An illumination model of the trachea appearance in videobronchoscopy images

Carles Sánchez[1] , Javier Sánchez[1] , Antoni Rosell[2] and Debora Gil[1] *

[1]Comp. Vision Center, Comp. Science Dep. UAB.
{csanchez,javier,debora}@cvc.uab.cat
[2]Pneumology Unit, Hosp. Univ. Bellvitge, IDIBELL, CIBERES.
arosell@bellvitgehospital.cat

Abstract. Videobronchoscopy is a medical imaging technique that allows interactive navigation inside the respiratory pathways. This imaging modality provides realistic images and allows non-invasive minimal intervention procedures. Tracheal procedures are routinary interventions that require assessment of the percentage of obstructed pathway for injury (stenosis) detection. Visual assessment in videobronchoscopic sequences requires high expertise of trachea anatomy and is prone to human error. This paper introduces an automatic method for the estimation of steneosed trachea percentage reduction in videobronchoscopic images. We look for tracheal rings , whose deformation determines the degree of obstruction. For ring extraction , we present a ring detector based on an illumination and appearance model. This model allows us to parametrise the ring detection. Finally, we can infer optimal estimation parameters for any video resolution.

Keywords: Bronchoscopy, tracheal ring, stenosis assesment, trachea appearance model, segmentation.

1 Introduction

Videobronchoscopy assessment focuses on three areas. First, it estimates stenosis percentage, that is, a reduction in air circulation in the tracheal airway. Second, it looks for pathological tracheal stent placement. Finally, it detects tumours [1]. In the two first cases, the physician has to infer measurements of the percentage of obstruction pathways (stenosis) and measurements for prosthesis length implantation by simple inspection of video frames images in perspective projection. Tracheal and bronchial rings are the main anatomical structures used to perform measurements. Although (tracheal) rings are cleary identified in bronchoscopy images, bronchologists lack of a automatic system providing objective measurements on their size. This is of extreme importance for determining the best prothesis in trachea and bronchi stenosis treatment. Current methods are

* This work was supported by the Spanish projects TIN2009-13618, CSD2007-00018 and TIN2009-10435. Debora Gil has been supported by the Ramon y Cajal Program. We would like to thank Albert Andaluz for his support.

based on either visual inspection or balloon placement. In the first case determining the true measurements strongly depends on the experience and the anatomical knowledge of the pulmonologist. In the second case the balloon is inflated to fit the trachea or bronchi to determine the size of the prothesis by comparison to a calibrated template. In any cases this semi quantitative methods result in a 30% of wrong prothesis choise and, thus, require recalibrate and repeat the whole procedure [2] [3]. Consequently, developing computer procedures for extracting the ring area from videobronchoscopy images would constitute a significant breakthrough in the field.

Due to the novelty of videobronchoscopy technology (it appears in clinical practice in the middle 90's) there are few works processing videobronchoscopy images. Most of them (or even all, as far as we know) address on navigation aid for bronchoscopy [4–8]. To our knowledge this is the first work addressing extraction of tracheal rings processing videobrinchoscopy images alone.

This is a challenging task due to the large variety of acquisition conditions, surgical artefacts and illumination artefacts. This paper contributes in two aspects. First, we describe the appearance of tracheal rings in images by means of illumination and 3D-2D projection models. Second, the model is used to define a parametric ring detector and infer the optimal parameters given the resolution. Experiments in sequences with different resolutions validate our strategy in terms of sensitivity and precision of the detected ring. A qualitative and quantitative results are presented.

2 Trachea appearance model

The trachea is a tabular structure located in front of the esophagus that connects the pharynx to the beginning of bronchial tree (carina). There are about fifteen to twenty incomplete C-shaped cartilaginous rings that reinforce the anterior and lateral sides of the trachea. In videobronchoscopy images the trachea appears as a tube in conical projection. If the camera is oriented in the axis of the trachea (so that the carina is on the center of the image) the conical projection of the trachea is given by a set of concentric circles corresponding to tracheal rings. Figure 2 shows the main anatomical structures that can be identified in a videobrinchoscopy frame adquired in central projection. Given that, out of center deviations collapse rings in a point close to the camera position. Those frames can not be provide reliable measures of the percentage of obstruction. Therefore, from now on, we will assume that our camera (projection) is centred at the carina and that the illumination is placed at the camera position.

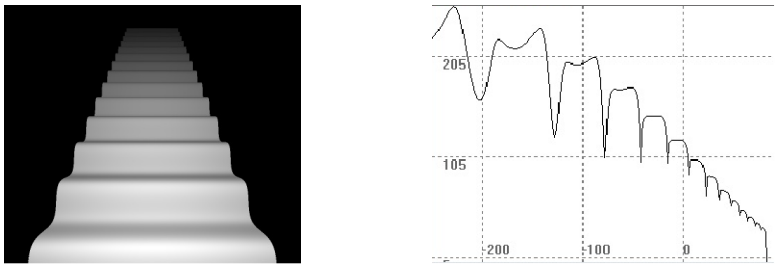
For the illumination appearance model we have chosen the Phong reflection model. According to this model, image intensity is given by the cosinus of the incident angle between light ray and the surface normal:

$$I = I_a K_a + fatt I_p (K_d \cos\theta + W(\theta)\cos\alpha) \tag{1}$$

where I is the light reflected by the surface to the camera, I_a is the intensity of ambient illumination, K_a is the ambient reflection coefficient, $fatt$ is the

attenuation factor of light distance, l_p is the spot light intensity, K_d is the diffuse reflection coefficient. θ determines the angle between the normal surface and the direction of the incidence spot light and α represents the angle between the normal surface and camera direction. Finally, $W(\theta)$ is the specular reflection coefficient.

Figure 1 shows the Phong reflection model applied to a sinusoidal surface simulating the 3D profile of tracheal rings illuminated from the camera point of view as it append in a bronchoscopy process. The left image is 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. Figure 2 shows a representative frame in central point of view and the radial intensity profile. We note that the radial intensity coincides with Phong illumination reflection pattern.



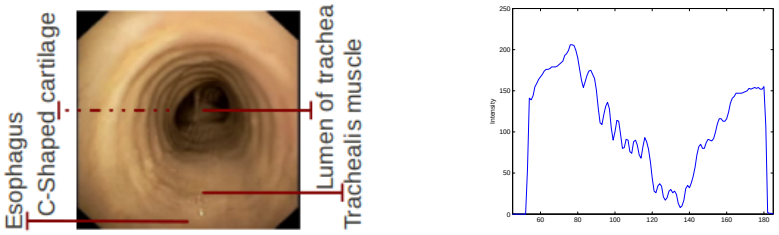
(a) (b)

Fig. 1. Phong model: simulation of the trachea surface (a) and plot of the intensity of the reflected illumination in each point of the surface (b).

By the analysis of (both) illumination and image radial intensity profiles, we conclude that the anatomical of the trachea can be described (in a computer vision words) as a concentric composition of valleys (between rings and shaded parts) and ridges (every ring and between valleys). Given that, since valleys present a more stable profile across the distance to the camera (radial direction in Figure 2), we will restrict to valley detectors.

In order to detect valleys we have choosed a method according to the non-uniform illumination of the images: Normalized Steerable Gaussian Filters (NSGF). The detector is defined by convolving the image with a second derivative of an anisotropic oriented gaussian kernel [9] :

$$g_\sigma^\theta = g_{(\sigma_x,\sigma_y)}^\theta = \frac{1}{(2\pi)\sigma_x\sigma_y} e^{-(\frac{\tilde{x}^2}{2\sigma_x^2}+\frac{\tilde{y}^2}{2\sigma_y^2})} \tag{2}$$



(a) (b)

Fig. 2. 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).

$$\left.\begin{aligned}\tilde{x} &= -x\sin\theta + y\cos\theta \\ \tilde{y} &= x\sin\theta + y\cos\theta\end{aligned}\right\} \tag{3}$$

for $(\tilde{x}, \tilde{y})$ the coordinates given by a rotation of angle θ and $\sigma_x > \sigma_y$. The second partial derivative along the axis constitutes the principal kernel for computing ridges or valleys:

$$\partial_y^2 g_\sigma^\theta = (\tilde{y}^2/\sigma_y^4 - 1/\sigma_y^4) g_\sigma^\theta \tag{4}$$

The response of the operator is calculated as the maximum response for a discrete sampling of the angle:

$$NSGF := \max_j (\partial_y^2 g_\sigma^{\theta_j} * I) \tag{5}$$

for $\theta_j = \{j\frac{\pi}{N}, \forall j = 1..N\}$ and applied to our image I.

We observe that the scale σ_y is related to the thickness of the valley and, thus, to the period of the sinusoidal approaching the intensity profile. In order to set its values we will use our illumination model. The interest of using the phong model is not the quantization of its parameters, whereas, a qualitative study of the profile of intensity reflected by the shape under a surface. Parameters as ambient illumination or attenuation factors modulate the amplitude of the sinusoid representing for the intensity value, but do not affect its period (related to the scale). It follows that, scales ranges of the bank of filters can be determined by the Phong model given by Formula 1. According to it (and due to the perspective projection), scales decrease for rings far from the camera (closer to the carina). According to it, σ_y ranges define the closest tracheal ring (max σ_y) and the farthest one (min σ_y). Furthermore, it suffices to determine the optimal scales for one sequence since scale ranges of any video will be given by the relation between their resolution and the resolution of the reference sequence.

3 Experiments

Our data set consist of four sequences of healthy trachea provided by Bellvitge hospital. Sequences include rigid and flexible bronchoscopies and were adjusted at different digital resolutions.

Scale parameters have been empirically set for the sequence of lowest resolution according to the illumination model. Ranges have been manually set for a subset of the sequence frames and then refined to give best results for the remaining frames. We have estimated the scale parameter for the remaining sequences by using scaling ratio between both resolutions. Scale ranges and image resolutions for all videos are given in Table 1.

	Seq1	Seq2	Seq3	Seq4
σ_{y1}	[0.5,2.5]	[1.8,6.4]	[1.4,5.2]	[1.8,6.4]
σ_{y2}	[0.7,2.7]	[2.0,6.6]	[1.6,5.4]	[2.0,6.6]
σ_{y3}	[0.9,2.9]	[2.2,6.8]	[1.8,5.6]	[2.2,6.8]
σ_{y4}	[1.1,3.1]	[2.4,7.0]	[2.0,5.8]	[2.4,7.0]
σ_{y5}	[1.3,3.3]	[2.6,7.2]	[2.2,6.0]	[2.6,7.2]
σ_{y6}	[1.5,3.5]	[2.8,7.4]	[2.4,6.2]	[2.8,7.4]
Resolution	192x144	512x288	360x288	512x288

Table 1. Scale ranges and resolutions for each sequence.

In order to validate the former strategy for scale detection, we will compare the accuracy of the rings detected using scale in Table 1 to a ground truth.

Comparison has been made over a set of 60 representative frames (see mosaic in Figure 3) manually segmented by a medical expert twice. Differences between ground truth and detected rings are measured in terms of sensitivity (true positives TP) and and precision (false positives FP).

$$Sens = \frac{\#(Detect \cap GT_{\oplus})}{\#GT} \tag{6}$$

$$Prec = 1 - (\frac{\#Detect - (\#(Detect \cap GT_{\oplus}))}{\#Detect}) \tag{7}$$

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 dilation.

Figure 4 shows the results for each sequence in terms of sensitivity and precision. We show the tendency of sensitivity and precision ranges for different range of scales (see Table 1) considered for each sequence. A remarkable issue is that, although ranges are different, all four sequences present equal 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.

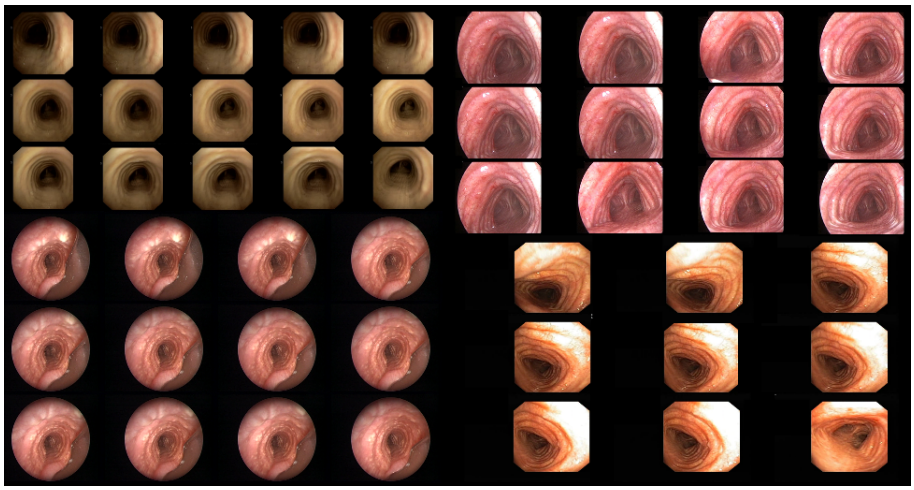
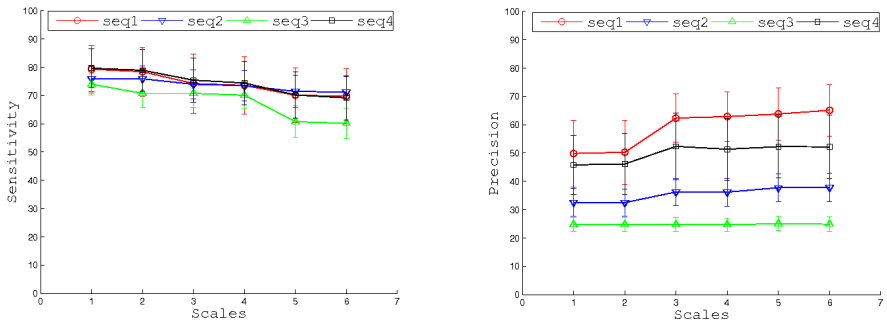


Fig. 3. Trachea dataset.

Figure 5 shows the results obtained for the best scale ranges. Green lines are the groundtruth and blue lines the automated detection. We observe that, as expected, all rings are detected, while there are some extra structures dropping precision. There are two main sources of precision dropping. The first problem is that there are false positive rings in the carina in all sequences. Also, 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. Finally, there are some rings (especially the most internal and external ones) properly detected by NSGF that were not identified by the expert.



(a) (b)

Fig. 4. Results for each sequence: sensitivity (a) and precision (b).

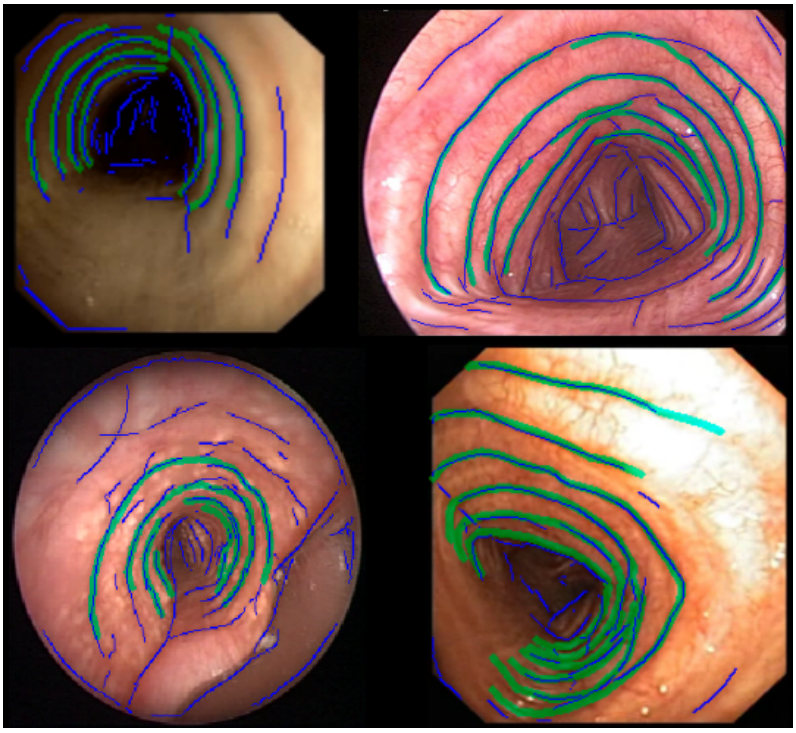


Fig. 5. Qualitative results for each sequence. Green lines are the groundtruth and blue lines our detection.

4 Conclusions

Tracheal ring detection for stenosi assessment in videobronchoscopy images is a challenging task hardly addressed, so far. Our experiments show that our detection approach is robust, as long as we have a valid reference. 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. Quantitative results prove that our detection can retrieve most clinically relevant tracheal rings.

However, there is still room for improvement, as a post-filtering stage would remove non-ring structures. In the future, we would like to translate our ring detection into the polar coordinates domain in order to remove orientation dependencies and use the concentric geometry of the rings. In this way, we will define an adaptive ring filter based on the luminance model. Finally, in order to minimize the impact of miss identifications in manual tracing, alternative validation protocols should be addressed.

References

1. C.T. Bolliger and P.N. Mathur, *Interventional bronchoscopy*, vol. 30, S Karger Ag, 2000.
2. S. Norwood, V.L. Vallina, K. Short, M. Saigusa, L.G. Fernandez, and J.W. McLarty, "Incidence of tracheal stenosis and other late complications after percutaneous tracheostomy," *Annals of surgery*, vol. 232, no. 2, pp. 233, 2000.
3. J.M. Vergnon, F. Costes, M.C. Bayon, and A. Emonot, "Efficacy of tracheal and bronchial stent placement on respiratory functional tests," *Chest*, vol. 107, no. 3, pp. 741–746, 1995.
4. D. Deguchi, K. Mori, M. Feuerstein, T. Kitasaka, C.R. Maurer, Y. Suenaga, H. Takabatake, M. Mori, and H. Natori, "Selective image similarity measure for bronchoscope tracking based on image registration," *Medical Image Analysis*, vol. 13, no. 4, pp. 621–633, 2009.
5. X. Luo, T. Kitasaka, and K. Mori, "Manismc: a new method using manifold modeling and sequential monte carlo sampler for boosting navigated bronchoscopy," *MICCAI 2011*, pp. 248–255, 2011.
6. T. Reichl, X. Luo, M. Menzel, H. Hautmann, K. Mori, and N. Navab, "Deformable registration of bronchoscopic video sequences to ct volumes with guaranteed smooth output," *MICCAI 2011*, pp. 17–24, 2011.
7. X. Luo, T. Kitasaka, and K. Mori, "Bronchoscopy navigation beyond electromagnetic tracking systems: a novel bronchoscope tracking prototype," *MICCAI 2011*, pp. 194–202, 2011.
8. Y. Schwarz, J. Greif, H.D. Becker, A. Ernst, and A. Mehta, "Real-time electromagnetic navigation bronchoscopy to peripheral lung lesions using overlaid ct images," *Chest*, vol. 129, no. 4, pp. 988–994, 2006.
9. W.T. Freeman, E.H. Adelson, Massachusetts Institute of Technology. Media Laboratory. Vision, and Modeling Group, "The design and use of steerable filters," *IEEE Transactions on Pattern analysis and machine intelligence*, vol. 13, no. 9, pp. 891–906, 1991.