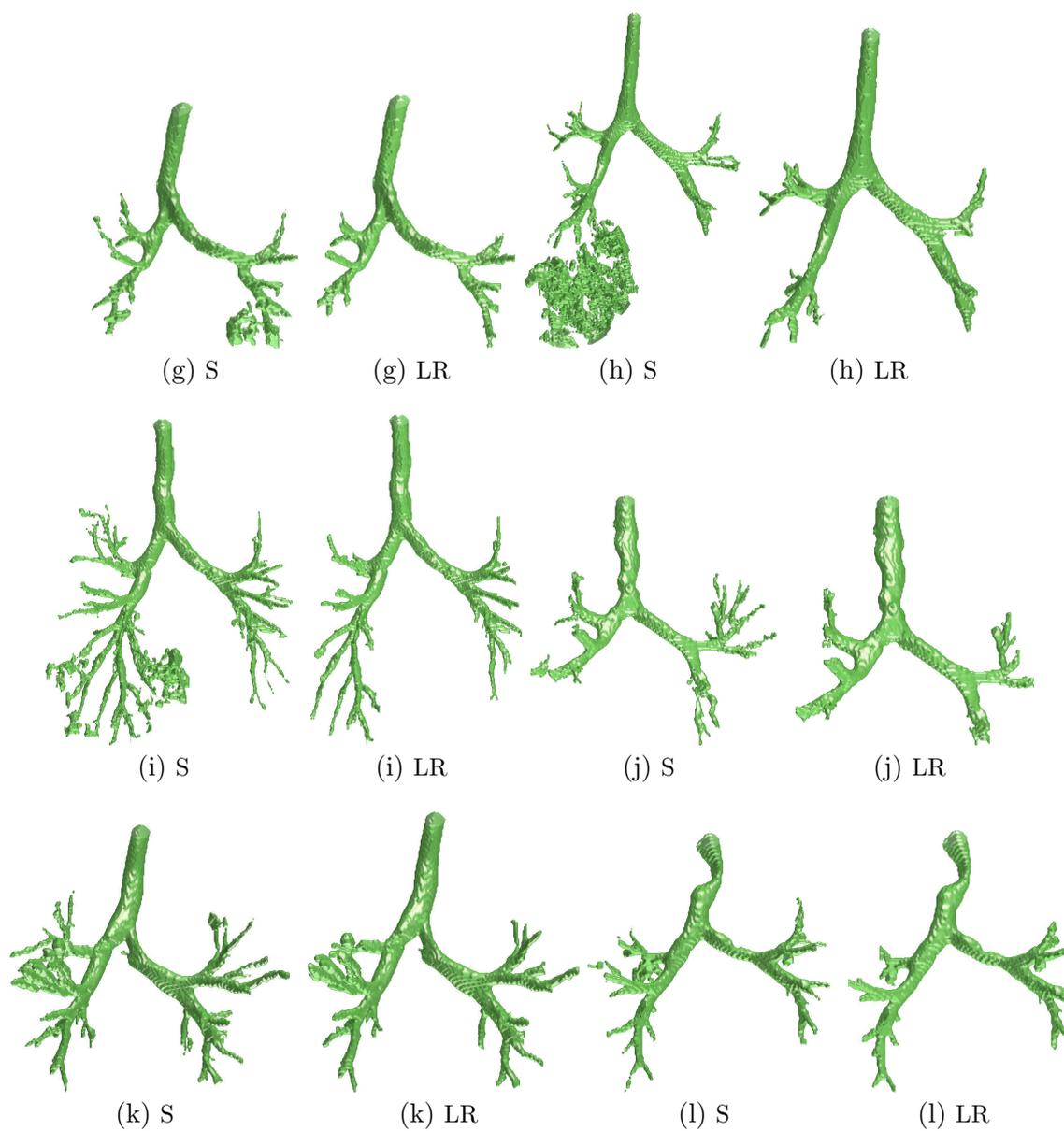


Figure 2.17: Examples from the Tygerberg 2012 non-TB paediatric evaluation set (age 2.4 ± 2.8 years)(S: Segmentation, LR: Segmentation with leak removal step)

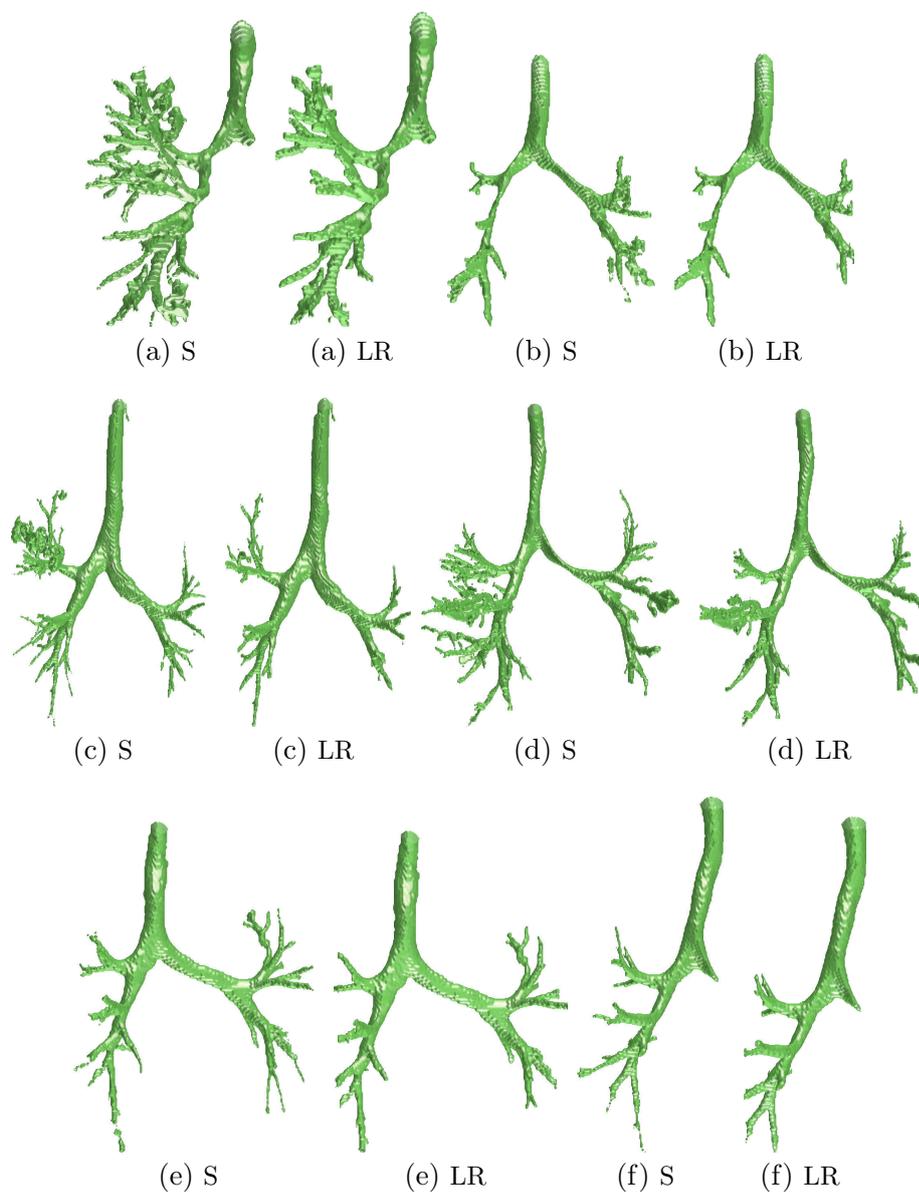


Figure 2.18: Examples from the Tygerberg 2012 TB paediatric evaluation set (age 3.1 ± 3.8 years)(S: Segmentation, LR: Segmentation with leak removal step). Some cases have complete obstruction of a bronchus

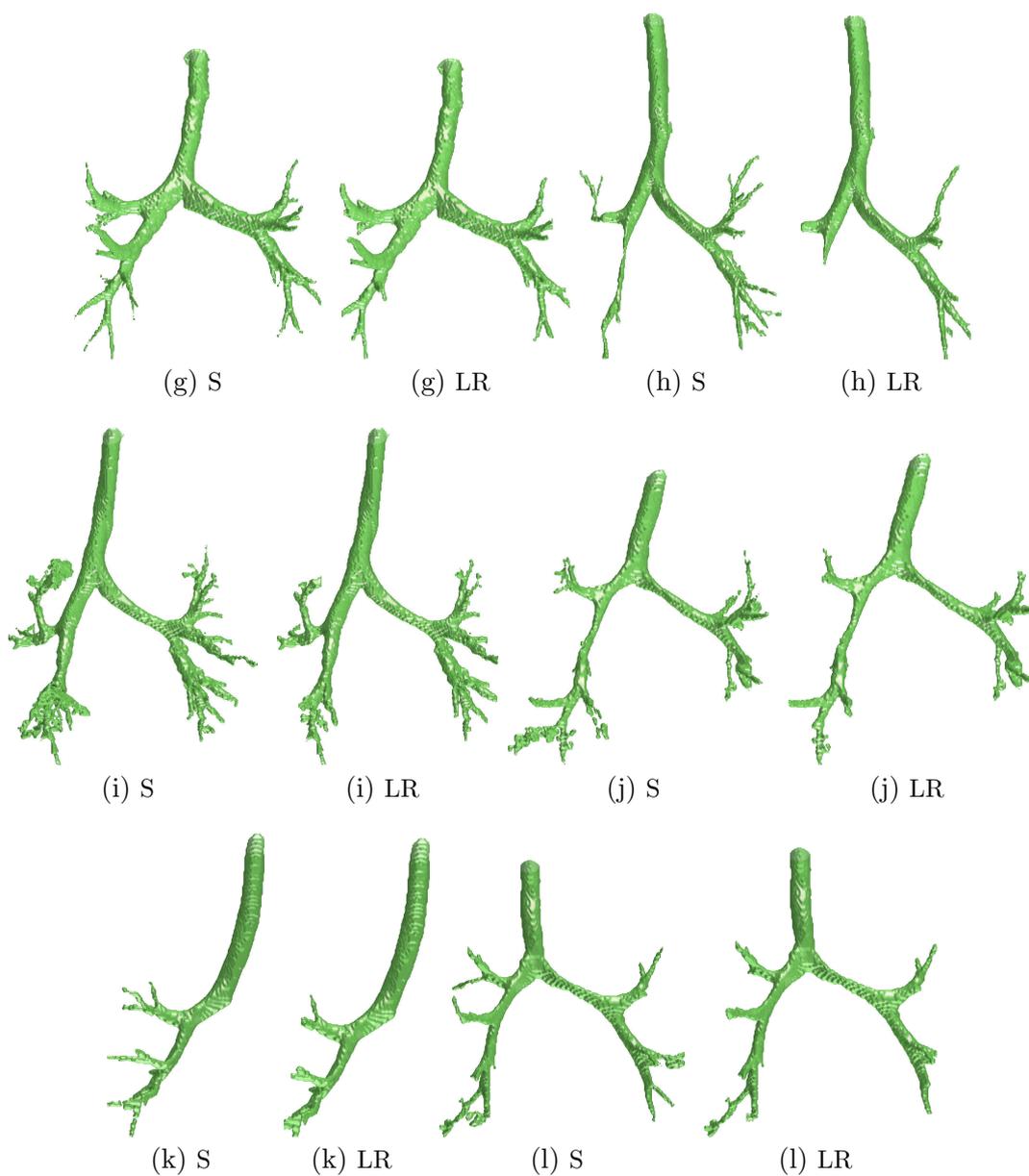


Figure 2.18: Examples from the Tygerberg 2012 TB paediatric evaluation set (age 3.1 ± 3.8 years)(S: Segmentation, LR: Segmentation with leak removal step). Some cases have complete obstruction of a bronchus

2.6 Discussion

This method was developed for the segmentation of the airways in paediatric patients and segmentations will be used in the modelling and detection of airway deformation and stenosis in later chapters. The method was adapted to adult airway segmentations by choosing suitable algorithm parameters using the EXACT'09 training set. Test set segmentations were submitted to EXACT'09 for evaluation.

In comparison to other “state-of-the-art” algorithms (Figures 2.15 and 2.16) our segmentation algorithm (Group 1) performed fairly well. Especially considering the early development stage of our algorithm at the time of submission. The plots show that at the FPR (lower is better) of 1.27%, out of the 15 teams, 4 teams performed better (lower FPR) and 10 teams performed worse. At this FPR, the branch detected was 43.5% (9 teams performed better and 5 teams performed worse). By removing (or reducing) the effect of the false branch removal step in the algorithm, the branch detected could be improved at the expense of the FPR. Therefore, for future studies, a ROC curve of branch detected vs FPR might be more helpful in evaluating algorithms. However, there could be difficulties modifying some algorithms to generate values at multiple FPRs.

The method was slow in the EXACT'09 challenge but modifications were made to the algorithm to considerably improve performance. The benefit of this method is the low false positive rate. This is required for paediatric TB cases because exudate in the lungs increases the likelihood of segmentation leaks. Exceptions are CASE31 and CASE32 that have relatively high leakage.

Some algorithms with better performance on the adult dataset are discussed in Section 2.5.3. These methods could be incorporated into the system and evaluated on paediatric Chest CT scans. However, sufficient performance is achieved with the algorithm for the paediatric airway analysis presented in later chapters.

A weakness of this method is in identifying airway cross-sections where partial volume effects and other artefacts give the appearance of airway wall discontinuities. The method could be improved by a more sophisticated region growing algorithm (after morphological filtering) that is able to use a number of consecutive slices at each step and is, thus, not affected by single slice discontinuities. This effect is apparent in some smaller branches. Small discontinuities in the segmentation for

CASE 22, 23, 32 and 35 removed a number of large branches when region growing was applied leading to poorer results. Later chapters focus on airway analysis of the first few generations of bronchi and, therefore, is not affected by this weakness and this is left to future developments of the algorithm.

The algorithm was applied to the paediatric airway dataset and used for further airway analysis. In later chapters, analysis is only performed on the early generations of bronchi (including the Trachea, RMB, RUL, BI, LMB, LUL and LLL) because these are the bronchi most likely to be affected by enlarged lymph nodes from paediatric pulmonary TB. Therefore, the segmentation algorithm is only required to segment the larger bronchi for the airway analysis performed in this thesis.

Qualitative observations of the paediatric segmentations showed good performance. However, paediatric airway segmentation is made more challenging by the considerably smaller airway size which means a larger impact due to voxel size, and, therefore, partial volume effects. Figure 2.18 shows airways from TB patients. Lymphadenopathy causes complete obstruction of the airways in some cases (LMB in cases a, f, k). The airway segmentation method is not able to segment beyond completely obstructed airways but a later chapter develops a method to segment disconnected airway branches. In future, a thorough evaluation of the performance of the segmentation algorithm on paediatric cases may be useful for comparison, if suitable “gold standard” reference segmentations could be obtained.

2.7 Conclusion

Airway segmentation from CT is a challenging topic given partial volume effects, image variability and difficulty establishing a gold standard segmentation. This chapter proposes a method to segment the airways using a morphology based segmentation algorithm. An efficient adaptive threshold method is used for the larger branches while 3D morphological filtering is used for the smaller branches. This approach is based on morphological filtering but introduces novel branch restriction and leak removal steps. A novel application of the filtering method to 3D is also presented by initially filtering in the coronal and axial directions, and mapping these enhanced regions to the greyscale value of the airway wall, in order to aid segmentation of branch points, during the axial filtering.

This algorithm, like most other airway segmentation algorithms, segments a connected airway region from the trachea. In a number of cases, severe pathology leads to complete obstruction of the major airways. A method to segment beyond complete obstructions is discussed in Chapter 5.

2.8 Chapter summary

This chapter presents an airway segmentation algorithm:

- Requirements
 - Segment paediatric airways from CT
 - Low false positive rate
 - Effective for scans with severe pathology
- Airway Segmentation
 - Initialise segmentation by identifying trachea
 - Segment main bronchi
 - Morphological filter and segmentation of smaller bronchi
 - Leak removal step
- Evaluation
 - Evaluated on adult dataset as part of EXACT'09 challenge
 - Applied to TB and non-TB paediatric dataset
- Application
 - Basis for airway analysis algorithm
 - Basis for obstructed airway segmented algorithm

3

Airway Structure

Abstract. This chapter presents methods to detect the airway branching structure from airway segmentations. An iterative thinning approach is used to extract the one-voxel thick skeleton of the airway. Bifurcations and individual branches are detected from the connectivity of each voxel in the skeleton. False branches are then removed using the length, location and connectivity of each branch. Once an accurate branching structure is obtained, heuristics are used to assign an anatomical label to each branch and a distance transform is used to label the surface mesh of the airway. This structure provides a basis for more advanced airway analysis.

The detection of the branching structure of the airways is an important step before performing more complex airway shape analysis. This chapter discusses the implementation of methods for airway centreline detection and branching structure detection from 3D airway segmentations. This can be used to guide branch measurements, airway registration and statistical airway models, which are discussed in later chapters.

A number of methods have been developed previously for airway centreline detec-

tion and branch labelling (outlined in Section 3.1). Our implementation is discussed in Section 3.2. The branching structure of the airways is identified by: skeletonisation to extract the branch centrelines (Section 3.2.1), bifurcation point analysis (Section 3.2.2), branch connectivity labelling (Section 3.2.3), false branch removal (3.2.4) and anatomical labelling of the seven main branches (Section 3.2.5). This chapter focusses on the implementation of structure analysis methods and lays the foundation for more sophisticated airway analysis in later chapters.

3.1 Background

3.1.1 Skeletonisation

Skeletonisation of the airway can be used to extract bifurcation points and branch centrelines – a one voxel thick approximation of a line that travels through the centre of any branch cross section. This is generally used to identify the branching structure of the airways and, therefore, to label each branch. There are also a number of other benefits to extracting the centreline, such as, directing virtual bronchoscopy and identifying airway cross sections. There are a number of skeletonisation methods. However, many cannot produce a 1-voxel thick centreline, which is required for branch point analysis. This section will focus on examples of algorithms that have been successfully applied to airway segmentations from CT images.

3.1.1.1 Region growing methods

Some methods identify the centreline as part of the airway segmentation step (Bülow et al., 2004; Schlathölter et al., 2002; Swift et al., 2002; van Ginneken et al., 2008). Swift et al. (2002) calculate points on the centreline by finding the centroid of boundary points in the cross section of the airway. Schlathölter et al. (2002) grow the skeleton as a product of a branch point detection step. Extraction of the centreline is performed in a similar way by Wood et al. (1995). The centreline is derived from the centroid of the wave-front, and bifurcation points are identified from the centrelines of the three branches forming the intersection. A seed point at the base of the trachea, in the segmented airway, is used to initialise the method and is labelled 0. Voxels

that are 26-connected to the 0 voxel are labelled 1 and unlabelled voxels that are 26-connected to voxels labelled 1 are labelled 2 etc. This is repeated until all voxels are labelled and the centroid is found, for each connected set of voxels with the same label, to create a centreline.

3.1.1.2 Thinning methods

Region-growing methods are not the only way to extract the centreline from the airway tree. Palágyi et al. (2003, 2006) developed an algorithm that extracts the centreline by iteratively thinning the segmentation. Each surface voxel is analysed in terms of orientation and connectivity, and voxels are removed that do not affect the structure of the shape, until a 1-voxel centreline remains. This method has also been used successfully by other researchers (Tschirren et al., 2002).

The algorithm proceeds as follows: a base point is found as the centre of the first slice containing the trachea and this point is preserved. Next, surface voxels of the segmentation are identified as points that are 6-connected to 0 voxels (background voxels) in one direction. They are grouped as U (up), N, E, S, W and D (down) voxels depending on the direction of the 6-connectivity to a background voxel. *Simple points* are points that can be removed without changing the topology of skeleton and *line-end points* are points that are 26-connected to only one other 1 voxel (an airway voxel).

Simple points that are not *line-end points* are selected and iteratively removed for each directional grouping U, N, E, S, W and D. For a voxel i in the current directional grouping:

- Check if i is a simple point or a line end point
- Remove i if it is a simple point and not a line end point
- Also if the point is a line-end point but the number of deleted 6-connected neighbours is greater or equal to t then also delete i

t controls the sensitivity of the algorithm to surface effects and branching. This algorithm acts to iteratively thin the airways down to a 1-voxel thick skeleton and is repeated until no more changes occur.

Surface irregularities can still cause false branches so a pruning method is used to remove false branches by cutting circularly connected regions and removing branches based on their length and distance from the airway surface.

Mori et al. (2000) also use 3D region thinning to find the centreline and label the branching structure. They overcome the problem of false branching by observing that generally false branches are short and have no child branches. They also found that branches with a large diameter tended to have more false branches. Therefore, branches that have no child branches are examined and if the branch is shorter than a threshold (t_{length}) or if the branch is connected to a branch that has a diameter greater than a threshold ($t_{diameter}$), then the branch is deleted. t_{length} and $t_{diameter}$ are found experimentally.

3.1.1.3 Bifurcation points

Bifurcation points can be found from the centrelines of the region growing methods or the skeleton of the thinning methods. The bifurcation region can be defined as the region where a bronchus divides into two child bronchi – where the bifurcation point is the centre of this region.

Extracting the centreline from the wavefront means that the branches are disconnected in the bifurcation region (see Figure 3.1a). Wood et al. (1995) overcome this problem by fitting lines to each of the three branches forming the bifurcation and use a minimisation algorithm to find the best intersection of these lines, which is labelled as the bifurcation point (see Figure 3.1b). The middle 70% of each branch is used to fit the line because of noise in the centreline around the bifurcation point. As thinning methods produce a 1-voxel thick skeleton, the branch points can be found from the connectivity of that skeleton.

However, there may be variability in the exact location of the bifurcation point because of variation between algorithms and airway shape. This will not affect the label of each bronchi but can affect the shape of the end of each branch (see Figure 3.6b), which can in turn influence a shape based model derived from the branches. Additional methods are proposed in Section 4.2.4 to overcome these challenges for statistical shape models.

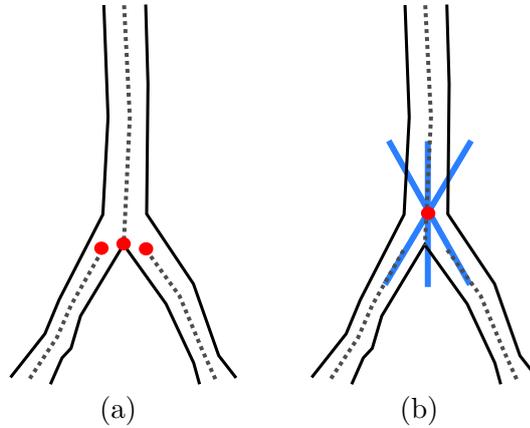


Figure 3.1: Bifurcation point detection from the wavefront a) the branches are disconnected in the bifurcation region b) a minimisation algorithm can be used to find the best point of intersection.

3.1.2 Anatomical branch labelling

Once the airway has been segmented and branch points found, branches can be matched to their anatomical labels. This is necessary for almost any analysis because it allows branch specific comparisons to be made. Examples include locating the position of abnormalities in the airway tree, and using the branching structure of the airway to identify lung structures.

Tschirren et al. (2005b) outline a method for matching airway branches to anatomical labels. They assume the first three or four generations of the airway can be matched and further generations differ for each individual. Their method uses the centreline to convert each airway tree into a graph where vertices represent branch points. An associative graph is then built between an airway of interest and a labelled anatomical model and used to find the best match between the trees.

The centreline and branch points are used to represent the airways as a directed acyclic graph where the vertices represent endpoints and branch points, and the vertices are connected by edges that are representative of the connectivity of the airway tree (as shown in Figure 3.2).

The association graph is used to match an airway tree to an anatomically labelled population average. An association graph (G_{assoc}) is a measure of the similarity be-

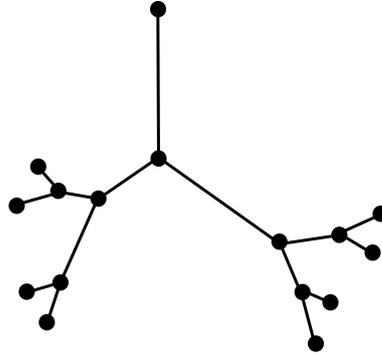


Figure 3.2: Graph representation of the airway tree

tween two graphs (G_1 and G_2) and has a vertex for every possible pairing of vertices consisting of vertices from G_1 and G_2 . An edge connects two vertices in G_{assoc} if the relationship between the corresponding vertices in G_1 is the same as that between the vertices in G_2 . Tschirren et al. (2005b) use inheritance and topological distance to define this relationship. Therefore, in the association graph *maximum clique* (the largest subset of vertices where every vertex is connected to every other vertex) represents the set of vertices of G_1 and G_2 that have the same relationship. This *maximum clique* represents the best match between the two trees and forms the basis of Tschirren et al. (2005b)'s algorithm. However, a number of steps are taken to reduce the computational complexity of the task of finding the maximum clique.

The anatomical model, which is matched to each airway, is a population average, created from a dataset of labelled segmentations. A number of measures are used to define the anatomical model, including individual measures of branch length and orientation, as well as measures between branches, of inheritance (e.g. parent and child), angle between segments, spatial relationship and topological distance.

Every branch is represented as a start vertex, end vertex and edge. Edges between vertices are only added to the association graph if the inheritance relationship is the same and the topological distance is similar for a branch of the tree and the population average. Vertices and edges in the association graph are then assigned weights according to the measures discussed earlier. Vertices and edges with low weights are removed to reduce the computational complexity of finding the maximum clique. The maximum clique is then found which represents the match of the branches

from the airway tree to the anatomical model. Therefore, using a population average, labels can be assigned to each branch of a segmented tree.

An alternative method is discussed by van Ginneken et al. (2008). They use a training set of manually labelled airway segmentations, and probabilities based on orientation, average radius and angle are then used to label branches. Starting from the trachea, each branch is labelled based on: the label of the parent and the probabilities of the child and grandchild branches.

Mori et al. (2000) also propose a template matching algorithm. Their method is similar to that of Tschirren et al. (2005b) in that they use branch position and hierarchy to match the branches. However, their labelling is based on minimising a global fitting function instead of using an associative graph.

Their approach is divided into two levels of detail: a) labelling of the right main bronchus (RMB) and left main bronchus (LMB) b) labelling the smaller bronchi. Candidate RMB branches and LMB branches are found by identifying branches with the trachea as the parent and an end point that is less than k from a template

$$\|Q_{b_i} - Q_{template}\| < k \quad (3.1)$$

where Q_{b_i} is the location of the end point of branch B_i and $Q_{template}$ is the location of the template end point. If more than one candidate exists then an evaluation based on the child branches is performed. Other potential branches exist because the skeletonisation method can create false branches (Mori et al., 2000) and this algorithm is designed to minimise the risk of breaking labelling at the RMB/LMB level.

Smaller branches are labelled by comparing each branch to a template branch using an energy function E , where:

$$E_{B_i K_j} = \frac{d^{B_i} \cdot d^{K_j}}{\|d^{B_i}\| \|d^{K_j}\|} \quad (3.2)$$

where B_i is the branch under consideration, K_j is the labelled template branch, and d^{B_i} and d^{K_j} are the directions of B_i and K_j , respectively. Therefore, E is greatest when B_i is in the same direction as K_j , and this is used to label a branch.

Therefore, starting from child branches of the main bronchi, each branch (B_i) is compared to template branches with the same parent branch. B_i is assigned the label which yields the highest E . If that label has already been assigned to another

branch then the branch with the highest E is assigned the label and it is removed from the other branch. This method is applied iteratively until all branches have been labelled. Finally, if more than one candidate LMB or RMB exists then each candidate is tested and all child branches matched. The candidate with the largest sum of all energy from the child branches is selected. This method successfully labelled 93 % of branches that were successfully segmented and skeletonised.

It is difficult to determine the best anatomical labelling method from the above studies. All studies achieve at least a 90% accuracy in assigning anatomical labels to segmented branches. Mori et al. (2000) achieve a 93% accuracy on a 14 patient validation set while Tschirren et al. (2005b) achieve a 97% accuracy on a 17 patient dataset containing normal and pathological cases. van Ginneken et al. (2008) use a much larger 170 patient dataset with a 90% accuracy. These methods only label the segmented branches and, therefore, the results are dependant on variation in the dataset and segmentation algorithm.

3.2 Method

The previous section discussed methods available for airway skeletonisation and branch labelling. In this section the implementation of methods to identify the centreline, bifurcation points and branch labels in our algorithm are discussed.

3.2.1 Airway Skeletonisation

Palágyi et al. (2003)'s skeletonisation method is used for the extraction of the centreline¹. There are two steps to the extraction of the medial skeleton: topological correction and centerline extraction. Morphological closing removes small holes, bays (surface deformations) and cavities in the segmentation². These small changes would cause unwanted changes such as false branches. Once the segmentation has been corrected, the centerline is extracted by thinning (as described in Section 3.1.1.2). Each surface voxel is analysed in terms of orientation and connectivity. If the surface voxel

¹Palágyi et al. (2003) provided C++ code for the skeletonisation.

²Morphological closing and opening was applied using a 6-connected structuring element. Larger structuring elements increased the risk of changing the topology of the segmentation for severely stenosed branches.

is classified as a simple point – a point that does not alter the topology of object – and is not an end point, then it is removed. This is repeated iteratively until no more voxels are removed. This results in a one voxel wide centreline. Examples of the airway centreline are shown in Figure 3.5.

3.2.2 Branch Point Detection

Branch points – the point at which the airway bifurcates – are found by analysing the connectivity of the centreline; the neighbourhood of each voxel is used to determine the connectivity. The input is the skeleton of the airway and the output is a list of branch points.

For each voxel in the skeleton, the 26-connected neighbourhood is found and the central voxel (the voxel of interest) is removed. If the number of 6-connected regions in the neighbourhood is greater than two, then a branch point has been found. Figure 3.3a illustrates a section of the airways in the 26-connected neighbourhood of a bifurcation voxel. Figure 3.3b shows the three 6-connected regions and the bifurcation point. Note that this method will only work for skeletons that are one voxel thick.

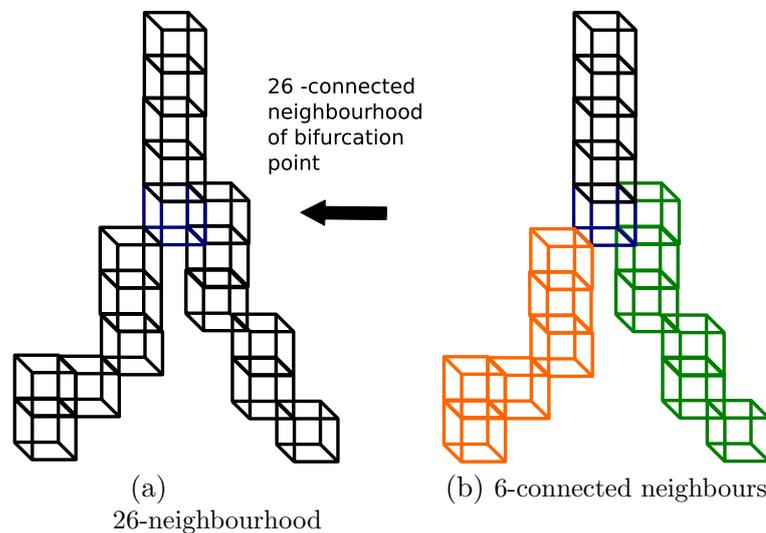


Figure 3.3: Representation a branch point in a one-voxel thick skeleton

3.2.3 Branch connectivity labelling

The skeleton and bifurcation points are used to assign a unique label to the centre-line voxels of each branch in the segmentation. Starting from the first voxel in the trachea, neighbouring voxels in the skeleton are given the same label unless there is a bifurcation, in which case, unlabelled neighbouring voxels are given a new label. An example of a labelled skeleton is shown in Figure 3.5. The order of the voxels in each branch is also recorded. Simplified pseudocode is presented for the branch labelling procedure using a one voxel thick branching centreline in Figure 3.4.

```
Start seed point
Add seed point to point list (PL)

WHILE PL is not empty
    current point (CP) is first point on PL

    extract the 26-neighbourhood (NE) of CP

    IF any NE are branch points (BP)
        FOR each 26-connected neighbour of the BP
            IF not labelled
                assign new label
                record label of parent branch
                add newly labelled points to PL
            END IF
        END FOR
    ELSE NE are not branch points
        FOR each NE
            IF not labelled
                assign same label as CP
                add newly labelled points to PL
            END IF
        END FOR
    END IF
    remove CP from PL
END WHILE
```

Figure 3.4: Simplified pseudocode for the branch labelling algorithm

3.2.4 Pruning

False branching sometimes occurs because of surface deformation (particularly when pathology is present). By identifying *end branches* (branches that end with a terminal vertex), a simple false branch removal algorithm can be developed that identifies end branches shorter than a given threshold, from the centreline, and removes them. This method assumes the false branches will be shorter than other branches. A problem with the assumption of branch length is that false branches from the main bronchi are often longer than true branches further down the tree. In our dataset we also have cases of complete obstruction of a branch – resulting in short centrelines. These need to be preserved in order to correctly identify the branching structure.

Therefore, the threshold was varied depending on the branch position. The algorithm was applied with 3 different thresholds to produce three pruned trees. Using knowledge of the tree connectivity, a new tree was constructed using the trachea from the first pruning and identifying the LMB and RMB labels; using the LMB and RMB from the second pruning and remaining branches from the third pruning (see Appendix B for a step by step example). After pruning, the branch relabels and connectivity are reassigned.

Figure 3.5 shows the labelled branch centrelines identified during the skeletonisation of the airway. Figure 3.5a and Figure 3.5c show two examples of the airways before pruning and Figure 3.5b and 3.5d show the relabelled skeleton after pruning. The numbers in the figure of the unique labels assigned to each branch for storing connectivity. Note that the false branch in Figure 3.5a may be longer than other end branches further down the airway tree – requiring the multi-threshold approach discussed in the previous paragraph. The false branch is caused by a surface artefact in the segmentation (shown in Figure 3.6b).

3.2.5 Anatomical branch labelling

Section 3.1.2 lists some sophisticated anatomical labelling techniques where each segmented branch is automatically assigned its anatomical label. This becomes very challenging when dealing with pathology that may cause completely obstructed branches (see Chapter 5). As discussed in Chapter 1, lymphadenopathy affects the main bronchi including the trachea, RMB, LMB, RUL, BI, LLL and LUL. These branches

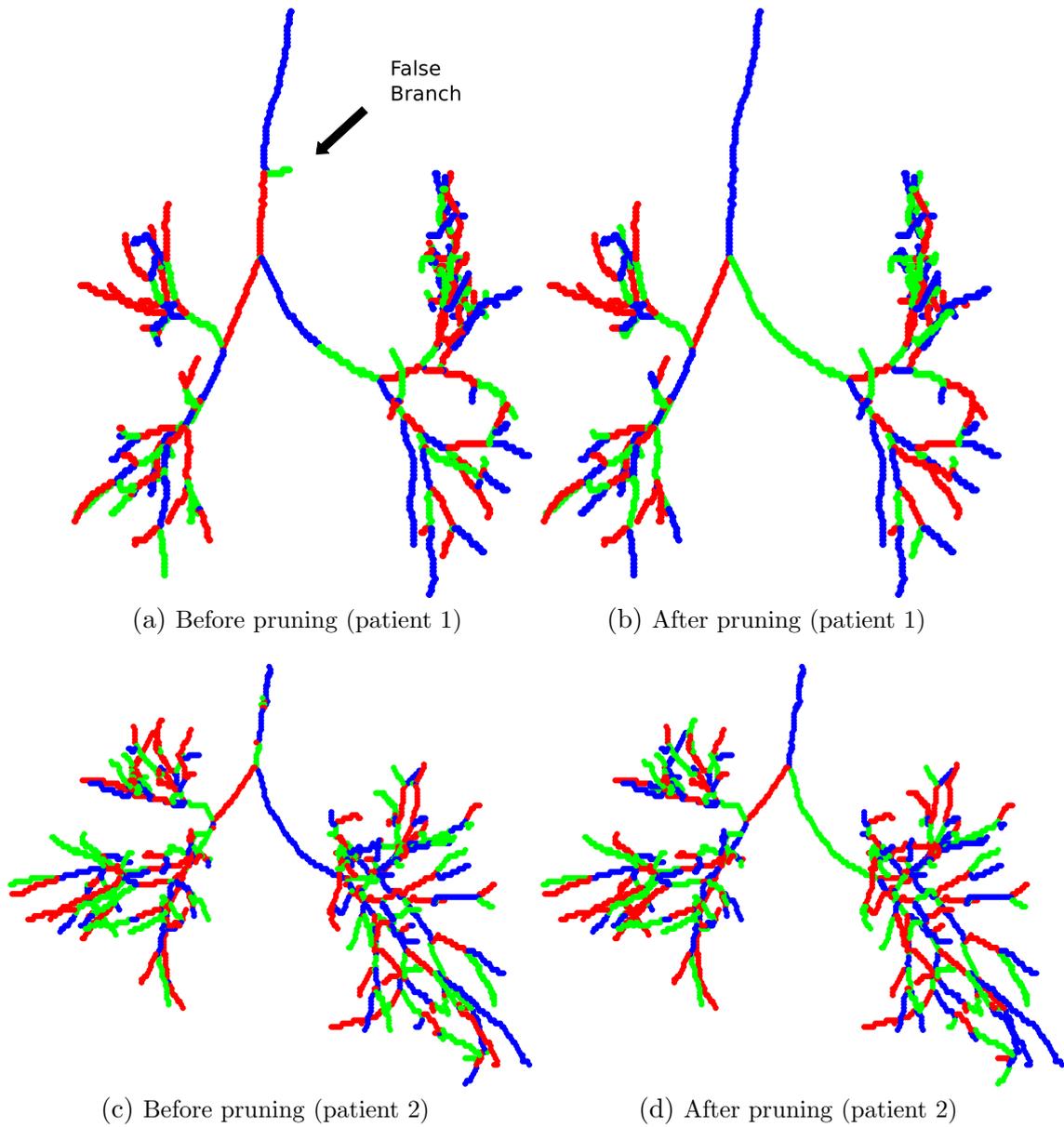


Figure 3.5: Centreline identification from a segmented airway tree (red, green and blue are used to illustrate distinct airway branches)

were labelled using a simple rule based method; the following rules were used to assign anatomical labels to these bronchi:

- the RMB and LMB are connected to the trachea
- the RMB is right of the LMB
- the RUL and BI are connected to the RMB
- the RUL is above the BI
- the LLL and LUL are connected to the LMB
- the LUL is above the LLL

This was sufficient to label the branches of interest for our case.

3.2.6 Airway surface labelling

In order to perform airway shape analysis, a mesh of the airway surface was used to represent each airway where each face was labelled according to the branching structure of the airway. The triangular surface mesh was constructed from the voxel segmentation using a marching cubes algorithm

Each mesh face ($m \in M$) was assigned a branch label by assigning the label of the closest voxel of the labelled skeleton (s_c):

$$s_c := \operatorname{argmin}_{s \in S} \|m - s\| \quad (3.3)$$

$$l(m) := l(s_c) \quad (3.4)$$

where M is the set of all mesh face centroid coordinates, S is the set of all voxels in the skeleton and each voxel has a branch label ($l(s) \in L$). The mesh representing each branch is then defined as $B_j := \{m | l(m) = l_{B_j}\}$ where l_{B_j} is the label of a given branch B_j

This method is effective for projecting the branching structure onto the surface. However, this labelling criterion introduces a number of issues when finding correspondence. Mesh faces in a bifurcation region are assigned to the connected branches

but these labels are influenced by branch shape – leading to variation at the bifurcation points. Figure 3.6 shows airway segmentations with branch labels (from the CT of a 20 month old patient and 6 year old patient). The red, green and blue colors are used to distinguish the connected branches. As shown in Figure 3.6b, in some cases, the closest point labelling approach leads to variability in the labels at bifurcation points (this is further developed in Chapter 4). Also shown, some cases may have some leaking in the segmentation in the peripheral branches; this does not affect the branching structure of earlier branches and can, therefore, be ignored.

3.3 Results and discussion

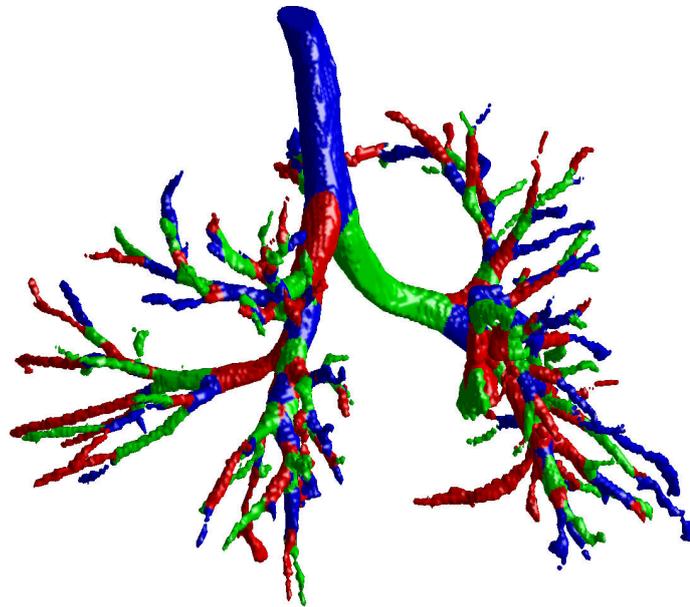
This chapter discusses the processing of a binary voxel segmentation in order to identify the airway structure. Figure 3.7 illustrates the smoothed centreline, airway segmentation and a coronal slice showing the lung fields. This structure is used as the basis for further analysis in the following chapters.

The pruning method uses a set of branch length thresholds to define false branches in each region of the airway. All end branches with a branch length of less than the defined threshold are removed. If more than one threshold is specified for a region then the threshold is repeated multiple times. This is useful when a false branch bifurcates to smaller false branches. Threshold values were determined qualitatively from the training set and are shown in Table 3.1. The rows represent the three extracted pruned trees and the columns represent the multiple thresholds for each region.

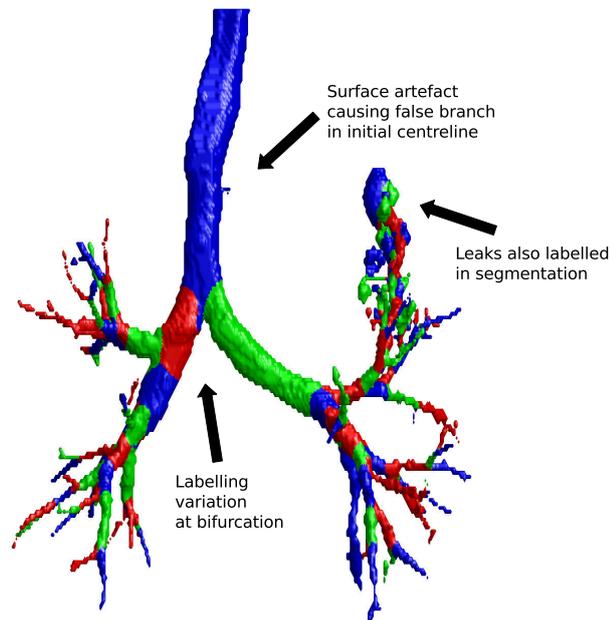
Region	$T_1(\text{voxels})$	$T_2(\text{voxels})$
Trachea	12	6
RMB-LMB	10	3
remaining branches	3	2

Table 3.1: Threshold values for pruning

The pruning method is effective for many cases and allows the centreline and branch points to be identified automatically. However 39 of the 179 cases still required some user interaction to remove a false branch or identify a missing branch, from the



(a) 20 month old patient (centreline shown in Fig 3.5d)



(b) 6 year old patient (centreline shown in Fig 3.5b)

Figure 3.6: Airway branch labelling from centreline (red, green and blue are used to illustrate distinct airway branches)

centreline, by choosing the branch to be removed or preserved. The majority of these cases are caused by pathology causing extreme changes to the airway shape (see Appendix B). Methods to automate detection of obstructed branches are discussed in Chapter 5.

Algorithms in later chapters require a skeleton with a direction that is representative of the direction of the branch. As discussed earlier, smoothing was used to overcome noise introduced by the voxel resolution. In the datasets used in this thesis, no kinks in the skeleton were found that may interfere with later analysis, and, therefore, no further smoothing was required.

As discussed in Section 3.1, airway skeletonisation and branch labelling is well established. This chapter implements a number of existing methods with some novel additions such as the multi-thresholded pruning approach. Implementing these existing algorithms in this computer assisted detection framework required a number of modifications to generate suitable results. However, the main purpose of the work in this chapter is to act as a bridge between airway segmentation and more advanced airway analysis techniques that are presented in later chapters.

3.4 Conclusions

Methods to detect the airway branching structure from airway segmentations have been outlined in this chapter. These include an iterative thinning approach to extract the skeleton, false branch removal, branch point detection and branch labelling. This structure provides a basis for more advanced airway analysis.

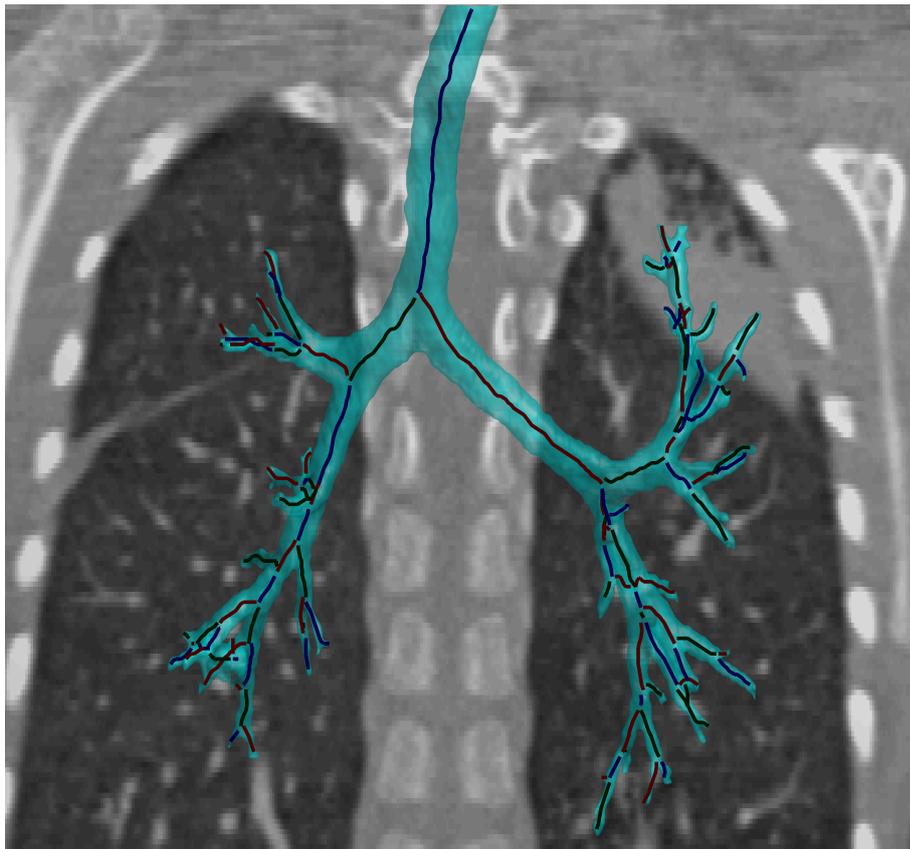


Figure 3.7: Smoothed centrelines with segmentation and coronal CT slice. The centrelines may not connect at bifurcation points because each branch is individually smoothed. The smoothing algorithm is discussed in Section 4.2.1 of the next chapter.

3.5 Chapter Summary

This chapter discusses the implementation of skeletonisation and branch point detection in order to extract the branching structure of the airway:

- Airway skeletonisation to identify a one voxel thick centreline
 - Palágyi et al. (2006) skeletonisation
- Branch point detection
 - Use the connectivity of the 1-voxel thick line
 - 26-neighbourhood of each voxel
 - 6-connectedness of each neighbour
- Branch labelling
 - Assign labels to the centreline voxels of each branch
 - Generate new labels at bifurcation points
- Prune false branches using a multi-threshold approach
 - Remove end branches with length less than a threshold T
 - Apply to regions of the airway for using three different threshold
 - Use branch connectivity to create final pruned tree from 3 prunings
- Use simple branch relations to assign anatomical labels to:
 - Trachea, LMB, RMB, RUL, BI, LUL, LLL

4

Airway registration and surface correspondence

Abstract. This chapter presents a novel approach to develop correspondence between multiple airway segmentations. The branching structure of the airway along with the orientation of the smoothed centreline is used to map local reference frames equidistantly along each branch. These reference frames are used to place landmark points on the airway surface. The landmark points along with a thin-plate-splines warp are used to align a template mesh to each airway. Local mesh alignment is then performed to match each vertex of the template to the airway surface. The optimal method was able to register a simplified template mesh to the airway dataset with a mean error of $2.1 \pm 0.9\%$. Errors larger than half a voxel contributed $0.22 \pm 0.49\%$.

Statistical shape models require correspondence in order to quantify shape variation. This chapter proposes a method to develop centreline and dense mesh correspondence between segmented airway trees. The input to this method is a branching centreline and surface mesh (Chapter 3) and the output is a new centreline and surface mesh where each centreline point and mesh vertex has a corresponding point and

vertex on every other airway – from which a statistical shape model can be built.

Section 4.1 discusses shape correspondence and Section 4.2 proposes a method for airway tree correspondence. Section 4.2.1 uses the resampled centreline to derive corresponding points onto the airway surface. These points are used to guide a thin-plate-spline warp of a template mesh onto each airway shape (Section 4.2.2). Finally point-to-point correspondence is developed by locally aligning the meshes (Section 4.2.3). Section 4.2.4 defines airway regions for developing local statistical shape models. Criteria for evaluating the template mesh registration are discussed in Section 4.3 and the results are presented in Section 4.4.

This is a fully automatic method that can be used to develop correspondence of the whole airway tree or just a single branch segment. Novel methods that are proposed in this chapter include: automation of landmark point generation, the application of a thin-plate spline warp to the branching airway tree, extension of closest point alignment for stenosed tubular structures, and a method to consistently identify local regions of the airway tree. The author is not aware of any other methods to develop mesh correspondence between segmented airways from different patients.

4.1 Background

In the previous chapter, the following topics have been discussed: skeletonisation, extraction of a surface mesh, and labelling the voxels and surface mesh according to the branching structure, and from this description of the airways we could extract branch measurements. For example, Palágyi et al. (2006) use the branching structure to perform simple measurements on each branch including: branch length, branch volume, branch surface area and average branch radius. These measures are calculated from the skeleton length, the number of voxels belonging to a branch and the voxel dimensions. The branch volume is taken as the sum of the number of voxels in a branch multiplied by the dimension of a voxels, and the length is calculated from the sum of voxels in the skeleton. These are useful measures but are restricted by using average measures for the whole branch. Tschirren et al. (2005a), on the other hand, perform a number of airway measurements along each branch by measuring the diameter of a section of the branch that is perpendicular to the skeleton. These forms of analysis do not require correspondence between airways.

Beyond simple branch measurements, there is potential to use statistical models to detect pathological deformation and stenosis in the airway tree from a dataset of airway segmentations. This is a relatively unexplored approach to airway tree analysis. Heimann and Meinzer (2009) provide an overview of statistical shape models in 3D. The main representations that are suitable for quasi-tubular objects – that is, objects showing similarities to tubes – include point distribution models (PDM) (corresponding points placed on the surface) and medial models (objects represented by centrelines and radii). Other methods include Fourier surfaces and spherical harmonics.

Active shape models (ASM), which apply PDMs to the problem of segmentation, are an example of a method for determining the prominent variation by modelling changes within a set of training shapes. Each shape in the training set is described by a set of corresponding landmark points. Therefore, a key requirement to create statistical shape models is accurate correspondence between the landmark points of each object in the training set. Otherwise, the model will include variation due to poor correspondence between shapes.

Determining shape correspondence is challenging due to the variation of biological structures. Often very different approaches are applied to different structures. The airways can be modelled as a branching quasi-tubular object and methods that determine the correspondence between tubular objects include the manual placement of landmark points, cylindrical parametrisation and a medial representation.

Hutton et al. (2003) and Paulsen et al. (2002) both use PDM models to analyse shape changes, in the face and ear canal respectively. They use a set of sparse landmarks which are selected manually and then perform thin plate spline (TPS) warping to a template to obtain a dense surface correspondence between the object surfaces. Dense correspondence is made by mapping the vertices of the template shape to the closest points on the surface of the warped objects. Hutton et al. (2003) use the mean landmarks as the template while Paulsen et al. (2002) use one of the training objects.

de Bruijne et al. (2003a,b) adapt ASMs for the analysis and segmentation of tubular objects. Landmark points are created by an expert placing equidistant points on edge contours of each slice of a training segmentation. While the focus of this method is not on the landmarking procedure, the disadvantage of this approach is

that, by only considering axial cross-sections, it assumes that the tubular object is orthogonal to the slice at all times. Manually drawn contours can be time consuming but in the case of ASMs, which are used to segment new objects, it is only required for creation of the training set. For evaluating shape changes (such as in our case), landmarks are also required for all cases and, therefore, it is beneficial to automate the landmarking procedure. In a related method, Deligianni et al. (2006) derive landmarks from the centroid of each axial slice in an airway segmentation. Each branch is manually defined, and the intersection between the surface and the cross-section from each centroid are used to define landmarks. This method is also limited by requiring manual interaction and choosing the branch centroids from the axial slice; accuracy is reduced when the branches are not perpendicular to the axial slices.

Huysmans et al. (2006) also use a PDM to analyse tubular objects but take a different approach when developing correspondence. Initially an iterative closest point (ICP) algorithm is used to align the shapes by obtaining a best match of the two shapes. They use cylindrical parametrisation to develop correspondence, whereby the mesh representing the object is iteratively collapsed, while minimising the error between the collapsed shape and the original shape, until the entire object is only represented by 6 vertices. The 6 vertices are then transformed to represent a cylinder and vertices are split to rebuild the complete mesh. All objects in the dataset are parametrised in this way to establish correspondence. Pinho et al. (2007) use the method of obtaining correspondence developed by Huysmans et al. (2006) to assess local tracheal stenosis. They do this by estimating the healthy trachea of a patient with tracheal stenosis. They apply the method to 9 shapes and placed 1024 landmarks on each shape. However, this method does not appear to be a practical approach for the analysis of branching tubular objects or large deformation. The assumption that the underlying shape can be approximated by a cylinder means that it is difficult to extend this analogy to a branching structure and severe deformation would lead to distortion of the correspondence.

An alternative representation is the medial model. Pizer et al. (2003a,b) use discrete medial representations (m-rep) of shapes. The basic component of a discrete m-rep model is the atom and a set of atoms represent a shape. An atom consists of a hub, which is the centre of the maximal sphere, and two spokes to represent the two radii whose end points are tangent to and intersect with the surface of the

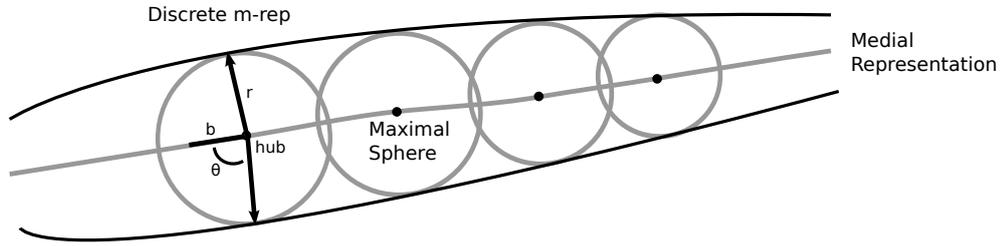


Figure 4.1: m-rep representation of a tubular object

object being modelled (as shown in Figure 4.1). Conditions restrict folding in the surface of this model. Principal component analysis (PCA) cannot be applied to this representation because it is not Euclidean, and therefore, Principal Geodesic Analysis (PGA) is developed, which applies PCA using geodesic distance on the manifold (Fletcher et al., 2004). This method shows potential for modelling complex shapes using a hierarchy of m-rep models but has not (as of yet) been applied to branching tubular structures.

Tubular structures would most likely be represented by a single chain of m-reps. The m-reps are defined by the two intersections of a maximal sphere; local surface variation as well as local stenosis would have a great effect on the representation, causing instability. The size of the m-rep also means that local protrusions may not be captured and will require additional subfigures (connected local m-rep models). Choosing constraints that will allow a robust representation of the airways and using subfigure m-reps consistently across the dataset could be challenging for branching tubular structures with severe pathology.

These methods have introduced different representations of biological shape that allow registration of multiple cases. The shapes can also be treated as point sets to be registered. Current state-of-the-art methods for registering sets of points include robust point matching (Zheng and Doermann, 2006) and coherent point drift (Myronenko and Song, 2010). Robust point matching uses a graph representation of the points to preserve local relationships during non-rigid registration. Relaxation labelling is then used to find the optimal fit. Coherent point drift registers two point sets by treating one set as Gaussian Mixture Model centroids, and fits these centroids using maximum likelihood and the posterior probabilities. A regularisation term is used to maintain coherence between neighbouring points during registration.

These methods are powerful for registering two point sets with an unknown transform. However, in the problem posed in this thesis, the branching structure and airway centreline can be used to provide corresponding landmark points that can be used to generate a transform, and therefore, better constrain the problem.

In this section, a number of methods to register tubular shapes have been discussed. These include methods that use correspondence of anatomical surface landmarks (Hutton et al., 2003; Paulsen et al., 2002), cylindrical parameterisation (Huysmans et al., 2006; Pinho et al., 2007), medial representations (Pizer et al., 2003a,b) and cross section landmark generation (de Bruijne et al., 2003a; Deligianni et al., 2006). Cylindrical parameterisation and medial representations are powerful techniques but are problematic when it comes to branching tubular structures. The airways lack consistent anatomical landmarks on the surface which could be used for registration of a surface mesh. Instead, the centreline, branch cross sections and bifurcation points define the airway structure for modelling distortion of the airways in space as well as local stenosis. Therefore, surface landmark points generated from the airway structure are the most useful for correspondence, which is an extension of the cross section landmark generation approaches. However, these previous approaches required manual labelling of landmarks or branch points, and assume branch cross sections are parallel to axial slices, which must be overcome for automated statistical airway analysis.

4.2 Method

A method for determining the correspondence between segmented airway trees is introduced, using the projection of pseudo-landmark points onto a surface mesh, and using the pseudo-landmarks to align a template mesh, and finally developing point to point correspondence by local alignment. This method takes advantage of the airway structure to create correspondence, allowing varying degrees of stenosis and deformation to be represented.

4.2.1 Surface point projection

A mesh of the airway surface was used to represent each airway and each face was labelled according to the branching structure of the airway. The triangular surface mesh was constructed from the voxel segmentation, and then smoothed using an implicit fairing method (Desbrun et al., 1999). The smoothing was required to remove noise due to the resolution of the CT scan. Voxel resolution has more of an impact in paediatric airway segmentation compared to adults because the smaller airway volume means that the airways are represented by fewer voxels.

Corresponding landmarks are required to align the airways in the dataset. However, the airways do not have clear landmarks between bifurcations. Therefore, pseudo-landmarks were generated using the topological structure and tubular shape of the branches. The centreline of the airways is represented as a one-voxel thick branching medial line (see Chapter 3). The voxel coordinates of the centreline were then interpolated by linearly interpolating between each set of neighbouring voxels to achieve more detailed representation of the centreline. A moving average smoothing filter was applied to the interpolated points to remove noise added to the centreline from the voxel resolution. The branch is then resampled evenly m times to form landmark points (p_i) on the centreline (see Figure 4.2). At each resampled point (p_i), the tangent to the centreline was calculated and the intersection between the vectors (orthogonal to the tangent) and the airway surface are calculated. This allowed corresponding landmarks to be generated on the surface of the airways based on airway topology and branch curvature .

The unit vector tangent to the centreline at each resampled point (p_i) was found from:

$$T(p_i) = \frac{p'_i(t)}{|p'_i(t)|} \quad (4.1)$$

$$p'_i(t) = \frac{p_i(t+h) - p_i(t-h)}{2h} \quad (4.2)$$

where each p_i can be represented as a function in the vicinity of the landmark and $p_i(t+h)$ is a small progression along the centreline.

Euler angles (x-convention) was then used to construct a rotation matrix that can

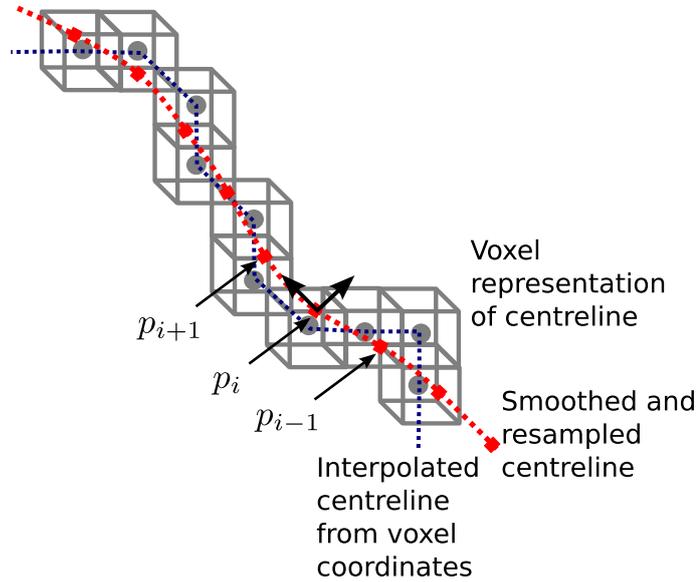


Figure 4.2: The voxel representation of the centreline, the interpolated centreline (blue), and the smoothed and resampled centreline (red). The tangent is calculated for each resampled point (p_i)

be used to map orthogonal vectors onto the centreline (see Figure 4.3). Any rotation can be specified by three angles (Weisstein, b) and, therefore, the general rotation (A) is described by three rotation matrices:

$$A = BCD \tag{4.3}$$

where B, C and D are:

$$B = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}$$

$$D = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where ϕ and θ are found from the tangent vector and $\psi = -\phi$ so that the unit vectors are consistent with the main figure axes. Therefore, a set of vectors (\mathbf{r}_i) on a plane with equidistant angles can be translated to the centreline points (p_i) and rotated so as to be orthogonal to the centreline. Figure 4.3 illustrates the translation and rotation of the 4 vector set \mathbf{r}_i to p_i .

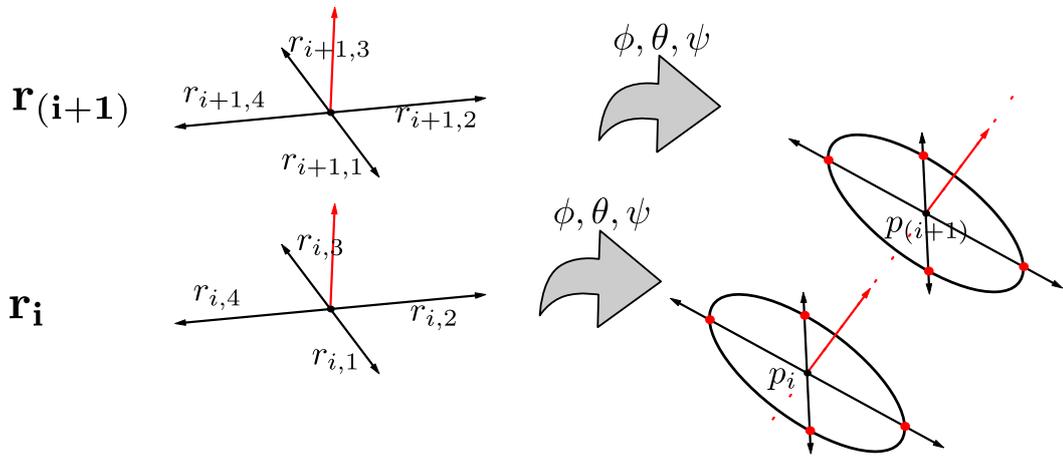


Figure 4.3: Mapping points to the airway surface by ray/triangle intersection of orthogonal vectors. Each set of vectors is translated to p_i and euler angles are used to align the central vector with the centreline tangent. The intersection between the vectors and the airway surface is then found.

The intersection of the vectors (r_{ij}) with the triangular surface mesh was found using Möller and Trumbore's (2005) ray/triangle intersection method. This method determines if the vector falls inside each face when intersecting a plane (with the plane normal defined by the face). Each face must be checked individually. To improve the speed of the search, only faces that fall within a sphere of intersection are analysed. The sphere is chosen as follows:

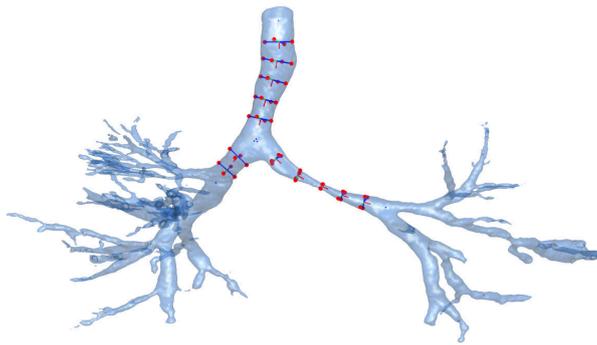
- The sphere of intersection is defined by two constants: sphere centre (α) and radius (β).
- The radius is chosen to be greater than the radii of all branches
- The centre is chosen along the vector of interest ($\alpha = p_i + \beta r_{ij}$).
- Therefore:
 - the point of intersection is guaranteed to fall in the sphere.
 - number of faces to be searched will be reduced

An example of a spherical search region is shown for one of the vectors in Figure 4.4c and Figure 4.5c.

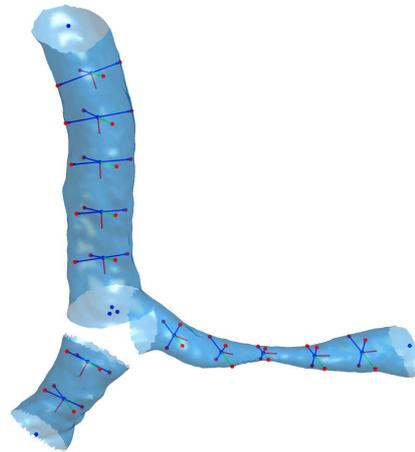
This representation allows a set of corresponding surface points to be mapped onto each branch of each segmented airway. Figure 4.4 and Figure 4.5 show the mapping of surface points onto two example airway segmentations; mapping for the Trachea-RMB-LMB and RMB-RUL-BI regions are shown. Figure 4.4a and Figure 4.4c show the points in relation to the complete airway tree while Figure 4.4b and Figure 4.4d only show an enlarged version with only the branches of interest.

The benefit of this representation is that it is based on the most consistent features of the airway tree, which are the bifurcation points and centreline; there are no useful surface features that can be used to establish correspondence.

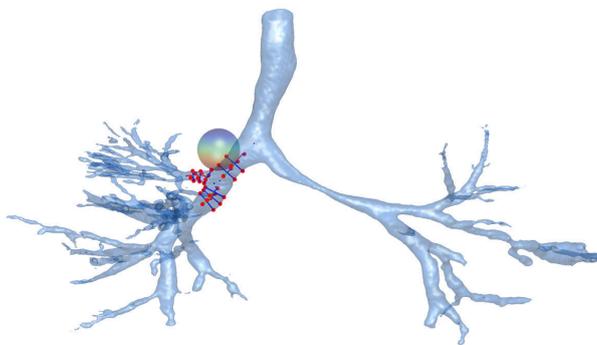
Statistical analysis methods could be applied directly to these corresponding landmark points. This would simplify the method because registration of an airway surface mesh template is not required. Statistical analysis applied directly to the landmark points is demonstrated in Chapter 5. However, using the vertices of a template mesh allows a more detailed representation of the airways. By only representing an airway



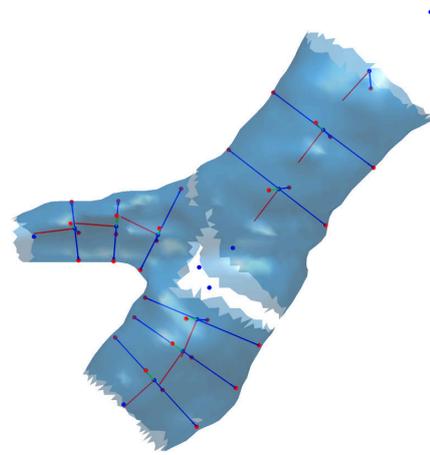
(a) Airway segmentation with landmark points for the Trachea-RMB-LMB region



(b) Enlarged region showing the point mapping to individual branches

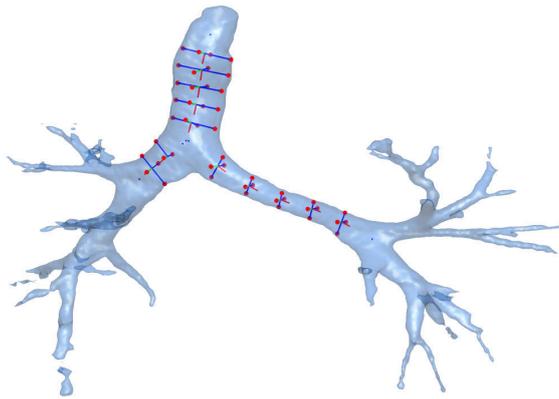


(c) Airway segmentation with landmark points for the RMB-RUL-BI region (with a single sphere showing the area of search for an example vector)

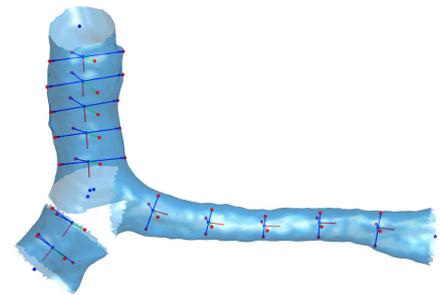


(d) Enlarged region showing the point mapping to individual branches

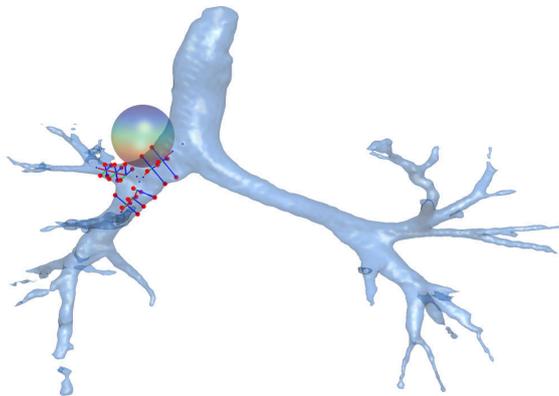
Figure 4.4: Example case 1, mapping corresponding points onto the surface mesh of an airway. Red points represent the surface landmark points obtained from the intersection of the vectors (blue) with the airway surface. Red vectors are tangents to the centreline.



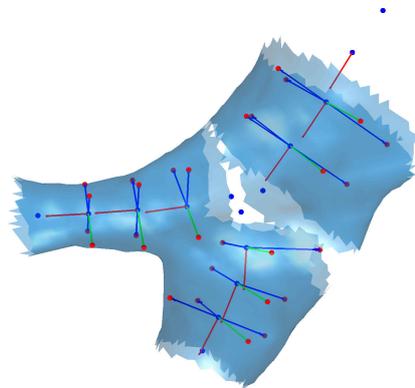
(a) Airway segmentation with landmark points for the Trachea-RMB-LMB region



(b) Enlarged region showing the point mapping to individual branches



(c) Airway segmentation with landmark points for the RMB-RUL-BI region (with a single sphere showing the area of search for an example vector)



(d) Enlarged region showing the point mapping to individual branches

Figure 4.5: Example case 2, mapping corresponding points onto the surface mesh of an airway

with landmark points projected from the centreline, concave regions will be represented by a large number of landmarks and convex regions will be represented by fewer landmarks; this is not ideal. In addition, if a large number of landmarks are chosen there is also a risk that the orthogonal radii will overlap causing folding in the surface representation. Using a small number of landmarks to align a template mesh overcomes these challenges.

The pseudo-landmark points were used to align a template mesh to each airway in the dataset.

4.2.2 Thin-plate-spline alignment

Thin-plate splines (TPS) is a method of interpolation that minimises the bending energy of the surface (Bookstein, 1989, 1997). This method is also useful as a method of non-rigid registration; a surface with one set of landmarks is warped to a corresponding set of landmarks in a physically realistic way.

TPS in 3D is defined by the minimisation of the bending energy E .

$$E = \int \int \int_{R^3} \left(\frac{\partial^2 f}{\partial x^2} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 + \left(\frac{\partial^2 f}{\partial z^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial y \partial z} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial z} \right)^2 dx dy dz \quad (4.4)$$

The function inside the integral defines the the bending energy due to a small deformation at a point. The energy is integrated over the entire space and minimised. The thin-plate spline function in 3D that transforms a point $\mathbf{p} = (x, y, z)$ by minimising the 3D version of Equation 4.4 is:

$$\mathbf{f}(\mathbf{p}) = \sum_{i=1}^k w_i U(|\mathbf{p} - \mathbf{P}_i|) + a_0 + a_x x + a_y y + a_z z \quad (4.5)$$

where $\mathbf{f}(\mathbf{p}) = [f_x(\mathbf{p}), f_y(\mathbf{p}), f_z(\mathbf{p})]$ is the new position of the point (\mathbf{p}) . \mathbf{P}_i are the k landmark points on the shape and w_{ij} are the weighting factors. $U(r) = r^2 \log(r^2)$ for the 2D case and $U(r) = |r|$ for the 3D case, where $r = \mathbf{p} - \mathbf{P}_i$.

The weighting factors w_{ij} are found by solving Equation 4.5 but replacing \mathbf{p} with each landmark point \mathbf{P}_i in turn and replacing \mathbf{f} with the transformed landmark $\tilde{\mathbf{P}}_i$. This problem can be rewritten as:

$$W = L^{-1}H \quad (4.6)$$

where

$$L = \begin{pmatrix} K & Q \\ Q^T & O \end{pmatrix} \quad H = [\tilde{\mathbf{P}}_1 \tilde{\mathbf{P}}_2 \dots \tilde{\mathbf{P}}_k 0 0 0 0]^t$$

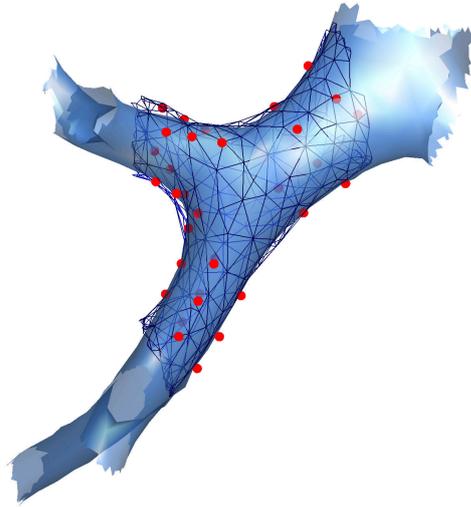
$$K = \begin{pmatrix} 0 & U_{12} & \dots & U_{1k} \\ U_{21} & 0 & \dots & U_{2k} \\ \vdots & \vdots & \dots & \vdots \\ U_{k1} & U_{k2} & \dots & 0 \end{pmatrix} \quad Q = \begin{pmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_k & y_k & z_k \end{pmatrix}$$

where O is a 4×4 matrix of zeros. Once the vector of weighting factors W is obtained, the warp of any point can be obtained using Equation 4.5.

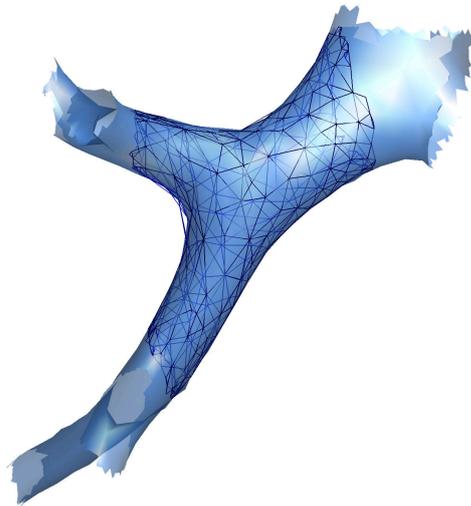
One airway from the dataset was selected as a template and a TPS warp was applied to align the template with each airway in order to represent each airway in the dataset with a corresponding mesh. Figure 4.6a and Figure 4.7a show templates that are warped by TPS to two regions of the airway.

4.2.3 Local mesh alignment

The TPS warp will only be exactly aligned at the landmarks and further matching is required so that the template mesh is exactly aligned with each target mesh (as shown in Figure 4.6 and Figure 4.7). The simplest method is to project the template mesh to the closest point on the target mesh (Hutton et al., 2003; Paulsen et al., 2002) but this can lead to unrealistic deformation and movement to the closest point while not representing small deformations. Figure 4.8a shows mesh misalignment, by projecting to the closest point, because of narrow sections caused by stenosis. Figure

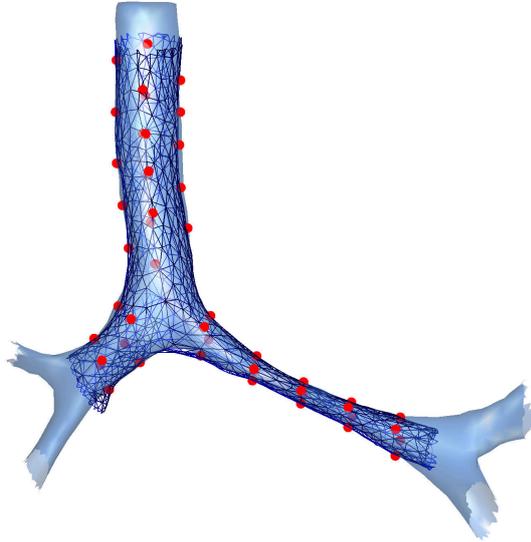


(a) Thin-plate spline warp by aligning landmark points (red) of a template mesh to the landmark points of a region of an airway

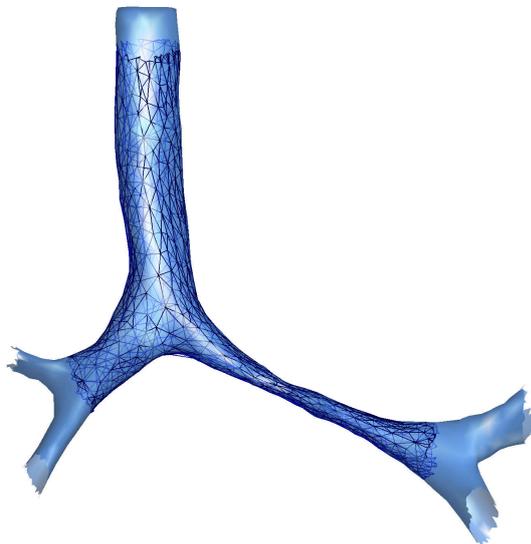


(b) Local alignment (performed after TPS warp) of each template vertex to the airway surface

Figure 4.6: Registration of a template mesh to an airway mesh for the RMB-RUL-BI region



(a) Thin-plate spline warp by aligning landmark points (red) of a template mesh to the landmark points of a region of an airway



(b) Local alignment (performed after TPS warp) of each template vertex to the airway surface

Figure 4.7: Registration of a template mesh to an airway mesh for the Trachea-RMB-LMB region

4.8b shows the improvement with the proposed method.

Meller and Kalender (2004) and Kaus et al. (2003) optimize the fit based on the distance between the meshes while an additional force preserves the mesh structure. Our method is based on these methods but is enhanced for deformed tubular objects by adding a third term based on the surface orientation. For each vertex (t_i) on the template mesh, a force ($F_{i,tot}$) is calculated to direct the warp. This consists of a component based on closest point on the object mesh (r_i) (Equation 4.7) and an internal forcing component is included to preserve the size of the faces. The force is calculated by the change in the distance of each of the m neighbouring vertices to a vertex t_i from the initial distance v_{0j} (Equation 4.8). These two forces are not enough to match to small indentations or protrusions found in the airways. Therefore, a local *inflation/deflation* force is also applied based on the normal of each vertex \hat{n}_i (calculated from the normal of the surrounding faces). The amount of inflation and deflation is controlled by the distance and direction of the target mesh to $F_{i,1}$ (Equation 4.9).

$$F_{i,1} = \vec{r}_i - \vec{t}_i \quad \text{nearest point ext. force} \quad (4.7)$$

$$F_{i,2} = \sum_j^p \hat{v}_j (||\vec{v}_j|| - ||v_{0j}||) \quad \text{where } \vec{v}_j = \vec{t}_j - \vec{t}_i \quad \text{internal force} \quad (4.8)$$

$$F_{i,3} = \hat{n}_i (\hat{n}_i \cdot F_{i,1}) \quad \text{normal ext. force} \quad (4.9)$$

$$F_{i,tot} = \alpha F_{i,1} + \beta F_{i,2} + \gamma F_{i,3} \quad (4.10)$$

In Equation 4.10, the forces are weighted with α , β and γ . This procedure is applied iteratively until stability is reached. Figure 4.8 shows the effectiveness of this method for narrowed and stenosed tubular regions. In Figure 4.8a, the closest point method fails in stenosed regions. In Figure 4.8b the proposed method overcomes these issues. The mesh after this local alignment is also shown in Figures 4.6b and 4.7b.

This method enables any segmented airway to be registered to a template mesh, and, thus, achieving vertex correspondence.

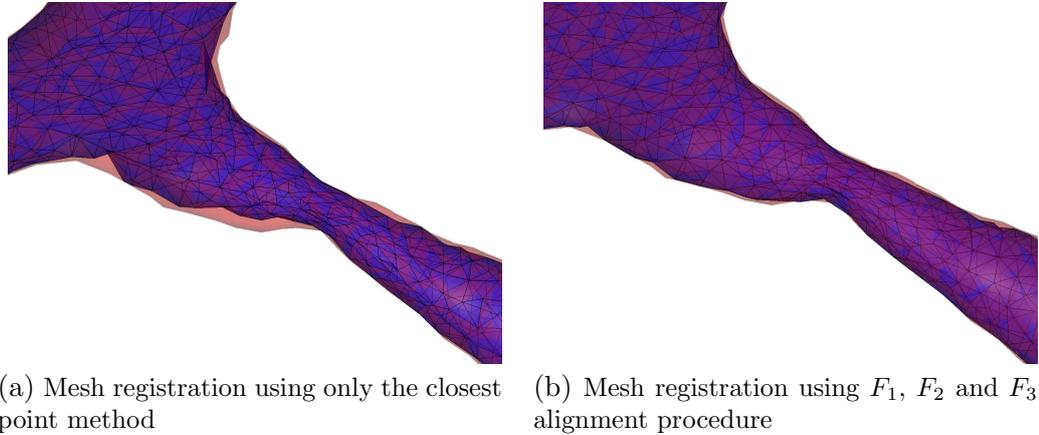


Figure 4.8: Comparison between the mesh representation of a single airway and the registered template mesh. The closest point method shows greater error when branches are narrowed from stenosis

4.2.4 Local shape models

The previous sections have introduced a method to register the surface mesh of segmented airway trees. This method can be applied to the whole airway tree or to a region of the airway. The Figures 4.4, 4.5, 4.6 and 4.7 illustrate regional based airway analysis. This section discusses a method to accurately extract a region of interest to aid analysis.

In Chapter 3, each mesh face ($m \in M$) was assigned a branch label by assigning the label of the closest voxel of the labelled skeleton (s_c):

$$s_c := \operatorname{argmin}_{s \in S} \|m - s\| \quad (4.11)$$

$$l(m) := l(s_c) \quad (4.12)$$

where M is the set of all mesh face centroid coordinates, S is the set of all voxels in the skeleton and each voxel has a branch label ($l(s) \in L$). The mesh representing each branch is then defined as $B_j := \{m | l(m) = l_{B_j}\}$ where l_{B_j} is the label of a given branch B_j

However, this labelling criterion introduces a number of issues when finding correspondence. Mesh faces in a bifurcation region are assigned to the connected branches

but this assignment is biased towards branches with a smaller diameter and is influenced by branch shape and orientation. This leads to irregular and noisy branch end points which impacts the shape analysis.

In order to improve the stability of the branch labelling, adjustments are made to this assignment. Using a neighbourhood ($neigh(B_j) := \{m | m \in B_j \vee m \in child(B_j) \vee m \in parent(B_j) \vee m \in sibling(B_j)\}$), defined as the meshes of all connected branches to the current branch, a subset of $neigh(B_j)$ are relabelled as the current branch (l_{B_j}):

$$\begin{aligned} \hat{B}_j := \{m | m \in neigh(B_j) \wedge \hat{n}_{j1} \cdot \hat{n}_{m1} > \alpha \wedge \hat{n}_{j2} \cdot \hat{n}_{m2} > \alpha \\ \wedge |m - e_{j1}| + |m - e_{j2}| < \beta |e_{j1} - e_{j2}|\} \end{aligned} \quad (4.13)$$

where e_{j1} and e_{j2} are the start and end points of the branch. n_{m1} , n_{m2} are the directions from e_j and end points to the centroid of face m (see Figure 4.9). α and β are constants identified from the training set. The first line of Equation 4.13 allows faces from neighbouring branches to be included, the second line limits the branch between the start and end point and the third line restricts the distance from the branch that voxels can be included (to avoid potential inclusion of disconnected regions from other branches). Each face (m) can have more than one label and m that are not assigned to a branch are labelled as bifurcation points. This method reduces the variance due to region selection by creating smooth and consistent edges between labels.

4.3 Evaluation

To evaluate the accuracy of the template registration, the registered template mesh was compared to the original meshes for each airway. The quality of the registered mesh was also assessed.

4.3.1 Mesh registration

Parameters of the mesh warping procedure were chosen to obtain an accurate warp of the template. The weighting factors (α , β and γ) of the optimisation forces were

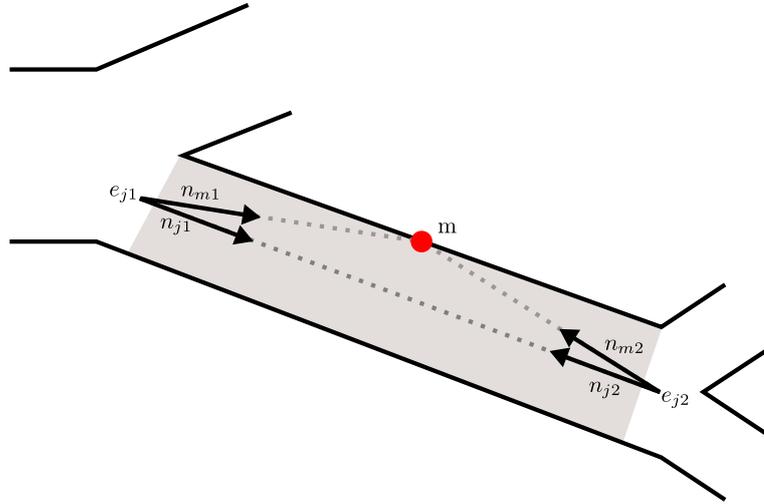


Figure 4.9: Relabelling the surface mesh where m is the face to be relabelled. m is labelled according to the face's position in relation to the branch end points e_{j1} and e_{j2}

varied using a grid search. The results were compared by generating a volume from each original mesh and template mesh, and calculating the difference:

$$V_{dif} = (V_{temp} \setminus V_{case}) \cup (V_{case} \setminus V_{temp}) \quad (4.14)$$

where \setminus is the relative complement and \cup is the union of the sets. By combining results for the entire dataset, the most suitable parameters could be chosen. The volumes V_{temp} and V_{case} were generated by re-voxelisation of the airway meshes; voxels inside the mesh were assigned to the volume¹. Figure 4.10 shows a section of the airway mesh and the generated volume.

A proportion of this error is due to small differences between the mesh and template mesh. This is because the template mesh is reduced in size and is made up of larger faces. This will not have a large effect on the model and we are more interested in avoiding large local errors. Some of the error is also due to the discrete voxel representation of the volume during the evaluation. Therefore, a morphological opening

¹The voxel density was chosen to be 5 times the original voxel density in each dimension, resulting in the volume containing 125 times as many voxels as the original segmentation. Voxel density was increased to better simulate the smoothed airway mesh

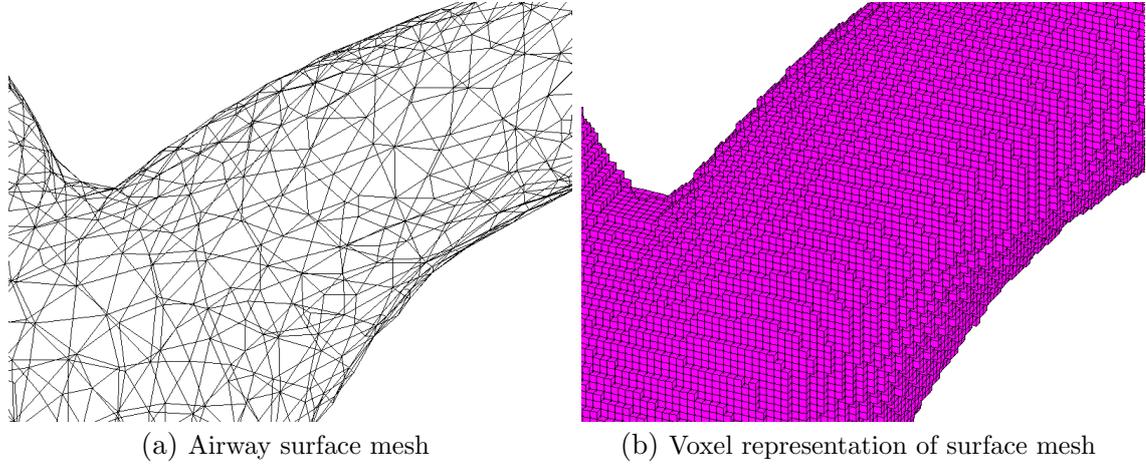


Figure 4.10: Generating a volume from an airway surface mesh. The volume is represented by voxels that each have a volume $1/125$ of the original voxels in the airway segmentation

was applied in order to obtain a second metric:

$$V_{open} = V_{dif} \circ K \quad (4.15)$$

where K is a 6-connected kernel. This metric only includes errors that have a *thickness* greater than $2/5$ of a voxel (2 resampled voxels). As the resolution of the airway is defined by the airway segmentation voxel size, removing errors less than half a voxel, was considered a more useful metric.

4.3.2 Mesh quality

Changes in mesh quality after registration are also considered in the evaluation. Two useful measures of mesh quality are: the ratio of sides of each triangle in the mesh and the ratio of the radii of the incircle and circumcircle (Pebay and Baker, 1991, 2003).

The radii of the incircle(r) and circumcircle (R) can be calculated from the sides of a triangle (a , b and c), as shown in Figure 4.11.

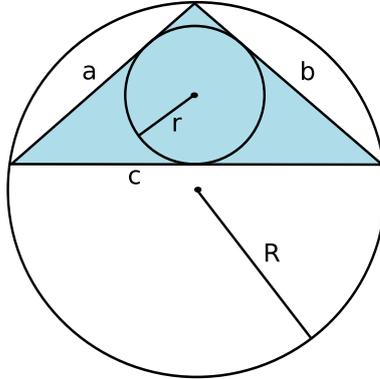


Figure 4.11: A triangle (sides a , b and c) with incircle (radius r) and circumcircle (radius R)

The incircle radius (r):

$$r = \frac{A}{s} \quad (4.16)$$

$$(4.17)$$

and the circumcircle radius (R):

$$R = \frac{abc}{4A} \quad (4.18)$$

where A is the area of the triangle given by Heron's formula and s is the semi-perimeter.

$$A = \sqrt{s(s-a)(s-b)(s-c)} \quad (4.19)$$

$$s = 0.5(a+b+c) \quad (4.20)$$

These values were used to calculate quality metrics for each triangular face where:

$$q_1 = \frac{\min(a, b, c)}{\max(a, b, c)} \quad (4.21)$$

is the ratio of the smallest and largest sides and

$$q_2 = \frac{2r}{R} \quad (4.22)$$

is the ratio of the radii (Pebay and Baker, 1991, 2003). Both q_1 and q_2 are equal to 1 for equilateral triangles.

Degenerate triangles are triangles with a volume of 0 and creation of these should be avoided when registering a mesh. q_1 and q_2 are useful measures as they provide tools for measuring the two types of nearly degenerate triangles: needle and flattened triangles. Pebay and Baker (2003) define a needle triangle as a triangle where one and only one angle is almost 0. A flat triangle is where one angle is almost π . $q_1 \rightarrow 0$ as the triangle becomes a needle and $q_1 \approx 2$ as the triangle flattens. q_2 is a measure of volume and, therefore $q_2 \rightarrow 0$ as the triangle flattens or becomes a needle². Therefore, these non-dimensional measures are 1 for equilateral triangles and tend to 0 for types of degenerate triangles, and provide tools for assessing the change in quality of an unstructured mesh during registration. The distribution of q_1 and q_2 was found for each mesh before and after registration and mean values were found for the entire dataset.

4.4 Results

This section presents results of the mesh registration by evaluating the accuracy of the registration and the quality of the mesh representation using the 89 case training set.

4.4.1 Algorithm parameters

As discussed earlier, the centreline (originally represented by voxels) was interpolated and smoothed. Linear interpolation between the voxel coordinates was performed and 100 points were sampled between each voxel. This was chosen to be suitably dense, from which, to derive resampled equally spaced landmark points. The centreline was

²In this thesis, an inverse of the metrics used by Pebay and Baker (2003) is used so that the measures lie between 0 and 1, instead of 1 and ∞

then smoothed using a moving average (as shown in Figure 4.12). The number of points used in the smoothing was qualitatively assessed using example cases from the training set. Five voxels were used to calculate the moving average at each point as this removed the appearance of individual voxel noise. Therefore, 500 points were used to calculate the smoothed centreline from the interpolated points.

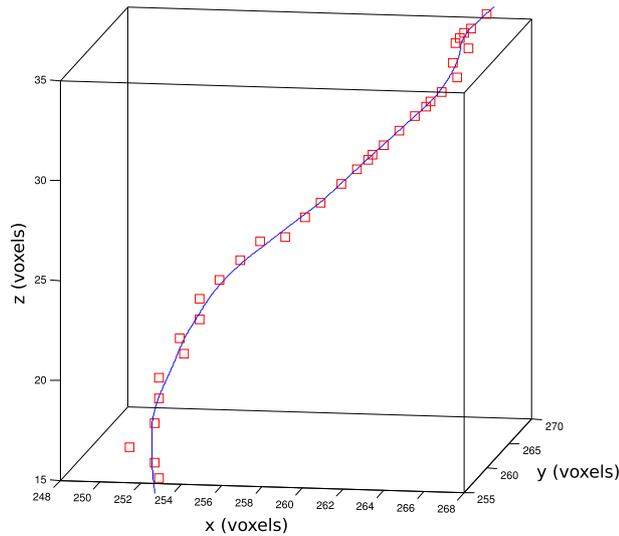


Figure 4.12: Section of the airway centreline showing the voxel coordinates (squares) and the smoothed and interpolated centreline (derived from the voxels)

A single case with no visible pathology (from the training set) was used as the template for registration. Choosing a different template had very little effect on the accuracy of the registration. A mean template generated from the training set could be used instead. However, as each airway segmentation is represented by meshes of varying sizes (dependent on the number of voxels representing each airway), generating a mean model would require a second registration procedure. The number of faces in the template was reduced to 20% of the initial mesh using Matlab's built-in `reducepatch` function. Therefore, the Trachea-RMB-LMB region was represented by 3286 faces and 1733 vertices, and the RMB-RUL-BI regions were represented by 1093 faces and 618 vertices. The number of vertices was decreased to reduce computational time.

The number of resampled points used, for radius based measurements, was [100 100 50] (number of points in each branch of a 3-branch section). At each of these

equidistant points, 4 points were projected to the surface to generate pseudo-landmark points. For registering the template mesh, many fewer points were used:

- Trachea-LMB-RMB [5, 5, 2] resampled points on each bronchi
- RMB-RUL-BI of [3, 3, 3] resampled points on each bronchi

These values were chosen so that the landmarks were suitably far apart that the orthogonal radii never intersected – which would cause folding in the surface representation.

4.4.2 Parameters and accuracy of mesh alignment

The accuracy of the mesh alignment was evaluated by comparing volumes generated from the template mesh and airway region of interest. Figure 4.13 shows contour plots of the proportion of error in the registration (V_{dif}) and a 1-voxel opening of the error (V_{open}) as a function of α and γ with $\beta = 1.0$ (Equation 4.10). For stability, only $\alpha < 1.0$ and $\gamma < 1.0$ were considered when choosing optimal values. Optimal values were found to be $\alpha = 0.1$ and $\gamma = 0.8$ with a mean $V_{dif} = 0.021 \pm 0.009$ and $V_{open} = 0.0022 \pm 0.0049$. The error for each airway registration, which is used to calculate the mean values, is shown in Table C.7 (Appendix C). Just using closest point mapping (Equation 4.7): $V_{dif} = 0.029 \pm 0.019$ and $V_{open} = 0.009 \pm 0.015$. Using closest point mapping and the mesh structure term (Equations 4.7 and 4.8): $V_{dif} = 0.027 \pm 0.018$ and $V_{open} = 0.006 \pm 0.011$. Therefore, including the *inflation/deflation* term improves the registration of the template and reduces the variance in the accuracy of the fit. This is particularly noticeable for V_{open} (larger local errors) and agrees with qualitative observations that adding the third term improves the registration for narrowed and stenosed branches. Results for all optimisations are available in Appendix C. This optimisation could be performed on an individual basis to choose parameters for each case. However, in this work the same fit parameters are used throughout the dataset.

Mesh quality was also considered before and after registration using the parameters (α , β and γ) chosen from the optimisation procedure. Figure 4.14 and Figure 4.15 show histogram plots of the mesh quality measures, q_1 and q_2 , respectively. These plots show a) the mean of the template before local alignment (after TPS warp) b) local alignment using closest point method c) local alignment using closest point

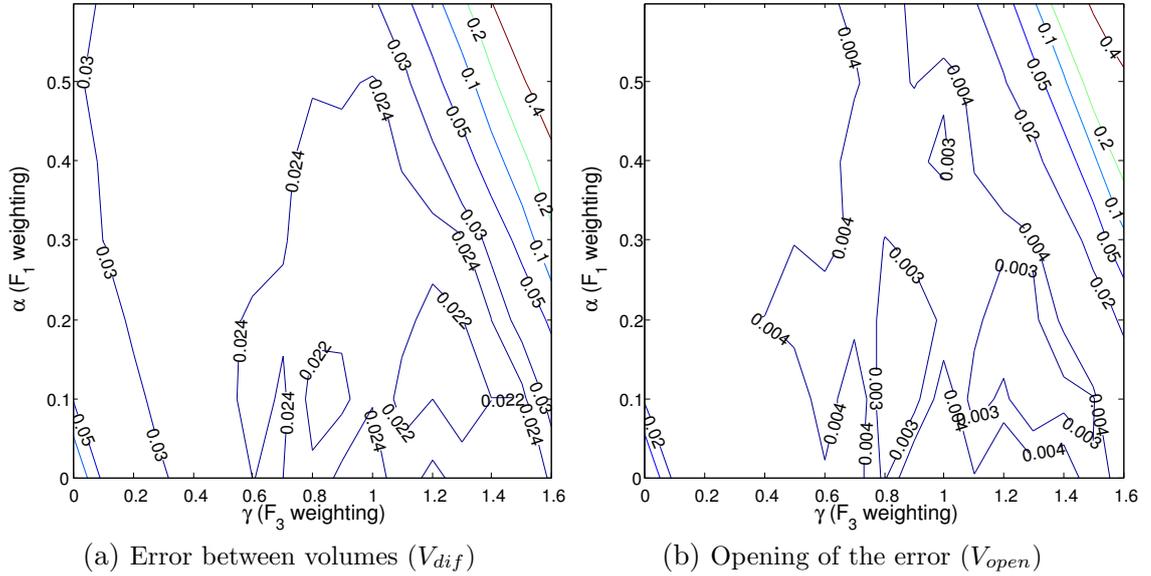


Figure 4.13: Contour plots showing the proportion of error when comparing the accuracy of the registered template mesh to the original mesh. The error is shown as a function of the weighting parameters α and γ of F_1 and F_3 , respectively ($\beta = 1$).

method with mesh structure preservation and d) the proposed method. The plots are calculated for each airway and from this the mean histogram is calculated (an example distribution for a single case is shown in Figure C.1 of Appendix C). Each mesh face in an airway tree varies in shape due to the mesh extraction and reduction procedures.

For consistency across airways, it is beneficial that a method that preserves the mesh shape as much as possible while optimising the fit. Particularly, the creation of nearly degenerate triangles (low q_1 and q_2) should be minimised. Figure 4.11 shows that while the distribution of the faces remain very similar, using the proposed method ($\alpha F_1 + \beta F_2 + \gamma F_3$) reduces the number of triangles with low q_1 and q_2 compared to using just closest point (F_1) or closest point with mesh shape preservation ($\alpha F_1 + \beta F_2$). Just using closest point introduces a large number of poor triangles. The differences are small but the figures show that using all three terms reduces that number of poor triangles.

Figures 4.16 and 4.17 show the registration of the template mesh to the entire test set. These cases contain TB and non-TB paediatric patients and are used in Chapter

6 for airway shape evaluation.

4.5 Discussion and Conclusions

In this chapter, a method to automatically create correspondence between airway surface meshes is introduced to allow modelling of airway shape variation. Initially, corresponding points are placed on the airway surface based on the airway branching structure and centreline. This is related to previous cross-sectional landmarking methods (de Bruijne et al., 2003a; Deligianni et al., 2006), however, the method extends and automates the previously semi-automatic methods by using the branching structure and 3D volume to identify the landmarks instead of axial cross sections. These landmark points are then used to guide the registration of a template mesh using TPS and local vertex alignment. Local vertex alignment is applied and modified in a novel way, for tubular objects with stenosis, by including a local *inflation/deflation* term to the optimisation, which improves results. This method is used to develop correspondence from airway segmentations where every vertex in each airway mesh corresponds to an equivalent vertex on every other mesh. The author is not aware of any previous methods to register airway surface meshes from multiple patients. Finally, a novel method to consistently extract local regions of the airway is introduced.

The optimal method was able to register a simplified template mesh to the airway dataset with a mean error of $2.1 \pm 0.9\%$. However, single voxel errors are likely to be caused by the re-voxelisation method used in the evaluation. Errors larger than 1 voxel (i.e. errors that could not be explained by the re-voxelisation procedure of the mesh faces) contributed to $0.22 \pm 0.49\%$. The local normalisation performed better than previous methods for the airway, both in terms of registration accuracy and mesh quality.

This airway correspondence method can be used to build statistical shape models of the airways.

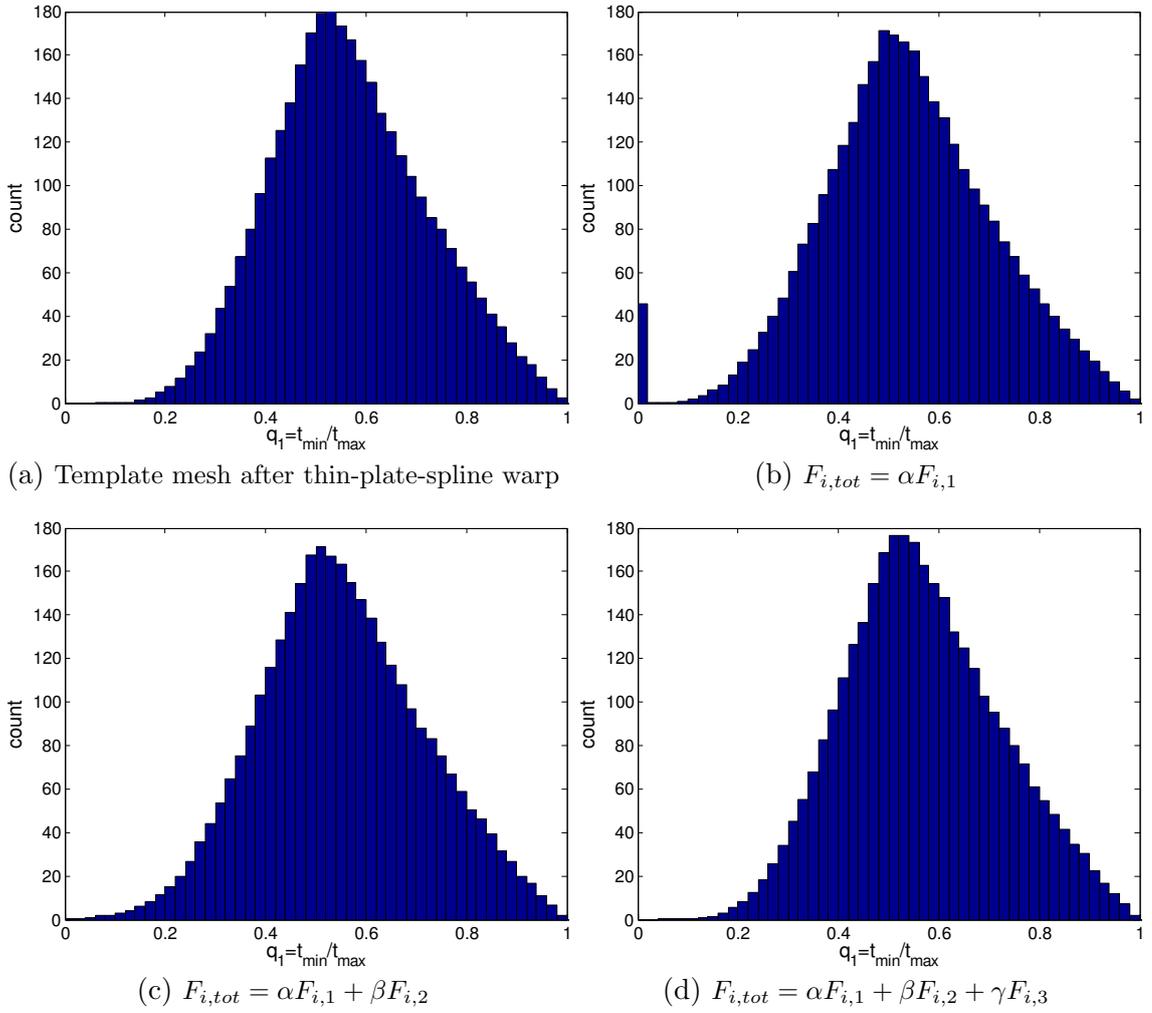


Figure 4.14: Mean histogram plots of the mesh quality measure q_1 (ratio of shortest and longest edge of each face) for the entire training set. a) shows the mean template before vertex alignment b) shows the mean template using the closest point method c) shows the mean template using the closest point with mesh preservation term d) shows the proposed method

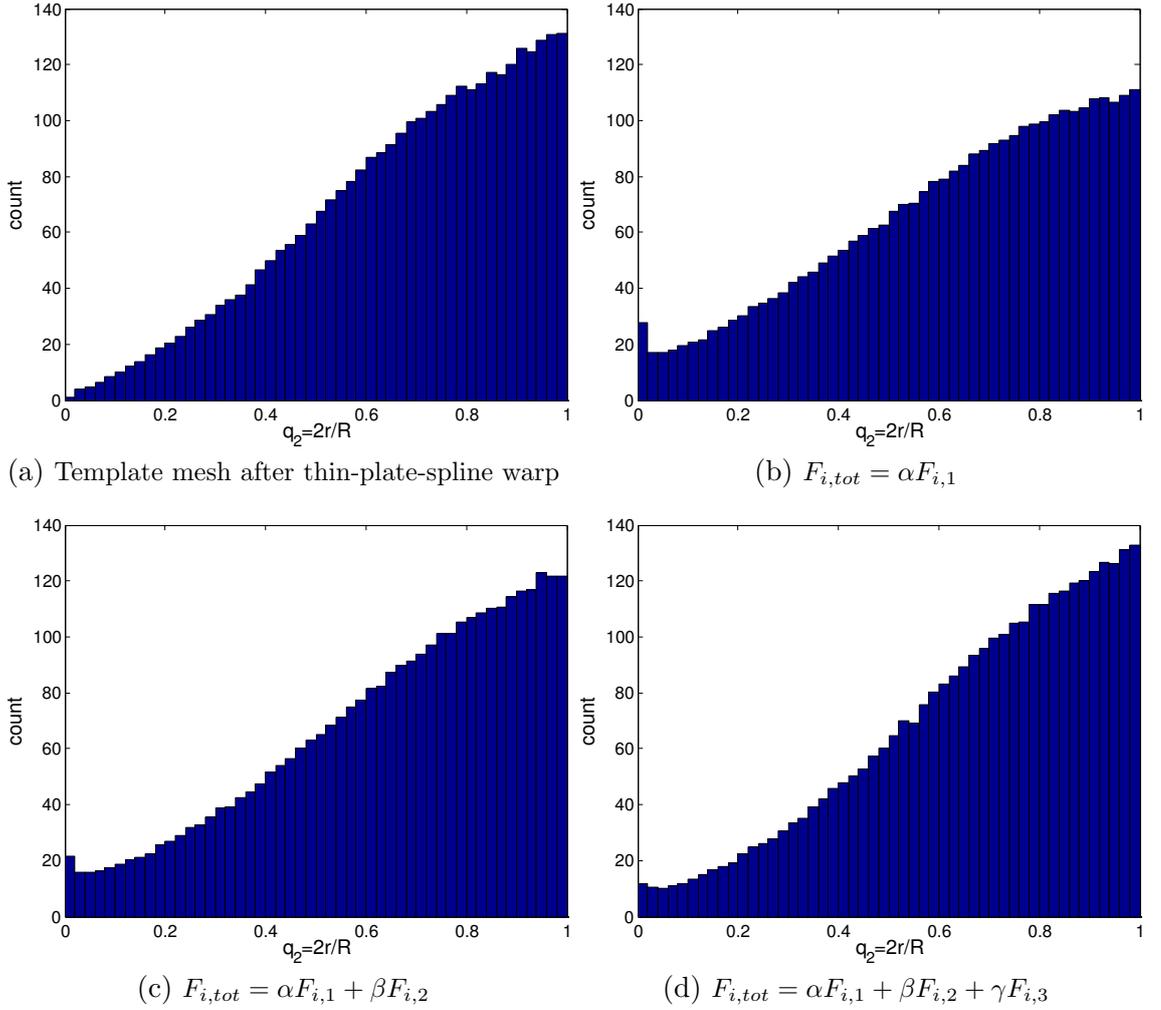


Figure 4.15: Mean histogram plots of the mesh quality measure q_2 (ratio of face incircle and circumcircle radii) for the entire training set. a) shows the mean template before vertex alignment b) shows the mean template using the closest point method c) shows the mean template using the closest point with mesh preservation term d) shows the proposed method

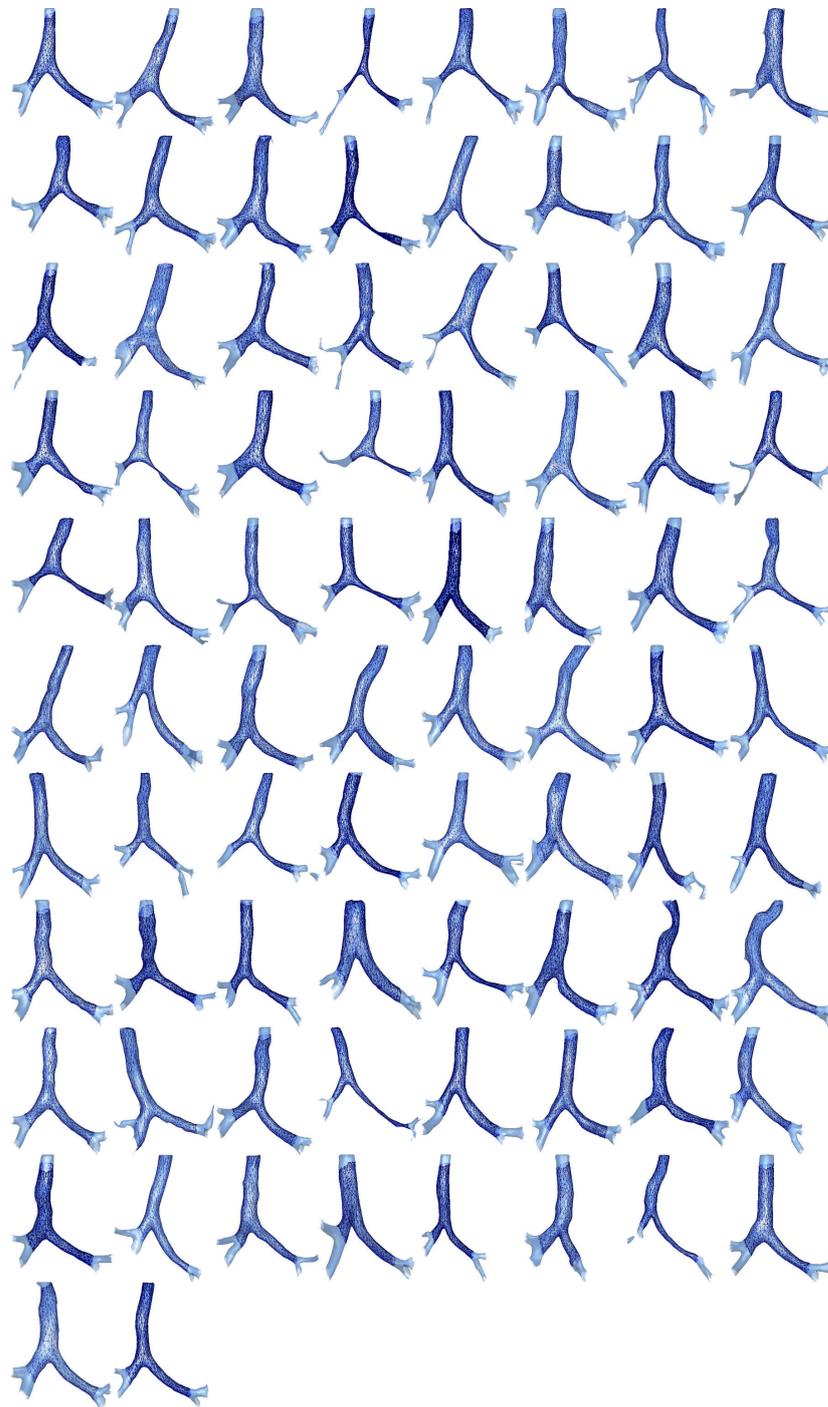


Figure 4.16: Registration of the template mesh to the trachea-RMB-LMB section for the entire test set. The first 36 cases are patients diagnosed with TB and the remainder are controls. Obstructed regions are automatically excluded (discussed in Chapter 6).

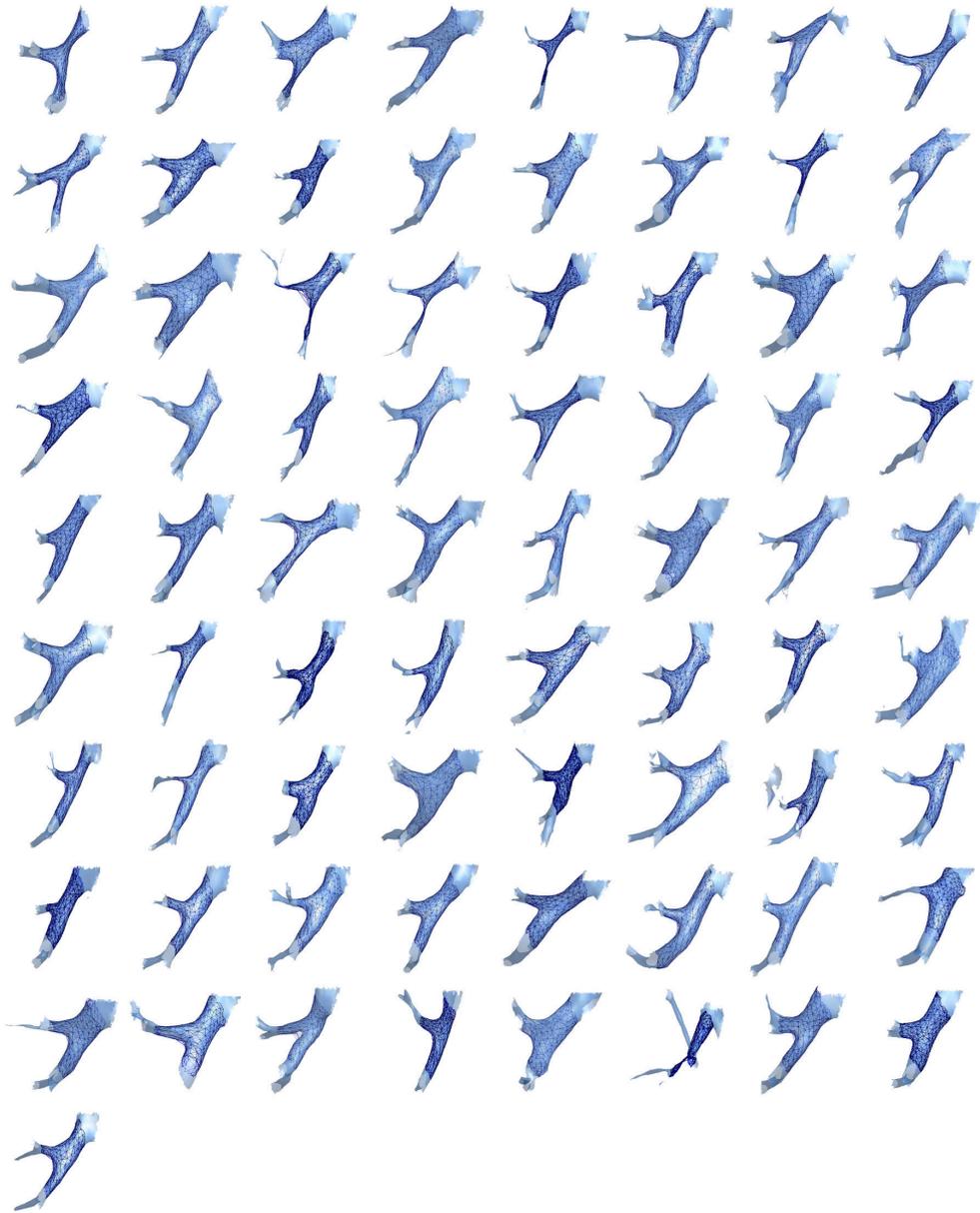


Figure 4.17: Registration to the template mesh to the RMB-RUL-BI section for the entire test set. The first 31 cases are patients diagnosed with TB and the remainder are controls. Obstructed regions are automatically excluded (discussed in Chapter 6).

4.6 Chapter summary

This chapter presents a method to develop vertex-to-vertex airway correspondence:

- Airway correspondence method consists of:
 - Centreline smoothing and resampling
 - Projection of orthogonal vectors from centreline
 - Intersection of vectors with the mesh using ray/triangle intersection
 - Alignment of the template mesh using the pseudo-landmark points
 - Local alignment of the template mesh to the airway using the terms:
 - * Closest point (F1)
 - * Mesh structure preservation (F2)
 - * Local expansion/contraction (F3)
- The method was evaluated using:
 - Analysis of the accuracy of the warp
 - Analysis of the mesh quality
- Results of the evaluation:
 - General error: 2.1 ± 0.9 % (influenced by revoxelisation procedure used in the evaluation)
 - Local error: 0.22 ± 0.49 %
- This method is used to build a point distribution model from the airway dataset in later chapters

5

Obstructed Branch Detection

Abstract. Most airway segmentation algorithms fail when there is severe pathology that leads to the appearance of complete obstruction in a bronchi. In this chapter a method is presented that can be added to existing airway segmentation algorithms to detect and segment beyond obstructions. The method builds a point distribution model of each bronchi, and support vector machines are used to classify regions that have missing branches. A region beyond the obstruction is then searched to detect additional airway regions. This method was evaluated on a dataset containing obstructions in the major bronchi and automatically identified 24 of the 26 obstructed branches; two branches could not be classified because of a lack of data. 18 of 19 disconnected airway regions were segmented.

As mentioned in previous chapters, a number of cases have severe airway compression so that the air filled airway region, usually visible in the CT examination, is discontinuous. This is problematic for airway segmentation algorithms as most rely on the airway being a connected structure.

In this chapter a method is proposed that takes into account the airway tree topology and branch shape, to identify and segment missing branches. This is added

to a previously tested segmentation algorithm described in Chapter 2. The goal is not only to improve the segmentation but also to link the missing branches to the structure of the airway tree – to allow further airway evaluation. In this chapter, the method is used in conjunction with a specific airway segmentation algorithm but has the potential to be applied to most existing algorithms.

The algorithm is evaluated on a paediatric chest CT dataset containing cases of airway compression caused by TB. The algorithm, including detection of missing branches, identification of the point of obstruction and segmentation beyond the obstruction is outlined in Section 5.2. Results from the evaluation of the method on a TB dataset are then presented in Section 5.3.

5.1 Background

As described in Chapter 2, there are a number of existing airway segmentation methods (Aykac et al., 2003; Fetita et al., 2004; Irving et al., 2009; Sonka et al., 1999; Tschirren et al., 2002; van Ginneken et al., 2008). However, these generally focus on adult airways and do not consider obstructions. In order to minimise the likelihood of segmentation of non-airway regions these methods generally require the airway to be completely connected. This is often performed using a region growing step from a seed point in the trachea. If a method is to be effective for the analysis of pathological airway trees then the method must be robust enough to segment branches beyond complete obstructions.

A few methods do not require the airways to be a connected structure and instead rely on the filtering step to minimise leaking. Sonka et al. (1999) propose a rule-based method that uses airway grey-scale, size and vessel adjacency to classify airways. They produce two representations: a 3D connected region and a representation that includes disconnected regions that are also classified as airways. Kitasaka et al. (2010) use filters to enhance tubular structures in the image. Fetita et al. (2004) use a morphological approach that can also be used to identify disconnected airway regions. However the method a requires compromise in the choice of morphological parameters that results in including nonbronchial structures of a similar shape with the disconnected regions. All these methods have the potential to segment disconnected airway structures but also result in considerably increased leaking because

non-airway disconnected structures are also included (Kitasaka et al., 2010; Sonka et al., 1999). Kitasaka et al. (2010) mentions that structures of a similar intensity and shape such as the oesophagus are included, and suggest statistical shape analysis as a possible approach to remove these objects.

Lo et al. (2009b) introduce a method that uses locally optimal paths to segment the airways. This method uses a trained appearance model based on local voxel intensities to determine a cost function. The cost function is then used with Dijkstra's algorithm to determine the optimal paths for airway segmentation. This method was shown to be effective for segmenting additional peripheral branches that have local occlusions, with only a small increase in false positives. However, this method relies on local greyscale values and is, therefore, unlikely to be effective for large obstructions in the larger bronchi. Obstructions of the larger bronchi are uncommon in most diseases affecting the airway but occur in a number of paediatric cases, including cases with TB. In these cases there may be the appearance of no bifurcation (see *missing branches* in Section 5.2.1) or more than half a bronchi missing, and the shape of the airway is expected to be a more useful indicator of missing branches than local greyscale values. Using the expected appearance of each bronchi (trained from a shape model) missing branches can be detected where there may be no other signs. Including expected airway shape in analysis has the potential to increase the robustness of segmentation algorithms, allow analysis of the airways beyond obstructions, and have a number of additional applications beyond airway analysis. There may also be potential to include texture analysis, with shape analysis, to identify the cause of the obstruction and further improve the detection of obstructions.

5.2 Method

This method builds on the airway segmentation, centreline identification and branch labelling algorithms of Chapters 2 and 3. Corresponding landmark points were then identified using the centreline and branch points as a reference (Chapter 4). Each branch is sampled 30 times along the centreline and 8 landmarks are generated on the airway surface for each sampled landmark point. In this Chapter a point distribution model was developed directly from the landmark points instead of from the vertices of a registered template mesh (as used in Chapter 6). This simplifies the algorithm

for shape variation involving missing branches but means that the representation is less detailed than later models.

Point distribution models consist of a set of points that represent each object in the dataset, where a point on one object has a corresponding point on each other object in the dataset (Cootes et al., 1995). The previous steps find a suitable point representation so that shape analysis can be performed to find obstructed branches. Three steps are used to detect new branches:

- Identify the existence of an obstruction
- Find the position of the obstruction
- Find airway beyond the obstruction

Obstructions can be separated into two categories: obstructions that occur late enough in the branch for the obstructed branch to be distinguished from the parent branch (which we will define as *incomplete branches*), and obstructions that occur close enough to the beginning of the branch for there to be no identifiable new branch on the main airway tree (*missing branches*) as shown in Figure 5.1. The separation between these categories depends on the sensitivity of the skeletonisation procedure to detect new branches. However, the more sensitive the skeletonisation procedure is to branching, the greater the chance that false branching will occur. Two separate methods are used to identify these categories of obstruction.

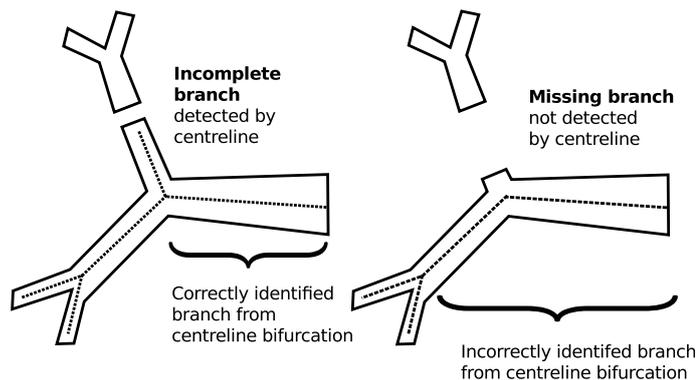


Figure 5.1: Missing and incomplete categories of obstructions. Branches are defined as incomplete if a centreline has been detected and missing if a centreline has not been detected

Figure 5.2 shows an outline of the obstructed branch detection algorithm. Statistical analysis is applied to each branch to identify *missing branches*, from the shape of the included region. Branch diameter is then used to detect the point of obstruction. If no *missing branches* are detected then connectivity of the skeleton is used to identify any *incomplete branches*. The CT volume beyond the obstruction is then searched for additional airway regions.

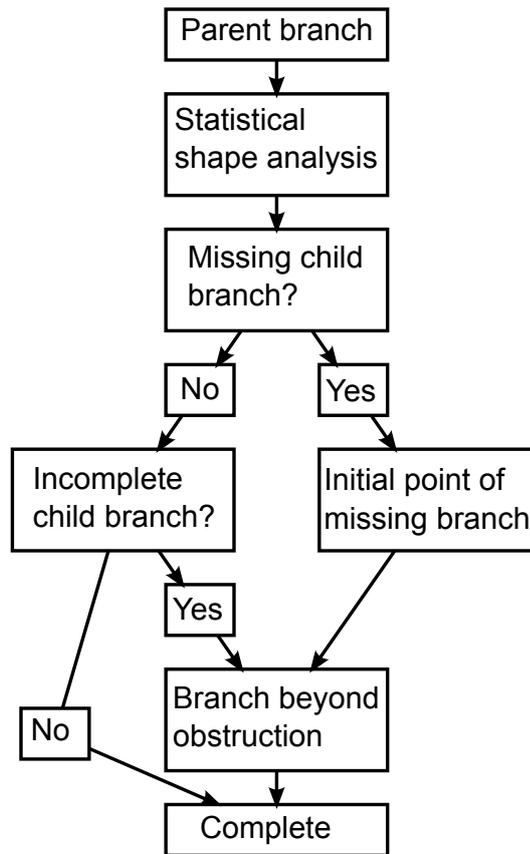


Figure 5.2: Diagram showing the procedure to identify missing child branches

5.2.1 Identification of obstruction

Incomplete branches are easier to detect than *missing branches*. An *incomplete branch* is identified as a branch with no child branches. Therefore, the branching structure of the segmentation is used to identify these obstructed branches. The point of

obstruction is defined as the last point of the centreline of the branch.

Identification of obstruction is more complicated for *missing branches* because the branch point can not be identified from the centreline and, therefore, two branches are detected as one (see Figure 5.1). In our algorithm, statistical shape analysis is used to distinguish normal branches from branch regions containing a missing branch. Principal component analysis (PCA) is applied to the branch landmark points in the dataset and the first four modes are used as a feature vector for classification. Branches in the dataset are manually labelled and used to train a support machine classifier (SVM). This trained classifier is used to identify obstructed branches¹. This results in branches that contain a missing branch being identified for further analysis.

5.2.2 Position of obstruction

Once branches have been identified as containing missing child branches from the classifier, an analysis of the radius of the branch is performed to detect the position of the obstruction. Four diameters, at each branch cross section, are calculated from the 8 landmark points, and the diameter profile of the branch (in these 4 directions) is used to identify the point of obstruction. Local maxima of the diameter profile were used to identify the position of the missing branch.

Figure 5.3 shows two patient profiles with BI obstruction. In both cases the BI is missing and the RMB is incorrectly identified as the *initial RMB region* shown in the figure. After statistical analysis this region is found to be incorrect. The 4 branch diameter profiles are then calculated for the region. By identifying the local maxima, the start point of the missing BI branch is found.

The branch is initially wide, therefore, only local maxima beyond the first third of the branch were used. The optimal direction for analysis was chosen for each region.

5.2.3 Segmentation of disconnected airway regions

Once the obstructed branches have been identified and the point of obstruction found then the next step is to identify the airway beyond the obstruction (*disconnected*

¹Each anatomical branch is trained and classified separately e.g. normal and missing LMB branches are used to train the classifier to detect unseen LMB branches

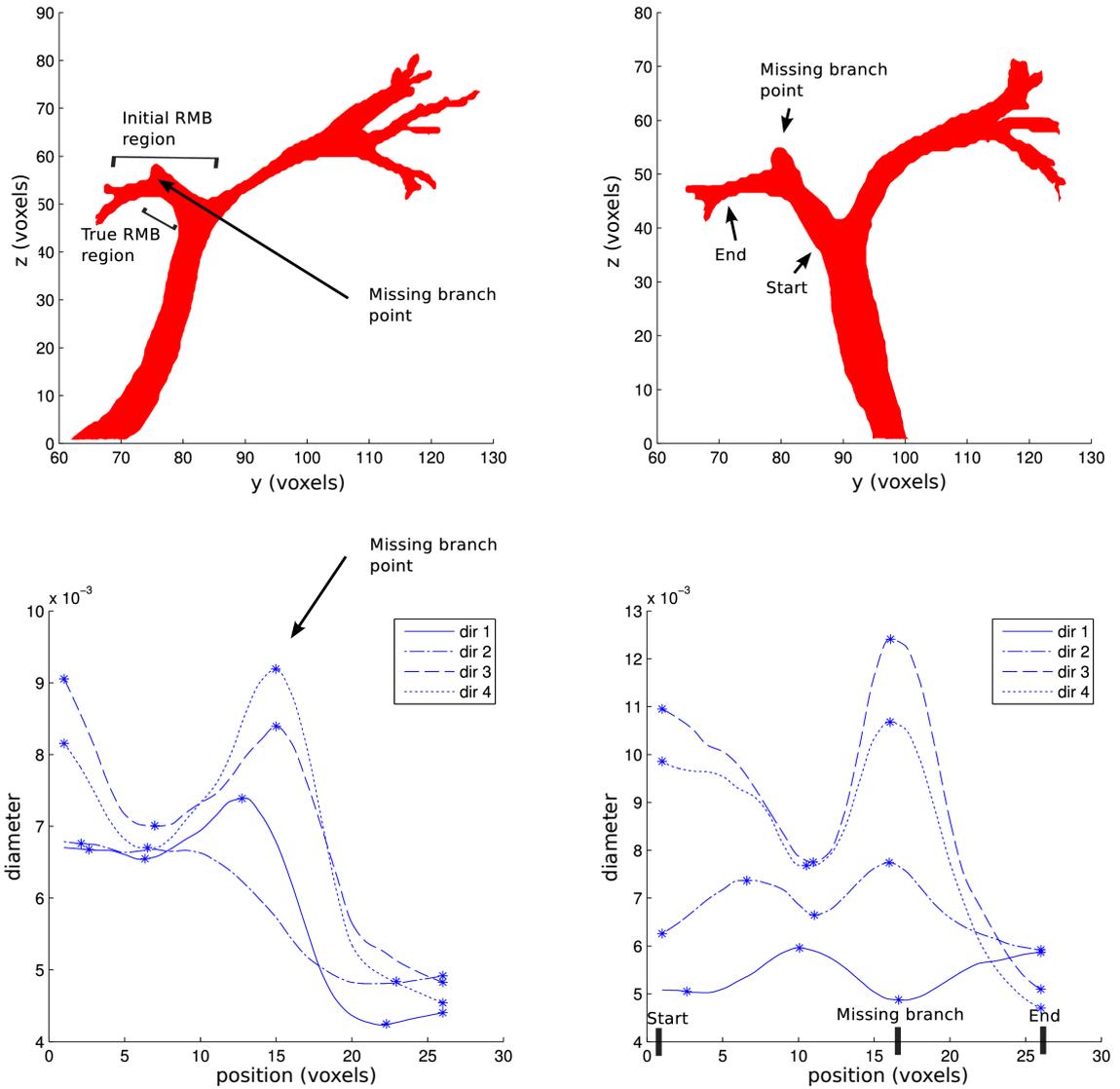


Figure 5.3: Airway shape and branch diameter profiles for obstructed branches for two patients. The top figures show a silhouette of the two patient airways and the bottom figures show branch radius for the RMB bronchus as a function of position along the centreline. The airways are normalised for size (more details in Equation A.2 of Appendix A.1) before branch diameter measurements so the diameters are relative measurements.

airway region). It is expected that if the direction branch is followed past the obstruction, then the remainder of the branch can be found. A cylinder is chosen as a simple way of defining a region where the branch may continue after obstruction. This cylinder is defined by the point of obstruction and the direction of the branch, and used to guide the branch search. Segmented objects that intersect the cylinder are identified and labelled. In most cases only one object is found, however, if more than one exists, a set of heuristics based on size and compactness are used to identify the correct object. The size of the cylinder and the direction of the search was chosen using the cases in the dataset. Cylinder dimensions were chosen as:

- radius = 20 voxels
- length = 27 voxels

The radius and length were chosen to be large enough to include all the disconnected regions in the obstructed dataset. Given more data these parameters should be evaluated on an independent test set.

The direction of the search was chosen as the approximate direction of each branch (the coordinate system is shown in Figure 5.3):

- LMB obstruction $v=[0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$
- BI obstruction $v=[0,0,1]$
- RUL obstruction $v=[0,-1,0]$

These parameters, which are used to search for obstructed branches, are a simple representation of branch directions.

5.3 Results and discussion

5.3.1 Dataset

The dataset consists of 49 chest CT scans from Tygerberg Hospital, Western Cape, South Africa. These are paediatric patients diagnosed with definite or probable TB from positive culture or scope and CT findings. The median age of a patient in

the dataset is 14 months and the maximum and minimum age is 108 months and 2 months, respectively. The scans were acquired on a number of units with voxel size varying from 0.3-0.4 mm in the XY plane and 0.7 -1.0 mm in the Z direction. This dataset contains various degrees of stenosis including complete obstructions. The age of the patients concerned meant that the airways are much smaller than adult airways, and therefore, with fewer branches segmented. Seven cases were removed from the dataset: 3 because of severe movement artefacts, 2 because of scan errors and 2 because a tube in the trachea affected the trachea shape. The algorithm was tested on the remaining 42 cases. Obstructions and disconnected airway regions were manually labelled from the thresholded airway regions of the segmentation – before region growing. These were also verified using the greyscale CT scans. These manually labelled images were used for training the SVM and for comparison. This method does not use the same data as the remainder of this research because a set of cases with obstruction were specifically selected by pulmonologists at Tygerberg Hospital for this research on obstruction.

The method was evaluated on the trachea, right main bronchus (RMB), left main bronchus (LMB), right upper lobe bronchus (RUL) and bronchus intermedius (BI). 2 cases had a missing LMB, 3 cases had a missing RUL and 9 cases had a missing BI. 1 case had an incomplete LMB, 1 case had an incomplete RMB, 3 cases had an incomplete BI and 7 cases had an incomplete RUL.

5.3.2 Detection of missing and incomplete branches

Shape analysis was performed on each branch to find missing child branches. Shape analysis was applied to the RMB to detect missing RUL and BI, and to the trachea to detect missing LMB and RMB. The first 4 PCA modes were used as a feature vector and a SVM with a polynomial kernel was used as a classifier. The SVM was implemented with PrTools (van der Heijden et al., 2004) and more detail on the SVM algorithm can be found in Appendix A.3. A SVM with polynomial degree 2 and $C = 70$ was used in this study. The first 4 modes accounted for 65% of the RMB variation in the dataset. Figure 5.4 shows the landmark representation of the RMB and the shape variation along first PCA mode.

Figure 5.6 shows the landmark representation of the airway and the identification

of the obstruction start point. Figure 5.5 is a plot of the dataset in feature space for the RMB and trachea. The plot uses multidimensional scaling (MDS) to present the 4 dimensions as a 2D plot. MDS is used to project data from a higher dimension to a lower dimension while attempting to preserve the pairwise distances between points (Bishop, 2006; Cox and Cox, 2001). This is only used in the figure to aid visualisation of the separability between classes and not used as part of the classification algorithm. The 3 classes for the RMB could be separated well using the 4 PCA modes where the classes represented: 1) RMB, 2) RMB with RUL “stump” or 3) RMB with BI “stump”. Using SVM with leave-one-out testing, all the RMB were classified correctly as shown in Table 5.1. This method was also applied to the trachea to identify missing LMB (as shown in Figure 5.5). The results look promising, however, with only two obstructed LMB cases and no obstructed RMB cases, a classifier could not be trained.

In future work, simulated data could be used to overcome the lack of a large enough dataset of clinical data. Simulated data could be created by adding obstructions to the mesh representations of existing airway segmentations. This would allow enough obstruction examples to be created for a complete analysis system of the major bronchi. However, the obstructions would need to be created in a way that is representative of the obstructions found in real cases.

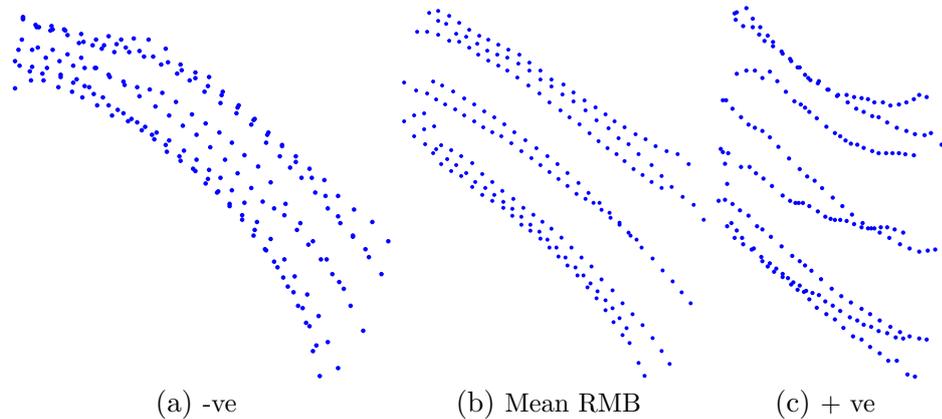


Figure 5.4: Variation along the first PCA mode of RMB

After missing child branches are identified, the algorithm checks the child branches of the remaining airways for further branching. If not then the child branches are incomplete (see Figure 5.2). All cases where the branches had not been labelled as

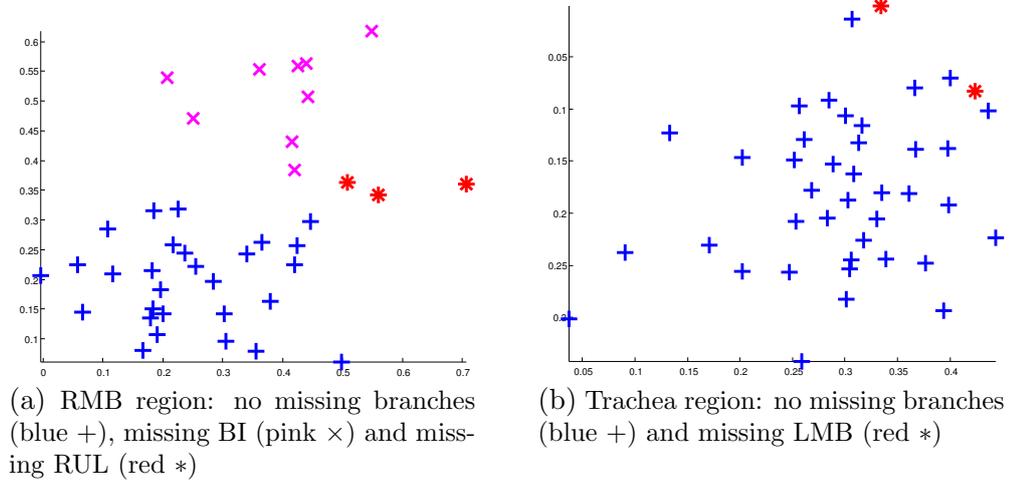


Figure 5.5: MDS reduction to 2 dimensions of 4 dimensional feature space.

True Labels	Estimated Labels			Totals
	1	2	3	
1	28	0	0	28
2	0	3	0	3
3	0	0	9	9
Totals	28	3	9	40

Table 5.1: Leave-one-out cross validation of SVM for the RMB. The class labels are as follows: no missing branches (class 1), missing RUL (class 2) and missing BI (class 3)

missing (by the classifier in the previous step) were checked and obstructions were detected for all cases (see Table 5.2).

5.3.3 Segmentation of disconnected airway regions

Table 5.2 shows the total number of cases with obstructions (*missing branches* and *incomplete branches*) and the total number of cases with disconnected airway regions. In some cases for RUL there are no disconnected branches after obstruction (see Table 5.2). In these few cases the possibility of a small region containing bronchi could not be excluded but were not identifiable during manual labelling. Two airways with obstructed LMB are not listed as being identified for the LMB. The statistical shape method is likely to work for these branches, however, there was not enough

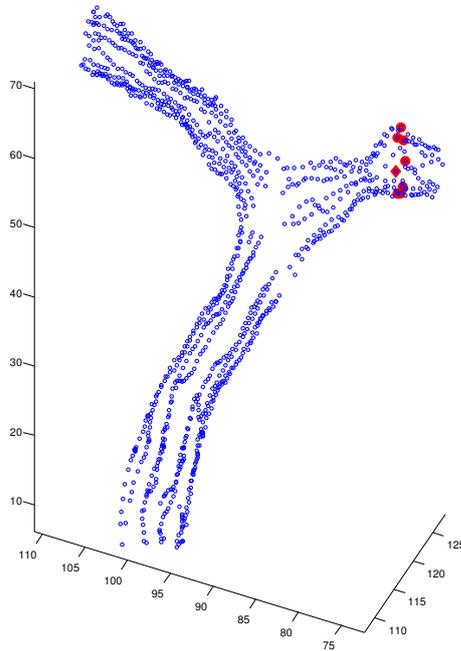


Figure 5.6: Point model with identified obstruction

data to train and test the model on this branch (see Figure 5.5). Assuming correct classification of all three LMB then the second step extracted all three disconnected airway regions correctly. Obstructions and disconnected branches for all 12 cases of the BI branch were identified correctly. “False obstructions” were identified in 3 cases where the airway was continuous in the scan but was not segmented completely because of severe stenosis. This algorithm was also able to identify these disconnected regions that were missed by the original segmentation algorithm. Figure 5.7 shows branch labelling of the original connected region and the addition of the disconnected region. Figure 5.8 shows example segmentations of disconnected airway regions.

In summary, this method automatically identified 24 of the 26 obstructed branches. Two branches could not be classified because of a lack of data. 18 of 19 disconnected airway regions were segmented.

Airway segmentation beyond obstructed regions will improve the effectiveness of segmentation procedures for severe pathology and will improve the viability of automated airway analysis. The authors are not aware of any other research that has attempted to identify obstructions in the airway and segment beyond the obstruction.

Branch	Manual labelling			Automated detection	
	Obstructed branch	Disconnected region	False Obstructions	Obstructed branch	Disconnected region
LMB	3	3	0	1 ^a	3
RMB	1	1	0	1 (+1 error) ^b	1
BI	12	12	3	12	12
RUL	10	3	0	10	2

^a2 excluded branches are likely to be detected by the SSM but classification could not be performed for these 2 cases because of lack of training data

^b1 branch was misclassified because of a lack of child branches

Table 5.2: Manually and automated labelling of obstructions and disconnected branches (no. of airways)

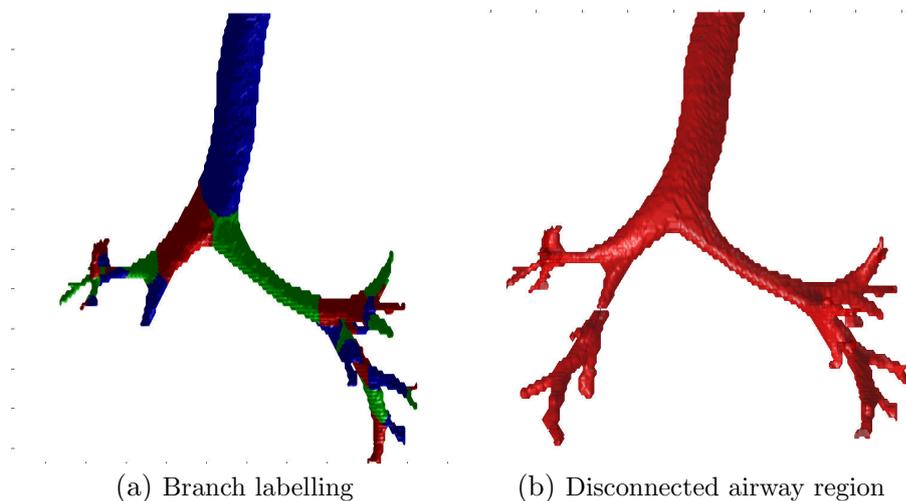


Figure 5.7: Branch labelling and identification of obstructed branch

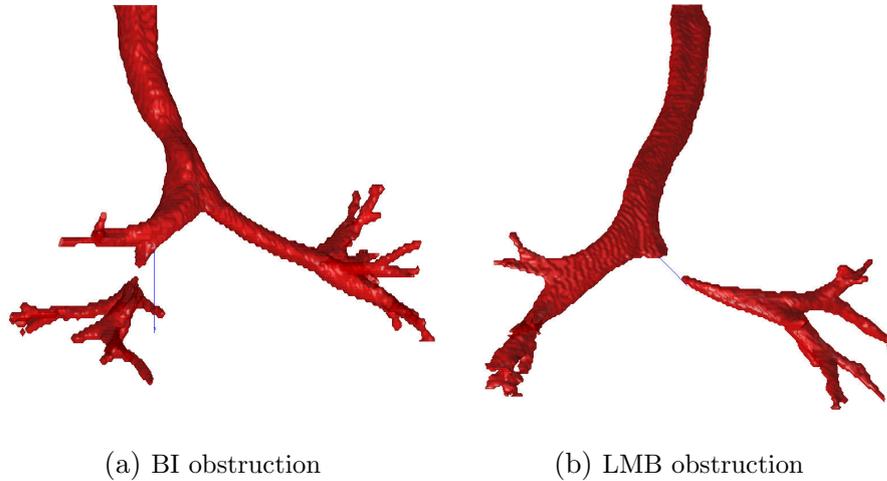


Figure 5.8: Segmenting past complete obstruction

A method has been presented that uses the airway topology and shape to identify disconnected airway regions – and can be applied to existing segmentation methods. This improves upon previous methods that did not take into account the relationship between the main airway tree and the disconnected airway regions, and applies statistical shape analysis for the first time to this problem. In addition, the topological relationship between the main tree and disconnected airway region is also found, which will aid further airway analysis. There is also potential to apply this method to other anatomy with a branching tubular appearance including blood vessels.

This method shows a lot of promise for actively detecting and segmenting beyond obstructed airways. However, both normal and missing branch examples are required to train the algorithm for each anatomical branch. Obstructions are common enough to be of importance for segmentation algorithms but it is challenging to acquire enough training data to evaluate all branches of interest. As shown in this study, 2 cases of LMB stenosis were not enough to train the classifier and this chapter contains an initial evaluation. In future, a dataset containing more examples of obstruction would be useful for an extended training, optimisation of algorithm parameters and evaluation².

²The method described in this chapter uses algorithms developed in the previous chapters. However, this method was not incorporated into the computer assisted detection system described in chapters 2, 3, 4 and 6, and is included in this thesis as an additional method for airway analysis.

5.4 Conclusion

A method was developed in this chapter that uses point distribution models of the bronchi, along with a classifier, to detect obstructions, and segment beyond the obstruction. This method has potential to extend and improve existing airway and vessel segmentation algorithms. However, the method requires cases of obstruction for each branch of interest, for training, and requires further evaluation using a larger dataset of obstructed branches.

5.5 Chapter summary

- Application
 - Segment beyond airway obstructions
- Algorithm
 - Classification of branch shape from PCA modes
 - Identify missing and incomplete branches
 - Find point of obstruction
 - Search beyond obstruction for new airway regions
- Future work
 - Requires many more cases of obstruction to extend method to other regions of the airway

6

Automated detection of paediatric tuberculosis from airway shape

Abstract. This chapter presents a method to identify abnormal airway shape variation, and uses airway shape to detect cases of paediatric pulmonary TB. The method develops statistical shape models of regions of the airway tree, and these models were evaluated in terms of compactness, specificity and generalisation. Features derived from the shape models were used for classification of airway pathology, along with alternative features based on branch cross-sections and orientation. The features were used along with a linear discriminant analysis classifier to detect cases with TB in the test set with leave-one-out cross validation. The best performing features were derived from the shape models of the airway surface and achieved an area-under-the-curve of 0.87(0.77 – 0.94) and 0.81(0.68 – 0.90) for the Trachea-LMB-RMB and RMB-RUL-BI regions, respectively.

CT and bronchoscopy are commonly used diagnostic procedures for paediatric TB patients with suspected airway involvement. Bronchoscopy is considered the “gold standard” for quantifying airway involvement but general anaesthesia is often required and the external cause of the involvement can not be seen. Recent studies

illustrate the possibility of using CT with volume rendering as an alternative to bronchoscopy (du Plessis et al., 2009) – offering the benefits of bronchoscopy while also allowing visualisation of the external cause of the airway involvement. There is potential to assist in the detection of paediatric tuberculosis (and other airway disease) by developing a system to detect abnormal airway changes. This could assist in identification of airway involvement.

Airway segmentation and visualisation in CT is an active field but only a few studies have focussed on the automated detection of airway pathology. The author is not aware of any other research on automated airway assessment in children or cases with TB. In this chapter a method is developed to model normal and pathological variation in local regions of the airway. This method is validated on CT scans of paediatric patients with TB but can potentially be applied to a wide range of disease affecting the airways. Previous chapters performed 3D airway segmentation on chest CT images and structure analysis on the segmented airway. A dataset of 3D airway segmentations from patients are then registered. Once this has been achieved we are in the position to analyse shape variation between airways in a dataset and use this variation along with a classified training set to make predictions about any pathology that might exist in the airways.

The classification methods, discussed in this chapter, derive features from both a corresponding dense mesh representation and a radius representation of each airway. These representations are discussed in Chapter 4 and build on airway segmentation (Chapter 2) and airway structure analysis (Chapter 3). The dense mesh representation is designed to flexibly represent any region from the entire airway to a branch section. The focus of this chapter is on the results of the entire airway analysis method and the accuracy of using this computer assisted detection method to detect paediatric pulmonary TB.

A method to extract features from a point distribution model (PDM) of local regions of the airway surface is presented in Section 6.2.1.1. This method is compared to an alternative approaches using radius based features (Section 6.2.1.3). Evaluation is a key goal of this chapter and the methods of evaluation as well as the dataset used are discussed in Section 6.3. Classification methods are then used to detect airway pathology using the PDM and radius feature vectors (Section 6.2.2). Finally the accuracy of this method for TB detection are presented in the results (Section 6.4).

6.1 Background

This section introduces machine learning and statistical concepts that are useful for airway analysis as well as surveying the literature for related shape modelling and airway analysis methods.

6.1.1 Introduction to feature extraction and classification

Machine learning has developed into a major area of research. The importance of the subject has increased with the availability of large quantities of data. There is an increasing need to model and extract useful information from large and complex datasets. A number of applicable concepts are briefly introduced.

Machine learning methods make use of a feature vector (or measurement vector) to represent each object in the dataset. This is a vector that describes measurements of an object in a way that is useful to distinguish the object from shapes in a different class or category. For example, the feature vector can represent sensor measurements, vertices of a mesh, or measurements of an object's geometry. Vertices of an object mesh are represented as a feature vector by including each dimension separately. Thus, n points $\{(x_1, x_2, x_3), \dots, (x_i, y_i, z_i), \dots, (x_n, y_n, z_n)\}$ representing the surface of an object is represented by the feature vector $\mathbf{f} = (x_1, \dots, x_n, y_1, \dots, y_n, z_1, \dots, z_n)$ of length $3n$.

Supervised learning and unsupervised learning are two important categories in machine learning. Both these classes of algorithms are designed to learn from samples in a dataset but supervised learning requires labelled samples, where the true category of a sample is known, while unsupervised learning can be used with unlabelled data (unknown categories) (van der Heijden et al., 2004); unsupervised methods attempt to find meaning from the feature vector by methods such as clustering and dimensionality reduction.

Acquiring a large dataset of medical images can be difficult. Therefore, the size of the feature vector can often be very large relative to the size of the training set. This is far from ideal and makes overfitting likely. There are a number of unsupervised learning methods that can be used to reduce the dimensionality of the dataset from the covariance of the features. These include principal component analysis (PCA), which

projects the features to an uncorrelated space, ordered by variance, and independent component analysis (ICA), maximising the independence of each component. Fisher discriminant analysis (FDA) is a supervised learning method, that reduces the data to an $(N-1)$ dimension by maximising the class separation, where N is the number of classes. The methods mentioned above are all linear methods, in cases where there is a nonlinear relationship between features, methods, such as principal curves and kernels, generalise and extend PCA (Bishop, 2006; Hastie et al., 2011). PCA has also been modified for a number of special cases including principal geodesic analysis for non-Euclidean space (Fletcher et al., 2004).

A classifier is a function that maps an object represented by a feature vector (\mathbf{f}) onto a class label (k). In our case we are trying to distinguish between “normal” airways and airways with pathology that causes deformation and stenosis. An example of a simple classifier would be the nearest mean classifier, which determines the mean in feature space of each class in the training set. A new shape is then classified according to its distance to each mean. k -nearest neighbour determines the class of the k nearest neighbours and uses this to classify the shape. This method is useful but has high memory and computation costs for large datasets and number of features (Bousquet et al., 2004; Russ, 2006). Linear discriminant analysis and support vector machines are two powerful classification techniques and are introduced and used later in the chapter (Section 6.2.2).

6.1.2 Statistical shape models

The previous section introduced some machine learning concepts; these methods can be used to model and classify complex biological structures. Cootes et al. (1995) introduced active shape models (ASM). These models represented each shape in a dataset by a series of corresponding landmarks. The point distribution model (PDM) is built using the covariance matrix of the landmark points in order to extract the principal components ordered by the eigenvalues of the covariance matrix. In ASM, these modes of variation are used to restrict the type of variation allowed by the model and, therefore, assist in the segmentation of new shapes. Greyscale information can be included to improve the segmentation via active appearance models (AAM). Mitchell et al. (2001) use a combination of ASM and AAM in a fully automated segmentation

method to detect the left and right cardiac ventricles from magnetic resonance images.

PDMs have been built to represent a number of biological shapes including atria in the heart (Lötjönen et al., 2004), the human ear canal (Paulsen et al., 2002), the pelvis (Meller and Kalender, 2004) and the prostate (Hu et al., 2011). de Bruijne et al. (2003a,b) introduce the idea of modelling the cross sections and the medial line, of a tubular object, separately. The medial line is constructed by finding the centroid in each slice and the tubular sections are straightened using the medial line as a reference before creating the PDM. These methods are commonly used to assist in segmentation of new cases or develop personalised shape model for a new patient using the statistical shape model (SSM) and a limited number of landmark points. Personalised 3D models have application during surgery; Rajamani et al. (2007) and Zachow et al. (2005) create shape models of the proximal femur and mandible respectively, which can be used along with manual landmarks to create personalised models.

PDMs can also be used to model normal and abnormal shapes in a dataset. Hutton et al. (2003) and Hammond et al. (2004) model the variation of faces to detect Noonan syndrome because Noonan syndrome can be identified by a number of facial features. They represent each face as a dense surface mesh, build a PDM and represent each face by a feature vector of principal components. Hammond et al. (2004) and Hutton et al. (2003) test number of classifiers for differentiating Noonan syndrome and controls, including: nearest mean, decision trees, neural networks, logistic regression and support vector machines (SVM). They found SVM to be the most effective classifier for this dataset with an average sensitivity of 92% and an average specificity of 93%. However, nearest mean, neural networks and logistic regression all had sensitivity and specificity greater than 83% and 86% respectively. While PDMs show potential to model airway shape, in order to use them, new methods are required to identify suitable landmarks, register the airways, identify relevant regions of interest for analysis, and extract useful features.

6.1.3 Airway Analysis

A number of studies have created methods that estimate automated branch diameter calculations, from segmented adult airways, to provide the clinician with additional

information (Kiraly et al., 2008; Palágyi et al., 2006; Tschirren et al., 2005a), including overlaid heatmaps of airway characteristics including broncho-arterial ratio for identifying COPD (Kiraly et al., 2008). A number of other methods have focussed on chronic obstructive pulmonary disease (COPD) and asthma identification using cross section and radial based measurements (Fetita et al., 2010; Petersen et al., 2010; Wiemker et al., 2004). These methods are useful but rely on identification of pathology from local airway diameter.

Detection of airway pathology due to external changes such as lymphadenopathy and congenital cardiac disease exhibit global changes that are difficult to identify from local airway analysis. PDMs show potential to model more complex airway shape variation. Deligianni et al. (2006) use a PDM of an airway phantom to model breathing. Pinho et al. (2011) built a PDM of healthy tracheas. This model was used to build personalised normal trachea models for patients with stenosed regions in unhealthy tracheas. The difference between the models can be used to assist with stent implants. This method focusses on modelling healthy tracheas but is not aimed at analysis of the bronchi given the global variability of the airway shape and complexity of a branching airway tree (Pinho et al., 2011).

Our computer assisted detection tool introduces a novel method to analyse regions of the airway tree using features from both statistical shape models of a dense surface mesh representations, and branch radius based representations of the airway. This method develops multiple PDM representations for regions of the airways using the corresponding airway meshes developed in Chapter 4. The classification of pathology using the PDM of these regions is performed in a similar way to the method that Hammond et al. (2004) used for facial morphometry, except, as the airway is a branching structure, local models are developed and classified independently for regions of the airway. Developing PDMs for the airways brings a number of unique challenges, including: accurate correspondence and combining of multiple regions in a final classification. Previously, automated airway analysis to assist in the detection of pathology, has focussed on local radius and wall thickness measurements (Fetita et al., 2010; Petersen et al., 2010; Wiemker et al., 2004) but does not consider global shape changes. Statistical shape models have been used to model the trachea in a few cases (Pinho et al., 2011) but these methods are not designed to model changes due to airway pathology or extend to other airway regions. While other studies have only

modelled breathing changes of the a single patient (Deligianni et al., 2006) and do not consider inter-patient variability, in this chapter a powerful method is introduced to build statistical shape models of local airway regions and use these models to identify airway pathology across a dataset.

6.2 Method

To classify airway pathology, features were generated from point distribution models (PDM) of the airway surface, as well as, from cross-sectional measurements (Section 6.2.1). These methods build on branch structure analysis and correspondence from Chapter 3 and Chapter 4. Classification methods – that take the feature vectors as input are used to identify new cases with airway involvement from TB (Section 6.2.2).

6.2.1 Feature Vectors

Feature vectors are derived from a PDM of the airway dataset and from cross sectional sampling of each branch in the airway tree.

6.2.1.1 Method 1: Airway point distribution model

Each airway is represented by a dense surface mesh. The corresponding vertices of each airway surface mesh can be used to build a PDM of the airway and extract a set of principal modes that can be used to represent the variation of each airway.

Some of the variation between airways in the dataset is due to size, position and rotation, which are related to patient age and scan position, and are not of interest for detection airway pathology. Therefore, the airways are first aligned by generalised procrustes analysis (GPA). GPA is used to iteratively align objects represented by a set of corresponding points using translating, scaling and rotation. The implementation of GPA in this study is outlined in Appendix A.1.

Principal component analysis (PCA) was applied to the aligned airways to extract the principal modes of variation of each airway region, which were used as features for classification. PCA is discussed in Appendix A.2. Therefore, an airway can be represented in terms of the mean shape $\bar{\mathbf{x}}$ and a displacement along each principal component Φ by \mathbf{b} :

$$\mathbf{x} \approx \bar{\mathbf{x}} + \Phi \mathbf{b} \tag{6.1}$$

The airway can be approximated by a set of m eigenvectors $\Phi = (\phi_1|\phi_2|\dots|\phi_m)$ and $\mathbf{b} = (b_1, \dots, b_m)^T$. This means that a dense surface mesh model of a shape can be reduced to a much shorter feature vector represented in terms of the principal components of shape variation.

PCA was applied to a dataset containing both TB and non-TB cases, which models airway variation from normal to pathological, and these features are used to perform a binary classification of pathology for an airway region. An alternative set of features could be constructed from separate PCA representations for both TB and non-TB cases. A measure of similarity such as Mahalanobis distance could then be used to determine the nearest class. This alternative method is worth exploring in future but may require careful selection of the similarity measure.

6.2.1.2 PDMs for local region analysis

Variation in airway shape could potentially be modelled at a number of levels from the variation in a single cross-section of a branch to global variation of the entire airway. In this study, local regions of the airway are modelled individually where each region consists of a parent branch, the bifurcation region and two child branches. Modelling each branch individually is not as effective because variation of a branch relative to neighbouring branches is important; lymphadenopathy can lead to branch deformation as well as stenosis (an example of tracheal displacement by lymphadenopathy is presented in Andronikou and Wieselthaler (2004)). Individual branch shape analysis also means that a larger portion of the model's variation is influenced by mesh artefacts at the beginning and end of the branch.

Alternatively, shape models can be applied to the entire airway for analysis. However, PCA, is a linear projection to a lower dimensional space and with each additional generation of a tree structure the variation of the airways becomes increasingly non-linear. While there are benefits to modelling the relationship between parent and child branches, there is limited value to modelling the relationship between the shape changes of second order connected branches. There are nonlinear formulations of PCA (including Kernel PCA in Schölkopf et al. (1997)), but because it is useful to

model the airway locally, modelling the airway as a set of 3-branch structures using PCA is proposed. An example 3-branch region is the RMB-RUL-BI region shown in Figure 6.1.

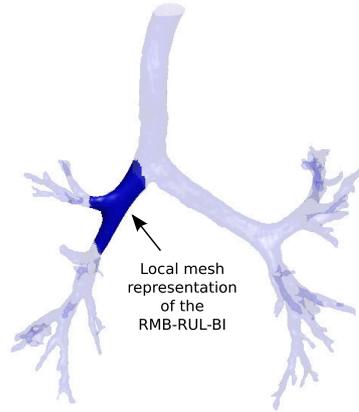


Figure 6.1: Shape models are generated from 3-branch regions (RMB-RUL-BI region in this example)

Therefore, this method creates a number PDMs of the airway for variation 3-branch regions. Each 3-branch region overlaps the previous region. For example, the RMB is represented in both the trachea-RMB-LMB and RMB-RUL-BI regions. These regions are automatically identified (Chapter 3 and Chapter 4), a point distribution model is built from the dataset and each 3-branch region is assigned a feature vector **(b)** that can be used for classification.

6.2.1.3 Method 2: Radius based features

The previous section describes a method to extract feature vectors from PDMs of local airway regions in order to detect pathology. For comparison, feature vectors based on cross-sectional measurements were also derived. A representation of the airway in terms of cross sectional measurements is described in Section 4.2.1 in Chapter 4. To recap, vectors orthogonal to the branch centreline were projected to the surface and the intersection between the mesh and the vector was found. Each branch can then be represented as a series of n cross sectional radii measurements where each cross section is sampled 4 times at 90° angles i.e. $B = \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n\}$ where $\mathbf{r}_i = (r_{i1}, r_{i2}, r_{i3}, r_{i4})$

at $[0, \frac{1}{2}\pi, \pi, \frac{3}{2}\pi]$. Figure 6.2 shows \mathbf{r}_i sampled along each branch (for visualisation purposes, fewer \mathbf{r}_i are shown than used in the analysis).

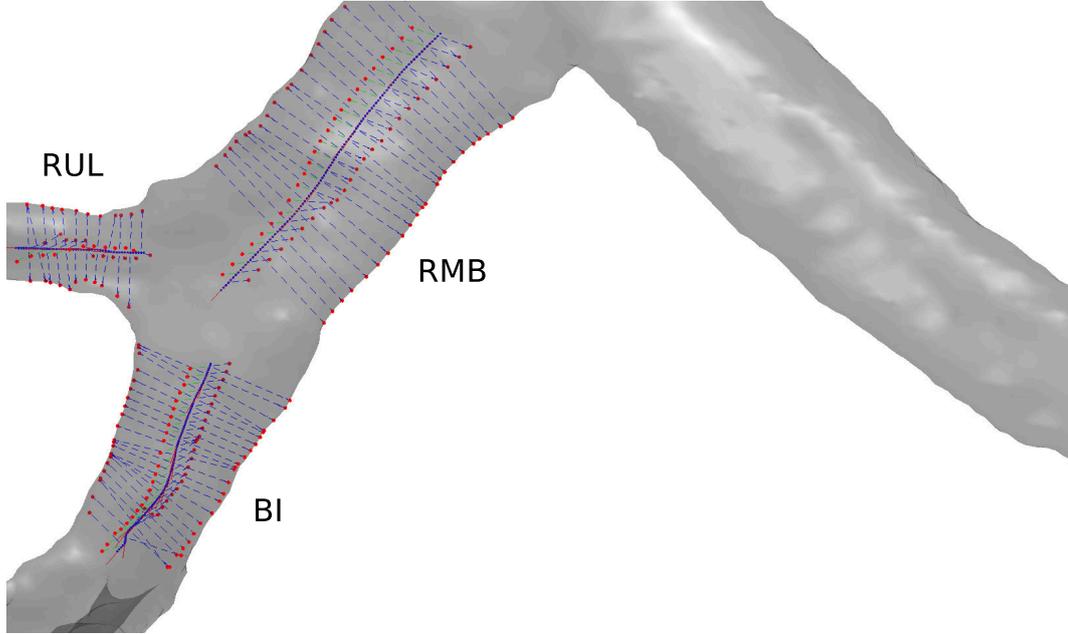


Figure 6.2: Branch radius measures using ray-surface intersection. 4 intersections are shown for each equidistant centreline point.

Diameter based features

The first set of features used are the mean branch diameters from 3 equal sections of each branch. The mean diameter of a branch section can be calculated from the representation as:

$$\bar{d} = \frac{1}{2(q-p+1)} \sum_{i=p}^q \sum_{j=1}^4 \mathbf{r}_{ij} \quad (6.2)$$

where p and q are the start and end points of the region. These features are normalised across airways by the branch length (l). As with the PDM derived features, the same 3-branch regions were used for this method (each region is represented by 9 features).

Features based on observed stenosis

An alternative set of radius derived features is also presented in this study. Three features were calculated for each branch:

- 1) the maximum ratio of the orthogonal diameters for each branch

$$f_1 = \max \left(\max \left(\frac{d_{i1}}{d_{i2}} \right), \max \left(\frac{d_{i2}}{d_{i1}} \right) \right) \quad (6.3)$$

where $d_{i1} = r_{i1} + r_{i3}$ and $d_{i2} = r_{i2} + r_{i4}$

- 2) the ratio of the average branch diameter and branch length

$$f_2 = \left(\frac{\bar{d}}{l} \right) \quad (6.4)$$

3) the local minima and maxima of the diameters d_{i1} and d_{i2} , where $i = (1, n)$, were identified from B' and B'' . The maximum ratio of local minima and neighbouring local maxima was used as a feature:

$$f_3 = \max \left(\frac{l_{max_a} + l_{max_b}}{2l_{min}} \right) \quad (6.5)$$

These features were based on clinical observations and were indicators of branch circularity, thickness and local stenosis, and were calculated for the trachea, RMB and LMB. The angle at bifurcation was also calculated for each airway. Therefore, feature vectors could be generated from each branch from the radii. For consistency the same three-branch regions were used as Section 6.2.1.2.

6.2.2 Classification

The training and test datasets were used to evaluate the algorithms ability to distinguish between TB and non-TB patients. Feature vectors, using the methods described in the previous sections, were used as input for a classifier and the classifier was evaluated on unseen cases. Linear Discriminant Analysis and Support Vector Machines were used for classification. These methods are outlined in Appendix A.3.

6.2.2.1 Classification of obstructed branches

A small proportion of cases have an obstructed branch in the three-branch region. An obstructed branch in the main bronchi is caused by complete stenosis of that branch. Obstructions considerably change the shape of the airway and, therefore, act as outliers that reduce the accuracy of the model. Instead, the branching structure (discussed in Chapter 3) was used to identify first or second generation bronchi without child branches. These obstructed regions were automatically identified and excluded from the classification. Assuming that completely obstructed branches imply pathology, these cases can be reincluded in the final classification result.

6.2.2.2 Combining classification results from multiple regions

Each region consisting of 3 branches and the connecting bifurcation region was evaluated independently, but combining multiple regions could be beneficial for a final disease classification. The features cannot be easily combined in the classification because a number of cases are automatically excluded due to obstructed branches in the region. Therefore, a number of cases will only have features for some of the regions being analysed.

However, the probability of disease from the classification of each region can be combined with a weight:

$$P(TB|X) = wP_T + (1 - w)P_R \quad (6.6)$$

where, for example, P_T and P_R are the posterior probabilities (generated from an LDA classifier) from the trachea-RMB-LMB and RMB-RUL-BI regions and a weight $w = 0.5$. This illustrates one method of combining multiple classifications. For cases where either P_T or P_R are missing due to an obstructed branch, only one posterior probability is used for final classification.

To improve the sensitivity at very high specificities (a likely operating point), a very high threshold could be applied to one region and used in combination with the classification of the second region. This would mean that only definite cases from the first region are added to the final classification:

$$P(TB|X) = P_T \vee (P_R > w) \quad (6.7)$$

where $w = 0.9$.

Therefore, the classifications are performed on three-branch regions of the airway and classifications from multiple regions can be combined. Obstructed branches can also be automatically identified and included to improve the generalisability of the classifier.

6.3 Evaluation Methods and Data

The airway analysis algorithm was evaluated on a paediatric chest CT dataset, and a set of metrics were used to evaluate performance.

6.3.1 Datasets

The paediatric TB datasets that were used in this research are discussed in Section 2.4 of Chapter 2. In order to clarify the results, a brief summary is presented here. The dataset used to train the model is made up of chest CTs of children with and without pulmonary TB. A set of paediatric non-TB cases was acquired from Great Ormond Street Hospital, London, UK (referred to in this study as *GOSH2010*) and TB cases from Tygerberg Hospital, Western Cape, South Africa (*TYG2010*). These cases were used to develop the model and choose parameters. In 2012, two dataset were collected Tygerberg Hospital, with and without TB (referred to as *TYGTB2012* and *TYGnonTB2012*). These were previously unseen sets used for evaluation. Several cases with artefacts from tubes or severe movement artefacts were excluded from the study. Table 6.1 shows the size and mean age of the patients of each dataset. Voxel size in the axial plane ranged from 0.3 - 0.5 mm and slice thickness 0.7 - 1 mm.

Non-TB patients in the dataset include children with the following conditions: parenchymal lung disease such as interstitial lung disease, congenital lung malformations which affect the parenchymal tissue and not the airways; suspected lung metastases; and infective conditions other than TB. This will include cases of bronchiectasis,

Table 6.1: Mean age and std deviation of each dataset

Dataset	Size	Age \pm Std (years)
<i>Training (GOSH2010)</i>	37	3.2 ± 1.8
<i>Training (TYGTB2010)</i>	52	1.8 ± 2.2
<i>Test (TYGTB2012)</i>	42	3.1 ± 3.8
<i>Test (TYGnonTB2012)</i>	48	2.4 ± 2.8

cystic fibrosis and pleural disease. Patients with suspected TB undergo a CT scan when there are signs or symptoms of airway involvement (du Plessis et al., 2009) and therefore, all TB cases in this dataset have suspected airway involvement.

Due to the challenges of detecting paediatric TB, a number of cases may have suspected TB but will not be definitively classified. If there is large uncertainty in the clinical labels, it is not possible to train or evaluate the accuracy of the method. Therefore, in this study, non-TB cases were compared to probable and definite TB cases. In future, an alternative evaluation could be performed, using cases originally suspected of TB, only some of which would later be confirmed as TB cases. This would either require the system to be incorporated into the clinical workflow, or knowledge of the clinicians initial suspicions (not available in this study).

In summary, a set of non-TB patients is compared to a set of patients with pulmonary TB (and suspected airway involvement). This computer assisted detection tool is aimed at detecting the approximately 40% of paediatric TB patients with airway involvement, to provide an additional test that can improve the overall detection of TB by detecting the subset with airway involvement.

In paediatric pulmonary TB, the most common locations of lymphadenopathy are subcarinal (90%), hila (85%), anterior mediastinum (79%) and paratracheal (63%) (Andronikou et al., 2004). Bronchial compression caused by lymphadenopathy is most apparent in the LMB, RMB, BI and trachea (Andronikou et al., 2004). Therefore for identifying TB, two regions were modelled (trachea-RMB-LMB and RMB-RUL-BI). This could be extended to the rest of the airway if useful for modelling additional airway pathology. The model is aligned using the branching structure and extracted mesh, and because three branches are modelled at any one time, no more complexity is added by analysing further generations.

6.3.2 Shape Model Evaluation

Before analysis of the accuracy of classification, the characteristics and performance of the point distribution model were analysed. Measures of compactness, generalisation and specificity, that are outlined by Styner et al. (2003), were used to evaluate the model.

6.3.2.1 Model Variance

A compact model can represent the variance within the dataset with only a few parameters. The variance can be calculated as a function of the number of shape parameters (modes of variation), as shown in Equation A.7 of Appendix A.

6.3.2.2 Model Generalisation

The model generalisation represents a model's ability to represent unseen cases (Styner et al., 2003). This was evaluated using leave-one-out reconstruction of the dataset. For each case in the dataset, the PDM was built without the case, and used to reconstruct the case using M modes of variation. The mean absolute distance (MAD)¹ of the vertices was then used to approximate the error of the PDM representation of each case. The mean and standard deviation of the MAD values for the entire dataset was calculated. Confidence intervals were found by the Central Limit Theorem.

6.3.2.3 Model Specificity

Model specificity evaluates the ability of the model to only generate shapes that are similar to those found in the dataset (Styner et al., 2003). 1000 new instances of the feature vector \mathbf{b}_i were generated for each PDM. Monte Carlo simulations were used to randomly generate \mathbf{b}_i from a multivariate normal distribution with standard deviation $\sigma_i = \sqrt{\lambda_i}$ and mean $\mu_i = 0$ (Hu et al., 2010). MAD from each instance to each case in the dataset was calculated, and the minimum distance was recorded, which represents to closeness of the instance to the nearest case in the dataset. The mean and standard deviation of the minimum MAD distances were calculated.

¹The mean absolute distance is defined as the mean distance between the vertices of the model representation and the vertices of the corresponding case (Styner et al., 2003)

6.3.3 Evaluation of Classification

Before the results of the airway classification algorithm are presented, a few methods, that will be used to quantitatively evaluate the accuracy of the computer assisted detection system, need to be introduced.

Most classifiers are trained using a labelled training set and these labels are commonly based on other clinical tests or expert annotations of the images. Each classifier finds a decision boundary, a boundary with as many dimensions as features, that defines the regions in feature space where a new sample would be labelled as a specific class.

Classification algorithms have varying degrees of complexity in the optimisation procedure and the decision boundary. If only a single set is used for training and validation then it is likely that the classification results will be “overfitted” – i.e. the decision boundary too closely models variation within an individual dataset and results will not be representative of the accuracy of classifying new data (van der Heijden et al., 2004). This can be overcome by using a training and test set, where the test set is previously unseen by the classifier.

For small datasets (size m), it is often not feasible to divide the dataset into a training and test set. Cross-validation (CV) divides the dataset (\mathbf{D}) into n equal subsets (d_i) of size m/n . The training and classification algorithm is run n times where the algorithm is evaluated on samples d_i and trained on samples $\mathbf{D} - d_i$. Therefore, all data is used for training and validation but each validation set is “unseen” by each classifier. The extreme of this method is leave-one-out cross validation (LOOCV) where $n = m$; training on all cases but one and testing on a single case in each step. This method allows the largest number of cases to be used for each training but is the slowest form of CV.

Some classifiers such as Kernel SVM allow parameters to be set within the algorithm. If these parameters are chosen using the whole dataset then this can also be a cause of overfitting. Therefore, if a cross-validation method is used to evaluate the classifier, a nested cross-validation should be applied. For each step in the CV, a nested loop is used to choose optimal parameters. This method shows very similar performance to an independent test set (Varma and Simon, 2006). This has a considerable computational performance cost but is only required for validation and not

classification of new cases.

In order to evaluate the performance of a classifier a good performance measure needs to be chosen. Sensitivity and specificity are useful measures. Sensitivity (Se) is the proportion of correctly identified positive cases ($Se = \frac{TP}{TP+FN}$), and specificity (Sp) is the proportion of correctly identified negative cases ($Sp = \frac{TN}{TN+FP}$), with the true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) (Altman et al., 2005). Therefore, a classifier has a perfect accuracy if $Se = 1$ and $Sp = 1$. These values can be calculated by comparing the true class label of each sample and the label identified by the classifier. The problem with a single measurement of sensitivity and specificity is that it only gives a snapshot of the methods performance. This is insufficient if the likely operating point – the point used in clinical practice – requires a different specificity or if different methods are being compared, which were evaluated at different specificities.

A receiver operating characteristic (ROC) curve is a plot of sensitivity versus (1-specificity) over the entire range of (0,1). A perfect classifier would have a curve along the left and top of the plot; a classifier that is no better than chance of detecting differences between groups would have a diagonal line running from [0,0] to [1,1]. This is a very useful visualisation of the performance of a classifier at various operating points but for comparing multiple systems it is also useful to have a single performance measure such as Area-Under-the-Curve (AUC) of the ROC (Altman et al., 2005). A perfect classifier will have an AUC=1 and a classifier that is no better than chance has an AUC=0.5. Many classification algorithms produce a probability associated with the likelihood of disease for each sample (the posterior probabilities in the case of Bayesian based classifiers). These probabilities can be used to generate the ROC curve. In other cases multiple classifications can be made at various specificities by modifying the cost function, prior probabilities or repeating samples from certain classes in the training set (van der Heijden et al., 2004).

With any performance measure, it is also useful to state the belief that we have in that measurement. Confidence intervals (CI) are a measure of the range of sample means of sets selected from a population. This is a useful measure of the range in measurements of sensitivity, specificity and AUC that we might expect on new datasets. A 95% confidence interval is a common choice and is the interval where the mean, of 95% percent of sets of samples from a population, falls. This is commonly

calculated by using the Central Limit Theorem to assume a normal distribution of the sampling mean and then approximating the standard error (S_e) from the sample standard deviation (Altman et al., 2005). The S_e along with the z-score can be used to find the CI. However, this can be problematic for small datasets and, therefore, a more direct approach is bootstrapping (Altman et al., 2005; Hastie et al., 2011). Bootstrapping generates a number of new sample sets from the original dataset by resampling with replacement. Therefore, each sample in the population can be picked more than once during the generation of a new set of samples. Statistics about the population including CI can be calculated directly from the generated samples. Altman et al. (2005) recommend 50-100 replications as being adequate for the CI of a mean and bias corrected and accelerated method (BCa) as the preferred approach. It is important to note that confidence intervals are not directly equivalent to statistical significance when comparing two ROC curves with CI. The confidence interval must be calculated from the difference of means to be appropriate (Altman et al., 2005). However, if confidence intervals from two means do not overlap then they are significant but if they do overlap then they may or may not be significant and a correct calculation of significance is required. These methods were used to evaluate the performance of the computer assisted detection algorithm at detecting signs of tuberculosis in paediatric airways.

6.4 Results

The previous sections of this chapter have proposed methods to create feature vectors for corresponding dense mesh representations of local regions of the airway. A number of algorithms are also introduced that can classify TB, and other airway disease that influence airway shape, based on the feature vectors and a labelled training set. The entire airway analysis algorithm was trained and evaluated on CT datasets obtained from hospitals in South Africa and the U.K. Results of the airway classification to detect tuberculosis are presented in this section based on training and test sets of paediatric patients with and without tuberculosis.

6.4.1 Shape model derived features

6.4.1.1 Point distribution model representation

Results for local airway regions using PDM derived features are presented first. The variation of the first 10 PCA modes for the trachea-RMB-LMB and RMB-RUL-BI regions are shown in Figure 6.3 - Figure 6.6. The centre column shows the mean airway ($\bar{\mathbf{x}}$) with variation ($\mathbf{x} = \bar{\mathbf{x}} \pm \Phi_i b_i$). Using $b_i = \pm 3\sqrt{\lambda_i}$ shows the variation of one mode three standard deviations from the mean in the dataset (Cootes et al., 1995). Dark regions in the figures show the local PDM model and the light region is an example airway that is registered to the model to aid visualisation. These principal components exhibit a range of changes including branch narrowing, local stenosis, deformation, length and angular changes. These modes also include variation that is considered pathological such as deformation and stenosis of branches. For example, the fifth mode of Figure 6.3 shows central narrowing of the LMB while the seventh mode of Figure 6.6 shows narrowing of the BI. Note that due to the perspective some of the 3D variation is not visible in the figures.

Figure 6.7 shows two example airway segmentations from a TB and non-TB patient each with a feature vector of principal component weights (\mathbf{b}):

$$\mathbf{b}(\text{nonTB}) = \begin{bmatrix} -0.0285 \\ 0.0781 \\ -0.0406 \\ -0.0316 \\ -0.0064 \\ -0.0161 \\ -0.0177 \\ -0.0091 \\ -0.0405 \\ -0.0198 \\ -0.0039 \end{bmatrix} \quad \mathbf{b}(\text{TB}) = \begin{bmatrix} -0.0330 \\ 0.0236 \\ 0.0578 \\ 0.0198 \\ -0.0069 \\ 0.0062 \\ -0.0026 \\ 0.0008 \\ -0.0306 \\ 0.0094 \\ -0.0005 \end{bmatrix}$$

Particularly evident in this example is narrowing of the LMB for the TB case. This can be observed in the third and fourth mode of variation (Figure 6.3) and is consistent with the third and fourth column of the feature vectors: $\mathbf{b}(\text{nonTB})$ and $\mathbf{b}(\text{TB})$. However, many more modes are required to distinguish the cases in the dataset. Small differences are also visible between the 11 component representation of the trachea-LMB-RMB region (dark blue) and the original airways (light blue).

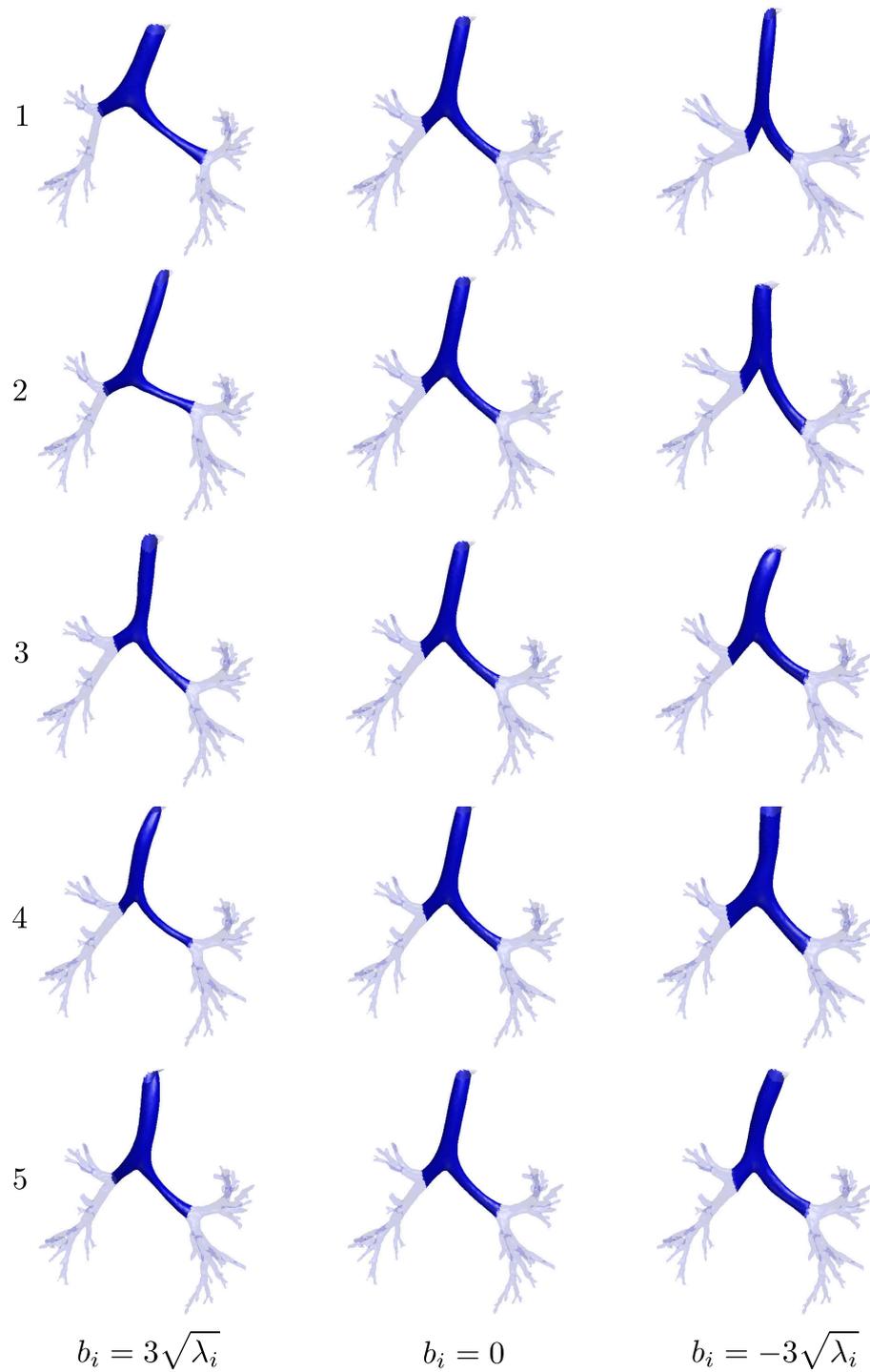


Figure 6.3: Modes of variation 1-5 for the trachea-LMB-RMB ($b_i = 0$ is the mean model and $b_i = \pm 3\sqrt{\lambda_i}$ represents the shape 3 standard deviations along each mode)

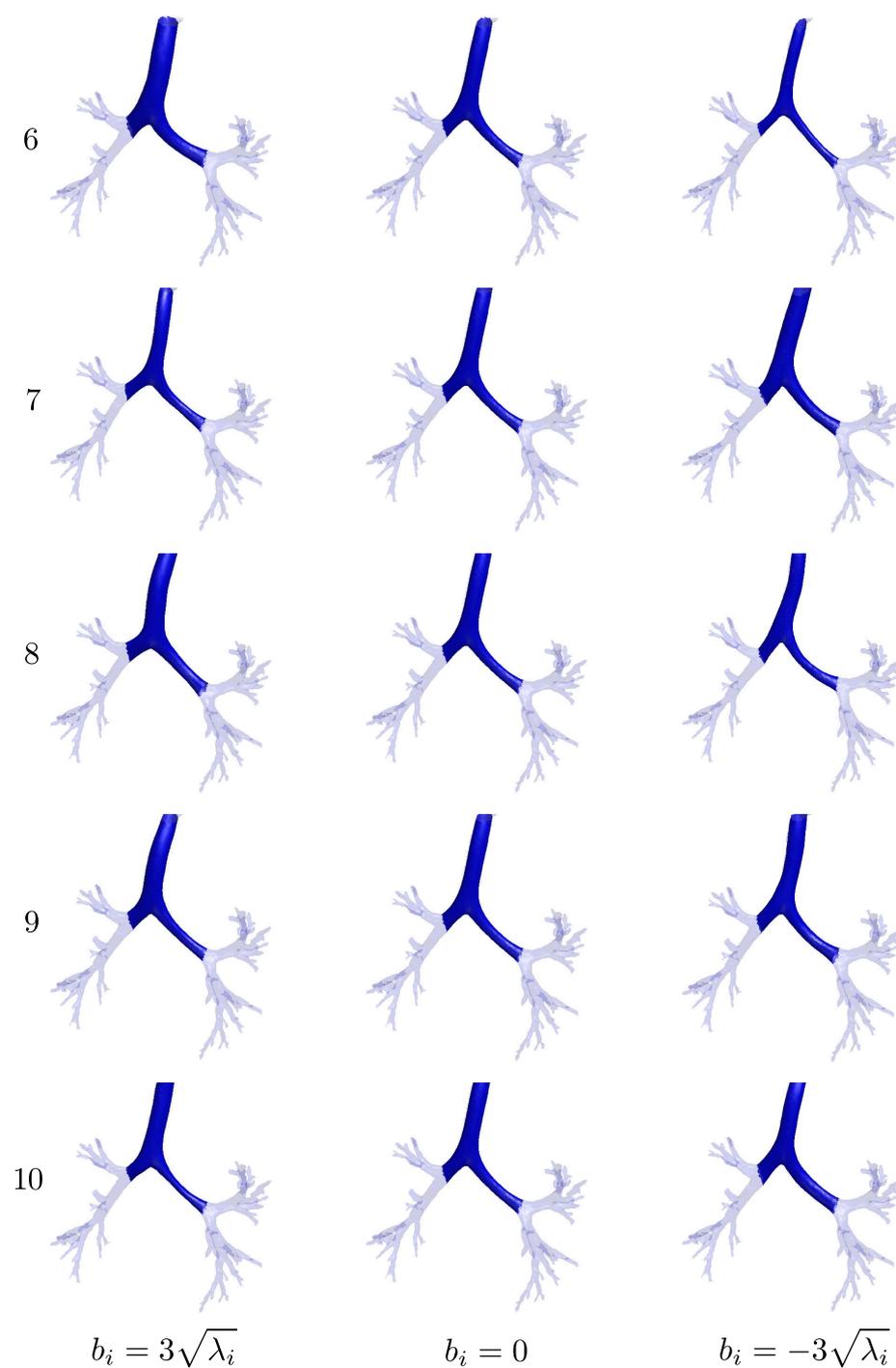


Figure 6.4: Modes of variation 6-10 for the trachea-LMB-RMB ($b_i = 0$ is the mean model and $b_i = \pm 3\sqrt{\lambda_i}$ represents the shape 3 standard deviations along each mode)

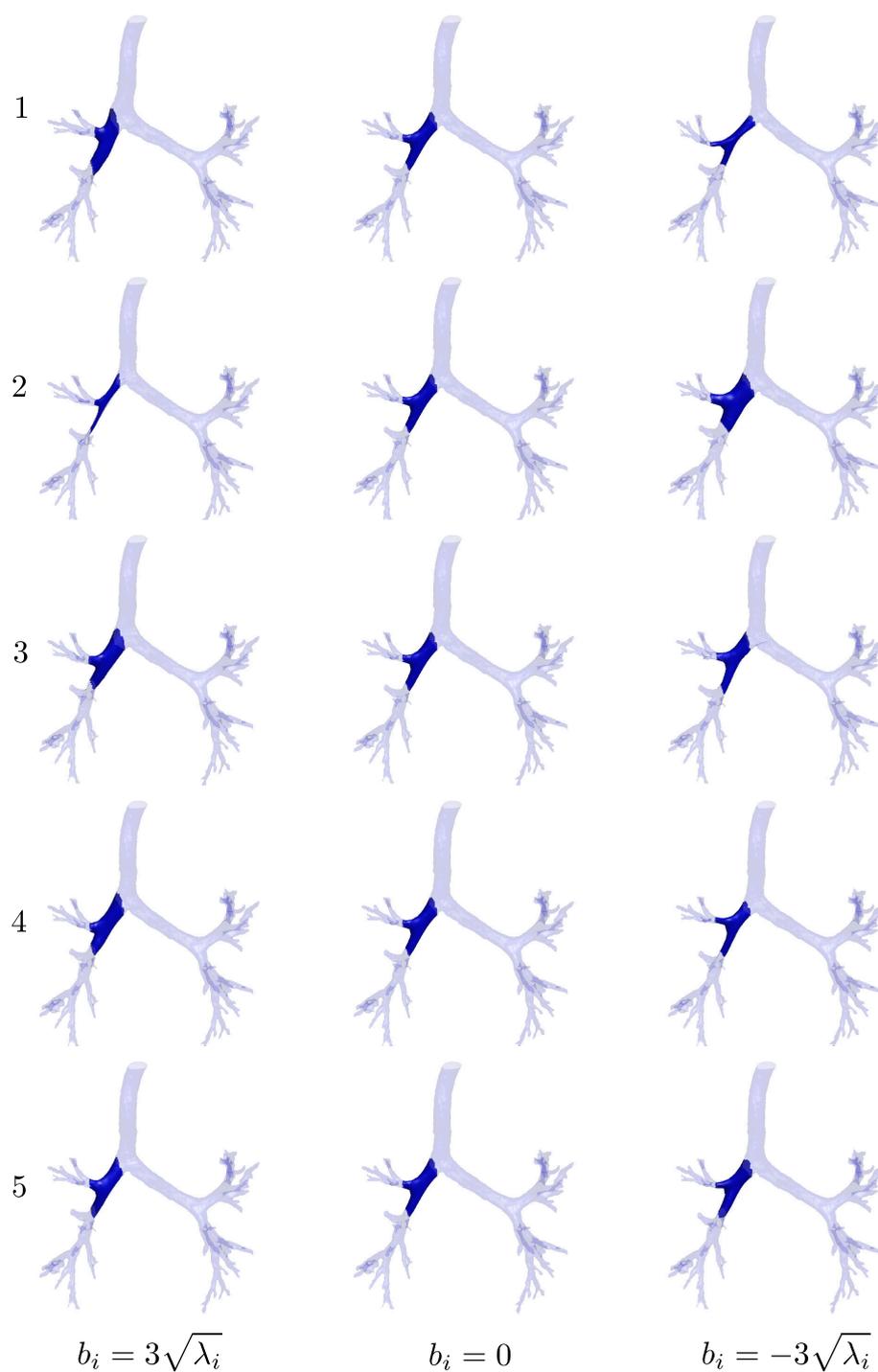


Figure 6.5: Modes of variation 1-5 of the RMB-RUL-BI ($b_i = 0$ is the mean model and $b_i = \pm 3\sqrt{\lambda_i}$ represents the shape 3 standard deviations along each mode)

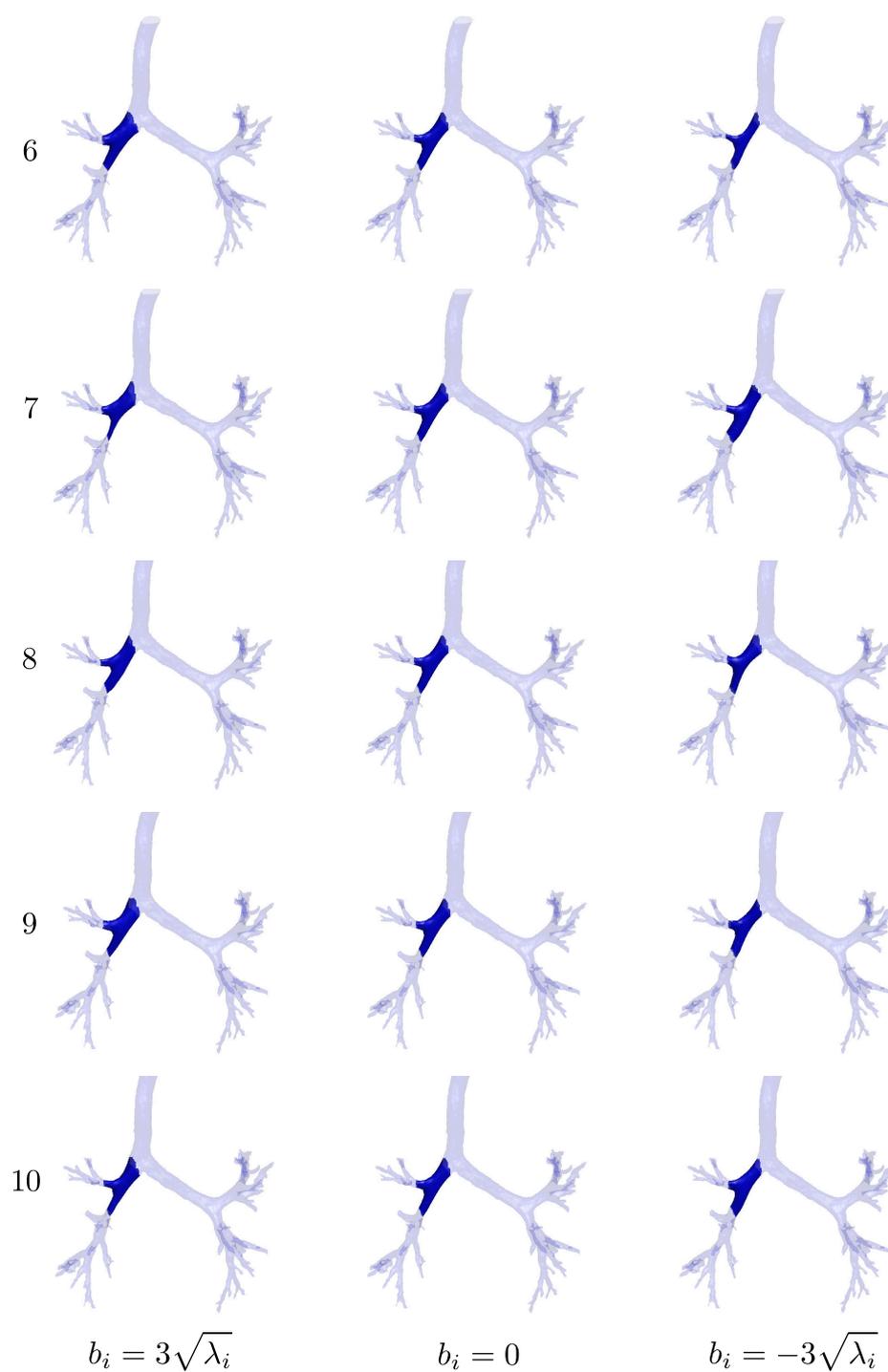


Figure 6.6: Modes of variation 6-10 of the RMB-RUL-BI ($b_i = 0$ is the mean model and $b_i = \pm 3\sqrt{\lambda_i}$ represents the shape 3 standard deviations along each mode)

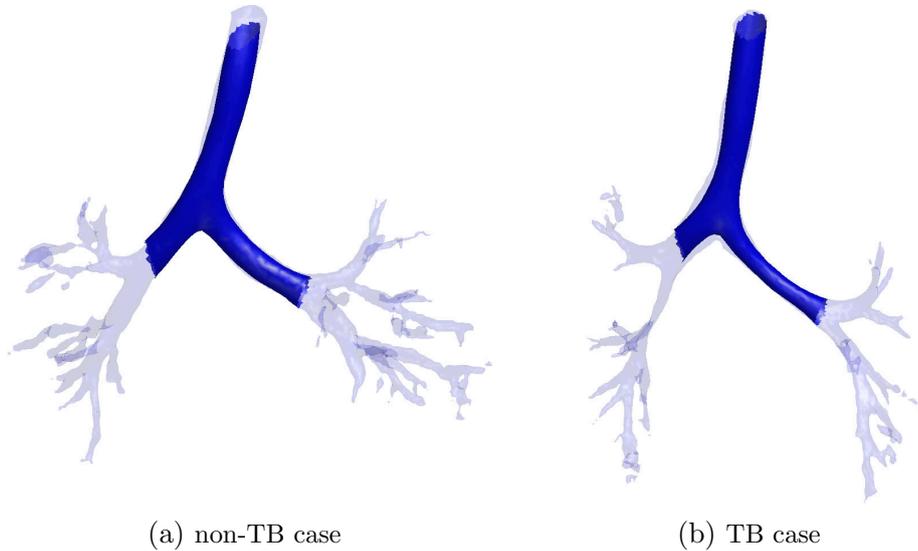


Figure 6.7: Example airways from the non-TB and TB datasets with overlaid 11 component representation of the first 3 branches. Small deviations of the model’s shape from the airway shape is explained by higher order modes that have been removed from the representation.

These differences are represented by the remainder of the principal components that are dropped from the model.

6.4.1.2 Model evaluation

PCA produces modes of variation that are ordered by the contribution to the total variance within the dataset. Figure 6.8 shows the contribution of each mode to the total variance. In both cases approximately 90% of the airway variation in the dataset is represented by the first 11 modes.

The model generalisation and model specificity for both PDMs are shown in Figure 6.9 as a function of the number of modes of variation (M). As expected the model generalisation – the ability of the model to represent unseen cases – performs better (decreases) with increasing number of modes. Also as expected, the specificity (the ability of the model to only represent similar cases) performs worse (increases) with increasing number of modes. The RMB-RUL-BI model outperforms the Trachea-LMB-BI model both in terms of generalisation and specificity. This is probably due

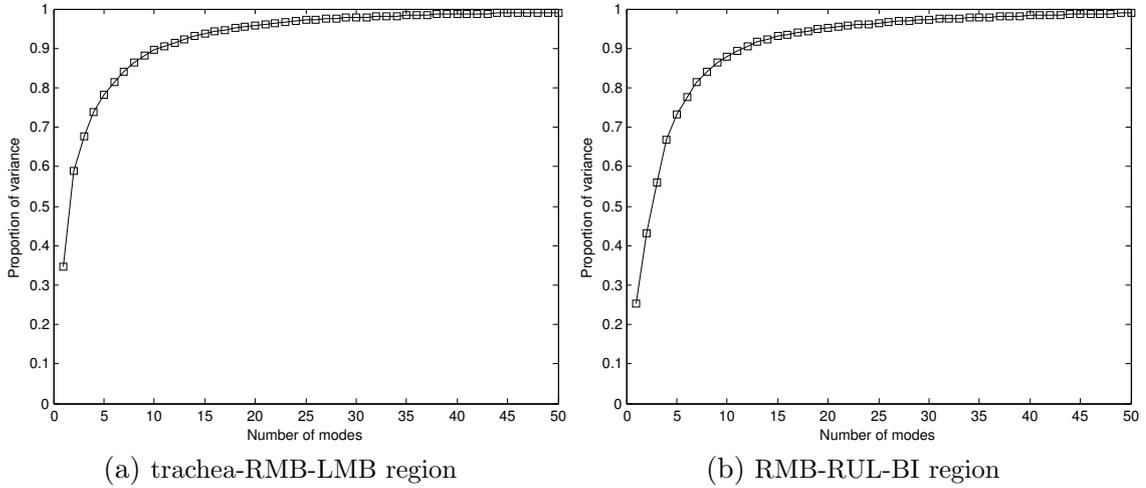


Figure 6.8: The proportion of the total variance in the dataset that is represented by n principal modes. Principal modes are ordered by variance and, therefore, higher modes will make smaller contributions. In this dataset, the first 11 modes represent $\approx 90\%$ of the variation.

to less variance in the region.

6.4.1.3 Detection of paediatric pulmonary tuberculosis

Fisher mapping is an affine mapping that maximises the ratio of the inter/intra class variability (van der Heijden et al., 2004). Plots of Fisher mapping provide a useful visualisation of the linear separability between classes for a set of features. Figure 6.10 and 6.11 show the separability of the datasets represented by 11 principal modes of variation (90% of the dataset variation) using Fisher mapping. The figures show the potential of the principal components to separate the classes of TB and non-TB patients. Good separability can be achieved for both the training and test sets. However, the training set (Figure 6.10b and 6.11b) show better separability than the test set (Figure 6.10c and 6.11c) for both regions. The feature vector of each case is shown in Appendix D and the mesh representation of each airway in the test set is shown in Figures 4.16 and 4.17 of Chapter 4.

Figure 6.12 shows ROC curves, for the two regions of interest, distinguishing TB from non-TB cases, using linear discriminant analysis (LDA) with leave-one-out cross validation. 95% confidence intervals for the sensitivity and AUC of the ROCs were

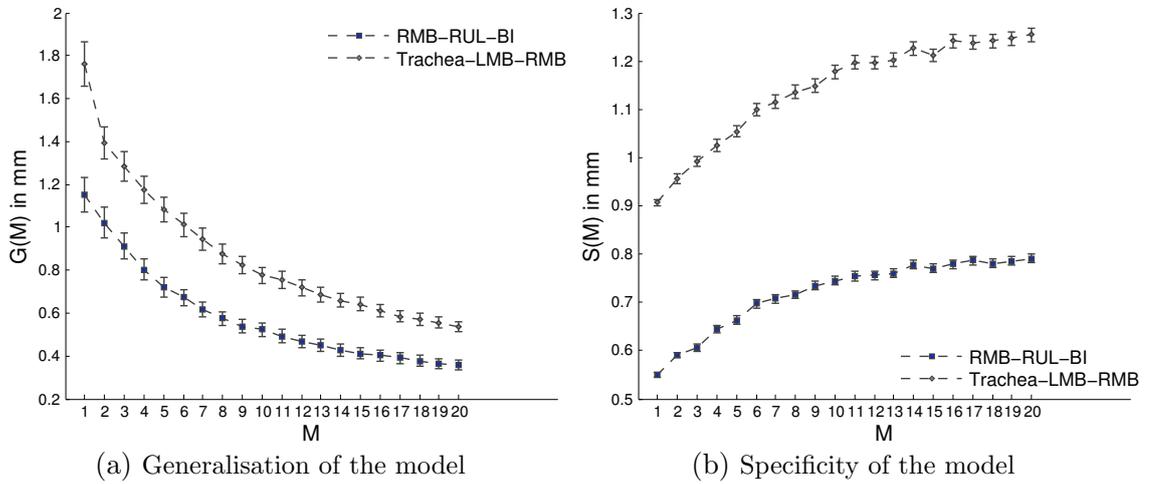


Figure 6.9: Generalisation $G(M)$ and specificity $S(M)$ of the model with confidence intervals, as a function of the number of modes (M), for the Trachea-LMB-RMB and RMB-RUL-BI regions.

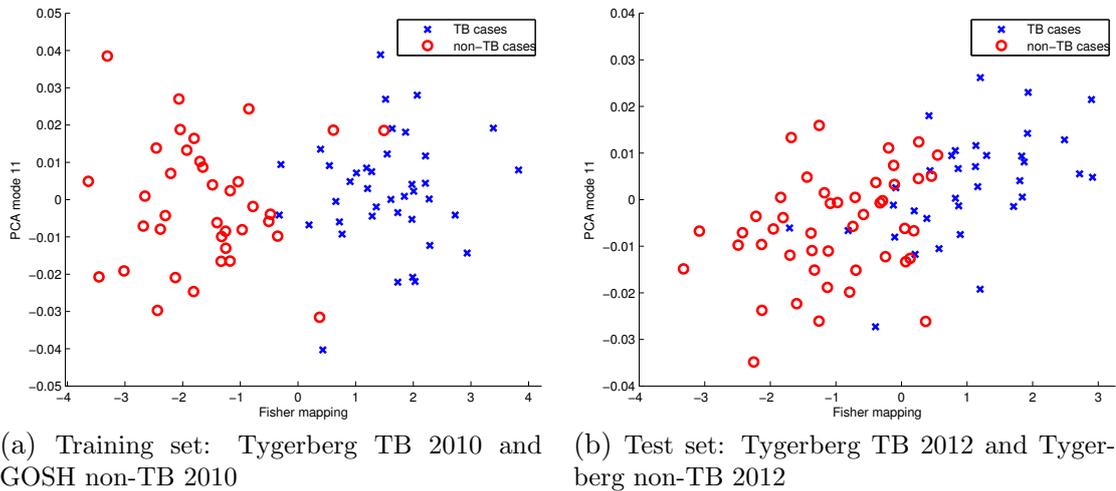


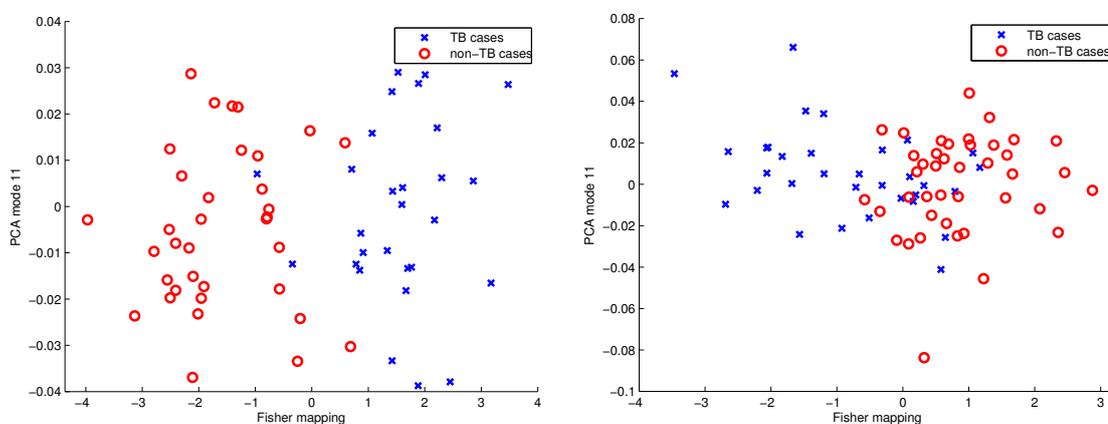
Figure 6.10: Fisher mapping of first 10 modes plotted against 11th mode for the trachea-RMB-LMB region

estimated using bootstrap BCa with 1000 replications. Figure 6.12a shows evaluations on the training set (*Tygerberg-GOSH 2010*) and test set (*Tygerberg 2012*) using the first three branches as the region of interest. Classification of TB patients in the training set (AUC: 0.96(0.86 – 0.99)) performs better than classifying TB patients in the test set (AUC: 0.87(0.77 – 0.94)). While the performance of the classifier is better for the training set, the difference between the AUCs is not significant ($p=0.05$). Figure 6.12b shows that second order bronchi can also play a role in TB detection with an AUC of 0.81(0.68 – 0.90) for the test set. While ROC curves of the validation set are the most useful for evaluating the algorithm, the other curves provide useful insights into the algorithm. The training set is made of of non-TB cases from GOSH and TB cases from Tygerberg. The improved accuracy of the training set could be explained by algorithm parameters being based on this set but it is also likely that the non-TB set from GOSH might have less severe pathology than the non-TB patients from Tygerberg (in the test set).

These ROC curves were generated using LOOCV and show that a high AUC can be achieved for classifying TB on the test set. The next set of results compares the relative performance of combining multiple regions and obstructed branches. As we are interested in relative performance for these next evaluations, the training and test sets were merged into a *Complete dataset* to work with a larger dataset and smaller confidence intervals.

The classification from multiple regions can be combined by weighting the prior probabilities (as discussed in Section 6.2.2.2) and indicates a trend in improved performance (Figure 6.13). Figure 6.13a uses the a weighted combination of the prior probabilities of the two regions (see Equation 6.6) while Figure 6.13b applies a high threshold to the second region so that only definite cases are added to improve the initial classification using the first three branches (see Equation 6.7). The first method improves the general performance while the second method improves performance at high specificities (a likely operating point). These methods offer small improvements but are considerable at certain specificities – a 10-20% improvement at a number of points. However, the increase in AUC is not significant ($p=0.05$); with a larger validation set, the CI would be smaller, and these improvements should become clearer.

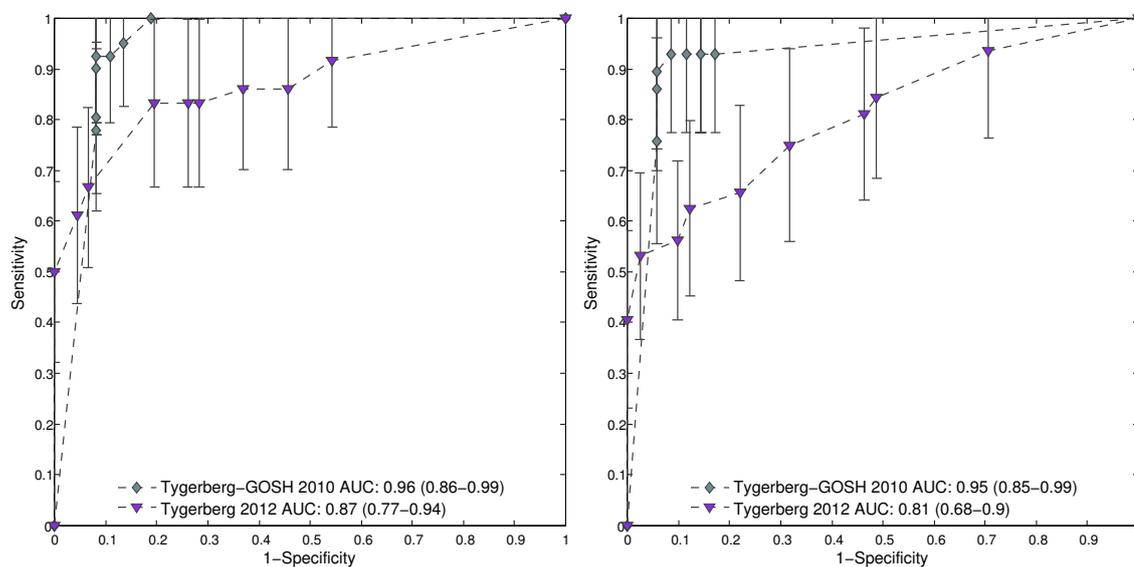
A number of cases have complete obstruction of a branch in the region being evaluated. In the trachea-LMB-RMB region there were 19 obstructions in the 160 cases



(a) Training set: Tygerberg TB 2010 and GOSH non-TB 2010

(b) Test set: Tygerberg TB 2012 and Tygerberg non-TB 2012

Figure 6.11: Fisher mapping of first 10 modes plotted against 11th mode for the RMB-RUL-BI region



(a) Trachea-RMB-LMB

(b) RMB-RUL-BI

Figure 6.12: ROC curves for the two regions of interest using the training and test sets (the error bars show the 95% confidence interval for the sensitivity)

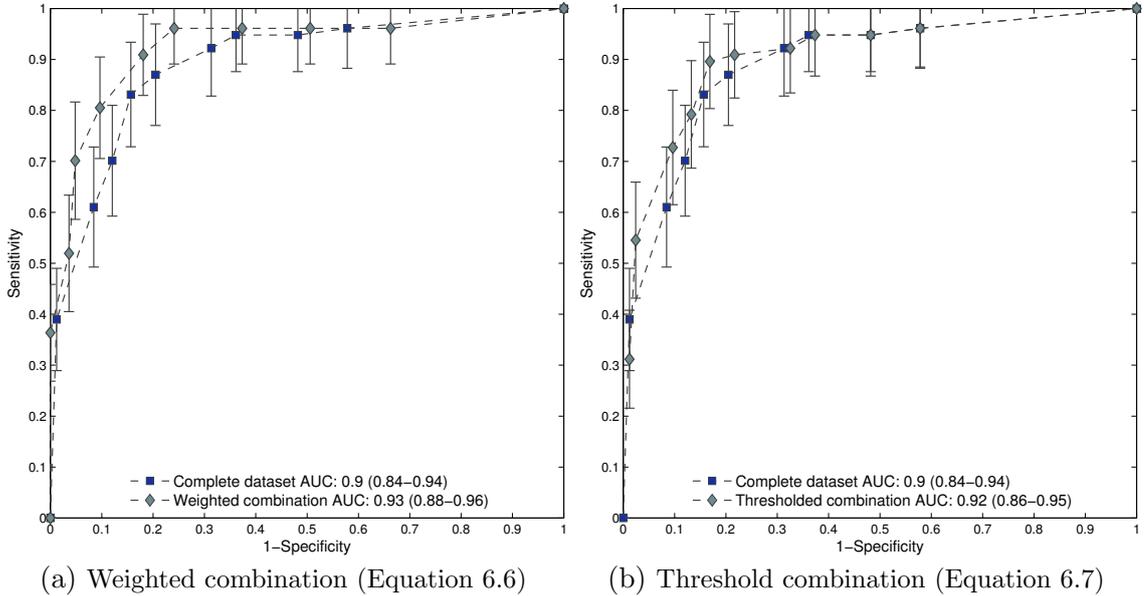


Figure 6.13: Classification based on 3-branch combinations (the error bars show the 95% confidence interval for the sensitivity)

(17 TB and 2 non-TB). In the RMB-RUL-BI region there were 42/160 obstructions (31 TB and 11 non-TB). These cases were automatically identified from the branching structure and excluded from the statistical model. By labelling these cases as TB cases and including them in the classification, a similar performance is achieved. Included obstructed branches for the trachea-LMB-RMB (AUC: 0.90 (0.85-0.94)) and for the RMB-RUL-BI region (AUC: 0.85 (0.78-0.91)), as shown in Figure 6.14. This allows more extreme variation to be included in the evaluation without impacting the statistical model. However, obstructions become less sensitive to pathology with smaller branches because the likelihood that the obstruction is caused by imaging artefacts increases.

The previous figures have shown ROC with the accuracy of distinguishing between TB and non-TB patients using airway shape features. These results show the evaluation of the airway detection algorithm on training and validation sets. All the methods use 11 principal components, representing 90% of the variation, for classification. Figure 6.15 considers the impact of the number of components on the accuracy of detecting paediatric pulmonary TB. Figure 6.15a and Figure 6.15b show

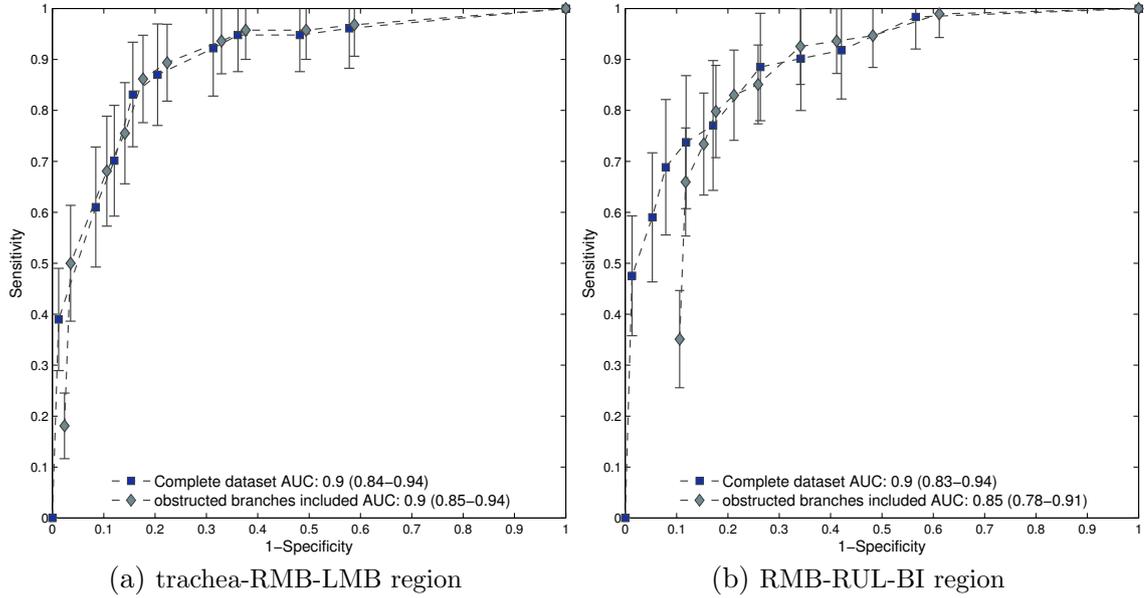
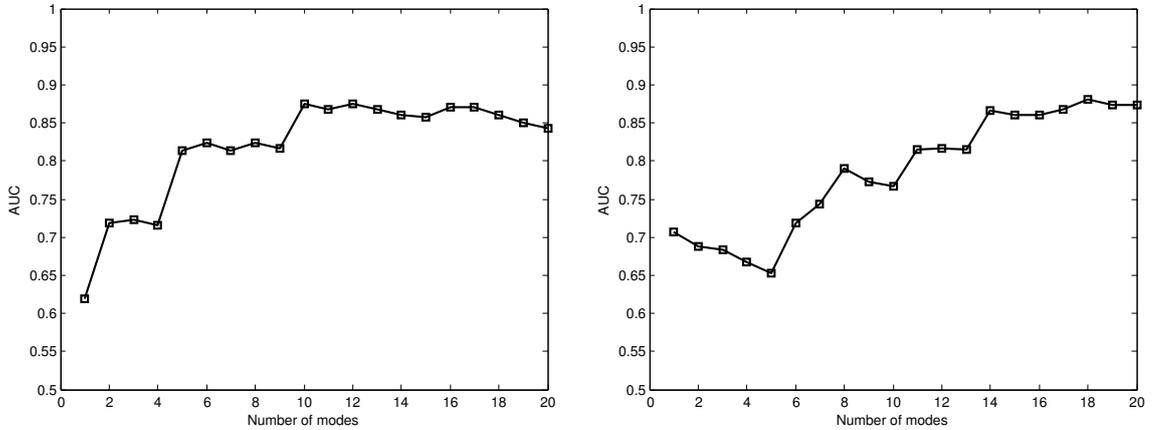


Figure 6.14: Including obstructed branches in airway region classification (the error bars show the 95% confidence interval for the sensitivity)

the AUC vs the number of components used for detecting pathology in the trachea-RMB-LMB and RMB-LMB-RUL regions. The AUC was calculated using LOOCV on the validation set.

Interesting to note, is that large improvements in performance are seen when adding certain components while others offer no improvement (or even lower the accuracy). The modes that allow TB and non-TB cases to be distinguished include the modes 1, 2, 5 and 10 for the trachea-RMB-LMB region and modes 1, 6, 7, 8, 11 and 14. Other modes may also contribute to the classification but only improve the classification when used in combination with one of the modes described above. Figure 6.15b shows that the AUC could be improved further if more modes than those contributing to the 90% variance were included. The plots also show that adding certain modes can worsen the performance of the classifier; fitting a classifier to a feature that does not contribute to the distinguishing the datasets has the potential to be penalised by the cross validation.



(a) trachea-RMB-LMB region. Including modes 1, 2, 5 and 10 improves the accuracy of the classification.

(b) RMB-RUL-BI region. Including modes 1, 6, 7, 8, 11 and 14 improves the accuracy of the classification.

Figure 6.15: AUC as a function of number of modes used in the classification.

6.4.1.4 Support vector machine classification and separate training and testing sets

In this study a separate training and validation set were used. The training set was used to determine parameters throughout the model while the previously unseen validation set was used for evaluation. Results were generated using the classifier with LOOCV on the validation set. A separate classification was made using the training set for comparison.

LOOCV was used instead of training the classifier on the training set because, more severe pathology appears to be present in the Tygerberg non-TB test set than the GOSH non-TB training set. This was also qualitatively observed. The classifier would, therefore, be biased to classifying many more cases in the test set as TB (because of the increase levels of pathology). In order to get in a suitable range of specificities, for this case, a very high threshold for classifying TB has to be applied to the posterior probabilities of classifier. This can also be seen from the differences in class separation between the two sets in Figure 6.10 and Figure 6.11. Using this form of test an AUC of 0.78 (0.66-0.87) was obtained for the trachea-RMB-LMB region and an AUC of 0.77 (0.64-0.86) was obtained for the RMB-RUL-BI. An almost 10% drop in performance. This performance was regained if the GOSH non-TB and Tygerberg non-TB were

divided between the training and test sets.

The results discussed in this chapter use a LDA classifier for evaluation. A SVM classifier was also applied to the dataset but obtained similar results on an initial evaluation (see Irving et al. (2011a)). Therefore, to improve algorithm speed LDA was used.

6.4.2 Radius derived features

The previous section discusses and evaluates the airway shape analysis method developed in this project. That version of the method uses features derived from the PDM of 3-branch regions of the airway. Alternate feature vector representations of the airway (discussed in Section 6.2.1.3) were also explored and results are presented here for comparison.

Figures 6.16a and 6.16b show results for the two regions using the average diameter based features. These features were derived from the average diameters of three equal sections of each branch. The method achieved a fair classification rate of 0.72(0.59 – 0.83) on the validation set for the first region but failed on the second regions 0.59(0.46 – 0.73). This suggests that branch diameter is not enough for the detection of airway shape variation from paediatric TB, and other methods to model variation – including PCA – can improve detection.

The second set of radius based features attempted to classify TB pathology based on observations of airway changes in the training set. This method used ratios of orthogonal diameters to represent branch compression, ratio of diameter to length to represent branch narrowing, detection of local minima to model local stenosis and a measurement of the carinal angle. These results improved on the previous method with a classification AUC of 0.83(0.73 – 0.91) and 0.75(0.64 – 0.85) on the validation set for the trachea-RMB-LMB and RMB-RUL-BI, respectively (Figures 6.16c and 6.16d). These results were not as good as the results from the modes of the point distribution model (as discussed in Section 6.4.1) of 0.87(0.77 – 0.94) and 0.81(0.61 – 0.90) (as shown in Figure 6.12 and presented again in Figures 6.16e and 6.16f for comparison).

The significance of the AUC improvement was calculated using bootstrapping (bias corrected and accelerated with 5000 replications using matlab’s `bootci` func-

tion) with a two-tailed Monte Carlo technique. Assuming a significance level of $\alpha = 5\%$, the PDM derived features were significantly better than the average diameter derived features for both the Trachea-RMB-LMB ($p=0.024$) and RMB-RUL-BI regions ($p=0.013$). While the PDM derived features are not significantly better than the second radius based method ($p=0.405$ and $p=0.465$ for the two regions respectively), there is a trend in performance improvement. The PDM modes also maintain a significantly better sensitivity ($p<0.05$) at very high specificities compared to both alternative methods, which is a useful operating point. The flexibility of the PDM to model variation automatically without having to “design” features for each type of object is also a considerable advantage.

6.4.3 Run time

This software was written using a combination of Matlab R2011a and C++. C++ was incorporated to improve the speed of various bottlenecks. The software was evaluated on system with a 2.80GHz Intel quad-core processor and 6GB of ram. The mean time for segmentation of each airway was 168 ± 57 s. Extraction of the centreline, pruning and branch point detection for each airway took a mean time of 22 ± 18 s. Once each airway was segmented and the structure extracted, then mesh generation, registration, training and classification of the entire 179 patient dataset took 1162s on a single run.

6.5 Discussion

The airway analysis method discussed in this thesis is able to accurately distinguish between TB and non-TB airways by examining regions of the airway likely to be affected by lymphadenopathy (AUC of 0.87(0.77 – 0.94) and 0.81(0.68 – 0.90)). The training set performed better on all regions (although the improvement in AUC is not significant $p=0.05$). This better performance in the training set could be due to the fitting of parameters but is more likely that the GOSH non-TB dataset used in this step contains less airway pathology. The test set performed better on the trachea-RMB-LMB region suggesting that this is more useful than the RMB-RUL-BI region. Combining the classification of both regions yielded the best results for the

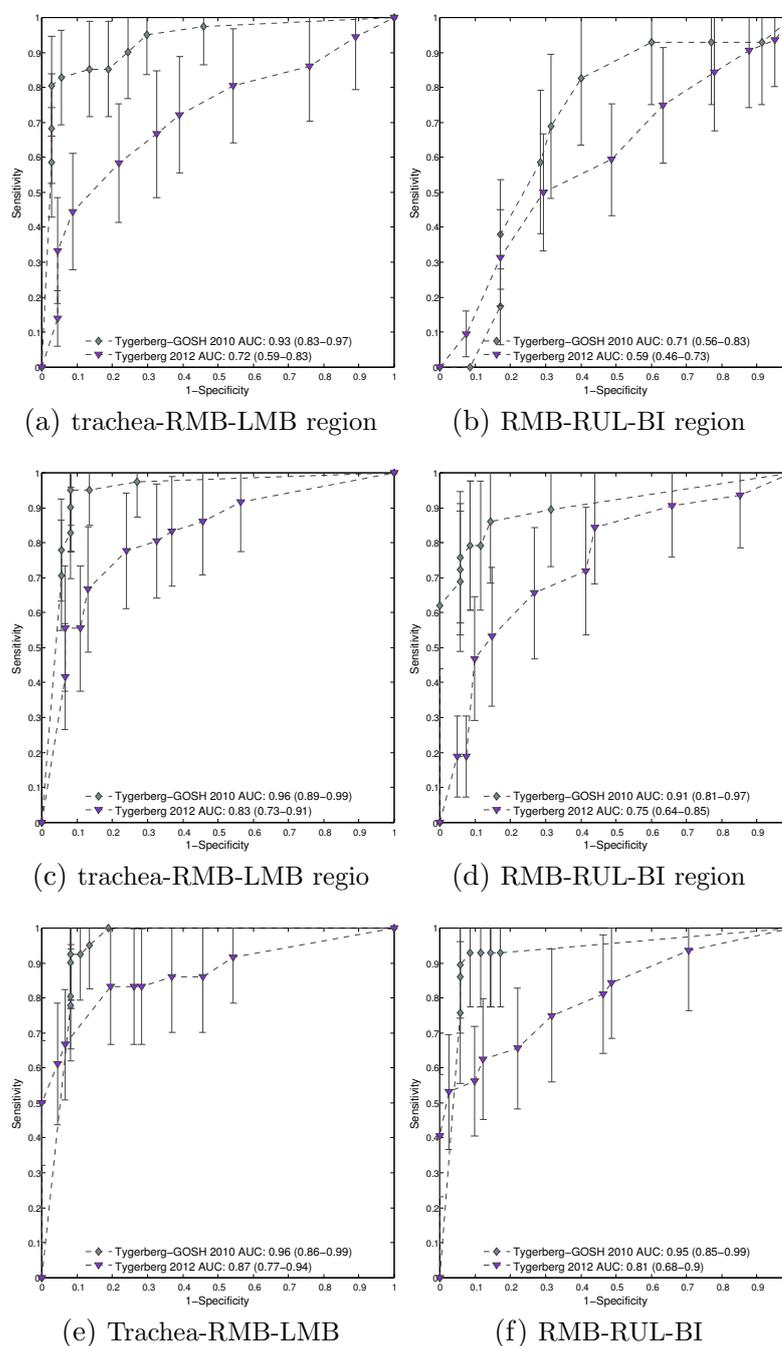


Figure 6.16: ROC curves for the two regions of interest using the training and test sets (the error bars show the 95% confidence interval for the sensitivity). a) and b) Average diameter based features (described in Section 6.2.1.3). c) and d) Features based on observed stenosis (described in Section 6.2.1.3). e) and f) PDM derived features (repeat of Figure 6.12).

entire dataset.

PDM based features performed better than two versions of radius based features on the validation set, particularly for the second generation region. Radius features 1 has AUC of 0.72(0.59 – 0.83) and 0.59(0.46 – 0.73). Radius 2 has an AUC of 0.83(0.73–0.91) and 0.75(0.64–0.85). The PDM derived features show more promise and this is likely to come from the ability to represent more complex variation in feature space than cross-sectional radius derived features, which focus on local regions of the airway. This would include spatial variation of branches in the region. An additional benefit of PDM based features is that the principal modes are adaptable to the training set and do not have to be specifically chosen with a particular type of deformation in mind.

The benefits of a more complex classifier were evaluated using Support Vector Machines with a polynomial kernel and nested cross validation for parameter selection and testing. This did not improve the performance of the classification and, in the interests of speed, LDA was used. All parameters were determined on the training sets (*TYGTB2010* and *GOSH2010*). Leave-one-out cross validation was used to evaluate the performance of the TB classification algorithm on the training and test datasets. Initially, the original datasets were used to train the linear classifier while the new datasets were used to test the classifier. This led to reduced performance (AUC = 0.78 (0.66-0.87)) probably because the nonTB Tygerberg dataset contains more severe non-TB pathology than the GOSH dataset and, therefore, different optimal decision hyperplanes (See Figure 6.10 and 6.11).

The methods were evaluated on local three-branch regions of the airway, which were chosen as the most likely to exhibit signs of deformation from lymphadenopathy and were used to classify each airway as *TB* or *non-TB*. The principal modes of each local region give information represent specific types of deformation and stenosis. Given a labelled dataset of the position of the lymphadenopathy then the classifier could be used to label lymphadenopathy e.g. hilar, mediastinal or paratracheal lymphadenopathy.

Classification aims to distinguish pathological shape variation from inter-patient variation, and variation due to age and breathing artefacts. The standard deviation of the patients' age is approximately 3 years (see Table 6.1). With a larger dataset, it would be possible to divide the dataset into age groups and possibly improve the

performance by removing noise caused by age variation. However, studies have shown that the proportions of the airway do not change considerably with age in children (Masters et al., 2006). Breathing artefacts could also add noise to shape-based airway classification as it is not possible to perform a breath hold scan on young patients. As discussed in Section 1.5, the bronchi lengthen and dilate during inspiration but these changes are expected to be distinguishable from TB pathology. The accuracy of the classifier also indicates that breathing variation does not have a considerable effect. However, the results could possibly be improved if breathing was explicitly modelled.

6.6 Conclusion

Paediatric TB is still difficult to diagnose and this study shows the potential of automated airway shape analysis to assist. The chapter uses a dataset of TB patients (with signs indicating possible airway involvement) and non-TB patients to develop a method that uses local airway deformation to detect new cases with TB. Features are derived from airway cross sections and point distribution models of the airway surface, and used to train a classifier. The method was evaluated on regions of the airway known to be affected by lymphadenopathy in paediatric pulmonary TB, which include the trachea, LMB, RMB, RUL and BI. The best performing features were derived from the surface point distribution model and achieved an AUC of 0.87(0.77 – 0.94) and 0.81(0.68 – 0.90) using the Trachea-LMB-RMB and RMB-RUL-BI regions, respectively.

This method could be used to provide additional automated visualisation and analysis for any patient undergoing a CT examination and might be applied to other diseases affecting the airway. Bronchoscopy is currently the “gold standard” for evaluating airway stenosis and deformation but is an invasive procedure. CT with techniques such as virtual bronchoscopy and virtual rendering, shows potential as an alternative to bronchoscopy. However, these methods require considerable manual interaction in the form of setting thresholds, viewing parameters, and manual assessment of the airways. The method discussed in this chapter has the potential to improve the analysis of airway shape deformation. An example application would be the identification and highlighting of particular cases – and regions of the airway in

those cases – that show signs of pathology. This would speed up assessment. Extensions of this method would also allow grading of the pathology in each region and identification of the lymph node of interest from the airway deformation.

The model could also be applied to 2D chest radiographs which would expand the application to routine TB screening. A 3D model of normal and abnormal pathology provides a detailed anatomical model that can be projected onto 2D radiographs, which have limited airway visibility. This could be used to assist in the airway segmentation and analysis of radiographs.

6.7 Chapter summary

This chapter presents feature extraction and classification techniques for airway analysis, and evaluates the algorithms performance:

- Feature extraction
 - First method uses point distribution model derived features
 - Second method uses cross section radius derived features
- Classification
 - Classification from feature vectors
 - Local 3-branch regions of the airway tree
 - Obstructed branches classified separately
 - Linear discriminant analysis
- Evaluation
 - Evaluated on TB and non-TB paediatric patients
 - Variables selected using the training set
 - Leave-one-out cross validation performed on the datasets
 - ROC and AUC with 95% confidence intervals
- Results
 - Best results using the PDM derived features
 - $AUC = 0.87(0.77 - 0.94)$ for the trachea-RMB-LMB region (validation set)
 - $AUC = 0.81(0.68 - 0.90)$ for the RMB-RUL-BI region (validation set)
 - Combining multiple regions improves performance

7

Model projection and airway segmentation in radiographs

Abstract. This chapter introduces a novel approach to the segmentation of the airways in 2D radiographs using a 3D statistical shape model to guide the segmentation. The silhouette of the 3D model is detected, projected onto a 2D plane and aligned to the airways using manual landmarks. The parameters of the 3D shape model are then optimised for each radiographic image in the test set using an energy function based on image gradient, greyscale shape kernels and a regularisation term. The optimisation is performed using the Nelder-Mead downhill simplex algorithm. The algorithm achieved a 6.8 ± 2.6 pixel error on the test set, a $25 \pm 17\%$ improvement on the initial rigid alignment using 4 manually-identified landmarks.

Previous chapters develop methods in airway segmentation, registration and analysis of CT scans that are used for the detection of paediatric TB. The results show that signs of TB can be detected by airway shape variation. These methods have only been applied to 3D representations of the airway tree; there is potential to extend this work to 2D imaging.

CT scans offer far better airway visibility than radiographs, which have poor airway visibility – particularly because anatomical structures overlap in a 2D projection. Therefore, it may be useful to project the 3D model to 2D to assist in the segmentation and analysis of 2D chest radiographs. In this chapter, a method is presented for the analysis of the airways in chest radiographs using the 3D statistical shape model that was generated in previous chapters.

Section 7.1 discusses the potential of using chest radiographs for airway analysis, introduces previous attempts to segment the airway tree, and surveys 3D to 2D projection methods. Section 7.2 proposes an algorithm to project and fit the model to chest radiographs, including: airway silhouette detection (Section 7.2.1), projection of the 3D model (Section 7.2.2), identification of image features (Section 7.2.3), and fitting the projected model to the 2D image (Section 7.2.4). The method is evaluated on a set of Lodox scans of paediatric patients with and without TB by comparing the accuracy of the segmentation to manual annotations (Section 7.3). The results are presented in Section 7.4.

7.1 Background

Due to the relatively low radiation dose and cost compared to CT scans, radiographs can be used for routine imaging and population screening. Including airway analysis in a screening tool, could improve the diagnosis of paediatric TB.

Radiographs usually have poor airway visibility but the Lodox linear slit scanning radiography unit offers improved visualisation of the trachea and main bronchi (Pitcher et al., 2008). In a study evaluating the airway visibility of the airways in 102 patients between 6 months and 13 years, the LMB and RMB could be seen and measured in over 90% of the cases by trained clinicians (Daya et al., 2010; Sanders et al., 2009). Sanders et al. (2009) also suggest that airway shape from radiographs could be used to detect abnormalities. The improvements in airway visibility could be explained by the higher kV used for the Chest AP examination and lower scatter of the slit scanning technology (Irving et al., 2008). An additional benefit of the linear slit scanning technology is the low dose compared to conventional radiography (Irving et al., 2008; Maree et al., 2007).

Few previous attempts have been made to segment the airways in radiographs

due to the poor airway visibility. Long (2008) developed a method that uses automatically identified landmarks to initialise the airway segmentation. He creates an airway template and applies affine and then piecewise warping to the template using landmark points on the lung border and airway. An expectation maximisation based method is then used to match the template to the airway in the radiograph. This method was able to label on average 50% of the airway region¹. This is too low to be clinically viable but, given that this method is completely automatic, the method demonstrates that it is possible to identify the airway region in a chest radiograph.

Tezoo and Douglas (2012) build a semi-automatic Active Shape Model (ASM) based method to segment the airways. The ASM includes the trachea and first bifurcation region and consists of 30 edge points built from manual annotations on a training set of chest radiographs. The method is semi-automatic and requires the user to input 9 landmark points on the airway edge to establish the initial alignment. The method was evaluated on 29 test images and achieved a mean Hausdorff distance of 18.9 ± 9.9 pixels.

As an extension of ASM, greyscale variation can be incorporated into the statistical model. Cootes et al. (2001) propose active appearance models (AAM) that incorporate both shape and greyscale variation of the training set in the modes of variation to identify new shapes. However, the airways overlap ribs, vertebrae and the heart. These structures have greater contrast and stronger edges than the airways, and the position of these structures relative to the airway is sensitive to small changes in the position of the patient. Therefore, greyscale variation was not considered to be a specific enough feature to use AAM.

The method proposed in this chapter uses a 3D statistical shape model of the airways for the LMB, RMB and trachea – the only part of the airways visible in most cases (Sanders et al., 2009). As discussed previously, bronchial compression from lymphadenopathy is most common in the LMB, RMB, BI and trachea (Andronikou et al., 2004). A 3D model is able to evaluate multiple views simultaneously, for example AP and lateral images.

Projection of a 3D model onto 2D radiographs has not been applied to the airways but has been used for bone structures including the rib cage (Benameur et al., 2005), the vertebra (Benameur et al., 2003), the femur (Zheng, 2011) and hip (Zheng, 2009).

¹(The percentage is calculated as $TP / (TP+FP)$)

Benameur et al. (2003) detect the silhouette of the 3D model, project the model onto a 2D surface and optimise the fit of the model to edges in the image. The edges are generated using a Canny edge detector. This model is simultaneously fitted to an AP and lateral radiograph and a regularisation term is used to constrain the 3D model.

Groher et al. (2007) and Ruijters et al. (2009) apply 3D registration of vessels to angiograms. Groher et al. (2007) use an initial segmentation of the vessels while Ruijters et al. (2009) use the vesselness measure to identify vessels. Vesselness uses the eigenvalues of the Hessian matrix to assign a probability to each pixel of belonging to a vessel. This is effective for angiograms but for chest X-rays, overlapping bone structures have a higher contrast and stronger edges, making vesselness ineffective and an initial airway segmentation very challenging.

The main contribution of this chapter is the development of a method to register a 3D statistical shape model to 2D radiographs, which has not been previously attempted, to assist in the segmentation of the airways in the radiograph. The 3D shape model along with novel 2D features are chosen to allow the airways to be detected while overlapping structures with stronger edges and higher contrast. A model developed from CT can include a number of generations of bronchi and is more complete than a model from radiographs where the RMB and LMB are often not completely visible. Thus, a 3D model can be used to build a more accurate representation of the airway tree to assist with segmentation in radiographs. Benefits of using a 3D model over a 2D model also include the potential to improve the segmentation by fitting multiple radiographic views simultaneously, develop a personalised 3D model from radiographs, and use the parameters of the 3D model to detect shape pathological shape variation (in a similar way to the method described in Chapter 6).

7.2 Method

This section outlines the 3D to 2D model projection and fitting method used for segmentation of the airways in chest radiographs. The 3D statistical shape model developed in Chapter 6 is used for the analysis. As described in Figure 7.1, a silhouette of the mean model is found (Section 7.2.1) and the silhouette points are projected onto a 2D plane and aligned to the airway region (Section 7.2.2). This 2D projection is used to generate an energy function to measure the fit of the model to airway

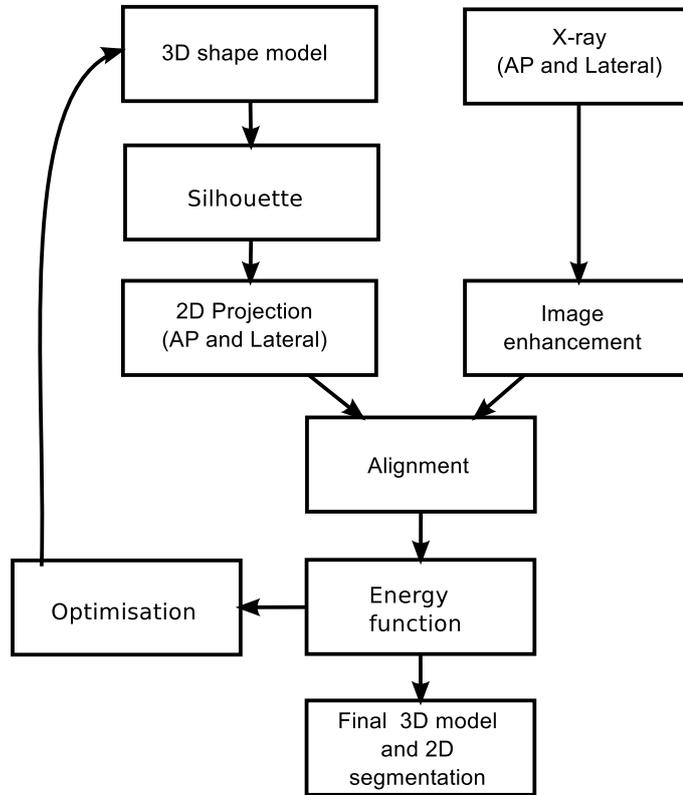


Figure 7.1: The proposed projection method to segment and model airway variation in chest radiographs. The 3D model is projected onto processed chest radiographs. The model parameters are then optimised with an energy function to obtain a locally optimal fit of the 3D model. The optimisation procedure updates the 3D shape, projected shape and energy function

features on the radiograph (Section 7.2.4). The parameters of the 3D model are then optimised to determine the best fit of the 3D model to the 2D image. Due to the ill constrained nature of optimising a 3D shape on a 2D projection, this optimisation is better suited to simultaneous optimisation of both AP and lateral radiographs. However, due to limited data, the method discussed in this chapter is only evaluated using a single perspective (chest AP radiographs).

7.2.1 Silhouette Detection

The silhouette of an object can be used to define the edges of the surface mesh that will be visible in a 2D projection. There are a number of methods to extract a

silhouette, although, for small models (meshes with less than 10 000 faces) a brute force approach has almost the same speed as more complex methods (Hartner et al.). A brute force approach for the detection of airway silhouette edges was implemented in this method to output the silhouette points for projection.

The normal (N) of each face in the airway surface mesh is found using the cross product of two edges². Silhouette edges in a mesh are defined as edges separating a face with a normal facing towards the viewer and a face with a normal away from the viewer (Hartner et al.). The inner product of the face normal (N) with the viewing direction (E) is used to detect the orientation of each face. If $N \cdot E > 0$ then the polygon faces frontwards (as shown in Figure 7.2). If $N \cdot E < 0$ then the polygon faces backwards and if $N \cdot E = 0$ then the face is orthogonal to the viewing direction.

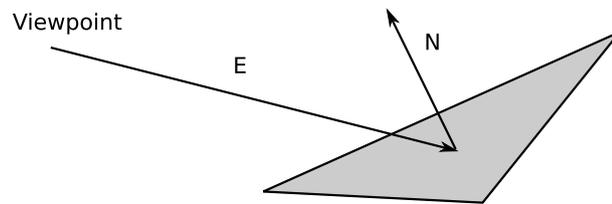


Figure 7.2: The viewing direction (E) and outward facing normal (N) of a triangle, which are used to define the direction of the face

Therefore the algorithm used to determine silhouette points is as follows:

- Determine face normals of the mesh
- Identify the two faces connected to each edge (f_1 and f_2)
- For each edge:
 - Find $d_1 = N_1 \cdot E_2$ for f_1 and $d_2 = N_2 \cdot E_2$ for f_2
 - If $d_1 d_2 \leq 0$ then the edge forms part of the set of silhouette points (**S**)

²The meshing procedure produces vertices of a consistent order. Therefore, the cross product of two edges will always lead to a consistent normal direction relative to the volume.

7.2.2 Model Projection and alignment

Once each silhouette point is identified³, the magnification of the shape (represented by the silhouette points) needs to be calculated in order to model radiographic projection onto the detector plane. The projection of a point onto a plane can be calculated using vector-plane intersection.

If a vector (\mathbf{v}) from the viewpoint (\mathbf{l}_0) to a silhouette point in direction \mathbf{l} is written as $\mathbf{v} = \mathbf{l}_0 + d\mathbf{l}$, then d at the point where the vector intersects a plane of interest can be shown to be:

$$d = \frac{(\mathbf{p}_0 - \mathbf{l}_0) \cdot \mathbf{n}}{\mathbf{l} \cdot \mathbf{n}} \quad (7.1)$$

where \mathbf{p}_0 is a point on the plane and \mathbf{n} is the normal of the plane.

Therefore, the projection of each silhouette vertex, from a viewpoint to a defined plane, can be calculated, to simulate radiographic projection of the 3D model onto the detector plane. Projection of both silhouette vertices and skeleton points is performed. Figure 7.3 shows the projection of the silhouette points onto a plane.

In our case we are simulating the projection of the airways onto a radiograph using the Lodox Statscan linear slit scanning radiography unit. Lodox uses a thin fan beam scanned across the patient and, therefore, the projected image is not magnified in the scanning direction. To take this into account, the z-position of each point is kept constant. A rough approximation of the Lodox Statscan dimensions was used by defining the source viewpoint as:

- X-ray source “view point” (\mathbf{l}_0)=[1000mm, 0mm, 0mm]
- Detector plane (\mathbf{p}_0)=[-100mm, 0mm, 0mm] with normal (\mathbf{n}) [-1,0,0]
- Airway model centred at [0mm, 0mm, 0mm]

Once the silhouette points have been projected onto the detector plane, this algorithm requires some manual interaction to align the model to the airways in the

³each silhouette edge is comprised of a start and end silhouette point.

image. Procrustes analysis (discussed in section 6.2.1.1) was used to align the model with 4 manually annotated landmark points. Figure 7.4 shows the alignment of the 4 bifurcation points of the projected 3D model with the 4 manual annotations on the image. The optimal scaling, translation and rotation, calculated from the procrustes analysis of the landmark points, are used to transform the silhouette and skeleton points.

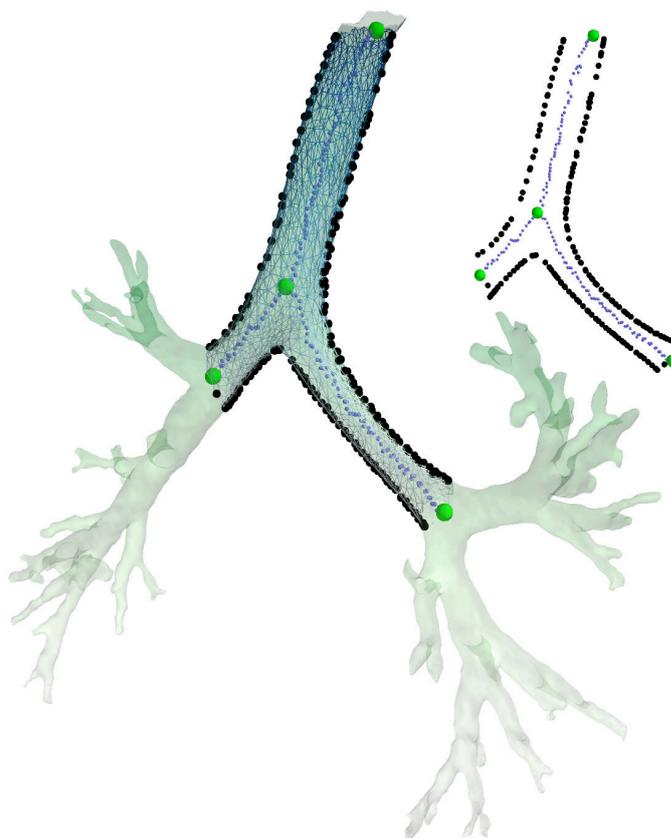


Figure 7.3: Airway silhouette detection and projection. Black points show the detected silhouette vertices on the trachea, LMB and RMB and the projected points onto a surface. The locations of the skeleton (blue points) and the branch start and end points (green points) are also projected.

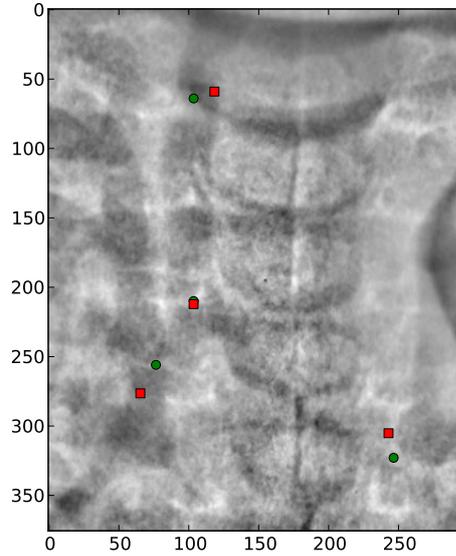


Figure 7.4: Procrustes analysis of the 4 manual landmarks of the image (defining the branch start and end positions) with the projected branch landmark points. Green dots show the manual landmarks on the airway of the image and red squares show the aligned landmarks from the 3D model

7.2.3 Image Features

An energy function derived from image features was used to fit the projected shape model to the airway region of the radiograph.

7.2.3.1 Preprocessing of the radiograph

There is large variation in the image intensity of the airway due to overlap with the ribs, vertebra and pulmonary structures. In order to reduce this variation, local normalisation was applied to the image, which has been used previously for airway enhancement (Long, 2008)

Local normalisation can be described by:

$$I_{ln} = \frac{I - G(I)}{\sigma_f} \quad (7.2)$$

where $\sigma_f = \sqrt{G((I - G(I))^2)}$ is the estimate of the standard deviation at each pixel

and $G(I)$ is the Gaussian blur function⁴ applied to image $I(x,y)$. Figure 7.5 demonstrates the effect of local normalisation on a chest radiograph.

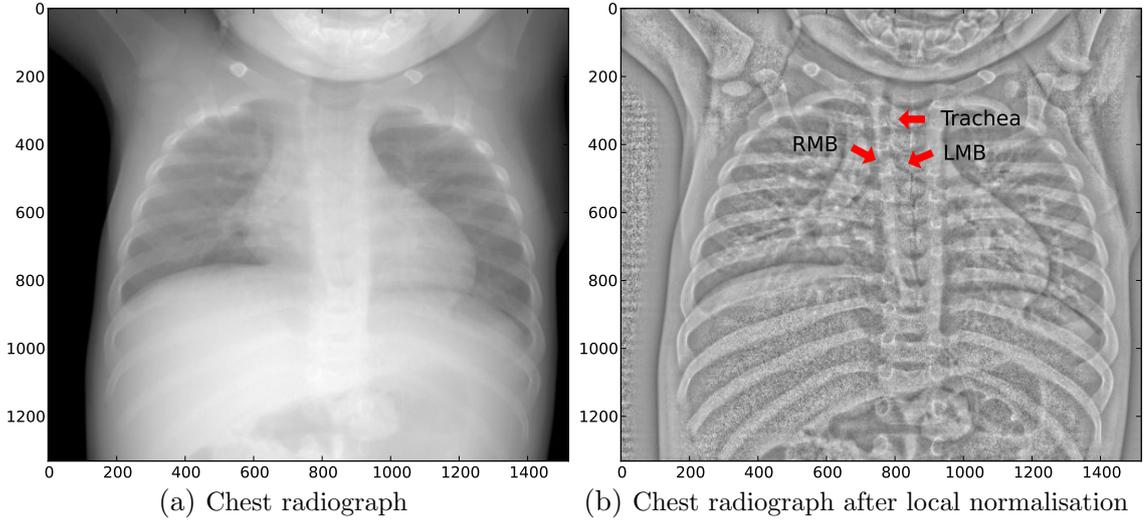


Figure 7.5: Improvement in airway visibility after the local normalisation filter is applied to the image

7.2.3.2 Gradient

Vesselness is a common method for detecting tubular structures in images and Canny edge detection is a common technique for edge detection. Both these methods were attempted but failed to provide suitable airway features because of overlapping structures with greater intensity and stronger edges. Instead, the smoothed gradient of the image was used as one set of image features. The Sobel filter was used to extract the gradient and a Gaussian convolution of $\sigma = 5$ was used to smooth the gradient.

The gradient derived energy function was calculated as the sum of the gradient at edge silhouette point:

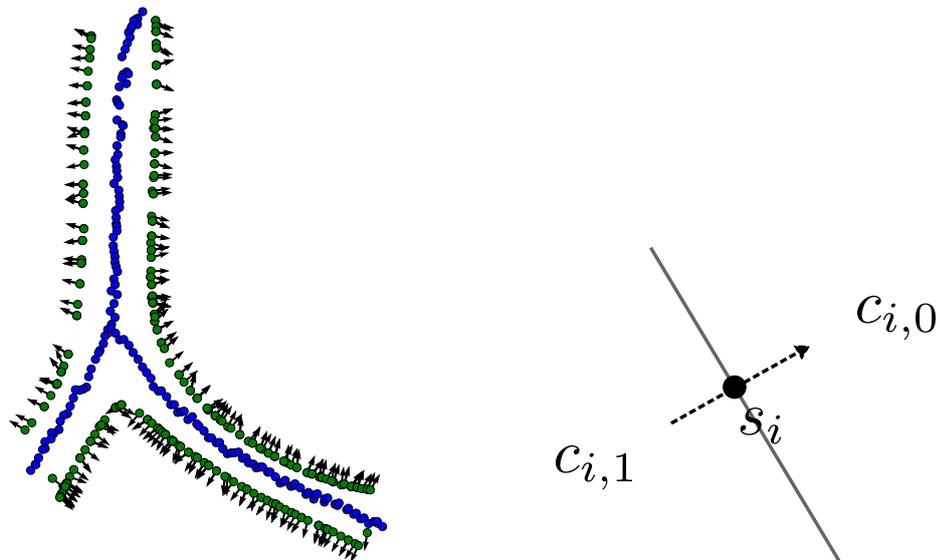
$$E_{grad} = - \sum_{i=1}^p |\nabla I(s_i)| \quad (7.3)$$

⁴Using the training set, $\sigma = 20$ pixels was used (approximately 80 % of the maximum airway radius). However, similar results on the test set were obtained for the range $\sigma = 10$ pixels to $\sigma = 30$ pixels.

where s_i are the p projected silhouette points on image I.

7.2.3.3 Greyscale

Edges detectors provide useful features for airway segmentation but if used alone will not always be effective when there are additional strong edges. A second characteristic of the airway is that it appears as a shadow over higher contrast structures. Therefore, while the image is often darker or brighter surrounding the airway, any structure overlapped by the airway should appear darker in the airway. Comparing the greyscale inside and outside the airway, can be used to optimise the fit. The outward direction of each silhouette point is defined as the vector from the closest skeleton point to the silhouette point, as shown in Figure 7.6.



(a) The outward direction of each silhouette point. The blue points represent the unsmoothed centreline, the green points represent the silhouette points of the first three branches, and the arrows represent the outward pointing vectors found from the centreline.

(b) Greyscale comparison inside and outside the airway. The mean greyscale value is found for $c_{i,1}$ and $c_{i,0}$ for each s_i .

Figure 7.6: Comparing greyscale of kernels inside and outside the airway region for each silhouette point

Once the orientation of each silhouette point is found then the mean greyscale

value of kernels just inside and just outside the airway can be compared. The orientation of the line connecting the two circular kernels is always perpendicular to the current airway wall. This method is related to a method used by Aylward et al. (2003) to estimate the cross section of a tubular object. However, because the region of interest is the airway projection and not a tube cross section, the kernels are positioned using the airway shape model instead of a defined circle.

The energy function was defined as:

$$E_{grey} = \sum_{i=1}^p \frac{c_{i,1} - c_{i,0}}{c_{i,0}} \quad (7.4)$$

where $c_{i,0}$ is the mean greyscale of the sampled circle outside the airway and $c_{i,1}$ is the mean greyscale of the sampled circle inside the airway at silhouette point s_i .

7.2.4 Optimisation

The projected silhouette points along with the energy function are used to optimise the fit of the 3D shape model to the airway regions in the radiograph. From Equation A.6 in Chapter 6, a shape can be described by variation along each eigenvector from the mean model:

$$\mathbf{x} = \bar{\mathbf{x}} + \Phi \mathbf{b} \quad (7.5)$$

where $\bar{\mathbf{x}}$ is a $n \times 3$ array of vertex positions of the mean model that are deformed by m eigenvectors $\Phi = (\phi_1 | \phi_2 | \dots | \phi_m)$ with weights $\mathbf{b} = (b_1, \dots, b_m)$. Therefore, \mathbf{b} can be used to choose a suitable model that is the best fit to a single airway case.

E_{edge} and E_{grey} are used to determine the optimal \mathbf{b} . A total energy (E) was created from a weighted combination of the energy functions:

$$E = \alpha_e E_{edge} + \alpha_g E_{grey} + r E_{reg} \quad (7.6)$$

where α_e and α_g are the weights. An additional regularisation term (E_{reg}) was included with weight (r). Regularisation is required because the 3D model is poorly constrained by optimising on a 2D projection. E_{reg} is chosen as:

$$E_{reg} = \sum_{j=1}^m \frac{|b_j|}{\lambda_j} \quad (7.7)$$

where b_j is the weighting factor of each eigenvector and λ_j is the eigenvalue of the j th eigenvector – equivalent to the variance of each mode within the dataset. This term penalises variation of the model away from the mean model, while penalising modes with larger variation in the dataset less than modes with smaller variation.

The local minimum of E was used to find the optimal \mathbf{b} , and therefore the best segmentation, for each chest radiograph. Most optimisation methods require a continuous function. We are working in discrete image space and so linear interpolation was used to determine the features used in E_{edge} and E_{grey} at any required value for s_i .

An optimisation algorithm was used to minimise E . At each step in the optimisation: \mathbf{b} is updated, the airway shape projected onto the image and E recalculated (shown in Figure 7.1 and Figure 7.7). Nelder-Mead downhill simplex optimisation method (Press et al., 2002) was used to find the local minimum of E from the initial mean model. A simplex is an object consisting of $m + 1$ points and connecting edges in m dimensional space (for the optimisation of a function with m parameters). After initialisation, the method iteratively moves the point, in the simplex, with the largest energy through the opposite face to a lower energy until a local minimum is reached. The simplex expands and contracts depending on the topology of the energy function. The downhill simplex method uses only function evaluations and does not require the function gradient to be calculated. It has a higher computational cost than other methods but is useful for prototyping (Press et al., 2002). Figure 7.3 shows initial mean model and Figure 7.8 shows the 3D model after optimisation to case 6.

7.3 Evaluation

The method was evaluated on a dataset of chest radiographs and the optimal fit was compared to manual annotations.

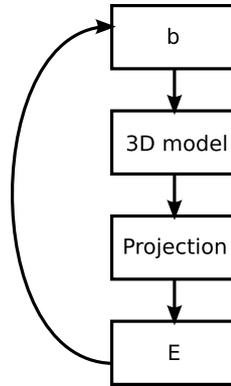


Figure 7.7: The optimisation method calculates a new \mathbf{b} at each step. A new 3D model is generated for this \mathbf{b} and projected onto the surface to recalculate \mathbf{E} . These steps are repeated until a local minimum is reached.

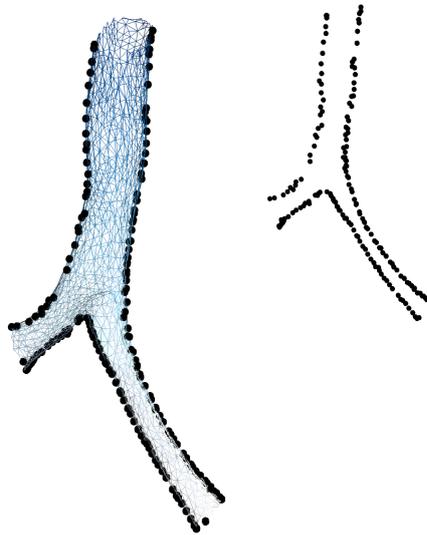


Figure 7.8: Case 6 from the training set as an example of the optimised 3D shape and projection. The optimisation is initialised with the mean model ($\mathbf{b}=[0,0,0,0,0,0,0,0,0,0]$). Final $\mathbf{b}=[0.0011566, -0.06434384, 0.09403463, -0.04700209, 0.0121405, 0.05619736, -0.00784436, -0.02942235, 0.04281598, 0.04363968, -0.01013572]$ after optimisation.

7.3.1 Dataset

A 41 patient chest X-ray dataset was obtained from Red Cross Children’s Hospital, South Africa. The images were acquired using the Lodox Statscan linear slit scanning radiography unit. The dataset contained patients that tested positive for TB from culture and patients without TB. The patient age ranged from 3 months to 60 months. The dataset was divided into a 10 patient training set (5 TB and 5 non-TB) and a 31 patient test set (11 TB and 20 non-TB). A small training set was used because the model is derived from a large set of CT images (see Chapter 6), and the 10 radiograph training set is only required to determine best weights (α_e , α_g and r) for the energy function.

7.3.2 Manual annotations

Manually drawn airway outlines were used to evaluate the method. As part of this study, a tool was created to allow a user to draw outlines by inputting a series of points. These line segments were then stored for each airway. Due to the amount of time required to perform each annotation, the author of this thesis created manual airway outlines for the evaluation of this algorithm. In future, a trained radiologist will be used to provide a second opinion. An example manual segmentation is shown in figure 7.9. The entire test set and the test set with manual outlines are shown, later in this chapter, in Figures 7.12 and 7.13. The method requires 4 manual alignment points, which were also input using the tool.

7.3.3 Comparison

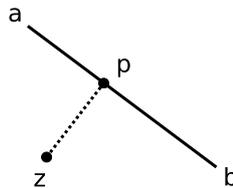
To evaluate the performance of the method, the optimised silhouette points were compared to the manual outlines of each test case. The shortest distance between each silhouette point and the manual outline was found by evaluating each segment of a manual outline.



Figure 7.9: The manually drawn outline of an example airway

The distance between a point and a line segment can be found by the orthogonal intersection of a vector from point \mathbf{z} with the line. A point \mathbf{p} on a line segment between \mathbf{a} and \mathbf{b} can be defined as:

$$\mathbf{p} = \mathbf{a} + u(\mathbf{b} - \mathbf{a}) \quad (7.8)$$



The shortest distance between \mathbf{z} and \mathbf{p} is where $(\mathbf{z} - \mathbf{p})$ is orthogonal to $(\mathbf{b} - \mathbf{a})$, that is $(\mathbf{z} - \mathbf{p}) \cdot (\mathbf{b} - \mathbf{a}) = 0$. Therefore,

$$u = \frac{(\mathbf{z} - \mathbf{a}) \cdot (\mathbf{b} - \mathbf{a})}{\|\mathbf{b} - \mathbf{a}\|^2} \quad (7.9)$$

Allowing use to obtain the closest point (\mathbf{p}) from Equation 7.8. If $u < 0$ or $u > 1$ then the closest point is \mathbf{a} or \mathbf{b} , respectively (Weisstein, a).

The distance between each silhouette point and the manual outline could be found by evaluating the closest distance to each line segment in the manual annotation and then choosing the minimum distance. The mean distance of all the silhouette points in each airway was used for evaluation.

7.4 Results

7.4.1 Airway segmentation

The 3D airway shape model was derived from CT scans as discussed in Chapter 6. The 10 training cases were used to find the optimal parameters for E (Equation 7.6). A grid search was performed to determine the weighting of edge energy (α_{edge}) and greyscale energy (α_{grey}) versus the regularisation parameter (r).

Table 7.1 shows mean error (pixels) for the training cases. This optimisation shows that there is a trend of poorer performance when α_{grey} tends to 0 or the r is large. Choosing the optimal value ($\alpha_e=0.4$, $\alpha_g=0.6$ and $r=0.003$) the performance of the segmentation algorithm was evaluated on the test set.

	$r=0.001$	$r=0.003$	$r=0.005$	$r=0.01$	$r=0.03$	$r=0.05$	$r=0.1$	$r=0.2$	$r=0.3$
$\alpha_e=1.0, \alpha_g=0.0$	7.734	7.363	7.502	7.307	7.659	7.908	8.073	8.24	8.276
$\alpha_e=0.8, \alpha_g=0.2$	6.275	5.556	5.793	5.417	6.068	6.796	7.443	7.889	8.238
$\alpha_e=0.6, \alpha_g=0.4$	6.022	5.864	5.467	5.365	5.795	6.065	6.835	7.49	7.672
$\alpha_e=0.4, \alpha_g=0.6$	5.474	5.353	5.632	5.655	5.43	5.55	6.279	7.366	7.506
$\alpha_e=0.2, \alpha_g=0.8$	6.018	5.842	5.72	5.982	5.556	5.690	6.030	6.854	7.422
$\alpha_e=0.0, \alpha_g=1.0$	5.869	6.227	6.134	5.712	5.897	5.709	5.762	6.716	7.355

Table 7.1: Selecting the optimisation parameters α_e , α_g and r using the training set. Lowest mean error was found at $\alpha_e=0.4$, $\alpha_g=0.6$ and $r=0.003$.

The mean distance from the silhouette points to the manual airway outline was calculated after the initial alignment (*initial error*) and after the optimisation of the statistical shape model (*final error*). The results for each case of the training set are shown in Figure 7.2 and the results for the test set are shown in Figure 7.3.

The mean error for all training cases was 5.4 ± 1.9 (mean and standard deviation), a 35 ± 18 % improvement on the original fit. The test set achieved a 6.8 ± 2.6 pixel error, a 25 ± 17 % improvement on the original fit. The method shows improvement for all cases of the test set except case 36. However, the method does not offer much

Case	Initial error (pixels)	Final error (pixels)	Percentage change (%)
1.	6.36	4.46	-30.
2.	9.71	3.43	-65.
3.	11.05	5.61	-49.
4.	7.03	4.86	-31.
5.	5.58	2.48	-56.
6.	6.76	4.20	-38.
7.	9.39	6.98	-26.
8.	10.48	6.06	-42.
9.	9.32	9.47	2.
10.	7.35	5.96	-19.

Table 7.2: The mean error after alignment (*initial error*) and SSM optimisation (*final error*). *Percentage change* shows the change after SSM optimisation

improvement on 4 of the cases: 20, 26, 30 and 31.

Figure 7.10 shows case 11 from the test set – an example of a good fit. The initial error of 6.67 pixels is improved to 3.10 pixels. Figure 7.11 shows a good fit except for failure on the one wall of the trachea. Figure 7.12 shows all filtered airway regions in the test dataset and Figure 7.13 shows the fit to these cases.

The robustness of the algorithm to variability of the landmark points was tested by assigning a second set of landmark points to the bifurcation regions in the training images. This tested the impact of intra-observer variability on the accuracy of landmark assignment. The mean error for all training cases was 5.4 ± 1.5 pixels (mean and standard deviation), a 27 ± 21 % improvement on the rigid registration of 7.5 ± 1.4 pixels. This compared to a fit using the previous landmarks of 5.4 ± 1.9 pixels from a rigid registration of 8.3 ± 1.8 pixels. The mean absolute difference between individual cases of the new and old landmarks was 2.0 pixels. This shows that the algorithm is fairly robust to intra-observer variability in the selecting of the landmark points. However, the second set of landmark points achieved a better initial fit, leading to a smaller improvement in the final fit.

The intra-observer variability in the drawing of the manual airway outline was evaluated by redrawing the outline for the training set of 10 images. The segmentation was then compared to the new outline to get an indication of the variability. The mean error for all training cases, using the new manual outlines as reference, was 6.1 ± 1.4 pixels (mean and standard deviation), a 24 ± 18 % improvement on

Case	Initial error (pixels)	Final error (pixels)	Percentage change (%)
11.	6.67	3.10	-53.
12.	5.08	4.14	-19.
13.	7.91	6.18	-22.
14.	12.55	7.70	-39.
15.	7.57	4.26	-44.
16.	6.96	4.50	-35.
17.	3.90	2.83	-27.
18.	5.79	7.90	36.
19.	5.15	3.58	-31.
20.	7.16	6.55	-9.
21.	10.15	6.47	-36.
22.	6.75	4.18	-38.
23.	12.63	9.71	-23.
24.	9.76	5.72	-41.
25.	6.52	4.25	-35.
26.	13.84	13.02	-6.
27.	5.33	3.54	-34.
28.	10.95	8.81	-20.
29.	8.20	6.17	-25.
30.	7.13	6.82	-4.
31.	12.23	11.18	-9.
32.	8.66	6.44	-26.
33.	8.68	5.65	-35.
34.	16.50	12.67	-23.
35.	11.10	7.81	-30.
36.	7.77	8.08	4.
37.	10.41	6.99	-33.
38.	12.25	9.16	-25.
39.	9.88	7.62	-23.
40.	9.80	7.44	-24.
41.	13.26	7.96	-40.

Table 7.3: The mean error after alignment (*initial error*) and SSM optimisation (*final error*). *Percentage change* shows the change after SSM optimisation

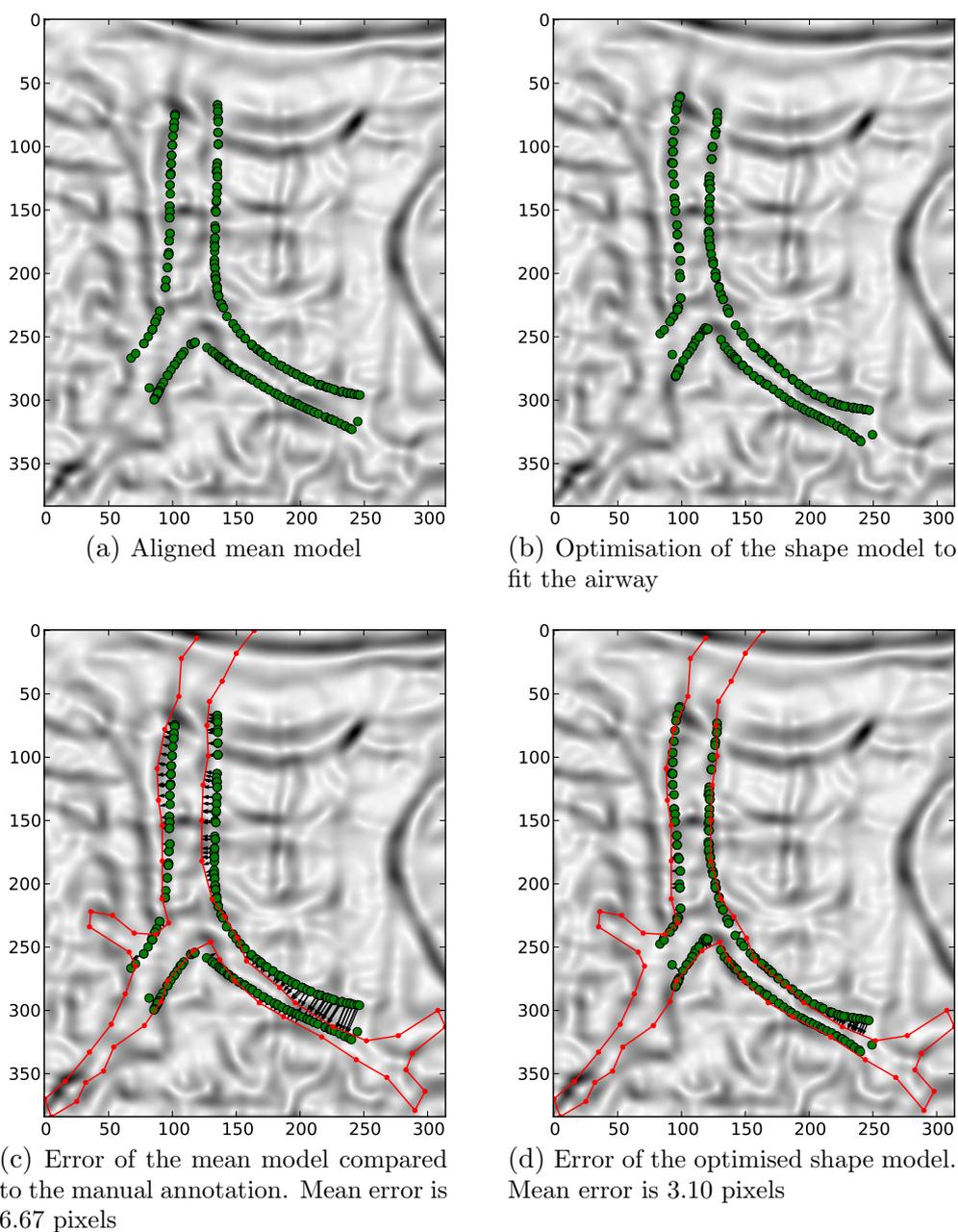


Figure 7.10: Case 11 showing fitting of the aligned shape model to the airway. a) and b) show the silhouette points before and after the shape model fit overlaid on the gradient image. c) and d) show the same fit but with the manual annotation overlaid. Arrows indicate the direction to the closest point on the manual annotation to measure the error.

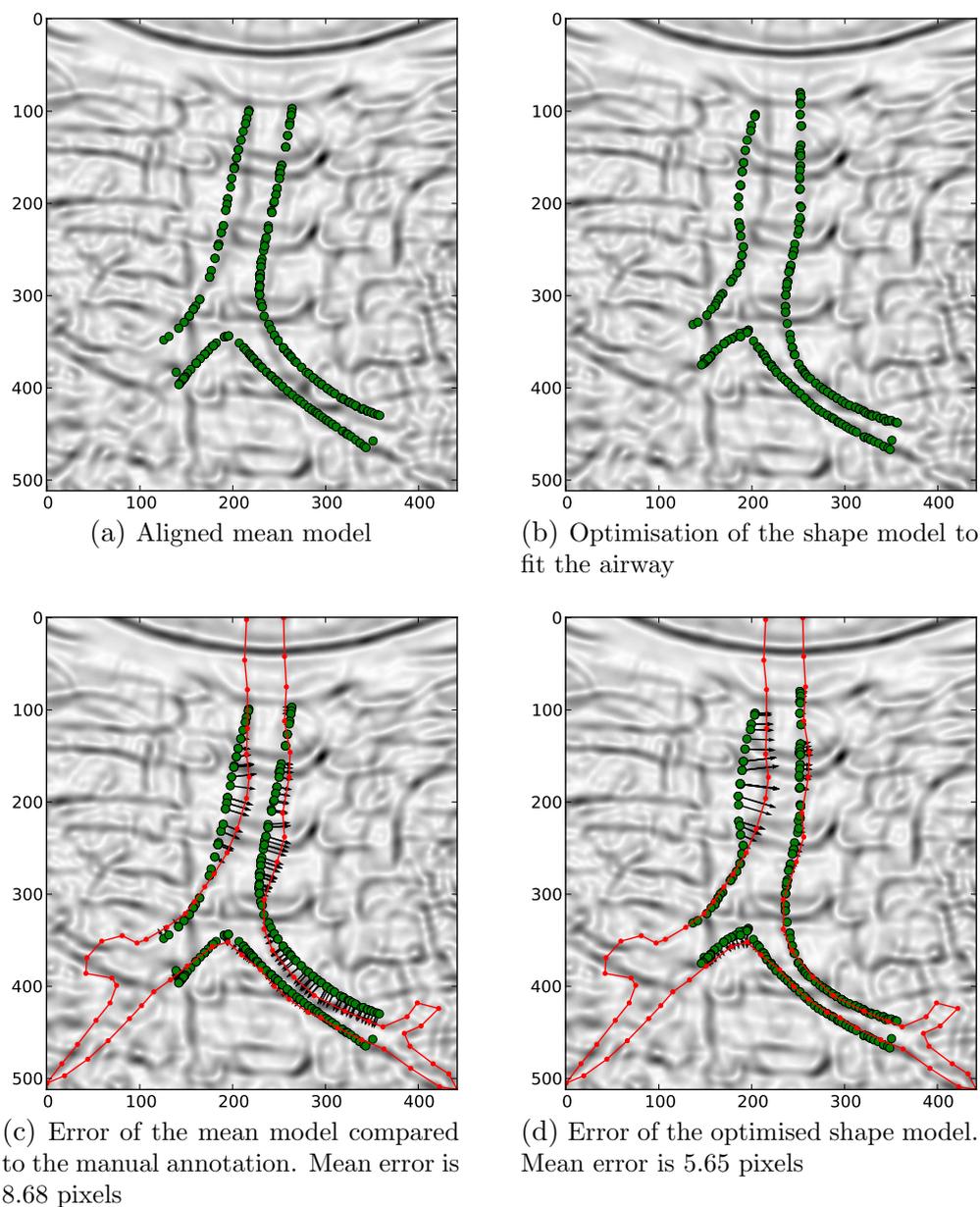


Figure 7.11: Case 33 showing fitting of the aligned shape model to the airway. a) and b) show the silhouette points before and after the shape model fit overlaid on the gradient image. c) and d) show the same fit but with the manual annotation overlaid. Arrows indicate the direction to the closest point on the manual annotation to measure the error.

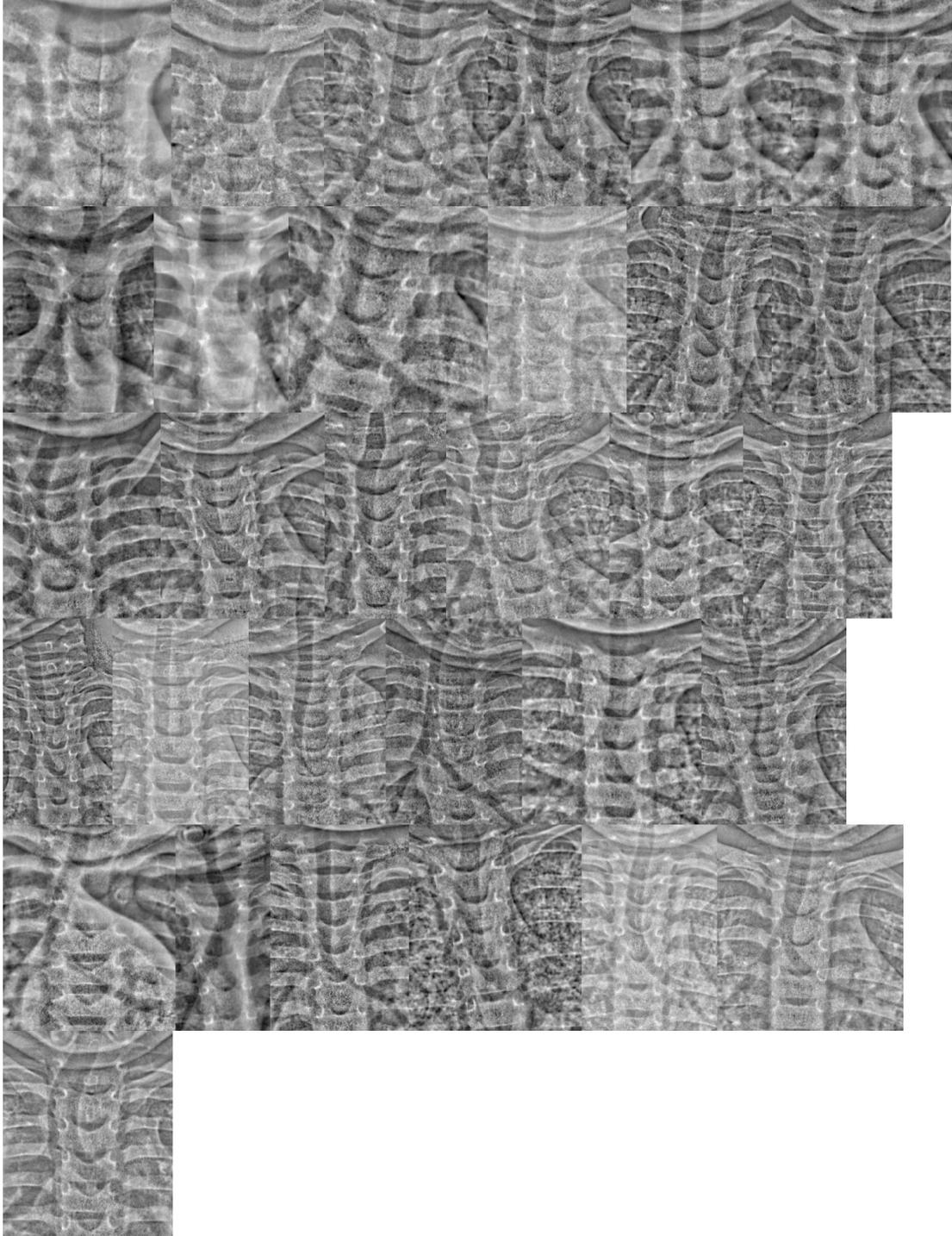


Figure 7.12: Airway region for all test images after local normalisation

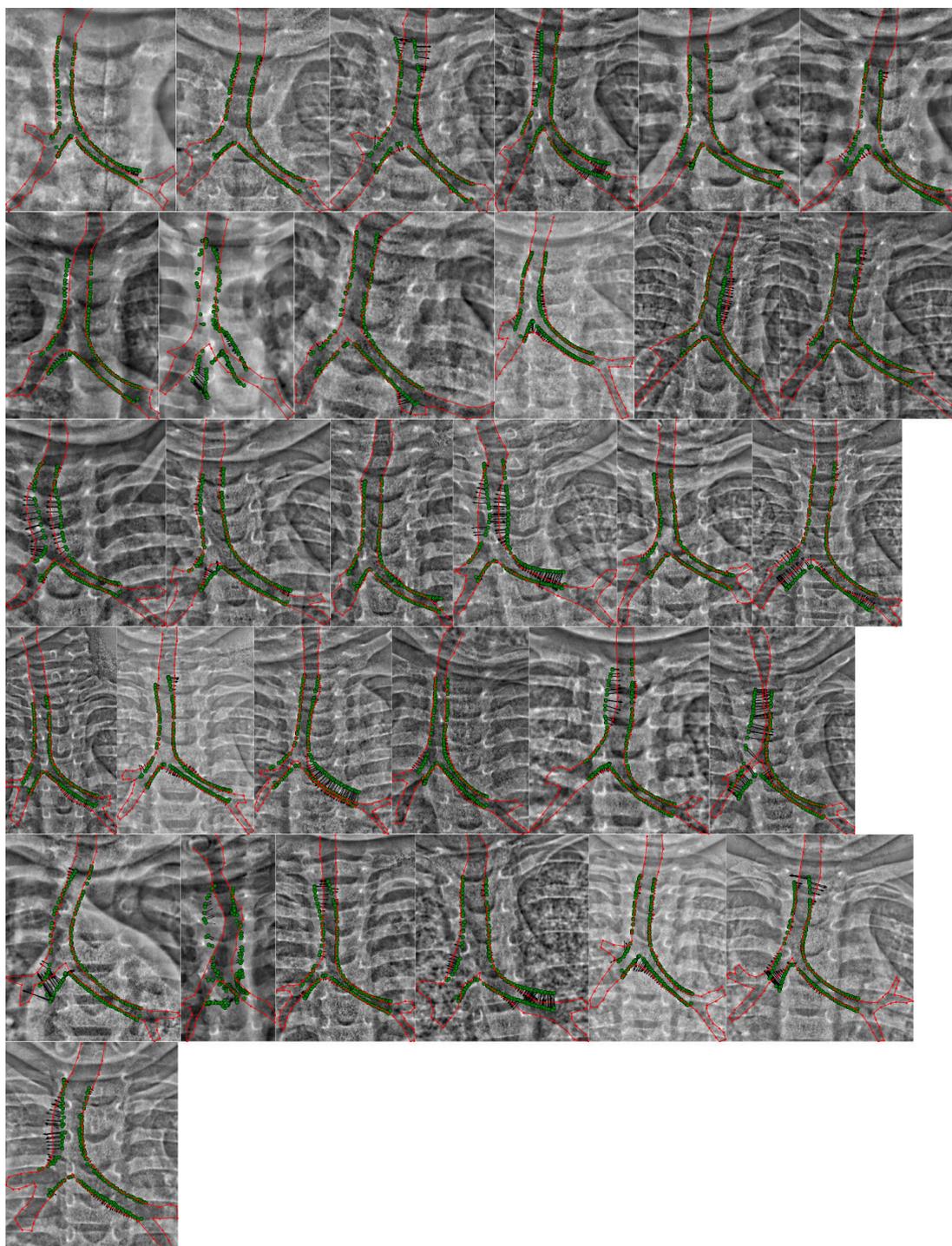


Figure 7.13: Airway region with segmentation outline (green) and manual annotations (red) for all test images

the rigid registration of 8.2 ± 1.8 pixels. The error of the segmentation algorithm when compared to the previous manual outlines was 5.4 ± 1.9 pixels from a rigid registration of 8.3 ± 1.8 pixels. This shows that the intra-observer variability when performing manual reference outlines does not have a considerable impact on the results. However, further evaluation using an experienced clinician is required.

As discussed earlier, there are few airway segmentation methods for radiographs. Tezoo and Douglas (2012) build a active shape model of the trachea and bifurcation region from annotations of 2D chest X-rays. To compare results, the maximum distance between the model and the manual annotations was also calculated for our method. The mean of this distance for all airways in the test set was 23.7 ± 9.9 pixels⁵. The Hausdorff distance could not be calculated because the manual annotations contain all visible airways in each image. However, maximum distance can be considered almost equivalent to the Hausdorff distance if the manual annotation does not have large concavities. Tezoo and Douglas (2012) achieved a mean Hausdorff distance of 18.9 ± 9.9 pixels – a comparable value given the standard deviation. They use 9 landmarks points on the surface of the airway to align the model, while we use 4 landmarks points at branch bifurcation. The additional landmark points may also mean that the initial alignment is closer to the true airway location. Importantly, they only consider the trachea and bifurcation region and not the whole RMB and LMB included in our model. The RMB and LMB plays a more important role in detection of tuberculosis from airway deformation than the trachea.

7.5 Discussion

Segmentation of the airways in chest radiographs is a very challenging topic because of the poorly defined borders and overlapping high contrast bone and heart structures. This chapter applies 3D to 2D shape model projection to segment the airways in radiographs, for the first time, using the 3D statistical shape model that has been developed in previous chapters. Novel airway airway shape based features are applied to the image to assist with the alignment.

The method requires 4 manual landmarks for the initial similarity transformation

⁵From the training set, a higher regularisation weight $r = 0.03$ was identified to obtain lower maximum distances

of the model (rigid transformation with scaling). The statistical shape model improves on this initial alignment by 25 ± 17 % to a mean distance between the model and the manual outline of 6.8 ± 2.6 pixels.

This initial exploration of the 3D to 2D airway segmentation problem only used AP chest radiographs to fit the model (the available data). Including a lateral view is expected to further constrain the optimisation while requiring a weaker regularisation term and, therefore, improve the segmentation. This method can take additional views with minimal modification.

The segmentation procedure was compared to a 2D ASM segmentation method. Our method achieves a similar error rate but can segment a larger region of the airway including the LMB and RMB. This is important for detection of lymphadenopathy.

The segmentation method discussed in this chapter takes advantage of a statistical shape model created from well defined 3D airways in CTs. There are a number of advantages to using a projected 3D shape model over a shape model derived from 2D landmarks in radiographs. The airways are much better defined in CT and, therefore, there is considerably less error in the segmentations used for training – producing a more accurate model. Regions of the airway may not be visible in chest radiographs, which would have to be ignored in a 2D model. Another advantage of the 3D model is that the segmentation can be further improved by adding additional radiographic views of the airways. This method also has the ability to construct 3D personalised airway models, which could be used by clinicians to identify stenosis, without having to expose paediatric patients to high radiation doses associated with CT. A disadvantage of this method over 2D shapes models is that the complexity of the algorithm is considerably increased.

7.6 Conclusion

This chapter introduces a novel approach to the segmentation of the airways in 2D radiographs using a 3D statistical shape model that was developed in previous chapters. The silhouette of the 3D model is detected and projected onto a 2D plane using the specifications of a X-ray machine to define the projection parameters. The parameters of the 3D shape model are then optimised for each radiographic image in the test set using an energy function based on image gradient, greyscale shape kernels

and a regularisation term. The optimisation is performed using the Nelder-Mead downhill simplex algorithm and the shape model is re-projected at each step of the optimisation to recalculate the energy function. The algorithm achieved a 6.8 ± 2.6 pixel error on the test set, a $25 \pm 17\%$ improvement on the initial rigid alignment using 4 manually-identified landmarks.

7.6.1 Future work

As discussed in the previous chapter, the model parameters can be used to classify pathological shape variation in paediatric TB cases. Classification of the airway shape has the potential to be extended to chest radiographs, but may require the lateral view to be included to suitably constrain the 3D model. The model was able to identify the airway in most cases, but did not succeed in improving the initial alignment for 5 of the 30 cases, and further improvements may be required before this method can be used for disease detection

8

Conclusions

8.1 Summary

This thesis presented methods to segment the airways, analyse the structure, detect obstructions, and classify airway pathology, from CT scans. The model is also used to segment the airways in radiographs.

Chapter 1 outlined the clinical motivation behind this thesis. It includes an introduction to paediatric pulmonary tuberculosis and the effect of lymphadenopathy on airway deformation. This motivates the development of a method to detect airway shape variation, which would assist in the detection of paediatric tuberculosis. While this method focuses on paediatric tuberculosis, airway shape analysis has wider applications.

In order to analyse airway shape changes in CT examinations, an airway segmentation method is needed. In **Chapter 2** a morphological airway segmentation method was developed. The lung fields were used to initialise a seed point in the trachea, and from this seed point, the trachea, LMB and RMB were segmented us-

ing a simple threshold and region growing method. In each axial slice, compactness and connectivity were used to monitor segmentation leaks and airway bifurcation, respectively. Morphological filtering using greyscale closing and reconstruction was applied to CT slices in the axial, coronal and sagittal directions, using a range of circular kernels, to enhance cross sections of smaller bronchi. A threshold was applied to extract potential airway locations for each of these circular kernels. The entire airway tree was then segmented by applying a region growing method from the initial segmentation to regions enhanced by the filter. A leak removal step was used to evaluate changes in the cross section of each bronchi and exclude false regions. This method was developed for paediatric airway segmentation, but given the lack of a “gold standard” segmentation, was evaluated as part of the EXACT’09 airway segmentation challenge on a 20 patient adult dataset. The algorithm achieved a fair result compared to other state-of-the-art airway segmentation algorithms. The mean number of branches detected was 43.5% and false positive rate (FPR) was 1.27%. Out of the 15 teams, 4 teams had a lower FPR (better) and 10 teams had a higher FPR (worse). At this FPR point, 9 teams had a higher sensitivity (better) and 5 teams had a lower sensitivity (worse). The algorithm with the highest sensitivity detected 76.5% of the airway with a FPR of 15.6%, demonstrating that there is potential to improve airway segmentation methods.

The next step in analysis, after segmentation of the airways, was to detect the branching structure of the airways and label each branch. In **Chapter 3** an existing airway skeletonisation method was used to extract the branching centreline of the airway tree. A multi-threshold pruning method was then used to remove false branches. The branching structure was used to label the surface mesh of each branch and identify bronchi.

Once the structure of each airway had been identified from the airway segmentation and centreline detection steps, it was necessary to register the surface mesh of each airway to compare variation between airway trees. **Chapter 4** introduced a method to register regions of the airway tree in a dataset of CT scans. Landmarks were projected onto the airway surface, and used to guide the thin-plate-spline warp of a template mesh onto each airway surface. The landmarks were generated by, firstly, detecting the start and end point of the smoothed centreline of each branch, then equidistantly sampling the centreline between these points, and placing points

on the surface of the airways using a reference frame derived from the coordinates of the centreline. A template mesh could then be aligned to each three-branch airway region using a thin-plate-spline warp and surface landmarks. A local matching algorithm was then used to align individual vertices of the template to each case while preserving the structure of the mesh. A number of cases had branch stenosis and it was found to be beneficial to add an inflation/deflation term, as well as a closest point and mesh preservation terms, to perform local alignment. The final alignment was evaluated in terms of the accuracy of the fit and the preservation of the mesh faces. The optimal method was able to register a template mesh to the airway dataset with a mean error of $2.1 \pm 0.9\%$. However, single voxel errors are likely to be caused by the re-voxelisation method used in the evaluation and the larger polygonal face size of the simplified surface mesh. Errors larger than 1 voxel (i.e. errors that could not be explained by the re-voxelisation procedure) contributed to $0.22 \pm 0.49\%$. The local normalisation performed better for airway registration than alternate methods used for other applications, both in terms of registration accuracy and mesh quality.

Due to severe pathology, a number of cases have main bronchi that appeared completely obstructed in the CT scan. These cases produce incomplete airway segmentations, and for visualisation and analysis, it is important to segment beyond the obstructions. In **Chapter 5**, we used both the airway structure defined by the labelled skeleton, and a statistical shape model of each branch of interest, trained on normal branches and missing child bronchi, to detect missing bronchi. The direction of the branch and point of obstruction were then used to search for additional airway regions. The method was evaluated on a limited number of cases with obstruction but showed promising initial results. This method automatically identified 24 of the 26 obstructed branches. Two branches could not be classified because of a lack of data. 18 of 19 disconnected airway regions were segmented.

Chapter 6 builds on the previous chapters to evaluate paediatric airway stenosis and deformation caused by tuberculosis. A statistical shape model for the Trachea-LMB-RMB and RMB-RUL-BI regions of the airway tree was constructed from a dataset of CTs with and without tuberculosis. The modes of variation from principal component analysis were used as features for classification. This method was compared to features derived from measurements of the branch diameter. These methods were evaluated on a validation set of patients from Tygerberg Hospital, South

Africa. The airway analysis method was able to accurately distinguish between TB and non-TB airways by examining regions of the airway likely to be affected by lymphadenopathy, resulting in an AUC of 0.87(0.77 – 0.94) and 0.81(0.68 – 0.90) for the Trachea-LMB-RMB and RMB-RUL-BI regions, respectively. This illustrated the potential of automatic assessment of airway shape in the detection of paediatric pulmonary tuberculosis. Combining the classification of both regions yielded the best results. These PDM derived features performed better than two versions of radius based features. Radius features 1 had an AUC of 0.72(0.59 – 0.83) and 0.59(0.46 – 0.73). Radius 2 had an AUC of 0.83(0.73 – 0.91) and 0.75(0.64 – 0.85). The PDM derived features showed more promise and this was probably because of the ability to represent more complex variation in feature space.

This airway assessment method was applied to radiographs. However, segmentation of the airways in radiographs is a challenging task because of poor visibility and overlapping bone and heart structures. In **Chapter 7** we investigated the use of a 3D statistical shape model of the airways in order to assist in the segmentation of the 2D airway region. Silhouette vertices of the mean airway model were detected and projected onto the 2D radiograph. This projection was initially aligned to the airway using four manual annotations. An energy function was created from the filtered image gradient and greyscale derived features, and used, along with a regularisation term, to fit the 3D shape model to the 2D radiograph. The statistical shape model improved upon this initial alignment by 25 ± 17 %. The mean distance between the model and the manual outline was 6.8 ± 2.6 pixels. The results showed that this method can be used to segment the trachea and main bronchi but could be improved by incorporating additional radiographic views to further constrain the optimisation.

As a research student at UCL, I was given this project to initiate and develop. Thus, I have developed all aspects of this detection system from scratch with the support of my supervisors.

8.2 General remarks and future directions

This thesis presents novel methods for airway analysis and detection of tuberculosis, and has been evaluated on a dataset of paediatric cases with and without tuberculosis. These methods also have potential to be applied to other airway or shape analysis

problems. This section highlights some other applications and future directions of this research.

Methods developed in this thesis could be applied to other diseases affecting the airways, such as congenital cardiac disease, which can cause airway obstruction by compression from enlarged cardiovascular structures. When there is airway involvement in congenital cardiac disease, it is necessary to investigate the cause and severity of involvement. There are also a range of diseases affecting the airways primarily in adults, which include: post-intubation stenosis, amyloidosis, relapsing polychondritis, rhinoscleroma and sarcoidosis. Signs of these diseases could all be detected using this method if training CT datasets of the disease could be acquired. There is also potential for this method to distinguish between different types of pathology by detecting differences in the type of airway deformation, and hence the variation in the shape model, associated with each disease. These methods may also be applied to blood vessels and other tubular structures.

This study investigates the effects of tuberculosis on the airway in paediatric patients and it would be useful to extend the method to investigate other causes of lymphadenopathy. Other causes of lymphadenopathy in paediatric patients are relatively uncommon in TB endemic areas such as South Africa. Therefore, as the diagnosis of paediatric TB often relies on the identification of lymphadenopathy and airway deformation, our aim has been to provide an automated approach to this detection, which would be used in conjunction with other clinical tests to exclude other causes of lymphadenopathy. In a future study it would be interesting to use this method to compare airway stenosis from lymphadenopathy in TB patients to other causes of lymphadenopathy.

The 3D shape model that has been developed for the detection of abnormal airway shape variation was built from a dataset of CT volumes. These volumes provide information about the airway shape at roughly a single point in time. As discussed earlier, the airway changes shape during breathing and movement, and therefore, to improve the specificity of the model, it would be useful to develop individual patient models of the normal shape variability during breathing. There is potential to capture this variability in future using dynamic CT, where a series of volumes are acquired over a period of time. Boiselle et al. (2006) and Baroni et al. (2005) use dynamic CT during breathing cycles or coughing to detect tracheomalacia – characterised by

flaccidity of the airway membranes. However, this method is limited to small volumes by the size of the MDCT detector.

This study uses the 3D airway shape model to segment the airways in 2D radiographs. It would be useful to also detect signs of pathology in these segmentations, because providing automated analysis of the airway shape in radiographs opens up the possibility for this method to be incorporated into screening tools. The 2D segmentation method shows potential but may require improved accuracy for analysis of pathology. A key problem is poor airway visibility in chest radiographs. Performance could be improved by using radiographs with optimised technique factors for better airway visibility, or using additional radiographic views to better constrain the 3D model and improve the fit to 2D projections.

In this thesis, the accuracy of distinguishing TB from non-TB cases is evaluated in regions of the airway by modelling airway stenosis and deformation caused by lymphadenopathy. Given a dataset of CT scans with annotated lymphadenopathy locations, the same model could be trained to identify the location of the offending lymph node. This could be used to provide annotations of lymphadenopathy to the clinician.



Statistical and classification methods

This appendix outlines the generalised procrustes methods (GPA), principal component analysis (PCA) and classification methods used in thesis.

A.1 Generalised procrustes analysis

Generalised procrustes analysis (GPA) can be used to align sets of corresponding landmark points by similarity transform (scaling, translation and rotation) (Cootes and Taylor, 2004). Each set of n points is made up of the vertices of an airway in 3D space i.e. $\mathbf{x} = [(x_1, y_1, z_1), \dots, (x_n, y_n, z_n)]$ and GPA alignment minimises the distance between these landmark points. GPA was incorporated into our method using the following steps:

1. Translate the centroid of each object to the origin
2. Scale each object
3. Choose an example shape as the initial mean

4. Align all objects to mean
5. Recalculate the mean from the aligned shapes
6. Align mean airway to the initial mean
7. Calculate difference between current and previous mean
8. Repeat from (4) if the difference is greater than a defined threshold

The three transforms in this algorithm are described below:

Centroid and translation: The centroid ($C(\mathbf{x})$) can be defined as the mean of all landmark points representing the object. Initially, the centroid is subtracted from the position of each point to translate the object to the origin.

$$C(\mathbf{x}) = (\tilde{x}, \tilde{y}, \tilde{z}) = \sum_{j=1}^n \frac{(x_j, y_j, z_j)}{n} \quad (\text{A.1})$$

Scale: The 2-norm $S(\mathbf{x})$ is used to define the size of an object. Each object is scaled by $S(\mathbf{x})$ to normalise the dataset.

$$S(\mathbf{x}) = \sqrt{\sum_{j=1}^n [(x_j - \tilde{x})^2 + (y_j - \tilde{y})^2 + (z_j - \tilde{z})^2]} \quad (\text{A.2})$$

where $(\tilde{x}, \tilde{y}, \tilde{z})$ is the centroid.

Rotation: The optimal rotation matrix between two sets of landmark points (from two shapes) can be found using singular value decomposition (SVD). The SVD of a matrix A has the form $A = UDV^T$. If \mathbf{x} and \mathbf{y} are two sets of landmark points, then by applying SVD to $\mathbf{x}^T\mathbf{y}$, the optimal rotation matrix can be shown to be VU^T (Stegmann and Gomez, 2002).

It is necessary to align each new mean airway shape to the initial mean (Step 6) otherwise the optimisation becomes ill-conditioned and will not converge (Cootes et al., 1995).

A.2 Principal component analysis

Principal component analysis (PCA) is an integral part of developing a point distribution model (PDM). PCA applies a linear transform that projects the vertices onto an uncorrelated space and can be used to extract relevant features (Cootes et al., 1995). PCA modes are ordered by the variance and, therefore, by selecting a subset of the modes with the most variance, each shape can be represented by a feature vector of lower dimensionality than the input feature vector.

As input into PCA, each object is represented as a $3n$ dimensional stacked vector of mesh vertices where n is the number of vertices in the mesh; for the Trachea-RMB-LMB regions $n \approx 1500$. Therefore, the GPA aligned 3D landmark points of the object $\mathbf{x} = [(x_1, y_1, z_1) \dots (x_n, y_n, z_n)]$ are now represented as a single vector $\mathbf{x}_i = (x_1, y_1, z_1, \dots, x_n, y_n, z_n)^T$ for object i .

PCA is computed from the covariance matrix:

$$\Sigma = \mathbf{X}\mathbf{X}^T \quad (\text{A.3})$$

Each column of \mathbf{X} is the mean subtracted $3n$ vector of each object

$$\mathbf{X} = (\mathbf{x}_1 - \bar{\mathbf{x}} | \mathbf{x}_2 - \bar{\mathbf{x}} | \dots | \mathbf{x}_n - \bar{\mathbf{x}}) \quad (\text{A.4})$$

where $\bar{\mathbf{x}}$ is the mean vector (i.e. the mean shape) in the dataset. It can be shown that the eigenvectors (Φ) of the covariance matrix can be used to project an airway (\mathbf{x}) into uncorrelated space, where the dimensions of the uncorrelated space are defined by the orthogonal eigenvectors. This projection is:

$$\mathbf{b} = \Phi^T(\mathbf{x} - \bar{\mathbf{x}}) \quad (\text{A.5})$$

where the projection matrix (Φ^T) is the transpose of the eigenvector matrix (Φ) and \mathbf{b} is the representation of the airway in the new space. Therefore, an airway can be represented in terms of the mean shape and a displacement along each principal component by \mathbf{b} :

$$\mathbf{x} \approx \bar{\mathbf{x}} + \Phi\mathbf{b} \quad (\text{A.6})$$

Of particular interest is that PCA results in an ordered set of eigenvectors, where the contribution of the variance of each eigenvector (ϕ_i) is represented by the eigenvalues (λ_i). A shape can, therefore, be approximated by a set of m eigenvectors $\Phi = (\phi_1|\phi_2|\dots|\phi_m)$ and $\mathbf{b} = (b_1, \dots, b_m)$. This means that a dense surface mesh model of a shape can be reduced to a much shorter feature vector represented in terms of the principal components of shape variation.

A common way of choosing the number of principal components (eigenvectors) that will represent the shapes is based on the amount of variance they represent. Typical choices range from 90% to 99.5% of the variance (van Ginneken et al., 2002a). The variance represented by the first m eigenvectors can be calculated from the sum of the first m eigenvalues over the sum of all $3n$ eigenvalues:

$$f = \sum_{i=1}^m \lambda_i / \sum_{i=1}^{3n} \lambda_i \quad (\text{A.7})$$

Therefore, each object can be represented by a vector \mathbf{b} of length m . These vectors represent the object shape (with an accuracy of up to the chosen variance f) and can be used to distinguish and classify airway shapes.

A.3 Classification methods

Two classification algorithms were used in this work: Linear discriminant analysis (LDA) and support vector machines (SVM). LDA is a widely used probabilistic classifier. This method uses Bayes' theorem to determine the posterior probability $P(G = k|X = x)$ of a class label $G = k$ given $X = x$. The class density, for a given x , is modelled by a multi-variate Gaussian:

$$f_k(x) = \frac{1}{(2\pi)^{p/2} |\sum_k|^{1/2}} e^{-\frac{1}{2}(x-\mu_k)^T \sum_k^{-1} (x-\mu_k)} \quad (\text{A.8})$$

where \sum_k is the covariance matrix. In LDA, the covariance is assumed to be the same for all classes. This simplifies the method and results in linear decision boundaries.

LDA benefits from its simplicity, low computational cost and low variance, and even with its simplicity it is still one of the top performing classification algorithms (Hastie et al., 2011).

SVM is a powerful discriminative approach, which finds a decision boundary that maximises the distance from the closest data point from any class. For a two class problem, the linear case of a discriminative function can be defined as (Bishop, 2006):

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0 \tag{A.9}$$

where \mathbf{x} is a feature vector and \mathbf{w} is the weighting vector. A vector \mathbf{x} is input and a threshold of the output $f(\mathbf{x})$ defines which class label the sample is given. These weighting factors are chosen during training of the SVM in order to maximise the margin between the decision boundary and samples (in the case of separable classes). The definition is extended for overlapping classes. This turns into a minimisation problem that can be solved by Lagrangian multipliers (Bishop, 2006). This linear method can be made nonlinear using the kernel trick – replacing \mathbf{x} with a nonlinear combination of features $\phi(\mathbf{x})$.

Various classifiers can often yield similar results on general data and feature extraction often plays a more important role. As an example, Hammond et al. (2004) developed a method to detect facial morphology and achieved very similar results using neural networks, support vector machines and logistic regression.

B

Structure analysis: additional examples and evaluation

There is considerable variability in the pathology present in both the non-TB and TB training sets. In the most extreme cases this leads to complete obstruction of bronchi. We require our method to be robust to these variations – in comparison to typical skeletonisation methods that appear to use healthy airways (Mori et al., 2000; Palágyi et al., 2006).

The pruning method (discussed in Section 3.2.4) applies an end branch removal algorithm at multiple thresholds. These thresholds allow larger false branches to be removed from the main tree while preserving smaller branches from the periphery. Figure B.1a shows the airway tree after skeletonisation, Figure B.1b shows the centreline after pruning with a low threshold and figure B.1c shows the centreline after pruning with a high threshold. This example has an obstructed RUL bronchi (leading to a short RUL centreline). The low threshold removes this branch and the high threshold preserves this branch but includes other false branches (Figure B.1c).

Reconstructing a new tree, from both these pruned trees, using branch connectivity, we can apply a high threshold to the trachea and a lower threshold to other branches (Figure B.1d).

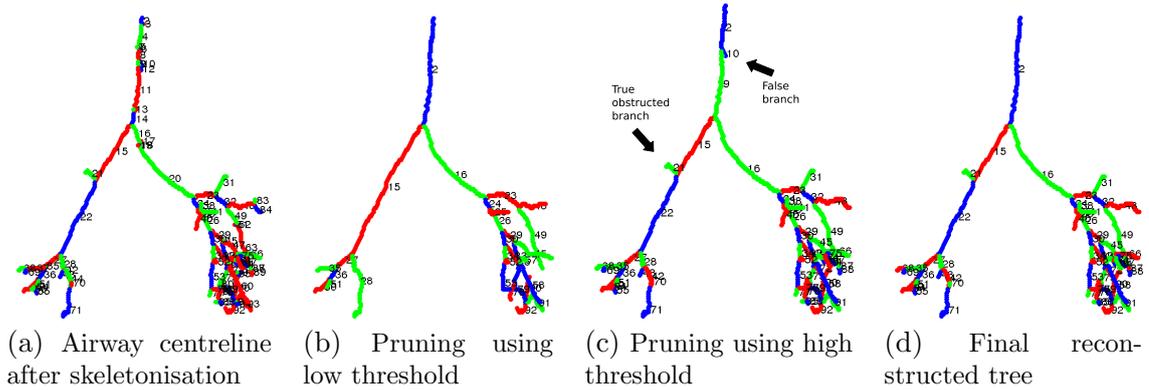


Figure B.1: Example of multithreshold pruning steps

Even with this method some cases require manual interaction (21 of the 90 test cases required some interaction). These cases are almost entirely due to severe pathology. Table B.1 shows the interactions required for these 21 cases. Branches can be kept, removed or branch points can be inserted. This is applied by listing the unique identifier assigned to the branch. Inserting a branch point requires two identifiers (branch and location on the branch).

Figure B.2 shows an example of complete obstruction of the LMB. In this case the pruning method will remove the remaining section of the LMB and manual interaction is required to preserve the branch (see t1 in Table B.1)

Figure B.3 shows a second example of manual interaction. As shown in the Figure, the structure of the LMB region is changed. This requires manual interaction to remove a section of the skeleton (see t8 in Table B.1)

TB dataset	remove	keep	insert 1	insert 2
t1	0	4	0	0
t2	33	0	0	0
t3	0	16	0	0
t4	11	15	0	0
t5	0	3	0	0
t6	6	0	0	0
t7	0	0	3	27
t8	48	0	0	0
t9	0	0	11	28
t10	0	0	2	73

Non-TB dataset	remove	keep	insert 1	insert 2
nt1	0	0	13	17
nt2	33	0	0	0
nt3	4	0	0	0
nt4	44	0	0	0
nt5	1	0	0	0
nt6	26	0	0	0
nt7	0	0	20	36
nt8	7	0	0	0
nt9	0	0	12	17
nt10	0	0	11	16
nt11	14	0	0	0

Table B.1: Manual interaction required for 21 of the 90 test cases using *keep*, *remove* or *insert* to edit skeleton. Each number is the unique label of the branch that requires interaction.

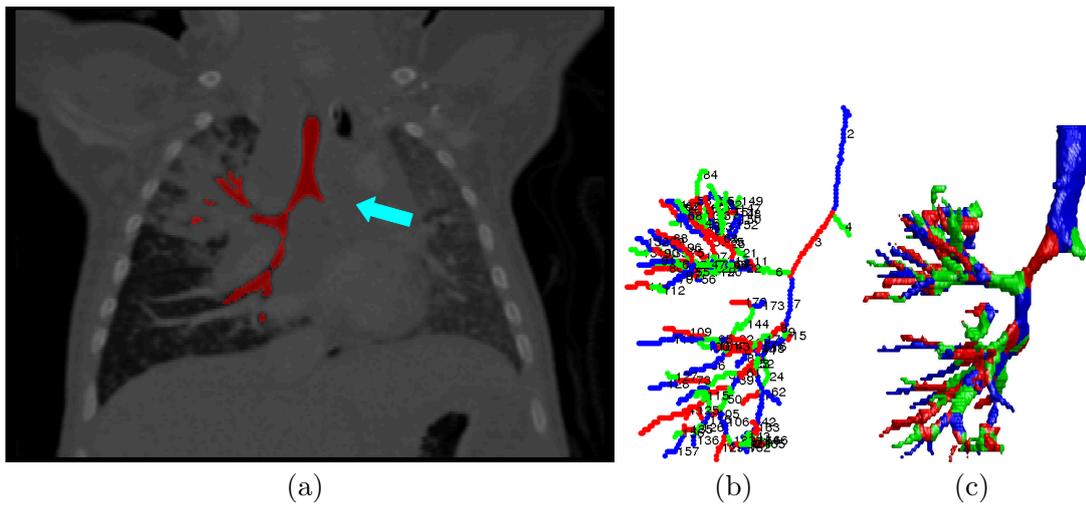
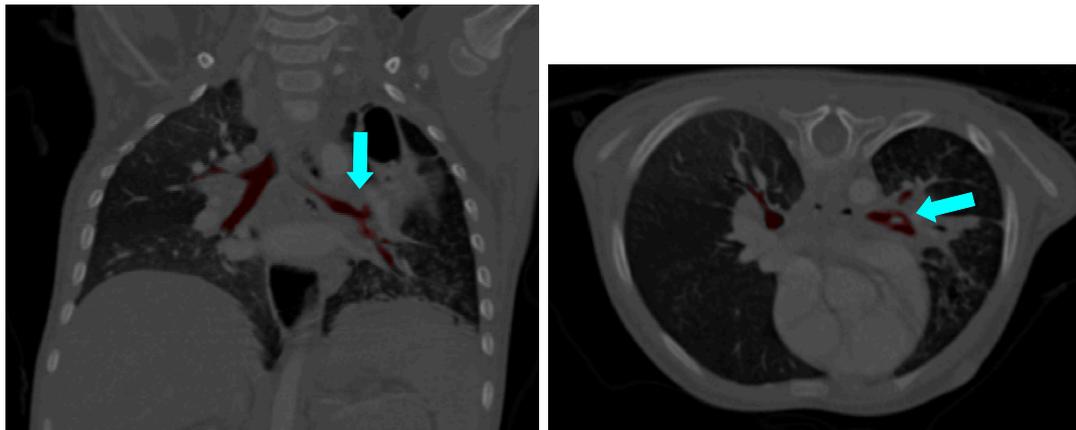
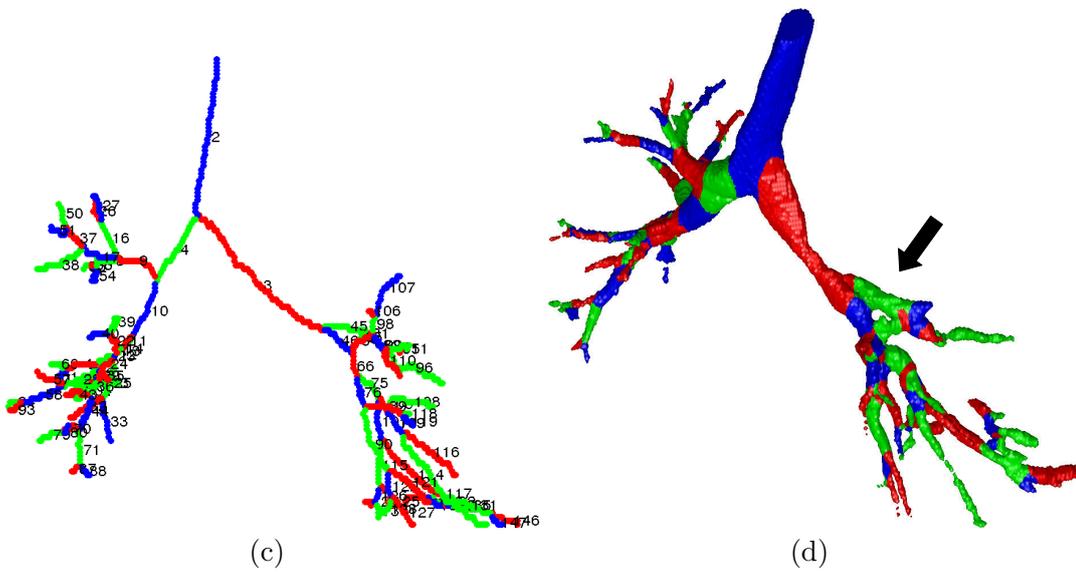


Figure B.2: Example T1



(a)

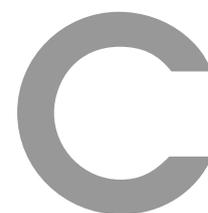
(b)



(c)

(d)

Figure B.3: Example T8



Data from registration analysis

This section contains additional data from the mesh quality analysis that was performed in Chapter 4. Table C.1 and Table C.2 show the mean accuracy of the registration in terms of V_{dif} and V_{open} , for parameters α and β . Table C.3 - C.6 show the optimisation values using all three parameters. A grid search was performed over α , β and γ . Each value on the grid search involves registering a mesh to all cases in the training set and takes approximately 30 minutes.

The optimal fit for the training set was found at $\alpha = 0.1$, $\beta = 1.0$ and $\gamma = 0.8$. Using these parameters, the optimal fit for the each case in the training set is shown in Table C.7. As discussed in Chapter 4, the mean values obtained are $V_{dif} = 0.021 \pm 0.009$ and $V_{open} = 0.0022 \pm 0.0049$.

Figure C.1 shows an example of the mesh quality analysis on a single airway. As expected more noise is present in a single airway mesh analysis than the entire airway tree. Figure C.1c shows a scatter plot of the relationship between q_1 and q_2 .

APPENDIX C. DATA FROM REGISTRATION ANALYSIS

α	$\beta=0.5$	$\beta=1$
0	0.1598	0.1598
0.1	0.0394	0.0452
0.2	0.0321	0.0371
0.3	0.0327	0.0326
0.4	0.029	0.0332
0.5	0.0293	0.0304
0.6	0.0312	0.0321
0.7	0.0305	0.0322
0.8	0.0292	0.0303
0.9	0.0292	0.0268

Table C.1: V_{dif} for closest point and mesh preserving terms (F1, F2)

α	$\beta=0.5$	$\beta=1$
0	0.0925	0.0925
0.1	0.0107	0.0148
0.2	0.008	0.0107
0.3	0.0097	0.0074
0.4	0.0073	0.0099
0.5	0.0077	0.0078
0.6	0.0098	0.0091
0.7	0.0095	0.0103
0.8	0.0084	0.0087
0.9	0.0083	0.0057

Table C.2: V_{open} for closest point and mesh preserving terms (F1, F2)

γ	$\alpha=0$	$\alpha=0.1$	$\alpha=0.2$	$\alpha=0.3$	$\alpha=0.4$	$\alpha=0.5$	$\alpha=0.6$
0	0.1598	0.0394	0.0321	0.0327	0.029	0.0293	0.0312
0.1	0.0357	0.0282	0.0284	0.0303	0.0283	0.0293	0.0305
0.2	0.0292	0.0257	0.0265	0.0274	0.0281	0.0288	0.029
0.3	0.0281	0.0245	0.0256	0.0267	0.0265	0.0272	0.0277
0.4	0.0261	0.0235	0.0253	0.0254	0.0258	0.0268	0.0269
0.5	0.0229	0.0226	0.0242	0.0258	0.0252	0.0266	0.0258
0.6	0.023	0.0225	0.0235	0.0252	0.0252	0.0256	0.026
0.7	0.0229	0.022	0.023	0.0244	0.0246	0.025	0.0254
0.8	0.0217	0.022	0.0225	0.0235	0.0239	0.0246	0.0246
0.9	0.0221	0.024	0.0225	0.0231	0.0243	0.0237	0.0243
1	0.0233	0.0224	0.0224	0.0231	0.0232	0.0232	0.0233
1.1	0.024	0.0236	0.0225	0.0222	0.0227	0.0228	0.0237
1.2	0.0236	0.0213	0.0225	0.022	0.0227	0.0232	0.0242
1.3	0.0223	0.021	0.0221	0.0217	0.0229	0.0232	0.0395
1.4	0.0218	0.0205	0.022	0.0229	0.0237	0.0373	0.1724
1.5	0.0237	0.0236	0.0226	0.0226	0.0351	0.1709	0.3599
1.6	0.0224	0.0209	0.0227	0.0371	0.1694	0.352	0.4371

Table C.3: V_{dif} for entire method (F1, F2, F3) ($\beta = 0.5$)

APPENDIX C. DATA FROM REGISTRATION ANALYSIS

γ	$\alpha=0$	$\alpha=0.1$	$\alpha=0.2$	$\alpha=0.3$	$\alpha=0.4$	$\alpha=0.5$	$\alpha=0.6$
0	0.0925	0.0107	0.008	0.0097	0.0073	0.0077	0.0098
0.1	0.0085	0.0058	0.0064	0.0087	0.0068	0.008	0.0094
0.2	0.0075	0.0043	0.0052	0.006	0.0068	0.0077	0.008
0.3	0.0078	0.0037	0.0048	0.0058	0.0057	0.0065	0.007
0.4	0.0061	0.0032	0.0043	0.0049	0.0052	0.0059	0.0064
0.5	0.0036	0.0028	0.0038	0.005	0.005	0.0063	0.0057
0.6	0.0036	0.003	0.0034	0.0046	0.0047	0.005	0.0054
0.7	0.0035	0.0027	0.0031	0.0041	0.0044	0.0046	0.005
0.8	0.0028	0.0028	0.003	0.0036	0.0041	0.0045	0.0046
0.9	0.0031	0.0049	0.0029	0.0034	0.0044	0.0039	0.0043
1	0.0043	0.0034	0.0028	0.0036	0.0034	0.0034	0.0035
1.1	0.005	0.0048	0.0032	0.0029	0.0031	0.0029	0.0037
1.2	0.005	0.0026	0.0032	0.0027	0.0033	0.0033	0.0041
1.3	0.0037	0.0022	0.0029	0.0025	0.0033	0.0034	0.0106
1.4	0.0033	0.0019	0.0029	0.0036	0.0039	0.0101	0.0785
1.5	0.0052	0.0048	0.0035	0.0032	0.0075	0.0766	0.241
1.6	0.0039	0.002	0.0031	0.0083	0.0765	0.2337	0.3182

Table C.4: V_{open} for entire method (F1, F2, F3) ($\beta = 0.5$)

γ	$\alpha=0$	$\alpha=0.1$	$\alpha=0.2$	$\alpha=0.3$	$\alpha=0.4$	$\alpha=0.5$	$\alpha=0.6$
0	0.1598	0.0452	0.0371	0.0326	0.0332	0.0304	0.0321
0.1	0.0379	0.0358	0.0332	0.03	0.0292	0.0292	0.0293
0.2	0.0304	0.0317	0.0287	0.0265	0.0288	0.0277	0.0277
0.3	0.0307	0.0277	0.0257	0.0274	0.027	0.027	0.0275
0.4	0.0272	0.0259	0.025	0.0256	0.0268	0.0269	0.0268
0.5	0.0258	0.0252	0.0248	0.025	0.0248	0.0262	0.0253
0.6	0.024	0.0227	0.0235	0.0253	0.0252	0.0257	0.0253
0.7	0.0241	0.0244	0.0236	0.0242	0.0243	0.0248	0.0246
0.8	0.0224	0.0213	0.0224	0.0232	0.0236	0.0241	0.0243
0.9	0.0247	0.0213	0.0225	0.0231	0.0234	0.0243	0.0245
1	0.0251	0.0239	0.0226	0.0232	0.023	0.0238	0.0268
1.1	0.0228	0.0212	0.0227	0.0234	0.0241	0.0251	0.0419
1.2	0.0246	0.022	0.0212	0.023	0.0258	0.0419	0.0666
1.3	0.0233	0.0204	0.0219	0.023	0.0379	0.066	0.1581
1.4	0.0237	0.022	0.0239	0.0326	0.0644	0.1572	0.3941
1.5	0.0224	0.022	0.0333	0.0573	0.1521	0.3819	0.542
1.6	0.0242	0.0332	0.0536	0.1503	0.3616	0.4993	0.6135

Table C.5: V_{dif} for entire method (F1, F2, F3) ($\beta = 1.0$)

APPENDIX C. DATA FROM REGISTRATION ANALYSIS

γ	$\alpha=0$	$\alpha=0.1$	$\alpha=0.2$	$\alpha=0.3$	$\alpha=0.4$	$\alpha=0.5$	$\alpha=0.6$
0	0.0925	0.0148	0.0107	0.0074	0.0099	0.0078	0.0091
0.1	0.009	0.0116	0.0093	0.0073	0.0065	0.0065	0.0068
0.2	0.0075	0.0095	0.0056	0.0047	0.0061	0.006	0.0062
0.3	0.0092	0.0067	0.0042	0.0055	0.0055	0.0057	0.0061
0.4	0.006	0.0053	0.004	0.0049	0.0054	0.0057	0.0057
0.5	0.0055	0.0052	0.0033	0.004	0.0043	0.0055	0.0045
0.6	0.0043	0.003	0.0032	0.0045	0.0043	0.0047	0.0042
0.7	0.0045	0.0052	0.0036	0.0037	0.0037	0.0041	0.0038
0.8	0.0028	0.0022	0.0028	0.003	0.0034	0.0035	0.0038
0.9	0.0053	0.0025	0.0028	0.0031	0.0032	0.0041	0.0041
1	0.0062	0.0049	0.0031	0.0036	0.0028	0.0031	0.0059
1.1	0.0041	0.0025	0.0033	0.0036	0.0041	0.0044	0.0183
1.2	0.0057	0.0033	0.0022	0.0033	0.0052	0.0183	0.032
1.3	0.0047	0.0018	0.0028	0.0032	0.0145	0.0326	0.0863
1.4	0.005	0.0025	0.0043	0.0108	0.0315	0.0816	0.2782
1.5	0.0031	0.0028	0.0108	0.0253	0.0766	0.2627	0.4222
1.6	0.0047	0.0119	0.0222	0.0763	0.2434	0.3797	0.4898

Table C.6: V_{open} for entire method (F1, F2, F3) ($\beta = 1.0$)

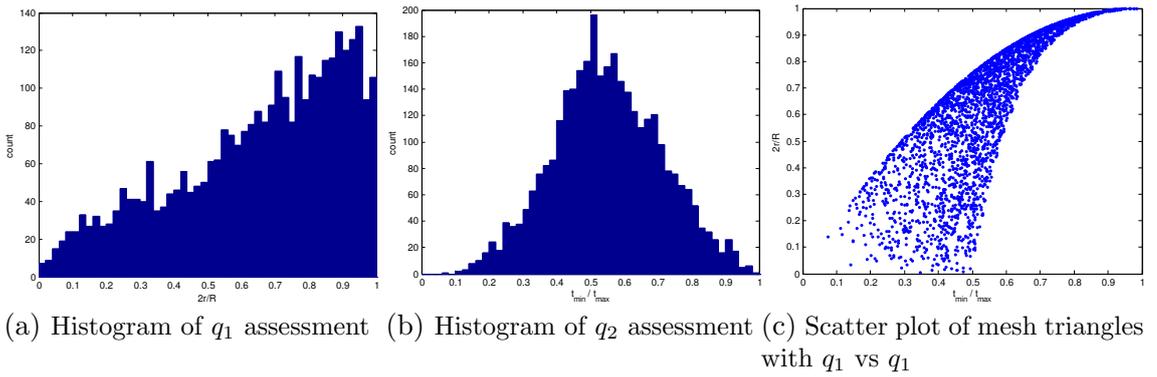


Figure C.1: Example quality assessment for a single airway

APPENDIX C. DATA FROM REGISTRATION ANALYSIS

Case	V_{dif}	V_{open}	Case	V_{dif}	V_{open}
C 1	0.05721	0.02620	C 41	0.02352	0.00004
C 2	0.00000	0.00000	C 42	0.02391	0.00000
C 3	0.01829	0.00220	C 43	0.01947	0.00003
C 4	0.01678	0.00000	C 44	0.01848	0.00000
C 5	0.02131	0.00041	C 45	0.02430	0.00227
C 6	0.03282	0.00789	C 46	0.02683	0.00000
C 7	0.01659	0.00000	C 47	0.02277	0.00141
C 8	0.02020	0.00000	C 48	0.02050	0.00404
C 9	0.00559	0.00000	C 49	0.02087	0.00000
C 10	0.02725	0.01051	C 50	0.02314	0.00000
C 11	0.02516	0.00085	C 51	0.03161	0.00483
C 12	0.01747	0.00000	C 52	0.03071	0.00857
C 13	0.02351	0.00477	C 53	0.01929	0.00000
C 14	0.00419	0.00000	C 54	0.02536	0.00000
C 15	0.01656	0.00003	C 55	0.02827	0.00006
C 16	0.00526	0.00000	C 56	0.02162	0.00071
C 17	0.01492	0.00012	C 57	0.00495	0.00000
C 18	0.01729	0.00014	C 58	0.01953	0.00000
C 19	0.01667	0.00000	C 59	0.01841	0.00028
C 20	0.02677	0.00546	C 60	0.02395	0.00011
C 21	0.01941	0.00002	C 61	0.02089	0.00010
C 22	0.02319	0.00467	C 62	0.02089	0.00042
C 23	0.02485	0.00604	C 63	0.02329	0.00000
C 24	0.01395	0.00002	C 64	0.02051	0.00000
C 25	0.03231	0.00000	C 65	0.04746	0.00726
C 26	0.01925	0.00040	C 66	0.02198	0.00000
C 27	0.00416	0.00009	C 67	0.02054	0.00471
C 28	0.01872	0.00000	C 68	0.01963	0.00041
C 29	0.02953	0.00804	C 69	0.01985	0.00002
C 30	0.01844	0.00007	C 70	0.02686	0.00071
C 31	0.00521	0.00015	C 71	0.04305	0.01952
C 32	0.01666	0.00000	C 72	0.02730	0.00054
C 33	0.01964	0.00466	C 73	0.02290	0.00000
C 34	0.02132	0.00071	C 74	0.01949	0.00111
C 35	0.00370	0.00000	C 75	0.02670	0.00000
C 36	0.01872	0.00029	C 76	0.00997	0.00040
C 37	0.01699	0.00257	C 77	0.02234	0.00000
C 38	0.02114	0.00132	C 78	0.02170	0.00000
C 39	0.02276	0.00083			
C 40	0.04192	0.02426			

Table C.7: Accuracy of the registered template mesh compared to the original mesh for each training case in the trachea-LMB-RMB region. Cases with obstruction are automatically excluded (discussed in more detail in Chapter 6).

D

Shape model features

Table D.1 and Table D.2 show the shape model derived features for the trachea-LMB-RMB and RMB-RUL-BI regions respectively. Each case is represented by 11 features (columns) and rows represent each case in the dataset.

If an obstructed branch is detected in the region of interest it is automatically removed from the feature based classification and processed separately. Therefore, the trachea-LMB-RMB regions consists of 160 cases (78 training and 82 test) with 19 cases (11 training and 8 test) automatically identified as having obstruction and not included. The RMB-RUL-BI region consists of 137 cases (64 training and 73 test) with 42 cases identified as having obstruction (23 training and 17 test).

APPENDIX D. SHAPE MODEL FEATURES

Table D.1: Principal components used for classification of the Trachea-RMB-LMB region. Rows are cases, columns are the 11 components used for classification. Gosh non-TB training set: C1-C37, Tygerberg TB training set: C38-C78, Tygerberg TB test set: C79-C114, and Tygerberg non-TB test set: C115-C160.

C 1	-0.04657	0.01323	0.05418	-0.04815	-0.04048	-0.00193	-0.01122	-0.02242	-0.01590	0.00871	-0.00329
C 2	0.03466	-0.07251	0.00189	-0.01655	-0.00444	0.01586	-0.01832	0.00710	-0.00047	-0.00585	-0.00800
C 3	-0.19104	-0.03292	0.02370	-0.03388	0.00209	-0.04407	-0.00878	-0.01070	0.00324	-0.00791	0.01366
C 4	-0.01770	-0.01839	-0.00540	-0.04704	-0.00974	0.00475	-0.01206	0.01883	-0.00834	0.02436	-0.00274
C 5	-0.05252	-0.06918	0.00675	-0.05244	0.01291	-0.03676	0.00221	0.02244	0.00448	-0.00990	-0.00911
C 6	0.01259	-0.14375	-0.02956	-0.02410	-0.04820	0.00041	0.00516	0.04674	-0.00483	-0.02073	-0.01062
C 7	-0.03197	-0.13428	-0.03608	0.01797	-0.03815	-0.00684	-0.00173	-0.01743	-0.02607	0.01026	-0.00141
C 8	0.03440	0.02990	-0.01909	-0.03143	-0.03480	0.02546	0.00673	0.02605	0.00130	-0.01302	-0.01717
C 9	-0.07890	-0.11889	-0.01773	0.00057	0.00760	-0.03592	-0.01333	-0.02331	-0.00425	-0.01647	0.01744
C 10	-0.00658	0.00948	0.01095	-0.01386	-0.03106	0.04407	-0.00357	-0.01955	-0.00233	0.00478	-0.00256
C 11	-0.09196	0.08446	0.00484	-0.02593	-0.01693	0.03301	0.02919	-0.01363	-0.01595	-0.02466	-0.01712
C 12	-0.16528	-0.06508	-0.02427	0.01913	-0.00476	0.01040	0.03118	0.02617	-0.02026	0.01880	-0.02108
C 13	-0.19482	0.05022	0.07578	-0.02338	-0.01605	-0.00731	0.02305	-0.06112	0.03065	-0.02972	-0.01955
C 14	0.08441	0.01320	0.00216	-0.02753	-0.00579	0.00845	0.00493	-0.00734	0.02024	-0.03154	0.00211
C 15	-0.07061	0.02669	-0.02829	-0.03823	-0.04761	0.00668	-0.01632	0.02161	-0.00083	0.01384	0.00659
C 16	0.01846	-0.10590	-0.01632	-0.02024	0.00270	0.01512	0.00701	-0.01941	0.00525	-0.00185	0.02996
C 17	-0.01270	-0.00254	-0.00079	-0.02676	-0.01575	0.02712	0.00855	0.00358	-0.00520	-0.00846	-0.00986
C 18	-0.02909	-0.11746	-0.00379	-0.02239	-0.03446	-0.02379	0.01689	-0.01178	-0.00977	0.01643	0.01448
C 19	-0.11031	-0.03864	0.00146	-0.00718	-0.02793	0.02730	-0.00292	-0.00147	-0.00761	0.00094	0.01589
C 20	0.00541	0.01398	-0.03090	-0.02923	-0.00322	0.01279	-0.00790	0.02845	-0.00544	-0.00390	-0.00422
C 21	-0.18604	-0.06840	-0.03600	0.00144	0.00505	-0.02804	-0.02113	0.03332	0.00714	0.00704	0.02299
C 22	-0.04920	-0.05555	0.01287	0.00478	-0.02519	0.00524	-0.00393	-0.00456	-0.01475	-0.00617	0.01179
C 23	-0.09930	-0.06349	0.01187	-0.01320	-0.01746	-0.00073	0.01100	0.01276	-0.01620	-0.00429	-0.02679
C 24	-0.06194	-0.00846	-0.05597	-0.07390	0.00645	0.03824	0.00384	0.01173	-0.00586	0.02699	-0.00980
C 25	-0.16393	-0.06260	-0.03013	0.03224	-0.00723	0.01217	0.01643	0.02244	0.02886	0.01326	0.00321
C 26	-0.01048	0.02373	-0.01355	-0.00191	0.00046	0.01937	-0.00891	0.01707	-0.00684	0.01864	-0.00707
C 27	0.00182	0.00021	-0.02989	-0.07527	-0.04194	0.01897	0.01791	-0.02065	-0.00241	-0.01912	-0.00019

Continued on next page

APPENDIX D. SHAPE MODEL FEATURES

Table D.1 – *Continued from previous page*

C 28	-0.03561	-0.03519	0.00235	-0.00607	-0.03233	0.01430	0.01815	0.00000	-0.01600	0.00397	0.00822
C 29	-0.02346	-0.12430	0.02581	-0.00732	-0.03005	-0.01018	-0.00865	-0.01190	0.00123	-0.02091	-0.01179
C 30	0.05895	-0.02944	-0.00587	-0.01127	-0.01032	0.05104	-0.00512	-0.00876	0.00164	-0.00979	0.00169
C 31	-0.09161	0.05020	0.03327	-0.02079	-0.02051	-0.01040	0.00631	-0.03004	-0.01314	-0.00807	-0.01069
C 32	-0.05913	-0.07156	0.03269	0.01954	-0.02940	-0.01697	-0.00299	-0.00038	0.01236	-0.01652	0.01419
C 33	-0.05040	-0.08696	0.03282	0.03343	-0.02722	0.02646	-0.00196	-0.00909	-0.00589	0.00239	0.02370
C 34	0.06522	0.06479	-0.01063	-0.02047	-0.02375	-0.02205	0.03976	0.01106	0.00820	0.01854	-0.00629
C 35	-0.02735	-0.07827	-0.02556	-0.02316	-0.04097	0.01217	0.00664	-0.01245	-0.01583	-0.00708	0.00960
C 36	-0.07979	-0.07675	-0.07824	-0.01942	-0.06579	0.00641	0.03182	-0.01251	-0.00883	0.03851	0.02242
C 37	-0.06751	-0.04200	0.00930	-0.05149	-0.05580	0.01935	0.03397	-0.01983	0.03877	0.00492	0.02025
C 38	0.04444	0.03698	0.02991	-0.00787	-0.02952	-0.01958	-0.01649	0.01700	0.01072	0.00486	-0.00062
C 39	0.08397	0.00605	0.00882	-0.03484	0.04952	-0.01973	-0.01201	0.01299	-0.02895	-0.00414	-0.01733
C 40	0.14723	-0.10949	-0.02567	-0.00125	0.04350	0.04008	-0.00393	0.01519	0.00112	-0.00529	0.01539
C 41	0.08844	-0.02469	0.03709	0.01259	-0.01043	0.00501	-0.00465	0.02050	-0.00427	0.02801	-0.00156
C 42	0.05444	-0.05215	0.04395	-0.04666	0.02067	-0.00517	0.04239	0.00672	0.01498	0.03888	-0.00970
C 43	0.01101	0.01338	0.06806	0.01357	0.00223	0.01480	0.01065	0.00549	0.00444	0.01223	0.00501
C 44	0.05809	0.09042	-0.02274	0.00130	-0.00634	-0.01606	0.00593	0.00825	0.00082	-0.02213	0.00551
C 45	0.04643	0.02175	0.02107	-0.02549	0.00994	-0.00272	-0.02139	-0.01630	0.00848	0.04349	-0.00606
C 46	0.06380	0.06264	-0.03329	-0.02361	-0.00763	-0.01027	0.00900	-0.01731	0.00961	-0.00197	0.02543
C 47	0.05156	0.05010	-0.02892	-0.02019	0.01740	0.00023	-0.00798	-0.02915	0.02361	0.00412	0.01298
C 48	-0.00519	-0.00028	0.05192	0.01903	0.00596	0.00597	-0.02008	0.00530	0.01563	0.01905	-0.00682
C 49	0.01806	0.10202	-0.00817	0.00592	-0.02597	-0.00026	-0.01887	-0.00087	-0.02026	0.00298	-0.01234
C 50	0.00648	0.00271	-0.04182	0.02725	0.01772	0.03572	0.00243	0.02787	-0.01379	0.00848	-0.01168
C 51	0.09886	0.01141	-0.03761	-0.00025	0.05076	-0.00270	-0.01900	-0.03539	0.00657	0.00796	0.00929
C 52	-0.02764	0.08985	0.00934	-0.00944	0.05364	0.01175	-0.02742	0.00320	-0.00677	-0.02082	0.01902
C 53	0.11147	-0.00660	0.00246	0.01359	-0.00743	-0.00346	0.01791	0.00866	0.00091	0.00440	0.01858
C 54	0.13276	0.01959	0.00271	-0.00385	-0.01784	-0.02747	0.01178	0.01507	-0.00294	0.00017	0.00566
C 55	0.04053	0.08718	0.03648	-0.00054	0.00662	0.03398	-0.00090	0.01424	-0.02441	-0.00347	0.00433
C 56	-0.03414	-0.07486	0.01921	-0.02909	0.02155	-0.00542	-0.00765	0.01386	0.00414	0.00941	-0.01079
C 57	0.18815	-0.05575	-0.02235	-0.03768	-0.02799	-0.03789	0.03547	-0.00827	0.00198	-0.00442	0.02726
C 58	0.03662	0.01956	-0.04906	0.02301	0.01279	0.00962	-0.00145	0.00063	-0.03519	0.00089	0.00096

Continued on next page

APPENDIX D. SHAPE MODEL FEATURES

Table D.1 – *Continued from previous page*

C 59	0.06510	-0.05679	-0.06419	-0.01470	0.02664	0.02048	-0.01072	0.00235	-0.02473	-0.00596	0.00027
C 60	0.05439	0.08651	0.02726	0.01045	0.03357	0.02912	0.02235	-0.01566	-0.01848	-0.01434	0.00676
C 61	-0.05850	0.03085	0.00674	0.01236	0.00454	0.02098	0.01678	-0.02461	0.01291	0.00912	-0.01052
C 62	0.03596	-0.02553	-0.02865	0.03841	-0.01674	-0.00079	-0.01307	0.01463	0.00296	0.02693	-0.00047
C 63	0.09397	-0.02730	0.03321	0.03811	-0.02673	0.02795	0.03234	-0.04924	0.02732	0.00007	0.00172
C 64	0.05176	-0.04429	-0.00238	0.04337	0.01546	0.01383	0.02686	0.01969	-0.00258	0.01172	0.01107
C 65	0.09425	-0.06044	0.08763	0.00701	-0.02803	-0.02484	0.03604	0.01090	-0.02137	-0.04032	0.00791
C 66	0.04415	0.07981	0.01453	0.00806	-0.01748	0.00458	0.01092	0.02044	0.01082	0.01807	0.00095
C 67	-0.01190	0.01273	0.00120	-0.01840	0.01929	0.04607	0.00298	-0.00378	-0.00916	0.01354	0.01091
C 68	0.05722	0.07836	0.01560	-0.02110	0.02223	0.00021	-0.02280	-0.01943	-0.01733	-0.01229	-0.00998
C 69	-0.10444	0.04964	-0.01503	-0.02786	0.05555	-0.02137	-0.00790	-0.02707	0.01231	0.00749	0.01590
C 70	0.09928	-0.03934	0.04214	0.00926	0.03486	0.00049	0.03119	-0.00004	-0.00120	0.01918	0.01305
C 71	-0.06708	-0.02379	-0.01650	0.02669	0.01917	-0.02775	-0.02223	-0.00003	-0.02435	-0.00924	-0.01102
C 72	0.12614	0.04783	0.01044	-0.00797	-0.01812	0.00555	-0.01668	0.01149	0.01577	0.00226	0.00262
C 73	0.04588	0.00511	-0.00831	-0.05818	-0.00263	0.01303	0.00326	-0.01463	0.01449	-0.00415	0.00108
C 74	-0.08721	0.03097	0.03130	0.03333	0.02470	0.05932	-0.01697	-0.03008	-0.00215	-0.00050	0.01143
C 75	0.07790	-0.03827	-0.04868	-0.01857	0.02117	0.02178	-0.02089	0.02713	-0.03110	0.00714	-0.01183
C 76	0.11565	0.02429	0.00176	-0.04348	-0.04122	-0.01194	-0.00095	0.00524	-0.00746	-0.00678	0.00731
C 77	0.08573	0.15048	-0.09975	0.02872	-0.02164	0.01047	0.02489	0.02170	0.04842	-0.02197	0.00331
C 78	-0.01587	0.06338	0.02082	-0.00996	0.01996	0.00385	-0.01007	-0.02704	-0.00353	0.04495	-0.01832
C 79	0.06643	-0.05878	-0.01301	-0.03036	0.00742	-0.04226	-0.00412	-0.03039	0.00248	-0.00664	-0.00211
C 80	0.03086	0.07824	0.01101	0.02096	-0.00835	0.01617	-0.01187	-0.01671	-0.00692	0.00054	0.00558
C 81	-0.04748	0.01826	0.05551	-0.00411	0.03568	0.04342	-0.00958	0.00571	0.02052	0.00405	-0.00038
C 82	-0.01889	0.02775	0.02090	-0.00930	0.02963	-0.05253	0.02758	-0.01658	-0.00555	-0.00803	0.02437
C 83	0.05149	-0.00594	-0.04637	0.01935	0.06434	-0.00607	0.03458	0.01210	0.01554	0.00808	0.01048
C 84	-0.02731	0.03100	0.04414	0.01987	0.01685	-0.00941	-0.01628	-0.01765	-0.00937	0.00279	-0.00842
C 85	0.00776	-0.01193	0.09021	0.03567	0.02399	-0.02168	0.05700	0.01598	0.00321	0.01051	-0.00898
C 86	-0.01290	-0.09688	0.00829	0.04525	-0.00263	0.02601	0.01375	-0.00492	0.00859	-0.00608	-0.01423
C 87	-0.00609	0.05358	-0.10398	-0.03975	0.03934	0.00142	0.03518	-0.00617	0.00503	-0.00133	0.00207
C 88	-0.06374	0.10846	-0.02917	0.00059	-0.02457	-0.00582	-0.00903	0.00438	0.00585	0.01800	-0.00345
C 89	-0.05365	0.07837	0.03671	0.01798	0.01624	0.02082	-0.00720	0.02384	0.00297	-0.02730	0.00641

Continued on next page

APPENDIX D. SHAPE MODEL FEATURES

Table D.1 – *Continued from previous page*

C 90	0.02381	0.06036	0.00817	-0.00455	0.01747	-0.01733	-0.03584	-0.02704	0.00267	0.01285	-0.00268
C 91	-0.03072	0.00367	0.01409	0.04760	-0.00241	-0.01598	-0.02151	-0.04454	0.01641	0.03257	-0.01471
C 92	0.15941	0.00751	0.02044	-0.02136	-0.03586	-0.01107	-0.02160	-0.00190	0.02541	0.01161	-0.02544
C 93	0.00251	0.01406	0.00710	-0.00123	0.00587	-0.00149	0.01913	0.02739	-0.01451	0.00256	-0.00074
C 94	0.10860	0.03859	-0.01595	-0.03240	0.01189	-0.01383	-0.02107	0.00314	-0.01927	0.02147	0.00515
C 95	0.00327	0.02356	0.01814	-0.02660	-0.00637	0.00319	-0.03453	0.00928	0.01995	-0.00117	0.01326
C 96	-0.07284	0.05345	0.03307	0.02328	0.00499	0.01534	-0.02540	-0.00682	-0.00227	-0.00405	0.00257
C 97	0.05668	0.09075	-0.00346	-0.01334	0.00254	0.01242	-0.00641	-0.00486	-0.00567	-0.01922	-0.01716
C 98	-0.08231	0.05619	0.02715	-0.01201	0.03183	0.00076	-0.01252	-0.01241	-0.03106	-0.00748	-0.00766
C 99	0.03580	0.01303	-0.00929	0.05223	0.00379	0.02607	-0.00416	-0.01697	-0.01282	0.00936	0.00357
C 100	0.11606	-0.07381	0.00200	-0.02031	0.05666	-0.00013	-0.01779	-0.00246	0.01615	-0.00146	0.00231
C 101	0.06919	-0.02049	0.00225	-0.01519	-0.01496	0.02355	-0.01073	-0.02425	-0.01490	0.00031	-0.01378
C 102	-0.00609	0.09458	-0.03099	-0.00077	-0.02025	-0.01974	0.00129	-0.01017	-0.00505	0.00622	-0.00985
C 103	0.06925	-0.00527	0.01015	-0.02691	-0.01614	-0.00910	-0.04522	0.00621	0.01606	0.02619	-0.00161
C 104	-0.00412	0.04225	0.02657	0.04295	0.03020	-0.00885	0.00233	0.00794	0.00623	0.01421	0.01736
C 105	0.01898	0.06397	-0.00372	-0.00414	-0.01090	0.00885	0.02188	-0.00094	-0.01917	-0.01175	-0.00986
C 106	0.07571	0.10465	0.02649	0.00660	0.01096	-0.00787	0.01072	0.00065	-0.00799	0.00479	-0.00044
C 107	-0.00256	-0.05214	0.02285	0.05891	0.03356	0.02491	-0.00717	-0.00117	-0.01574	0.00948	-0.00356
C 108	-0.03295	0.02361	0.05778	0.01982	-0.00689	0.00624	-0.00255	0.00081	-0.03057	0.00944	-0.00047
C 109	-0.00662	0.05792	0.00592	-0.02920	0.01854	-0.00762	0.02898	0.00075	-0.02933	-0.01052	0.00655
C 110	0.00892	0.04965	-0.05746	0.04413	0.00745	-0.00401	0.01376	-0.02116	-0.00671	0.00710	-0.01727
C 111	0.14406	0.08474	-0.04401	0.01456	0.01093	-0.01468	0.02272	-0.01013	0.03112	0.00554	0.00922
C 112	-0.00534	-0.02490	0.01203	0.04129	0.02800	0.00456	0.01341	0.00273	-0.00254	-0.00241	-0.00845
C 113	-0.00042	0.03715	0.05550	-0.00870	0.02532	-0.03206	0.01128	0.02179	0.00120	0.00666	0.01526
C 114	0.12270	-0.00880	0.03301	-0.00899	0.00267	-0.01818	-0.00854	0.01858	0.01132	0.02301	-0.01683
C 115	-0.05109	-0.02588	0.02913	-0.01329	0.02806	-0.01032	-0.02845	-0.01302	0.01407	-0.01883	0.00894
C 116	-0.03018	-0.01320	0.05649	0.00548	-0.00396	0.01590	-0.04669	0.00991	0.02375	-0.00319	-0.00268
C 117	0.09738	-0.05306	0.01096	0.04697	-0.02011	0.03723	0.01550	0.00992	0.00699	0.00453	0.00397
C 118	-0.02509	0.03148	0.04270	0.02634	0.01074	-0.01362	0.04052	0.00514	-0.00506	0.00502	-0.00834
C 119	-0.07569	0.07721	0.02207	0.02122	-0.02172	0.00347	-0.00990	-0.00553	-0.00090	0.00325	0.00441
C 120	0.04379	-0.12169	-0.03343	0.08037	0.00397	0.00005	-0.01229	-0.02681	-0.01116	-0.01192	0.00589

Continued on next page

APPENDIX D. SHAPE MODEL FEATURES

Table D.1 – *Continued from previous page*

C 121	-0.03173	0.05736	0.01318	0.00241	-0.02500	0.01290	-0.02722	0.00707	0.02477	-0.02233	-0.00259
C 122	-0.12990	0.04038	-0.01610	0.02718	-0.02319	-0.00004	0.00173	-0.00350	0.02343	0.01591	0.00897
C 123	0.02765	-0.02916	-0.01945	0.02071	0.00311	-0.00666	-0.00620	0.01557	-0.03701	-0.01511	0.00712
C 124	-0.06724	0.09489	-0.02825	0.02854	-0.01869	0.00071	-0.00006	0.01478	0.01528	0.00048	0.00079
C 125	0.06140	0.01920	0.01427	-0.02762	0.01026	-0.03339	-0.00681	0.01095	0.00488	-0.00618	-0.00630
C 126	-0.00451	-0.03473	0.00061	-0.02174	0.02338	-0.05616	-0.00113	-0.01273	-0.01589	0.01109	-0.00824
C 127	-0.13886	-0.02558	-0.09154	0.02235	0.04088	-0.01195	0.04596	-0.00270	0.01190	-0.00710	-0.02762
C 128	-0.15523	0.06098	0.02351	0.01943	0.02010	-0.01600	0.00623	0.00982	0.01630	-0.00063	0.00469
C 129	-0.02722	0.05599	-0.00411	0.03005	-0.00730	-0.00744	0.04001	-0.00643	-0.01985	0.00956	0.00217
C 130	0.05896	-0.04671	-0.02414	0.04724	-0.02686	-0.01268	-0.02523	-0.02012	-0.00112	-0.00081	-0.00798
C 131	0.04245	0.02044	0.01590	-0.02526	-0.01843	-0.01394	-0.02091	0.01639	0.00264	0.01238	0.00294
C 132	0.03869	-0.02714	-0.05555	0.04311	-0.03160	-0.00701	-0.00190	-0.00058	-0.00958	-0.00963	-0.00262
C 133	-0.07154	-0.09259	0.01004	-0.01576	-0.00220	-0.01005	-0.02412	-0.00844	-0.01012	0.01333	0.00498
C 134	-0.08146	0.01763	-0.01475	0.05686	0.00703	-0.00899	-0.01305	-0.02262	-0.00374	0.00734	-0.00015
C 135	0.04900	-0.02205	-0.00545	-0.02441	0.00425	-0.03520	-0.01878	0.00684	-0.01410	-0.01102	-0.00140
C 136	0.03568	0.02253	0.04984	-0.03577	-0.01161	-0.02118	-0.00117	0.00041	0.00263	-0.01514	-0.01237
C 137	-0.03986	0.01092	0.02351	0.01225	0.00200	0.00896	0.01166	0.00873	-0.00504	-0.00069	0.00643
C 138	0.04019	-0.12452	-0.15183	0.00517	0.05958	0.02856	-0.03971	-0.03012	0.01804	-0.02606	0.00184
C 139	0.07356	-0.03744	0.01292	0.04268	-0.04605	-0.01892	0.00277	-0.00292	0.00703	-0.00385	-0.00971
C 140	0.03828	-0.03682	-0.00537	-0.01639	-0.00011	0.03691	0.00235	-0.03312	0.00946	-0.01261	0.00032
C 141	0.00852	0.06844	-0.05178	-0.01697	-0.00520	-0.03064	0.01867	-0.03457	-0.01063	-0.01225	-0.00951
C 142	-0.05567	0.09992	-0.09109	0.04836	-0.05050	-0.04443	-0.04744	0.01109	0.03194	-0.00675	0.00225
C 143	-0.02468	0.01415	0.01336	-0.02290	0.01118	-0.02295	0.01433	-0.00253	-0.00257	-0.00024	0.00155
C 144	-0.15130	0.08253	-0.09371	-0.00111	0.03034	-0.03011	0.01290	0.04259	0.02223	-0.00357	-0.01463
C 145	-0.01640	0.02715	0.05979	0.00113	0.04706	0.02228	0.02552	0.02295	0.01039	-0.02613	0.01355
C 146	0.06819	-0.13216	0.04053	0.06027	0.04673	-0.06385	0.02778	-0.00663	0.00275	-0.01093	-0.00495
C 147	-0.02324	-0.06212	-0.02319	0.01621	0.01123	0.00109	0.00897	0.00646	-0.01969	0.00148	-0.00714
C 148	-0.02851	0.07814	-0.04062	-0.03157	-0.00640	-0.01611	-0.01775	-0.00910	-0.04049	-0.01985	-0.00386
C 149	0.03936	0.02825	-0.00951	0.07792	-0.03432	0.00936	0.01832	0.02524	0.01374	-0.02374	-0.02108
C 150	0.00107	-0.04282	0.04197	0.04202	-0.02238	-0.02648	-0.05464	0.05013	0.01058	-0.01486	0.01107
C 151	0.03828	0.00038	0.04330	-0.01055	-0.02215	0.01751	-0.00096	0.02494	0.01962	-0.03485	-0.00517

Continued on next page

APPENDIX D. SHAPE MODEL FEATURES

Table D.1 – *Continued from previous page*

C 152	-0.03486	-0.00742	0.01956	0.04718	-0.00988	-0.01236	-0.04208	0.00614	-0.01989	-0.00719	0.01106
C 153	-0.02987	-0.02160	0.01370	0.01125	0.01743	0.01025	-0.00796	0.01421	0.04221	0.00367	0.01449
C 154	0.02896	-0.20451	0.01027	-0.02354	0.03555	-0.00347	-0.00854	-0.01275	0.05238	-0.00977	-0.06025
C 155	-0.09162	-0.06059	0.02961	0.00362	0.01149	0.01378	0.00664	0.01123	-0.00086	-0.00629	-0.00099
C 156	-0.17603	-0.01751	-0.00500	-0.09175	0.04694	0.00193	-0.00066	0.03788	0.00459	0.00049	0.01416
C 157	0.02923	0.05911	-0.02248	0.08178	-0.02898	-0.03599	-0.02036	-0.00453	-0.01865	-0.00574	0.00468
C 158	0.00815	-0.00294	0.03851	-0.02365	0.01001	0.02476	-0.02054	0.02119	-0.00020	-0.00669	0.00429
C 159	-0.02788	0.00391	-0.01739	0.00773	0.04884	0.01369	-0.00275	0.01839	0.00737	-0.01333	0.00203
C 160	-0.06451	0.02320	-0.03819	0.03704	-0.01396	-0.01097	0.01603	0.00613	-0.00199	0.00483	-0.00462

Table D.2: Principal components used for classification of the RMB-RUL-BI region. Rows are cases, columns are the 11 components used for classification. Rows are cases, columns are the 11 components used for classification. Gosh non-TB training set: C1-C35, Tygerberg TB training set: C36-C64, Tygerberg TB test set: C65-C96, and Tygerberg non-TB test set: C97-C137.

C 1	0.09973	0.01411	0.04256	-0.06621	-0.03136	0.04565	0.02455	0.01551	-0.03670	-0.00881	-0.00699
C 2	0.01968	0.03215	-0.01063	-0.00847	-0.02177	0.04012	-0.00869	0.01774	0.02081	-0.03345	-0.00797
C 3	0.03961	0.14930	-0.04182	0.03354	0.03791	-0.00112	0.00410	0.00784	-0.00143	0.01217	-0.01286
C 4	0.03030	0.00624	0.01364	0.04105	-0.04035	0.01005	-0.00096	0.00971	0.01781	-0.00267	-0.00427
C 5	0.12056	-0.00410	0.02183	0.07195	0.00938	-0.00593	0.07089	-0.00899	-0.04712	0.00193	-0.03722
C 6	0.13224	0.01088	0.06705	0.00155	-0.04874	-0.00938	0.02924	0.01661	-0.02094	0.02872	0.01681
C 7	0.02979	-0.09625	0.07819	0.03869	-0.04392	0.02305	0.03670	-0.01788	-0.02750	0.01377	0.02655
C 8	0.07483	0.07148	-0.02531	0.01462	0.07182	-0.06226	-0.01651	-0.01789	-0.00209	0.02172	-0.00391
C 9	0.08502	-0.07805	-0.02766	-0.00969	0.00148	-0.00717	-0.00119	-0.00678	0.01313	-0.00227	0.01066
C 10	0.06824	-0.00592	-0.04708	0.05031	0.04398	0.01563	-0.04218	-0.01187	0.01089	0.00377	-0.01050
C 11	0.15702	0.03235	0.10353	-0.04212	-0.06744	-0.03759	0.03224	-0.02603	-0.02935	-0.02320	-0.00378
C 12	0.01503	0.13229	-0.17238	0.08522	-0.02273	0.01879	-0.01579	0.04683	-0.02053	-0.01811	0.00101
C 13	0.01326	-0.07935	-0.09091	0.04272	0.04148	-0.02631	0.00523	-0.01104	-0.01382	-0.02421	-0.00468
C 14	0.12998	-0.00559	0.01391	0.00076	-0.00881	-0.00392	0.02920	0.01295	0.02753	-0.01510	-0.01518

Continued on next page

APPENDIX D. SHAPE MODEL FEATURES

Table D.2 – *Continued from previous page*

C 15	0.05237	-0.02037	-0.02647	0.05716	-0.01292	-0.01809	0.00183	0.00534	0.01127	0.02151	-0.01065
C 16	0.09172	0.09257	0.00383	0.01762	-0.00852	-0.02795	-0.01537	0.01314	0.02429	0.00662	-0.00074
C 17	0.11911	0.00104	0.01497	0.07082	-0.03350	-0.02971	0.01614	-0.01140	-0.00233	-0.01973	-0.01648
C 18	0.15906	-0.07214	-0.01856	0.01671	0.06651	-0.00951	0.01101	-0.00719	0.00911	-0.01982	0.03223
C 19	0.16995	0.05595	0.09533	-0.01712	0.02198	-0.03591	-0.02720	-0.04690	-0.00931	-0.01730	0.00994
C 20	0.16108	0.01890	-0.01060	-0.07026	0.01545	-0.02136	0.02455	0.01086	-0.01624	-0.00895	-0.00708
C 21	0.17205	0.03649	-0.00895	-0.01077	-0.01634	-0.01696	0.05474	0.01930	-0.00211	-0.02363	-0.00393
C 22	0.02809	-0.11533	0.01549	0.07632	0.06833	-0.05838	0.01199	0.02017	0.04849	0.01096	0.00807
C 23	0.13443	0.07066	0.04207	-0.04415	-0.10060	0.01161	-0.00387	-0.00495	0.01244	-0.03693	0.01231
C 24	0.05469	-0.07270	0.06864	0.03159	-0.04096	0.01258	0.01007	-0.02050	0.00982	0.01638	0.01812
C 25	0.00053	-0.03677	-0.03665	0.11711	0.05241	-0.02222	0.00064	0.00589	0.01762	-0.00057	-0.00715
C 26	0.13712	0.00901	-0.03037	0.01734	-0.01964	-0.01659	0.01175	0.00685	-0.00976	0.01245	-0.00751
C 27	0.13183	0.03712	-0.02673	-0.02009	-0.01448	0.00376	0.03199	0.02863	0.02346	-0.00796	0.00864
C 28	0.07877	-0.01227	-0.06879	0.00837	0.02006	-0.00233	-0.01905	0.03586	0.00632	0.02243	0.01593
C 29	0.03623	-0.02558	-0.07287	-0.00425	0.00502	-0.01095	0.02202	0.00550	-0.01064	-0.01783	-0.03501
C 30	0.15185	0.09930	-0.03818	0.00077	-0.00348	0.00341	-0.01387	0.00390	-0.00454	-0.00969	0.02200
C 31	0.07655	0.11309	-0.05787	-0.00438	-0.05534	0.01754	0.02275	0.02457	-0.00133	-0.00273	-0.00103
C 32	-0.03006	-0.03587	-0.03257	0.05931	0.02421	-0.00588	-0.01679	-0.02773	0.00690	-0.03028	-0.01903
C 33	0.15217	0.03385	-0.03306	0.02446	0.04778	-0.02697	0.07172	-0.00244	0.00928	-0.01586	-0.00351
C 34	0.19800	0.00695	0.06918	0.02731	-0.04658	-0.08021	0.07492	0.00328	0.00570	-0.00288	-0.00360
C 35	0.12576	-0.08268	-0.12289	0.10912	0.09811	-0.00655	0.03820	0.01710	0.00123	-0.00496	0.02483
C 36	0.02243	-0.05623	-0.00721	-0.07688	0.01915	0.03442	-0.00416	-0.03326	-0.00085	-0.01817	-0.00459
C 37	-0.02105	-0.02004	0.00448	-0.00032	-0.06047	0.02294	-0.01363	-0.04349	0.00473	-0.00950	0.01005
C 38	-0.05278	-0.02363	-0.12471	0.01008	-0.04112	-0.02943	0.01682	-0.03117	-0.03758	-0.00576	-0.01959
C 39	-0.12726	-0.12862	-0.04795	-0.13374	0.01877	-0.11036	0.04589	0.03822	0.01289	0.03300	-0.00144
C 40	-0.05050	0.00678	-0.00967	0.02535	-0.02881	0.05756	0.01031	-0.05789	0.03018	0.00621	0.00254
C 41	-0.04040	0.04021	-0.03229	0.01182	0.04887	0.00086	-0.00160	-0.00544	0.01145	-0.01377	-0.03332
C 42	0.01456	0.09953	0.00355	-0.03118	-0.02156	0.10887	0.00384	-0.04618	-0.01799	0.02660	0.03393
C 43	-0.09703	-0.03011	0.12425	-0.06114	0.03279	0.05110	0.03484	0.02096	0.01843	-0.01655	0.02839
C 44	0.01366	-0.01448	0.05386	-0.00350	-0.06790	0.00961	-0.00251	-0.05845	0.03174	0.01584	0.03149
C 45	-0.09214	0.07567	-0.00889	-0.01240	0.05800	0.02484	0.06480	0.01156	0.02235	-0.03871	-0.02969

Continued on next page

APPENDIX D. SHAPE MODEL FEATURES

Table D.2 – *Continued from previous page*

C 46	-0.13828	0.05510	-0.00815	0.02757	0.07290	0.01849	0.03145	-0.00936	-0.00789	0.00554	-0.01335
C 47	-0.09652	-0.00411	-0.02202	0.01887	0.03818	-0.01787	0.00824	-0.03159	0.01133	0.01701	0.00499
C 48	-0.01169	-0.02990	0.01270	-0.04706	0.00928	0.00747	-0.02150	0.00946	-0.02050	-0.00995	-0.01134
C 49	-0.01616	-0.06091	-0.02495	0.01746	0.00673	-0.00664	-0.02391	-0.02120	-0.00757	-0.01244	0.00127
C 50	0.06748	-0.04344	0.02856	0.06537	0.06432	0.04461	-0.02438	0.03499	0.02560	0.00704	0.05306
C 51	-0.05486	-0.04972	0.06235	-0.02886	0.01343	0.04042	-0.03625	0.04898	-0.03358	0.00330	-0.02066
C 52	-0.05264	-0.06979	0.02128	-0.00213	-0.04714	0.00791	0.03962	-0.02395	0.02137	-0.01312	-0.02133
C 53	-0.06378	0.02615	0.07767	-0.01817	0.02378	-0.02822	-0.01681	-0.01425	-0.00427	0.00410	-0.01891
C 54	-0.09984	0.06644	-0.02953	0.03769	-0.03305	-0.00989	0.01099	-0.01949	0.01373	0.02484	-0.00287
C 55	0.03608	-0.18576	-0.01417	-0.02267	0.05437	-0.01967	-0.00453	-0.05824	0.00085	-0.03332	0.03355
C 56	-0.10194	0.05243	0.02192	-0.06118	0.05230	0.04128	0.05538	-0.01622	-0.00873	0.02637	0.01757
C 57	-0.19222	0.01562	0.02398	0.07546	0.00051	-0.01964	-0.02485	-0.00158	-0.01350	-0.03789	-0.01341
C 58	-0.08925	0.02537	0.11174	0.05811	0.02251	-0.01175	-0.02987	-0.00345	0.01897	0.02901	-0.00415
C 59	-0.07141	-0.09528	-0.00425	0.05054	0.00812	0.01808	-0.00915	-0.01492	-0.04655	0.02847	0.02164
C 60	0.00375	-0.07743	-0.08795	0.10261	0.04937	0.03133	0.01872	0.01927	0.01530	-0.01244	0.01903
C 61	-0.00723	0.03171	0.11879	0.00705	0.08800	0.02616	0.05819	-0.01536	0.02839	0.00044	0.01077
C 62	-0.03331	-0.10862	0.02700	0.04022	-0.05222	-0.01013	0.00580	-0.06306	-0.03327	-0.01337	-0.02160
C 63	0.00548	-0.09892	0.04212	0.03591	-0.01256	0.00517	0.03207	-0.00905	-0.04300	0.00806	-0.01979
C 64	-0.04750	-0.06724	-0.01691	0.02146	0.01681	-0.00968	0.00852	-0.06979	0.04294	-0.00292	0.01308
C 65	-0.06026	0.04559	-0.06925	-0.01218	0.06094	0.00513	0.01301	-0.04709	-0.00752	-0.00299	0.01555
C 66	-0.05167	0.06786	0.00919	0.04810	-0.01165	-0.00612	-0.02520	-0.00333	0.00792	-0.00065	-0.02914
C 67	-0.03446	0.04989	-0.00717	-0.01506	0.04157	0.06089	0.04989	0.00961	0.02870	0.00532	-0.00632
C 68	-0.05739	-0.08282	-0.04592	0.04254	-0.04679	0.06380	0.01788	0.03473	-0.01348	0.01497	0.02118
C 69	-0.04101	0.14949	-0.05376	-0.03938	0.05933	-0.01270	0.02372	-0.01033	-0.01307	0.03402	0.01316
C 70	-0.00884	-0.00402	-0.01540	-0.11892	0.02407	-0.04087	-0.05386	-0.00154	-0.01221	0.00808	-0.01175
C 71	-0.09184	0.05199	-0.13842	-0.11134	-0.04140	-0.04327	0.00590	0.02327	-0.00315	-0.04119	-0.02551
C 72	-0.13449	-0.06211	-0.01186	-0.00278	0.00885	-0.05101	0.00157	0.04910	-0.02061	0.02129	0.00280
C 73	-0.13373	-0.01284	-0.00961	-0.04097	-0.03627	0.04185	0.06849	0.01565	0.03047	-0.00969	0.00508
C 74	-0.03776	-0.11458	0.01073	-0.00535	-0.02731	0.03437	-0.04354	0.01791	-0.03347	0.00359	0.00677
C 75	-0.02172	-0.01314	0.01058	0.02937	-0.00725	0.02756	-0.06157	0.00803	-0.02702	-0.00349	-0.00517
C 76	0.00640	0.00444	0.03935	0.00487	-0.02899	0.00585	-0.03911	-0.04701	0.03054	-0.00678	-0.01457

Continued on next page

APPENDIX D. SHAPE MODEL FEATURES

Table D.2 – *Continued from previous page*

C 77	-0.00908	-0.13540	-0.00593	-0.00017	-0.05438	0.02769	0.01133	0.00622	0.03422	-0.01628	0.00378
C 78	0.01752	-0.03644	-0.00384	-0.04257	0.05346	-0.01618	-0.04158	-0.04016	-0.02620	-0.00831	0.01180
C 79	-0.11920	0.04294	0.03205	-0.08062	0.05038	-0.01102	0.04007	-0.00016	0.03007	0.01783	0.00132
C 80	-0.01697	0.05590	-0.00338	0.00300	-0.04927	0.10657	0.03094	0.00504	0.02815	-0.02423	-0.01825
C 81	-0.03528	-0.02310	0.02023	-0.00860	-0.02868	0.01912	-0.02081	-0.02857	-0.00956	-0.00149	-0.01994
C 82	-0.03176	-0.05317	-0.01542	-0.01898	0.02013	0.04071	0.00820	0.04538	0.00232	-0.00052	-0.00765
C 83	-0.10427	0.01783	-0.07102	-0.14684	-0.01898	-0.03171	0.02565	0.01939	-0.00230	0.06608	0.01047
C 84	-0.10219	0.09159	-0.02147	-0.03809	0.00096	0.01474	0.03900	-0.00289	0.01497	0.01752	-0.00762
C 85	-0.10768	-0.05462	-0.03456	0.07282	0.03742	0.00175	-0.00222	-0.04540	-0.00606	0.01576	-0.00236
C 86	-0.12337	0.05706	0.04018	-0.02998	0.04135	0.01989	-0.00244	0.00818	-0.00482	-0.02126	-0.01727
C 87	0.04671	-0.11153	0.00608	-0.04253	0.00184	0.02360	-0.01534	0.01003	0.03350	-0.02565	0.00686
C 88	-0.05034	0.05659	0.00590	0.00951	0.02673	0.02460	0.03694	-0.06285	-0.01054	0.05335	0.01005
C 89	-0.02473	-0.01055	0.00531	0.00631	0.01120	0.02015	0.00956	0.02692	0.00484	0.01652	-0.03777
C 90	0.03313	-0.01366	0.01530	0.00064	-0.01467	-0.01728	-0.01378	0.01082	0.02459	0.01506	-0.03425
C 91	-0.02832	0.10075	0.00955	0.10237	0.01312	0.05277	0.04013	0.02855	-0.02097	0.00494	0.00983
C 92	-0.08564	0.03818	0.00694	0.03649	-0.03981	0.01202	0.03996	-0.01186	0.04053	0.00034	-0.02300
C 93	-0.10268	0.04490	-0.09040	-0.02740	-0.02515	-0.01425	0.01403	-0.02545	-0.01031	0.01339	-0.01045
C 94	-0.04471	0.00747	-0.06003	-0.01277	-0.05292	-0.01962	0.01626	-0.04100	-0.02049	0.00502	-0.03060
C 95	0.00225	-0.00621	0.02397	0.05243	-0.02023	-0.00920	0.02608	-0.00716	-0.05357	-0.00512	-0.01696
C 96	0.01482	-0.05754	0.00201	-0.05133	-0.05047	0.01637	-0.01064	-0.05416	-0.01316	0.03536	0.00817
C 97	-0.03353	0.10275	-0.04742	0.01788	0.00003	0.06378	-0.03565	0.02395	-0.01349	-0.00621	-0.00092
C 98	0.04394	0.03734	-0.04642	-0.02613	-0.03175	0.01532	-0.00912	-0.00637	0.02569	0.01384	-0.00478
C 99	0.03833	0.06128	-0.09394	-0.11105	-0.01526	-0.00323	-0.01823	0.00106	0.00277	0.01482	-0.00266
C 100	0.05892	-0.05443	-0.03605	-0.09135	0.04944	-0.03037	-0.03434	-0.02871	0.03402	-0.00525	0.01114
C 101	-0.07345	0.10978	0.03836	0.02572	-0.06765	-0.01484	-0.02345	-0.01449	0.00338	-0.01893	-0.00549
C 102	0.02287	-0.08567	0.01417	0.00535	-0.00670	0.03212	-0.07047	-0.00192	-0.00933	0.01225	0.01964
C 103	-0.06406	-0.03649	0.01244	0.03418	-0.07045	-0.02231	-0.01379	0.01693	0.07116	-0.02493	0.00813
C 104	-0.05330	-0.02609	0.12249	0.00331	-0.00823	-0.01561	-0.02759	0.00946	-0.03768	0.00816	0.00074
C 105	0.00099	-0.05991	0.07285	-0.08856	0.07537	0.00563	0.01281	-0.01177	-0.02024	-0.00747	0.00478
C 106	-0.00163	0.01365	-0.05841	-0.02221	-0.02335	0.00235	-0.00570	-0.00452	-0.02347	0.02626	-0.00601
C 107	-0.10849	0.05303	0.04059	0.06550	0.00281	-0.03486	-0.03630	0.01151	0.03023	0.00883	-0.00496

Continued on next page

APPENDIX D. SHAPE MODEL FEATURES

Table D.2 – *Continued from previous page*

C 108	0.03417	0.05266	0.12976	0.03011	-0.03784	-0.06128	0.01811	0.02714	0.02618	0.02090	0.00938
C 109	-0.08464	-0.05666	0.01634	0.00100	0.00218	-0.00439	-0.05425	0.01689	0.04012	-0.00599	-0.00347
C 110	0.00123	0.05866	0.03434	0.05290	0.02625	-0.01577	-0.04146	0.00416	0.03675	0.02184	-0.00076
C 111	0.04958	0.00424	-0.02178	-0.11603	-0.01489	0.01797	-0.01566	-0.01240	0.00581	-0.01501	-0.00255
C 112	0.06034	0.10952	0.14333	-0.02251	0.04863	0.01357	-0.02560	0.04621	-0.06344	-0.00296	0.00139
C 113	0.07193	0.09781	0.00913	-0.03914	0.01005	-0.00930	-0.06518	-0.02689	0.01191	0.00492	0.02923
C 114	0.00754	0.13167	0.01547	-0.02536	-0.03873	-0.02136	-0.05421	-0.01777	-0.00479	0.01889	-0.00359
C 115	-0.04610	0.01636	0.05167	0.02437	0.04024	0.01377	-0.03436	0.02291	-0.01098	-0.04562	-0.02751
C 116	-0.00226	-0.03851	-0.03517	-0.06676	0.07097	-0.00869	-0.02414	0.01910	0.00480	0.00968	-0.01637
C 117	0.04458	-0.04787	0.01463	0.04289	-0.00667	-0.03768	-0.00842	0.01460	-0.01831	0.01408	-0.00038
C 118	0.07409	-0.01540	-0.02129	-0.08121	-0.00262	-0.00111	-0.01101	0.01863	0.03237	0.01894	-0.02081
C 119	-0.10284	-0.00449	0.09418	0.13882	-0.11018	-0.04914	0.00539	0.04215	0.00402	0.04403	0.01251
C 120	-0.00583	0.03555	-0.03248	-0.07481	-0.00633	-0.01619	-0.06770	0.03181	0.01641	-0.02328	-0.01490
C 121	0.03742	0.03996	-0.05368	0.13517	0.02257	-0.01965	-0.04400	-0.00763	-0.02450	-0.00660	-0.00025
C 122	0.05085	-0.01046	0.04027	-0.05433	0.03040	-0.01627	-0.00273	0.01903	-0.01722	0.01022	-0.00169
C 123	0.02242	0.06333	0.12552	-0.00468	0.02126	0.00246	-0.03658	0.02208	0.02271	-0.01185	0.00279
C 124	-0.02178	0.00614	-0.03467	0.02725	-0.02176	0.02288	-0.01659	-0.01654	-0.01274	-0.01310	-0.01624
C 125	-0.08212	-0.09467	-0.12173	0.03206	-0.05041	-0.01865	-0.04046	-0.00171	-0.00114	-0.02882	0.01287
C 126	0.03237	0.11618	0.04847	0.02458	0.05063	0.00131	-0.05851	0.03913	0.02951	0.00560	0.02597
C 127	-0.09578	0.06784	0.10792	0.06429	0.03758	-0.03422	-0.02728	0.00309	-0.03750	-0.00594	-0.01607
C 128	0.02620	-0.04416	0.09514	-0.13309	0.02408	0.07752	-0.01528	0.00600	-0.01436	-0.02704	0.00840
C 129	0.04935	-0.11774	-0.08141	-0.01839	-0.01576	0.03236	0.01628	0.07011	-0.02661	0.01940	0.01628
C 130	-0.14562	-0.15029	0.10822	-0.07912	-0.04275	-0.06761	0.02159	0.05677	0.01340	-0.02374	-0.00090
C 131	-0.03949	-0.07295	-0.09071	0.01230	-0.09642	0.02284	-0.03059	0.05218	-0.00218	0.02101	0.02858
C 132	0.12483	0.03442	-0.12996	-0.06071	-0.03118	-0.01817	-0.04971	-0.01582	-0.00427	0.02150	0.01399
C 133	-0.00434	-0.08240	0.03004	-0.03944	0.04569	-0.01538	0.00755	0.01243	-0.05450	0.00601	-0.01635
C 134	-0.25465	0.18578	-0.03386	-0.01176	-0.03472	-0.10464	0.04495	0.00429	-0.04825	-0.08375	0.12316
C 135	0.03454	-0.02767	-0.03262	0.02801	-0.01151	0.05891	0.00328	0.04200	-0.01258	0.02476	0.01385
C 136	-0.01417	-0.08272	0.03550	0.02835	0.00164	0.01830	-0.00550	0.00650	-0.02551	-0.02593	-0.00590
C 137	0.06387	0.00837	0.04533	-0.01325	-0.03352	-0.04333	-0.00199	-0.00548	0.01419	0.03225	-0.00781

References

- D.G. Altman, D. Machin, T.N. Bryant, and M.J. Gardner. *Statistics with confidence*. BMJ Books, London, 2nd edition, 2005.
- S. Andronikou and N. Wieselthaler. Modern imaging of tuberculosis in children: thoracic, central nervous system and abdominal tuberculosis. *Pediatric Radiology*, 34:861–875, 2004.
- S. Andronikou, E. Joseph, S. Lucas, S. Brachmeyer, G.D. Toit, H. Zar, and G. Swingler. CT scanning for the detection of tuberculous mediastinal and hilar lymphadenopathy in children. *Pediatric Radiology*, 34:232–236, 2004.
- S. Andronikou, B. Irving, T. Hlabangana, T. Pillay, P. Taylor, P. Goussard, and R. Gie. Technical developments of post processing of paediatric airway imaging. *Pediatric Radiology*, 2012. (In Press).
- D. Aykac, E.A. Hoffman, G. McLennan, and J.M. Reinhardt. Segmentation and analysis of the human airway tree from three-dimensional X-ray CT images. *IEEE Transactions on Medical Imaging*, 22:940–950, 2003.
- S.R. Aylward, J. Jomier, S. Weeks, and E. Bullitt. Registration and analysis of vascular images. *International Journal of Computer Vision*, 55(2):123–138, 2003.
- U. Bagci, M. Bray, J. Caban, J. Yao, and D.J. Mollura. Computer-assisted detection of infectious lung diseases: A review. *Computerized medical imaging and graphic*, 36:72–84, 2012.

- R.H. Baroni, S. Ashiku, and P.M. Boiselle. Dynamic ct evaluation of the central airways in patients undergoing tracheoplasty for tracheobronchomalacia. *American Journal of Roentgenology*, 184:1444–1449, 2005.
- C. Bauer, H. Bischof, and R. Beichel. Segmentation of airways based on gradient vector flow. In *Proc. of Second International Workshop on Pulmonary Image Analysis*, pages 191–201, 2009.
- S. Benameur, M. Mignotte, S. Parent, H. Labelle, W. Skalli, and J. de Guise. 3d/2d registration and segmentation of scoliotic vertebrae using statistical models. *Computerized Medical Imaging and Graphics*, 27:321–337, 2003.
- S. Benameur, M. Mignotte, F. Destrempes, and J.A. De Guise. Three-dimensional biplanar reconstruction of scoliotic rib cage using the estimation of a mixture of probabilistic prior models. *IEEE Transactions on Biomedical Engineering*, 52: 1713–1728, 2005.
- C.M. Bishop. *Pattern recognition and machine learning*. Springer, New York, 4th edition, 2006.
- P.M. Boiselle, K.S. Lee, S. Lin, and V. Raptopoulos. Cine CT during coughing for assessment of tracheomalacia: preliminary experience with 64-MDCT. *American Journal of Roentgenology*, 187:W175–W177, 2006.
- P.M. Boiselle, J. Catena, A. Ernst, and D.A. Lynch. Tracheobronchial stenoses. In Phillip M. Boiselle and David A. Lynch, editors, *CT of the Airways*, Contemporary Medical Imaging, pages 121–149. 2008.
- F.L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11: 567–585, 1989.
- F.L. Bookstein. Shape and the information in medical images: A decade of the morphometric synthesis. *Computer Vision and Image Understanding*, 66:97–118, 1997.
- S.J. Bourke. *Respiratory Medicine*. Blackwell Publishing, Oxford, 7th edition, 2007.

- O. Bousquet, U. von Luxburg, and G. Ratsch. *Advanced lectures in machine learning*. Springer, Berlin, 2004.
- R.A.L. Brewis, B. Corrin, D.M. Geddes, and G.J. Gibson. *Respiratory Medicine*. W.B. Saunders Company, London, 2nd edition, 1995.
- T. Bülow, C. Lorenz, and S. Renisch. A general framework for tree segmentation and reconstruction from medical volume data. In *Proc. Medical Image Computing and Computer Assisted Intervention*, pages 533–540, 2004.
- J. Collins and E.J. Stern. *Chest Radiology: The Essentials*. Lippincott Williams & Wilkins, Philadelphia, 2007.
- T.F. Cootes and C. J. Taylor. Statistical models of appearance for computer vision. Technical report, University of Manchester, 2004.
- T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, et al. Active shape models-their training and application . *Computer Vision and Image Understanding*, 61:38–59, 1995.
- T.F. Cootes, G.J. Edwards, and C.J. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:681–685, 2001.
- J.B.S. Coulter. Diagnosis of pulmonary tuberculosis in young children. *Annals of Tropical Paediatrics*, 28:3–12, 2008.
- T.F. Cox and M.A.A. Cox. *Multidimensional scaling*, volume 1. CRC Press, 2001.
- R. Daya, R. Pitcher, L. Workman, T. Douglas, and V. Sanders. A pilot study evaluating erect chest imaging in children, using the Lodox Statscan digital X-ray machine. *South African Journal of Radiology*, 13:80–85, 2010.
- M. de Bruijne, B. van Ginneken, M.A. Viergever, and W.J. Niessen. Adapting active shape models for 3D segmentation of tubular structures in medical images. In *Proc. Information Processing in Medical Imaging*, volume 2732, pages 136–147, 2003a.
- M. de Bruijne, B. van Ginneken, M.A. Viergever, and W.J. Niessen. Three-dimensional point distribution models for tubular objects. Technical report, Utrecht University, Utrecht, 2003b.

-
- F. Deligianni, A.J. Chung, and G.Z. Yang. Nonrigid 2-D/3-D registration for patient specific bronchoscopy simulation with statistical shape modeling: Phantom validation. *IEEE Transactions on Medical Imaging*, 25:1462–1471, 2006.
- M. Desbrun, M. Meyer, P. Schröder, and A.H. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. *Proc. SIGGRAPH*, pages 317–324, 1999.
- L.F. Donnelly, K.H. Emery, A.S. Brody, T. Laor, V.M. Gylys-Morin, C.G. Anton, S.R. Thomas, and D.P. Frush. Minimizing radiation dose for pediatric body applications of single-detector helical ct. *American Journal of Roentgenology*, 176(2):303–306, 2001.
- J. du Plessis, P. Goussard, S. Andronikou, R. Gie, and R. George. Comparing three-dimensional volume-rendered CT images with fiberoptic tracheobronchoscopy in the evaluation of airway compression caused by tuberculous lymphadenopathy in children. *Pediatric Radiology*, 39:694–702, 2009.
- G. du Toit, G. Swingler, and K. Iloni. Observer variation in detecting lymphadenopathy on chest radiography. *International Journal of Tuberculosis and Lung Disease*, 6:814–7, 2002.
- A. El-Baz and J.S. Suri. *Lung Imaging and Computer Aided Diagnosis*. Taylor and Francis, 2011.
- C. Fetita, M. Ortner, P.Y. Brillet, Y.O. Hmeidi, and F. Prêteux. Airway shape assessment with visual feed-back in asthma and obstructive diseases. In *Proc. SPIE*, volume 7625, pages 76251E–76251E–12, 2010.
- C.I. Fetita, F. Preteux, C. Beigelman-Aubry, and P. Grenier. Pulmonary airways: 3-D reconstruction from multislice CT and clinical investigation. *IEEE Transactions on Medical Imaging*, 23:1353–1364, 2004.
- P. Fletcher, C. Lu, SM Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23:995–1005, 2004.

- R. Gie. *Diagnostic Atlas of Intrathoracic Tuberculosis in Children: A Guide for Low Income Countries*. Paris: International Union against Tuberculosis and Lung Disease, 2003.
- P. Goussard and R. Gie. Airway Involvement in Pulmonary Tuberculosis. *South African Medical Journal*, 97:986–988, 2007.
- R.G. Grainger and D.J. Allison. *Diagnostic radiology: an Anglo-American textbook of imaging*. Churchill Livingstone, Edinburgh, 1986.
- M. Groher, F. Bender, R.T. Hoffmann, and N. Navab. Segmentation-driven 2d-3d registration for abdominal catheter interventions. In *Proc. Medical Imaging Computing and Computer Assisted Intervention*, pages 527–535, 2007.
- P. Hammond, T.J. Hutton, J.E. Allanson, L.E. Campbell, R.C.M. Hennekam, S. Holden, M.A. Patton, A. Shaw, I.K. Temple, M. Trotter, et al. 3D analysis of facial morphology. *American Journal of Medical Genetics*, 126:339–348, 2004.
- A. Hartner, M. Hartner, E. Cohen, and B. Gooch. Object space silhouette algorithms.
- T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, 2nd edition, 2011.
- T. Heimann and H.P. Meinzer. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*, 13:543–563, 2009.
- WS Holden and GM Ardran. Observations on the movements of the trachea and main bronchi in man. *Journal of the Faculty of Radiologists*, 8:267–275, 1957.
- Y. Hu, R. van den Boom, T. Carter, Z. Taylor, D. Hawkes, H. U. Ahmed, M. Emberton, C. Allen, and D. Barratt. A comparison of the accuracy of statistical models of prostate motion trained using data from biomechanical simulations. *Progress in Biophysics and Molecular Biology*, 103:262 – 272, 2010.
- Y. Hu, T. Carter, H. Ahmed, M. Emberton, C. Allen, D. Hawkes, and D. Barratt. Modelling prostate motion for data fusion during image-guided interventions. *IEEE Transactions on Medical Imaging*, 30:1887–1900, 2011.

- T.J. Hutton, B.F. Buxton, P. Hammond, and H.W.W. Potts. Estimating average growth trajectories in shape-space using kernel smoothing. *IEEE Transactions on Medical Imaging*, 22:747–753, 2003.
- T. Huysmans, J. Sijbers, F. Vanpoucke, and B. Verdonk. Improved Shape Modeling of Tubular Objects Using Cylindrical Parameterization. In *Proc. Medical Imaging and Augmented Reality*, volume 4091, pages 84–91. Springer, 2006.
- B. Irving, P. Goussard, R. Gie, A. Todd-Pokropek, and P. Taylor. Identification of paediatric tuberculosis from airway shape features. In *Proc. Medical Image Computing and Computer Assisted Interventions*, volume 14, pages 133–140, 2011a.
- B. Irving, P. Goussard, R. Gie, A. Todd-Pokropek, and P. Taylor. Segmentation of obstructed airway branches in ct using airway topology and statistical shape analysis. In *Proc. IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 447–451, 2011b.
- B. Irving, A. Todd-Pokropek, and P. Taylor. Airway segmentation and analysis from computed tomography. In A. El-Baz and J.S. Suri, editors, *Lung Imaging and Computer Aided Diagnosis*. CRC press, 2011c.
- B.J. Irving, G.J. Maree, E.R. Hering, and T.S. Douglas. Radiation dose from a linear slit scanning X-ray machine with full-body imaging capabilities. *Radiation Protection Dosimetry*, 130:482–489, 2008.
- B.J. Irving, P. Taylor, and A. Todd-Pokropek. 3D segmentation of the airway tree using a morphology based method. In *Proc. of Second International Workshop on Pulmonary Image Analysis*, pages 297–307, 2009.
- S. Katsuragawa and K. Doi. Computer-aided diagnosis in chest radiography. *Comput. Med. Imaging Graph.*, 31:212–223, 2007.
- M.R. Kaus, V. Pekar, C. Lorenz, R. Truyen, S. Lobregt, and J. Weese. Automated 3-D PDM construction from segmented images using deformable models. *IEEE Transactions on Medical Imaging*, 22:1005–1013, 2003.

- W.S. Kim, J.I. Choi, J.E. Cheon, I.O. Kim, K.M. Yeon, and H.J. Lee. Pulmonary tuberculosis in infants: radiographic and CT findings. *American Journal of Roentgenology*, 187:1024–1033, 2006.
- A.P. Kiraly, W.E. Higgins, G. McLennan, E.A. Hoffman, and J.M. Reinhardt. Three-dimensional human airway segmentation methods for clinical virtual bronchoscopy. *Academic Radiology*, 9:1153–1168, 2002.
- A.P. Kiraly, B.L. Odry, M.C.B. Godoy, B. Geiger, C.L. Novak, and D.P. Naidich. Computer-aided Diagnosis of the Airways: Beyond Nodule Detection. *Journal of Thoracic Imaging*, 23:105–113, 2008.
- T. Kitasaka, H. Yano, M. Feuerstein, and K. Mori. Bronchial region extraction from 3d chest ct image by voxel classification based on local intensity structure. In *Proc. Third International Workshop on Pulmonary Image Analysis*, 2010.
- S.L. Lee, Y.F. Cheung, M.P. Leung, Y.K. Ng, and N.S. Tsoi. Airway obstruction in children with congenital heart disease: assessment by flexible bronchoscopy. *Pediatric Pulmonology*, 34:304–311, 2002.
- N.L. Leung, A.N. and Muller, P.R. Pineda, and J.M. Fitzgerald. Primary tuberculosis in childhood: radiographic manifestations. *Radiology*, 182:87–91, 1992.
- P. Lo, J. Sporning, and M. de Bruijne. Multiscale vessel-guided airway tree segmentation. In *Proc. of Second International Workshop on Pulmonary Image Analysis*, pages 323–332, 2009a.
- P. Lo, J. Sporning, J. Pedersen, and M. Bruijne. Airway tree extraction with locally optimal paths. In *Medical Image Computing and Computer-Assisted Intervention*, volume 5762, pages 51–58. 2009b.
- P. Lo, B. van Ginneken, J. Reinhardt, and M. de Bruijne. Extraction of airways from CT. In *Second International Workshop on Pulmonary Image Analysis*, 2009c.
- P. Lo, B. van Ginneken, J.M. Reinhardt, T. Yavarna, P.A. de Jong, B. Irving, C. Fetita, M. Ortner, R. Pinho, J. Sijbers, M. Feuerstein, A. Fabijanska, C. Bauer, R. Beichel, C.S. Mendoza, S. Zayed, R. Wiemker, J.Lee, A/P. Reeves, S. Born,

- O. Weinheimer, E.M. van Rikxoort, J. Tschirren, K. Mori, B. Odry, D.P. Naidich, L. Hartmann, E.A. Hoffman, M. Prokop, J.H. Pedersen, and M. de Bruijne. Extraction of Airways from CT (EXACT'09). *IEEE Transactions on Medical Imaging*, 2012. (In Press).
- Lodox. Statscan critical imaging system: Product specifications and physical dimensions. Technical report, Lodox Systems, 2006.
- M. Long. Segmentation of body regions and airways in full-body digital paediatric X-ray images. Master's thesis, University of Cape Town, 2008.
- J. Lötjönen, S. Kivistö, J. Koikkalainen, D. Smutek, and K. Lauerma. Statistical shape model of atria, ventricles and epicardium from short- and long-axis MR images. *Medical Image Analysis*, 8:371 – 386, 2004.
- B.J. Marais. Childhood tuberculosis–risk assessment and diagnosis. *South African Medical Journal*, 97:978–982, 2007.
- B.J. Marais, R.P. Gie, C.C. Obihara, A.C. Hesselning, H.S. Schaaf, and N. Beyers. Well defined symptoms are of value in the diagnosis of childhood pulmonary tuberculosis. *British Medical Journal*, 90:1162–1165, 2005.
- B.J. Marais, R.P. Gie, H.S. Schaaf, N. Beyers, P.R. Donald, and J.R. Starke. Childhood pulmonary tuberculosis: old wisdom and new challenges. *American Journal of Respiratory and Critical Care Medicine*, 173:1078, 2006.
- G.J. Maree, B.J. Irving, and E.R. Hering. Paediatric dose measurement in a full-body digital radiography unit. *Pediatric Radiology*, 37:990–997, 2007.
- I.B. Masters, R.S. Ware, P.V. Zimmerman, B. Lovell, R. Wootton, P.V. Francis, and A.B. Chang. Airway sizes and proportions in children quantified by a video-bronchoscopic technique. *BMC Pulm. Med.*, 6:5–13, 2006.
- Y. Masutani, T. Schiemann, and K.H. Hoehne. Vascular Shape Segmentation and Structure Extraction Using a Shape-Based Region-Growing Model. *Lecture Notes in Computer Science*, 1496:1242–1249, 1998.

-
- S. Meller and W.A. Kalender. Building a statistical shape model of the pelvis. In *Proc. Computer Assisted Radiology and Surgery*, volume 1268, pages 561–566, 2004.
- S.C. Mitchell, B.P.F. Lelieveldt, R.J. Van Der Geest, H.G. Bosch, JHC Reiver, and M. Sonka. Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac mr images. *IEEE Transactions on Medical Imaging*, 20:415–423, 2001.
- T. Möller and B. Trumbore. Fast, minimum storage ray/triangle intersection. In *Proc. ACM SIGGRAPH*, 2005.
- K. Mori, J. Hasegawa, J. Toriwaki, H. Anno, and K. Katada. Recognition of Bronchus in Three-Dimensional X-ray CT Images with Application to Virtualized Bronchoscopy System. In *Proc. International Conference on Pattern Recognition*, volume 3, pages 528–532, 1996.
- K. Mori, J. Hasegawa, Y. Suenaga, and J. Toriwaki. Automated anatomical labeling of the bronchial branch and its application to the virtual bronchoscopy system. *IEEE Transactions on Medical Imaging*, 19:103–114, 2000.
- K. Moses, J.C. Banks, P.B. Nava, and D. Petersen. *Atlas of Clinical Gross Anatomy*. Elsevier Mosby, Philadelphia, 2005.
- A. Myronenko and X. Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262–2275, 2010.
- L.J. Nelson and C.D. Wells. Global epidemiology of childhood tuberculosis. *The International journal of Tuberculosis and lung Disease*, 8:636–647, 2004.
- K. Palágyi, J. Tschirren, and M. Sonka. Quantitative analysis of intrathoracic airway trees: methods and validation. In *Proc. Information Processing in Medical Imaging*, volume 2732, pages 222–233. Springer, 2003.
- K. Palágyi, J. Tschirren, E.A. Hoffman, and M. Sonka. Quantitative analysis of pulmonary airway tree structures. *Computers in Biology and Medicine*, 36:974–996, 2006.

- W. Park, E.A. Hoffman, and M. Sonka. Segmentation of intrathoracic airway trees: a fuzzy logic approach. *IEEE Transactions on Medical Imaging*, 17:489–497, 1998.
- R. Paulsen, R. Larsen, C. Nielsen, S. Laugesen, and B. Ersboll. Building and testing a statistical shape model of the human ear canal. In *Proc. Medical Image Computing and Computer Assisted Intervention*, volume 2489, pages 373–380, 2002.
- P.P. Pebay and T.J. Baker. A comparison of triangle quality measures. In *Proc. National Academy of Sciences*, pages 9653–9657. Plenum Press, 1991.
- P.P. Pebay and T.J. Baker. Analysis of triangle quality measures. *Mathematics of Computation*, 72:1817–1840, 2003.
- S.S. Peng, P. Chan, Y. Chang, and T.T. Shih. Computed tomography of children with pulmonary mycobacterium tuberculosis infection. *Journal of the Formosan Medical Association*, 110:744 – 749, 2011.
- J. Petersen, P. Lo, M. Nielsen, G. Edula, H. Ashraf, A. Dirksen, and M. De Bruijne. Quantitative analysis of airway abnormalities in ct. In *Proc. SPIE*, volume 7624, page 76241S, 2010.
- R. Pinho, J. Sijbers, and T. Huysmans. Segmentation of the human trachea using deformable statistical models of tubular shapes. In *Proc. 9th international conference on Advanced concepts for intelligent vision systems*, pages 531–542, 2007.
- R. Pinho, K.G. Tournoy, and J. Sijbers. Assessment and stenting of tracheal stenosis using deformable shape models. *Medical Image Analysis*, 15:250–266, 2011.
- C. Pisupati, L. Wolff, W. Mitzner, and E. Zerhouni. Segmentation of 3D pulmonary trees using mathematical morphology. *Mathematical morphology and its applications to image and signal processing*, pages 409–416, 1996.
- R. Pitcher, A. van As, V. Sanders, N. Wieselthaler, A. Vlok, S. Paverd, T. Kilborn, H. Rode, H. Potgieter, and S. Beningfield. A pilot study evaluating the “STATSCAN” digital X-ray machine in paediatric polytrauma. *Emergency Radiology*, 15:35–42, 2008.

- S.M. Pizer, P.T. Fletcher, S. Joshi, A. Thall, J.Z. Chen, Y. Fridman, D.S. Fritsch, A.G. Gash, J.M. Glotzer, M.R. Jiroutek, et al. Deformable M-reps for 3D medical image segmentation. *International Journal of Computer Vision*, 55:85–106, 2003a.
- S.M. Pizer, P.T. Fletcher, A. Thall, M. Styner, G. Gerig, and S. Joshi. Object models in multiscale intrinsic coordinates via m-reps. *Image and Vision Computing*, 21: 5–15, 2003b.
- W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 2nd edition, 2002.
- F. Prêteux, C.I. Fetita, A. Capderou, and P. Grenier. Modeling, segmentation, and caliber estimation of bronchi in high resolution computerized tomography. *Journal of Electronic Imaging*, 8:36–45, 1999.
- K.T. Rajamani, M.A. Styner, H. Talib, G. Zheng, L.P. Nolte, and M.A.G. Ballester. Statistical deformable bone models for robust 3d surface extrapolation from sparse data. *Medical Image Analysis*, 11(2):99–109, 2007.
- G.D. Rubin. 3-D imaging with MDCT. *European Journal of Radiology*, 45:37–41, 2003.
- D. Ruijters, B.M. ter Haar Romeny, and P. Suetens. Vesselness-based 2D–3D registration of the coronary arteries. *International Journal of Computer Assisted Radiology and Surgery*, 4:391–397, 2009.
- J.C. Russ. *The Image Processing Handbook*. CRC Press, Boca Raton, 5th edition, 2006.
- G.E. Salazar, T.L. Schmitz, R. Cama, P. Sheen, L.M. Franchi, G. Centeno, C. Valera, M. Leyva, S. Montenegro-James, R. Oberhelman, et al. Pulmonary Tuberculosis in Children in a Developing Country. *Pediatrics*, 108:448–453, 2001.
- V. M. Sanders, R. D. Pitcher, T. S. Douglas, M. A. Kibel, R. B. Daya, and A. B. van As. Digital radiographic measurement of the main bronchi: A pilot study. *Annals of Tropical Paediatrics*, 29:209–216, 2009.

- A. Sandgren, L.E. Cuevas, M. Dara, R.P. Gie, M. Grzemska, A. Hawkridge, A.C. Hesselning, B. Kampmann, C. Lienhardt, D. Manissero, et al. Childhood tuberculosis: progress requires advocacy strategy now. *European Respiratory Journal*, 40: 294–297, 2012.
- T. Schlathölter, C. Lorenz, I.C. Carlsen, S. Renisch, and T. Deschamps. Simultaneous segmentation and tree reconstruction of the airways for virtual bronchoscopy. In *Proc. SPIE*, volume 4684, pages 103–113. SPIE, 2002.
- B. Schölkopf, A. Smola, and K.R. Müller. Kernel principal component analysis. In *Artificial Neural Networks - ICANN*, volume 1327, pages 583–588. 1997.
- M.J. Siegel. Multiplanar and three-dimensional multi-detector row ct of thoracic vessels and airways in the pediatric population1. *Radiology*, 229:641–650, 2003.
- M. Sonka, W. Park, and E.A. Hoffman. Rule-based detection of intrathoracic airway trees. *IEEE Transactions on Medical Imaging*, 15:314–326, 1996.
- M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision*. International Thomson publishing, London, 1999.
- M.B. Stegmann and D.D. Gomez. *A Brief Introduction to Statistical Shape Analysis*. Informatics and Mathematical Modelling, Technical University of Denmark, 2002.
- M. Styner, K. Rajamani, L.p. Nolte, G.l Zsemlye, G. Székely, C. Taylor, and R. Davies. Evaluation of 3d correspondence methods for model building. In Chris Taylor and J. Noble, editors, *Information Processing in Medical Imaging*, volume 2732 of *Lecture Notes in Computer Science*, pages 63–75. Springer Berlin / Heidelberg, 2003.
- S. Swaminathan and B. Rekha. Pediatric tuberculosis: global overview and challenges. *Clinical Infectious Diseases*, 50:S184–S194, 2010.
- R.D. Swift, A.P. Kiraly, A.J. Sherbondy, A.L. Austin, E.A. Hoffman, G. McLennan, and W.E. Higgins. Automatic axis generation for virtual bronchoscopic assessment of major airway obstructions. *Computerized Medical Imaging and Graphics*, 26: 103–118, 2002.

- G.H. Swingler, G. du Toit, S. Andronikou, L. van der Merwe, and H.J. Zar. Diagnostic accuracy of chest radiography in detecting mediastinal lymphadenopathy in suspected pulmonary tuberculosis. *Archives of Disease in Childhood*, 90:1153–1156, 2005.
- T. Tezoo and T.S. Douglas. Interactive segmentation of airways from chest x-ray images using active shape models. In *Engineering in Biology and Medicine*, 2012. (Accepted).
- AC Theart, BJ Marais, RP Gie, AC Hesselning, and N. Beyers. Criteria used for the diagnosis of childhood tuberculosis at primary health care level in a high-burden, urban setting. *International Journal of Tuberculosis and Lung Disease*, 9:1210–1214, 2005.
- J. Tschirren, K. Palágyi, J.M. Reinhardt, E.A. Hoffman, and M. Sonka. Segmentation, skeletonization, and branchpoint matching—a fully automated quantitative evaluation of human intrathoracic airway trees. In *Proc. Medical Image Computing and Computer Assisted Intervention*, volume 2489, pages 12–19. Springer, 2002.
- J. Tschirren, E.A. Hoffman, G. McLennan, and M. Sonka. Intrathoracic airway trees: segmentation and airway morphology analysis from low-dose CT scans. *IEEE Transactions on Medical Imaging*, 24:1529–1539, 2005a.
- J. Tschirren, G. McLennan, K. Palágyi, E.A. Hoffman, and M. Sonka. Matching and anatomical labeling of human airway tree. *IEEE Transactions on Medical Imaging*, 24:1540–1547, 2005b.
- J. Tschirren, T. Yavarna, and J. Reinhardt. Airway segmentation framework for clinical environments. In *Proc. of Second International Workshop on Pulmonary Image Analysis.*, pages 227–238, 2009.
- F. van der Heijden, R.P.W. Duin, D. de Ridder, and D.M.J. Tax. *Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB*. Wiley, Chichester, 2004.

- B. van Ginneken, A.F. Frangi, J.J. Staal, B.M. ter Haar Romeny, and M.A. Viergever. Active shape model segmentation with optimal features. *IEEE Transactions on Medical Imaging*, 21:924–933, 2002a.
- B. van Ginneken, S. Katsuragawa, B.M. ter Haar Romeny, Kunio Doi, and M.A. Viergever. Automatic detection of abnormalities in chest radiographs using local texture analysis. *IEEE Transactions on Medical Imaging*, 21:139–149, 2002b.
- B. van Ginneken, W. Baggeman, and E.M. van Rikxoort. Robust segmentation and anatomical labeling of the airway tree from thoracic CT scans. In *Proc. Medical Image Computing and Computer Assisted Intervention*, pages 219–226, 2008.
- E.M. van Rikxoort, B. de Hoop, M.A. Viergever, M. Prokop, and B. van Ginneken. Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. *Medical Physics*, 36:2934–2947, 2009.
- S. Varma and R. Simon. Bias in error estimation when using cross-validation for model selection. *BMC bioinformatics*, 7:1–8, 2006.
- L. Vincent. Morphological grayscale reconstruction: definition, efficient algorithm and applications in image analysis. In *Proceedings of Computer Vision and Pattern Recognition*, volume 92, pages 633–635, 1992.
- L. Vincent. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *IEEE Transactions on Image Processing*, 2:176–201, 1993.
- E. W. Weisstein. Point-line distance–2-dimensional. from mathworld–a wolfram web resource. <http://mathworld.wolfram.com/Point-LineDistance2-Dimensional.html>, a. (Accessed: February 2012).
- E.W. Weisstein. Euler angles - mathworld - a wolfram web resource. <http://mathworld.wolfram.com/EulerAngles.html>, b. (Accessed: January 2010).
- WHO. Summary for TB-HIV estimates for 2005 by WHO regions. www.who.int/tb, 2012. (Accessed: July 2012).

-
- R. Wiemker, T. Blaffert, T. Bülow, S. Renisch, and C. Lorenz. Automated assessment of bronchial lumen, wall thickness and bronchoarterial diameter ratio of the tracheobronchial tree using high-resolution ct. In *Proc. Computer Assisted Radiology and Surgery*, volume 1268, pages 967 – 972, 2004.
- S. A. Wood, E. A. Zerhouni, J. D. Hoford, E. A. Hoffman, and W. Mitzner. Measurement of three-dimensional lung tree structures by using computed tomography. *Journal of Applied Physiology*, 79:1687–1697, 1995.
- S. Zachow, H. Lamecker, B. Elsholtz, and M. Stiller. Reconstruction of mandibular dysplasia using a statistical 3d shape model. In *International Congress Series*, volume 1281, pages 1238–1243, 2005.
- G. Zheng. Statistically deformable 2D/3D registration for accurate determination of post-operative cup orientation from single standard X-ray radiograph. In *Proc. Medical Image Computing and Computer Assisted Intervention*, pages 820–827, 2009.
- G. Zheng. Personalized X-ray reconstruction of the proximal femur via intensity-based non-rigid 2D-3D registration. In *Proc. Medical Image Computing and Computer Assisted Intervention*, pages 598–606, 2011.
- Y. Zheng and D. Doermann. Robust point matching for nonrigid shapes by preserving local neighborhood structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):643–649, 2006.