



**Universitat
Autònoma
de Barcelona**

MASTER IN COMPUTER VISION AND ARTIFICIAL INTELLIGENCE

REPORT OF THE MASTER PROJECT

OPTION: COMPUTER VISION

Tracheal ring detection in bronchoscopy

Author: **Carles Sánchez**

Advisor: **Javier Sánchez,**

Debora Gil

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 Agnes Borrás and Albert Andaluz 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.

ABSTRACT

Endoscopy is the process in which a camera is introduced inside a human. Given that endoscopy provides realistic images (in contrast to other modalities) and allows non-invasive minimal intervention procedures (which can aid in diagnosis and surgical interventions), its use has spread during last decades. In this project we will focus on bronchoscopic procedures, during which the camera is introduced through the trachea in order to have a diagnostic of the patient. The diagnostic interventions are focused on: degree of stenosis (reduction in tracheal area), prosthesis or early diagnosis of tumors. In the first case, assessment of the luminal area and the calculation of the diameters of the tracheal rings are required. A main limitation is that all the process is done by hand, which means that the doctor takes all the measurements and decisions just by looking at the screen. As far as we know there is no computational framework for helping the doctors in the diagnosis. This project will consist of analysing bronchoscopic videos in order to extract useful information for the diagnostic of the degree of stenosis. In particular we will focus on segmentation of the tracheal rings. As a result of this project several strategies (for detecting tracheal rings) had been implemented in order to compare their performance.

Keywords: *Bronchoscopy, tracheal ring, segmentation*

Contents

1	Introduction	1
1.1	The evolution of endoscopic techniques	1
1.2	The bronchoscopy	3
1.3	Computer vision in bronchoscopic diagnosis	7
1.4	Motivations and goal	8
1.5	Our contributions	8
2	Main features of bronchoscopy images	10
2.1	Anatomy description in images	10
2.2	Typology from the image processing point of view	14
3	Detecting tracheal rings : workflow	16
3.1	Image pre-processing	17
3.1.1	Frame suppression	17
3.1.2	Filtering	17
3.2	Representation	18
3.2.1	RGB space	19
3.2.2	CIELAB space	19
3.2.3	HSI space	20
3.2.4	Principal component analysis	20
3.3	Ring detector	20
3.3.1	Steerable Gaussian Filters (SGF)	21
3.3.2	Level Set Geometry (LSG)	22
3.4	Binarization	22
3.4.1	Non-maximal suppression (NMS)	22
3.4.2	Hysteresis	23
3.5	Strategies definition	24

4	Experimental settings and results	28
4.1	Validation protocol	28
4.1.1	Usage scenario conditions for medical procedure	28
4.1.2	Groundtruth for validation	28
4.1.3	Metric for evaluation	29
4.2	Results	29
5	Conclusions and future work	36
5.1	Conclusions	36
5.2	Future work	37

Chapter 1

Introduction

1.1 The evolution of endoscopic techniques

Since the ancient civilizations of Greece and Rome, humans have developed many techniques to look inside the human body with medical purposes. Endoscopy is the process in which a tube is introduced inside a patient to explore the organs and the cavities. The first prototype was found in the ruins of Pompeii and consisted in a three-bladed vaginal speculum that was very similar to the modern one [1]. In order to obtain a better visualization of the internal anatomy, light was incorporated to the endoscopic practices.

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 (Figure 1.1a). In 1853, Antoine Jean Desormeaux was the first one who introduced an improved "Lichtleiter" of Bozzini to the body of a patient. This instrument consisted of a system of mirrors and a lens, and where the improved light source was a lamp flame (Figure 1.1b); the endoscope burned a mixture of alcohol and turpentine. Liechester's was mainly used for urologic cases. Its main complication 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]. A fundamental thing is that all of this first endoscopes were rigid.

In 1932 Dr. Rudolph Schindler created the first flexible gastroscope which consisted of a flexible device with several lenses located along the tube and a mini-light attached (Figure 1.1c) [1]. This gastroscope allowed to bend the tube (with 75 cm of length and 11 mm of diameter) some degrees and this lead to an improvement on the rigid device's since it allowed the physician to look at more directions from the same place. From this moment, 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.

In 1950 appeared the first endoscope registering what the physician was looking at (Figure 1.2a). The device was equipped with a thinner tube and a photographic lens located at the end point of the flexible tube. The images were captured in a monochromatic film where they could be rewound by pulling a cable. This technology allowed registry procedures but unfortunately image/video could not be edited. Also, the diameter of the tube had to

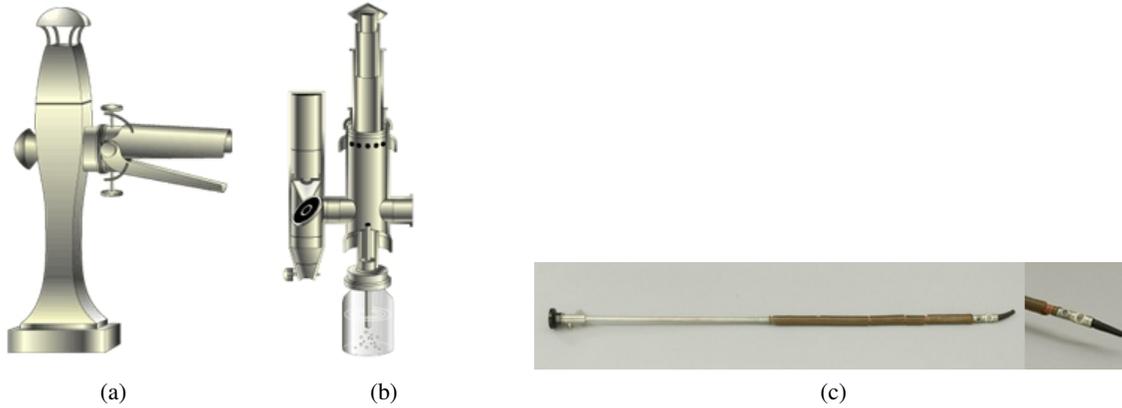


Figure 1.1: Optical endoscopes : Rigid models (a,b), flexible model and detail of the light bulb attached at its tip (right) (c).

be large enough for storing the camera and film cassette (see Figure 1.2b and 1.2c) [1]. As a consequence the endoscope could not reach deep inner parts of the human body.

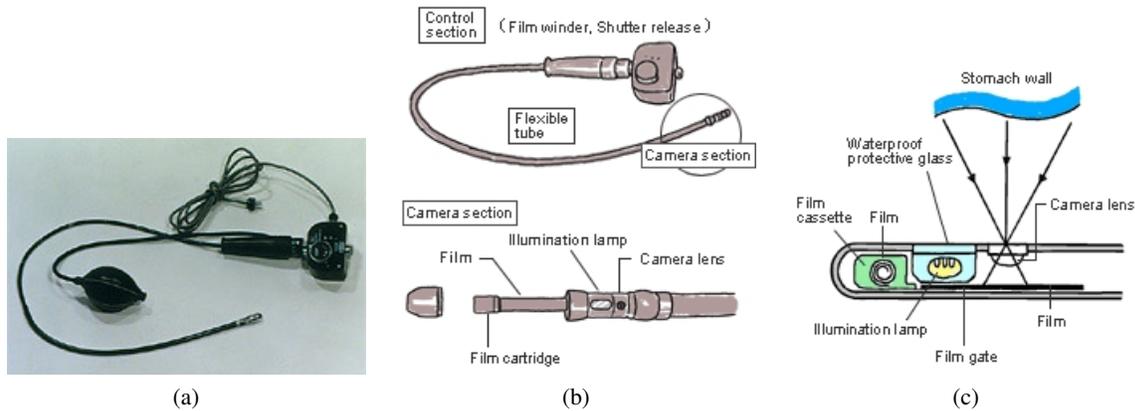


Figure 1.2: Flexible endoscope: Real gastrocamera (a), Scheme of gastrocamera (b), Gastrocamera tip structure scheme (c).

The introduction of fiberglass in 1964 allowed endoscopes to transmit the light along the tube even if it was bended. With this new material and also new ocular cameras (thanks to the new technologies) the camera was now allowed to be placed outside the tube. In this manner, the end point section of the endoscope could be smaller than the gastrocamera and, thus, achieve more depth positions inside the human body. In 1975 the era of gastrocameras finished, when they were replaced by the fiberscope (Figure 1.3a) [1]. At this time the endoscope begins to be used in several fields of diagnostic medicine such as esophagus, duodenum, large intestine, bronchus and gallbladder. The main limitation of fiberscope was that the quality of analogical (each pixel represented by one fiber) images was not good enough to properly identify anatomical structures and, thus, an efficient use require high quality

training of the physician.

In 1990 the development of first High Resolution CCD is developed, making possible to build smaller color video cameras with resolution which lead to the development of modern videoscopes. A videoscope (electronic scope) is an endoscope with a built-in video camera using CCD. It allows smaller caterers and deeply analysis into the patient. Videoscope converts images into electric signal for display on a TV monitor (Figure 1.3b). While only one skilled doctor at a time was able to observe the interior condition of an organ with the conventional instrument, the new device allowed several doctors and nurses to examine the condition simultaneously by watching the screen. Also, the video can be seen as many times as they want. In this manner, the safety of the patient was improved because of a higher accuracy in its diagnosis and thanks to the good quality of the images. In addition, the image-processing feature can make focal sharpness adjustments through electric signal control in order to highlight the lesion by boosting specific color signals for easier viewing. New features such as these have expanded the horizons of endoscopic possibilities still further [1].

Finally in 2002 the new HDTV Technology was introduced in the videoendoscopy process (Figure 1.3c).The safety and precision were more improved using this system because of the color ,resolution or zoom of the images (which provide more realistic images).

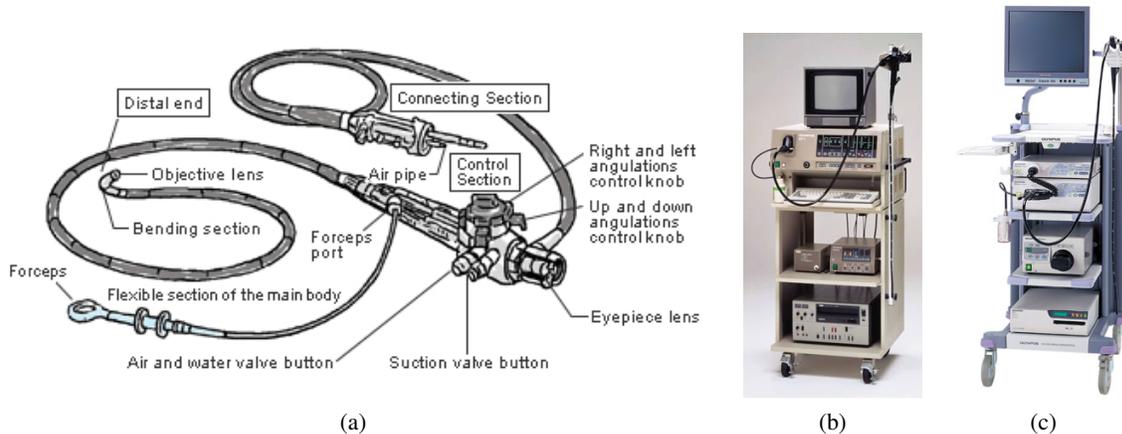


Figure 1.3: Videoendoscopic systems: fiberscope (a), A CCD endoscopy system (b), A HDTV endoscopy system (c).

1.2 The bronchoscopy

Bronchoscopy consists in visualizing the inside of pulmonary airways for diagnostic and therapeutic purposes. Bronchoscopy was first performed in 1897 by the German physician Dr. Gustav Killian, which performed a rigid bronchoscopy. Rigid bronchoscopy was the only available technique for bronchoscopy until 1979, when fiber optic technology was applied successfully to bronchoscopy and make it possible to use flexible bronchoscopic devices (Figure 1.4). In modern bronchoscopy as the instrument progresses through the patient, the camera captures

the videoframes and the doctor can observe the interior of the bronchial tree in a TV monitor. Rigid bronchoscopy is being performed frequently although still far less often than flexible bronchoscopy.

Rigid bronchoscopy and flexible bronchoscopy:

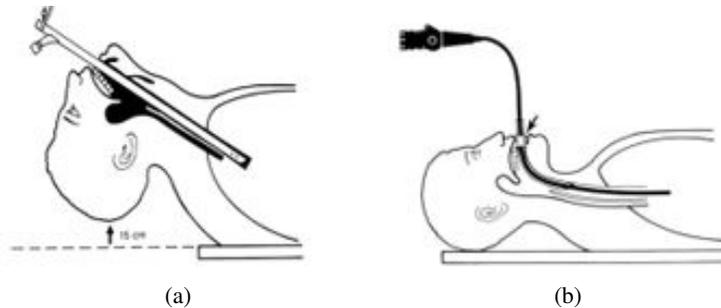


Figure 1.4: Types of: Rigid bronchoscopy position (a), Flexible bronchoscopy position (b).

- **Rigid bronchoscopy.** This method uses a rigid and straight hollowed (Figure 1.4a) metal pipe which allows the insertion of additional equipment into the lower airways such as ultrasound probes or laser lights, and placement of airway stents. The large lumen of the rigid bronchoscope restricts its use to the larger pathways (trachea) for performing interventional procedures such as tracheal prostheses implants and treatment of tumors. Rigid bronchoscopes permit the insertion of additional equipment into the lower airways such as ultrasound probes or laser lights, and placement of airway stents (Figure 1.5a). This kind of exploration is routinely done under general anaesthesia in the surgical operating room. This method needs of much experience to handle and therefore it is not widely used (Figure 1.5) [3] [4].

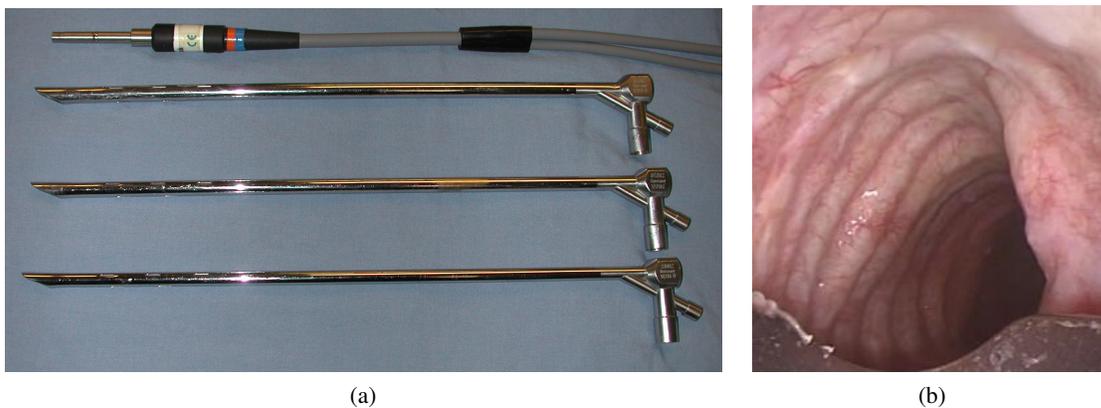


Figure 1.5: Rigid procedure: Rigid bronchoscopy instrument (a), Rigid bronchocopy image (b).

- **Flexible bronchoscopy.** Flexible bronchoscopy uses an instrument which is longer and thinner than the rigid bronchoscope (Figure 1.6a). It can contains a: fiberoptic system that transmits an image from the tip

of the instrument to an eyepiece or video camera at the opposite end; or small camera at the tip of the instrument. Using Bowden cables connected to a lever at the hand piece, the tip of the instrument can be oriented, allowing the practitioner to navigate the instrument into individual lobe or segment bronchi (Figure 1.6b). Most flexible bronchoscopes also include a channel for suctioning or instrumentation, but these are significantly smaller than those in a rigid bronchoscope. Flexible bronchoscopy causes less discomfort for the patient than rigid bronchoscopy and the procedure can be performed easily and safely under moderate sedation. Nowadays, flexible bronchoscopy is more currently used than rigid bronchoscopy. (Figure 1.6) [3] [4].

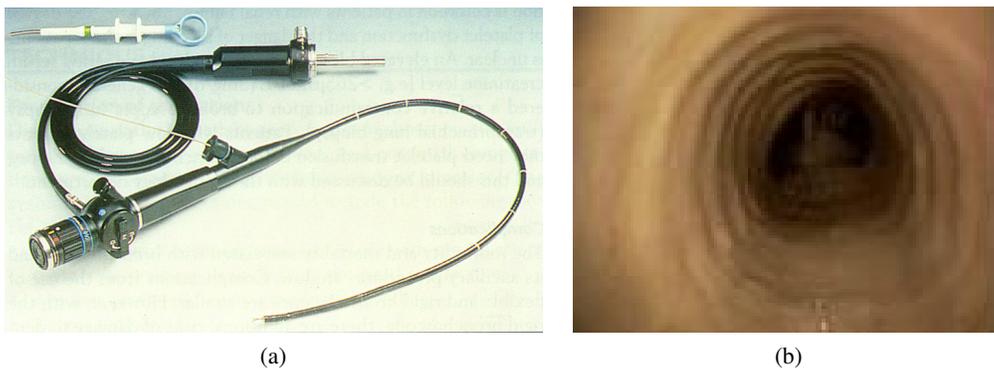


Figure 1.6: Flexible procedure: Flexible bronchocopy instrument (a) Flexible bronchocopy image (b).

Even though both modalities of bronchoscopy are currently used, the flexible one is more common than the rigid one. During the exploration process, the physician operates the bronchoscope by referring to his anatomical knowledge. The physician inserts the bronchoscope into the patient's nose or mouth and explores four zones (see Figure 1.7) in order to provide a complete diagnosis of the patient [4]:

Zones of bronchoscopic exploration:

- **From nose/mouth to larynx (upper-left image in Figure 1.7).** The scope advances from the nose to the larynx with the help of local anaesthesia.
- **From larynx to trachea (lower-left image in Figure 1.7).** It enters in the subglottic area and once the vocal cords are passed the scope is slightly flexed downwards (if from the head).
- **From trachea to carina (lower-right image in Figure 1.7).** The tracheal path is one of the most explored areas. The trachea is almost a straight pipe composed of several equidistant tracheal rings. The area enclosed by tracheal rings is the air path and is called luminal area. The end of the trachea, where it is divided into two main bronchus, is called carina.
- **From carina to bronchial system (the right/left main bronchus, lower-right image in Figure 1.7).** Main bronchus is the second more explored area. From the carina the left/right main bronchus 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 depending on the size of the camera.

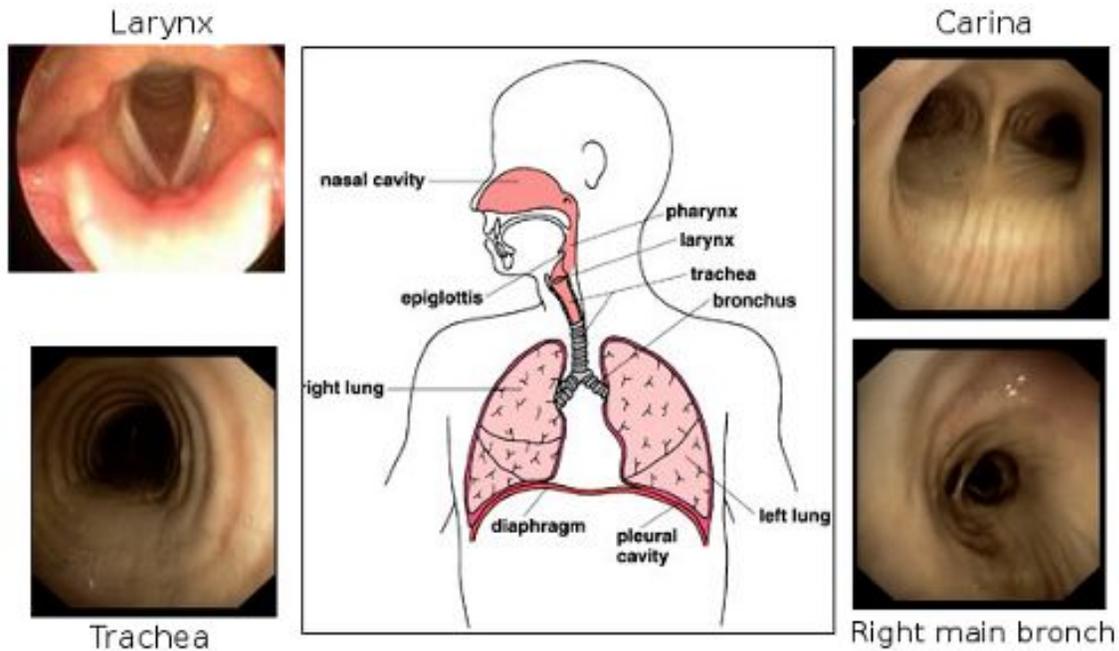


Figure 1.7: Different regions of the bronchoscopy process.

Diagnostic interventions are routinary processes, so standard hospitals such as Bellvitge achieves an average of three interventions per week. Bronchoscopic exploration cover several kind of clinical topics:

Types of diagnosis according to bronchoscopic exploration:

- **Assessment of the degree of stenosis (reduction in bronchial area).** The narrowing of the bronchial and tracheal path way is known as stenosis. The assessment of the luminal area and tracheal rings diameters is required to estimate the degree of obstruction or stenosis. Measures are visually assessed from inspection of bronchoscopy videos. They are concretely taken from closest ring to the camera that appears entirely in the image. Variability between observers and quantitative measures are a clear handicap.
- **Prosthesis implantation.** In order to choose the most appropriate prosthesis the physician has to estimate the diameter and length of the obstructed tracheal segment. As in the case of stenosis assessment, a main limitation is the visual assessment of images to estimate the diameter of the obstructed path. However, in this case an extra source of variability that the doctor has to infer a 3D measurements of the prosthesis length by simple inspection of the 2D video frames (2D images in perspective projection). Determining the true 3D length strongly depends on the experience and the anatomical knowledge of the doctor. Thus, a high inter-observer variability often implies a remodelation of the prosthesis.

- **Early diagnosis of tumors.** Identification of tumor regions in standard bronchoscopic videos recorded using white light is a difficult task that requires extraction of a tissue sample for biopsy analysis. This fact has motivated the development of Narrow Band Imaging (NBI). 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 1.8 we can observe an example of NBI.

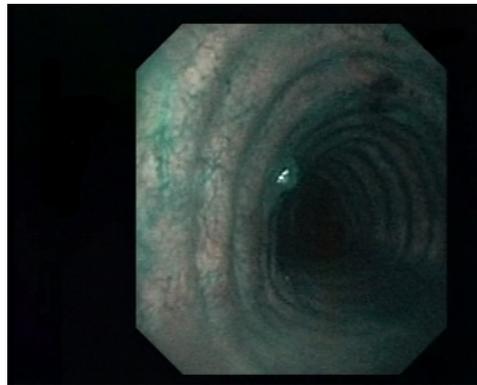


Figure 1.8: Trachea NBI.

1.3 Computer vision in bronchoscopic diagnosis

The limitation following virtual assessment, have encouraged developing computer tools for helping in diagnosis. However, since modern video bronchoscopy is a relatively new technology (it appeared in the middle 90's), there is not much computational framework yet.

- **Works related to the measurement of degree of stenosis and prosthesis implantation.** The aim of these works is to help the physician in the calculation of 3D measurements. There are many works in virtual bronchoscopy. Although these methods can be useful to navigate and perform 3D models, virtual bronchoscopy is not feasible in clinical practice. Some works are based on bronchoscopy tracking using virtual bronchoscopy or 3D models [5] [6] [7] [8] [9].

Virtual bronchoscopy or endoscopy can be done by using CT-images (Computed tomography) [10] [11] and put it into correspondence with the real-time images. 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 [12]. But, with this technique there is a big problem, the irradiation. It is not possible to have a tracing diagnostic of a patient because excessive exposure to x-rays.

- **Work on tumor detection.** The introduction of new technologies such as autofluorescence bronchoscopy, narrow band imaging, endoscopic ultrasound, endobronchial ultrasound, electromagnetic navigation, optical coherence tomography, and confocal fluorescent laser microscopy can help the physician [13] in tumor detection. In this manner highly irrigated tissue, which is a good indicator of tumors, can be easily identified. The main limitation in NBI and all of the others new techniques is that they are still an experimental approach. As far as we know nowadays there is no correspondence between the appearance of this kind of light and the nature of the tissue. For that reason this techniques are not used in clinical practice. Exist one work from 2002 that does a segmentation in endoscopic image processing [14] for the lumen detection but nothing in tumor detection.

1.4 Motivations and goal

The evaluation of stenosis degree is a common practice in bronchoscopic exploration. To do so, the computation of the percentage of airway trachea reduction needs to be done in bronchoscopical images. Consequently, the detection of the tracheal rings plays a meaning roll in the process. Nevertheless, in the literature, there is a lack of works that assess this problem using computer vision tools. This way, in this project we will focus on measuring the degree of stenosis in the tracheal part. To do so, we will need to evaluate the percentage of airway trachea reduction according to bronchoscopical images. Since this is a pioneer work in the bronchoscopic scope we have made a study to design the best image processing tool to detect tracheal rings. This preliminary study is the first step to going a some computer vision tool that help the physician in having an objective measurement of stenosis degree. With that, it can be deleted the variability of human factor taken measurements.

1.5 Our contributions

We have divided the problem of tracheal ring detection in a set of phases and in each one of these phases we provide a set of contributions. That we can summarize in three main points:

- First we present an study of the content of the bronchoscopic images. In this study we have observed the principal problems on the images according to the image processing techniques. Moreover we have developed a categorization of videoframes. It means to recognize the different parts of a bronchoscopy and choose the frames for our image database set. This categorization help us to create the usage scenario conditions for medical procedure.
- In a second phase we have studied the best strategy to detect the tracheal rings by answering some questions:
 - Does color/texture convey useful information?

- Which is best way to convert to gray level minimizing lost of anatomic information?
 - Which is best image filtering that preserves our structures of interest?
 - Which is best tracheal ring detector?
- Finally, in order to answer the above questions, the output of several techniques has to be assessed. We have manually segmented a representative sampling of our database twice, in order to account for intra-observation variability.

The contents of the master are organized as follows. In chapter 2 we report the frame categorization (from an anatomical and image processing point of view) and identify the main artefacts of the images. Chapter 3 explains the different image processing tools considered in this study. Chapter 4 reports the validation protocol and the quantification results achieved by each strategy. Finally, chapter 5 exposes conclusions, answers the questions and outlines future lines.

Chapter 2

Main features of bronchoscopy images

2.1 Anatomy description in images

The first step of the project is to analyse how bronchoscopy images look like.

The images of the project belong to the trachea. The trachea is a tube that connects the pharynx or larynx to the lungs and lies in front of the esophago. The trachea has an inner diameter of about 21 to 27 millimetres (0.83 to 1.1 in) and a length of about 10 to 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 [15].

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 of protecting and maintaining the airway equally spand 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) [15]. The histologic cut of Figure 2.1 shows the main elements of a trecheal section: C-shaped rings, interconnectivity muscle and lumen. In Figure 2.3 we can observe the same elements in a real frame.

The videos used in this project were recorded for surgical purposes. In those videos trachea can be seen as tube in a conical projection (Figure 2.2a). If the camera is oriented in the axis of the trachea (carina is on the center) the conical projection goes to a set of concentric circles (Figure 2.2b). Moreover, a vertical cut of the trachea make us able to see how the C-shaped cartilaginous rings looks like (Figure 2.2c).

First of all we have selected trachea frames for each sequence and we have organized them according to its anatomical content. Every sequence presents lots of difficulties/problems that are a clear handicap. There are three types of artefacts in broncoscopic videos:

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

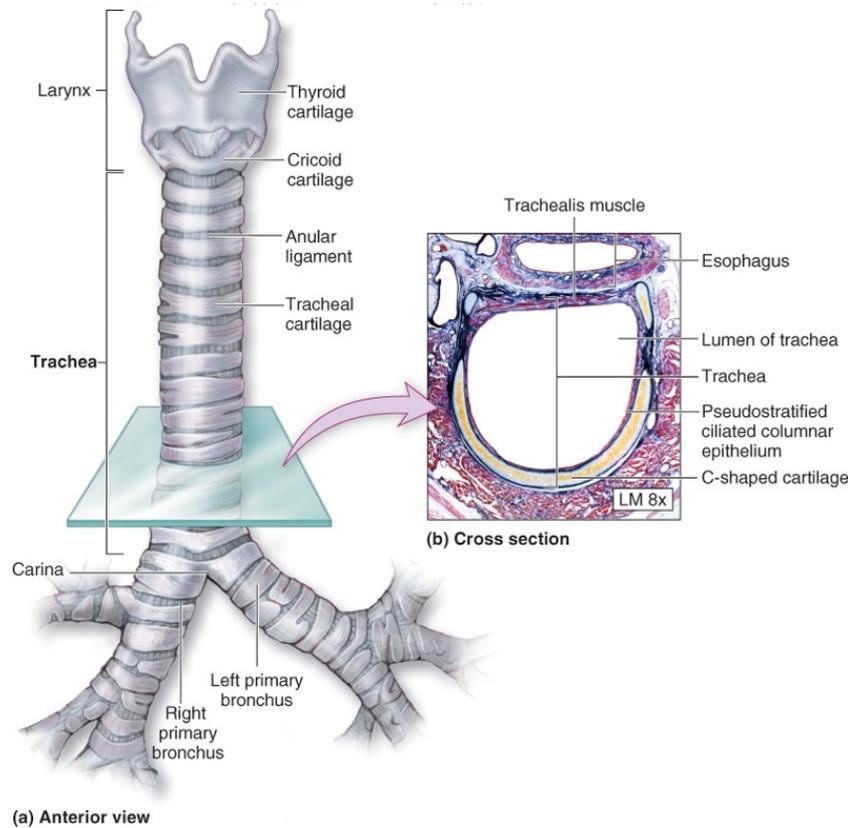


Figure 2.1: Trachea anatomical elements and histologia cut.¹

- **Worsening of images quality.**

- **Non-uniform illumination** (see **Figure 2.4**). The illumination is not the same in all the images because the light of the camera is not always in the same position. It implies, sometimes, that some parts are more illuminated while others present more shadows.
- **Blurring** (see **Figure 2.4**). It appears when the camera moves too fast and also when the camera objective is fogs because the breath of the patient.
- **Specular highlights**. Because of the structure of the trachea or alien elements such as: mocus and blood vessels , bubbles or part of the rigid bronchoscopy.
- **Interlacing**. Consists of two sub-fields taken in sequence, each sequentially scanned at odd and even lines of the image sensor [16]. If camera does not moves you cannot see the effect, but in our case camera is always moving so in most of the frames interlacing appears (borders of the structures in the image are indented) (see **Figure 2.5**).

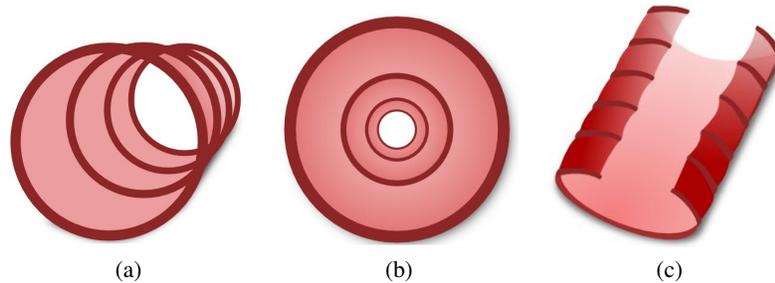


Figure 2.2: Trachea scheme: Conical projection (a), Concentric rings, carina on the center (b), Vertical cut (c).

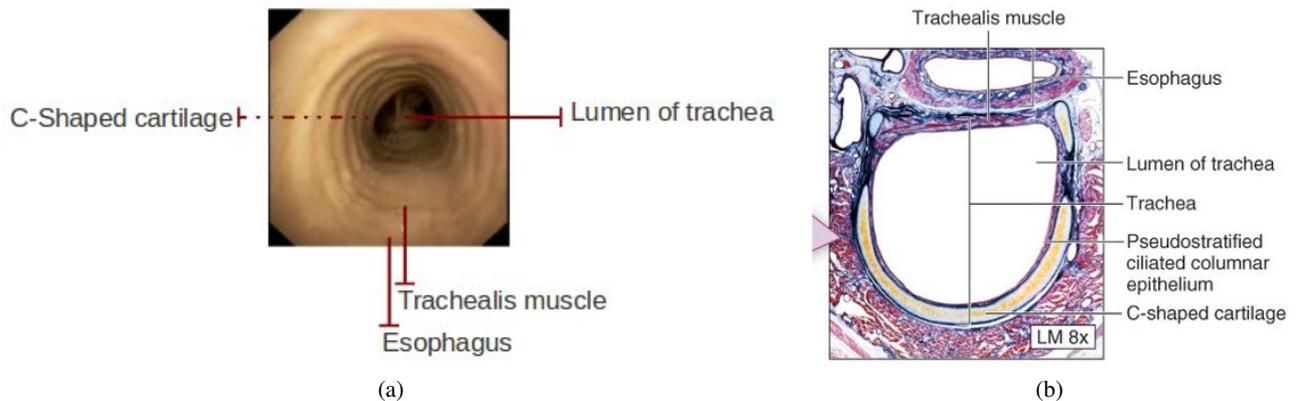


Figure 2.3: Trachea anatomical elements in a real frame (a), histologia cut scheme (b).

- **Alien elements not belonging to the tracheal rings structure.**

- **Lumen.** The dark part of the trachea frames. Part of the trachea tube that is farthest of the camera.
- **Mucus, blood and bubbles.** Substances or liquids that can appear in human trachea.
- **Carina** (see **Figure 2.6**). This structure can be seen when the bronchoscope is getting to the end of the trachea. Is the beginning of the bronchial system.
- **Rigid bronchoscope** (see **Figure 2.6**). Is a rigid material that the physician use when needs to guide the flexible bronchoscope or do some surgery. Explained before in the introduction chapter (Figure 1.5).

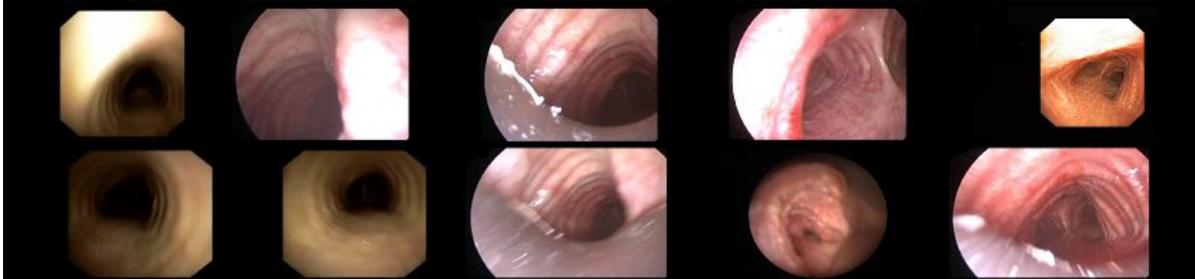


Figure 2.4: Image quality artefacts: first row : ilumination problems, second row: blurred problems.



Figure 2.5: Interlacing problem example.

- **3D-2D projection geometric distortion.** 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 structures (tracheal rings) are well defined. Out of center deviations introduce two main artefacts in 2D images:
 - **Camera focused on some trachea wall (see Figure 2.7).** We cannot see the tracheal rings or we only see part of them.
 - **Rings appear as collapsed due to the projection (see Figure 2.7).** 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.

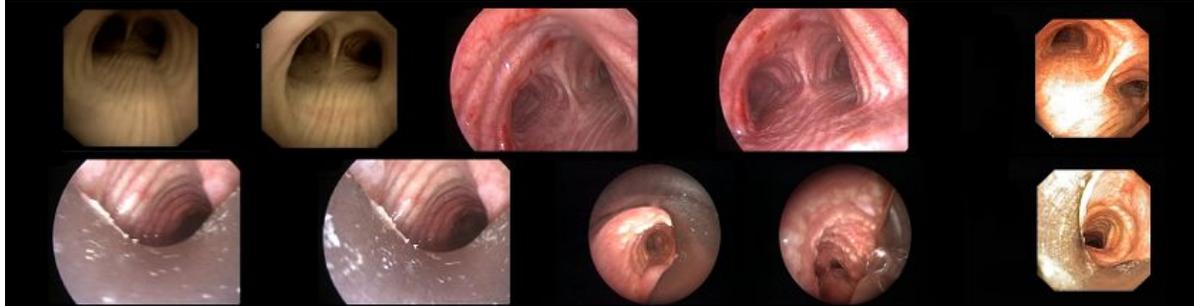


Figure 2.6: Non anatomical artefacts: first row : carina artefact, second row: rigid bronchoscope artefact.



Figure 2.7: Geometric artefacts: first row : trachea wall, second row: collapsed tracheal rings.

2.2 Typology from the image processing point of view

We have already seen the typology of the anatomic elements that appears in our frames of bronchoscopy procedures. Once we have selected the representative dataset, it is time to analyse what is the structure of the elements that we want to recognize, in terms of image processing. As we said before, trachea is a tube composed by cartilage rings equidistantly separate. If we take this tube and we put a camera in front, what we will see will be a set of concentric tracheal rings, big rings at the front and smaller rings as we look further (Figure 2.2). Moreover, if we do a radial cut section of the trachea, what a shuffler (interlaced) composition of valleys (between rings) and ridges (every ring and between valleys). Figure 2.8 shows a frame with concentric rings and the profile of the radial cut (green line on the image) on the left plot. Dots indicate the position of ridges and crosses the position of valleys. We note that in this example the illumination is not uniform, that is why in the dark side of the frame, the valley-ridge shape is different than the illuminated side.

As we said our goal is to detect the rings that appears in each frame. In order to do that, in computer vision words, our task will be to recognize structures represented by the concentric composition of [...valley-ridge-valley-ridge...]. Therefore we will focus ridge/valley detectors with all the preprocessing needed to clean the image as much as possible (as explained in the methods' chapter). Figure 2.9 how is the pattern that represent the tracheal rings.

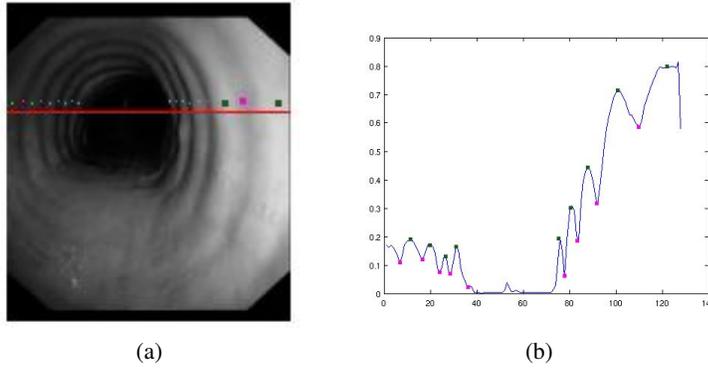


Figure 2.8: Ridge-valley structure of tracheal rings: real image with a line showing the radial cut (a) and plot of the intensity of the radial cut (b).

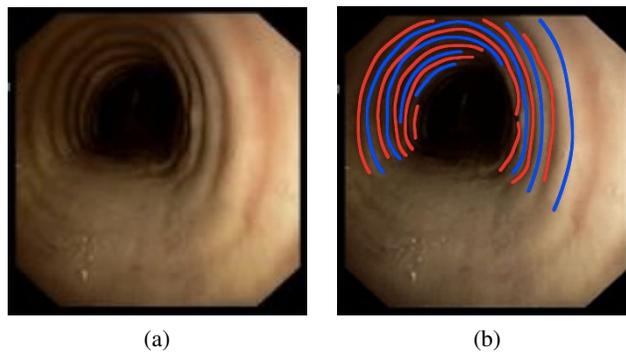


Figure 2.9: Pattern [...valley-ridge-valley-ridge...] : real image (a), real image with the pattern , blue for ridges and red for valley (b).

Chapter 3

Detecting tracheal rings : workflow

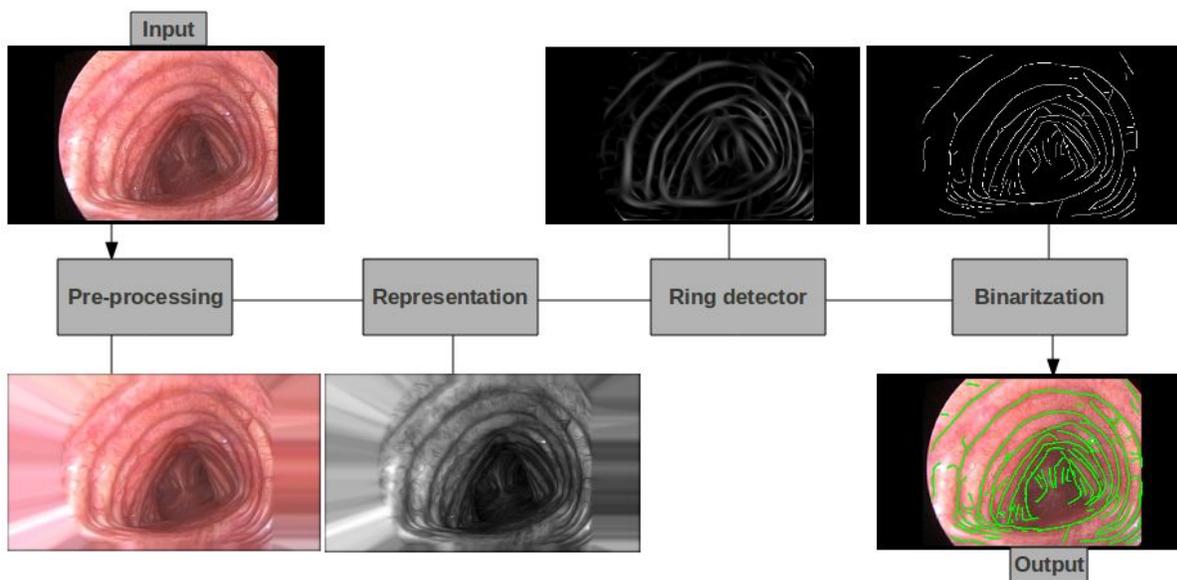


Figure 3.1: Workflow for segmenting tracheal rings

The strategy we propose to detect tracheal rings covers four main stages:

1. **Pre-processing.** In this phase we are going to prepare the image to be processed by cleaning the small structures and suppressing the frame.
2. **Image representation.** In this step we are going to represent the image in color or grayscale by using different color spaces. We will compare how the results change by using different color representations.
3. **Ring (ridge/valley) detector.** We are going to process the image by using different ridge/valley detector methods, because it is the way that we already defined to represent tracheal ring.

4. **Binarization.** We are going to binarize the response of the ridge/valley methods. We will have a detection binary image to compare with our groundtruth.

The scheme given in Figure 3.1 sketches workflow of the method that we are going to use to do our tracheal ring detector. For each step there are several different options that we are going to explain in the following sections.

3.1 Image pre-processing

3.1.1 Frame suppression

The given images have a black frame that probably will interfere in the detection of valleys and ridges. In order to minimize its impact, the values of the image border are extended to fill the whole black framework (Figure 3.2). we have considered a nearest neighbourhood extension. For each framework pixel, we consider its nearest pixel on the image border and we use its color value. Any valley-ridge detector would give large response at the borders because of the black frame.

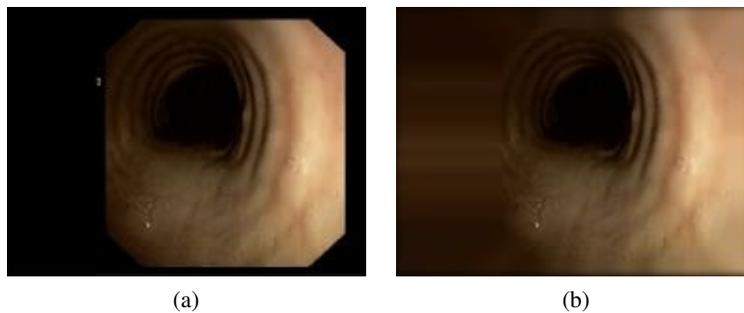


Figure 3.2: Mark suppression example : Given image with mark (a), Pre-processed real image (b).

3.1.2 Filtering

Filtering is perhaps the most fundamental operation of image processing in computer vision. In the broadest sense of the term "filtering", the value of the filtered image at a given location is a function of the values of the input image in a small neighbourhood of the same location [17]. In this case filtering step is going to clean the image trying to preserve the structures that we are interested on. Two methods have been considered:

Bilateral filtering

Bilateral filtering was proposed by Tomasi and Manduchi in 1998 as a non-iterative method for edge-preserving smoothing in a color or gray images [18]. The function of this method is based on a combination of domain filtering and range filtering. It means that the filtered image is taken into account two things: the pixels values in

a neighbourhood domain and also the range of values of these pixels in the input image. Bilateral filtering can be very useful because it keeps strong edges while smoothing the image. Good filtering behaviour is achieved at boundaries, thanks to the domain component of the filter, and crisp edges are preserved at the same time, thanks to the range component [18]. The performance of the bilateral filtering strongly depends on the size (σ) of the Gaussian used. Figure 3.3 shows the same image filtered using increasing σ 's.

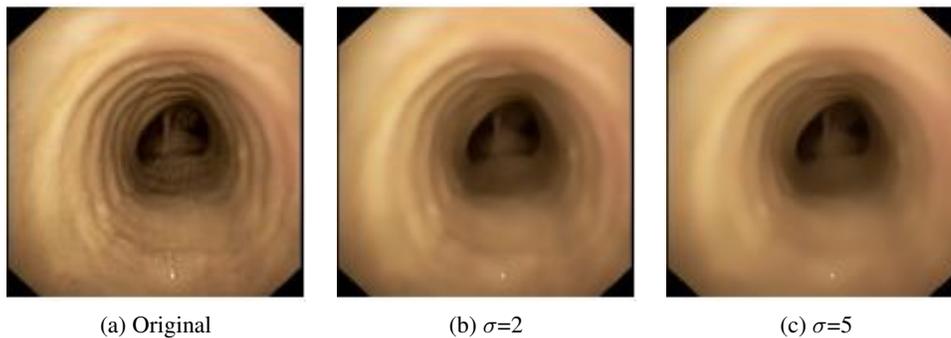


Figure 3.3: Performance bilateral filtering with different σ parameter.

Structure Preserving Diffusion (SPD)

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, while restoring contours consistent with anatomical structures.

As D.Gil explains in [19] Anisotropic diffusion operators are based on image appearance discontinuities and might fail at weak inter-tissue transitions. The method (SPD) that this paper presents is based on the propagation of the structural tensor. In this manner and according to the physical inspiration, SPD is going to give a non-uniform intensity image presenting homogeneous inter-tissue transitions along anatomical structures because homogeneous propagation in the tensor. But smoothing background texture because non-homogeneous propagation of the structural tensor [19].

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 3.4 it preserves our interested structures while it is smoothing around them. Given that in our case the valley-ridge profile is given by thin structures, SPD is the most appropriate filtering for preserving it.

3.2 Representation

Bronchoscopic images are in color and thus, the impact of color and gray level represent spaces must be explored. We can deal with different color spaces or we can use the grayscale representation. According to

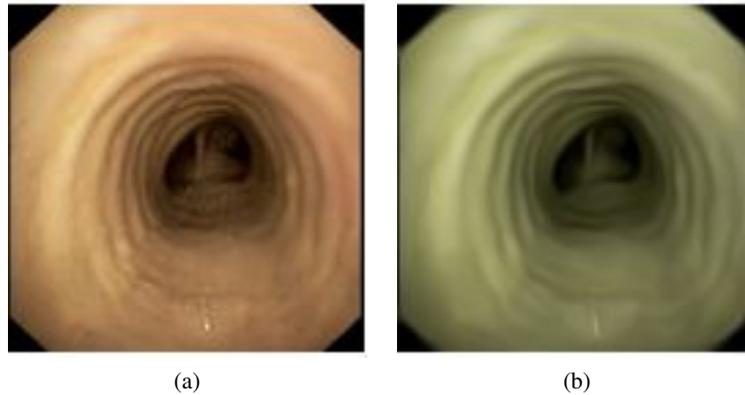


Figure 3.4: SPD filtering example : Real image (a), Filtered image (b).

color, we have selected to work with Red-Green-Blue (RGB). Moreover we have processed the information of the CIELAB and the HSI representations using a Principal Component Analysis (PCA). Finally we have used the grayscale information obtained from the RGB information. So at the end we will have one color representation (RGB) and four gray scale representation (RGB+PCA, HSI+PCA, CIELAB+PCA and intensity gray from RGB). Next we present a brief explanation of the color spaces we have used.

3.2.1 RGB space

The RGB model represents every color in terms of intensity of each of primary colors (red, green and blue). It is a model of color that is based on additive synthesis, that is, it represent any color by linear combination of the primary light colors. Given that the values of intensity for each primary colors are in the $[0..255]$ range, the representation of RGB model is a cub. An advantage of this color space is that, is the standard digital format. It is used in screens, projectors, scanners or cameras. In the other hand it has a clear disadvantage, RGB space is able to represent a limited range of colors in the spectrum that we can perceive [20].

3.2.2 CIELAB space

CIELAB is a chromatic and perceptual model normally used to describe all the colors at way human perceives them [21]. The transformation from RGB to CIELAB is non-linear. The three parameters of the model represent: the light color or illumination (L , $L=0$ black and $L^*=100$ white). The other two channels a and b use a opponent color representation. The position between red and green (a , green for negative values and red for positives) and the position between yellow and blue (b , blue for negative values and yellow for positives). A disadvantage is that most of the devices uses a RGB color representation, so it has to be converted to make it possible work with it.

3.2.3 HSI space

The HSI color representation model has three components: hue (H),saturation (S), intensity (I). The transformation from RGB to HSI is non-linear. The Hue component is an angle between $[0,360]$ degrees, that represents the color itself. With for instance, 0 degree for red, 120 for green 240 for blue, 60 degrees for yellow and 300 degrees for magenta. The Saturation component represents the amount of white color and its range is $[0,1]$. The Intensity is the amount of light that the color contains and its range is between $[0,1]$ where 0 means black and 1 means white [22]. An advantage of this color space is that the environment colors are differentiable between them through the hue component (this does not happend in the RGB representation). A disadvantage is that has to be converted when is acquired from a device.

3.2.4 Principal component analysis

Principal component analysis (PCA) is an orthogonal transformation to convert a set of different observations of maybe correlated variables into a set of principal components. The number of principal components is equal or lower than the original variables. Given this transformation, the variance of each component (of the principal components) is as high as possible [23]. In our case we are going to use PCA to have a grey image from a color image. Using this procedure we expect to have the most variable things of each channel in the first principal component (gray image). As Figure 3.5 shows, given that image our principal component should be the red line. So, to get our gray image we just have to project all the points to the new 1D (red line) space (Figure 3.5c).

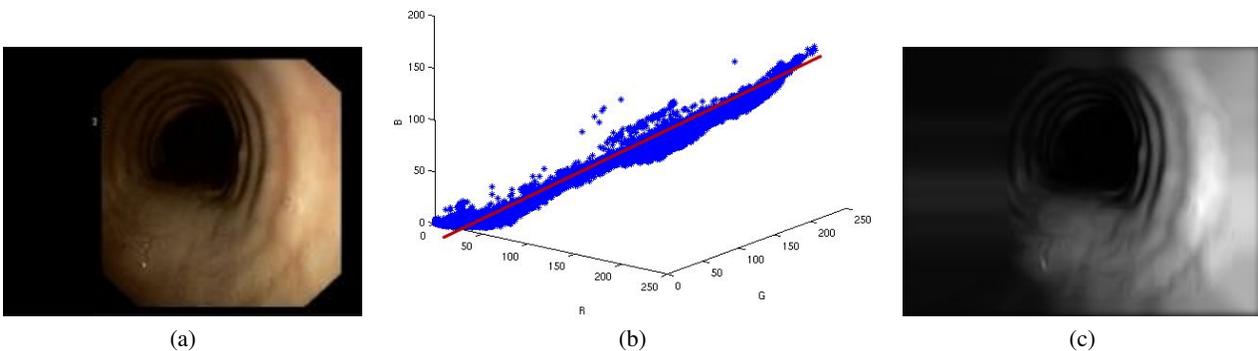


Figure 3.5: Given image (a) , RGB plot (b) , Projection to the first principal component of PCA (with all the pre-processing done) (c).

3.3 Ring detector

In this step we are going to detect tracheal rings. We are going to do that in such a way according on what we have defined as a tracheal ring. As we said in the second chapter tracheal ring is a composition of ridge and valley, so what we need are ridge-valley detector methods. We have chosen two different methods according to the

non-uniform illumination of the images. The methods chosen are invariant to the illumination. Those methods have also been applied in gray and color representation. Next, we explain some ideas of how these methods work. Figure 3.6 shows how the valley-ridge detectors work.

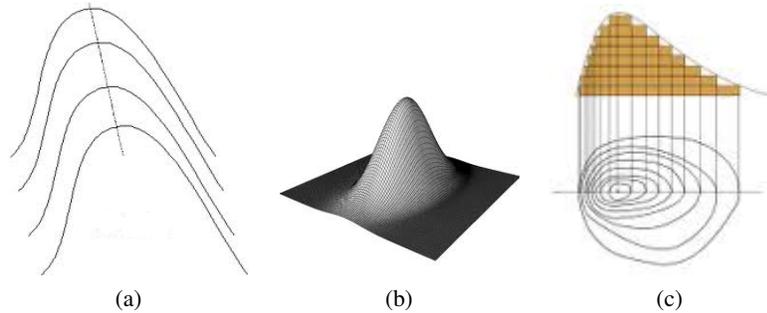


Figure 3.6: Ridge detection example in different detectors: Ridge (a), Steerable Gaussian Filters (b), Level Set Geometry (c).

3.3.1 Steerable Gaussian Filters (SGF)

Given the represented image SGF detector is going to identify valleys and ridges. This method is applied to the image by using the second derivative of an anisotropic Gaussian oriented. Number of orientations and range of scales are the parameters of the method. What the method does, is find the best intensity pattern that matches 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 (Figure 3.7). SGF gives the maximum value of matching between all the scales and orientations (Figure 3.6b shows a Gaussian that has to be matched in the ridge Figure 3.6a). At the end we get a two energy image with the ridges and valleys detected. The detected structures depend on the scale. Figure 3.8 shows an example detecting ridges and valleys.

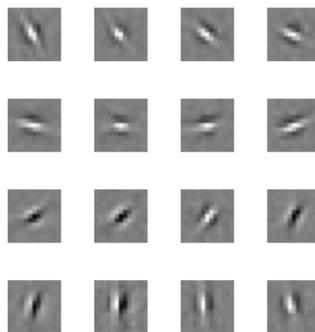


Figure 3.7: Different anisotropic Gaussian filters with different orientations.

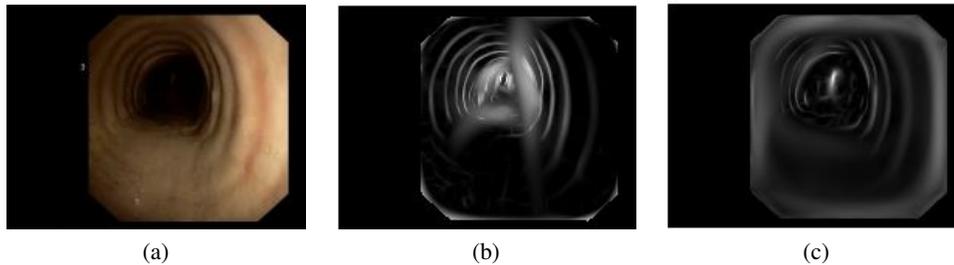


Figure 3.8: Example of SGT output: Original image (a), Valley detection (b), Ridge detection (c).

3.3.2 Level Set Geometry (LSG)

Given the represented image LSG detector will identify valleys and ridges based on the level set curves. This method is explained in [24]. As Figure 3.6c shows the method is based on finding the maximum value of level curves (for ridges) and the minimum value of level curves (for valleys). This method is very sensitive with small structures. The method returns an image with the valley and ridge detections. Valley detections with negative values and ridge detections with positive values, all of them in the range between $[-1,5 \dots 1,5]$. Figure 3.9 shows an example detecting ridges and valleys.

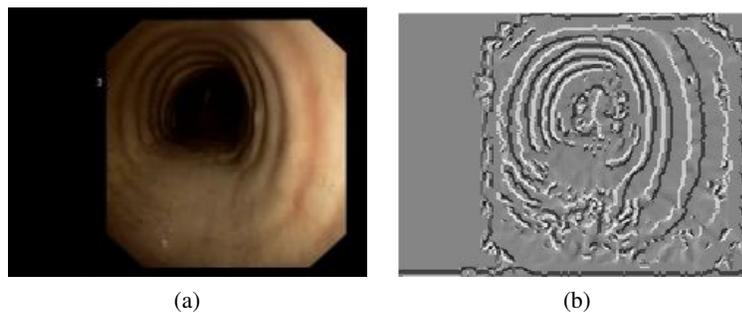


Figure 3.9: Example of LSG output: Original image (a), Valley-ridge detection (b).

3.4 Binarization

Given the energy image given by the valley-ridge detector, we extract tracheal rings by a two-step binarization: non-maximal suppression followed by hysteresis thresholding.

3.4.1 Non-maximal suppression (NMS)

Non-maximal suppression is used as an intermediate step in many computer vision algorithms. Given a gray image, NMS scans along it using the gradient direction and also some neighbouring. What NMS does is setting

to zero all the pixels that are not part of the local maxima (Formula 3.1).

$$I_o(x, y) = \begin{cases} I(x_n, y_n) \leq I(x, y) > I(x_m, y_m) & I(x, y) \\ \text{otherwise} & 0 \end{cases} \quad (3.1)$$

It can be done by comparing the intensity of a pixel ($I(x, y)$) with the neighbour pixels ($I(x_n, y_n)$, $I(x_m, y_m)$) along the direction of the gradient of the current pixel $I(x, y)$. If any of the neighbouring pixels has a value higher than the current pixel, the algorithm is going to eliminate (put it to zero) that pixel. Otherwise this pixel is going to be a candidate. At the end NMS suppresses all the pixels that are not candidates, therefore, all the pixels that are not a local maxima. The result is a thinned one-pixel mask of the dominant ridges-valleys, as shown in Figure 3.10.

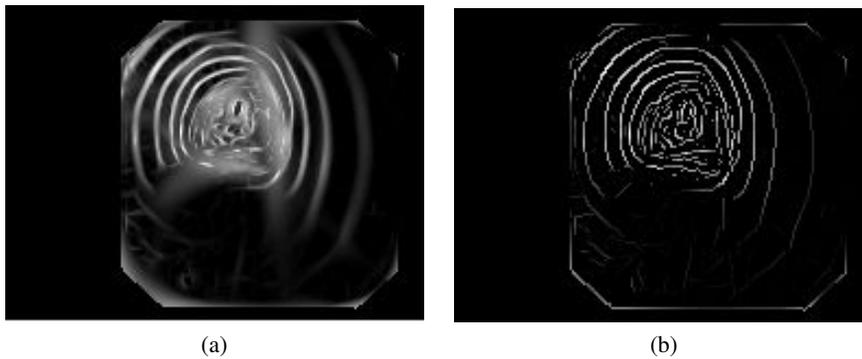


Figure 3.10: Energy image (Ring detector) (a) , NMS of energy image (b).

3.4.2 Hysteresis

Once we have thinned the energy image with our detected rings is the moment to binarize it. Hysteresis is a good method for binarizing that deals with a double threshold (α_1, α_2). Taken the NMS image $I(x, y)$ the hysteresis consists in applying the high threshold $I(x, y) > \alpha_1$ and using the low threshold to reconstruct the binary image. All the pixels $I(x, y) > \alpha_2$ that are neighbour to some pixel included in the first threshold will be in the final binarized image. The final binary image will also contain the pixels that are included in the second threshold whenever they are connected to any pixel that belongs to the high threshold. Figure 3.11 shows what is the benefit of using hysteresis binarization against a single threshold.

The results of the tracheal detection process rely on a good selection of the hysteresis thresholds. This way, we have introduced a process to automatize the selection of these values according to the information contain in the histogram of the energy image after NMS. Figure 3.12 shows histograms of different NMS images (from different sequences). The first row contains histograms computed from the detection of LSG ridge-valley detector and the second row is related to the responses of the SGF ridge-valley detector (in both cases the valley detection). We observe that the histogram information of the LSG case can be represented as three gaussians and the SGF

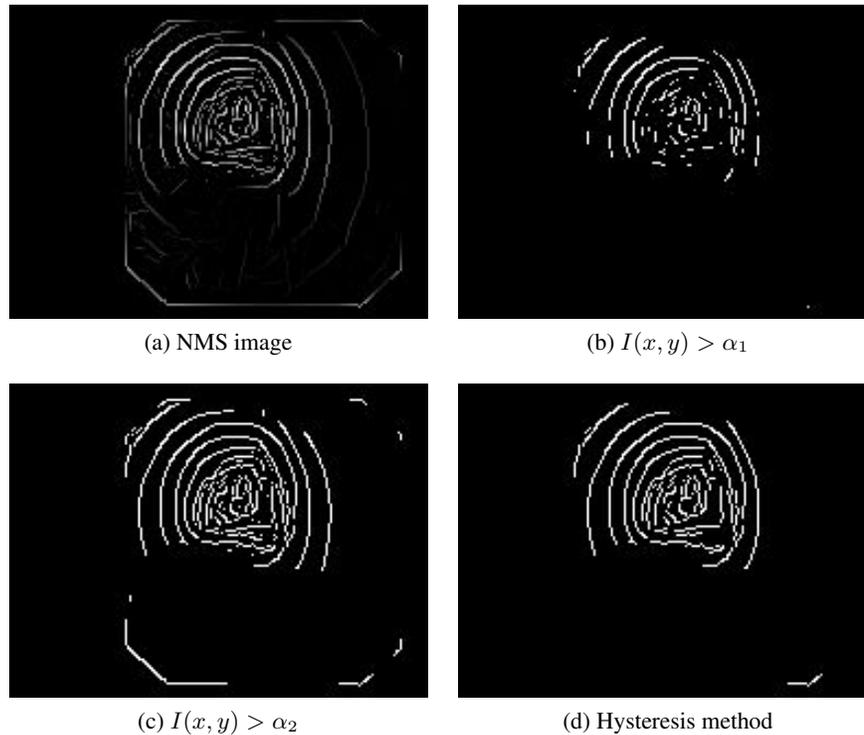


Figure 3.11: Hysteresis thresholding: NMS image (a), mask obtained by thresholding (b)(c) and hysteresis (d).

case with two. Next step is to know what represent every gaussian. As can be seen in Figure 3.13 high values in the histograms mean high detection of valley/ridge. We have made a previous approach by setting the thresholds according red lines in the Figure 3.13 the binarized image contains most of the tracheal ring detection. To automatize this process we have used a expectation maximization (EM) algorithm to fit a number of gaussians (gaussian mixture) in a distribution (our histograms). Once we have the μ and σ of all the gaussians we are looking for, we can just play with this values to find the best combination of values for the thresholds. For instance, the threshold in the first row of Figure 3.13 could be: μ of the second gaussian for the high threshold and $\mu - 2\sigma$ for the lower one. In experimental settings section we will show the obtained results by setting this thresholds.

3.5 Strategies definition

Once we have seen all the possible methods in every step, we are going to define which will be our strategies to evaluate. Figure 3.14 shows all the possible methods that we can use in each step. Notice that the filtering step can be done into the color representation or after applying PCA. As can be seen bilateral filtering has been discarded. We decided not to use this sort of filtering because it is a method based on edges but does not deal well with ridge and valley structures (in Figure 3.3 it can be seen how ridge and valley structures are disappearing).

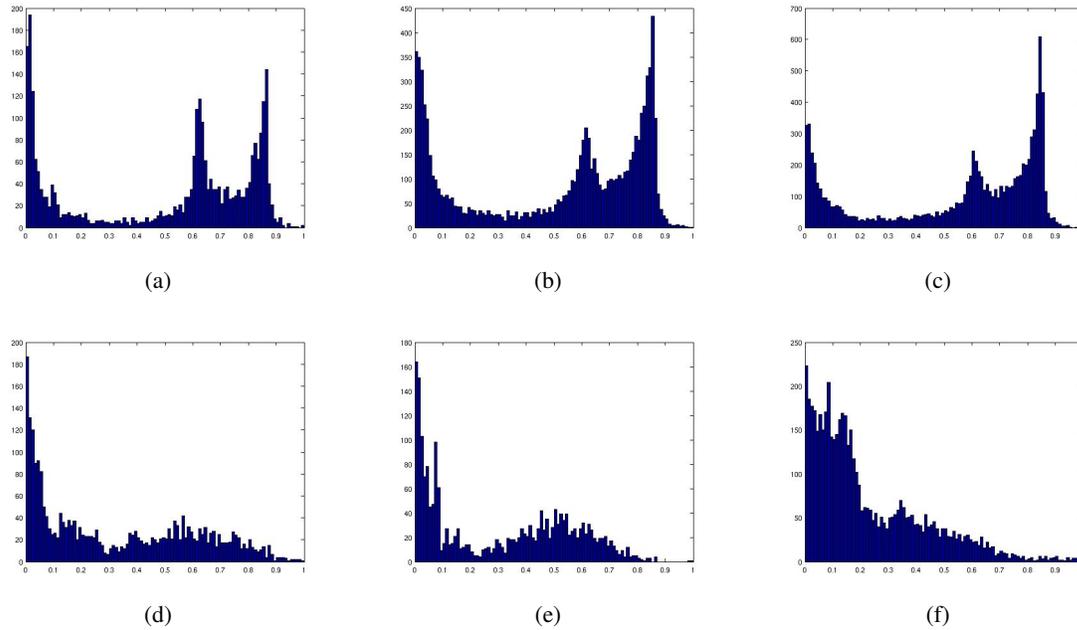


Figure 3.12: Automatic selection of thresholds. Histograms of NMS image of valley detectors. First row from LSG valley detector and second row from SGF valley detector.

Also it is quite aggressive with small structures like morphological filters. In order to preserve (thin) tracheal rings while smoothing background, we have used the SPD introduced in filtering section. HSI color representation, also has been discarded. This color space was discarded because does not have a euclidean structure, it means that using this color space we cannot use cartesian operators. PCA and our ridge-valley detectors use cartesian operators.

In order to have an idea of which are our different strategies, Table 3.1 shows what are the methods used in each one of them.

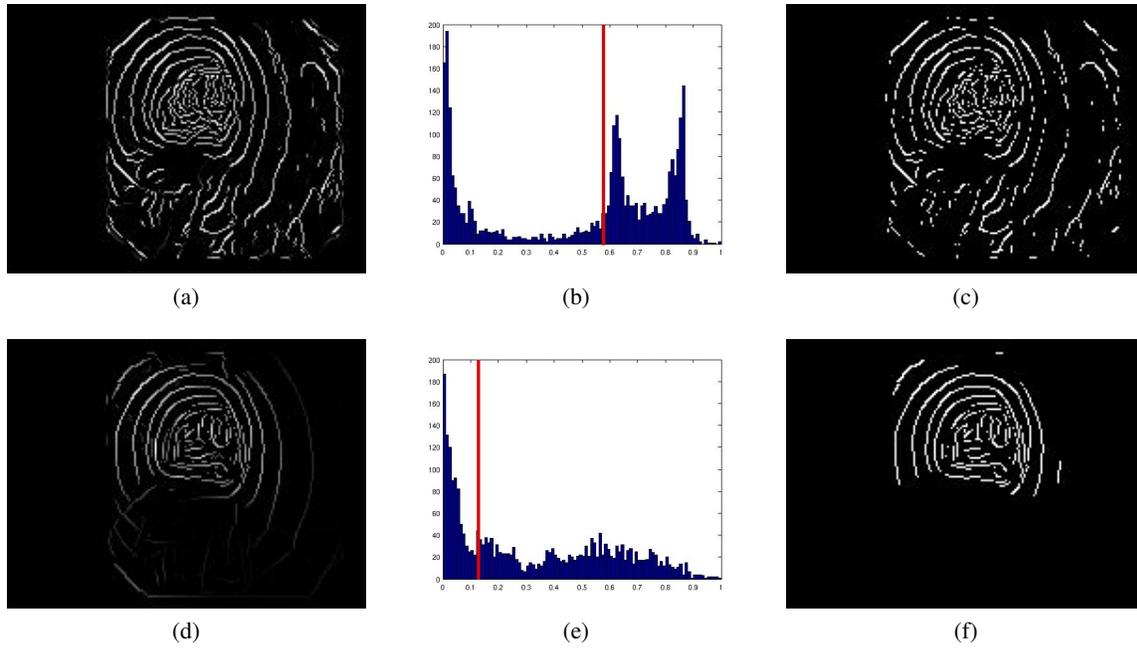


Figure 3.13: Histograms of NMS image of valley detectors. Last gaussian identification.

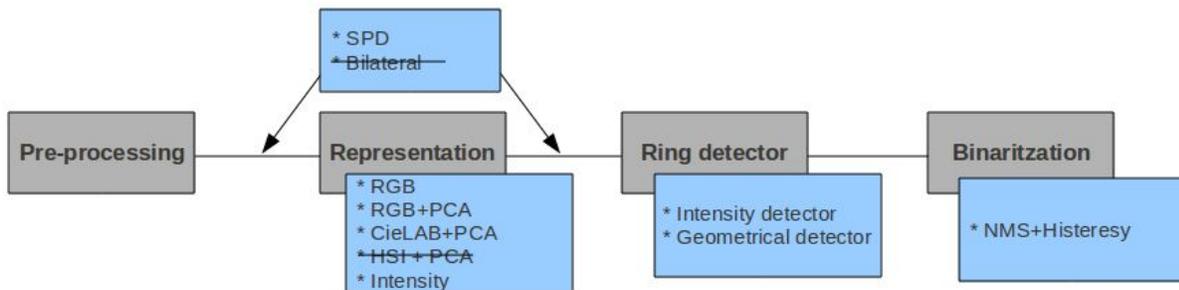


Figure 3.14: Different methods that we can use in every step.

Strategy	Color representation	Filtering	PCA	Filtering	Ring detector	Binarization
Strategy1	Intensity	SPD	No	No	LSG	Yes
Strategy2	RGB	SPD	No	No	LSG	Yes
Strategy3	RGB	SPD	Yes	No	LSG	Yes
Strategy4	CIELAB	SPD	Yes	No	LSG	Yes
Strategy5	Intensity	SPD	No	No	SGF	Yes
Strategy6	RGB	SPD	No	No	SGF	Yes
Strategy7	RGB	SPD	Yes	No	SGF	Yes
Strategy8	CIELAB	SPD	Yes	No	SGF	Yes
Strategy9	RGB	No	Yes	SPD	LSG	Yes
Strategy10	CIELAB	No	Yes	SPD	LSG	Yes
Strategy11	RGB	No	Yes	SPD	SGF	Yes
Strategy12	CIELAB	No	Yes	SPD	SGF	Yes

Table 3.1: Different strategies combining methods.

Chapter 4

Experimental settings and results

In this section where we are going to comment what is the data that we are going to evaluate and which methodology we are going to follow. In the final section of this chapter we present a table that summarizes the results.

4.1 Validation protocol

4.1.1 Usage scenario conditions for medical procedure

We told in the two first chapters the importance of having a centered viewpoint of the trachea. Only in this case we obtain a concentric view of the tracheal rings. Knowing that, we should put some constraints to the physician. There are two important conditions on how to get frames. We know that the bronchoscopy procedure takes very few time and during the exploration the camera is always moving. Nevertheless for a good tracheal ring detection the physician has to maintain a camera position few seconds in the direction of the tracheal axis looking to the carina. The doctor have to do this process in two different places, one focusing to a healthy tracheal ring and the second one focusing to a obstructed ring. Because this project will continue to the calculation of the degree of stenosis the physician should accomplish the conditions at least on the conflicting zones.

4.1.2 Groundtruth for validation

Four sequences of healthy bronchoscopy patients are given by the Bellvitge hospital. These sequences contain two rigid bronchoscopy and two flexible bronchoscopy. Images taken in rigid bronchoscopy have more quality (high resolution) than flexible bronchoscopy. As we explained before a bronchoscopy sequence goes from the larynx to the four/five level of bronchial system. We have only considered those frames where (usage scenario conditions) the camera is oriented in the axis of the airway of the trachea and viewing to the carina (centred). In this frames we can see the tracheal rings (our [...valley-ridge-valley-ridge...]) completely. We have selected all the frames that meet the above condition. We got more than 3000 frames from all of the sequences. In order

to validate our detection methods 60 representative frames have been selected and segmented manually to have the groundtruth. Figure 4.6 shows a representative subset of the dataset. We have segmented the dataset twice. Segmentations have been done using Osirix [25], a program for medical image visualization which includes several annotation and marking tools. In this program you can mark points that form segments (our tracheal rings) and obtain a cvs file with their location information. Parse code has been made to pass from the cvs file to a binary mask images for matlab. Our detection method results will be compared with this groundtruth masks.

4.1.3 Metric for evaluation

We have validated the different strategies of our method according to several metrics. We have compared the binary images that result from all the strategies with the binary image that conforms the groundtruth. Formula 4.1 shows the metric that computes the sensitivity (true positives TP) and Formula 4.2 shows the one related to precision (false positives FP).

$$Sensitivity = \frac{\#(Detect \cap GT_{\oplus})}{\#GT} \quad (4.1)$$

$$Precision = 1 - \left(\frac{\#Detect - (\#(Detect \cap GT_{\oplus}))}{\#Detect} \right) \quad (4.2)$$

Where *Detect* is the final binarized image of every strategy, GT_{\oplus} is the mask of groundtruth dilated in two pixels and *GT* is the mask without dilatation. The symbol # means the number of white pixels. We dilate the mask with a disk of 2x2 because the segmentations are in one pixel and the valley-ridge are thick, so the detections could be very close but not in the same place. An ideal segmentation should have 100% of both sensitivity and precision. That is, percentage of detected valleys that are in the GT mask for the sensitivity and the percentage of good detections instead of noisy for the precision.

4.2 Results

In this section we present the ring detection obtained with different strategies. The results are done by validating valley detection. We assume that the process for ridge detection should be similar. Figure 4.1 shows the possible strategies combining the different methods in the boxes of the different steps. The SPD (filtering step) can be done before or after applying PCA. At the end we got 12 different strategies to validate that we already seen in the last chapter. The results of those strategies have been introduced into a table which contain:

- **Strategy.** (Rows) Name of each strategy is referring to the representation and ring detector used. The definition of those names are done in Table 3.1.
- **Experts.** (Columns) Since we have segmented twice our dataset, we have a result for each segmentation that we call here expert (exp).
- **Metric values.** (Columns) For each expert and strategy, we report the sensitivity(%) and precision(%) ranges for the 60 frames in the groundtruth.

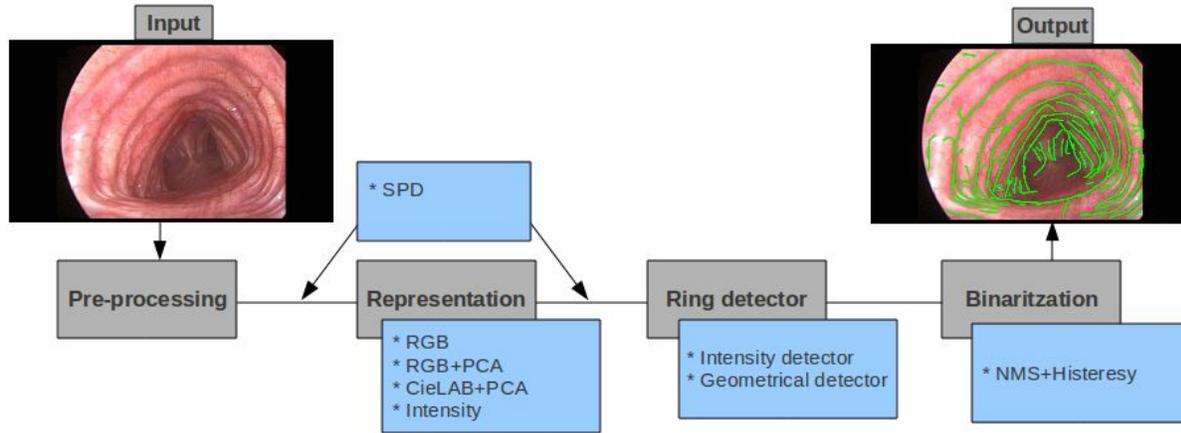


Figure 4.1: Different strategies.

Table 4.1 show the results for each strategy defined in the last chapter of the report. To have an idea of which are de best strategies we have tuned all the parameters. Finally Table 4.2 shows the variability between experts or observers.

Strategy	Exp1		Exp2	
	Sensitivity(%)	Precision(%)	Sensitivity(%)	Precision(%)
Strategy1	78.94 ± 6.63	27.09 ± 11.74	82.41 ± 5.46	24.19 ± 9.55
Strategy2	83.67 ± 5.55	20.61 ± 8.71	86.67 ± 4.40	18.23 ± 6.82
Strategy3	78.29 ± 6.83	25.97 ± 11.02	81.71 ± 5.32	23.14 ± 8.75
Strategy4	78.59 ± 6.13	27.14 ± 10.93	81.82 ± 5.15	24.15 ± 8.62
Strategy5	74.70 ± 8.80	36.02 ± 13.76	79.54 ± 7.29	33.37 ± 13.10
Strategy6	77.65 ± 7.22	31.37 ± 9.95	81.91 ± 5.18	28.72 ± 9.15
Strategy7	73.86 ± 8.64	35.25 ± 11.80	78.77 ± 6.32	32.72 ± 11.22
Strategy8	70.70 ± 11.71	39.33 ± 15.43	75.77 ± 10.58	36.71 ± 15.03
Strategy9	78.18 ± 6.74	25.41 ± 10.67	81.68 ± 5.28	22.66 ± 8.43
Strategy10	78.98 ± 6.01	26.24 ± 10.35	82.24 ± 4.85	23.36 ± 8.18
Strategy11	72.51 ± 9.14	36.27 ± 12.03	77.58 ± 6.90	33.82 ± 11.64
Strategy12	71.64 ± 11.21	38.48 ± 15.87	76.77 ± 10.06	35.95 ± 15.56

Table 4.1: Results in ring detector applying SPD before PCA.

Looking at the tables, an important thing is that the LSG ring detector detect better but also introduce more noise (less precision), for any representation space (Figure 4.2 shows two different detection examples for each ring detector). SGF ring detector seems more clean but also lose some part of the valley.

We observe that both ring detection perform better in the color representation. Finally, result do not change significantly applying SPD filtering before or after PCA, this might be attributed to the nearly linear nature of

	Exp1		Exp2	
Strategy	Sensitivity(%)	Precision(%)	Sensitivity(%)	Precision(%)
Exp1	100	0.00	76.35 ± 13.41	66.61 ± 15.07
Exp2	66.42 ± 15.30	76.02 ± 13.09	100	0.00

Table 4.2: Results variability of the experts.

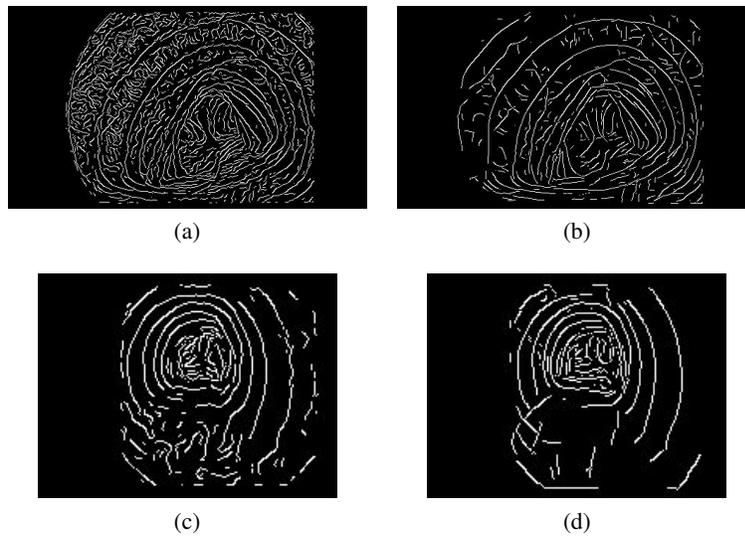


Figure 4.2: Examples of noisy detection: LSG detector (a)(c), SGF detector (b)(d).

SPD.

A common trend is a substantially low precision in our detection and this should be analysed. Whenever sensitivity increases precision decreases, that is, much good detections are at a cost of increase of noise. The problem of segmentation images motivates us to analyse which can be the variability between different observers. As all the tables show, the expert 1 is more restrictive than expert 2. Last table shows the similarity between the two experts. As you can observe in Table 4.3 expert segmentations are quite different, so it is difficult to have an exact metric to evaluate our strategies because there is a subjective dependence. If we take a look at the numbers, we can see that there is a 30% of difference in precision between each expert, therefore, the discrepancy between experts makes impossible has a precision high than 70%.

But what happens with the rest of the precision value? Where are we losing precision and why?

First of all we are going to check different results but just having a look to one strategy, i.e Strategy7. First hypothesis is that the methods are applied to four different sequences that have different resolution. Since we have tuned the parameters, it can affect to the valley detections. Figure 4.4 shows the different value of precision (RGBIntens strategy), for each sequence (seqn_1) of videobronchoscopy and using the same scale that are in the

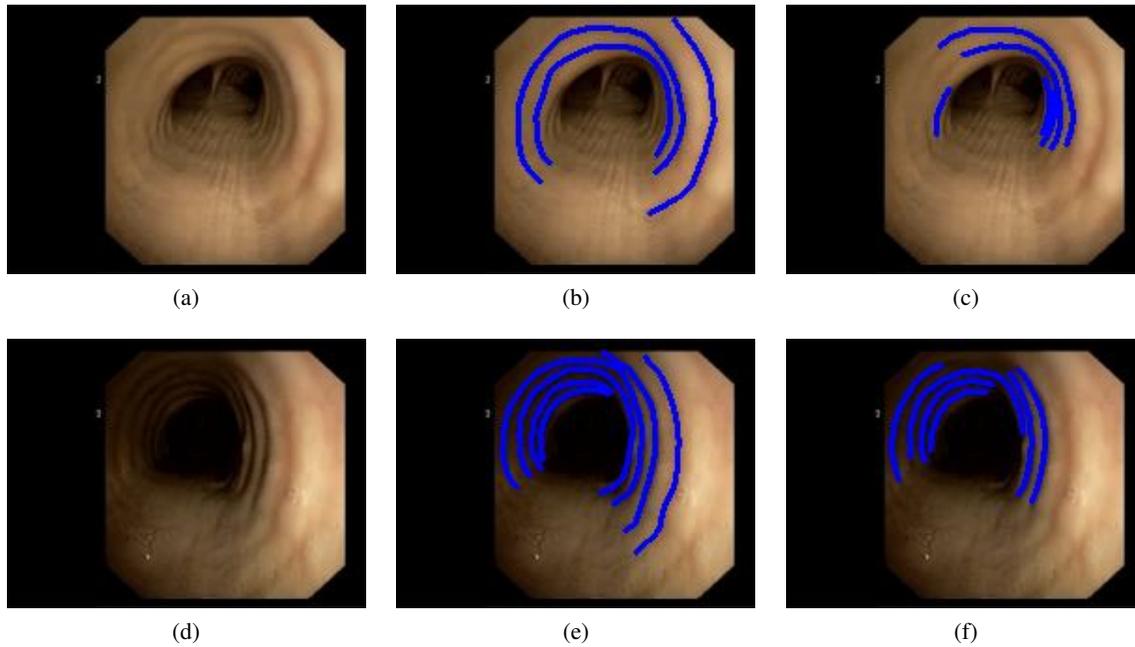


Figure 4.3: Segmentation of tracheal rings: original image (a)(d), expert 1 (b)(e), expert 2 (c)(f).

presented results. Figure 4.4 shows also the value of precision (seqn_h) but in this case we raise up the scale parameter in the SGF ring detector. Figure 4.5 shows how looks the detection using this different scales for each sequence of video. If we look at the graphic and the result images of these figures we can notice some sequence problems if we raise up the scale parameter:

- Sequence 1 loses precision because this video has a low resolution. As can be seen in the first row of Figure 4.5 if we use a large scale there is less noisy (it still is) but also there is less ring detection, so precision goes down because few detection of rings.
- Sequence 2,3,4 gain precision because this video has a higher resolution. As can be seen in the second row of Figure 4.5 if we use a large scale there is less noisy (small veins are removed) but ring detection is still doing well, so precision goes up. For the sequence 2 and 4 is a significant improvement.

We have seen a scale problem that can explain the small values in the precision of the detection. By the way, we can have some general comments that can still raising up the precision rate:

- High response of the SGF ring detector in the lumen/carina. It is shown in the second (Figure 4.5) column of all the sequences.
- There are detected rings but these are not in the groundtruth.
- High response of the SGF ring detector in the rigid artefact (third row in Figure 4.5).

- High response of the SGF ring detector in the borders of the image (first, second and last row in Figure 4.5).

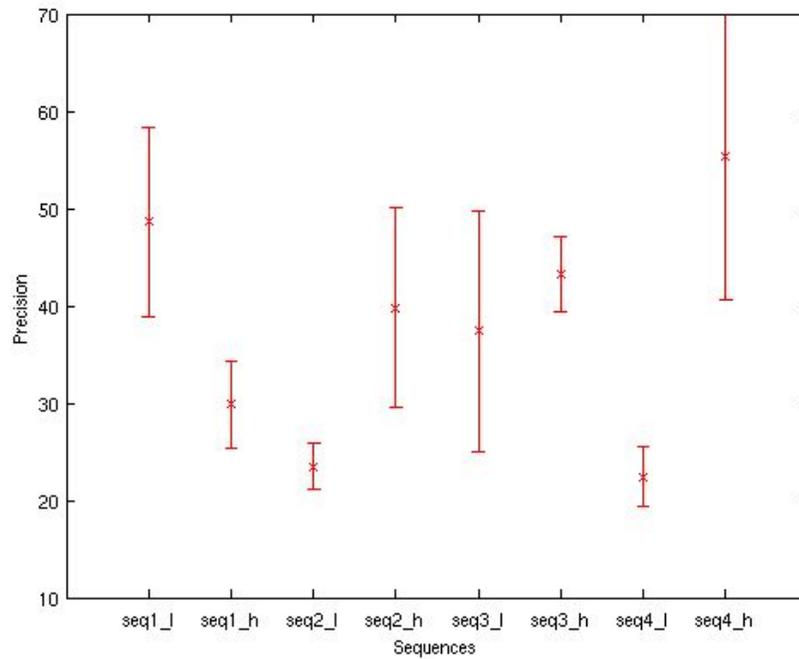


Figure 4.4: Precision value of using different scales: small (l) and large (h), for the different sequences.

Most of this problems can be solved if we take into account the structure of a tracheal ring but this was not a goal for this project. If we assume that we will have goods methods to clean detection image (that will be explained in the future work), the strategy that we are looking for is the high sensitivity one (ColorGeo).

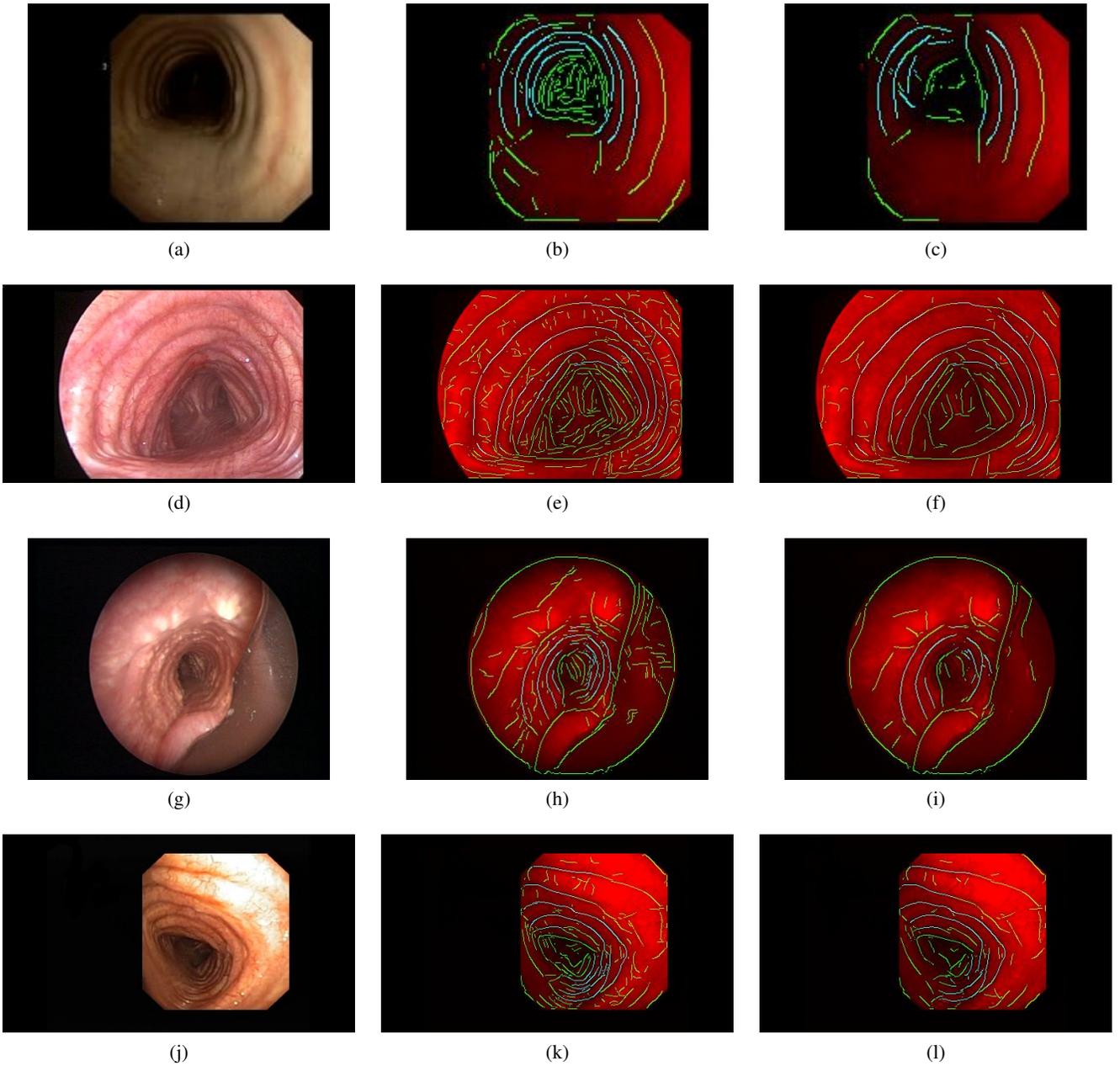


Figure 4.5: Ring detections by using different scales. Blue lines are good detections and green lines noisy detections. First column has the real image. Second column has the detected rings by using small scale. Third column has the detected rings by using large scale. Each row has a different sequence.

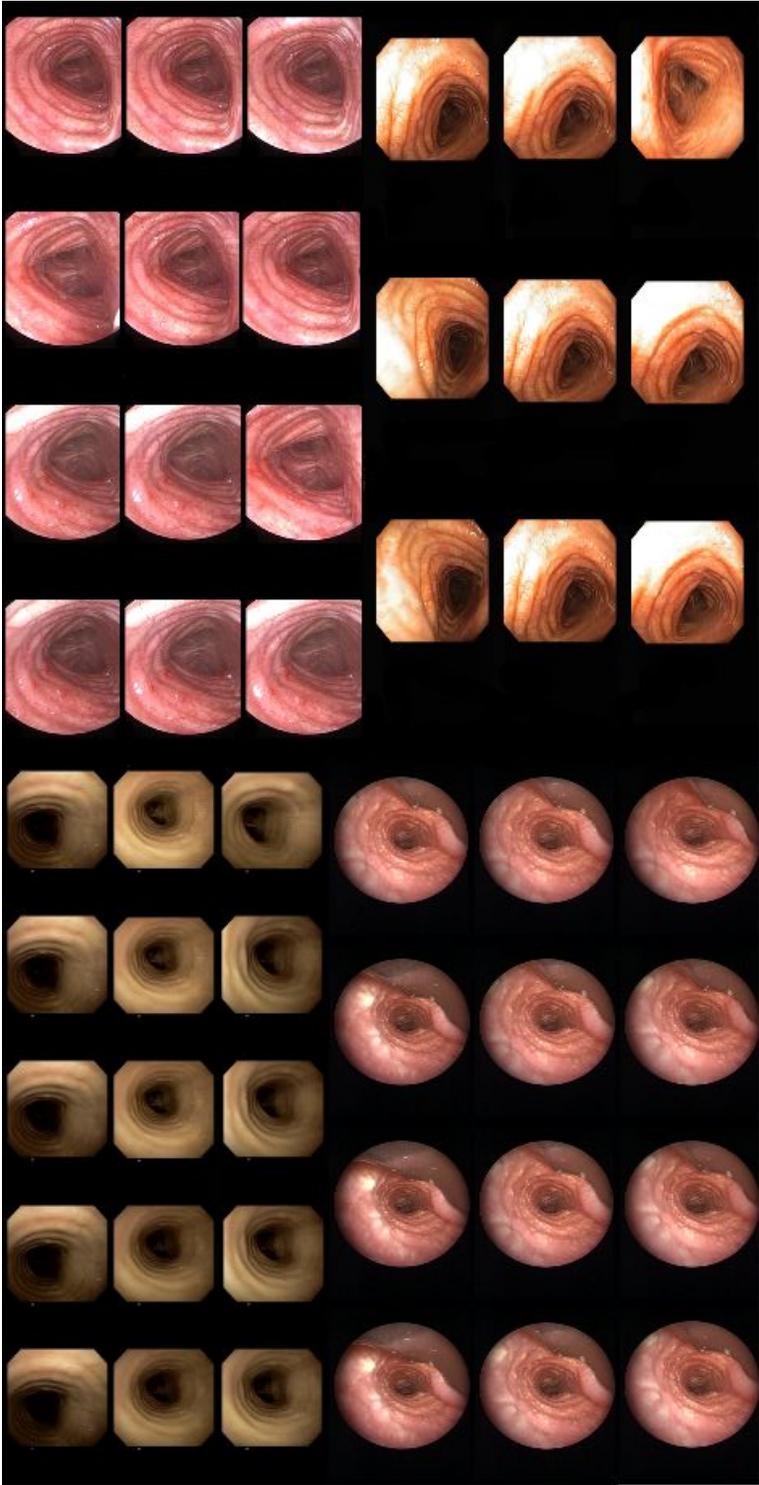


Figure 4.6: Trachea dataset.

Chapter 5

Conclusions and future work

5.1 Conclusions

Several phases have been done during the project. First, an analysis and categorization of video bronchoscopy frames have done. This helped us to put some constrains and focus the project in the tracheal part. Once we did this categorization we were able to create a dataset for validation looking for the frames that meets the conditions.

The acclimatization of the hysteresis method improve the results than using a simple and tuned threshold. Hysteresis thresholding is based in the histogram of each image (as we explained in workflow chapter).

Several strategies have been tested and the results can answer the questions that we propose at the begin of the report (our contributions). Since the project is a preliminary study, what we want is to know which way to take.

As we said in results chapter the variability between experts is high $23 - 33 \%$, so the values of sensitivity and precision are not accurate. But in the results there is trend that is following in all of them, that can help us to answer our goal questions.

Looking at the results the best strategy in terms of ring detection is using RGB color + LSG ring detector, Strategy2 (expert 1, 83.67 ± 5.55 , expert 2, 86.67 ± 4.40). It all most cover all the rings that we have segmented. Although, this strategy includes a lot of noisy (low precision). Taken into account the high intra-variability between experts we can say that, the sensitivity differences between strategies is not significant. This is because all the strategies has more sensitivity if we compare with the sensitivity between experts. In this way, our relevant metric should be the precision, therefore we are going to focus on the strategies that have more precision. This strategies are those that SGF valley detector is applied. Moreover, results make us to thing that our validating system is not a good way, and it can be a good new contribution to the problem.

Talking about color information, if we have a look to the results it can be seen that strategies with color have better results (both ring detectors), so we will say that color can be useful than the grayscale strategies. Talking about the best way of converting to gray images, PCA is worse than using grayscale image.

Finally, since the best strategies are those that precision is high, we have to find filtering strategies as a post-

process to raise up our values in precision, which means remove as much as possible noisy produced by: the resolution of the videos, carina/lumen, borders of the image or rigid artefact. As we said in the results chapter most of this noise will be removed when we apply some tracheal ring structure constrains in the process, that we can call it post-filtering process. Possible methods of post-filtering are explained as future work in the next chapter.

5.2 Future work

As we already seen, post-filtering step should be the future work to improve our precision in the strategies. As a first step, we have to detect the ridges in our images. Then we can combine valley and ridges to look for the structures that we have been explained in chapter 2. In this section we will explain some ideas to clean the images maintaining our regions of interest (tracheal rings) that follow this special structure. This ideas could be processed in parallel or in a single way.

- **SGF ring detector scales according to the resolution.** As we have seen before scale parameter has to be dependent with the resolution of the sequence.
- **Combining SGF and LSG ring detectors.** Good option can be to use the image given by SGF detector, that is a cleaner image, and reconstruct the rings by using the detected image from the LSG detector.
- **Noisy suppression following the normal of our detection.** Using the normal vector of our detections we can generate a new image drawing lines following the direction of this normal. The intensity of this lines will be the accumulative sum of all the pixels that are included in this line. The new image should has a white zone (a lot of lines pass through this zone) where all the concentric rings will have the center. Now we can generate a new image where we only will paint the pixels that are in the lines that pass through this white zone. With this process some noisy should eliminate, all the noisy that does not follow the structure of the concentric tracheal rings.
- **Noisy suppression by keeping [...valley,ridge,valley...] pattern.** This is an other option to remove noisy. We know that a tracheal ring is a structure of valleys and ridges following the explained pattern. As in the previous case we can use the normal vector of our detections and plot the distances between pixels classified as a valley and ridge. When there is a concentric rings we will find a constant distances in the plot instead of the noisy that will be distances non constant. Once we got this distance map we should remove those pixels that are not included in constant distances.
- **Curvature tracking.** We can analyse the detected rings by tracking pixels and its vicinity. If this vicinity follows some curvature we will maintain this pixels as a candidates to be a tracheal ring. There is a lot of noisy that does not follow any curvature, so noisy is going to be removed.

All of this methods can help us on removing noisy, specially in the detection problems of the lumen/carina and rigid artefact.

An other improvement can be to adapt the parameters for each sequence. It means that not all the videos has the same properties such as resolution or colors, so this system should adapt the parameters to the video conditions.

Finally, the validation protocol has to be examined by putting more constrains to achieve good and constant segmentations.

Bibliography

- [1] <http://www.olympus-global.com/en/corc/history/story/endo/origin/>.
- [2] http://www.alexea.org/main/index.php?option=com_content&view=article&id=23&Itemid=13.
- [3] <http://en.wikipedia.org/wiki/Bronchoscopy>.
- [4] <http://www.bronchoscopy.org>.
- [5] X. Luó, M. Feuerstein, T. Reichl, T. Kitasaka, and K. Mori, "An application driven comparison of several feature extraction algorithms in bronchoscope tracking during navigated bronchoscopy," *Medical Imaging and Augmented Reality*, pp. 475–484, 2010.
- [6] D. Deguchi, K. Mori, Y. Suenaga, J. Hasegawa, J. Toriwaki, H. Takabatake, and H. Natori, "New image similarity measure for bronchoscope tracking based on image registration," *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003*, pp. 399–406, 2003.
- [7] D. Koppel, Y. Wang, and H. Lee, "Image-based rendering and modeling in video-endoscopy," in *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*. IEEE, 2004, pp. 269–272.
- [8] K. Mori, J. Hasegawa, Y. Suenaga, and J. Toriwaki, "Automated anatomical labeling of the bronchial branch and its application to the virtual bronchoscopy system," *Medical Imaging, IEEE Transactions on*, vol. 19, no. 2, pp. 103–114, 2000.
- [9] I. Wegner, J. Biederer, R. Tetzlaff, I. Wolf, and H. Meinzer, "Evaluation and extension of a navigation system for bronchoscopy inside human lungs," in *SPIE Medical Imaging*. Citeseer, 2007.
- [10] S. M. â. B. E. Z. S ZILÁGYI, L.ĂŞ S ZILÁGYI, "Medical image segmentation for virtual endoscopy," *Proc. 16th IFAC World Congress*, no. paper1775, pp. 1–5, 2005.
- [11] S. ZILÁGYI, "Medical image processing methods for the development of virtual endoscope," *Periodica polytechnica ser. el. eng.*, vol. 50, no. 1-2, pp. 69–78, 2006.
- [12] <http://www.wisegEEK.com/what-is-ct-imaging.htm>.
- [13] G. A. El-Bayoumi, Ezzat; Silvestri, "Bronchoscopy for the diagnosis and staging of lung cancer," *Semin Respir Crit Care Med*, vol. 29, no. 3, pp. 261–270, 2008.

- [14] P. Wang, S. Krishnan, Y. Huang, and N. Srinivasan, "An adaptive segmentation technique for clinical endoscopic image processing," in *[Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] EMBS/BMES Conference, 2002. Proceedings of the Second Joint*, vol. 2. IEEE, 2002, pp. 1084–1085.
- [15] http://en.wikipedia.org/wiki/Vertebrate_trachea.
- [16] <http://en.wikipedia.org/wiki/Deinterlacing>.
- [17] http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MANDUCHI1/Bilateral_Filtering.h
- [18] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 839–846.
- [19] D. Gil, A. Hernández-Sabaté, M. Burnat, S. Jansen, and J. Martínez-Villalta, "Structure-preserving smoothing of biomedical images," in *Computer Analysis of Images and Patterns*. Springer, 2009, pp. 427–434.
- [20] <http://altfoto.com/2010/08/espacios-de-color-el-ojo-de-las-maquinas>.
- [21] http://es.wikipedia.org/wiki/Espacio_de_color_Lab.
- [22] <http://www.scribd.com/doc/16617255/HSI-Color-Space>.
- [23] http://en.wikipedia.org/wiki/Principal_component_analysis.
- [24] 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*, vol. 21, no. 4, pp. 327–335, 1999.
- [25] A. Rosset, L. Spadola, and O. Ratib, "Osirix: An open-source software for navigating in multidimensional dicom images," *Journal of Digital Imaging*, vol. 17, pp. 205–216, 2004, 10.1007/s10278-004-1014-6. [Online]. Available: <http://dx.doi.org/10.1007/s10278-004-1014-6>