Tracheal Structure Characterization using Geometric and Appearance Models for Efficient Assessment of Stenosis in Videobronchoscopy

A dissertation submitted by **Carles Sánchez Ramos** at Universitat Autònoma de Barcelona to fulfil the degree of **Doctor of Philosophy**.

Bellaterra, November 9, 2014

Director	**Dr. F.Javier Sánchez Pujadas** Dept. Ciències de la Computació & Centre de Visió per Computador Universitat Autònoma de Barcelona
Co-director	**Dra. Debora Gil Resina** Dept. Ciències de la Computació & Centre de Visió per Computador Universitat Autònoma de Barcelona
Co-director	**Dr. Jorge Bernal del Nozal** Dept. Ciències de la Computació & Centre de Visió per Computador Universitat Autònoma de Barcelona
Thesis committee	**Dr. Laura Igual** Universitat de Barcelona Barcelona, Spain **Dr. Fernando Vilarino** Universitat Autónoma de Barcelona Bellaterra, Spain **Dr. Joaquim de Ciurana** Universitat de Girona Girona, Spain

This document was typeset by the author using LaTeX 2 .

The research described in this book was carried out at the Computer Vision Center, Universitat Autònoma de Barcelona.

ISBN: 978-84-940530-7-8

Printed by Ediciones Gráficas Rey, S.L.

Dedicat a la familia i amics per ser-hi sempre

Acknowledgements

I would like to thank the supervisors Javier Sanchez and Debora Gil for their help and guidance throughout this Master Thesis. I would also thank the help of Jorge Bernal and all the people of CVC for their support. Finally, I would like to thank Toni Rosell from the Bellvitge hospital for their knowledge that helped me to understand the problem. Thank you! You re the best!

Abstract

Recent advances in the optics of endoscopic devices have increased their use for minimal invasive diagnostic and intervention procedures. Among all endoscopic procedures (which include exploration of patient internal organs from mouth to anus), bronchoscopy is one of the most frequents with around 261 millions of procedures per year. Although the use of bronchoscopy is spread among clinical facilities it presents some drawbacks, being the visual inspection for the acquisition and assessment of anatomical measurements the most prevalent of them. One of the basic interventions implies to obtain the percentage of airway obstruction, which is known as stenosis. Errors in this measurement have a direct impact in the treatment the patient will be prescribed. Therefore an objective computation of tracheal stenosis in bronchoscopy videos would constitute a breakthrough for this non-invasive technique and a reduction in operation cost for the public health service due to the decrease in interventions because of bad visual measurements. The calculation of the degree of stenosis is based on the comparison of the region delimited by the lumen in an obstructed frame and the region delimited by the first visible ring in a healthy frame. In order to obtain a good delimitation of these regions a reliable characterization of the needed tracheal structures: rings and luminal area is needed. We propose in this thesis a parametric strategy for the extraction of lumen and tracheal rings regions based on models of their geometry and appearance that guide a deformable model. In order to ensure a systematic applicability we present a statistical framework to select the optimal parameters of our method. We have validated our method in two ways: accuracy of the anatomical structures extraction and comparison of the performance of our automatic method over manually annotated stenosis. Experiments report results within inter-observer variability for both quality metrics. The performance of our method is comparable to the one provided by clinicians and its computation time allows for a on-line implementation in the operating room.

Resum

Els aventatges recents en les òptiques dels dispositius endoscopis han fet que hi hagi un creixement en el seu us per a diagnostics i intervencions no ivasives. Dins de tots els procediments endoscopics (els cuals s incloeix l exploració dels organs interns del pacient des de el nas al anus), la broncoscopia es un dels mes frequents amb 261 millions d intervencions a l any. Encara que el seu us s exten per per la majoria de les cliniques presenta alguns problemes o inconvenients com per exemple el més important de tots: el mesurament de forma visual de les estructures anotòmiques. Una de les intervencions més bàsiques implica l obtenció del percentatge d obtrucció anomenat estenosis. Enrrors en aquesta mesura tenen un impacte directe al tractament del pacient, per tant, un mesurament objectiu de les estenosis traqueals en videos de broncoscòpia constituiria una roptura en la utilització d aquesta tècnica no invasiva i una reducció en el cost de les operacions per la sanitat pública per un decreixement en les operacions repetides per mesuraments visuals errornis. El càlcul del grau d estenosis esta basat en la comparació de la regió delimitada per l area luminal (en un frame patològic) amb la regió delimitada per el primer anell visible de referència (en un frame sa). Per aconseguir tenir una bona delimitació d aquestes regions, una bona definició de les estructures traqueals (anells i lumen) es necessària. En aquesta tesis proposem una estrategia paramètrica per a l extracció de les estructures traqueals necessaries per al càlcul de l estenosis. Per assegurar una aplicabilitat del sistema presentem un protocol estadístic per seleccionar els paràmetres més òptims. Els experiments demostren que el nostre mètode està dins de la variabilitat dels metges, per tant el rendiment es comparable a un metge amb la particularitat que el temps de computació permet una implementació on-line en el moment del diagnòstic.

Resumen

Los últimos avances en la óptica de los dispositivos de endoscopia han motivado un incremento en su uso en procedimientos mínimamente invasivos de diagnóstico y exploración. La broncoscopia, con más de 261 millones de intervenciones al ano, es una de las técnicas endoscópicas -exploración de órganos internos desde la boca hasta el ano- más extendidas. Aunque su uso se ha popularizado en las instalaciones sanitarias aún dista de funcionar de manera perfecta, presentando algunas carencias como la imposibilidad de obtener medidas anatómicas exactas debido a que éstas son tomadas mediante inspección visual. De entre todas las intervenciones broncoscópicas nos centramos en el cálculo de estenosis que permite obtener el porcentaje de obstrucción traqueal. Un error en la estimación del grado de obstrucción tiene impacto en el tratamiento que se le prescribe al paciente, por tanto un cálculo objetivo de la estenosis supondría un salto cualitativo para la técnica así como una reducción de costes para el sistema sanitario al evitar repeticiones de operaciones debido a toma de medidas erróneas. El cálculo del porcentaje de estenosis se basa en la comparación de la región delimitada por el lumen en una imagen de obstrucción traqueal y la región delimitada por el primer anillo visible en una imagen normal de tráquea. Para obtener una definición adecuada de dichas regiones necesitamos de una caracterización eficiente de las estructuras traqueales involucradas: anillos y área luminal. En esta tesis proponemos una estrategia paramétrica para la extracción de dichas estructuras basándonos en un modelo de geometría y apariencia que guía un modelo deformable. Con el fin de potenciar una aplicabilidad real de nuestra propuesta proponemos un marco estadístico para elegir de manera automática los parámetros que optimizan el rendimiento de nuestro método. Nuestra propuesta es validada tanto a nivel de caracterización de estructuras anatómicas como a nivel global de porcentaje de estenosis. Dicha validación se ha realizado sobre bases de datos generadas en el contexto de esta tesis, validadas y anotadas por personal clínico. Los experimentos realizados certifican un rendimiento comparable al de los expertos clínicos, suponiendo nuestro método un coste computacional asumible de cara a una prometedora futura integración en los protocolos actuales.

Contents

1 Introduction **1**
1.1 Motivation and Goal 1
1.2 State of the art and challenges 6
1.3 Contributions 9

2 Geometric and Appearance Model of the Trachea **13**
2.1 Physical Models 13
2.2 Computational Models 21
  2.2.1 Tracheal Lumen Computational Model 21
  2.2.2 Tracheal rings computational model 23

3 Tracheal structures segmentation for stenosis assessment **27**
3.1 Preprocessing 28
3.2 Structure Characterization 33
  3.2.1 Luminal Region 33
  3.2.2 Tracheal ring 35
3.3 Structure Segmentation 38
  3.3.1 Obstructed Lumen 38
  3.3.2 Reference Ring Segmentation 39
  3.3.3 ANOVA design for optimal parameter setting 42

4 Experimental Settings and Databases **45**
4.1 Structure characterization 45
  4.1.1 Lumen centre detection 45
  4.1.2 Tracheal ring 46
4.2 Structure segmentation 48
  4.2.1 Obstructed lumen segmentation 48
  4.2.2 Reference ring segmentation 50
4.3 Stenosis assessment 50
4.4 Anova parameter tuning for stenosis assessment 52

5 Experimental Results **57**
5.1 Structures characterization 57
  5.1.1 Lumen centre detection 57
  5.1.2 Tracheal Ring segmentation 58
5.2 Structures segmentation 62
  5.2.1 Obstructed lumen segmentation 63
  5.2.2 Reference ring Segmentation 64

5.3 Stenosis assessment 67

6 Discussion **71**
6.1 Structure characterization 71
6.1.1 Lumen centre detection 71
6.1.2 Tracheal rings 72
6.2 Structure segmentation 73
6.2.1 Obstructed lumen segmentation 74
6.2.2 Reference ring segmentation 74
6.3 Stenosis assessment 75

7 Conclusions and Future Work **77**
7.1 Conclusions 77
7.2 Future research lines 78

List of Figures

1.1 Optical endoscopes : Rigid models (a,b), flexible model and detail of the light bulb attached at its tip (right) (c). 1
1.2 Bronchoscopic instruments: Flexible bronchocopy (a), rigid tubular device (b), interconnection channels (c) and instruments (d). 2
1.3 Different regions of the bronchoscopy process. 3
1.4 White light image with a tumor (right) and same image with a NBI light (left). 4
1.5 Graphical example of the desired output of our method. The healthy reference ring on the left image and compared to the obstructed lumen show on the right image. 5
1.6 Examples of variability in lumen appearance: single (a) and multiple (b) bronchoscopy image; centred (c) and biased (d) colonoscopy image. 7
1.7 Image quality artefacts: first row : ilumination problems, second row: blurred problems. 8
1.8 Interlacing problem example. 8
1.9 Non anatomical artefacts: first row : carina artefact, second row: rigid bronchoscope artefact. 9
1.10 Geometric artefacts: first row : trachea wall, second row: collapsed tracheal rings. 9
1.11 Graphical example of the process and desired output of our method. The healthy reference ring calculation on the top process and compared to the obstructed lumen process show on the bottom process. 10

2.1 Trachea schema (left) and main anatomical elements in an histological cut.[1] . 14
2.2 Composition of tracheal structures in a videobronchoscopy frame (a) an in a histological cut (b). 15
2.3 Frame valid for stenosis assessment (left), frame not valid for stenosis assessment (right). 16
2.4 Decomposition of the light reflected by the surface into the three components: ambient; diffuse and specular. 16
2.5 Synthetic Model of the Trachea: sinusoidal tracheal surface (left) and longitudinal cut showing the generating sinusiodal profile, (right). 17
2.6 Trachea illumination Appearance in synthetic trachea: main anatomic structures in left image and radial intensity profile in right plot. 18
2.7 CT Model of the Trachea: bronchial system surface (left) and longitudinal cut showing the trachea profile, (right). 19
2.8 Trachea illumination Appearance in CT: main anatomic structures in left image and radial intensity profile in right plot. 19

2.9 Trachea illumination Appearance in Original Images: main anatomic structures in left image and radial intensity profile in right plot. 20
2.10 Other tracheal structures such as veins (first zoom) and carina (second zoom) that have the same valley profile as tracheal rings (third zoom). 20
2.11 Graphical explanation of DGA algorithm: Original synthetic image (a); Corresponding gradient vectors superimposed to the image (b); Example of the extension of gradient vector lines (c), and resulting DGA accumulation map (d). 22
2.12 Real frame, result of applying DGA and its surface representation (first row). Gaussian smoothing applied to the result of DGA and its surface representation (second row). 22
2.13 Graphical explanation of DRI algorithm: Original bronchoscopy image (a); Smoothed images with σ: 1 8 (b); 1 16 (c); 1 24 (d); 1 32 (e); 23
2.14 Challenges in lumen appearance and geometry. Differences in acquisition conditions (first column), differences in appearance within the same sequence (second column) and differences in number of lumens (third column). 23
2.15 Modelling of concentric disposition of tracheal rings in the polar domain: steerable filters in the original cartesian domain (right) and in the polar transform (left) 25
2.16 Modelling of tracheal rings radial thickness by analyzing the profile of maximal scales. 26

3.1 Stenosis assessment workflow. 28
3.2 Graphical example of the extension of the M mask. 31
3.3 Alternatives to solve the apparition of the black mask 31
3.4 Vector field representing level curves of an angiography for a vessel and a background structure-less area 32
3.5 SPD filtering example : Real image (a), Filtered image (b). 33
3.6 Adequation of our feature space to videobronchoscopy. 34
3.7 Workflow for the lumen centre processing. 35
3.8 Workflow for tracheal ring characterization. 36
3.9 Consecutive lumen segmentations (l^{k1}, l^{k2}) of the feature space, top of the image; likelihood maps (LK_l^{k1}, LK_l^{k2}) and snake computation (L^{k1}, L^{k2}) in polar coordinates in the middle images and snake computation in cartesian domain in the bottom image. 40
3.10 Workflow for the tracheal ring segmentation. 1. original image, 2. detected rings in white, candidate ones in yellow and the reference ring in red over the image in polar coordinates, 3. shows the result of ACC closing (grey) and snake completion (blue) and 4. represents the final reconstructed ring in the cartesian domain. 41

4.1 Examples of images from our database: real images (first and third) and its respective ground truth (second and four). 46
4.2 Lumen dataset. Examples of bronchoscopy (three first rows) with one lumen (first row), multilumen or bronchial lumens (second row) and no lumen presence (third row). 47
4.3 Example of our database: bronchoscopy image (first), expert 1 groundtruth (second) and expert 2 groundtruth (third). 48
4.4 Tracheal ring dataset. 49

4.5 Examples of our obstructed lumen database: real obstructed lumen frames (first column) and its manual segmentation (second column) made by one of the experts. 50
4.6 Obstructed lumen dataset. Some examples of bronchoscopy frames used to validate our method of obstructed lumen segmentation. 51
4.7 Examples of our reference ring database: four examples of manual segmentation made by one of the experts. 52
4.8 Reference ring dataset. Some examples of bronchoscopy frames used to validate our method of reference ring segmentation. 53
4.9 Examples of our stenosis database: pair of images (each row) with area of healthy (first column) and obstructed (second column) airway (first) segmented by expert 1 (yellow line) and expert 2 (red line). 54
4.10 Stenosis dataset. Some examples of bronchoscopy frames used to validate our method of stenosis assessment. For each patient pair of images (each row) of healthy (first and third column) and its respective obstructed airway (second and fourth column). 55

5.1 Qualitative lumen centre detection results. Good detections marked with green crosses and bad or missing ones with green circles. 58
5.2 Results for each sequence: sensitivity (a) and precision (b). 59
5.3 Inter-observer variability in manual segmentations. 61
5.4 Visual Assessment of Automatic Ring Detection. 61
5.5 Qualitative results for each sequence. Green lines are the groundtruth and blue lines our detection. 62
5.6 Multicomparison plot for the train set of the lumen segmentation. 63
5.7 Multicomparison plot for the test set of the lumen segmentation. 64
5.8 Qualitative results for the lumen segmentation with l_k parameter fixed in $k = 21$. 65
5.9 Multicomparison plot for the train set of the reference ring segmentation. . . 66
5.10 Multicomparison plot for the test set of the reference ring segmentation. . . . 66
5.11 Qualitative results for the tracheal ring segmentation with interval parameter $[\sigma_j \ \sigma_{j+k}]$ fixed in $j = 2$ and $k = 15$. 67
5.12 Multicomparison plot for the test of stenosis assessment. 68
5.13 Qualitative result for the calculation of the stenosis index. Normal airways in the first row and its paired abnormal frame in the second row. Blue lines correspond to the automatic method and yellow triangles and red squares correspond to the ground truth of the different experts. 69

6.1 Advantages of combining appearance and geometry (1st row) of the tracheal rings compared to intensity-based approaches (2nd row, images courtesy of [55]). 73
6.2 Example of approximate ring annotation for the two experts (yellow and red lines) and the ring detected by our methodology. 76

List of Tables

4.1 Description of the groundtruth of the lumen database. 46
4.2 Main features of our dataset: Scale ranges and bronchoscope type, pathologies and resolutions. 47

5.1 Precision and recall results on lumen centre detection. 57
5.2 Comparison to Inter-observer Ranges. 60
5.3 Wilcoxon signed-rank test and confidence intervals for difference in means between AUT and IO and $p - val$. 60
5.4 Statistical analysis results for the lumen segmentation method. 64
5.5 Statistical analysis results for the ring segmentation method. 65
5.6 Statistical analysis results for the systematic stenosis quantification. 67
5.7 Description of tracheal stenosis quantification. 68

6.1 Description of lumen database. 72
6.2 Precision and recall results on lumen centre detection in colonoscopy. 72

Chapter 1

Introduction

1.1 Motivation and Goal

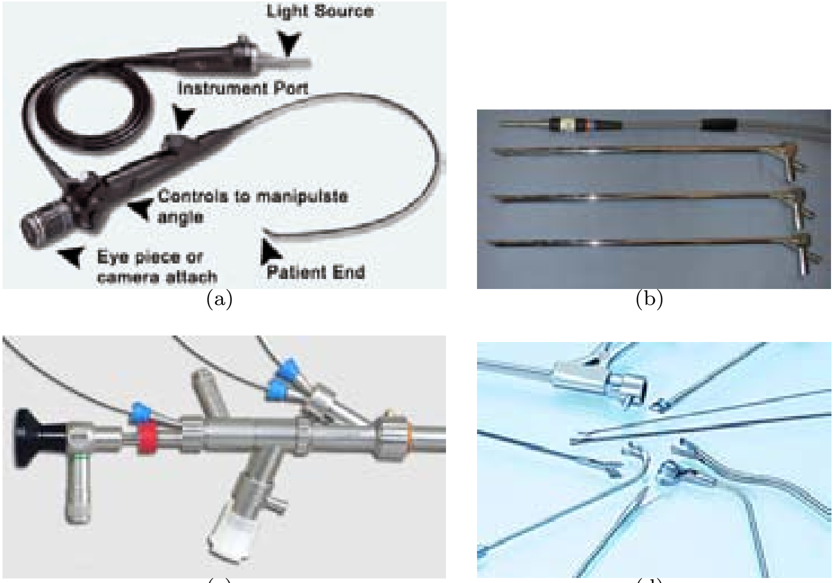
In-vivo exploration of human internal organs provides valuable information for pathology identification and diagnosis. The introduction of a tube inside the patient orifices has been known since the ancient Greece to be a simple way to explore diagnose diseases. However, its use was restricted to very specific (annal) exploration due to a lack of significant development of the very first rudimentary endoscopic devices.

Philip Bozzini was the first to create a rigid tube known as a Lichtleiter (light guiding instrument) for studying the urinary tract, rectum and pharynx in 1806 (Fig. 1.1 (a)). In 1853, Antoine Jean Desormeaux introduced an improved Lichtleiter of Bozzini to the body of a patient. This instrument consisted of a system of mirrors and a lens, in which the improved light source consisted of a lamp flame (Fig. 1.1 (b)); the endoscope burned a mixture of alcohol and turpentine. Liechester s was mainly used for urologic cases. Its main drawback was the apparition of burns inside the patient. For many, Desormeaux is considered the Father of Endoscopy . He was the first to use the term Endoscope [1] [2].

One common feature of the before mentioned devices is that all of them were rigid. In 1932 Dr. Rudolph Schindler created the first flexible gastroscope which consisted of a flexible

(a) (b) (c)

Figure 1.1: Optical endoscopes : Rigid models (a,b), flexible model and detail of the light bulb attached at its tip (right) (c).



(a) (b)

(c) (d)

Figure 1.2: Bronchoscopic instruments: Flexible bronchocopy (a), rigid tubular device (b), interconnection channels (c) and instruments (d).

device with several lenses located along the tube and a mini-light attached (Fig. 1.1 (c)) [1]. This gastroscope allowed to bend the tube (with 75 cm of length and 11 mm of diameter) some degrees, which lead to an improvement over the rigid device's since it allowed the physician to look at more directions from the same place. From this moment on, research was oriented to develop more sophisticated gastroscopes by building very small lenses, using stronger light sources, and investigating new materials to develop more flexible tubs. (see Fig.1.1).

Since the introduction of High Definition Camera Compression (HDCC) in the early 90's [40], endoscopic devices have been significantly improved in order to provide detailed high resolution information about humans internal anatomy using scopes of reduced size. Such an improvement in endoscopic quality and size has generalized its use in a wide range of minimal invasive procedures, such as laparoscopic surgery, colonoscopy diagnoses or videobronchoscopy.

Videobronchoscopy is the endoscopic technique that allows interactive navigation inside the respiratory pathways and facilitates the performance of minimal invasive interventions such as tracheobronchial procedures. Bronchoscopic procedures are performed routinely, with about 261 millions of interventions around the world per year. Videobronchoscopy consists of visualizing the inside of pulmonary airways for diagnostic and therapeutic purposes. A bronchoscope has a fiberoptic system that transmits an image from the tip of the instrument to an eyepiece or video camera at the opposite end (see Fig. 1.2 (a)). The tip of the instrument can be oriented using Bowden cables connected to a lever at the hand piece, allowing the practitioner to navigate the instrument into individual lobe or segment bronchi for disease diagnosis [51] [3]. Aside, the scope can be inserted into a rigid tubular device (Fig. 1.2 (b)) that includes different channels (Fig. 1.2 (c)) for introducing the instrumentation (such as needles, small forceps, suctionary aspiration tools) needed for a given

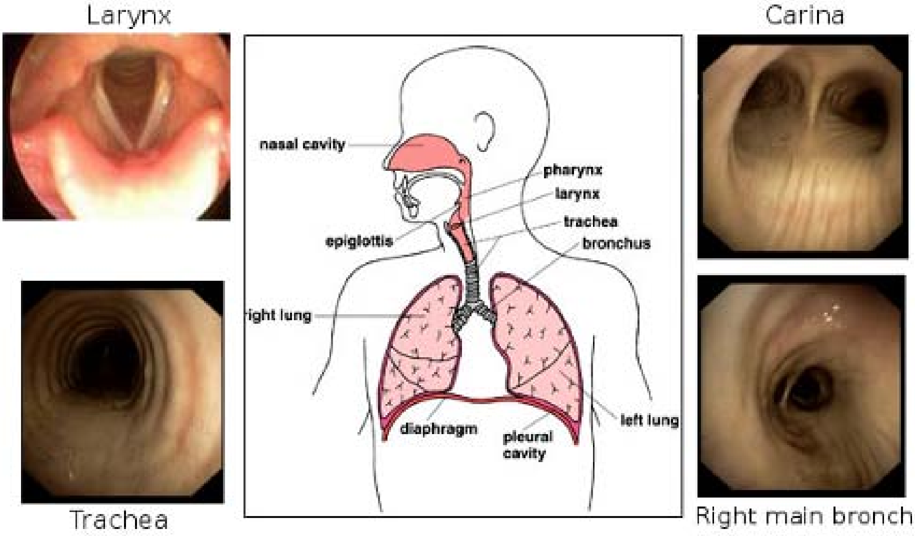


Figure 1.3: Different regions of the bronchoscopy process.

intervention procedure (Fig. 1.2 (d)).

During the exploration process, videobronchoscopy allows interactive navigation across respiratory airways (see Fig. 1.3) as the physician operates the bronchoscope from the patient s nose or mouth to the main airway anatomical segments [3]. There are four main explorations areas: larynx, trachea and main bronchi. The larynx (upper-left image in Figure 1.3) is the first segment explored in brochoscopy as the scope enters in the subglottic area and goes through the vocal cords. The next segment is the tracheal path from larynx to the main bronchi bifurcation (known as carina). The trachea (lower-left image in Figure 1.3) is almost a straight pipe composed of several equidistant tracheal rings in which are semicircular structures that are concentric in brochoscopy frames. The area enclosed by tracheal rings is the air path and is called lumen. The trachea ends at the carina, where it divides into two main bronchus as shown in top right image of Fig.1.3. Main bronchi are the second most explored area after the trachea. Starting at the carina the left/right main bronchus (top-right image in Figure 1.3) is entered just by twisting the wrist (control device) to the left/right and advancing for 1 -2 cm. From this point, and only for flexible bronchoscopy, the doctor can achieve four or five levels of the bronchial tree (named as secondary bronchi) depending on the diameter of the scope. From these levels of bronchi it branch into smaller tubes, known as bronchioles where just ultra-thin bronchoscopes can reach (lower-right image in Figure 1.3).

The depth in airways levels that current devices can achieve with high quality image recording and the development of instrumentation of minimal size have enlarged the number and diversity of diagnosis and intervention protocols formed with bronchoscopes:

- **Diagnostic Explorations**. Bronchoscopy allows direct visualization of the airway interior with minimal risk of trauma and discomfort for the patient. Therefore, it is a unique tool used to investigate the source of obstructive diseases (such as bronchitis, asthmatic inflamation or tracheal tumours) causing a narrowing of airways, which leads to increased work of breathing. The narrowing of airways is known as stenosis. Depending on the stenosis degree (slight, moderate or severe) different treatments will

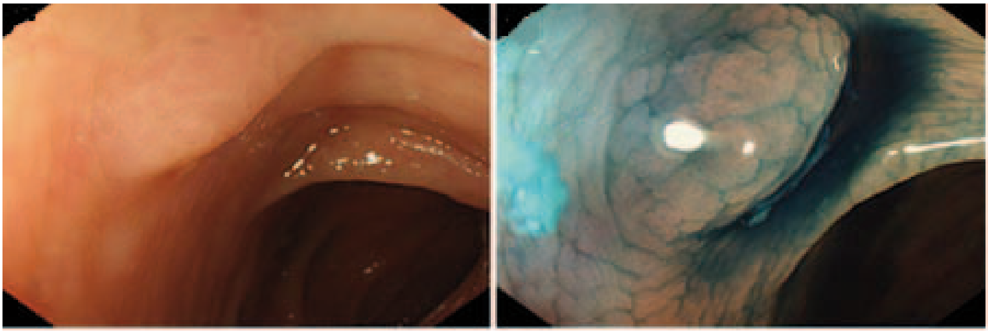


Figure 1.4: White light image with a tumor (right) and same image with a NBI light (left).

be applied. In particular, severe obstructions might require placing a stent or, even, a prothesis at the injured segment [34]. Therefore, accurate measures are key for the diagnostic yield and further treatment. Visual assessment of the stenosis degree in videobronchosocpic explorations has been reported to suffer from almost 50% of misclassification regardless of the bronchoscopist expertise [41].

Another potential use of the bronchoscopic internal view of airways is the identification of tumor regions. However, in white light standard bronchoscopic videos this is a challenging task that requires the extraction of a tissue sample for biopsy analysis. This fact has motivated the development of Narrow Band Imaging (NBI) [56] imaging technology. NBI is an optical image enhancement technology that enhances vessels in the surface mucosa by employing the light absorption characteristic of hemoglobin at a specific wavelength. NBI uses two types of narrow spectrum light: 415 nm light (blue) which is absorbed by capillary vessels in the surface layer of mucosa, and 540 nm light (green) which is absorbed by blood vessels located below the capillary vessels in the surface layer of the mucosa. In Figure 1.4 we can observe an example of NBI. Unfortunatelly, NBI is not available for all commercial devices (it is exclusive of Olympus devices) and, thus, its use is restricted to experimental assays.

- **Tracheal Procedures**. Rigid bronchoscopy may also be used therapeutically for several procedures such as for removing inhaled foreign bodies, resecting tumours or placing stents to overcome severe airway obstruction. Among the usual interventions, tracheal implants are one of the most expensive (average cost of 2200 euro per intervention) and time-consuming ones [37]. In order to choose the most appropriate prothesis the physician has to accurately estimate the diameter and length of the obstructed tracheal segment. As for the case of stenosis assessment, a main limitation is related to visual inspection impact on obtaining accurate measurements. Moreover, in this case there is an extra source of variability related to clinicians having to infer 3D measurements of the prosthesis length by only means of a simple inspection of the 2D video frames (2D images in perspective projection). The determination of the true 3D length strongly depends on the experience and the anatomical knowledge of the doctor. Thus, not having objective measurements could imply a remodelation of the prosthesis to fit better the patient, with the additional cost that this supposes.
- **Biopsy Collection**. Another usual intervention is focused on biopsying tumor regions. In order to achieve this a reliable navigation path that leads the clinician through bronchial system to reach the tumour is needed. Navigation guidance is only

done by using confocal technology and its use is currently at its experimental stage.

Although the use of bronchoscopy is spread among clinical facilities it presents some drawbacks, being the visual inspection for assessment of anatomical measurements the most prevalent of them. Measurement extraction is required for customization of tracheal stents and implants and stenosis diagnosis. In the first case, errors in the measures obtained by visual exploration choose the wrong prothesis in 25% of the cases, which might lead to the repetition of the intervention [47,62]. Being 2200 euro the average cost of tracheal procedures, this repetition rate has an impact in the healthcare system efficiency as some of the resources dedicated to bronchoscopy are not used to explore new patients.

Concerning stenosis, a recent study [41] demonstrates that its subjective assessment is highly prone to errors and that the quality of this assessment does not depend on bronchoscopist experience. There are two type of errors identified related to whether the degree of stenosis has been under or over estimated. In the first case such underscoring has several consequences to both the patient and the healthcare system. The patient will be misdiagnosed and a consequently an new exploratory intervention will be necessary. This implies a delay in providing the adequate treatment to the patient which compromises its safety. In particular this problem appears for almost a quarter of the measurements. Regarding the overestimation of stenosis degree, the patient will have undergone to a treatment that was not necessary incurring in a waste of resources. As it can be seen in both cases there is an economic cost associated either to the repetition of the intervention or to the waste of resources by providing an erroneous treatment. Finally there is an emotional cost of the patient that should also be taken into account which may affect the reputation of the clinical institution in charge of the patient.

Therefore, a method for objectively quantifying the percentage and size of the obstructed airway in the operating room would allow immediate and accurate diagnosis and treatment [42,52]. Stenosis assessment is achieved by comparing the area of a reference healthy airway to the area of the obstructed segment (as illustrated in Fig.1.5). The area of the healthy segment is defined by the most external and complete tracheal ring, while the degree of obstruction is given by the luminal area. Thus, a system aiming at a deployment in clinical

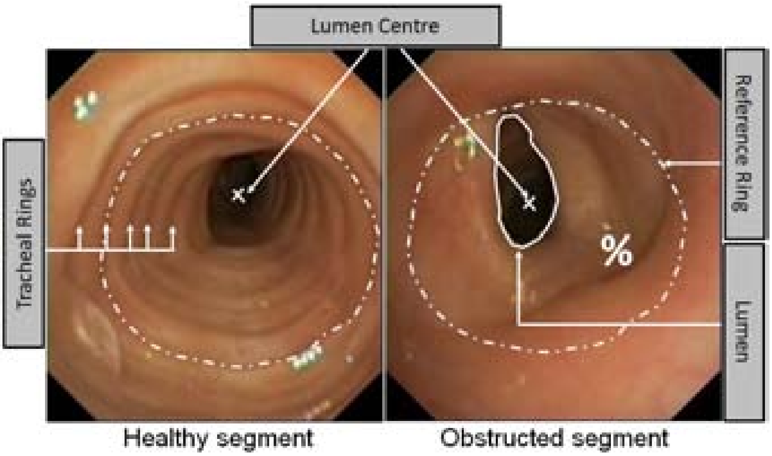


Figure 1.5: Graphical example of the desired output of our method. The healthy reference ring on the left image and compared to the obstructed lumen show on the right image.

practice should be able to characterize on-line the main tracheal structures needed for the measurements (lumen and tracheal rings) from diagnostic videos in the operating room.

1.2 State of the art and challenges

Although some methods have already been proposed either in the literature or as commercial solution, we have to take into account that videobronchoscopy is a relatively new technology (it appeared in the middle 90 s) and consequently there is not much computational framework yet. In this section we review existing works (either hardware or software) which aim to improve the stenosis diagnostic yield. Stenosis degree can be assessed using either Computed tomography (CT) scans or videobronchosocpy explorations.

Computed tomography (CT) imaging, also referred to as a computed axial tomography (CAT) scan, involves the use of rotating x-ray equipment, combined with a digital computer, to obtain images of the body. Using CT imaging, cross sectional images of body organs and tissues can be produced. Though there are many other imaging techniques, CT imaging has the unique ability to offer clear images of different types of tissue. CT imaging can provide views of soft tissue, bone, muscle, and blood vessels [59]. However, the quantification of airway narrowing is difficult in some patients for several causes such as, dyspnoeic or uncooperative patients [32], patients with cognitive dysfunction and patients hospitalized in the intensive care unit. Another important problem is that CT scan visualization is affected by the secretions which could be confused as airway narrowing [50].

Nowadays CT imaging is used in bronchoscopy to create a planning of the intervention in a methodology that is known as virtual bronchoscopy (VB). Unfortunately VB does not directly allow to locate the clinician inside the trachea, it does not also consider patient breathing dynamics or significant intrasubject variations in airway lumen measurements across sites [46] and involves radiating the patient which make these test poorly suited for repeat testing, specially in younger patients [60]. Finally, virtual bronchoscopy is not practically feasible in intervention time as processing a stack of CAT images is computationally hardware costly and the need of experienced interpreter of radiographical findings limits their application in daily practice [45].

Current methods for measuring the size of tracheal airway using bronchoscopy can be split into contact and non-contact procedures:

Contact procedures determine the diameter of tracheal rings by inserting endotracheal tubes with increasing stepwise size [25, 43]. The endotracheal tube is inflated to fit the tracheobronchial airway to determine the size of the prothesis. A main concern is that the introduction of endotracheal tubes can cause lesions to the soft tissues of the tracheobronchial airway and also the patient has to be intervened two times: first for usual diagnostic and second for the insertion of the endotracheal tube. Another concern is that such semi-quantitative methods choose a wrong prothesis in a 25% of the cases, which leads to the repetition of the whole procedure [47, 62]. Such procedure repetition is given because a wrong positioning of the endotracheal tube or the low precision of the inflated tube is 5mm.

Existing *non-contact procedures* can be divided according to their application into bronchoscopic device improvement procedures and videobronchoscopic computer image analysis ones. In the first case, a suitable modification of the bronchoscope optics ([23, 28]) can provide accurate measurements with micrometer precision. However, these technologies are still experimental and far from being commercially available. Regarding stenosis assessment from videobronchoscopic computer analysis, it requires reliable extraction of the main anatomical structures (lumen and tracheal rings) involved in the computation of the degree of stenosis. The area of the healthy segment is defined by the most external and complete tracheal ring,

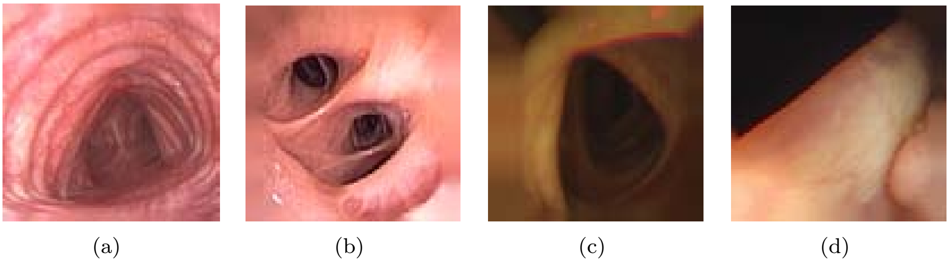


Figure 1.6: Examples of variability in lumen appearance: single (a) and multiple (b) bronchoscopy image; centred (c) and biased (d) colonoscopy image.

while the degree of obstruction is given by the luminal area.

The majority of the relevant work in lumen localization and detection is related to gastrointestinal image analysis. Under the assumptions that the largest dark blob of the images usually correspond to lumen [58] and it is always present in the images [48], there are several works that segment the lumen using a region growing approach over the image grey level [4]. These approaches are accurate as far as the initial seed for the region growing is placed inside the luminal area and their performance decrease in the presence of shadows or low contrast images. Recent approaches use contrast changes to account for local differences in image intensity. The authors in [17] characterize the luminal region in wireless capsule videos by means of Haar features followed by a supervised boosting for detecting the probability of having the lumen in a given frame. A main drawback for its application to standard bronchospcopy procedures is that its usual central navigation illuminates the luminal area and, thus, reduces contrast changes (compare images in figure1.6(a) and (b)).

A common limitation is that most methods can not handle having more than one lumen in an image, which is quite frequent in bronchoscopy videos. The recent approach in [65] detects multiple lumen areas by using mean shift [9]. Although it provides information about multiple lumen, it might fail in the absence of any luminal area and it has a high computational cost not suitable for its use in intervention time. Other approaches for multiple lumen detection in bronchoscopy [11, 14, 35, 36] are semi-automatic procedures which are applied off-line.

Concerning, the works on tracheal rings segmentation in videobronchoscopy, they are even more sparse. As far as we know, there is no specific processing tools on tracheal ring segmentation and the only work is the preliminary study reported in [55] assessing the feasibility of a reliable extraction of rings. This study identified the main issues that make segmentation of tracheal structures (luminal area and rings) using image processing methods a challenging task. Such artifacts were grouped into:

- **Challenge 1: Appearance Artifacts**.

  The illumination is not the same for all the images from a given video due to the light of the camera is not always in the same position. This implies that in some cases, some parts of the trachea are more illuminated while others present more shadows (see Figure 1.7 for an example of non-uniform illumination). In this way, specular highlights (bright spots of light) can be a problem as their appear as a result of the illumination on shiny objects. Another issue is the blurring of the images (see Figure 1.7). This effect appears when the camera moves too fast also when the camera objective fogs because of the breath of the patient. In this way, interlacing appears when the camera is also moving fast. Interlacing consists of two sub-fields taken in

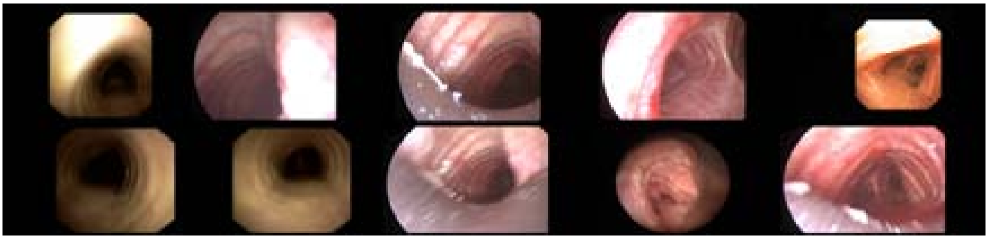


Figure 1.7: Image quality artefacts: first row : ilumination problems, second row: blurred problems.

sequence, each sequentially scanned at odd and even lines of the image sensor [10]. If camera does not moves fast you cannot see the effect, but in our case camera is always moving and sometimes is moving fast, so in some of the frames interlacing appears (borders of the structures in the image are indented) (see Figure 1.8). Finally taking a look at the main anatomical structures such as tracheal rings and lumen, in both cases we can have different appearance due to different image acquisition methods or different appearances across the same sequence. In the case of tracheal rings image intensity appearance is not enough to discriminate rings to other structures like veins or carina structures.

- **Challenge 2: Alien Artifacts**. Videobronchoscopy scene can contain some elements which may difficult tracheal structure characterization: for instance we can have elements such as mucus, blood and bubbles which are substances or liquids that can appear in human trachea or structures such as blood vessels which are present always in the trachea. Regarding blood vessels, they can have in some cases the same appearance of the intensity valley that appears between tracheal rings. When the bronchoscope is getting to the end of the trachea, the beginning of the bronchial system starts to be seen (carina). This structure also contains rings which are related to the main bronchis and consequently their identification is out of our target and

Figure 1.8: Interlacing problem example.

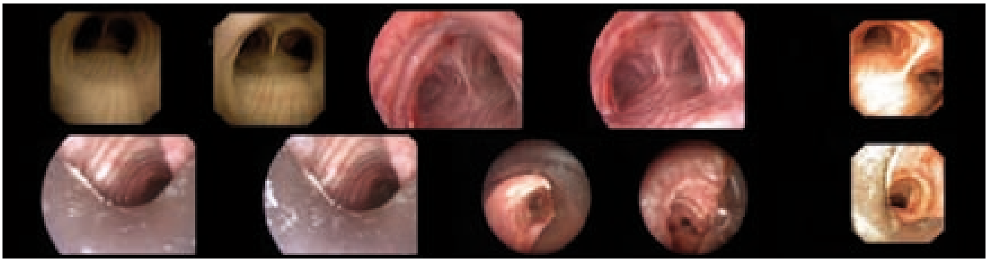


Figure 1.9: Non anatomical artefacts: first row : carina artefact, second row: rigid bronchoscope artefact.

their potential detection should be discarded. Finally in some images some part of the instrumental is present (rigid bronchoscope). The presence of this instrumental also conveys with valley information which may alter the performance of our methods.

- **Challenge 3: Geometric Artifacts**. In a 3D-2D projection, the 3D geometry of the object can be severely distorted depending on the point of view. If camera is centred in the axis of the trachea the tracheal structures can be easily defined. Out of center deviations introduce two main artefacts in 2D images. The first type of artifacts happens whenever the camera is focused on the trachea wall (see Figure 1.10). In this case, we cannot see the tracheal rings or we can only see part of them. We also have to consider that for stenosis assesment we also need to have a clear and complete view of the lumen. The second types is due to tracheal rings appearing as collapsed due to the projection (see Figure 1.10). Since the camera is not well centred in the tracheal axis, all the tracheal rings that we can see in the image are collapsed in a certain (close to the camera position) point. Another geometrical issue is related to the number of luminal area. When the camera is through the trachea only one lumen is present in the scene whereas when the camera is passing through bronchial system we can have several lumens in the image.

1.3 Contributions

The goal of this thesis is the development of an automatic method to estimate the degree of stenosis in brochoscopy images. In order to achieve this and following clinical practice, we need an accurate characterization of tracheal structures -tracheal rings and lumen- which

Figure 1.10: Geometric artefacts: first row : trachea wall, second row: collapsed tracheal rings.

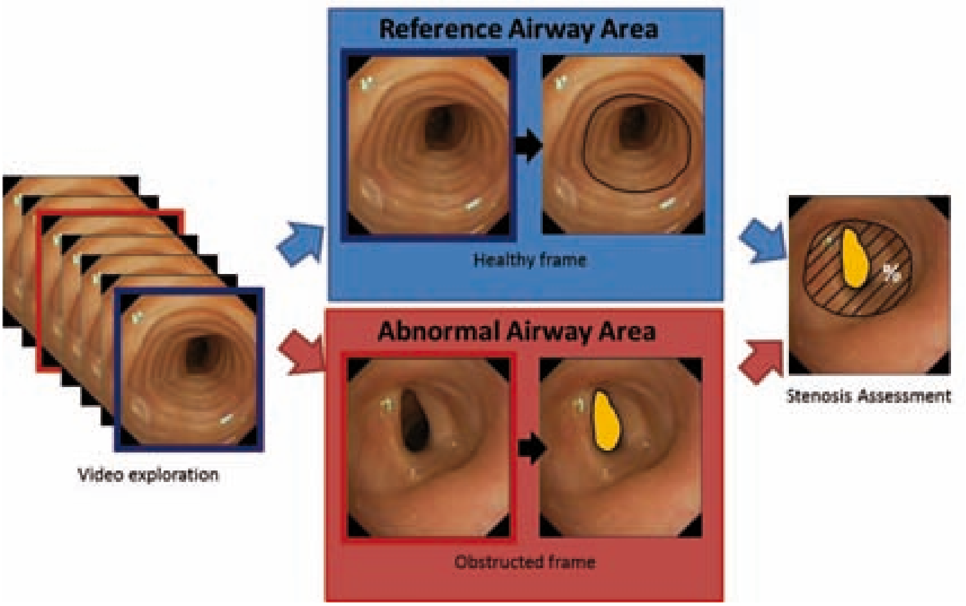


Figure 1.11: Graphical example of the process and desired output of our method. The healthy reference ring calculation on the top process and compared to the obstructed lumen process show on the bottom process.

play a key role in the process (see Fig.1.5). We introduce in Figure 1.11 the general processing scheme for stenosis assessment.

The main contributions of this thesis are:

- **Contribution 1: System for automatic stenosis degree assessment**

  Stenosis degree is defined as the percentage of trachea obstructed and consequently, our methodology needs of a definition of a healthy and an obstructed frame from which to calculate the final percentage of obstruction. Tracheal rings characterization is used mainly to characterize the healthy airway size whereas lumen characterization is needed for the obstructed airway size. Finally we compute the stenosis degree by completing and selecting the reference ring for the healthy frame and segmenting the obstructed lumen for the obstructed one.

- **Contribution 2: Appearance and geometrical model for characterizing anatomical tracheal structures**

  In order to obtain a reliable characterization of the needed tracheal structures we have developed both physical and computational models for tracheal ring and luminal area. These models combine geometrical and appearance features that describe those anatomical structures in a way that discriminates them from other structures that we have already presented in the challenges part.

- **Contribution 3: Segmentation of main anatomical structures**.

  We propose a multi-step strategy for reference ring segmentation and obstructed lumen segmentation in order to select the best parameters for its segmentation. For this reason we also propose a novel statistical framework based on Analysis of Variance (anova) [39] for selecting the optimal parameters which ensure within inter-observer variability performance.

- **Contribution 4: Validation protocol and public databases**.

  One of the main drawbacks of all the proposed methods is the difficulty when comparing their performance. Up to our knowledge, there is no public annotated database of tracheal rings, lumen regions or stenosis assessment. This constitutes a major flaw for the development of generic algorithms able to achieve accurate results in a wide range of images. We introduce of different fully annotated databases to validate the performance of the developed methods.

The remains of this thesis are structured as follows: Chapter 1 includes the introduction to to videobronchoscopy by following with the state of the art, challenges and contributions of the thesis. Chapter 2 is dedicated to the introduction and explanation of the models used to describe the endoluminal scene and Chapter 3 presents the methodology and computer vision tools used to use these models and segment the important endoluminal scene elements for a final stenosis assessment. Chapter 4 introduces the databases, validation protocol and a description for each experiment that have been carried out for the evaluation of the used methods in the scope of the research. The results (quantitative and qualitative) of this experiments are explained in Chapter 5 and the discussion of its validation, limitations and applicability is done in Chapter 6. Finally, Chapter 7 closes the thesis by exploring the main conclusions that can be extracted along with sketching the opened future lines of research.

Chapter 2

Geometric and Appearance Model of the Trachea

Stenosis assessment in diagnostic bronchoscopic explorations requires a reliable extraction of trachea main structures (lumen and tracheal rings). In order that structure identification is robust against the variability of artefacts that are present in exploratory bronchoscopic videos, the image processing algorithms used for their extraction will be based on geometric and appearance features that characterize them from other structures of the tracheal scene. To do so, we will first characterize tracheal structures in videobronchoscopy on the grounds of the physics that are involved in the bronchoscopic exploration. Such physical model will provide with a description of anatomical structures that discriminates them from other elements of the tracheal scene, such as intervention instrumentation and blood vessels. Second, physical model describing tracheal structures features in videoframes will be implemented using image operators able to cope with videobronchoscopy appearance and geometric variability.

In this chapter we describe the geometric-appearance model of tracheal structures based on acquisition videobronchoscopic physical principals. First, we describe our physical model from how videobronchoscopy frames are acquired, and, specifically, how main structures appear in the projected frames. Second, we derive the image processing operators that best match the physical description of the main structures (lumen and tracheal rings) in the bronchoscopic operations.

2.1 Physical Models

Trachea geometry, camera position, orientation and illumination, Phong s illumination model and the light attenuation are the bases for modelling trachea images using a bronchoscope. In this section, we will first describe the tracheal anatomy reported in morphological and histological studies, second its geometric features after conical projection from the scope camera point of view and, finally, its intensity in images using a spot light source located at the scope camera. In order to validate our physical model, we will compare synthetic images generated using our camera-illumination model on two tracheal geometric models to bronchoscospy video frames. After its validation, we will use our physical model to characterize the geometry and appearance of tracheal structures in videobronchoscopy.

Anatomically, the trachea is a tube lying in front of the esophagus that connects the

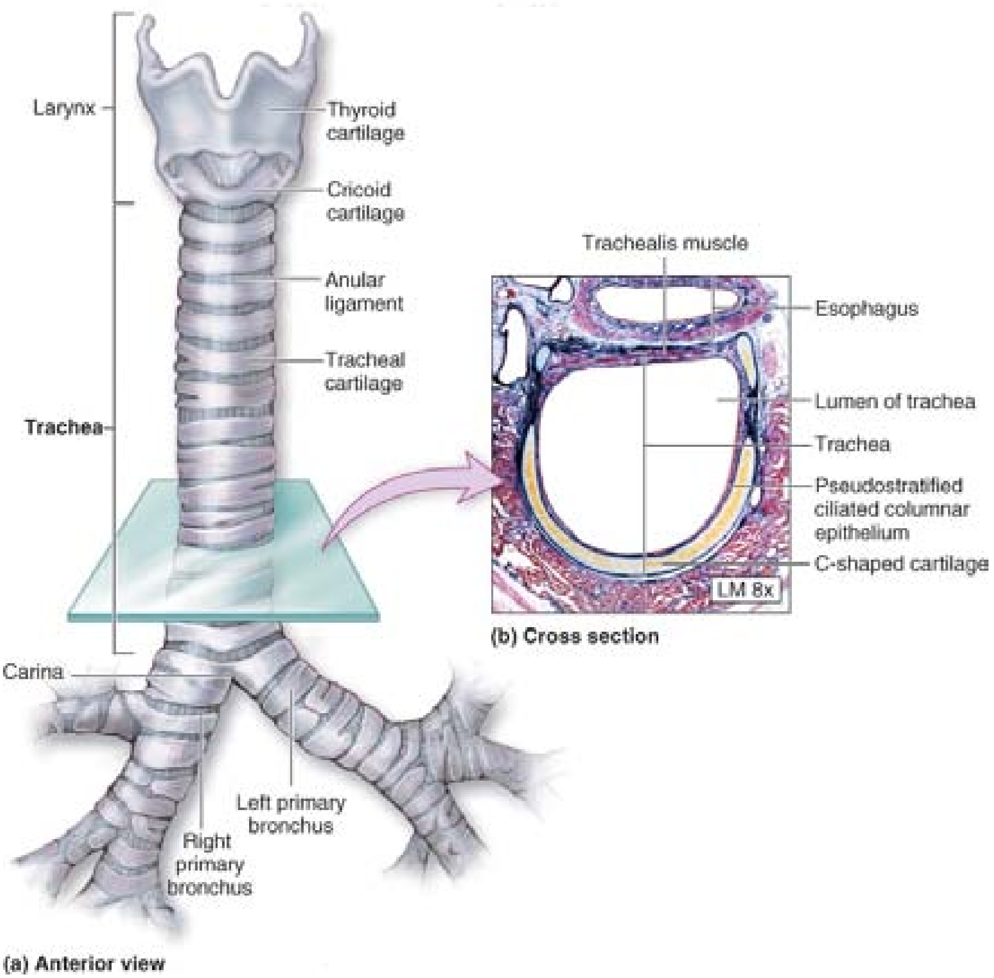
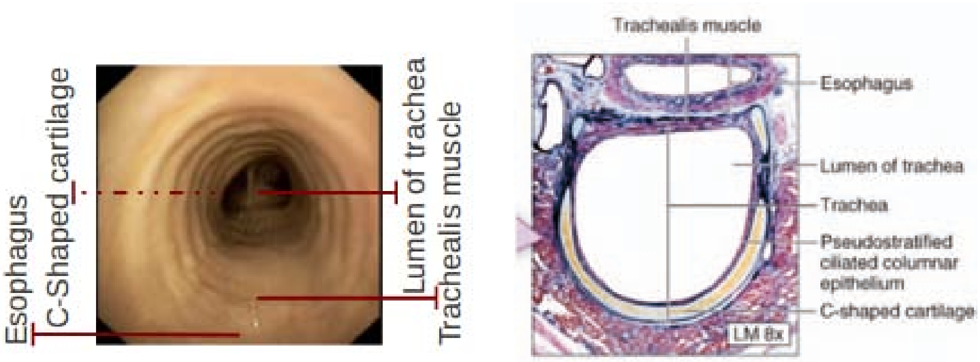


Figure 2.1: Trachea schema (left) and main anatomical elements in an histological cut.[2]

pharynx or larynx to the bronchial system (as shown in the sketch in Fig.2.1). The trachea has an inner diameter of about 21 to 27 millimetres (0.83 to 1.1 inches) and a length within 10 and 16 centimetres. It begins at the larynx in the fifth level vertebra and bifurcates into the primary bronchi at the vertebral level of T4/T5 [21]. A main structural element are tracheal rings [21]. There are about fifteen to twenty incomplete C-shaped cartilaginous rings that reinforce the anterior and lateral sides of the trachea. The aim of the tracheal rings is to protect and maintain the airway equally span along the tracheal tube. The trachealis muscle connects the ends of the incomplete rings and contracts during coughing, reducing the size of the lumen of the trachea to increase the air flow rate. The cartilaginous rings are incomplete in order to allow the trachea to collapse slightly so that food can pass down the esophagus (posterior to the trachea) [21]. The histologic cut of Fig. 2.1 shows the main anatomical elements of a tracheal section: C-shaped rings, interconnectivity muscle and trachea lumen.



(a) (b)

Figure 2.2: Composition of tracheal structures in a videobronchoscopy frame (a) an in a histological cut (b).

The main structural elements observed in histological cut are also observed in videobronchoscopy frames as the clinician navigates through the tracheal lumen, as it can be observed in an actual frame -Fig 2.2-. However, we would like to note that bronchoscopy frames are acquired using a C-MOS video camera [63], thus, 3D structures are projected using a conical projection. Such projection might distort the appearance and geometry that trachea anatomy has in cross sectional histological cuts. The geometric appearance of the trachea tubular structures in conical projection depends on the scope camera position and point of view. In case the camera position is central to the lumen and it is oriented along the tracheal medial axis, the conical projection looks like histological cross sectional cuts. Tracheal rings appears as a set of concentric circles that preserve their c-shape, the trachialis muscle that closes them at the esophagus region is flat and lumen corresponds to the darkest central circular shape. Fig 2.2 compares the geometric appearance of the main structures, between a bronchoscopi frame taken in a central projection and a cross sectional histological cut. Moreover, a vertical cut of the trachea allows us to see how the C-shaped cartilaginous rings looks like, which is needed for an accurate tracheal structure characterization. Considerable deviation of the camera from central navigation distorts the cross-sectional geometry. Tracheal rings are not concentric any more and collapse in a certain pathway point and lumen can be partially occluded by other structures, thus, loosing its circularity. In this case, stenosis measurements would not be reliable and, thus, these images are automatically discarded by experts (see Fig.2.3).

The intensity of tracheal structures in videobronchoscopy frames depends on the camera illumination and the incident angles between camera light and trachea surface. The bronchoscope has a spot light source and a camera attached to it in a way that both camera and the source of light are in the same direction. This allows the modelling of the image intensity as a pinhole camera following a given illumination model. We have chosen Phong's reflection model [49], because is the most simple and basic illumination model and it is enough for having an idea on which is the reflection between the incidence light and trachea surface. Regarding the propagation of a spot light we consider that it is attenuated according the inverse of the square of the distance.

Phong's reflection model defines the illumination of each surface point as a combination

---

[1] http://academic.kellogg.edu/herbrandsonc/bio201_mckinley/Respiratory%20System.htm

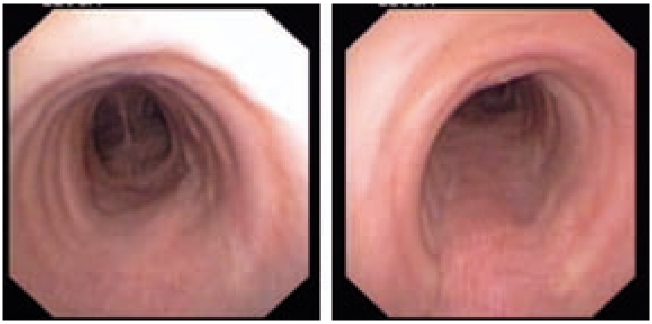


Figure 2.3: Frame valid for stenosis assessment (left), frame not valid for stenosis assessment (right).

of the diffuse reflection of rough surfaces with the specular reflection of shiny surfaces. It is based on Bui Tuong Phong's informal observation that shiny surfaces have small intense specular highlights, while dull surfaces large highlights that fall off more gradually. Phong's illumination model also includes an ambient term to account for small amount of light that is scattered about the entire scene. The graphical example of Fig.2.4 illustrates the decomposition of the light reflected into the three components: ambient, diffuse and specular. According to this model, the image intensity at a pixel is given by the cosinus of the incident angle between light ray and the surface normal as:

$$I = I_a K_a + fatt I_p (K_d \cos\theta + W(\theta) \cos\alpha) \tag{2.1}$$

where I stands for the light reflected by the surface to the camera, I_a for the intensity of ambient illumination, K_a for the ambient reflection coefficient, $fatt$ for the attenuation factor of light distance, l_p for the spot light intensity and K_d for the diffuse reflection coefficient. In this case, θ determines the angle between the normal surface and the direction of the incidence spot light whereas α represents the angle between the normal surface and camera direction. Finally, $W(\theta)$ is used to represent the specular reflection coefficient.

By applying Phong's illumination model to a 3D models of the trachea main structures, we can derive their physical models of their intensity in a central projection. We have considered two different geometrical meshes of the trachea: computational synthetic models and a mesh extracted from a CT scan. Scope navigation along the two geometric models and Phong model illumination was simulated using 3D Studio Max (3Ds Max) to obtain the cross sectional images in conical projection.

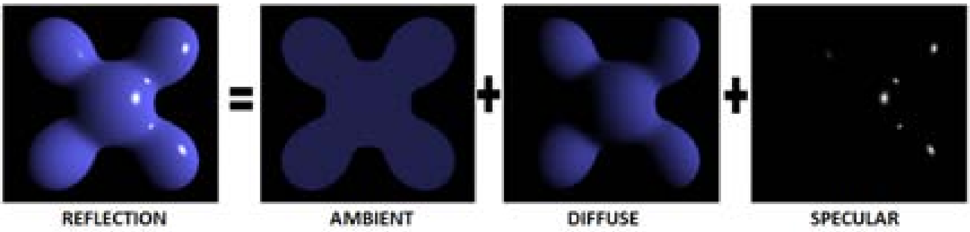


Figure 2.4: Decomposition of the light reflected by the surface into the three components: ambient; diffuse and specular.

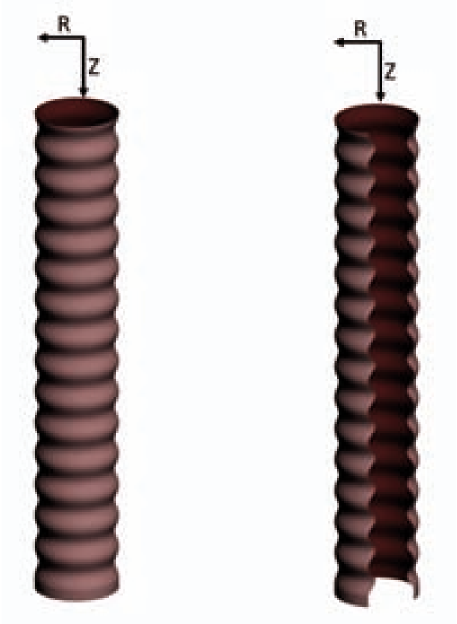


Figure 2.5: Synthetic Model of the Trachea: sinusoidal tracheal surface (left) and longitudinal cut showing the generating sinusiodal profile, (right).

In central projection, the intensity of rings is determined by the profile of a longitudinal cut, which is repeated along the c-shape angular range. Such longitudinal profile can be modelled as a suface of revolution generated by rotation of a sinusoidal curve around the z-axis. The sinusoid represents a radial coordinate (depending on z), which amplitude and frequency models the cartilaginous tracheal rings. The angle of rotation is limited to the angular range of the ring c-shape. Our computational model of the trachea is given in cylindrical coordinates as (θ, z) as a functions in the z-coordinates given by:

$$\Phi(\theta, z) = (R(z)cos(\theta), R(z)sin(\theta), z) \tag{2.2}$$

for $(\theta, z) \in [0, 2\pi] \times [0, L]$, being L the trachea average length in adults, and the radius $R(z) = Rmin + (RMx - Rmin)cos(\lambda(z/180 * \pi))$. The parameters $Rmin$, RMx are computed so that $R(0)$, $R(360/\lambda)$ achieve the average trachea maximum and minimum diameters and λ is set to have approximately the average number of tracheal rings in a segment of length L.

Figure 2.5 shows one of our computational models simulating the 3D profile of the trachea with the main parameters used for its computation depicted in the left plot. Figure 2.6 shows the Phong's reflection model applied to the synthetic tracheal surface illuminated from a central camera point of view. The left image shows the projected surface and the right plot is the intensity along a central line. Bright pixels correspond to cartilage tracheal rings, while dark shaded ones are inter-ring spacing. We observe that the width of bright zones decreases with the distance to the camera due to the perspective projection. This narrowing effect is significantly reduced for shaded areas between rings.

In order to obtain a more realistic cross sectional profile, the physical model (camera navigation and Phong illumination) describing the acquisition of videobronchoscopy frames,

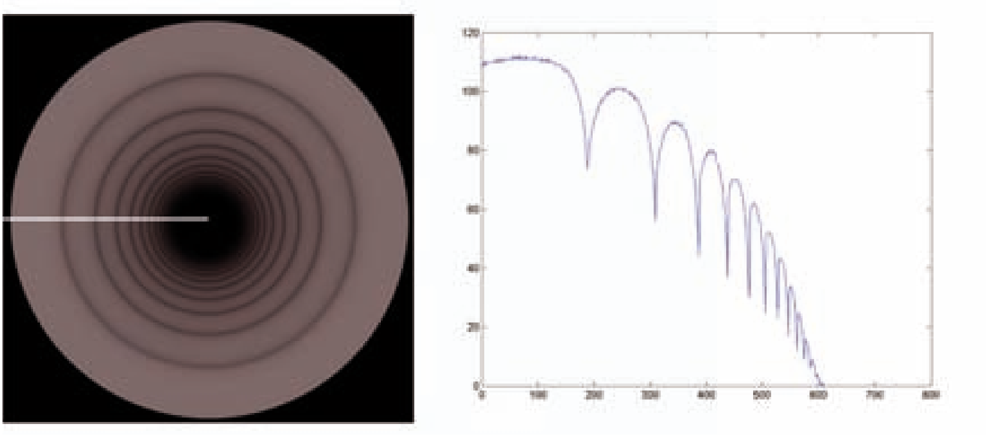


Figure 2.6: Trachea illumination Appearance in synthetic trachea: main anatomic structures in left image and radial intensity profile in right plot.

has been applied to a mesh of the airways extracted from a CT scan of an adult volunteer [3]. CT scan was acquired in axial orientation. The DICOM volume was segmented using 3DSeg. In this case, central navigation was done using the skeleton of the segmented airway volume.

Figure 2.7 shows the airway mesh segmented from CT scans and a longitudinal cut of the segmented trachea. We note that this longitudinal profile matches the sinusoidal wave used to generated our synthetic computational models. Figure 2.8 shows the Phong's reflection model applied to the segmented trachea illuminated from a central camera point of view. We can see that the intensity profile of the anatomical structures matches with the one analysed with the synthetic computational model.

The validity of our model in real images can be visually assessed by comparing synthetic images and radial profiles to those extracted from videobronchoscopy frames. Figure.2.9 shows a bronchoscopy frame and the grey intensity profile along a radial line, $L1$, depicted in white. As can be seen in the Fig.2.9, pixels in the boundaries between C-Shaped cartilage (tracheal rings) have low grey level whereas pixels belonging to the tracheal ring appear brighter. Consequently we have a jump in intensity in the border which separate consecutive tracheal ring regions which constitute an intensity valley. We can also observe a region of the image with low gray level which constitutes the luminal area which is defined by the farthest points from the camera because of the light attenuation.

The analysis of both illumination and image radial intensity profiles obtained for the synthetic and CT tracheal models, leads us to the following physical characterization of tracheal structures (rings and lumen) in videobronchoscopy frames:

1. **Tracheal Lumen Physical Model.** As illustrated in Fig. 2.9, the amount of light that falls on the scene decreases approximately according to the inverse of the square of the distance between the light source and each 3D point. Consequently the farthest parts of the image, such as the lumen, are poorly lighted. The fact that the amount of light increases from the centre of the lumen outwards allows to incorporate geometric gradient-based features to our characterization of the lumen. Our model of appearance for lumen uses the former characterizes the lumen centre as the dark region of the image

---

[3] this data is courtesy of Bellvitge Hospital

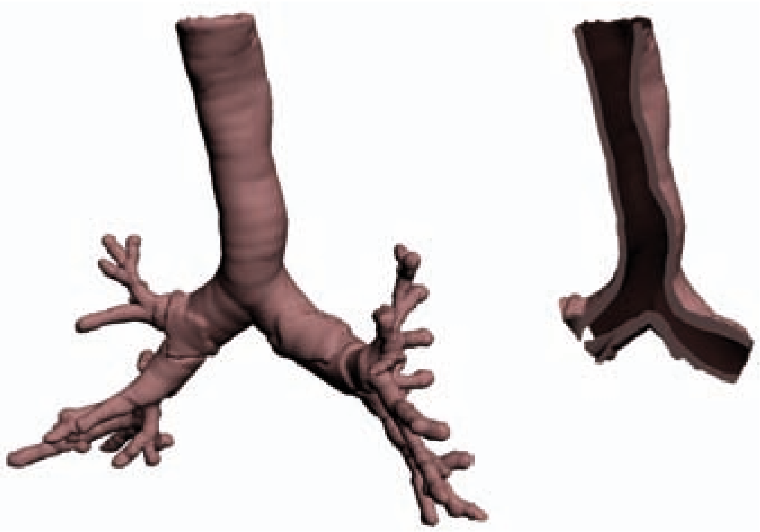


Figure 2.7: CT Model of the Trachea: bronchial system surface (left) and longitudinal cut showing the trachea profile, (right).

(appearance) which centre is the hub of image gradients (geometry). In this way, we can also define the centre of the lumen as the darkest and gradient centre point of the lumen area.

2. **Tracheal Rings Physical Model.** Tracheal rings can be locally described as a composition of valleys (between rings and shaded parts) and ridges (every ring and between valleys). However, tracheal rings are not the only source of valley information in bronchoscopy images. Other structures, such as surgical devices or other

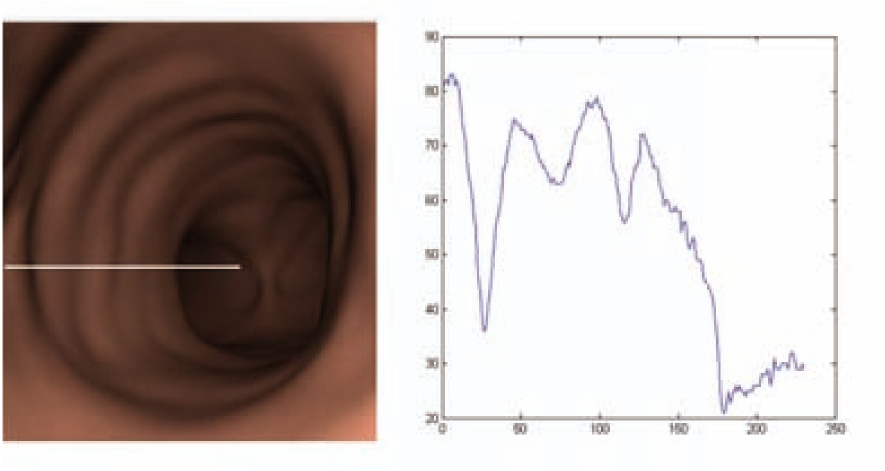


Figure 2.8: Trachea illumination Appearance in CT: main anatomic structures in left image and radial intensity profile in right plot.

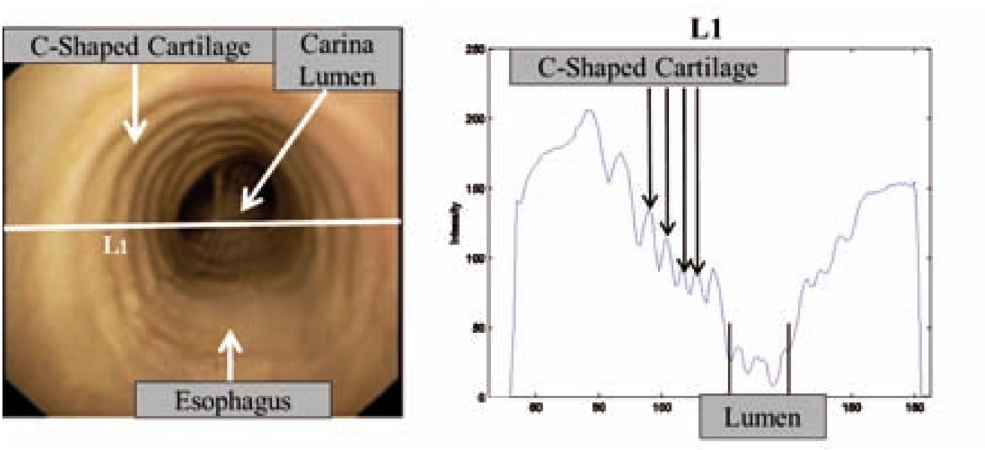


Figure 2.9: Trachea illumination Appearance in Original Images: main anatomic structures in left image and radial intensity profile in right plot.

tracheal structures such as veins or the carina, have also a similar local intensity profile (as illustrated in fig.2.10). What allows us to discriminate rings from such alien elements is their global geometric profile identified using our physical model. Using their global disposition, we will identify tracheal rings as circles with a concentric disposition around the protection axes which is the centre of the lumen. Moreover, there are two aspects referring to the valley pattern across its concentric disposition. First, the thickness of the rings decrease a far as they are from the camera, this effect appears because of the conical projection. Second, the intensity of these valleys is reduced as far as the rings are from the camera, this effect appears because of the light attenuation with the distance.

Figure 2.10: Other tracheal structures such as veins (first zoom) and carina (second zoom) that have the same valley profile as tracheal rings (third zoom).

2.2 Computational Models

The physical description of the geometry and appearance that each tracheal structure has in video frames should be implemented using image processing operators in order to systematically extract them. In this section, we provide the computational formulation of the image features that characterize lumen and tracheal rings. The computational tools are designed to minimize some of the geometric and appearance artifacts. In particular, we will define parametric operators to account for variability of tracheal structures geometry and normalised operators giving a uniform response under illumination changes.

2.2.1 Tracheal Lumen Computational Model

The luminal region is characterized as the darkest and most circular blob in images. This features are implemented using two different descriptors. Circularity is characterized based on the idea that the lumen centre is the source of all image gradients using a Directed Gradient Accumulation (DGA). Meanwhile, dark blobs are detected by means of a convolution with a Gaussian kernel, which is referred as Dark Region Identification (DRI).

Directed Gradient Accumulation (DGA) is based on the assumption that gradient lines tend to have the origin inside the luminal area, facing towards outside of the lumen. Considering this, DGA value for each point is calculated as the number of gradient-directed lines that cross it. These lines have the same direction than the gradient and they are created by extending gradient lines to cover the whole frame. Image gradient centralness is formulated using the accumulation of the gradient lines of all image pixels. This operator is called Directed Gradient Accumulation (DGA) and is computed as:

$$DGA(x\ y) := \qquad \underset{\lambda\ (x\ y)=(x_0\ y_0)+\lambda\ \ I(x_0\ y_0)}{\qquad\qquad} I(x_0\ y_0)$$

where $(x_0\ y_0) + \lambda\ \ (x_0\ y_0)$ corresponds to the parametric formulation of line through $(x\ y)$ oriented across the image gradient $\ \ I(x\ y)$, λ is the free parameter of the gradient line equation and # denoting the number of elements in a set.

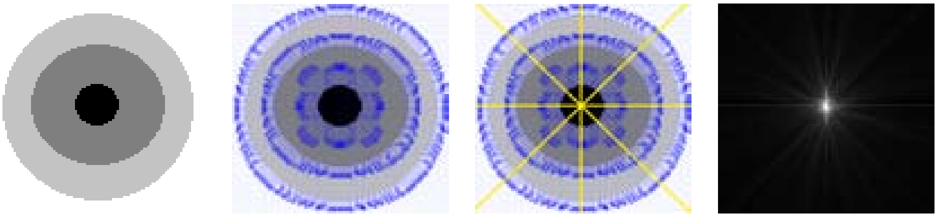
If a given image point is at the centre of a tubular structure, by Phong s illumination model, image normal lines will accumulate around this point. It follows that DGA achieves maximum values at either darker (i.e. lumen) or brighter (i.e. specular highlights) regions. The synthetic images in Fig. 2.11 illustrate how DGA works. In this example all gradient vectors are directed from the centre of the image (darkest part) to the brightest external part and, thus, DGA maximum response corresponds to the centre of the image (Fig. 2.11 (d)). It is worth mentioning that the direct output of DGA is not smooth and presents several discontinuities. In order to overcome this we apply a gaussian smoothing to the DGA output in order to obtain a more continuous DGA profile, as can be seen in Fig.2.12.

Dark Region Identification (DRI) maps are calculated by applying a smoothing using a gaussian kernel which σ is related to the scale of the lumen and it is determined using a training set. We characterize dark areas of the image using our DRI given by a convolution of the original image, $I = I(x\ y)$, with a gaussian kernel, g_σ, of size σ:

$$DRI(x\ y) := g_\sigma * I = \frac{1}{(2\xi\ )\sigma^2} e^{-\left(\frac{x^2}{2\sigma^2}+\frac{y^2}{2\sigma^2}\right)} * I(x\ y)$$

for $*$ the convolution operator.

The response to DRI enhances dark values and, thus, the luminal area. Fig. 2.13 shows the output of DRI for several scales. If we decrease the scale we increase the size of the



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

Figure 2.11: Graphical explanation of DGA algorithm: Original synthetic image (a); Corresponding gradient vectors superimposed to the image (b); Example of the extension of gradient vector lines (c), and resulting DGA accumulation map (d).

expected luminal area up to the point when this region stays stable for a given number of consecutive scales. In this case the final output of DRI will correspond with the final smoothed image with the stable scale.

As for DGA maps we show how DRI algorithm works by a graphical example. As can be seen, if we increase the scale (reducing consequently the size of the resulting dark blobs) we go from having a big dark blob (Figure 2.13 (b)) to a smaller one which matches the lumen region (Figure 2.13 (e)). As can be observed the final accumulation map could already be used as an approximation of the luminal region. This method is also valid to differentiate shadows from luminal regions as an smoothing with a small scale will increase shadow pixel intensity values when combined with the structures shadows are originated from.

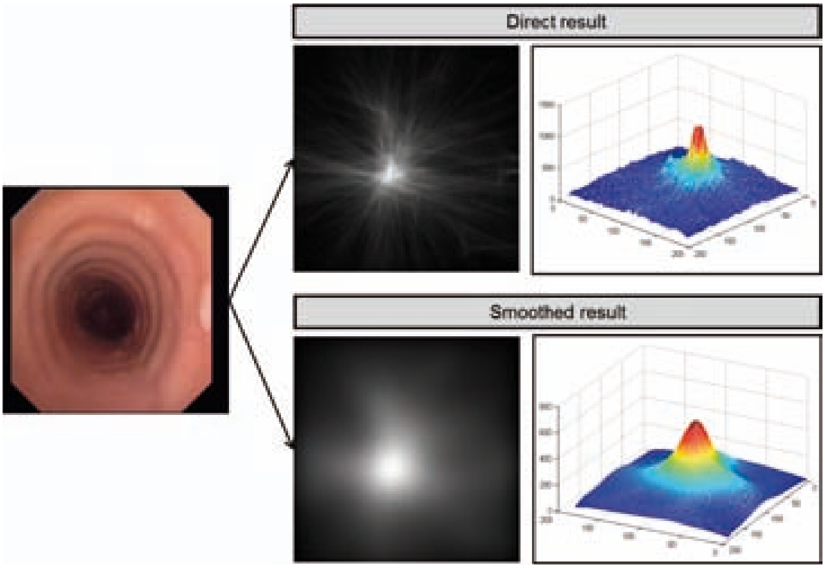


Figure 2.12: Real frame, result of applying DGA and its surface representation (first row). Gaussian smoothing applied to the result of DGA and its surface representation (second row).

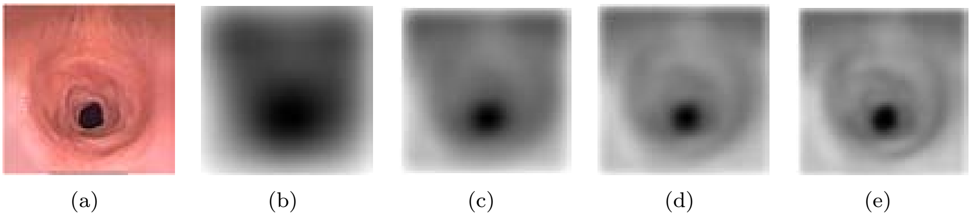


Figure 2.13: Graphical explanation of DRI algorithm: Original bronchoscopy image (a); Smoothed images with σ: 1/8 (b); 1/16 (c); 1/24 (d); 1/32 (e);

Our model aims to be applicable for a wide range of images and lumen pixel characterization is based on both geometric and appearance cues which make it generalizable. Considering our two different descriptors we define that a pixel belonging to the luminal area will have a low DRI value and a high DGA whereas pixels belonging to tracheal rings and folds will have high DRI and DGA value.

Our computational lumen model also acknowledges some of the challenges that videobronchoscopy frames present which can be grouped into three categories: differences in appearance due to different image acquisition methods, different lumen appearances within the same sequence and the possibility of having more than one luminal area -or none- in a given image. Examples of these challenges are shown in Fig.2.14.

2.2.2 Tracheal rings computational model

The physical model of section 2.1 characterizes tracheal rings as a set of sinusoidal signals of increasing radial wave length centred at the lumen. Given the tracheal ring physical model, since valleys present a more stable profile across the distance to the camera (radial direction in Figure 2.9), we will restrict to valley detectors for a tracheal ring detection. Among all

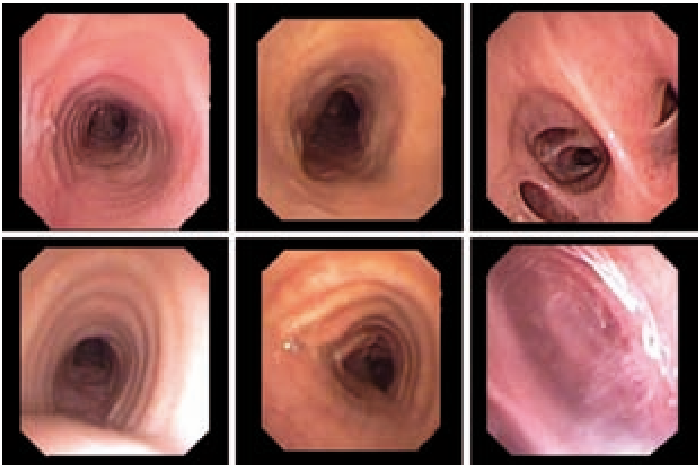


Figure 2.14: Challenges in lumen appearance and geometry. Differences in acquisition conditions (first column), differences in appearance within the same sequence (second column) and differences in number of lumens (third column).

the valley detectors we can classify them into two main groups: geometrical and gaussian detectors. In one hand, geometrical valley detectors are based on the level set curves. As the authors in [33] explain the method is based on finding the minimum value of level curves (valleys). This method is very sensitive with small structures and it does not depend on the orientation. On the other hand, gaussian detectors are based on applying second derivative of an anisotropic gaussian oriented filters into the images. In this case the parameters needed are the number of orientations and range of scales. This method aims to find the best intensity pattern that match with the defined gaussians. These gaussians have a multi scale parameter (σ) between the range defined and also a number of orientations. The orientation is the degree of position of the gaussian. Gaussian valley filter gives the maximum value of matching between all the the scales and orientations. Consequently geometrical detectors introduce more noise and gaussian detectors seems to be more clean and also parameters such as scale and orientation can be controlled which is better for our ring modelling. We have chosen to use Normalized Steerable Gaussian Filters (NSGF) which are implemented as convolution of the image with a second derivative of an anisotropic oriented gaussian kernel [15]. Oriented anisotropic gaussian kernels are given by:

$$G_{\Sigma\ \theta} = G_{(\sigma_x\ \sigma_y)\ \theta} = \frac{1}{(2\xi\ )\sigma_x\sigma_y} e^{-\left(\frac{\tilde{x}^2}{2\sigma_x^2}+\frac{\tilde{y}^2}{2\sigma_y^2}\right)} \tag{2.3}$$

for $(x\ \ y)$ the coordinates given by a rotation of angle θ:

$$\begin{aligned} x &= -x\sin\theta + y\cos\theta \\ y &= x\sin\theta - y\cos\theta \end{aligned} \tag{2.4}$$

and the scale $\sigma_x > \sigma_y$ usually given by $\sigma_x = 4\sigma_y$ [16]. Thus, the anisotropic bank of filters is simply formulated as:

$$G_{\sigma\ \theta} = \frac{1}{(2\xi\ )4\sigma^2} e^{-\left(\frac{\tilde{x}^2}{2(4\sigma)^2}+\frac{\tilde{y}^2}{2\sigma^2}\right)} \tag{2.5}$$

The second partial derivative along the y axis constitutes the principal kernel for computing ridges and valleys:

$$\overset{2}{\tilde{y}}G_{\sigma\ \theta} = (y^2\ \ \sigma^4 - 1\ \ \sigma^4)G_{\sigma\ \theta} \tag{2.6}$$

In order to account for non-uniform illumination artifacts, we use a normalized convolution operator:

$$NSGF_{\sigma\ \theta} := \frac{\overset{2}{\tilde{y}}G_{\sigma\ \theta} * I}{\overset{2}{\tilde{y}}G_{\sigma\ \theta}\quad I} \tag{2.7}$$

for $\cdot$ the L^2 integral norm and $*$ denoting the convolution operator. The response of the operator is calculated as the maximum response for a discrete sampling of the angle and scale domains:

$$NSGF := \max\left(NSGF_{\sigma\ \theta}\right) \tag{2.8}$$

There are two main geometrical aspects of tracheal rings that allow their discrimination from other anatomical structures: an increasing thickness across the radial direction and a concentric disposition around the carina. These two features bound the angular and scale ranges of the NSGF bank. Concentric disposition implies that, for each pixel, the orientation of filters should be perpendicular to radial rays emerging from the carina/luminal area. Meanwhile, increasing radial thickness implies that the scale of the filter achieving the maximum response in Eq. (2.8) increases along each radial ray (see $L1$ profile in figure 2.9). The left scheme in figure 2.15 illustrates the specific filter design modelling the geometric features of tracheal rings in the original image cartesian domain. To better model the

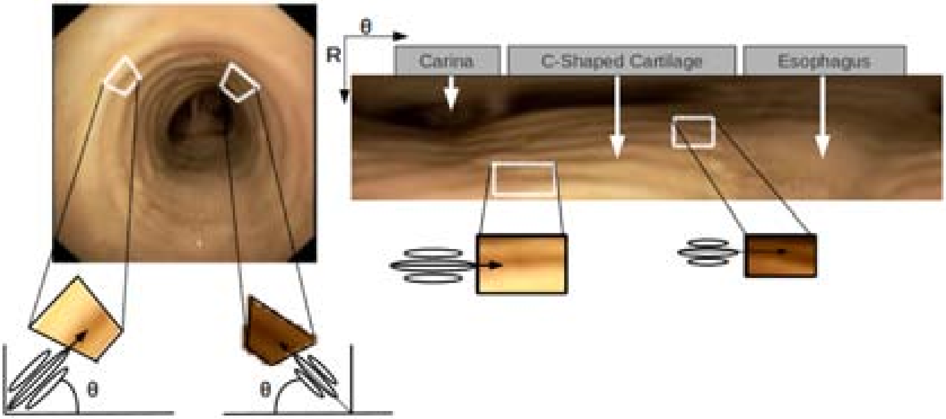


Figure 2.15: Modelling of concentric disposition of tracheal rings in the polar domain: steerable filters in the original cartesian domain (right) and in the polar transform (left) .

geometric features (concentric and radial deceasing scales), images are transformed to polar coordinates. This transformation needs of a definition of a centre of coordinates, which is as identified by the physical model, the centre of the lumen. In the polar domain the ridge-valley profile of tracheal rings follows approximately an horizontal line in which carina becomes to a wide strip at the top of images as shown in the right image of figure 2.15. Rows correspond to the radial coordinate and columns to the angular one.

Due to the perspective projection, the thickness range (scale of the valley) is always within the same range for a given video camera and digital resolution. As scale range depends on image size, a given range is not valid for every bronchoscopy sequence. In this case, differences in resolution between videos can be easily solved by applying a correcting factor depending on the resolution of the training and testing videos.

In order to reduce the impact of alien structrues in tracheal ring characterization we have to consider that the maximum response of Eq. (2.8) includes most tracheal rings but also carina components and some surgical devices (as illustrated in the left images of figure 2.16). Structures not belonging to tracheal rings are removed by forcing rings increasing thickness across the radial direction.

In order to achieve this, we explore the radial profile of the scales achieving the maximum value in (2.8). The bottom left image in figure 2.16 shows the maximal scales at points achieving a local maximum of NSGF. Higher intensity values correspond to larger scales as indicated by the colorbar. The right plot shows the radial profile along the white line labelled $L1$ and shown also in the close-up in the left part of the figure. The cross in the line $L1$ and its radial profile indicates the beginning of the carina.

We observe that, in the absence of non-ring artifacts (L1 profile), such profile keeps decreasing starting from the most external detections as rings are traversed. This is illustrated by the scales of higher radius in the $L1$ profile corresponding to outermost rings in the left image close-up. However, surgical devices and the carina cause the profile of the maximal scales to increase, thus, violating the radial increasing thickness of rings (as illustrated by at the minimal radius in L1 profile). Therefore, structures not belonging to tracheal rings are removed by discarding the first encounter with increasing scale (as indicated by the vertical dotted line in L1 profile). This greedy approach chooses a set of candidate points on tracheal

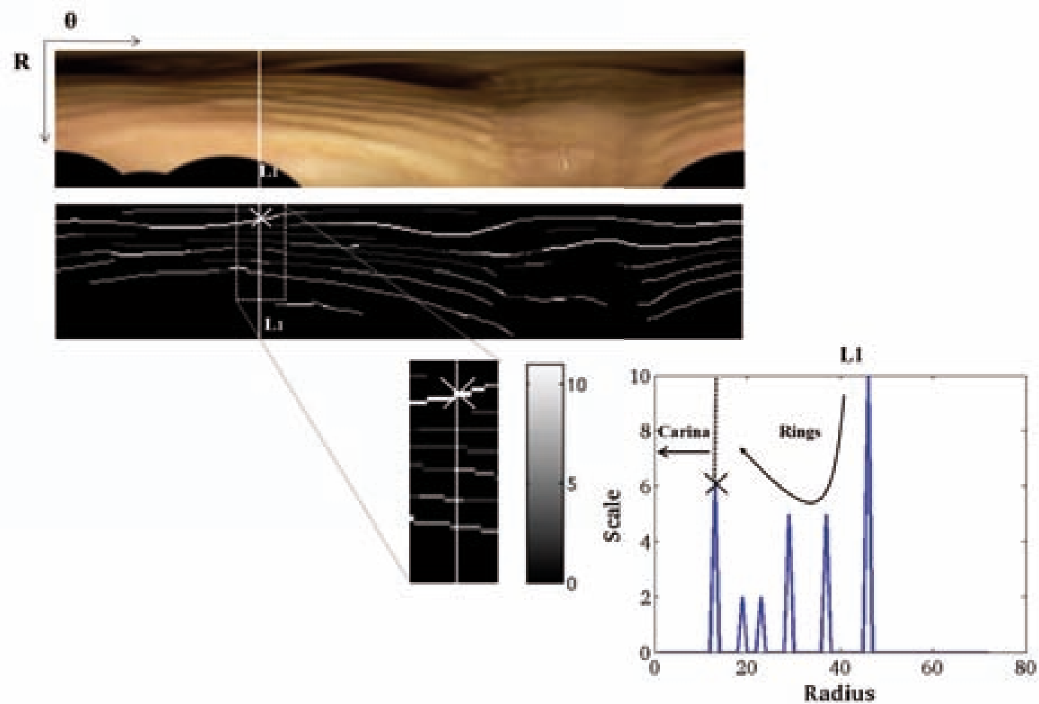


Figure 2.16: Modelling of tracheal rings radial thickness by analyzing the profile of maximal scales.

rings, which are completed by hysteresis on the maximum response (2.8).

Once physical and computational models of tracheal structures have been presented we will proceed in the next chapter on explaining how these structures can be accurately segmented.

Chapter 3

Tracheal structures segmentation for stenosis assessment

We present in this chapter our strategy for efficient and reliable quantification of the degree of stenosis from videobronchosocpic explorations. According to clinic literature [41], such degree can be assessed using the Stenosis Index (SI) metric. Stenosis Index degree is calculated by the physicians from information given by two different frames selected by themselves, one showing a healthy tracheal view whereas the other shows an obstructed tracheal view. The calculation of the final stenosis index is based on the comparison of the area delimited by the first external ring in a healthy frame against the obstructed area delimited by the lumen. The Stenosis Index can be further quantized according to the percentage SI [41] in 3 degrees: Slight ($SI < 50\%$), Moderate ($SI \quad (50\% \; 70\%)$) and Severe ($SI > 70\%$). This partition is currently perfomed by clinicians by means of visual inspection and it is used to decide further treatment actions.

Following ([41,42]) SI can be calculated as:

$$SI = (\frac{AA_{Ref} - AA_{Lumen}}{AA_{Ref}}) \cdot 100 \tag{3.1}$$

for AA_{Ref} the airway area of a reference normal segment and AA_{Lumen}, the airway area of the abnormal one. We observe that being SI an area ratio, it is independent of device and image resolutions. The area AA_{Ref} is computed as the area enclosed by the first complete tracheal ring in a healthy segment, while AA_{Lumen} is given by the luminal area of the obstruction. Therefore, such tracheal structures have to be efficiently detected in bronchoscopy video frames. To achieve a system ready for a deployment in clinical practice, we should develop a strategy able to cope with the main challenges of interventional videos. We propose a multi stage strategy (sketched in Fig.3.1) approaching each of the challenges identified in Chapter 1:

- **STAGE 1: Image Preprocessing**. In order to mitigate some of the problems that we have already pointed out in the previous section, a pre-processing step of the images is needed. This step will clean and prepare the images to be processed in the next stages of our methodology (first step in Fig.3.1).
- **STAGE 2: Structure Characterization using our Physical Model.** In this stage we derive the image processing operators that best match the physical description of the main structures (lumen and tracheal rings) defined in Chapter 2. With this

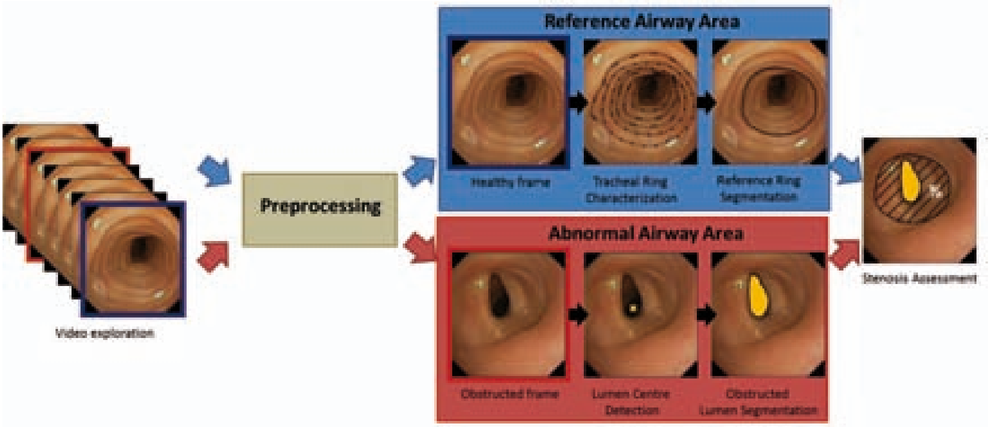


Figure 3.1: Stenosis assessment workflow.

characterization we can reduce the impact of alien structures such as veins or responses at the carina (second step in Fig.3.1).

- **STAGE 3: Structure Segmentation with Optimal Parameter Setting.** Once we have characterized the main structures used for stenosis assessment we segment in this stage the obstructed lumen (pathological frame) for the computation of the abnormal area and the most external and complete ring (healthy frame) for the computation of the normal area. In both cases we use a snake in polar coordinates to fit the shape we are looking for (third step in Fig.3.1). Finally a framework for optimally adjust the parameters for stenosis assessment is used.

Being both structures convex curves centered at the lumen, we will model them in polars as radial functions that depend on the angle $\theta \in [0, 360]$. We will note $R = R(\theta)$, $L = L(\theta)$, the angular functions that segment the reference ring and the lumen in the polar domain. AA_{Ref}, AA_{Lumen} are then given by the integral of each radial function, approximated it by:

$$AA_{Ref} = \sum_{i=1}^{N_\theta} \frac{R(\theta_i)^2}{2} h_\theta \tag{3.2}$$

$$AA_{Lumen} = \sum_{i=1}^{N_\theta} \frac{L(\theta_i)^2}{2} h_\theta \tag{3.3}$$

for $\theta_i = ih_\theta = \frac{i360}{N_\theta}$ a uniform sampling of the interval $[0, 360]$ of size N_θ using as step h_θ.

The remains of this chapter are structured as follows. First we explain which is the preprocess we apply to all images for mitigate the impact of alien structures. Then, we describe the structure characterization of the luminal area which involves the calculation of the the centre and we apply the computation model for tracheal ring characterization which involves valley detector ans some postfiltering. Finally we explain the segmentation of the main structures (obstructed lumen and reference ring) using its characterization.

3.1 Preprocessing

- **Interlacing suppression** Interlacing appears due to the image acquisition process.

In order to increase resolution and reduce acquisition time, video frames are recorded twice at half resolution and then combined to obtain a frame at double resolution. The final frame is obtained by distributing and interlacing each half frame rows into the odd and even rows of the full frame. Camera motion and anatomy dynamics (breathing in our case) introduce small changes between the two half resolutions acquisitions that make odd and even lines not match. So that, we may find aliased contours or edges in images. Although we suspect that for the majority of the frames the difference will be minimal, we have decided to address the interlacing problem in a rather drastic way [10]. In order to prevent the possible effects of time misalignment between odd and even lines, we will only take one of each two lines and we will resize the image in order to maintain the proportions. The process is as follows: Imagine we have an image with size 600x400. If we take only one of each two lines, we obtain two images of size 600x200. We discard one of them and, in order to keep the proportion we have to pass from size 600x200 to 300x200, which can be done bye resizing the image along the horizontal direction by a 0.5 factor.

- **Image black frame inpainting** Camera lens introduces a black frame that might distort the output of the image operators introduced in Chapter 2. In order to minimize the impact of the black frame in image operators without losing anatomic information, we will extend (inpaint) image values to the black mask [6]. Inpainting is performed in a two-stage process. First, image values at the boundary of the mask are extended to the whole mask by using a diffusion process. Second, in a neighbourhood of the black frame boundary extended values are blended with the original values in order to obtain a smooth transitions for valley operators.

In the diffusion stage we diffuse values from the original image into pixels with no value which are under the detection mask M. We track the positions of the pixels under M and, for each of them we perform as follows: we obtain a 3×3 neighborhood around the pixel and change its original value by the mean value of the valid neighbors. Valid neighbours are those pixels which either do not belong to the original M mask or that have already being modified by the diffusion process. This process is repeated until every pixel under M has a new value. Once this happens, we repeat the process until the difference between the new and the previous value of pixels under M is smaller than a threshold values_{th}. We can see that for the calculation of the diffused value of the pixels under M, which are painted in white, we only use information from valid neighbors, painted in orange in the image.

The complete diffusion algorithm is:

Algorithm 1 Inpainting diffusion algorithm

```
Data: Diffusion($I$,$FM$,$MC$)
Arg:($I$: original image, $M$: detection mask, $MC$: minimum change threshold)
    Result: Diffused image($I_d$)
Initialization of valid neighbors mask $VNM = \neg M$
Calculation of diffused values for pixels in $M$
repeat while the image is modified over $s_{th}$
    $stop = true$;
    forall the $x \in I : M(x) == 1$ do
        Definition of a neighborhood around a pixel $Neigh = \{p|p \in Neighborhood(x)\ VNM(p) == 1\}$
        if $\#Neigh > 0$ then
            Calculation of the diffused value $nv = \frac{\sum_{p \in Neigh} I_d(\vec{p})}{\#Neigh}$
            Calculation of the stop flag
            if $VNM(x) == 1$ then
                if $|nv - I_d(x)| > s_{th}$ then $stop = false$ ;
            else
                $stop = false$
            end
            Actualization of the diffused image value $I_d(x) = nv$
        end
    end
until $stop == true$;
```

In order to obtain the final image we take into account that we want to assure that the pixels under M do have their new value on the final image but we also have to consider that if we do a direct substitution there will still remain a clear frontier between pixels inside and outside the final image. In order to solve this we create an extended mask which ponders the way we combine I_o and I_d in the I_{np} image. This M_1 mask is created by dilating the original M mask with a circle structural element and later convolving the result with a gaussian kernel (see Eq. 9).

$$M_1 = M \oplus C_\sigma * G_\sigma \tag{3.4}$$

Once this mask is obtained the final inpainted image I_{np} is calculated in the following way:

$$I_{np} = M_1 \cdot I_o + (1 - M_1) \cdot I_d \tag{3.5}$$

As it can be seen in Figure 3.2, I_o value of pixels under the original M mask are completely replaced by their corresponding values in the I_d. On the other hand, as we depart from the original M mask, the contribution of the original I_o values increases. where $I_o(x\ y)$ and $I_n(x\ y)$ correspond, respectively, to the original image and the image where the values of the pixels belonging to pixels under the original M mask have been changed and α corresponds to the decay factor of the mask. The α factor is used to weight in the final image the contribution of the original version of the image and its smoothed version. By doing this, pixels close to the boundary of the mask will have more content of $I_n(x\ y)$ image and pixels further from the mask will keep their original value.

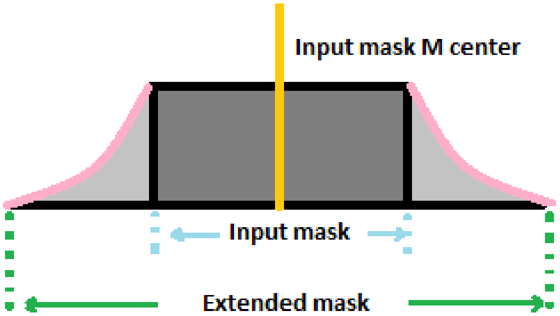


Figure 3.2: Graphical example of the extension of the M mask.

We only show here some graphical examples of the inpainting method applied to the black mask in Fig.3.3.

- **Structure Preserving Diffusion (SPD)**

  Bronchoscopy frames have two main artefacts hindering performance of image processing operators. On one hand, non-uniform illumination may highligh some parts of the trachea while others present more shadows. In the other hand, the impact of other structures not belonging to the main structures (tracheal rings and lumen) should be reduced. In this way, we propose to use structure preserve diffusion to keep valleys from pixels in which the orientation of the gradient is continuous. Pixels with this orientation profile should only correspond to tracheal rings. Consequently filtering will aim to mitigate those short structures with non continuous gradient orientation such as veins. In order to enhance response at anatomical structures, while smoothing texture and noise we will use Structure Preserving Diffusion (SPD) described in [19].

  Diffusion is a mathematical foundation inspired in such a way that physics describe as the propagation of heat on materials. In image processing, diffusion is a technique to reduce image noise by preserving gray-level transitions between adjacent tissues,

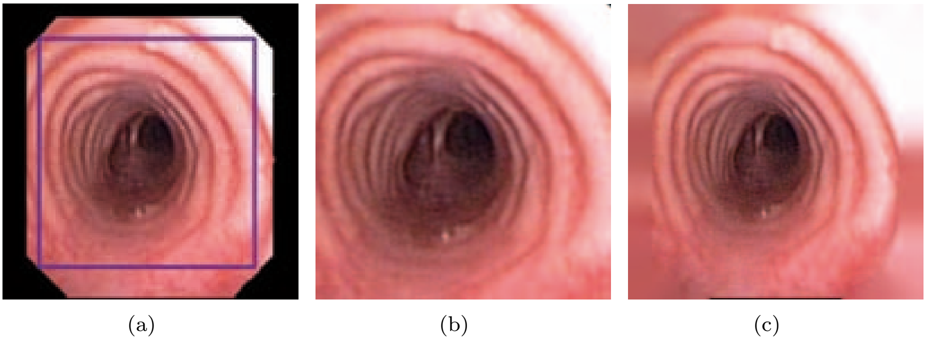

(a) (b) (c)

Figure 3.3: Alternatives to solve the apparition of the black mask: (a) Original image, where the biggest portion of the image without any black mask information is marked by a green square; (b) Result of cropping the image by means of the green square. (c) Inpainting of the area under the black mask.

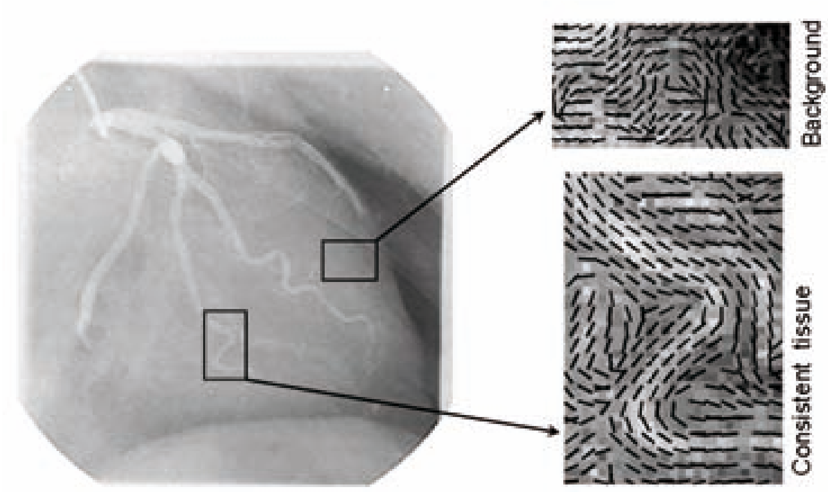


Figure 3.4: Vector field representing level curves of an angiography for a vessel (bottom-right image) and a background structure-less area (upper-right image).

while restoring contours consistent with anatomical structures. The structure preserving diffusion restricts diffusion along the image gradient to preserve anatomical information.

The second moment matrix [38] or Structure Tensor [27] provides a good description of local image structures. The Structure Tensor tensor matrix, $ST_{\rho,\sigma}$ describes the gradient distribution in a local neighborhood of each pixel by averaging the projection matrices onto the image gradient:

$$ST_{\rho,\sigma} = g(\rho) * \left[\begin{pmatrix} I_x(\sigma) \\ I_y(\sigma) \end{pmatrix} (I_x(\sigma), I_y(\sigma)) \right] = \begin{pmatrix} g(\rho) * I_x^2(\sigma) & g(\rho) * I_x(\sigma) I_y(\sigma) \\ g(\rho) * I_x(\sigma) I_y(\sigma) & g(\rho) * I_y^2(\sigma) \end{pmatrix}$$

Image derivatives are computed using gaussian kernels, g_σ, of variance σ (differentiation scale):

$$I_x(\sigma) = g(\sigma)_x * I \text{ and } I_y(\sigma) = g(\sigma)_y * I$$

The projection matrix onto the image gradient, $\nabla I = (I_x(\sigma), I_y(\sigma))$ is averaged using a gaussian of variance ρ (integration scale). Since $ST(\rho, \sigma)$ is the solution to the heat equation with initial condition the projection matrix, its eigenvectors are differentiable (smooth) vector fields that represent image level sets normal (principal eigenvector, ξ) and tangent (secondary eigenvector, $\xi^\perp$) spaces. In the absence of corners (like anatomical contours in bottom right image in fig.3.4), the vector $\xi^\perp$ is oriented along image consistent contours (in the sense of regular differentiable curves [57]). At textured or noisy regions, $\xi^\perp$ is randomly distributed (upper right image in fig.3.4).

The Structure-Preserving Diffusion (SPD) is given by:

$$I_t = (Q \Lambda Q^t \nabla I), \quad I(x, y, 0) = I_0(x, y) \tag{3.6}$$

with:

$$Q = \left(\xi^\perp, \xi \right) \text{ and } \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

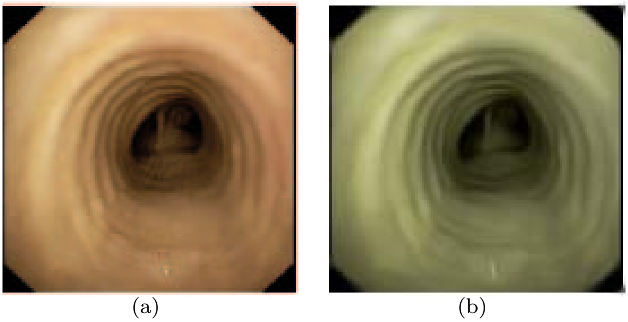

(a) (b)

Figure 3.5: SPD filtering example : Real image (a), Filtered image (b).

for ξ the principal eigenvector of $ST(\rho, \sigma)$. By $\xi^{\perp}$ distribution (fig.3.4), SPD smoothes image grey values along regular structures (bottom right image in Fig.3.4) and performs like a gaussian filter at textured and noisy regions (upper right image in fig.3.4). Its geometric nature makes the restricted diffusion evolution equation converge to a non trivial image that preserves the original image main features as curves of uniform gray level [18]. In this manner, SPD output achieves a uniform response to local image descriptors suitable for a further detection and segmentation of image (anatomical) regions. In our case, we apply SPD to each RGB component to get a color diffused image. Fig. 3.5 shows the benefits of SPD for the extraction of anatomical structures based on ridge operators. Fig. 3.5 (a) shows the original color image and Fig. 3.5 (b) shows its SPD version. The response to NSGF is shown in bottom images. The original SPD defined on greyscale images we apply it to color images by diffusing each channel using its structure tensor. As it can be seen Figure 3.5 it preserves our interested structures while it is smoothing around them.

3.2 Structure Characterization

The computational physical model described in Chapter 2 are used to characterize and discriminate tracheal structures present in the videobronchoscopy scene as follows.

3.2.1 Luminal Region

The operators given in Chapter 2 that describe luminal geometry and appearance in images define a 2-dimensional feature space that characterizes several elements of the endoluminal scene. In particular, pixels belonging to the lumen have a low value of DRI and a high DGA value, specular highlights and other bright protruding surfaces have high DGA and DRI and structures like folds and rings (which generate shadows) have low DGA and DRI values. Therefore, a classification of the (DRI,DGA) feature space into this three classes should characterize the luminal region. In order to account for possible non-gaussianity and skip tedious manual labelling in a supervised scheme, the partition of the feature space into this three classes is obtained by unsupervised 3-means clustering over a training set. Such training set consist of a set of 30 images randomly sampled from 20 representative videos of our dataset. These images were uniformly sampled to obtain a set of pixels that cover the 2D feature space and are the input for the k-means classification. The cluster with the

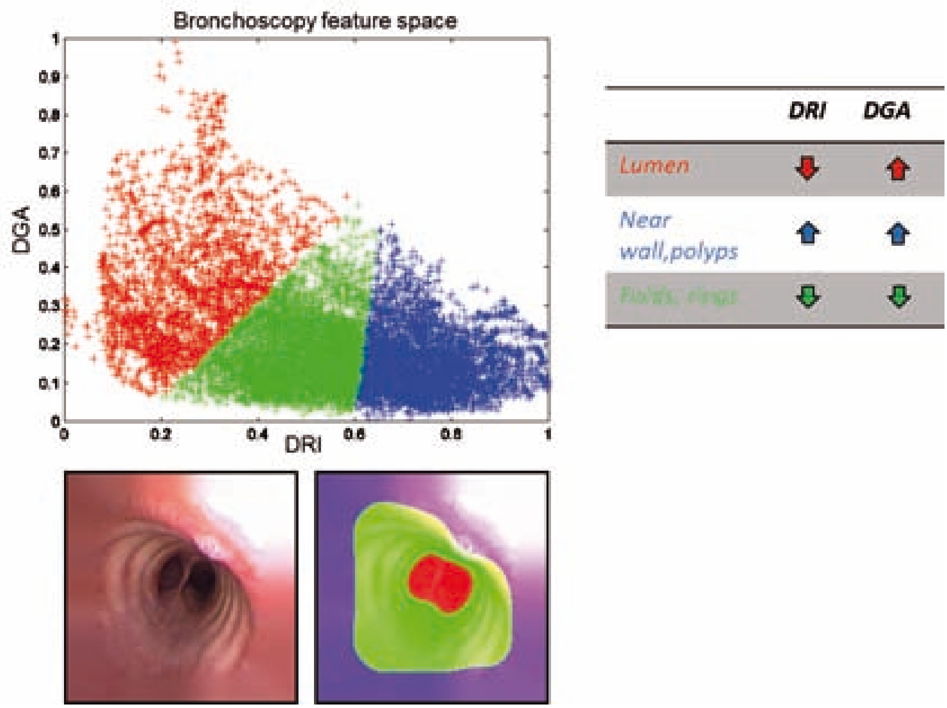


Figure 3.6: Adequation of our feature space to videobronchoscopy.

lowest DRI and highest DGA values is selected as being the luminal one.

In order to have comparable values, the output of both descriptors has been normalized in $[0, 1]$ range. This normalization has been obtained by means of the maximum and minimum values achieved for the training set. Figure 3.6 shows the (DRI, DGA) normalized feature space for the training set classified in the three clusters. Red pixels correspond to the lumen, green pixels to near walls and instruments, and blue pixels belongs to the folds and tracheal rings. Bottom images in Fig.3.6 show one of the training frames and the classification using k-means. Given the visual quality of the luminal cluster (quantitatively assessed in Experiment 5.1.1 before proceeding to further steps), we did not consider necessary to use advanced supervised classifiers (such as Support Vector Machine).

The distance of a pixel to the borders given by the clustering defines a likelihood map of its belonging to the each of the classes. In our application this border has been approximated by a linear plane of origin (DRI_0, DGA_0) and normal direction (V_{DRI}, V_{DGA}), so, for each feature point (DRI, DGA) its likelihood map LK is defined by:

$$LK(DRI, DGA) = (DRI - DRI_0)V_{DRI} + (DGA - DGA_0)V_{DGA} \tag{3.7}$$

The likelihood map LK can be used to detect the lumen center and also to define a external energy guiding the final snake that segments the lumen (as described in Section 3.3.1).

The center of the lumen can be calculated as the local maxima of the LK map. In this case we use the local maxima and not the global maximum as we aim our model to be general so it must contemplate that for some bronchoscopy frames -representing bifurcations in the

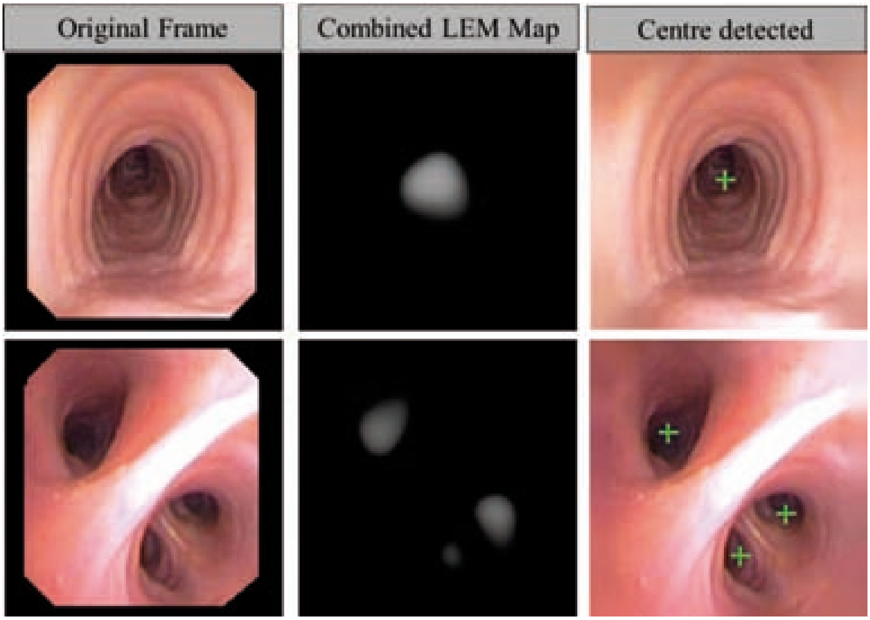


Figure 3.7: Workflow for the lumen centre processing.

bronchial tree- there is more than one luminal area. We have to note that as we are training our classifier our model is able to cope with lumen absence: in the case of no luminal area there will be no pixel which falls in the red region of the feature space. Figure 3.7 shows different examples of how the centre is obtained. We can see that our methodology allows us to cover all possible lumen presence cases: our method is able to detect one lumen (first and third row), multilumen (second row) and gives no output where there is no lumen in the image (last row).

3.2.2 Tracheal ring

Following the physical model described in Chapter 2, tracheal rings are characterized in terms of image valleys computed using our NSGF operator given by convolution to a bank of second derivatives of Gaussian kernels $\partial_y^2 G_{\sigma,\theta} = (\tilde{y}^2/\sigma^4 - 1/\sigma^4)G_{\sigma,\theta}$. The maximum response for a discrete sampling of the angle and scale domains defines the valley feature:

$$NSGF := \max_{i,j} \left(NSGF_{\sigma^i,\theta^j} \right) \tag{3.8}$$

for $\|\cdot\|$ the L^2 integral norm and σ^i,θ^j discrete scale and orientation sampling. To account for the global concentric-profile, NSGF is computed in the image polar domain. The transformation to polars is computed using the lumen centre (extracted following Section3.2.1) as origin of coordinates.

Figure 3.8 shows a frame in cartesian domain and its polar transformation. In the polar image, rows correspond to the radial coordinate (with origin at the top row) and columns correspond top the angular coordinate in the range $[0, 360]$. We show in the polar image the axis indicating the angular and radial coordinates. We observe that in the polar domain, rings are the only horizontal structures covering the full angular range. Therefore, in such

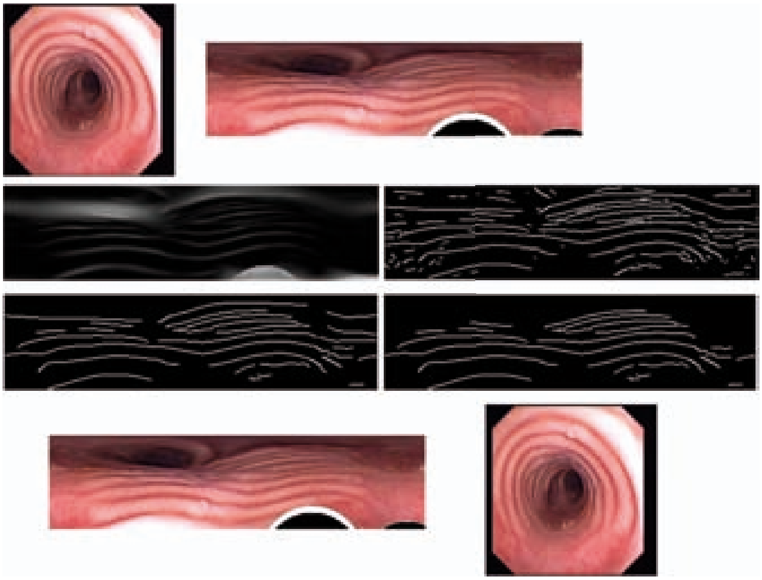


Figure 3.8: Workflow for tracheal ring characterization.

polar domain, the orientation and scales used to compute (3.2.2) are set to cover the scale ranges of ring thickness and restricted to horizontal structures. It follows that, in the polar domain, the pixel-wise radial orientation of the NSGF filters is reduced to an horizontal orientation given by $\theta = 0$. In order to account for any deviation in the circular profile of rings (see images shown in figure 4.4 for an example) the sampling of the angle defining $NSGF(\sigma^i, \theta^j)$ considers a small range around 0 given by $[-\epsilon, \epsilon]$:

$$\theta^j = \{\frac{2\epsilon j}{N_\theta} - \epsilon, \forall j = 0 \dots N_\theta\} \tag{3.9}$$

for N_θ the number of sampled angles.

Concerning scale sampling, we have to consider that σ is related to the thickness of the valley. For a given sequence, let the range be defined as $[\sigma^m, \sigma^M]$ and the scale sampling will be given by a uniform sampling of the former interval:

$$\sigma^i = \{\sigma^m + \frac{i(\sigma^M - \sigma^m)}{N_\sigma}, \forall i = 1 \dots N_\sigma\} \tag{3.10}$$

for N_σ the number of sampled scales.

The response to NSGF using the restricted parameter ranges for Gaussian filter orientation and scales is shown in the second polar image of fig. 3.8. We observe that there is almost no response at other valley-like structures present in the image. Such suppression of artifacts is better appreciated in the close-ups showing an area inside the carina and an region containing a vessel.

Tracheal rings correspond to the local maxima of the response to NSGF. Local maxima of NSGF response are computed using Non-Maxima Supression [44], NMS. Non-maxima are suppressed by keeping only those pixels such that NSGF value is maximum along the orientation, θ_{NSGF}, of the filter that achieved the maximum response in (3.2.2). In order to keep valley geometric information, NMS values are given by the scales, σ_{NSGF}, of the filter that achieved NSGF maximum value, so that we define:

$$NMS_{NSGF}(\theta\ \ R) = \tag{3.11}$$

$$= \begin{array}{ll} \sigma_{NSGF} & \text{if } NSGF > \max(NSGF_{V+}\ \ NSGF_{V-}) \\ 0 & \text{otherwise} \end{array}$$

for $V = (\cos(\theta_{NSGF})\ \sin(\theta_{NSGF}))$, $NSGF_{V+} = NSGF(x + V_x\ \ y + V_y)$ and $NSGF_{V-} = NSGF(x - V_x\ \ y - V_y)$. $NSGF_{V+}$ and $NSGF_{V-}$ images are computed using linear interpolation.

As a result, NMS should suppress all the pixels that are not candidates to be part of the skeleton, therefore, all the pixels that are not a local maxima. The result is a thinned one-pixel mask of the dominant valleys, as shown in Figure 3.8.

In order to suppress invalid responses due to anatomical elements such as veins and carina, we apply a length and scale filters. Our length filter assess NMS_{NSGF} connected components of length below a given threshold, LCC. NMS_{NSGF} segments length is computed as:

$$CC^i = \quad NMS_{NSGF}(x\ \ y)^i \quad _{nc} \tag{3.12}$$

$$AreaCC^i = \qquad R^i_j(\theta^i_j - \theta^i_{j+1}) \tag{3.13}$$

for $R_j\ \ \theta_j$ the radial and angular coordinates of each NMS_{NSGF} component in the polar domain. Therefore, the length filtered NMS_{NSGF} is defined as:

$$NMS_{AF}(x\ \ y) = \tag{3.14}$$

$$= \begin{array}{ll} CC^i & \text{if } AreaCC^i > \overline{AreaCC} \\ 0 & \text{otherwise} \end{array}$$

Where CC^i are all the nc connected components, $AreaCC^i$ is the polar area (with radius R and angle θ) of each of the CC and NMS_{AF} is the final mask removing those components that its area is higher than the mean area.

Although the length filter already removes some response at the carina, the length of some carina components can be close to 2ξ in polar domain and, thus, LCC in cartesians. To remove such carina responses we use rings physical property that sets that they should follow a decreasing thickness pattern in the polar domain that is if a external ring is detected in scale σ_1, scales of inner rings should be at lower scale values. The scale filter is computed by analysing the radial profile of $NMS_{NSGF(\theta\ R)}$ for each angle. Let $NMS(\theta_o\ \ R)$ the plot of scales for a given angle θ_o and $R_SF = \min_R NMS(\theta_o\ \ R)$ the radius achieving the minimum scale. We define our scale filter, NMS_{SF} for each angle θ_o as:

$$NMS_{SF}(\theta_o\ \ R) = \tag{3.15}$$

$$= \begin{array}{ll} NMS(\theta_o\ \ R) & \text{if } R < R_{SF} \\ 0 & \text{otherwise} \end{array}$$

3.3 Structure Segmentation

In order to obtain the curves that enclose the obstructed lumen and the reference ring areas we will use active contour or snakes. A snake schema is defined as far as the external energy and the initial curve are defined and the internal energy parameters are defined. In order to ensure convergence to the target contour, the external energy should be convex or the initial snake close enough to the target. Also, internal parameters might influence snake convergence and often require fine tuning. Image noise and non-convex profiles are visual sources of snakes premature convergence that force an accurate initial curve. In this section we will describe two strategies for ensuring convergence to obstructed lumen and reference ring, independent of internal energy settings. Finally we present a novel framework based on the use of Anovas to obtain the optimal parameter values that allows our methods for stenosis assessment to reach inter-observer variability performance.

3.3.1 Obstructed Lumen

In this case the formulation of the snake $L = L(\theta)$, that segments the lumen region is:

$$L_t = \ \ _r(L_{Ext}) + \alpha \ \ _\theta(L_{t-1}) + \beta \ \ _{\theta\theta}(L_{t-1}) \tag{3.16}$$

for α, β corresponding, respectively, to the weights for the stiffness and elasticity snake terms. In this case, the external energy L_{Ext} is given in terms of the linear classifier as follows:

$$\begin{aligned} LK_l \quad & := \\ & DRI \cdot V_{DRI} + DGA \cdot V_{DGA} - \\ & -(DRI_0 \cdot V_{DRI} + DGA_0 \cdot V_{DGA}) \ = \\ & = \ DRI \cdot V_{DRI} + DGA \cdot V_{DGA} - l \end{aligned} \tag{3.17}$$

for $\ \cdot \ $ the absolute value. The linear classifier defines a 1-parametric family of likelihood maps depending on the intercept, l. The top plot in Fig. 3.9 shows the $(DRI \ \ DGA)$ feature space with two different classifiers at two different intercept values, l_k and l_{k+n} (dashed lines).

The values LK_l can be interpreted as the distance (in the feature space) to the set of pixels that define the border (given by $LK_l = 0$) between non-lumen and lumen points and, thus, correspond to a local minima of LK_l. We note that values within $l \quad [min(V_{DRI} \ \ V_{DGA}) \ \ max(V_{DRI} \ \ V_{DGA})] = [l_m \ \ l_M]$ produce likelihood maps that have a well defined curve of minimal points. Outside this interval, LK_i does not have any local minima because for such values all image pixels would be classified as either lumen or non-lumen. The set of LK_l mimimal points progressively approximate the lumen border as the interval $[l_m \ \ l_M]$ is swept. The optimal intercept value, l_j that yields the best approximation will be computed using the ANOVA design described in Section 3.3.3.

The solution to (3.16) using $L_{Ext} = LK_l$ would approximate the lumen border provided that L_0 is closed enough to the target curve or LK_l is convex. Otherwise, the solution could get trapped in a local minima which does not correspond to the lumen border. In order to ensure convergence to the lumen border and also speed-up the process, we will consider a family of snakes, $(L^j)_{j=1}^{j=N_l}$, that solve (3.16) for a sampling of size N_l of the interval $[l_m \ \ l]$ given by $l^j = l_m + j \ \ N_l(l - l_m)$. Each snake L^j is formulated as the steady state of:

$$L_t^j = \ \ _r(LK_{l^j}) + \alpha \ \ _\theta L_{t-1}^j + \beta \ \ _{\theta\theta} L_{t-1}^j \text{ with } L_0^j = L^{j-1} \tag{3.18}$$

being L^{j-1} the solution for $L_{Ext} = LK_{l^{j-1}}$ and $L^0 \equiv 1$. This progressive scheme overcomes premature convergence of the snake and also minimizes the impact of its initialization and internal energy parameters α, β. The snake for $j = N_l$ is our final lumen segmentation in the polar domain. The optimal intercept, l, defines the snake L^l to be used to segment the lumen. l will be set using the ANOVA design described in Section 3.3.3.

The middle images in Fig. 3.9 show two different LK_{l^j} energies in polar coordinates for consecutive values l^{k_1} and l^{k_2}. The snake solving (3.18) for l^j and l^{j+n} is plotted on the polar representation of the original image shown under each LK_l map and on the cartesian bottom image. The curve corresponding to $l = l^{k_1}$ is plotted in blue while the curve for $l = l^{k_2}$ is plotted in white.

3.3.2 Reference Ring Segmentation

The output of tracheal ring characterization offers a set of lines -in the polar domain- which are candidates to be the reference ring. The reference ring corresponds to the most external complete ring and we segment it in a three step strategy sketched in Fig.3.10. First, a set of candidate points (yellow lines in the first central image in Fig.3.10) are selected from the response to the operator $NMS_{SF}(\theta \ R)$ filtered at suitable scales. Second, these unconnected segments are completed using the image geometric information. The result of this step is shown as a grey line in the second central image in Fig. 3.10. Finally, a snake is used to model a closed curve (blue line in both the second central and the right image of Fig. 3.10) for AA_{Ref} computation.

The candidate rings to be the reference one are selected by filtering NMS_{SF} in a sub-range of scales from the ones used in the original NSGF calculation. This sub-range of scales can be set to detect rings at a desired scale range but in our case we would look for the limits of the optimal sub-range of scales, $[\sigma^j \ \sigma^{j+k}]$ that best separate reference ring from the others. The obtention of this optimal sub-range will be set using the ANOVA design described in Section 3.3.3.

The mask of candidate rings within the sub-range will be noted by θ_{SF}:

$$\theta_{SF}(x \ y) = \tag{3.19}$$

$$= \begin{matrix} 1 & \text{if } NMS_{SF}(x \ y) \quad [\sigma^j \ \sigma^{j+k}] \\ 0 & \text{otherwise} \end{matrix}$$

The candidate rings (as illustrated in the central top image of Fig.3.10) tend to appear as a fragmented set that should be completed in order to compute AA_{Ref}. Completion is achieved in a two-stage process that takes into account as much anatomical information as possible.

First, we apply the Anisotropic Contour Closing, ACC, described in [20], to complete each candidate ring connected component using the gradient information of the original image in polar coordinates. The ACC operator is defined as the following restricted heat operator:

$$u_t = (Q\Lambda Q^t \quad u) \qquad u_0 = \theta_{SF} \tag{3.20}$$

with Dirichlet conditions on $\theta_{NSGF}(x \ y)$ equal to 1:

$$u \quad_{\theta_{SF}(x \ y)>0} = 1 \tag{3.21}$$

The symbol in (3.20) stands for the divergence operator and the diffusion tensor $Q\Lambda Q^t$ is defined as:

$$Q := \begin{pmatrix} \partial & \partial \end{pmatrix} \text{ and } \Lambda = \begin{matrix} 1 & 0 \\ 0 & 0 \end{matrix} \tag{3.22}$$

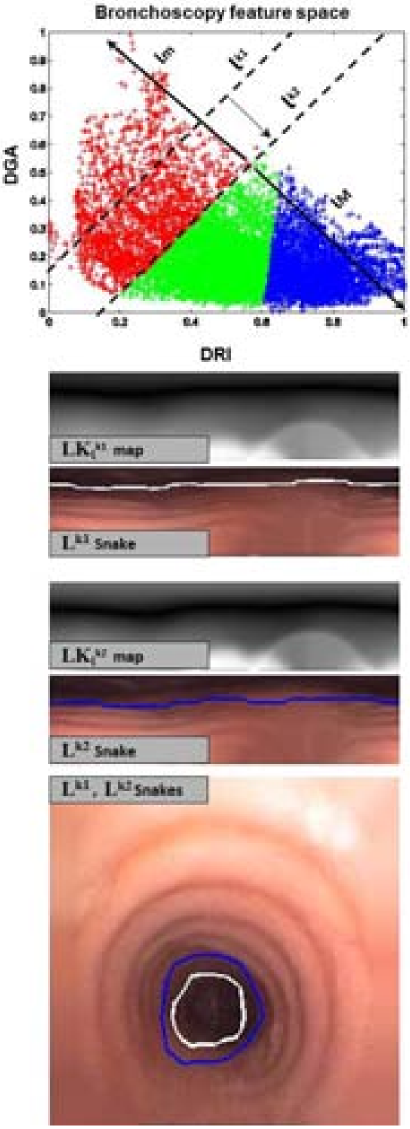


Figure 3.9: Consecutive lumen segmentations (l^{k1}, l^{k2}) of the feature space, top of the image; likelihood maps (LK_l^{k1}, LK_l^{k2}) and snake computation (L^{k1}, L^{k2}) in polar coordinates in the middle images and snake computation in cartesian domain in the bottom image.

where ∂, ∂ are the eigenvectors (ordered in decreasing eigenvalues) of the structure tensor, $ST(\rho_1 \ \rho_2)$, computed over the original image in polar coordinates:

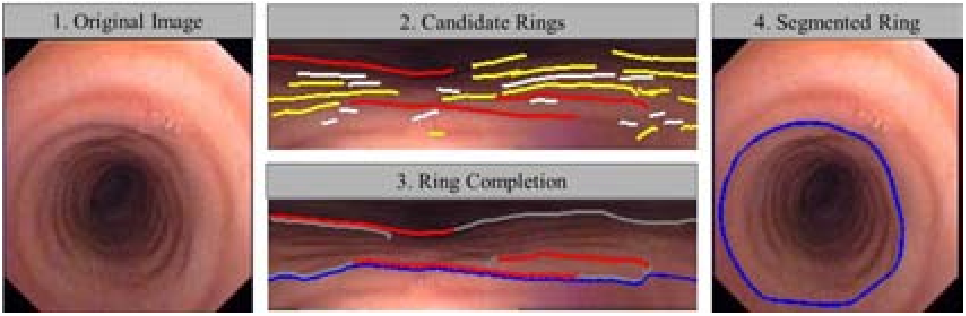


Figure 3.10: Workflow for the tracheal ring segmentation. 1. original image, 2. detected rings in white, candidate ones in yellow and the reference ring in red over the image in polar coordinates, 3. shows the result of ACC closing (grey) and snake completion (blue) and 4. represents the final reconstructed ring in the cartesian domain.

$$ST(\rho_1\ \ \rho_2) := g_{\rho_1} * \begin{matrix} {}_xI_{\rho_2} \\ {}_yI_{\rho_2} \end{matrix} \quad {}_xI_{\rho_2} \quad {}_yI_{\rho_2} \quad =$$

$$= \begin{matrix} g_{\rho_1} * \ {}_{xx}I_{\rho_2} & g_{\rho_1} * \ {}_xI_{\rho_2}\ {}_yI_{\rho_2} \\ g_{\rho_1} * \ {}_xI_{\rho_2}\ {}_yI_{\rho_2} & g_{\rho_1} * \ {}_{yy}I_{\rho_2} \end{matrix}$$

for g_{ρ_1} a gaussian kernel of variance ρ_1 and image derivatives being computed by convolution with the derivatives of another gaussian kernel, g_{ρ_2}, as:

$${}_xI_{\rho_2} = \ {}_xg_{\rho_2} * I \text{ and } \ {}_yI_{\rho_2} = \ {}_yg_{\rho_2} * I \tag{3.23}$$

Since $ST(\rho\ \sigma)$ is the solution to the heat equation with initial condition the projection matrix, its eigenvectors are differentiable (smooth) vector fields that represent image level sets normal (principal eigenvector, ∂) and tangent (secondary eigenvector, ∂) spaces. Therefore, the solution to (3.20) is a binary map of a closed model of the uncompleted initial contour. Intuitively, the restricted operator serves to integrate the vector ∂ , so that it interpolates the unconnected curve segments along it. Based on this interpretation, we can derive a numeric implementation [26] that is computationally efficient and, thus, could be used in operation room time. Figure 3.10 central bottom image shows the ACC completion (gray lines) of the fragmented curves shown on the top image. We observe that ACC is able to restore the original ring profile, while stopping at axes where no geometric information is available.

The ACC completion step gives connected contours that match the C-shape geometry of tracheal rings and, thus, are open curves that end at the trachea lower part limiting with the esophagus. In order to obtain as a result a closed curve, we use again a deformable model [29] guided by the distance map to the reference ring. The reference ring is the most complete ring covering approximately the 180 degrees of the C-shape arc. In case of having more than one complete ring, the most external one having the maximum radial coordinate is chosen. If we note by R_{Ext} the distance map that acts as the snake external energy, and by $\ {}_\theta R$, $\ {}_{\theta\theta}R$ the first and second derivatives with respect to the angle θ, then the curve that segments the ring is the steady state of:

$$R_t = \ {}_r(R_{Ext}) + \alpha\ \ {}_\theta(R_{t-1}) + \beta\ \ {}_{\theta\theta}(R_{t-1}) \text{ with } R_0 \equiv 1 \tag{3.24}$$

for α, β corresponding, respectively, to the weights for the stiffness and elasticity snake terms. The blue line in the central bottom image of Fig. 3.10 shows the final snake that closes the selected ring completed with ACC (gray line). The ring in the Cartesian domain is shown in the right most image of the same Fig. 3.10.

3.3.3 ANOVA design for optimal parameter setting

In medical diagnosis a new method is accepted as a diagnosis tool provided that it can substitute the gold standard [64]. The usual criterion to assess such interchangeability is to compare the output of the new tool to the one provided by the gold standard. In the context of image processing for the extraction of anatomical structures, the gold standard (or ground truth, GT) is provided by experts manual annotations. Meanwhile, the comparison is given by one metric quantifying differences between manual and automatic regions or curves [24].

In case of having a unique GT, it is common to choose the method or parameter configuration achieving the least average error (metric) for a training set. However, in order that this conclusion can be generalized to the whole data set with a given confidence, it is mandatory to use some kind of statistical test for detecting multiple differences across the tested methods [13].

We propose to detect significant differences across parameter configurations using Analysis of variance (ANOVA). ANOVA s [8] are powerful statistical tools for detecting differences in performance across methodologies, as well as the impact of different factors or assumptions. We can apply ANOVA in case our data consists of one or several categorical explanatory variables (called factors) and a quantitative response of the variable. The variability analysis is defined as soon as the ANOVA quantitative score and the different factors and methods are determined. Training data (individuals) is grouped according to such factors and differences among quantitative response group mean are computed. ANOVA provides a statistical way to decide whether such differences are significant enough with a given confidence level α. Given that different observers produce different GTs, a method is optimal if its metric ranges are comparable to such inter-observer variability.

We will use ANOVA to compare multiple methods to inter-observer variability and explore sources of methods errors as follows. First we perform a 2-way ANOVA with factors given by the several methods and the observers to check if methods rankings varies across experts. The individuals and ANOVA variable are taken as before. The desired result of this test would be a significance in the method s factor, possibly a significance across observers and, most important, no significant interaction. A non-significant difference across observers would indicate that there is no evidence of an inter-observer variability for the chosen set of experts and, thus, that any of them could be used as GT. In case of significant interaction, a 1-way ANOVA with the combined method-observer factor should be used to detect the sources of bias. Otherwise, we can compare the methods output to observer ranges by using the following 1-way ANOVA.

The final comparison to inter-observer ranges in case of no interaction will be given by a 1-way ANOVA with factor groups given by the different methods and the observers evaluated against each other in a single group. For the methods, the ANOVA variable will be the metric averaged over all observers, while for the observers control group we will average the metrics obtained by evaluating each observer against the remaining ones. Given that in this case, the observer group acts like a control group we want to compare methods to, the correction used for the multicomparison test should be Dunnet [12]. The optimal methods configuration will be the ones such that Dunnet test does not find significant differences.

In the context of parameter tuning, ANOVA groups would be defined by the different parameter settings and the ANOVA variable would be given by the region or curve comparison metric for a random sampling of the annotated data set. In case of significant differences across ANOVA groups, a multiple comparison test with Tukey correction [61] would detect those configurations that are significantly worst. The remaining methods are good candidates to substitute the observer that performed the manual annotations. However, it is a fact that there are significant differences across experts annotations and, thus, the variability across them should be taken into account in the validation process [31].

Multi-comparison to inter-observer variability is not an easy task [30]. A usual strategy consists in averaging the quality metrics obtained using each observer as GT and comparing this score to the average metrics obtained by evaluating the observers against each other. The method is considered to perform within inter-observer variability if the average ranges are comparable. We note that this is a fair comparison provided that the methods ranking is the same across observers and that agreement occurred by chance is corrected [22]. Ignoring the impact of variability across observers in methods rankings could hide sources of variability and error bias in either methods performance or experts annotations that should be further investigated. This framework will be applied for the different structure segmentation experiments that will be presented in the next chapter.

Chapter 4

Experimental Settings and Databases

We present in this chapter the different experimental setups that we have created to assess the performance of the different methods developed within this thesis. In the one hand, we have tested the accuracy of our methodology for the characterization of tracheal structures -tracheal rings and lumen centre-. On the other hand, we propose different experiments for assessing main structures segmentation -reference ring and obstructed lumen- which are key for the validation of automatic stenosis assessment. For each exeriment we report its goal, databases used and validation protocol and metrics.

One main contribution of this thesis has been the introduction of the only available fully annotated databases, which are also presented for each experiment. All the frames belonging to these databases have been provided by the Hospital de Bellvitge. Frames have been extracted from several sequences following the specific requirements of each experiment.

4.1 Structure characterization

4.1.1 Lumen centre detection

The aim of this experiment is to validate our lumen centre characterization. Given an input frame we apply our method for lumen centre detection and we check if this centre is inside of the lumen region annotated by the expert (groundTruth). In order to be useful for validating a wide range of algorithms, an annotated database should fulfil the following requirements:

1. The selected frames should be different enough in order to have the maximum variability available of lumen appearance.
2. The database should also contain examples of frames both with multiple lumen (more than one lumen centre) and without lumen (absence of lumen centre).

Taking these constraints into account, we have built up a database of 125 images[1] extracted from 20 different videobronchoscopy sequences. Table 4.1 gives a description of the different groups, Fig. 4.1 shows an example with its segmentation and Fig. 4.2 shows a mosaic with some examples of the database.

---

[1] http://iam.cvc.uab.es/downloads/

Index	Type	Resolution
1 − 100	Bronchoscopy [2]	[144 × 144 288 × 288 186 × 186]
201 − 225	Bronchoscopy (15 multi-lumen and 10 no lumen) [2]	[144 × 144 288 × 288 186 × 186]

Table 4.1: Description of the groundtruth of the lumen database.

The lumen centre detection has been validated in terms of true localizations (TL), false localizations (FL) and no localizations (NL). We have used Precision and Recall scores to summarize the performance:

$$Prec = \#TL/(\#TL + \#FL) \tag{4.1}$$

$$Rec = \#TL/(\#TL + \#NL) \tag{4.2}$$

Precision results will measure how good are our centres and Recall will give a sense of the lumens we are missing.

4.1.2 Tracheal ring

The objective of this experiment is to see how good we are detecting just tracheal ring structures. The importance of this stage relies on finding whether the method is able to detect only either full rings or part of them for a further use in recovering the reference one. This experiment has two different parts: the first one explores the relationship between the scale parameter of our ring detector and the resolution of the input image; once this scale parameter is set we aim to assess the performance of our tracheal ring segmentation method. Considering this, given an input frame we detect the rings and we compare our results with the two manual segmentation made by two experts.

Our data set consists of four sequences of healthy and pathological trachea cases provided by the Bellvitge hospital. Sequences have been obtained with either rigid and flexible bronchoscopes at different resolutions. Scale parameters have been tuned empirically for the sequence of lowest resolution and estimated for the remaining sequences by using the scaling ratio between both resolutions. Scale ranges and video main features for all videos are given in Table 4.2.

---

[2] Bellvitge Hospital Barcelona

[3] Beaumont and St. Vincents Hospistal Dublin

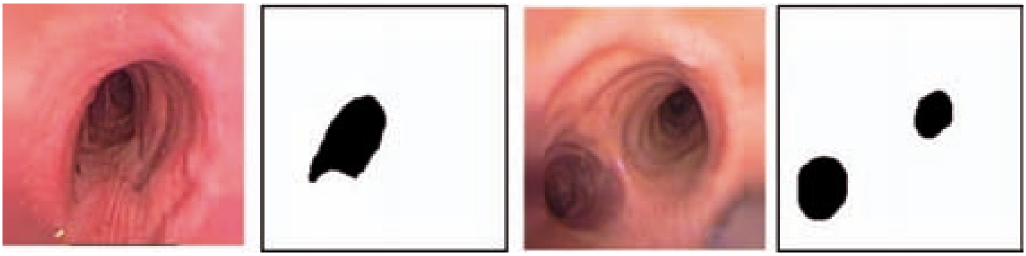


Figure 4.1: Examples of images from our database: real images (first and third) and its respective ground truth (second and four).

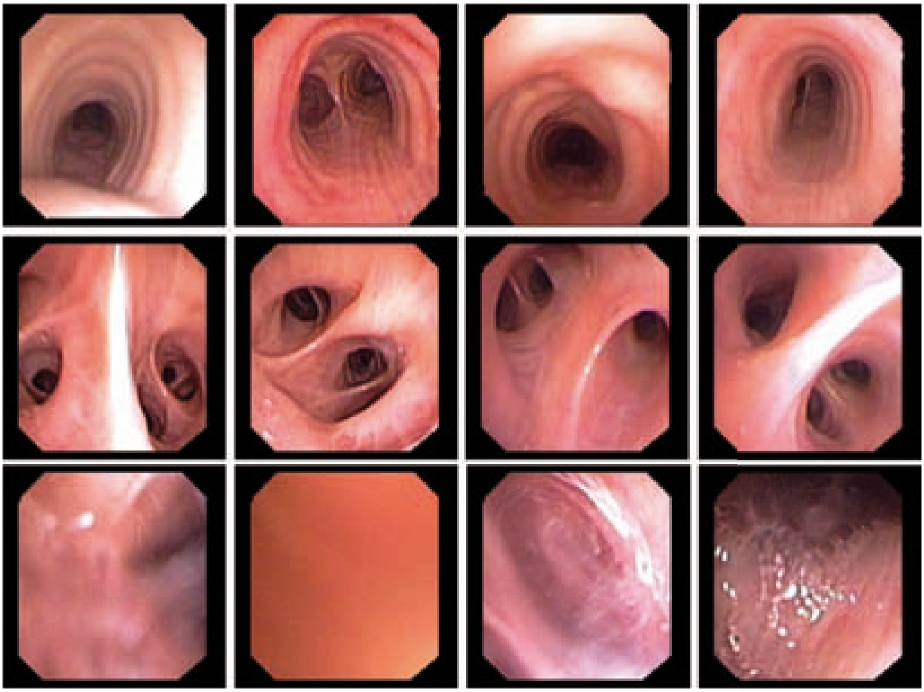


Figure 4.2: Lumen dataset. Examples of bronchoscopy (three first rows) with one lumen (first row), multilumen or bronchial lumens (second row) and no lumen presence (third row).

In order to assess our geometric appearance model, we have compared automatically detected rings with manual segmentations. Comparison has been made over a set of 60 representative frames uniformly sampled from each video. Figure 4.3 shows an example of segmentation of two frame by two different experts and Fig. 4.4 shows some of the frames chosen for each of the sequences.

Difference between manual and automatically detected rings has been quantified in terms of true positives, TP, and false positives, FP. We have chosen this goodness measures instead of standard distances to curves because they are more sensitive and discriminative

Table 4.2: Main features of our dataset: Scale ranges and bronchoscope type, pathologies and resolutions.

	Seq_1	Seq_2	Seq_3	Seq_4
Bronchoscope type	Flexible	Rigid	Rigid	Flexible
Pathology	No	Yes	Yes	No
Resolution	192x144	512x288	360x288	512x288
σ	[0.9,2.9]	[2.2,6.8]	[1.8,5.6]	[2.2,6.8]

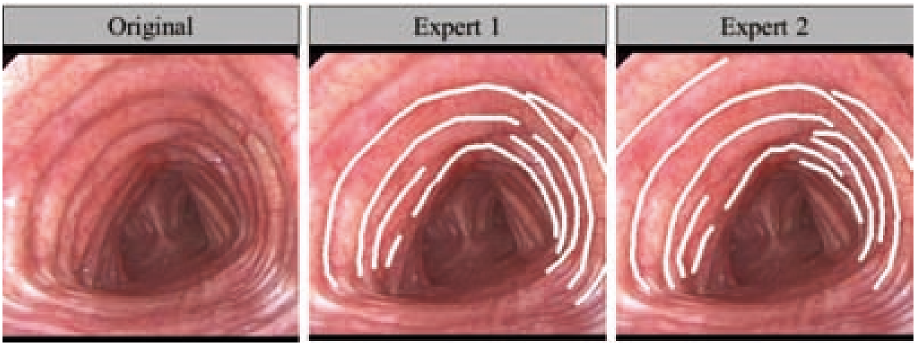


Figure 4.3: Example of our database: bronchoscopy image (first), expert 1 groundtruth (second) and expert 2 groundtruth (third).

for the amount of lost rings (given by TP) and alien structures included in detections (given by FP). In order to account for accuracy in ring location, TP are defined as those points 1 pixel distance away from manual curves. The percentage of TP and FP are reported in terms of sensitivity and precision scores:

$$Sens = 100\frac{TP}{TP+FN}, \; Prec = 100\frac{TP}{TP+FP} \tag{4.3}$$

$$Prec = 100\frac{TP}{TP+FP} \tag{4.4}$$

where FN is the number of false negatives.

Two experts have annotated our data set in order to account for inter-observer variability. Comparison to inter-observer variability was computed as follows. Each expert was used as ground truth for computing $Sens$ and $Prec$ scores achieved by automatic detections. Inter-observer variability was computed by taking one of the experts as ground truth and the other one as detection output for computation of $Sens$ and $Prec$. The differences of $Sens$ and $Prec$ between observers and the automatic detections were analyzed using a Wilcoxon signed-rank test.

4.2 Structure segmentation

4.2.1 Obstructed lumen segmentation

The objective of this experiment is to validate our automatic luminal area segmentation method. Given an input frame we segment the obstructed lumen area and the we compare the curve that limits this area against curves annotated by two different experts. The obstructed lumen database contains 80 frames selected from 20 different explorations (see Fig.4.6 with some examples of the database). The database covers different appearances of obstructions and balances the presence of examples of all the different SI categories visually classified as 15 cases of Slight ($SI < 50\%$), 15 cases of Moderate ($SI \in (50\%, 70\%)$), 30 cases of Severe ($SI > 70\%$) and 20 frames of healthy trachea. For this database we have two different annotations made by two different experts in order to be able to compare the results with inter-observer variability. Fig.4.5 shows two examples of obstructed lumen frame and its manual segmentation.

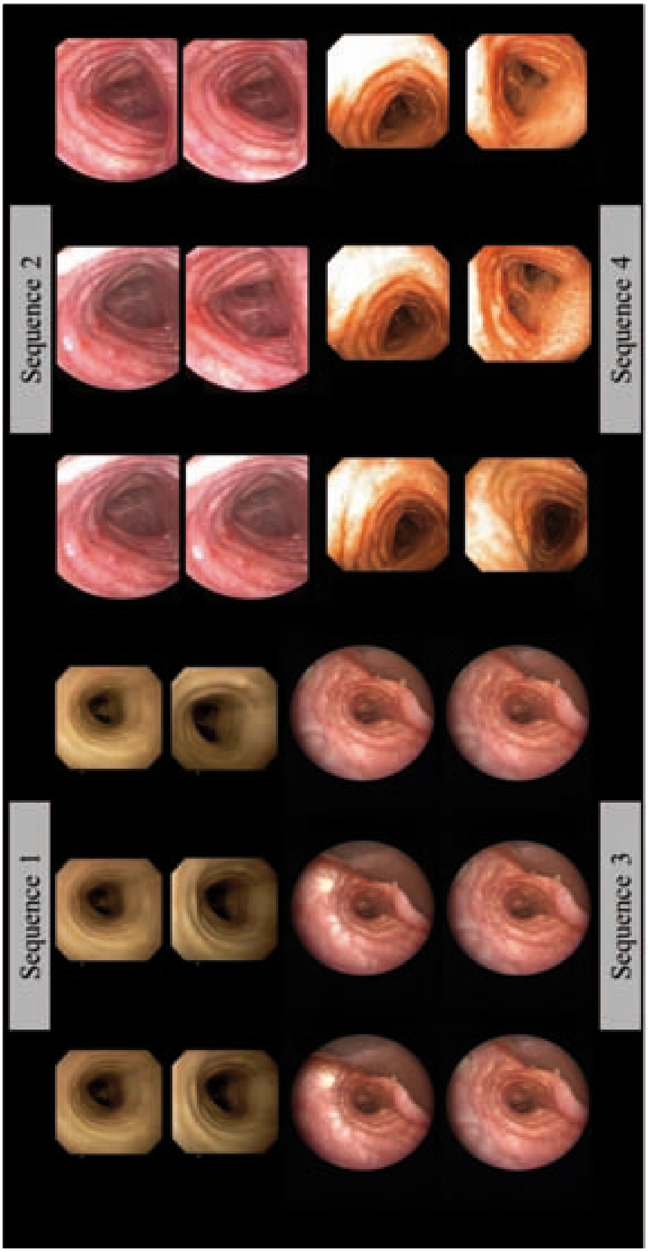


Figure 4.4: Tracheal ring dataset.

Manual curves annotated by experts were compared to our automatic segmentations by means of the average distances computed across curve angular parameter. Given two different curves in the polar domain, $\gamma_A(\theta) = (\theta, r_A(\theta))$, $\gamma_B(\theta) = (\theta, r_B(\theta))$, the average

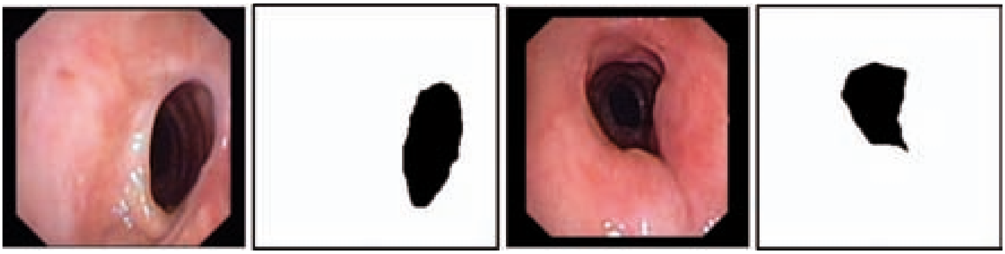


Figure 4.5: Examples of our obstructed lumen database: real obstructed lumen frames (first column) and its manual segmentation (second column) made by one of the experts.

distance between them is computed as:

$$AvD_B^A = ((\frac{1}{N_\theta}\sum_i abs(r_A(\theta_i) - r_B(\theta_i)))/s_x) \cdot 100 \tag{4.5}$$

for $\theta_i = 360i/N_\theta$ a uniform sampling of the interval $[0, 360]$ of step $h_\theta = 360/N_\theta$ and s_x the height dimension of the original image to account for differences in device resolution. For a given curve annotated by observer O_k, we will note by $AvD_{Aut}^{O_k}$ as the distance between the curve of the observer and the output of our method.

4.2.2 Reference ring segmentation

The objective of this experiment is to validate our automatic segmentations of the reference ring. Given an input frame our method provides as output the reference ring and we compare the curve obtained against the ones given by the two experts. The performance of this method is related with the output of tracheal ring segmentation, as part of the input of our reference ring segmentation comes from the output of ring segmentation.

The reference ring database contains 80 frames selected from the same 20 explorations used in the obstructed lumen database (see Fig.4.8 with some examples of the database). The database covers different appearances of rings. In order to build this database we imposed the constraint that the camera has to be always always centred at the carina. For this database we have two different annotations made by two different experts in order to be able to compare the results with interobserver variability. Fig.4.7 shows four different examples of healthy trachea with the first most external ring manually segmented.

Following the same strategy than for obstructed lumen segmentation, the validation of our reference ring segmentation will consist of comparing curves provided by our method and annotations made by experts. For a given curve annotated by observer O_k, we will note by $AvD_{Aut}^{O_k}$ as the distance between the curve of the observer and the output of our method. $AvD_{Aut}^{O_k}$ is calculated as depicted in Eq. 4.5.

4.3 Stenosis assessment

The objective of this experiment is to provide an automatic stenosis degree by using information from reference ring segmentation and obstructed lumen segmentation. Given two frames, one with and obstructed lumen segmented (output of the obstructed lumen experiment) and a healthy frame with the reference ring segmented (output of the reference ring

Figure 4.6: Obstructed lumen dataset. Some examples of bronchoscopy frames used to validate our method of obstructed lumen segmentation.

segmentation) we compare the two areas to calculate the percentage of obstruction. Finally, we compare this percentage with the one made by the two experts, in a visually way and using an off line program (such as imageJ).

The data set used for this experiment is composed by 20 new cases of patients with several degrees of stenosis ^{2} different from the ones for obstructed lumen and reference ring segmentation. Out of all this 20 cases, 3 were visually classified as 3 slight, 10 as moderate

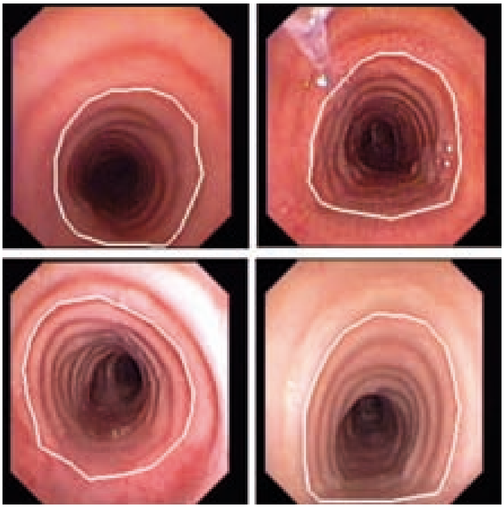


Figure 4.7: Examples of our reference ring database: four examples of manual segmentation made by one of the experts.

and as 7 severe obstructions. A pair of images showing a normal and an abnormal airway view were extracted from each of the cases. A mosaic showing the pair of images for some of the cases is shown in Fig. 4.10.

Experts assured that both normal and abnormal views were obtained with a same distance from the scope to either the obstructed area or the healthy one. The scope was kept centred so that both the entire circumference of the airway lumen for obstruction and a complete external ring for the case of a healthy shot can be visualized. Ground truth corresponding to this experiment was generated by two different experts by means of ImageJ software [42] -see Fig. 4.9.

The accuracy of the stenosis index SI has been quantified by comparing the output of our method obtained by means of 4.6 to the percentage obtained by the manual annotations of obstructed lumen and reference ring made by the same two experts. In this case AA_{Ref} refers to the airway area of a reference normal segment and AA_{Lumen} to the airway area of the abnormal one. AA_{Ref} is computed as the area enclosed by the first complete tracheal ring in a healthy segment, while AA_{Lumen} is given by the luminal area of the obstruction.The stenosis index determined by each expert measures will be noted by SI_{O_1} and SI_{O_2} respectively.

$$SI = \left(\frac{AA_{Ref} - AA_{Lumen}}{AA_{Ref}}\right) \cdot 100 \tag{4.6}$$

4.4 Anova parameter tuning for stenosis assessment

Stenosis degree assessment is based on the calculation of two different areas, one for obstructed luminal area and another for the area enclosed by the reference ring. The segmentation of these structures is validated against curves annotated by experts: in this case anova parameter setting is used to optimize the output of our segmentation methods in order to decrease the distance of our curves to the ones provided by experts.

---

[2] Bellvitge Hospital Barcelona. Sampling size for clinical trial.

Figure 4.8: Reference ring dataset. Some examples of bronchoscopy frames used to validate our method of reference ring segmentation.

Regarding tracheal structure segmentation for each of the data set proposed, a random sampling of 25 frames was used in the training stage. This stage aims to adjust the optimal scale, $[\sigma_j, \sigma_{j+k}]$ for reference ring segmentation and the optimal intercept, l for obstructed luminal area segmentation. The optimal parameter setting will be tested in the remaining 55 images of each database, using first, a 1-anova test over the test set (as in the training data), and second, the best optimal parameter will be tested using a t-test for paired data

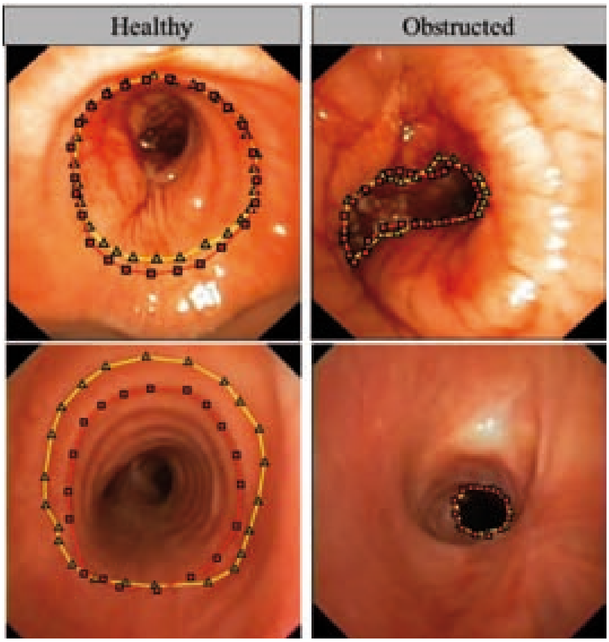


Figure 4.9: Examples of our stenosis database: pair of images (each row) with area of healthy (first column) and obstructed (second column) airway (first) segmented by expert 1 (yellow line) and expert 2 (red line).

to check the actual inference. All tests have been done at significant level $\alpha = 0.05$.

The random variable for statistical tests is a measure of segmentation accuracy against the annotations made by two different experts using ImageJ [42]. For a given curve annotated by observer O_k, we will note by $AvD_{\rho}^{O_k}$ -see Eq. 4.5- as the distance between the curve of the observer and the output of our method using a given parameter ρ. The parameter ρ represents the scale interval $[\sigma_j, \sigma_{j+k}]$ for tracheal ring segmentation and to the intercept l_k for lumen segmentation.

The variable and group of factors for the anova parameter setting is defined as follows. For the 2-anova stage assessing interaction between method ranking and observers, the anova groups are the observers for the first factor and the different parameters for the second one. The anova variable is, thus, $AvD_{\rho}^{O_k}$ computed for the 25 training frames. For the 1-anova test, groups are defined by the different parameters and an average observer. In this case the anova variable is given by averaging distance for the two observers: $AvD_O = (AvD_{O_2}^{O_1} + AvD_{O_1}^{O_2})/2$ for the observer group and $AVD_{\rho} = (AvD_{\rho}^{O_1} + AvD_{\rho}^{O_2})/2$ for the method ones.

The samplings defining the different anova groups for the method are given by uniform samplings of size N_σ for ring scales and N_l for lumen intercept:

$$\begin{aligned} (([\sigma_j, \sigma_{j+k}])_{j=1}^{N_\sigma - k})_{k=1}^{N_\sigma};\ \sigma_j &= \sigma_m + j(\sigma_M - \sigma_m)/N_\sigma \\ l_k &= l_m k/N_l \cdot (l_M - l_m) \end{aligned} \tag{4.7}$$

For the case of final stenosis calculation, in order to compare the performance of our method against the ground truth we have calculated the SI from each observers' annotation.

Figure 4.10: Stenosis dataset. Some examples of bronchoscopy frames used to validate our method of stenosis assessment. For each patient pair of images (each row) of healthy (first and third column) and its respective obstructed airway (second and fourth column).

The comparison between SI and SI_{O_k} was done using a 1-way anova with groups given by SI, SI_{O_1} and SI_{O_2}. As depicted before, tests were done at significant level $\alpha = 0.05$ for either the 1-way anova and the final t-test validation experiment.

Chapter 5

Experimental Results

We present in this chapter results of our automatic stenosis assessment system. As tracheal characterization structure (lumen and tracheal rings) and structure segmentation (obstructed lumen, reference ring) are part of the global stenosis degree calculation method, we first present a break down of the results of each stage of our strategy.

5.1 Structures characterization

Lumen segmentation is used to obtain obstructed airway area and in order to achieve this we first need to obtain a characterization of image pixels according to lumen content -using LEM maps- and then lumen centre is obtained to guide segmentation of the final luminal area.

5.1.1 Lumen centre detection

Lumen centre detection is a key point in our stenosis assessment system as it is used for both obstructed airway area and reference airway area. The experiment involved the partition of the feature space created with DRI and DGA. The final partition threshold -$Th_{LK} = 0$ 12- is set during a training stage using 30 frames of the database. DGA does not need of any parameter value and the optimal parameter value needed for DRI -$\sigma = 1$ 24- has also been set from the before mentioned training stage.

We present Precision and Recall results of lumen centre detection in Table 5.1. We use TL, FL and NL metrics as explained in Section 4.1.1. As it can be seen, Precision and Recall scores are over 93% regardless of the presence and the number of lumens in the images. We can also observe that only 2 lumens are missed (all in multi lumen images).

Type of image	**TL**	**FL**	**NL**	**Prec**	**Rec**
Bronchoscopy (single lumen) $\#lumen = 76$	76	1	0	98.70%	100%
Bronchoscopy (multilumen) $\#lumen = 30$	28	0	2	100%	93.33%
Bronchoscopy (no lumen) $\#total-images = 10$	0	2	10	80.00%	80.00%

Table 5.1: Precision and recall results on lumen centre detection.

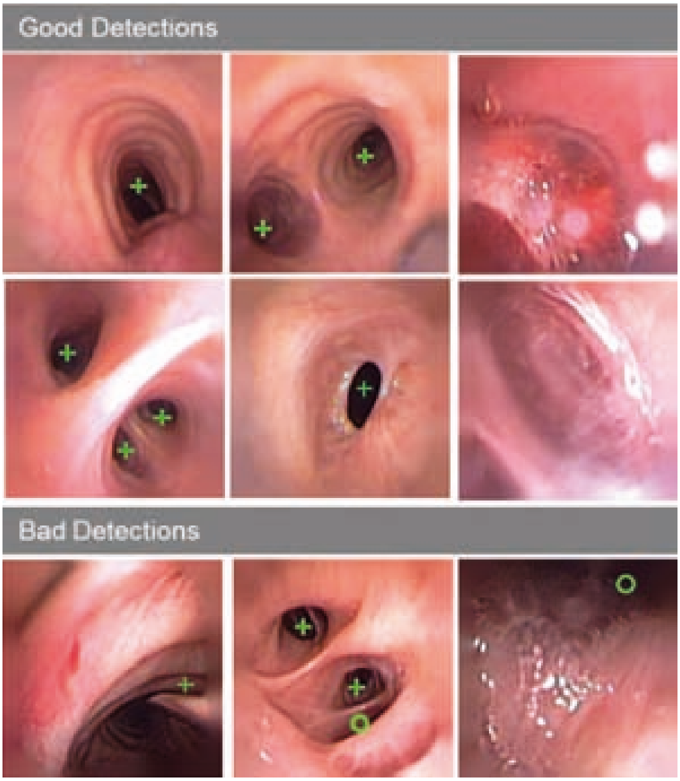


Figure 5.1: Qualitative lumen centre detection results. Good detections marked with green crosses and bad or missing ones with green circles.

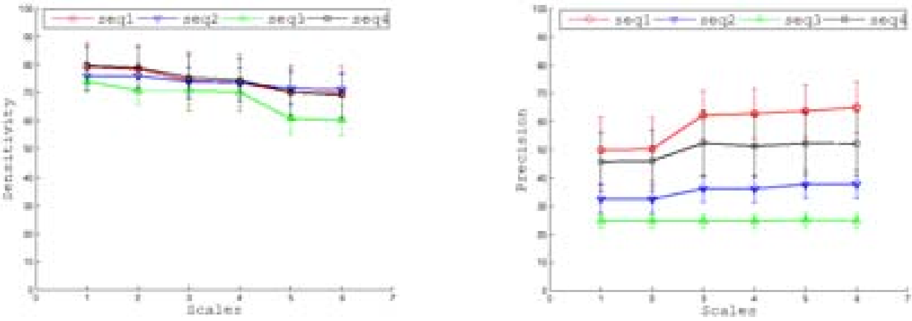
Another interesting result from this experiment is that our methodology can be used to assess lumen presence in the images: more precisely our method does not detect any center in 80% of the images without lumen.

Fig. 5.1 shows qualitative results including good and bad detections. The first 2 columns in the Figure show examples of good lumen centre detection in single and multi lumen images. Column 3 shows two examples of the potential of our method on detecting lumen presence: we can observe that no centre point is marked in the image.

Erroneous detections (marked with a green circles) of our method are shown in the bottom image. First column show a bad identification of the centre of the lumen, second column a missing centre in a multi lumen image and third column a show a detection when there is no lumen in the image. It is worth to mention that in some cases like the ones shown Fig. 5.1 it is unclear if our algorithm has not really performed well due to the fact that when making a ground truth sometimes there is a great variability on delimiting the lumen region -even the presence/absence of lumen in certain images depend of the experts' criteria-.

5.1.2 Tracheal Ring segmentation

This first stage is crucial to the performance of the overall system, as in the case that some rings are missed from this stage the posterior selection of one of them as reference ring may



(a) (b)

Figure 5.2: Results for each sequence: sensitivity (a) and precision (b).

lead to erroneous stenosis results if the chosen one was not really the reference ring. Before starting with tracheal rings segmentation we will expose the results of a preliminary study on checking which scales are optimal for this segmentation and the relationship of this scale value with image resolution.

Preliminary study on automatic scale definition This experiment checks for existing relations between the scale range selected for *NSGF* filtering and the resolution of the frames of a given video. The experiment consists of defining a set of scale ranges for a given reference sequence. This same scale range is used for the rest of the sequences but pondered according to the difference in resolutions between the reference and the new sequence -see Table 4.2-. The values for the reference sequence are set using a Phong illumination model [49] onto a synthetic trachea model simulated using sizes reported for human adults [51].

The main objective of this experiment was not to obtain the best scale ranges which offer best tracheal ring segmentation performance but to observe whether the use of a given scale range -pondered if necessary- leads to obtain a same performance trend for all the sequences. Thanks to this experiment [54] we can also infer the optimal scale range in one reference frame and we can apply it to the rest of the sequence.

Figure 5.2 shows the results for each sequence in terms of sensitivity and precision. We show the tendency of sensitivity and precision interval for different range of scales (see Table 4.2) considered for each sequence. A remarkable issue is that, although these scale ranges are different -although equal if the ponderation is supressed-, all four sequences present equal performance profiles and attain these maximum values simultaneously. This validates our illumination model for trachea appearance as a tool for automatically selecting NSGF scales in terms of the bronchoscope resolution.

It has to be noted that this experiment was focused on the definition of the best scale range and how this value, when calculated for a sequence, can be easily inferred to other sequences. Considering this we have not included in this processing operations such as area filtering or decreasing scale-pattern which are part of our tracheal ring segmentation method.

Tracheal Rings Segmentation Once the best scale range has been selected we can proceed to obtain tracheal rings segmentation using the methodology depicted in Section 3.2.2.

Table 5.2: Comparison to Inter-observer Ranges.

	AUT		*IO*	
Validation	*Sens*	*Prec*	*Sens*	*Prec*
Seq_1	[71 16 96 20]	[57 67 80 08]	[71 21 99 61]	[71 70 98 87]
Seq_2	[66 11 94 01]	[52 22 67 72]	[56 13 72 84]	[56 32 72 59]
Seq_3	[60 29 83 53]	[51 25 60 97]	[59 56 78 41]	[59 70 78 23]
Seq_4	[57 87 92 72]	[51 16 75 90]	[62 66 86 35]	[62 84 86 08]

Table 5.2 and 5.3 report the statistical analysis comparing automatic detections (labelled as *AUT*) and inter-observer variability (labelled as *IO*) for the four video sequences (rows). Table 5.2 reports *Sens* and *Prec* ranges (given by mean ± standard deviation, computed for the two experts) and table 5.3 the Wilcoxon signed-rank test $p-val$ and confidence intervals for the difference in means between *AUT* and *IO*.

First, we observe that there is not a high agreement between experts, especially for Seq_2 and Seq_3. Expert disagreement is mainly due to unmarked rings at most inner (Seq_3 and Seq_4 in figure 5.3) and outer parts (Seq_1 and $Seq2$ in figure 5.3) of the trachea (as illustrated in figure 5.3). Such variability in most outer and inner rings significantly increases for pathological cases (Seq_2 and Seq_3).

Second, it is worth noticing that our detection includes the union of rings segmented by the two manual trials. This increases sensitivity scores for sequences with a lower agreement between experts, such as the pathological Seq_2 and Seq_3. It follows that sensitivity results compares to inter-observer variability for Seq_1, Seq_3 (although with positive bias in differences) and Seq_4 ($p-val > 0\ 3$ in table 5.3), but for Seq_2 our method has sensitivity ranges above inter-observer agreement.

Third, concerning precision ranges, our methodology might drop its performance up to a 20% compared to inter-observer ranges. This is mainly due to an over-detection of rings rather than inclusion of non-anatomic artifacts. Figure 5.4 shows the results obtained as well as ground truth for four representative frames from four different sequences. For a better visualization, our detections are shown in black lines and ground truth in thicker white lines. We observe that, as expected, all rings are detected and include the union of the rings marked by experts (as comparison to images in figure 5.3 shows). This is the main source of precision

Table 5.3: Wilcoxon signed-rank test and confidence intervals for difference in means between *AUT* and *IO* and $p-val$.

	$Wilcoxon(CI\ \ p-val)$	
Validation	*Sens*	*Prec*
Seq_1	[−11 75 4 74] 0 4856	[−23 58 −12 08] 0 0004
Seq_2	[7 94 24 14] 0 0006	[−11 06 −1 48] 0 0183
Seq_3	[−0 70 6 55] 0 3481	[−17 09 −8 61] $9\ 27 * 10^{-6}$
Seq_4	[−6 70 8 27] 0 9031	[−18 18 −3 68] 0 0056

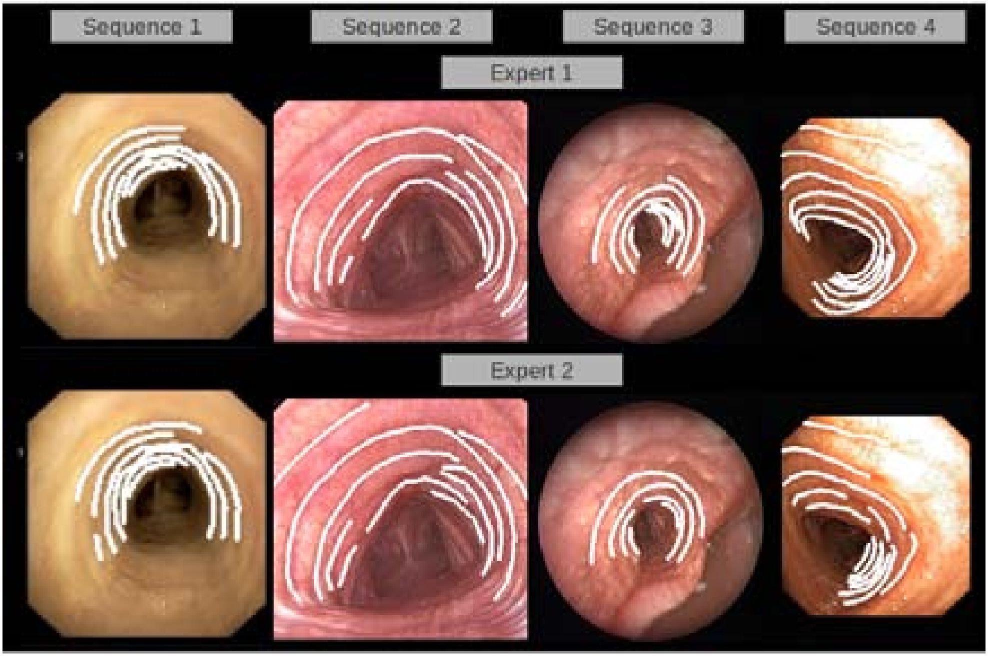


Figure 5.3: Inter-observer variability in manual segmentations.

dropping in $Seq1$. Still, for some cases there are some extra structures decreasing our global precision. In particular, part of the end probe, as well as, calcified spots are included at some frames of Seq_3. Also our method might include the ring continuation on the esophagus (bottom part of Seq_2 images in figure 5.4) which is generally discarded by clinical experts.

Figure 5.5 shows the results obtained for the best scale ranges. Green lines are the ground truth and blue lines the automated detection. We observe that, as expected, all rings are detected, while there are some extra detected structures which have a negative impact in precision results. There are two main sources of this performance decrease. The first problem is that there are false positive rings in the carina in all sequences. Also,

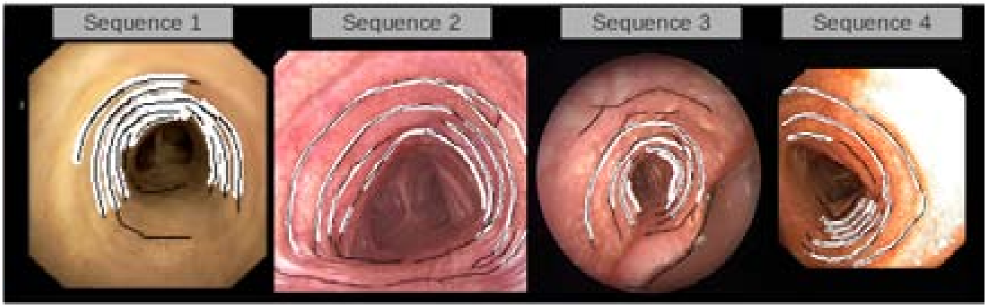


Figure 5.4: Visual Assessment of Automatic Ring Detection.

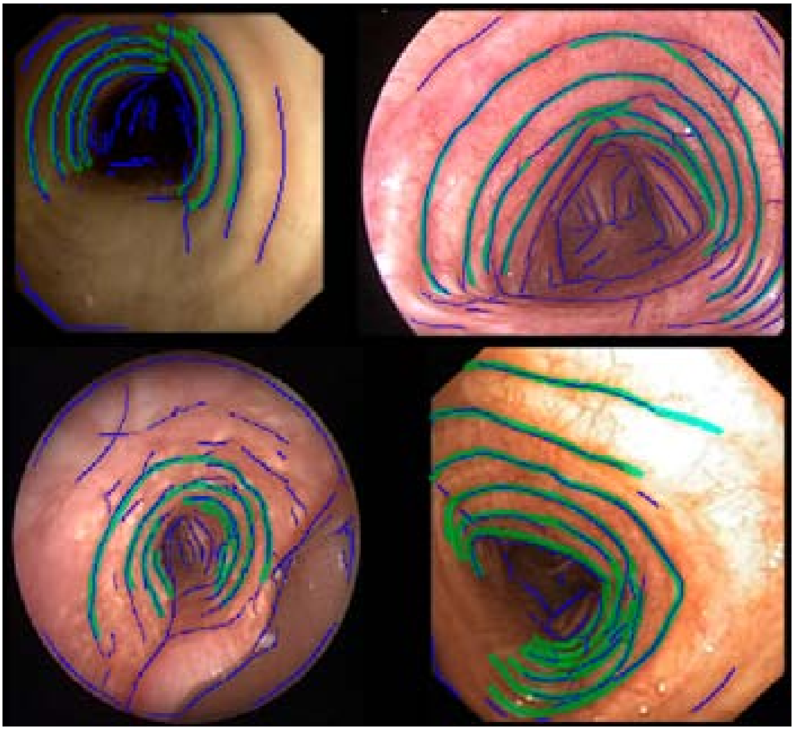


Figure 5.5: Qualitative results for each sequence. Green lines are the groundtruth and blue lines our detection.

in the lowest precision sequence, there is a rigid artifact that we detect as a ring. Such surgical device drastically drops precision graphics in Figure 5.5. Finally, there are some rings (especially the most internal and external ones) properly detected by NSGF that were not identified by the expert.

As it can be seen, our experiments show that our tracheal ring segmentation method can retrieve most clinically relevant tracheal rings. This is a promising result that indicates the value of the proposed methodology as a tool for helping bronchoscopists in the operating room in the sense that it will be further used to automatic select and reconstruct a reference ring to be used as part of the final stenosis assessment.

5.2 Structures segmentation

Structure segmentation is used to obtain reference airway area: our processing stage has two different steps: the first involves obtaining all the tracheal rings in a given input image and the second involves the selection of the reference ring -the most external- and the posterior segmentation of the area that it encloses.

We present in this section results of the segmentation of tracheal structures which also include the use of ANOVA-based parameter setting explained in the previous chapter.

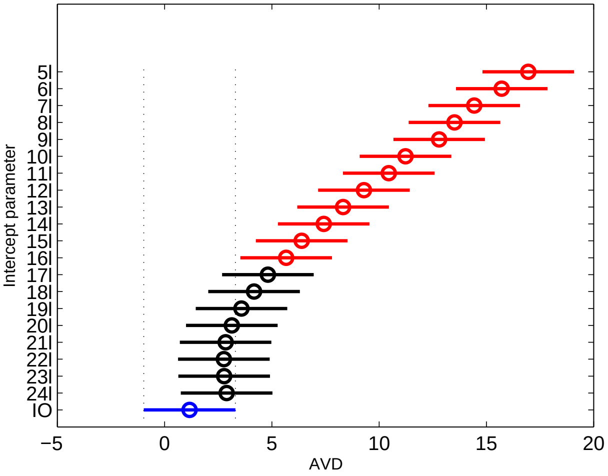


Figure 5.6: Multicomparison plot for the train set of the lumen segmentation.

5.2.1 Obstructed lumen segmentation

Regarding obstructed airway segmentation we will follow validation strategy using ANOVA as explained in Chapter 4.4. In this case we compare the performance of our method with two different observers. For this experiment, the anova groups are the observers for the first factor and the different parameters for the second one.

Results of 2-anova test show that there is not interaction between observers and intercept value (with $p - val = 1$), no significant differences between observers (with $p - val = 0\ 9653$) and detects significant differences across intercept values (with $p - val < 10^{-16}$).

The first result implies that the optimal intercept value does not depend on the observer: this implies that 1-anova can be applied to obtain the optimal intercept l. Results of this 1-anova test are significant ($p - value < 10^{-16}$) and best performers are selected by using a Dunnet multicomparison test. Fig. 5.6 shows the results of the multicomparison test comparing the performance of each intercept (labelled lk) to inter-observer (IO). The plot shows intervals for mean differences as in Fig. 5.9. Each level mean is represented as horizontal lines centered at the level mean and vertically distributed according to the level. The bottom line corresponds to the observer group and the remaining to the sampling l_k. We do not show results obtained with $k > 24$ as the likelihood map has not a minimum for all frames. Fig. 5.7 shows the results of the multicomparison over the test set. We observe that a same trend is kept for both training and test sets: the most permissive we are when characterizing lumen pixels -frontier of the feature space closer to the one provided initally by k-means-, the less precise we are when segmenting the lumen region. Conversely, the most strict we are with the threshold value the less non-lumen information we include in our final region. This trend is kept up to a given intercept value -$l20$ to $l23$-: from this intercept values and go on the final luminal area segmentation also gets worse.

The second result indicates that any of the observers can be used as ground truth to which compare our method: considering this we have built up a new observer as the mean aggregation of the two original observers AvD_O.

The final result from 2-anova test directly implies that there is an optimal parameter

l_k	AvD_O	AvD	$AvD_O - AvD$	$p - value$
l_{20}	$[1.43, 1.88]$	$[1.78, 2.20]$	$[-0.55, -0.10]$	0.0042
l_{21}	$[1.43, 1.88]$	$[1.61, 1.93]$	$[-0.28, 0.05]$	0.1874
l_{22}	$[1.43, 1.88]$	$[1.63, 1.95]$	$[-0.27, 0.01]$	0.0632
l_{23}	$[1.43, 1.88]$	$[1.75, 2.11]$	$[-0.45, -0.09]$	0.0032

Table 5.4: Statistical analysis results for the lumen segmentation method.

configuration with yields with better segmentation results. Table 5.4 reports the statistical analysis for the t-test comparing AvD_O with the best visual performer of Fig. 5.6 and Fig. 5.7, which in this case corresponds to AvD_{l21}. We show the same statistics as in Table 5.5. There are not significant differences ($p - value = 0\ 18$) between observers and our method. Out of all the possible intercept values we will use for the following stages $l21$, as it is the one which leads to obtain a higher $p - value$, meaning this that our method is closer to observer for this intercept value. Consequently we will use $l21$ for SI computations.

Finally we present in Fig. 5.8 some qualitative results using the same color coding as in Fig. 5.11. We observe that our method achieves performance results comparable to the ones provided by both observers.

5.2.2 Reference ring Segmentation

Regarding reference ring segmentation we will follow the same validation strategy applied for obstructed airway segmentation. In this case we also compare the performance of our method with two different observers. For this experiment, the anova groups are the observers for the first factor and the different parameters for the second one.

According to the 2-anova test there is no interaction between observers and scale ranges ($p - val \quad 1$), no significant differences between observers ($p - val = 0\ 9917$) but a significant

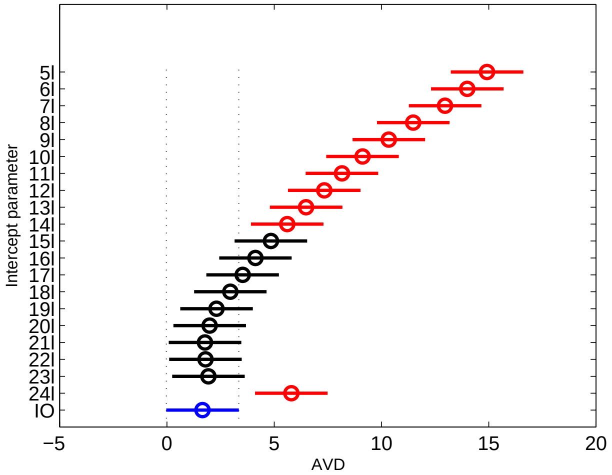


Figure 5.7: Multicomparison plot for the test set of the lumen segmentation.

$[j, j+k]$	AVD_O	$AVD_{[2\ 15]}$	$AVD_O - AVD_{[2\ 15]}$	$p - value$
$[1, 15]$	$[2.68, 4.71]$	$[3.65, 5.22]$	$[-1.58, 0.11]$	0.0883
$[1, 16]$	$[2.68, 4.71]$	$[3.71, 5.32]$	$[-1.66, 0.03]$	0.0606
$[2, 15]$	$[2.68, 4.71]$	$[3.65, 5.22]$	$[-1.58, 0.11]$	0.0883
$[2, 16]$	$[2.68, 4.71]$	$[3.71, 5.22]$	$[-1.66, 0.03]$	0.0606
$[3, 15]$	$[2.68, 4.71]$	$[3.83, 5.68]$	$[-2.01, -0.09]$	0.0324
$[3, 16]$	$[2.68, 4.71]$	$[3.90, 5.77]$	$[-2.09, -0.17]$	0.0219
$[4, 15]$	$[2.68, 4.71]$	$[3.81, 5.64]$	$[-1.97, -0.07]$	0.0351
$[4, 16]$	$[2.68, 4.71]$	$[3.87, 5.74]$	$[-2.06, -0.15]$	0.0238

Table 5.5: Statistical analysis results for the ring segmentation method.

difference across scale values ($p - val < 10^{-16}$).

These results indicate that the performance of our method does not depend on the observer so that best performers can be selected by comparing to an average observer using the 1-anova test described in Section 4.4. This 1-anova test detects a significance ($p - value < 10^{-16}$) in segmentation accuracy depending on the chosen scale. This significance is further explored using a Dunnet correction for a multicomparison test.

Figure 5.9 shows the results of the multicomparison plot for the training set of reference ring segmentation. For the sake of notation each scale range $[\sigma_j\ \ \sigma_{j+k}]$ is labelled as $[j\ \ j+k]$ and the average inter-observer is labelled as IO. The plot shows intervals for mean differences for a representative set of the scale ranges. Each level mean is represented as horizontal lines centred at the level mean and vertically distributed according to the level. Those scale ranges that are not significantly different from IO (blue line) are plotted in black, while scales with a significant different performance are depicted in red. We can observe from Figure 5.9 that our method offers comparable performance to IO when scale ranges are [1 14] [1 15] [1 16] [2 15] [2 16] [3 15] [3 16] [4 14] [4 15] [4 16]. It has to be noted that the best scale ranges tend to include higher scale values as we aim to search for the most external ring in the image.

Fig. 5.10 shows the multicomparison between $[j\ \ j+k]$ and IO for the test set with the same labelling used in Fig.5.9. This test is done to confirm the trends indicated by the previous multicomparison test. In this case we can observe that the trend in significant differences in performance is kept -best ranges are again [1 14] [1 15] [1 16] [2 15] [2 16] [3 15] [3 16] [4 14] [4 1
.

Table 5.5 reports the statistics for the t-test comparing AvD_O to $AvD_{[j\ j+k]}$ for the ranges mentioned before that visually achieved the least difference as observed in Fig.5.10. We show confidence intervals for the means of AvD_O, $AvD_{[j\ j+k]}$ and the difference $AvD_O - AvD_{[j\ j+k]}$, as well as, the p-value for the t-test. In spite of a small negative bias, average distances are within average inter-observer variability ranges with $p - value > 0\ 05$.

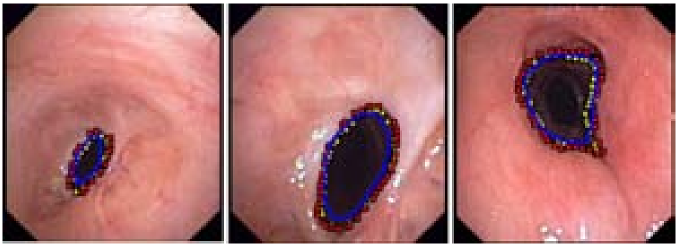


Figure 5.8: Qualitative results for the lumen segmentation with l_k parameter fixed in $k = 21$.

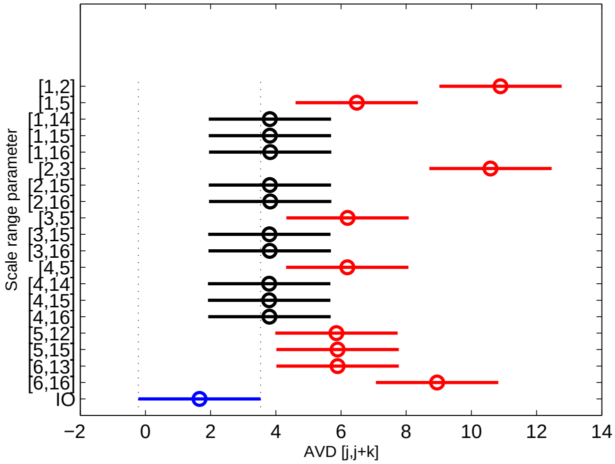


Figure 5.9: Multicomparison plot for the train set of the reference ring segmentation.

Although we could use any of the 8 scale sub-ranges, we will use [2 15] as it offers the best performance for the testing set and equal performance to the best candidates in the training set. Consequently, this scale range configuration will be used to compute SI.

Figure 5.11 shows some qualitative examples of the tracheal ring segmentation. The blue line corresponds to our method whereas the yellow triangles and red squares correspond to each of the observers. The first two images show two examples in which our method accurately approximates the curves outlined by experts. The last one is an example of a

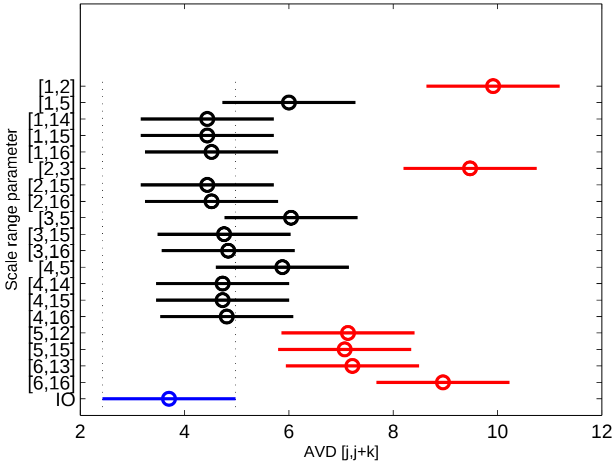


Figure 5.10: Multicomparison plot for the test set of the reference ring segmentation.

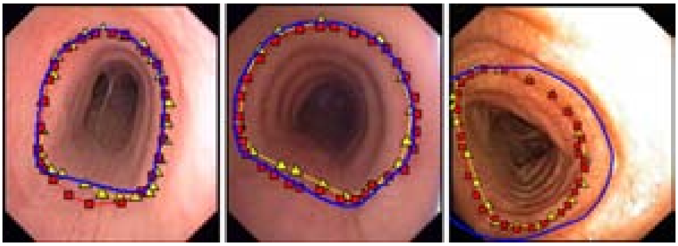


Figure 5.11: Qualitative results for the tracheal ring segmentation with interval parameter $[\sigma_j \ \ \sigma_{j+k}]$ fixed in $j = 2$ and $k = 15$.

point of view not centred at the carina with the most external ring partially out of the frame. Following our annotation convention, experts outlined the first fully complete ring, while our method detected the most external one as it only had a small portion outside the image frame.

5.3 Stenosis assessment

Regarding stenosis assesssment we will only show performance results considering that we are using the optimal intercept value l_{21} and scale range [2 15] which best approximate the response of our tracheal structure characterization to the one provided by experts. In this case we will only perfom a 1-anova test as we are not aiming to find the best parameter configuration of our method but how it does compare with human observers.

According to this 1-anova we observe that there are not significant differences across observers and our automatic SI index ($p - value > 0 \ 83$). This is also confirmed in the multicomparison plot of Fig. 5.12. Finally a t-test for paired data comparing SI_{AUT} with the average expert score SI_O could not find any significant differences ($p - value = 0 \ 96$).

Table 5.6 reports confidence intervals for SI_O, SI_{AUT}, the difference $SI_O - SI_{AUT}$ and the corresponding p-value. It is worth noticing SI_{AUT} agrees to observer scores ±10%, which is clinically acceptable according to the literatue [41].

In order to present our results in more detail we present in Table 5.7 the SI values for each of the different cases/patients and its corresponding SI index value for each of the mentioned ground truths. Table 5.7 describes for each patient (first column) the visual degree of airway narrowing (second column) and the measurements of the stenosis index using Image J software by different experts (third and forth column) and finally our automatic detection.

By observing Table 5.7 we can see that there is a big discrepancy between experts for some concrete cases such as 1 3 8 12 18 where there is a difference up to 20% in the final stenosis index.

We present in Fig. 5.13 some representative examples of the output of our method in comparison with manual annotations for normal airway (top images) and tracheal stenosis (bottom images). It is worth noticing the high agreement achieved among observers and our

SI_O	SI_{AUT}	$SI_O - SI_{AUT}$	$p - value$
$[72.81, 83.01]$	$[69.25, 87.03]$	$[-10.3, 9.8]$	0.962

Table 5.6: Statistical analysis results for the systematic stenosis quantification.

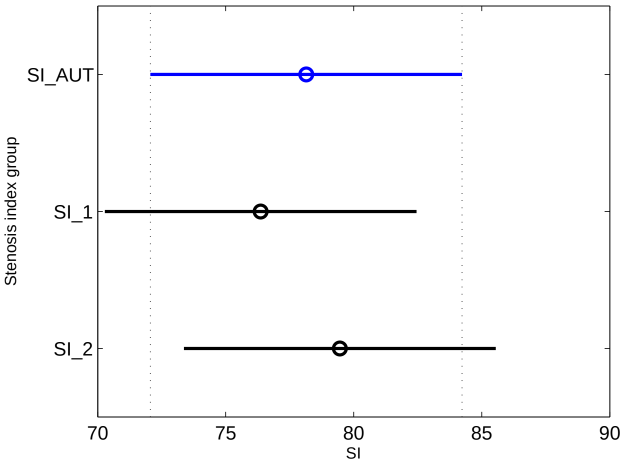


Figure 5.12: Multicomparison plot for the test of stenosis assessment.

Patien	**Visual**	**SI Obs1**	**SI Obs2**	SI_AUT
1	Moderate	57%	74%	81%
2	Moderate	71%	76%	95%
3	Moderate	56%	97%	84%
4	Severe	91%	80%	91%
5	Slight	84%	83%	78%
6	Severe	91%	96%	50%
7	Moderate	82%	91%	77%
8	Slight	33%	76%	66%
9	Moderate	67%	76%	96%
10	Moderate	76%	73%	91%
11	Moderate	80%	94%	100%
12	Severe	80%	43%	79%
13	Slight	75%	88%	76%
14	Moderate	87%	88%	69%
15	Moderate	79%	83%	85%
16	Severe	80%	79%	69%
17	Severe	80%	80%	82%
18	Severe	86%	64%	15%
19	Moderate	69%	50%	80%
20	Severe	95%	92%	93%

Table 5.7: Description of tracheal stenosis quantification.

method.

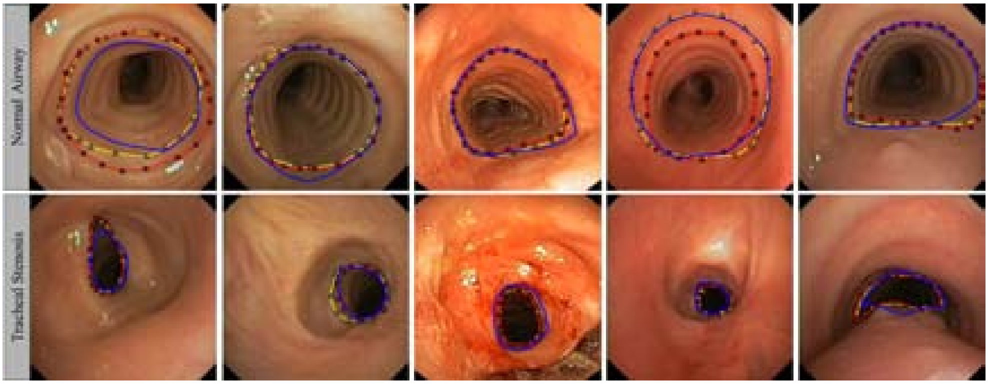


Figure 5.13: Qualitative result for the calculation of the stenosis index. Normal airways in the first row and its paired abnormal frame in the second row. Blue lines correspond to the automatic method and yellow triangles and red squares correspond to the ground truth of the different experts.

Chapter 6

Discussion

We have presented the main results of the different methods developed in the context of this thesis in the previous chapter. In this chapter we will deepen the details of some of these results, following the same order than the used in their explanation, that is, we will start by discussing some interesting results from structure characterization methods, followed by structure segmentation and ending with a general discussion on the performance of the final stenosis assessment method.

6.1 Structure characterization

The two main structures that we need to characterize are the tracheal rings and the luminal region. Both structures are characterized by considering the defined geometrical and appearance physical models of the trachea. In this section we will discuss in depth some of the results along with exposing some other domains where the developed techniques may be of use.

6.1.1 Lumen centre detection

As it has been explained in previous chapters, the posterior segmentation of both structures need of a good definition and detection of the lumen center therefore the performance of this stage is crucial to the performance of the global method.

Our method has been proven to perform accurately for the possible cases we may encounter, that is, presence of multiple lumen -bifurcations of the bronchial tree- or in absence of lumen -exploratory frames-. The performance of our method depends on the definition of the feature space, more precisely in how we define the border which separates lumen and non-lumen pixels. In this case we use as borders the ones provided by k-means clustering method and the only source of errors related to this could appear in uniformly illuminated scenes where we do not have strong gradients which make it possible a differentiation using DGA values.

Regarding this method, it is worth to mention that the detection of the lumen centre can be useful for several applications, such as scene description, 3D reconstruction processes or helping in computer aided diagnosis. Moreover, by detecting accurately the lumen centre we can potentially obtain the navigation path inside the organ which could be useful for quality assessment purposes or the following-up of injured tissues. Results of our lumen center

Index	Type	Resolution
101 − 200	Colonoscopy [3]	[500 × 577]
226 − 250	Colonoscopy (no-lumen) [3]	[500 × 577]

Table 6.1: Description of lumen database.

detection method can also be used to automatically detect bifurcations, as our method is able to detect accurately the presence of multiple lumens.

Finally, it is also important to mention that our algorithm can also be applied to other type of endoscopic images offering similar performance results. More precisely we presented in [53] results of lumen center detection in colonoscopy images. It is clear that in order the method to be applicable a new feature space specific for colonoscopy images should be created from a training set. The database used for this experiment was the same as the one used in our lumen centre detection but with colonoscopy frames. Table 6.1 describes which are the frames and its configuration. We show in Table 6.2 results for lumen center detection in colonoscopy. As it can be seen, our method reaches good performance values for colonoscopy frames although in this case we have to consider that there are differences in how lumen appears in colonoscopy and bronchoscopy. More precisely, lumen in bronchoscopy tends to appear as circular whereas its shape can vary largely for colonoscopy frames, as the clinician has more freedom to choose the point of view on the scene and consequently we cannot guarantee a frontal view and the gradient profile is less concentric

6.1.2 Tracheal rings

Tracheal ring detection has been proven as a challenging task because there are many structures, not necessarily anatomical, that have similar appearance in images to tracheal rings. We have introduced in this thesis a geometrical and appearance model aiming to differentiate and enhance ring structures from other valley structures in the image. More precisely our geometrical structure avoids any response at the carina and minimizes the impact of surgical devices.

Our contribution in this field supposes an important breakthrough with respect to intensity-based methods applied into transformed grey level bronchoscopic images. Regarding this, Figure 6.1 shows two frames segmented using only intensity (second row) and our geometrical appearance model (first row). Detections exclusively based on image intensity include many responses at vessels and the carina. These responses do not follow the ring geometric characteristics and, thus, are suppressed in our model.

Another interesting contribution is that our whole tracheal rings characterization methodology is robust as long as we can obtain a valid reference frame that we can use for the whole

Type of image	TL	FL	NL	Prec	Rec
Colononoscopy $\#lumen = 77$	75	4	2	94.94%	97.40%
Colononoscopy $\#nolumen = 18$	0	2	16	87.5%	87.50%

Table 6.2: Precision and recall results on lumen centre detection in colonoscopy.

---

[3] Beaumont and St. Vincents Hospital Dublin

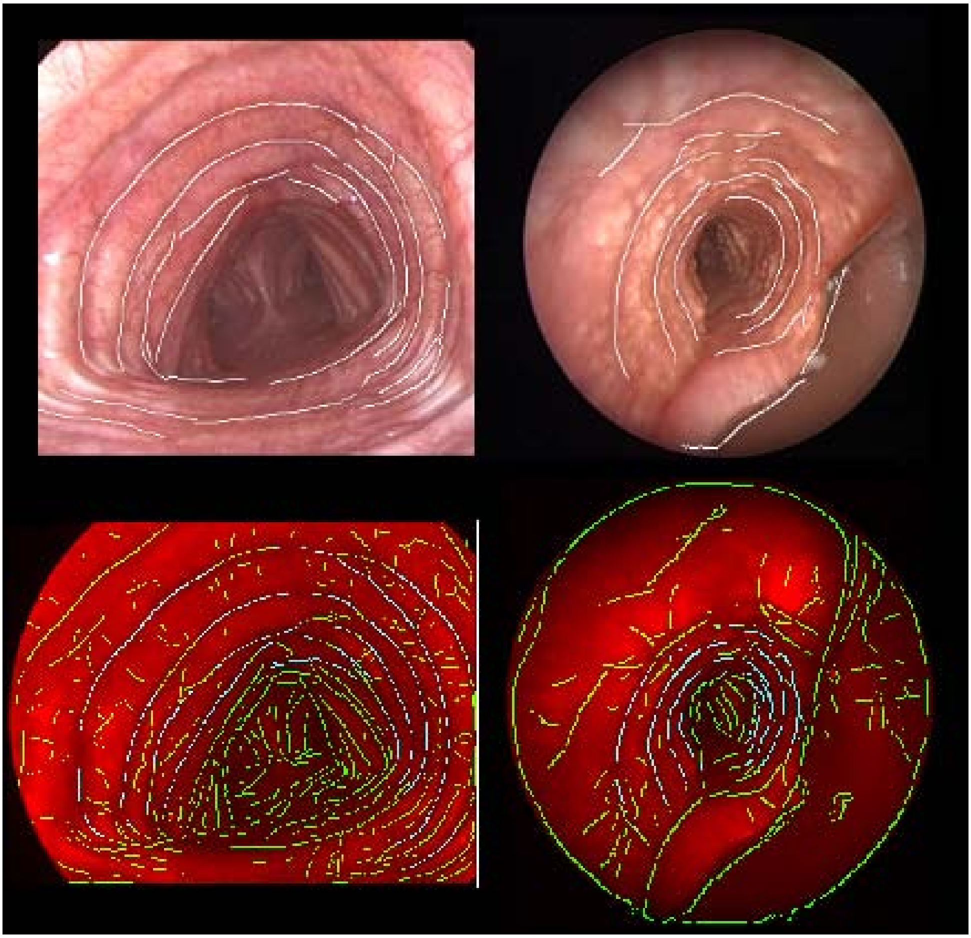


Figure 6.1: Advantages of combining appearance and geometry (1st row) of the tracheal rings compared to intensity-based approaches (2nd row, images courtesy of [55]).

sequence. Moreover, we have also proven that there is a strong correlation between the scales of the rings and the resolution of the video frames. For this reason, we can automatically infer the scale parameter in new cases from a given sequence where we have obtained the optimal scale parameters of the method.

6.2 Structure segmentation

Before starting with the discussion on how structure segmentation methods perform it is important to mention that their performance depends greatly on the first characterization stage, for instance, if lumen centre is misdetected it will be impossible to obtain an accurate lumen segmentation. Consequently and regarding the discussion of segmentation methods, we will analyse their performance isolating those error sources that are not attainable to them but to previous stages.

6.2.1 Obstructed lumen segmentation

With respect to obstructed lumen segmentation, we have observed that our method offers a comparable performance to clinicians although we have observed that there are some discrepancies in the segmentation results. In this sense we can divide sources of errors in two different groups: those related to how ground truth has been created and those directly attainable to our method.

The first one is related on the validation of our methods. Our experience has proven that the definition on the lumen region itself is indeed a challenging task and even clinicians do not agree when delimiting the area that the lumen occupies in a given frame. In this case we have to consider that our method can be close to one of the two experts but as our validation is against a mean observer, our performance score may be damaged by the great difference between two observers.

Regarding errors directly related with how our method has been designed we have discovered that the main source of errors is related to the use of a unique threshold value (l_k) for all the frames regardless of the sequence they come from. As can be thought, a given threshold value may provide with optimal segmentation results for all the frames from the same sequence but this parameter configuration may not provide same performance levels for other sequences.

One possible solution could be to automatically adapt the threshold for each sequence by choosing a representative frame, which can coincide in time with the selection of the obstructed frame. Another possibles solution, which was tested on colonoscopy frames for Non Informative Region definition, could consist of the definition of an optimal threshold value for each frame: this optimal value would be the one that separates better lumen region from the rest of the image. In this case the optimal border of the luminal region will be the one which offers maximum difference of energy between the two regions, as explained in [5].

6.2.2 Reference ring segmentation

Concerning reference ring segmentation we have identified two main groups of sources of errors, as for the case of obstructed lumen, one related on the definition of reference rings by clinicians and another directly related with our method. In this case both sources of error are directly related as it will be exposed next.

Regarding direct errors from our method it is worth to mention that the performance of our method is related to whether the external ring appears completely and partially isolated in the image or not. The performance of our method is clearly comparable to the one provided by experts although two special cases were identified: 1) Overestimation and 2) Underestimation.

The first group of errors is again related on how experts have annotated our database. There are some case where the first ring is not clearly visible but clinicians are inherently able to identify them; in this case our method needs of stronger evidence of this presence to detect the ring and will provide as reference ring the next visible one -second reference ring-, as it is shown in Figure. In these cases we will offer an underestimation of the reference ring.

For cases where the external ring surpasses image borders our method will tend to close the structure outside the frame: in this case we will have an overestimation of the tracheal ring (first and second case of the second row in Fig. 5.11). This is quite a frequent issue as can be observed in Table 5.5 with the negatives values in the CI intervals.

Finally we have another scenario where the first ring is not clearly visible due to either an incomplete profile lying outside the frame or an appearance deviation from the valley model. Incomplete profiles mainly arise from a camera point of view not centred at the carina, which distort the concentric disposition of rings. The impact of such cases is easily minimized by standardizing the acquisition protocol for stenosis assessment, imposing how healthy and abnormal frames should be selected.

6.3 Stenosis assessment

We have presented in this thesis the first, up to our knowledge, automatic tool for the systematic quantification of tracheal stenosis. Our methodology presents a breakthrough for stenosis assessment as it provides for the first time an accurate measurement of tracheal obstruction from the analysis of videobronchoscopy images, which has a potential to objectively compare and track stenosis, decide when it reaches a critical lumen and obtain the measurement to choose the right stent. Our proposal allows a fast and accurate calculation of stenosis degree which can be performed by clinicians in an intuitive way without altering the intervention protocol. Therefore, our methodology needs very little human interaction, although it requires a carefully choosing of the healthy and abnormal frame, especially regarding to whether the distance from the scope to the target is the same for both frames. If this happens, the performance of our method has been proven to be comparable to expert clinicians in a off-line methodology.

One key feature of our methods is that it does not need to rely on a previous CT scan to obtain stenosis degree as our calculations are performed only by means of the analysis of videobronchoscopy frames. Our methodology seems especially appropriate for the case of stenosis analysis in children, as it allows for an accurate calculation without needing to radiate the patient [25]. Moreover, our methodology is also capable to cope with dynamic stenosis, a condition in which airflow is decreased or even stopped as intrathoracic preassure increases. It can be affect only the posterior wall of the trachea or all the cartilage support, in which case it is , also referred as tracheomalacia [7]. In these cases patients do not present a fix obstruction degree but a variable one which can change as the patient breathes or coughs. Current procedures based on CT scan information can only provide a static image at maximum inspiration and at maximum expiration. Our methodology allows multiple data acquisition for stenosis calculation to obtain accurate obstruction intervals for a given patient, which can be key to certify tracheomalacia condition.

The analysis of the results presented show that our methodology reaches performance comparable to physicians in tracheal structures characterization and stenosis assessment in a computational time that could reach intervention time. This is a main advantage over off line methods based on manual annotations using ImageJ. Still there are some topics that deserve some discussion.

The computation of SI relies on the identification of a complete ring in images of a healthy segment. This can be difficult to obtain in severely ill patients (like the one shown in Fig. 6.2) with pathologies such as diffuse tracheal inflammation (post radiotherapy), infiltration (amyloidosis) or cartilage ring dismorfia (Tracheobronchopathia osteochrondoplastica). In such cases, our method could under estimate the degree of obstruction, yielding an optimistic diagnose. On one hand, these cases are not common in tracheal stenosis diagnosis. On the other hand, they could be solved by computing the reference area using the lumen segmented from a non-obstructed segment. From the clinical application point of view this would only require asking the physicians to mark as healthy frame one where the lumen can be observed without obstruction and indicate to the system that reference is lumen-like.

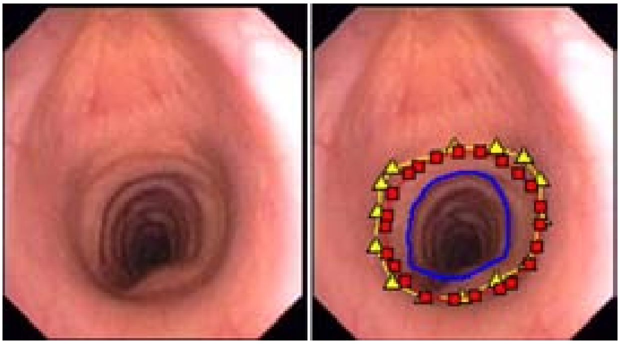


Figure 6.2: Example of approximate ring annotation for the two experts (yellow and red lines) and the ring detected by our methodology.

Regarding tracheal ring segmentation, we have observed that it is very difficult to obtain adjacent healthy tracheal rings in sever complex stenosis. Clinicians in these rare cases calculate stenosis degree by approximating the original lumen region from the obstructed image. Our method inherently will adapt to the first available visible and complete ring (see Fig. 6.2), as our algorithm is designed to recover the most external and complete ring in the healthy frame in a straight regular lumen. Consequently, our method will not offer a good performance in those cases where there is no visible ring or the lumen of the airway is not straight and regular, as in main or lobar bronchi. Nevertheless, our proposal is already prepared to cope with this problem, which could be easily in the following way: instead of asking the physicians to mark as the healthy frame one in which tracheal rings are visible, we can ask him/her to mark as healthy frame one where the lumen can be observed without obstruction.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

Image processing methodologies are becoming an important tool for using in medical imaging for diagnosis, treatment and intervention. We have presented in this thesis the main contribution based on a **novel strategy for automatic real time stenosis degree calculation** from analysis of videobronchoscopy frames. The calculation of the stenosis index is based on the comparison of the area delimited by the first external ring in a healthy frame against the obstructed area delimited by the lumen. We have assessed the performance of each tracheal structure characterization and segmentation by dealing with four main challenging problems: Appearance Artifacts, Alien Artifacts, Geometric Artifacts. This challenging problems have been solved by developing a multi-stage system that involve: Image Preprocessing, Structure Characterization using our Physical Model and Structure Segmentation with Optimal Parameter Setting. The main conclusions for the others contributions of the novel strategy for stenosis assessment are presented next:

- **Appearance and geometrical model for characterizing anatomical tracheal structures**

  We have developed an appearance and geometrical physical and its computational models for the elements involved in the stenosis assessment framework. These elements are tracheal rings and luminal area which we have described using appearance and geometric features that describe those anatomical structures in a way that discriminates them from other challenging structures.

- **Segmentation of main anatomical structures**

  We have developed a strategy for the segmentation of main anatomical structures needed for stenosis assessment. In the one hand, we have proposed a three step strategy for reference ring segmentation where our method selects the most complete and external one from all the characterized rings. In the other hand, we have presented a recursive step strategy for obstructed lumen segmentation where we are able to distinguish luminal region from the rest of the videobronchoscopic scene using a 2D feature space. Finally we have presented a validation protocol for evaluating each of the strategy steps in order to ensure that the final calculation of the degree of stenosis will not depend on the previous steps, so errors are not propagated. For this reason we have proposed a novel statistical framework based on Analysis of Variance (anova) [39] for selecting the optimal parameters which ensure within inter-observer variability performance in the segmentation of reference ring and obstructed lumen.

- **Validation protocol and public databases**
  The main difficulty when comparing performance of the methods is the lack of, up to our knowledge, public annotated databasse of tracheal rings, lumen regions or stenosis assessment. This constitutes a major flaw for the development of generic algorithms able to achieve accurate results in a wide range of images. We have introduced of different fully annotated databases to validate the performance of the developed methods. Databases which contain a balanced presence of all the possible appearances that we may find in future procedures, including extremely abnormal examples.

Experimental results show that the performance of our method is comparable to clinicians for structure characterization, structure segmentation and stenosis degree assessment. Aside our method has a low computational cost that could have a real-time implementation to be used in the operating room. Therefore, the output of our method has potential to be useful for helping the doctor as part of a computer-aided diagnosis (CAD) system, as having an accurate calculation of the degree of obstruction may lead to a reduction in the number of repetitions of tracheal implant procedures. Our methodology could also help the doctor to choose the appropriate stent or prosthesis diameter size. A main advantage is that it does not requires information from CT scans and, thus, it reduces patients radiation.

7.2 Future research lines

Regarding future work, our methodology could be adapted to cope with other respiratory system stenosis such as laringe or bronchi. An important remark is that, for bronchi, images have a conic view that distort image proportions along a given branch. Such distortion should be corrected for a practical use of a stenosis index computed from single images and would eventually require a sort of 3D reconstruction from videos. The methods proposed in this PhD thesis pave the way for several studies:

- **Navegation System** One of the algorithms developed in the context of this thesis is related with lumen centre detection. Using this information it is somewhat easy to imagine a navigation system using lumen centre detection to identify bifurcations. In this case we can see a clear connection between VB planning systems and our method for on-line navigation, as once the bifurcation is detected we could use planning information to indicate the clinician the path he/she has to take. Lumen centre detection could also be used to estimate how deep we are inside a given bronchi level, as if we only see a single luminal area we can expect to be near to the beginning of the level and as more luminal areas are identified -lumen centres- we would be close to the end of the level.
- **Volume recovering using tracheal rings or lumen** As we are able to segment accurately both tracheal rings and luminal region in separate frames, it would also be possible to extend these segmentations to a whole video aiming to obtain a 3D volume representing trachea shape. In this case we would use the correlation between consecutive frames to obtain a first representation.
- **3D reconstruction using optical flow** Another objective of this thesis was to define accurate anatomical landmarks that could be used to identify interesting tracheal structures. If the detection of these landmarks is robust we could use two different frames from the same scene to obtain a rough 3D reconstruction using the correspondence between features combined with some of state-of-the-art methods such as

Structure from Motion, Optical Flow or SLAM. Once we have a 3D reconstruction of the scene we would be able to obtain absolute, accurate and objective measurements which, as has been shown throughout the thesis, is one of the main drawbacks of current videobronchoscopy interventions. For instance and regarding tracheal implant we could provide with a complete characterization of the prosthesis needed, giving both diameter and length measures.

List of Publications

Submited Journals

- Carles Sánchez, Jorge Bernal, F. Javier Sánchez Debora Gil. (2014) On-line Systematic Quantification of Tracheal Stenosis from Videobronchoscopy. Medical Image Analysis (*submitted*)

International Conferences

- Carles Sánchez, F. Javier Sánchez, Antoni Rosell, Debora Gil. (2012). An illumination model of the trachea appearance in videobronchoscopy images. In Image Analysis and Recognition (Vol. 7325, pp. 313320). LNCS. Springer Berlin Heidelberg.
- Carles Sánchez, Debora Gil, Antoni Rosell, Albert Andaluz, F. Javier Sánchez. (2013). Segmentation of Tracheal Rings in Videobronchoscopy combining Geometry and Appearance. In Sebastiano Battiato and José Braz (Ed.), Proceedings of VISAPP 2013 (Vol. 1, pp. 153161). LNCS. Portugal: SciTePress.
- Carles Sánchez, Jorge Bernal, Debora Gil, F. Javier Sánchez. (2014). On-line lumen centre detection in gastrointestinal and respiratory endoscopy. In Klaus Miguel Angel and Drechsler Stefan and González Ballester Raj and Wesarg Cristina and Shekhar M. G. and O. L. Marius and Linguraru (Ed.), Clinical Image-Based Procedures. Translational Research in Medical Imaging (Vol. 8361, pp. 3138). LNCS. Springer International Publishing (*Best Student Paper Award*).
- Jorge Bernal, Debora Gil, Carles Sánchez, F. Javier Sánchez. (2014). Discarding Non Informative Regions for Efficient Colonoscopy Image Analysis. In 1st MICCAI Workshop on Computer-Assisted and Robotic Endoscopy (*Best Paper Award*).

Bibliography

[1] `http://www.olympus-global.com/en/corc/history/story/endo/origin/`.

[2] `http://alexea.org/main/`.

[3] `http://www.bronchoscopy.org`.

[4] K. V. Asari. A fast and accurate segmentation technique for the extraction of gastrointestinallumen from endoscopic images. *Medical engineering & physics*, 22(2):89 96, 2000.

[5] J. Bernal, D. Gil, C. Sanchez, and F. J. Sanchez. Discarding non informative regions for efficient colonoscopy image analysis. In *1st MICCAI Workshop on Computer-Assisted and Robotic Endoscopy*, 2014.

[6] J. Bernal, J. Sánchez, and F. Vilarino. Impact of image preprocessing methods on polyp localization in colonoscopy frames. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 7350 7354. IEEE, 2013.

[7] K. A. Carden, P. M. Boiselle, D. A. Waltz, and A. Ernst. Tracheomalacia and tracheobronchomalacia in children and adultsan in-depth review. *CHEST Journal*, 127(3):984 1005, 2005.

[8] J. Cohen. *Statistical powe analysis for the behavioral sciences.* Lawrence Erlbaum Associates, 1988.

[9] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):603 619, 2002.

[10] G. De Haan and E. B. Bellers. Deinterlacing-an overview. *Proceedings of the IEEE*, 86(9):1839 1857, 1998.

[11] W. Dorffel, I. Fietze, D. Hentschel, J. Liebetruth, Y. Ruckert, P. Rogalla, K. Wernecke, G. Baumann, and C. Witt. A new bronchoscopic method to measure airway size. *ERJ*, 14(4):783 788, 1999.

[12] C. W. Dunnett. A multiple comparison procedure for comparing several treatments with acontrol. *Journal of American Statistical Association*, 50:1096 1121, 1980.

[13] M. Everingham, S. M. Ali Eslami, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman. Assessing the significance of performance differences on the pascal voc challenges via bootstrapping. Technical report, 2013.

[14] L. Forkert, H. Watanabe, K. Sutherland, S. Vincent, and J. Fisher. Quantitative videobronchoscopy: a new technique to assess airway caliber. *American journal of respiratory and critical care medicine*, 154(6):1794, 1996.

[15] W. Freeman and Adelson. The design and use of steerable filters. *PAMI*, 13(9):891 906, 1991.

[16] D. Gabor. Theory of communication. *J. Inst Elect. Eng.*, 93:429 457, 1946.

[17] G. Gallo and A. Torrisi. Lumen detection in endoscopic images: a boosting classification approach. *International Journal On Advances in Intelligent Systems*, 5(1-2):127 134, 2012.

[18] D. Gil. *Geometric Differential Operators for Shape Modelling.* PhD thesis, Universitat Autónoma de Barcelona, 2004.

[19] D. Gil, A. Hernàndez-Sabaté, M. Burnat, S. Jansen, and J. Martínez-Villalta. Structure-preserving smoothing of biomedical images. In *Computer Analysis of Images and Patterns*, pages 427 434. Springer, 2009.

[20] D. Gil and P. Radeva. Extending anisotropic operators to recover smooth shapes. *Comp. Vis. Imag. Unders.*, 99:110 125, 2005.

[21] H. C. Grillo. Surgery of the trachea. *Archives of Surgery*, 112(12):1508, 1977.

[22] K. L. Gwet. *Handbook of Inter-Rater Reliability (2nd Edition).* 2010.

[23] A. Hayashi, S. Takanashi, T. Tsushima, J. Denpoya, K. Okumura, and K. Hirota. New method for quantitative assessment of airway calibre using a stereovision fibreoptic bronchoscope. *British journal of Anaesthesia*, 108(3):512 516, 2012.

[24] T. Heimann, B. van Ginneken, M. Styner, Y. Arzhaeva, and V. A. others. Comparison and evaluation of methods for liver segmentation from ct datasets. *IEEE Trans. Med. Imag.*, 28(8):1251 1265, 2009.

[25] E. Hein and M. Rutter. New perspectives in pediatric airway reconstruction. *International anesthesiology clinics*, 44(1):51, 2006.

[26] A. Hernàndez-Sabaté. *Exploring Arterial Dynamics and Structures in IntraVascular UltraSound Sequences.* PhD thesis, Universitat Autònoma de Barcelona, 2009.

[27] B. Jahne. *Spatio-temporal image processing@ARTICLEMiko05, author = K. Mikolajczyk and T. Tuytelaars and C. Schmid and et. al., title = A Comparison of Affine Region Detectors , journal = Int. J. Comp. Vis., year = 2005, volume = 65, pages = 43-72, owner = csanchez, timestamp = 2014.10.01* . Lecture Notes in Comp. Science, 1993.

[28] N. Jowett, R. Weersink, K. Zhang, P. Campisi, and V. Forte. Airway luminal diameter and shape measurement by means of an intraluminalfiberopticprobe: a bench model. *Archives Otolaryngology- Head and Neck Surgery*, 134(6):637, 2008.

[29] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321 331, 1988.

[30] B. KW. Validation of medical image analysis techniques. In *SPIE*, pages 567 606, 1999.

[31] J. Landis and G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33:159 174, 1977.

[32] K. S. Lee, M. R. Sun, A. Ernst, D. Feller-Kopman, A. Majid, and P. M. Boiselle. Comparison of dynamic expiratory ct with bronchoscopy for diagnosing airway malaciaa pilot evaluation. *CHEST Journal*, 131(3):758 764, 2007.

[33] A. Lopez, F. Lumbreras, J. Serrat, and J. Villanueva. Evaluation of methods for ridge and valley detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(4):327 335, 1999.

[34] M. Manno. Pediatric respiratory emergencies: upper airway obstruction and infections. 2010.

[35] I. Masters, M. Eastburn, and W. et al. A new method for objective identification and measurement of airway lumeninpaediatric flexible videobronchoscopy. *Thorax*, 60(8):652, 2005.

[36] P. McFawn, L. Forkert, and J. Fisher. A new method to perform quantitative measurement of bronchoscopic images. *ERJ*, 18(5):817 826, 2001.

[37] E. L. Melgoza, L. Serenó, A. Rosell, and J. Ciurana. An integrated parameterized tool for designing a customized tracheal stent. *Computer-Aided Design*, 44(12):1173 1181, 2012.

[38] K. Mikolajczyk, T. Tuytelaars, C. Schmid, and et. al. A comparison of affine region detectors. *Int. J. Comp. Vis.*, 65:43 72, 2005.

[39] R. G. Miller Jr. *Beyond ANOVA: basics of applied statistics.* CRC Press, 1997.

[40] H. P. Moreton and J. Smith. Hdcc: a software-based compression algorithm for video conferencing. In *(SPIE)*, volume 2187, pages 190 195, 1994.

[41] S. Murgu and H. Colt. Subjective assessment using still bronchoscopic images misclassifies airwaynarrowing in laryngotracheal stenosis. *Interactive cardiovascular and thoracic surgery*, 16(5):655 660, 2013.

[42] S. Murgu and H. G. Colt. Morphometric bronchoscopy in adults with central airway obstruction: caseillustrations and review of the literature. *The Laryngoscope*, 119(7):1318 1324, 2009.

[43] C. Myer 3rd, D. O connor, R. Cotton, et al. Proposed grading system for subglottic stenosis based on endotracheal tubesizes. *The Annals of otology, rhinology, and laryngology*, 103(4 Pt 1):319, 1994.

[44] A. Neubeck and L. Van Gool. Efficient non-maximum suppression. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 3, pages 850 855. IEEE, 2006.

[45] K. Neumann, J. Winterer, M. Kimmig, D. Burger, A. Einert, K.-H. Allmann, M. Hauer, and M. Langer. Real-time interactive virtual endoscopy of the tracheo bronchial system: influence of ct imaging protocols and observer ability. *European journal of radiology*, 33(1):50 54, 2000.

[46] C. J. Newth, M. J. Lipton, R. G. Gould, and M. Stretton. Varying tracheal cross-sectional area during respiration in infants and children with suspected upper airway obstruction by computed cinetomography scanning. *Pediatric pulmonology*, 9(4):224 232, 1990.

[47] S. Norwood, V. Vallina, and K. e. a. Short. Incidence of tracheal stenosis and other late complications after percutaneoustracheostomy. *Annals of surgery*, 232(2):233, 2000.

[48] S. Phee, W. Ng, I. Chen, F. Seow-Choen, and B. Davies. Automation of colonoscopy. ii. visual control aspects. *Engineering in Medicine and Biology Magazine, IEEE*, 17(3):81 88, 1998.

[49] B. T. Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311 317, 1975.

[50] R. Polverosi, M. Vigo, S. Baron, and G. Rossi. [evaluation of tracheobronchial lesions with spiral ct: comparison between virtual endoscopy and bronchoscopy]. *La Radiologia medica*, 102(5-6):313 319, 2000.

[51] U. B. Prakash. *Bronchoscopy.* Raven Press, 1994.

[52] H. Rozycki, M. Van Houten, and G. Elliott. Quantitative assessment of intrathoracic airway collapse in infants andchildrenwith tracheobronchomalacia. *Pediatric pulmonology*, 21(4):241 245, 1996.

[53] C. Sánchez, J. Bernal, D. Gil, and F. Sánchez. On-line lumen centre detection in gastrointestinal and respiratory endoscopy. In *CLIP-MICCAI*, volume 8361 of *LNCS*, pages 31 38, 2014.

[54] C. Sánchez, F. Sánchez, A. Rosell, and D. Gil. An illumination model of the trachea appearance in videobronchoscopy images. In *ICIAR*, 2012.

[55] C. Sánchez, J. Sánchez, and D. Gil. Detecting tracheal rings in videobronchoscopy images. In *CVCRD2011*, pages 132 135, 2011.

[56] P. Sharma, A. Bansal, S. Mathur, S. Wani, R. Cherian, D. McGregor, A. Higbee, S. Hall, and A. Weston. The utility of a novel narrow band imaging endoscopy system in patients with barrett s esophagus. *Gastrointestinal endoscopy*, 64(2):167 175, 2006.

[57] M. Spivak. *Calculus.* 3rd edition, 2006.

[58] L. E. Sucar and D. F. Gillies. Knowledge-based assistant for colonscopy. In *Proceedings of the 3rd international conference on Industrial and engineeringapplications of artificial intelligence and expert systems*, volume 2, pages 665 672, 1990.

[59] W. Sureshbabu and O. Mawlawi. Pet/ct imaging artifacts. *Journal of nuclear medicine technology*, 33(3):156 161, 2005.

[60] M. Tubiana, S. Nagataki, L. Feinendegen, et al. Computed tomography and radiation exposure. *N Engl J Med*, 358(850):852 853, 2008.

[61] J. W. Tukey. Comparing individual means in the analysis of variance. *Biometrics*, 5:99 114, 1949.

[62] J. Vergnon, F. Costes, M. Bayon, and A. Emonot. Efficacy of tracheal and bronchial stent placement on respiratory functionaltests. *Chest*, 107(3):741 746, 1995.

[63] N. H. Weste and K. Eshraghian. Principles of cmos vlsi design: a systems perspective. *NASA STI/Recon Technical Report A*, 85:47028, 1985.

[64] S. Wieand, M. H. Gail, B. R. James, and K. L. James. A family of nonparametric statistics for comparing diagnostic markers withpaired or unpaired data. *Biometrika*, 76(3):585 592, 1989.

[65] X. Zabulis, A. A. Argyros, and D. P. Tsakiris. Lumen detection for capsule endoscopy. In *IROS. IEEE/RSJ InternationalConference on*, pages 3921 3926. IEEE, 2008.

This work was supported in part by a research grant from Universitat Autonòma de Barcelona 471-01-2/2010, by the Spanish Government through the founded projects COLON-QA $(TIN2009-10435)$ and FISIOLOGICA $(TIN2012-33116)$.