16

Towards Intelligent Systems for Colonoscopy

Jorge Bernal, Fernando Vilariño and Javier Sánchez
Computer Vision Center and Computer Science Department,
Universitat Autònoma de Barcelona
Spain

1. Introduction

Colorectal cancer is one of the leading causes of cancer related deaths. Colorectal cancer's survival rate depends on the stage in which it is detected, decreasing from rates higher than 95% in the first stages to rates lower than 35% in stages IV and V (Tresca, A. (2010)); hence the importance of detecting it on its early stages by using screening techniques, such as colonoscopy (Hassinger, J.P., Holubar, S.D. et al. (2010)), which is still considered the gold standard for the screening of patients for colon cancers and lesions. Classical focal colonoscopy has been proved to be a successful tool for colon screening, although other tools are being also used for this purpose, such as Virtual Colonoscopy, Computed Tomography Colonoscopy, Chromoendoscopy or Wireless Capsule Video Endoscopy, among others.

In this chapter we present tools that can be used to build intelligent systems for colonoscopy. The idea is, by using methods based on computer vision and artificial intelligence, add significant value to the colonoscopy procedure. Intelligent systems are being used to assist in other medical interventions. For instance, we can build systems that can be used to develop the knowledge bases used by expert systems, such as KARDIO (Bratko et al. (1990), which was developed to interpret electrocardiograms. Another example can consist of developing a system that, in the context of anesthesia, provides a robust/reliable control system that could determine the optimal infusion rate of the drugs (muscle relaxant, anesthetic, and analgesic) simultaneously, and titrate each drug in accordance to its effects and interactions. Such a system would be a valuable assistant to the anesthetist in the operating theater. An example of such a system can be found in the work of Nunes et al. (2005). More close to our topic of interest, colonoscopy, we can find many examples of intelligent systems build to assist in cancer detection. Such is the case of breast cancer detection (Wei et al. (2011)) or prostate cancer detection (Viswanath et al. (2011)).

The question that arises now is: how can intelligent systems help in colonoscopy? What kind of applications these systems can be built for? In Figure 1 we depict some of the potential areas related to colonoscopy where an intelligent system can play a key role.

As shown in Figure 1, we foresee four different areas where an intelligent system can be introduced and add significant value to the colonoscopy procedure:

1. The most manifest application of this kind of systems could be the **assistance in the diagnosis procedure** during the intervention or in post-intervention time. This could be very useful in order to reduce the miss rate associated to polyp identification.

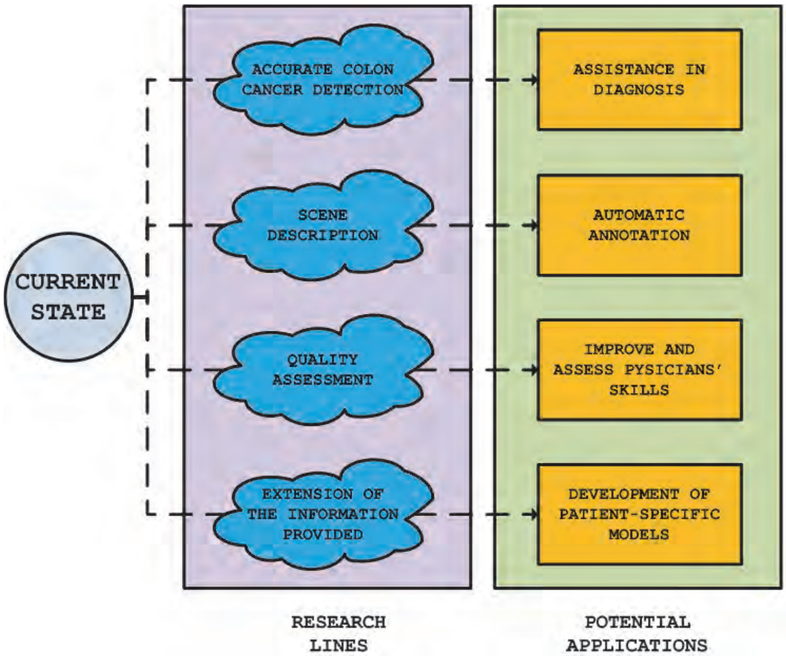


Fig. 1. Research lines and potential applications in the scope of intelligent systems for colonoscopy.

2. We can make use of the scene description provided by an automatic system -including the presence of salient traits, such as informative frames, anatomical structures, insertion or withdrawal phases, etc.- in order to **automatically annotate colonoscopy videos**. This would potentially provide a very efficient way of case annotation, with multiple uses in different applications.
3. In addition, an intelligent system may offer a quality assessment of the colonoscopy intervention, which could provide a non-subjective way of assessment. This could also be used as a way to train physicians in a way such they can **assess and improve their skills** without the cost associated to a real interventions, and it would allow to compare different performance metrics objectively.
4. We can also build intelligent systems that extend and provide additional information from colonoscopy data. Belonging to this area we can think of applications such as the **development of patient-specific models**, that can be re-used later, when a new study arrives, to check for coincidences that can help in the diagnosis and enrich in this way a final case report.

From the examples mentioned above, we can see that intelligent systems can indeed play a role in the future of colonoscopy. The objective of this chapter is to give an insight into this future, to expose what tools could be used to build an intelligent system for colonoscopy. In the next section we present the object of analysis in colonoscopy video by introducing the endoluminal scene. In Section 3 we deploy the different methods proposed to extract information from the endoluminal scene by explaining the techniques for image processing, the methods for scene description, the current proposals for polyp detection, the approaches pointed for quality assessment, and an insight into the potentiality of developing patient specific models. Next, we address the problem of the construction of databases of colonoscopy video and the acquisition of the ground truth. Finally, in Section 4 we fully develop our contribution with one example of a polyp segmentation algorithm.

2. Endoluminal scene

Before starting with the explanation of the several methods that can be used in building intelligent systems for colonoscopy, it is necessary to identify what is present on the endoluminal scene. This is compulsory because by knowing what we may find in images, we can develop methods that are devoted or attracted to the individual parts or objects of interest in the image.

As presented in Figure 2, the colonoscopy endoluminal scene shows the intestinal lumen, i.e., the inner region of the colon. The colon presents a tubular shape defined by the intestinal walls. Throughout a colonoscopy screening, the appearance of the intestinal walls varies in color and texture not only among different subjects but also among some of the different regions of the colon (ascending, transverse, descending and also sigmoid colon). Regardless of this intrinsic variability, the endoluminal scene consistently shows the lumen (1), the folds and wrinkles of the intestinal walls (2), the blood vessels (3), and eventually, diverticulosis

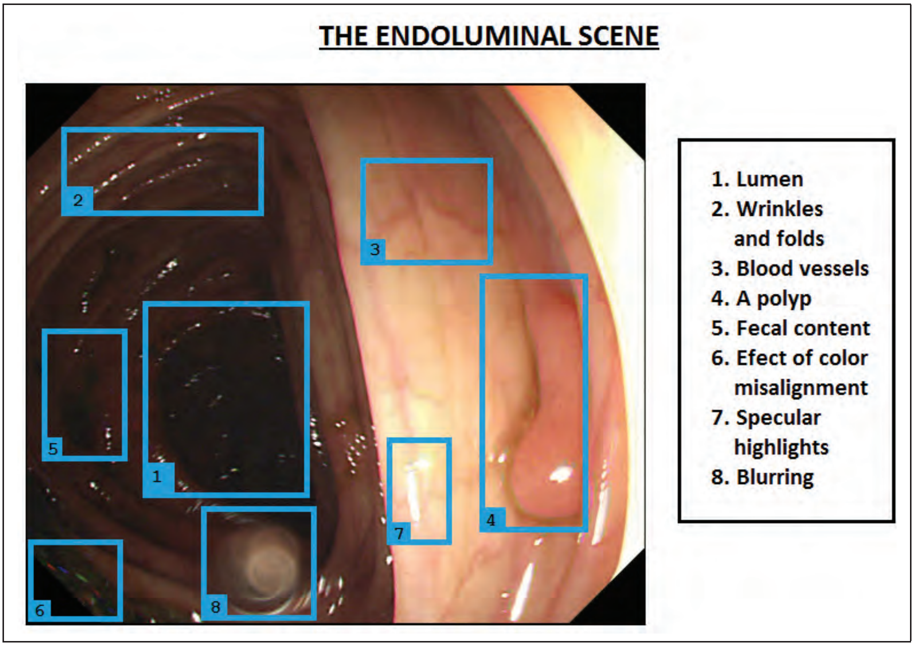


Fig. 2. Complete endoluminal scene: 1) Lumen, 2) wrinkles and folds, 3) blood vessels, 4) a polyp, 5) fecal content, 6) effect of the color misalignment, 7) specular highlights, and 8) blurring.

and lesions, such as ulcer, localized bleeding or polyps (4). Due to the flexible and extendible nature of the colon, and in part owed to the impact of the probe insertion or withdrawal in its deformation, it is difficult to find a perfect tubular appearance in the colon lumen because intestinal walls can be bent and folded. In addition, the wrinkles associated to the colon physiological structure appear in the scene as radial protuberances which modify the flat surface of the intestinal walls. On the intestinal walls, blood vessels are observed with their characteristic tree ramifications, presenting a certain variability associated to their width. Diverticulosis are shown as cavities or holes in the intestinal wall. The lesions related with bleeding are generally identified by its characteristic color. Polyps present a large variety in shapes, and seldom show a discriminative change in texture and/or color in comparison to the surrounding area. Despite a preparation is required for most of the colonoscopy interventions -with the aim of eliminating all fecal matter so that the physician conducting the colonoscopy can have a clear view- in many cases intestinal content is still present after the preparation procedure, and this intestinal content will hinder the right visualization of the intestinal walls. The procedure of elimination of the remaining fecal matter (5), consisting of the direct injection of water through the colonoscope in order to dilute the intestinal contents, turns out into the blurring of the video sequence and the appearance of bubbles. Finally, during the time of intervention, some tools used by the physician for different tasks -i.e., biopsy, cauterization, etc-. can be part of the visual scene.

In addition to the difficulties associated to the characterization of the colonoscopy scene due to its high variability and complexity, there are many visual artifacts the impact of which should be taken into account in order to tackle a robust system for the automatic analysis of colonoscopy video. The nature of these artifacts is twofold: 1) On the one hand, image acquisition systems are usually based on a tree-channel RGB image sensor which acquires each color component at a different time. In the event of motion, this delay on the acquisition time can yield to color phantoms (6) (Dahyot et al. (2008)). In addition to the former, and in order to optimize the dimensions of the image sensor and minimize its price, the video sequence is usually a composition of two interlaced fields (odd and even), corresponding to a compound of the RGB channels at different times. The result is an interlacing effect which is perceptually filtered by the observer during the video screening, but which is present at a frame level, carrying out the characteristic textured pattern of interlaced video. Finally, a different artifact, technically designated as specular highlights (7), is produced when the camera is frontally focusing a reflecting surface, the outcome of which is a white saturation in the image frame. 2) On the other hand, due to the limited bandwidth and frame rate of the video capture device, fast movements of the colonoscope can give away blurred images with diffused visual features (8). All these artifacts can have a potential impact into the automatic systems, and for this reason these artifacts should be addressed efficiently in order to guarantee an optimal performance.

Diverse endoscopy techniques, such as capsule endoscopy (both for small bowel and colon), bronchoscopy, gastroendoscopy, etc. show different endoluminal scenes, each of them with particular features. Besides, there is a variety of imaging methods to be used to enhance particular physiological targets, which is the case for narrow band imaging or chromoendoscopy, just to mention a few. This situation sets up a heterogeneous scenario from the perspective of automatic analysis using computer vision, and makes it not feasible to tackle the endoscopic image problem as a whole. However, it is possible to take some of the methods used in a given technique and adapt them to the specific particularities of colonoscopy video. For example, the automatic detection of intestinal content is a topic addressed in the bibliography of capsule endoscopy (as it can be seen in Vilariño (2006)) by means of the analysis of color distribution and texture, and its equivalent to the detection of intestinal content in colonoscopy would require relatively minor modifications. The identification of elliptical shapes in the image in order to detect potential polyps is a recurrent subject in virtual colonoscopy, and evolved versions of some of these techniques can be tried to colonoscopy video by adapting the search context, as we propose in Section 4.

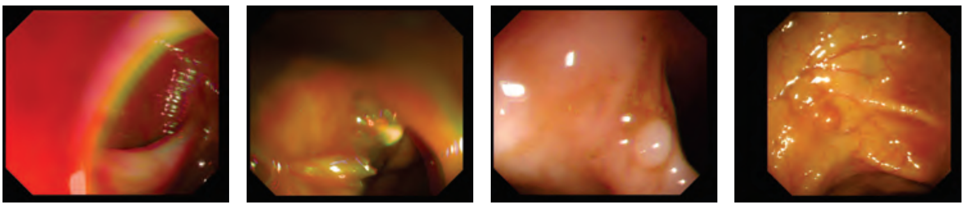
3. Building up an intelligent system for colonoscopy

Once we know what can appear in the endoluminal scene, which is where we will extract all the information from, we present in this section several tools that can provide information from the image frames. Since the number of works that apply computer vision and artificial intelligence is large, we have grouped them into the four potential areas suggested in Section 1. As an introduction, we first analyze which are the preliminary steps that should be performed in colonoscopy video.

3.1 Image preprocessing

Image preprocessing is needed in order to eliminate or minimize the impact of image artifacts associated to colonoscopy video, which fundamentally consist of the color phantoms and the presence of specular highlights.

The problem of color phantoms associated to the temporal misalignment of the color channels has been addressed in the work of Dahyot et al. (2008). This happens because most colonoscopy devices use monochromes cameras, in which the RGB components are taken at different times. This causes a worsening in the quality of the images, as it can be seen in Figure 3 a,b), which may difficult posterior image analysis tasks. The method proposed involves both color channels equalization and the estimation and compensation of the camera motion. The experimental results that can be found in Dahyot et al. (2008) show a global improvement in the quality of the images, failing only in cases when the quality of the original image is very low, although the evaluation is done qualitatively.



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

Fig. 3. Examples of: (a-b) color channel misalignment (c-d) specular highlights.

Adding a solution to a different problem, the work of Arnold et al. (2011) presents a method to improve the quality of colonoscopy videos. More precisely, the authors offer a solution for two problems, the already mentioned color channel misalignment and the apparition of specular highlights. Specular highlights appear on the intestinal surface as an effect of frontal illumination, as can be seen in Figure 3 c,d). The issue of specular highlights is addressed using two methods: 1) segmentation based on nonlinear filtering and a posterior color image thresholding and 2) fast inpainting method. We have also faced the problem of correction of specular highlights, as it can be seen in Bernal, J. and Sánchez, J. and Vilariño, F. (2011). In our case our objective was to provide a fast method that replaces the specular highlights with an interpolative approximation of what could really be under them, based on the information that surrounds the specular highlights.

3.2 Scene description for automatic video annotation

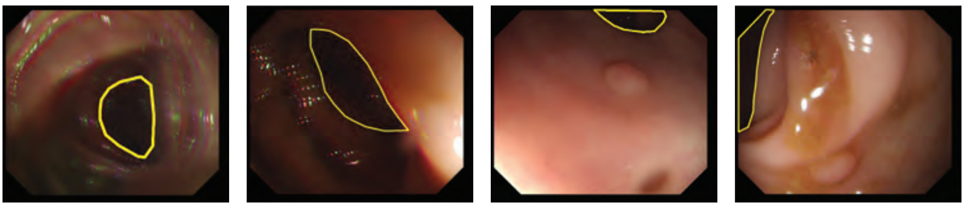
The majority of the existing literature devoted to scene description in colonoscopy video can be grouped according to three different topics, namely: 1) segmentation of the lumen; 2) definition of non-informative frames, and 3) polyp detection. Taking into account the importance of the latter we will analyze the first two topics, and later on we will deploy the different approaches to polyp detection in a full subsection.

3.2.1 Lumen segmentation

The detection of the lumen and its position can be crucial, for example, in post-intervention video processing. Frames where the proportion of lumen out of all the image is large can be related to the progression of the colonoscope through the patient (Figure 4 a-b)). On the other hand, frames where the amount of lumen presence is low (Figure 4 c-d)) may potentially indicate areas of the image where the physician has paid more attention. In addition to that, an efficient lumen segmentation may lead to remove great part of the image for a further computational analysis.

Several works are centered on lumen detection, such as the work of Gunduz-Demir et al. (2010), which aims at decomposing the tissue image (where lumen may be present) into a set of primitive objects and segment glands making use of the organizational properties of these objects. In this approach, an image is first decomposed into its tissue components that, as they are difficult to locate, are approximately represented transforming the image into a set of circular objects (nucleus and lumen objects). Following a similar line of research, the work of Tian et al. (2001) presents an automatic segmentation algorithm for lumen region and boundary extraction from endoscopy images which uses an adaptative Iris filter applied to the Region of Interest (segmented by an adaptative progressive thresholding method).

The work of Zhang et al. (2010) presents an approach for lumen segmentation which uses the information that contrast agents provide. Following this trend, the work of Bevilacqua et al. (2009) addresses lumen segmentation by first estimating the centerline, which can be achieved by first removing the background and then extracting air regions with a threshold filter.



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

Fig. 4. Examples of lumen (surrounded by a yellow boundary): a) and b) full view, c) and d) partial view.

Although used in CT and virtual colonoscopy respectively, the extension of these methods to focal colonoscopy may lead to more efficient lumen segmentation techniques due to their bidimensional nature.

3.2.2 Definition of non-informative frames

The content of the image itself can lead to the definition of informative and non-informative frames. In this domain of application, non-informative frames are those that, either their quality is so damaged (by the artifacts, hindering intestinal content, etc.) that it is difficult to extract information from them, or they are clinically uninteresting for a given task. For instance, frames where the instrumental take up great part of the image may not be relevant for polyp detection tasks. An accurate detection of the non-informative frames could also lead to a great reduction in the processing time of a stored colonoscopy intervention. Fundamentally, this information may be used for automatic video annotation and efficient video indexing and retrieval.

There are a few works that are centered on the **identification of non-informative frames**. The work of Arnold et al. (2009) addresses the identification of clinically uninteresting frames by analyzing the energy of the detail coefficients of the wavelet decomposition of a given image, which is used as the input to the classification system. In this case non-informative frames are those which do not carry any useful clinical information, such as those that occur when the camera is covered with liquids or when it is very close (even touching) the mucosa. These cases do occur frequently in colonoscopy procedures leading to extremely blurry images. This method is based on the 2D discrete wavelet transform which results in a set of approximation and detail coefficients. The approximation coefficients represent the low frequency content of the image while the detail coefficients hold the complementary high frequency information. The authors use detail coefficients to distinguish between informative and non-informative frames holding on the fact that the norm of the detail coefficients will be lower for low contrast images, making them more likely to be classified as non-informative.

The work of Cao et al. (2007) presents a method that extract those frames which correspond to a diagnostic or therapeutic operation, following work done in other domains (i.e., detecting important semantic units such as scenes and shots). This work takes profit of several characteristics that colonoscopy videos present, such as the presence of many blurred frames due to the frequent shifts of the camera position while it is moving along the colon. The identification of the operation shots is based on the detection of diagnostic or therapeutic instruments. In this case the authors map the problem of detecting instruments to the problem of detecting the cables of these instruments as they are present in the operation, regardless of their type. The architecture scheme shown in Figure 5 consists of five different steps which involve: 1) image preprocessing, to remove the effects of the specular highlights; 2) identification of the insertion direction of an instrument; 3) region filtering, where regions that are not part of the cable are removed; 4) region merging, which combines regions where parts of the instrument appears and 5) region matching, which matches the candidate regions in the image with the cable and without the cable.

Apart from the two methods presented related to the identification of non-informative frames, other approaches have been carried out such as the work of Oh et al. (2003) where a measure called the *isolated pixel ratio* (IPR) is used to classify the frames into informative, ambiguous and non-informative. The IPR measure is calculated from the edges of the image: an edge pixel that is not connected to any other edge pixel is defined as an isolated pixel. Those

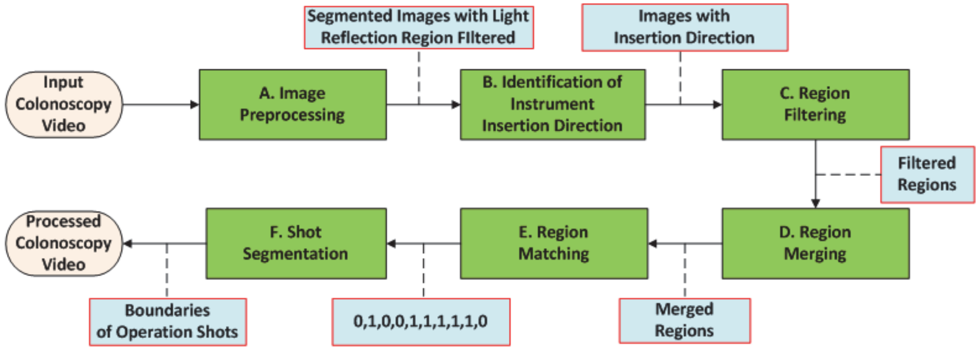


Fig. 5. A system architecture for operation shot detection as described in Cao et al. (2007).

isolated pixels are counted for each frame and are put in relation with the total number of edge pixels to obtain the IPR ratio.

Finally, an example of an endoscopic full multimedia information system for video annotation implementing many of these approaches is described in the work of Liu et al. (2007).

3.3 Automatic polyp detection

The main objective of the colonoscopy procedures is to check the status of the colon and to find possible lesions and cancer polyps on it. The general appearance of the polyps has been covered widely by medical bibliographic sources. However, there is a great variability in polyp appearance in colonoscopy videos, since there are some challenges that hinder polyp detection, namely: 1) the non-uniform appearance of polyps (see Figure 6 a-b); 2) their shape, flat or peduncular (Figure 6 c-d); 3) the effects of image acquisition, such as changes in pose, blurring, occlusions, specular highlights (Figure 6 e-g), and 4) the high similarity between the tissues inside and outside the polyp, which disables the possibility of relying only on texture or color cues (Figure 6 f), just to mention a few. The direct application of the methods presented in this section can be the potential assistance in the diagnosis, both during and in post-intervention time.

In the case of polyp detection, the great majority of the approaches is based on the analysis of features detected in the image. In the context of image processing, features can be defined as singular visual traits, associated to the visual primitives that constitute an object, such as edges, corners or lines, among others. The usual procedure is to use *feature detection* methods to locate the potential ROIs of the image and then describe them using one or many *feature descriptors*. After doing an extensive research on the different types of feature descriptors (Bernal, J. and Vilariño, F. and Sánchez, J. (2010)), we have divided them into four groups: shape, texture, color and motion.

Shape-based approaches

These approaches observe the structure of polyps as they appear in images and find the shapes which polyps commonly have. In general, polyps present two different shapes: flat or peduncular, as can be seen in Figure 6 a-b). What makes polyp detection by shape features difficult is that, in many times, we do not have a perfect shot of the polyp but an image where its pose, size and appearance can vary largely. Thus, many of the approaches presented try to

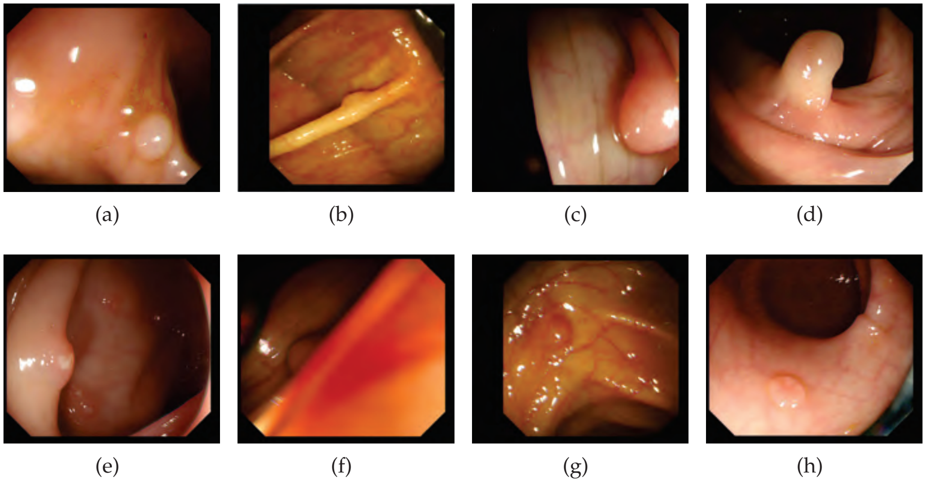


Fig. 6. Challenges in polyp detection: (a-d) non uniform appearance, e) partial (lateral) views, f) blurred images, g) specular highlights, and h) uniform texture and color inside and outside the polyp.

detect polyps not by detecting its whole shape but by detecting parts of the image that may indicate polyp presence.

For instance, flat polyps are meant to have elliptical shapes so one way to detect polyps is trying to find which structures in the image are surrounded by boundaries that constitute ellipses. The difficulty in this case is that in many occasions we do not have complete boundaries or the point of view makes it difficult to fit elliptical shapes (such is the case of lateral views). The works presented in this subsection could also be classified into two categories, namely: 1) detection by curvature analysis and 2) detection by ellipse fitting.

As it has been mentioned, **curvature analysis** is used to perform polyp detection. We can use curvature in different ways. For instance, we can check the curvature profile of the boundaries that appear in the image, that may have been detected by means of an edge detector. An example of the former can be consulted in the work of Krishnan et al. (1998), where the authors present an approach to polyp detection by means the extraction of contours. These contours are smoothed afterwards to make them suitable for curvature computation.

The work of Zhu et al. (2010) elaborates on the use of curvature-based shape measures (such as the shape index, curvedness or mean curvature) to analyze the local shapes in the colon wall. The problem in this case can appear in spurious calculations indicating high curvature, which is observed when the kernel contains two surfaces. Another problem that can appear is the discontinuities in curvature, which is shown when the gradient magnitude necessary to calculate the curvature vanishes. One possible solution to this is the redefinition of curvature as the magnitude of the change of the surface normal.

A different approach can be found in the work of van Wijk et al. (2010). The authors present a method that enables automated detection and segmentation of colorectal polyps, proposing a method that measures the amount of protrudedness of a candidate object in a scale adaptative fashion. The work uses the second principal curvature flow as a way to remove protruding elements from a curve in 2-D. We can suppose that the the points on the convex region of the polyp (the polyp head) are iteratively moved inwards, flattening the object. The convex region expands during the process and will ultimately include the polyp neck. After a certain amount of deformation, the surface flattening is such that the protrusion is completely removed, as if the object was never there. The idea here is to observe the second order differential properties of the implicit surface embedded. If we consider the colon as a long elongated structured tube, for a perfect perfect cylinder shape the principal curvatures are smaller than or equal to zero everywhere. However, the colon contains many folds, i.e., structures which are bended only in one direction: the first principal curvature is larger than zero, whereas, the second principal curvature is close to zero. Protruding objects, such as polyps, have positive values for the first and second principal curvature. Therefore, an operator is designed to affect only on points with a positive second principal curvature and in such a way that the second principal curvature decreases. The distinction between the head and neck is made by the sign of the second principal curvature. On the line connecting the inflection points, which in this case separate head to neck, the curvature is 0. It is clear that this method works well for images where we have clearly identified what is tissue and what is not so, in this case, the use of curvature measures can lead to good results in polyp detection.

The other category that collects many works is polyp detection by **ellipse fitting**. The general idea is, once we have a set of boundaries in the image, try to fit elliptical shapes in them. Belonging to this group we can observe in the work of Kang, J. and Doraiswami, R. (2003), which performs an edge detection in each of the R, G and B channels after applying a contrast enhancement algorithm. In order to classify the several regions (connected by closed edges) this method uses area, color and elliptical shape.

An approach that combines both curvature and ellipse fitting can be found in the work of Hwang, S. and Oh, J.H. and Tavanapong, W. et al. (2007). The method presented consists of fitting ellipses into the frontiers obtained after a first segmentation, and then classifying candidate regions by considering curvature, distance to edges and intensity value. Without entering into many details, in order to detect the ellipses an edge image is needed, where desirable edges should be grouped. Taking into account the challenges that colonoscopy images present, only some parts of the polyp boundary will have strong edge information so, based on this, the method uses the marker-controlled watershed (Vincent, L. , Soille, P. (1991)) algorithm for polyp segmentation because it can handle the gap between broken edges properly. Then, using the edges in each segmented region, the method generates an ellipse using an ellipse fitting method. Finally the number of final ellipses is reduced by removing those which do not represent actual polyps filtering by curve direction and curvature, by edge distance and by intensity value.

There are some works that cannot be assigned to an specific category because they use methods that appear in both curvature and ellipse-fitting categories. For instance, the work of Dhandra et al. (2006) also starts with a watershed segmentation but it performs its detection scheme by using color information. As it will be presented later for the case of texture descriptors, in the work of Coimbra & Cunha (2006), MPEG-7 descriptors are used in polyp detection tasks. In the field of shape descriptors, *region-based shape descriptor* is presented. RBS descriptor belongs to the broad class of shape-analysis techniques based on moments. A set of separable angular radial transformation (ART) basis functions is defined that classifies shape along various angular and radial directions. The RBS descriptor obtains 35 coefficients from the ART transform.

Finally, the work of Krishnan, S.M. and Goh, P.M.Y. (1997) is devoted to describe polyp appearance. Several parameters are evaluated, such as the response in the red channel of the image (which may indicate the presence of malignant tumors), the perimeter, the enclosed boundary area or the the *form factor*, which can give indication of possible presence of abnormalities in the colon (the more irregular the shape of the lumen, the smaller the value of the *form factor*).

Our work (Bernal, J. and Sánchez, J. and Vilariño, F. (2011)) cannot be easily assigned to a single group too, because it also starts with a basic segmentation but in this case it is based on the definition of a model of polyp appearance. This model defines a polyp as a region enclosed by intensity valleys. While we will explain it in more depth in Section 4, our results show that our model is a valid starting point that can be used in polyp detection, applicable for most types of polyp appearance.

Texture-based approaches

The use of texture descriptors on polyp detection has been gaining interest during last years. There is a number of works that are based on the use of **wavelet descriptors**. In this case the wavelet transform is calculated for each frame and the attention is put on the detail and approximation coefficients.

The work of Li et al. (2005) takes into account that, when detecting abnormalities in colonoscopic images, the location, shape and size of the abnormal regions in the image are unknown and vary across images therefore it is difficult to determine the appropriate patch-size to use for searching. In this case the solution is to use multi-size patches and ensemble them in order to achieve good performance. The features extracted from these patches are taken from both approximating and detail coefficients from wavelet decomposition of the image patches in the three channels of the CIE-Lab color space.

Also in the context of texture-based approaches we can observe the works of Karkanis et al. (2003). In these works the first operation that is done to the image is wavelet transformation, which is combined with other texture descriptors, such as co-ocurrence matrices or local binary patterns. The same group of researchers developed of a tool to detect colorectal lesions in endoscopic frame, which was named CoLD (colorectal lesions detector, Maroulis et al. (2003)). This tool provides a graphical user interface so both novice and experts user can take advantage of its use. In this case wavelets' information is used to discriminate amongst regions of normal and abnormal tissue.

There are some other texture descriptors that have been used to develop polyp detection method, such as the already mentioned local binary patterns or co-ocurrence matrices. The work of Ameling et al. (2009) combine both of them, with the novel use of local binary patters in opponent color space. As the authors state, texture can be seen as a local property and therefore, each image is divided into small image patches and four different methods were implemented, which combine co-ocurrence matrices (using different statistical measures such as energy, homogeneity or entropy) and local binary patterns. The analysis of the performance results points out that the inclusion of color characteristics (in this case, in local binary patterns) gives better results so color should be considered as an important feature for polyp detection.

As in the case of shape-based approaches, MPEG-7 also offers texture descriptors that can be used to build polyp detection methods. In the work of Coimbra & Cunha (2006), although applied to a different type of endoscopic process, several texture and color descriptors are presented. In the sub-field of color descriptors, methods such as *dominant color*, *scalable color* or *color structure* are presented (see Bernal, J. and Vilariño, F. and Sánchez, J. (2010) for a further explanation of them). Related to texture descriptors, *homogeneous texture* and *local edge histogram* are introduced. These methods are evaluated in a big database and, in order to quantify the performance of each descriptor, several measures were used such as descriptor's redundancy or the variation of the descriptors' value. The experimental results show the superiority of *scalable color* over other color descriptors due to its higher resolution. On the other hand we have the apparently strong *local edge histogram* that performs worse than other simpler approaches, such as *homogeneous texture*, since it pays too much attention to the small texture variations in the image.

Not all the texture-based methods are built on the use of a certain descriptor. The work of Tjoa et al. (Tjoa et al. (2002) and Tjoa & Krishnan (2003)) introduces the concept of texture unit (TU) and texture unit number (NTU). Texture units characterize the local texture information for a given pixel and its neighborhood, and the statistics of all the texture units over the whole image reveal the global texture aspects. Without entering into details, each pixel value is compared with the value of the pixels in its neighborhood and then the value for this pixel in the TU matrix is assigned according to the comparison. The texture information is presented in the texture spectrum histogram, which is obtained as the frequency distribution of all the texture units. Six statistical measures are used to extract new features from each texture spectrum, which include energy, mean, standard deviation, skew, kurtosis and entropy.

3.4 Quality assessment

Currently, there are several metrics for the assessment of the quality of the colonoscopy intervention, such as the insertion time and withdrawal time. For instance, current ASGE (American Society for Gastrointestinal Endoscopy) and ACG (American College of Gastroenterology) guidelines suggest that, on average, withdrawal time should last a minimum of 6 minutes. Other works propose the use of additional metrics that include the quality of preparation, among others (Morán et al. (2009)). In the case of Europe, a very good work on quality assessment in colonoscopic interventions can be found in the work of Segnan et al. (2011), which defines from how to prepare conveniently the patient to an intervention to a classification of the polyps that can be found. These metrics can be potentially used into training programs for physicians, in order to assess their skills.

3.4.1 Obtention of metrics from colonoscopy videos

Unfortunately, there is not a lot of information about what metrics could be extracted from a colonoscopy video in terms of computer vision analysis. One interesting approach can be found in the work of Hwang et al. (2005), which was extended in Oh et al. (2009). These works presents a method to measure automatically the quality metrics for colonoscopy videos, based on analysis of a digitized video file created during colonoscopy and produces information regarding insertion time or withdrawal time. The process to calculate these metrics involve some results that have been already presented in previous subsections, such as: 1) detection of non-informative frames; 2) estimation of the camera motions, performed to find a boundary between insertion and withdrawal phases; 3) segmentation of the colonoscopy video based on camera motions (proximal and distal direction), with the objective of finding the end of insertion phase (cecum); 4) lumen identification, performed to decided if an informative frame contains the colon lumen or not and 5) computation of seven metrics: insertion time, withdrawal time, clear withdrawal time and ratio, number and ratio of camera motion changes and clear operation-free withdrawal time. The authors have considered a combination of the metrics in a single quality score but the idea has been abandoned due the fact that several components affect the final quality of a colonoscopy, some patient-related (such as the preparation of the colon), some equipment-related (quality, color or contrast of the image) and some endoscopist-related.

3.4.2 Applications for training

Intelligent systems can be used to provide information oriented to build up training systems for the physicians to improve and test their skills. The work ofVilariño et al. (2009) proposes the evaluation of the skills of the trainees, and their evolution during learning processes, by using eye-tracking methodologies as a tool for the assessment of abilities such as active visual search and reaction time to the presence of polyps, among others. This study presents a novel method which compares visual search patterns between the skilled specialists and the trainees. This is done by tracking the eye position of two groups of physicians (experts and novices) while they are shown a set of colonoscopy videos. Several measures were computed by analyzing the eye-tracker results, such as eye movement speed or number of fixations. The obtained results show that colonoscopy experts and novices show a different behavior in their visual search patterns, and therefore the proposed eye-tracking based procedure can provide automatic and objective measures for their evaluation. A method similar to the one presented in Vilariño et al. (2009) can be potentially used both for assessment of the skills of the trainees during their learning process or to assess the quality of the whole procedure in intervention time. In addition, the inclusion of the models of appearance and the item categorization from the tools for scene description can provide an objective ground-truth against which to check the abilities of the trainee. This can potentially be implemented by analyzing the extent to which the trainee identifies the regions of interest.

The link between the attention models that guide the physicians visual search and the visual descriptors which discriminate between the regions of interest is an open and most paramount line of work, that can help in a twofold way: On the one hand, by providing essential information for the deepening in the understanding of the processes associated to the target abilities in colonoscopy (such as the active search of polyps or the accurate navigation throughout the colon), and on the other hand, by allowing the simulation of these processes in computer models that can be potentially used for assessment.

3.5 Development of patient-specific models

Approaches identified in the previous sections deploy the state-of-art about a novel conception regarding the role that computer-assisted technologies may potentially play in the discipline of colonoscopy. These contributions aim at increasing the quality of the intervention by adding in complementary information which is not currently at the reach of the hand of the physician. Since the described methods allow the detection, segmentation and characterization of anatomical structures, lesions and physiological behavior, there is a manifest potential to use these strategies in order to endow current techniques with architectures ready to work with patient-specific models. The patient-specific approach has been one of the main trends in clinical research lately and it has been one of the pillars of the research funding schemes for Information and Communication Technologies related to health care in Europe during the last Framework Programs (ICT Programme Commitee (2011)). The patient-specific orientation focuses on the adaptation of existing methodologies so that they can take profit of the particular information, traits, clinical details or characteristics associated to each patient. Thus, the patient-specific viewpoint aims at the focalization of the (general) outcomes provided by each technique onto the (particular) specificities of each case. The extent to which this perspective can be exploited by using intelligent systems in colonoscopy is an open field of work. Here, we expose only as a matter of example a tentative list of a few prospective ideas:

- On the one hand, the use of feature detection in colonoscopy video could provide a way to the characterization of the inner walls of the colon, based on the identification of unique traits. This could be used for the tagging or annotation of physiological features as markers, and apply this information in a further step for the identification of the exact place of the situation of region close to a polyp.
- A system storing the visual traits of the colon from a given patient could make use this information in order to find those very specific locations when a new colonoscopy intervention is performed on that patient. This could provide a method for a precise spatial localization of regions of interest. The straightforward application of this potential implementation would be oriented to the registration and study of evolution of lesions in time (or whatever other item of interest) in the sequential routine interventions carried out on a particular patient, by automatically providing the specialist with a measure of certainty about the location of those lesions.
- The generalization of this methodology could be addressed towards the definition of a patient-specific atlas of the colon, in a way in which the specialist could have track of landmark positions in intervention time. This perspective presents a scenario in which the specialist is endowed with a road map for the navigation in intervention time, allowing the specialist to address specific targets with high reliance, reduced time and a potential shrinking of the miss rates.

Finally, and since the former ideas can be intrinsically implemented at the intra- and inter-patient levels, one of the major challenges for a patient-specific approach consists of the efficient definition of the video database, together with a solid ground truth against which obtaining qualitative and quantitative performance assessments of the different methods. The next section deals with the headlines related to the building-up of such databases.

3.6 Ground-truth and database building

In order to carry out an objective assessment of a given method or system, a ground-truth must exist. The ground truth consists of set of samples from a given number of case studies, with the corresponding annotations provided by an expert or group of experts. In our context, a video annotation can be of different natures, among which we can highlight, only to mention a few: 1) a whole frame, indicating that it is that frame which contains a particular event -e.g., the first image in a sequence showing a polyp-; 2) a given region of interest (ROI) -e.g., indicating the bounding box surrounding the polyp itself-; 3) any textual information -e.g., a qualitative assessment of the clinical relevance of a polyp-, etc. These annotations are used to check the performance of a new expert or a new method against the results provided by the annotator, who is considered the reference. In the ideal case, the annotation procedure should be repeated by each expert, in order to get a intra-observer variability measure, and by different experts, in order to get a inter-observer variability measure. A good database with

a solid ground-truth is an invaluable resource and a key point for the objective assessment of different methods under a common context of evaluation.
Unfortunately, databases of annotated colonoscopy videos are scarce, and even the access to small databases is very restricted (few examples can be found at National Cancer Institute (2011)). The reason of this (without taking into account the natural motivations related to ethical and administrative issues) has to do with the generalized fact that colonoscopy video interventions are not routinely saved, since no a-posteriori analysis is needed after the intervention. In many cases, the only image saved consist of a single picture of the ileo-cecal valve, which serves as a prove of the its achievement during the phase of introduction and indicates the start of the withdrawal phase (Malik (2010)). In the computer vision bibliography, some authors proposed pilot approaches that were validated in a few frames, with no significant inference for the case of a large video. In other cases, when the number of cases was higher, the database used for the results was not available. We provide our own test set of annotated videos with polyp sequences, which is accessible to the interested researchers by direct email petition to the authors. This dataset will be described in Section 4.

3.6.1 Building up of a database

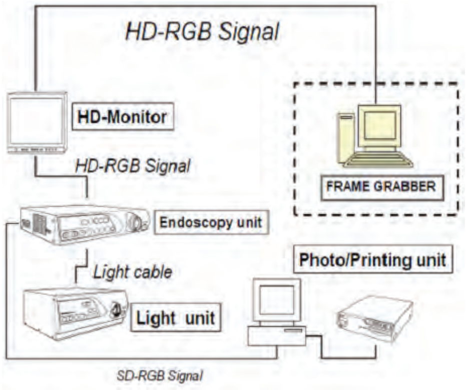
The building-up of a colonoscopy database consists of two different parts, namely: 1) The video acquisition system, and 2) the video annotation procedure. The video acquisition system must be able to grab HD frames from the colonoscopy source and store them to hard disk, with no lose of frame rate or frame quality. Although the posterior analysis of the frames must not need HD size, by storing the HD version of the video we assure the maximum image quality provided by the device. In order to capture the HD frames, a HD frame grabber must be installed into a PC which will play the role of data repository. Finally, in order to keep the frame rate and video quality, the frames must be compressed with a fast lossless compression codec, and stored with a high speed RAID 1 configuration (parallel disc write). Figure 7 a) shows a graphical representation of the set-up.

3.6.2 Annotation of colonoscopy video

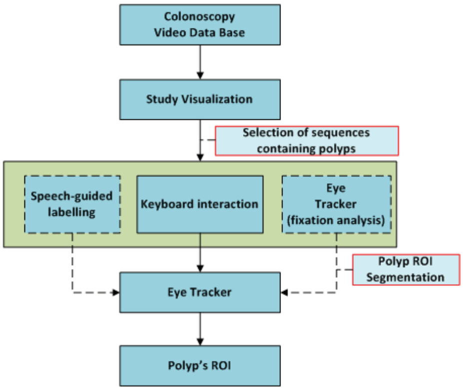
The video annotation procedure can be performed in different ways. In the case of frame annotation, a keyboard interaction can be potentially enough to select the desired frames. A navigation system must be implemented if the the expert is allowed to go forward and backwards in the video sequence. If the annotation task consist of the definition of ROIs, a mouse, a digital pen, or a tactile device can be used. More sophisticated techniques, such as the use of eye-tracking (Vilariño et al. (2007)), can be implemented in case that the video is to be annotated by using attention/perception models -see Figure 7 b) for a general scheme.

4. Polyp detection by means of a model of polyp appearance

Our current work, as it has been stated in Section 3, can be enclosed into polyp detection by means of shape descriptors. The difference between our work and the works by other authors is that in order to develop our approach we started by defining a model of polyp appearance. This model is based on how polyps are shown in colonoscopy videos, and the aim of our work is to build up our polyp detection system by taking into account as many types of polyp appearances as possible. Hence, we rely on methods that are based on the detection of the elliptical shape of flat polyps but we also consider the detection of peduncular polyps.



(a)



(b)

Fig. 7. a) Schematics of the HD colonoscopy data acquisition system. b) Data annotation scheme.

Our processing scheme consists of three stages (see Bernal, J., Sánchez, J., Vilariño, F. (2011)), namely: *region segmentation*, *region description* and *region classification*. The segmentation stage is stated as follows: Given an input image showing a polyp, our objective is to divide the image into meaningful separate regions, provided that one and only one of these regions will contain the whole polyp. The resulting regions from the segmentation stage will be then analyzed for region description. In the *region description* stage, our plan is to find the descriptor that best characterizes the region in terms of likelihood for polyp presence. The descriptions of the segmented regions will be introduced into the *region classification* stage whose aim is to decide, considering the data that receives and previous knowledge, if a region contains a polyp or not. This section focuses exclusively on the region segmentation stage, since its good performance is crucial for the consecutive phases of the method.

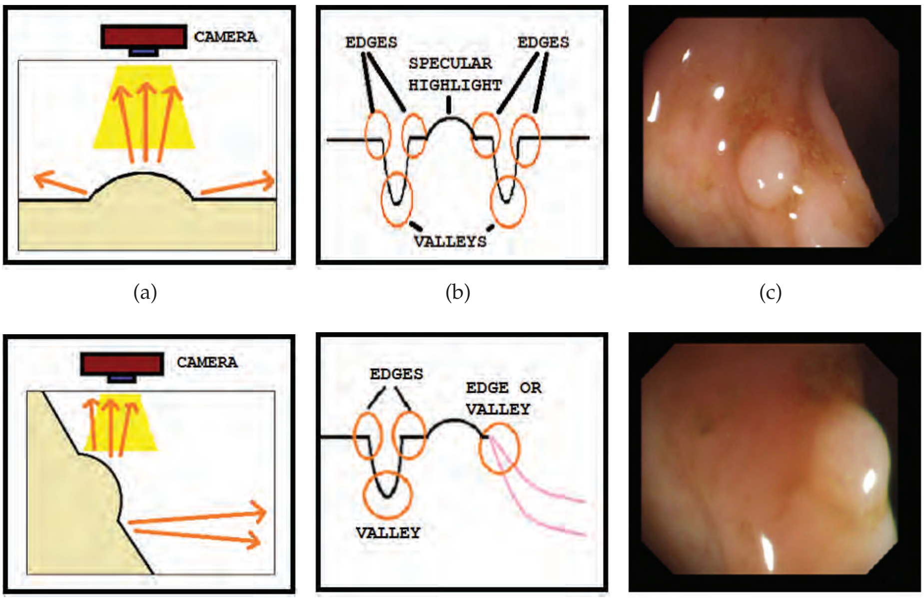
4.1 Region segmentation by means of a model of polyp appearance

As mentioned above, our model of polyp appearance is based on how polyps are shown in colonoscopy videos. More precisely, the analysis of different frames from our database leads to the conclusion that the lighting of the probe gives us hints about how the polyps appear. As light falls perpendicularly to the colon walls, it creates shadows around the surfaces, and when the light falls into a prominent surface it creates a bright spot surrounded by darker areas. Finally, the surrounding shadows define boundaries, and boundaries are identified as edges and intensity valleys in the processed images. We can see in Figure 8 a graphical example of our method and how it can be valid for both lateral and overhead views.

Although this model of prominent surface appearance under the lighting conditions of the colonoscope appears to be valid, there are some challenges to be overcome which are those previously exemplified in Figure 6, namely: non-uniform appearance of polyps, different point of view, specular highlights, image blurring and color channel misalignment, among others. Taking all this into consideration, we base our segmentation method on a model of polyp appearance where we define a polyp as *a prominent shape enclosed in a region with presence of edges and valleys around its frontiers*.

A complete example of how our *region segmentation* method works can be seen in Figure 9. Recall that the objective of this stage was: given an input image, divide it into a minimum number of informative regions. In our case, the term *informative regions* is used here for regions candidate to contain polyps, which will be classified in a later stage into polyp- vs. non-polyp. Finally, one and only one of the informative regions will contain the whole polyp. Conversely, *non-informative regions* are those assumed not to contain a polyp inside, and therefore there will be no need of further processing for polyp detection (Bernal, J., Sánchez, J., Vilariño, F. (2010))). The proposed region segmentation scheme consists of 3 different stages:

1. **Image Preprocessing:** Before applying any segmentation algorithm there are some operations that should be done: 1) converting the image into gray-scale (Figure 9 b)), 2) de-interleaving (since our images come from a high definition interleaved video source), 3) correction of the specular highlights (Figure 9 c)), and 4) inverting the grey-scale image (Figure 9 d)).
2. **Image Segmentation:** Following some of the approaches presented in Section 3, we apply watersheds as a first segmentation method. The improvement of our contribution is that instead of using the original image we use gradient information, which makes the boundaries between the regions obtained this way closer to the boundaries that separate the different structures (see Bernal, J., Sánchez, J., Vilariño, F. (2010)).
3. **Region Merging:** Since it is difficult to obtain perfect segmentation results from a basic segmentation process, the output of the previous stage will be a high number of regions



(a) (b) (c)

(d) (e) (f)

Fig. 8. a) and d): Simulation of an illuminated prominent surface. b) and e): Grey-scale profile. c) and f): Example images.

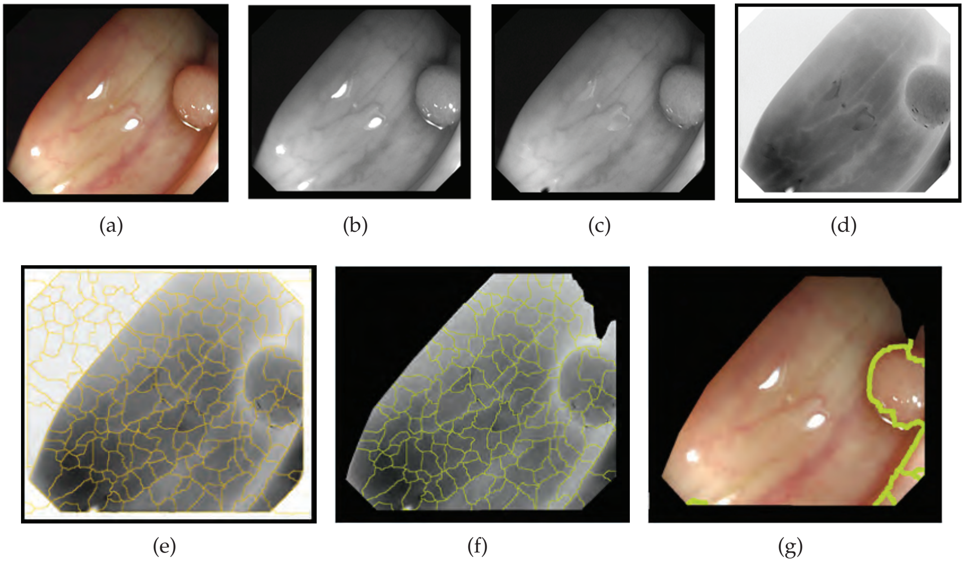


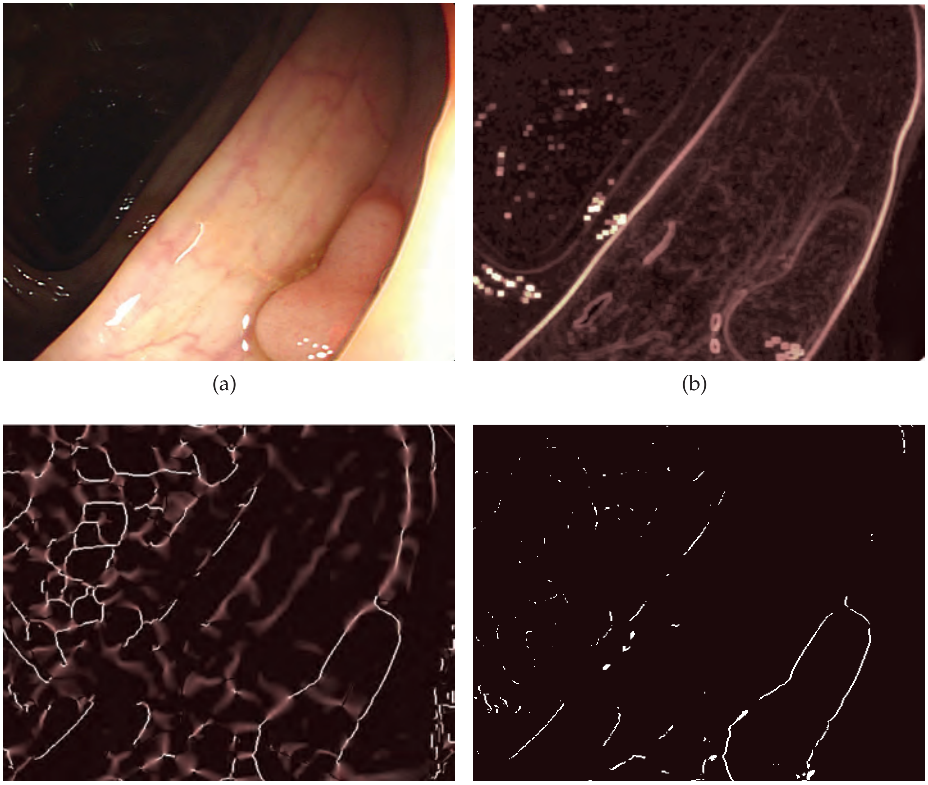
Fig. 9. Complete example of segmentation: a) Original image, b) grey-scale image, c) correction of reflections, d) complemented image, input to watershed segmentation, e) first watershed segmentation, f) segmentation results after region information-based region merging, and g) final segmentation

that should be reduced. In this case our region merging method takes into account the previously-defined model of polyp appearance to combine neighbor regions into larger ones. We apply a joint strategy: 1) Region information-based region merging, and 2) Depth of valleys-based region merging.

a) *Region information-based*: We first calculate the neighborhood map of the segmented image (Figure 9 e)) and identify the frontiers between each pair of regions. Next, we categorize the regions and frontiers in terms of the amount of information that they contain (Bernal, J., Sánchez, J., Vilariño, F. (2010)). For instance, a low information region will have a very high (or very low) mean grey level, and also a very low standard deviation of grey level. We will only merge regions: 1) with the same kind of information, and 2) separated by weak frontiers. In this case, in order to consider a frontier as weak we propose a frontier weakness measure as defined in Equation 1. This measure combines the information of the mean gradient of the frontier pixels (weighted by the coefficient α) and the strength of the frontiers (weighted by the coefficient β). The latter is measured as the percentage of frontier pixels remaining after applying two median filters of increasing order -this helps removing spurious regions created by blood vessels-. We merge and categorize regions until their number is stabilized or there are no weak frontiers left (Figure 9 f)).

$$FrontierWeakness = \alpha * gradient + \beta * strength \tag{1}$$

b) *Depth of valleys-based*: We define a *depth of valleys* measure that combines the output of a ridges and valleys detector (see López, A.M., Lumbreras, F. et al. (1999) for details) with the



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

Fig. 10. Creation of the depth of valleys image: a) Original image. b) morphological gradient image, c) valleys image, and (d) depth of valleys image.

output of the morphological gradient. This gives information about the depth of the pixels in the valley image. The pixels that constitute the frontier of the region will have high values for both the valley and gradient magnitudes (both measures will be smaller for the inner pixels, as can be seen in Figure 10). Using this information we can continue merging regions, keeping only those whose frontiers are strong in terms of *depth of valleys*. We merge regions until there are no weak frontiers according to the depth of valleys threshold or when the number of regions is stabilized (Figure 9 g)).

In order to assess the quality of our *region segmentation* method, we have created a database with different studies of polyp appearance. We were provided 15 random cases, in which the experts (physicians) annotated all the sequences showing polyps, and a random sample of 20 frames per sequence was obtained. The experts guaranteed that all these 20 frames showed a significantly different point of view within the scene by rejecting similar frames. This allows us to maximize the variability of the images used, while not jeopardizing any bias at all.
We will evaluate the performance of our method by using two different measures: Annotated Area Covered (AAC) and Dice Similarity Coefficient (DICE). Both measures are

complementary as the former calculates the amount of annotated polyp area while the latter complements it with the amount of non-polyp information that is kept in the region. We will compare our final segmentation results with the ones obtained using one state-of-the-art method such as *normalized cuts*.
We also have to consider that colonoscopy images have black borders around them. Our region segmentation method consider their presence and use the results of non-informative region identification (which borders of the image are part of) to avoid further processing of this areas. In order to test the effect that these borders have in the segmentation results, we have also created a database that eliminates the borders of the images, although it may lead to a loss of information.
In Table 1 we show results for polyp region detection, comparing the performance of our method with the performance achieved by normalized cuts, using the same number of final regions. We get better results than normalized cuts in terms of AAC and DICE. This means that our method outperforms normalized cuts by providing regions that do not separate the polyp, and the polyp is always present in only one region.
The elimination of the borders in the image, in terms of AAC, has almost no effect for our method but it has more incidence for normalized cuts. DICE results improve for both methods by using the image without borders (better as the threshold value increases), although normalized cuts results are better. But, as it can be seen in Figure 11, normalized cuts non-polyp regions tend to be larger than our non-polyp regions (in our case we know that the larger region corresponds always to the background).

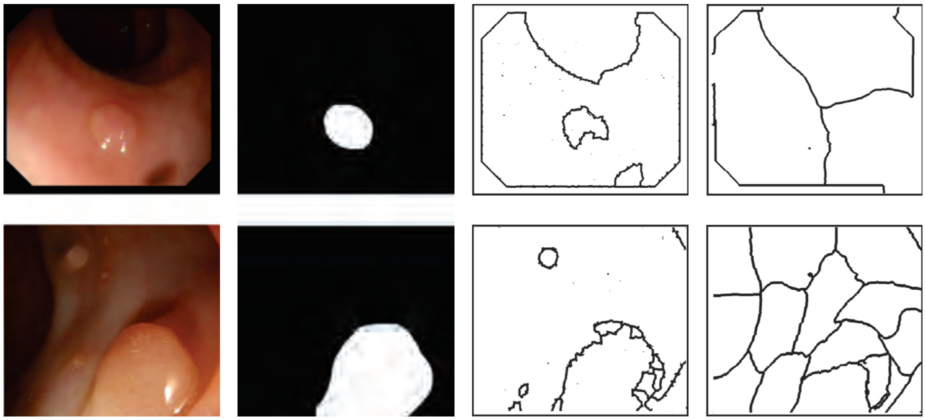
With Borders						
Measure / Method	Ours	NCuts	Ours	NCuts	Ours	NCuts
Threshold Value	0.6		0.7		0.8	
AAC	61.91%	63.66%	70.29%	69.06%	75.79%	70.86%
DICE	55.33%	44.97%	44.6%	37.75%	36.44%	34.01%
Without Borders						
Measure / Method	Ours	NCuts	Ours	NCuts	Ours	NCuts
Threshold Value	0.6		0.7		0.8	
AAC	60.71%	60.2%	70.29%	63.98%	74.32%	64.24%
DICE	55.68%	63.15%	48.01%	61.84%	45.01%	56.73%

Table 1. Comparison between the results obtained by our method and normalized cuts with respect to the value of the depth of valleys.

In Figure 11 (a-d) we can see examples of the output of each method. It can be seen that the final regions that our method provides (c) are closer the polyp mask -the ground-truth segmented by experts-.

4.2 Analysis of results

As it has been shown in the previous subsection, we have developed a region segmentation method which takes into account the model of polyp appearance. The preliminary results obtained indicate that our model can be valid for polyp segmentation, although there are some problems that need to be overcome. Although the numerical results are satisfactory in terms of accuracy, DICE results must be improved. The number of final regions should be



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

Fig. 11. Comparison of segmentation results: a) Original images, b) polyp masks, c) output of our method, and d) output of normalized cuts.

reduced to the minimum possible if we do not want to hurt, in terms of computation time, the following stages in the pipeline. One possible solution will be to start discarding regions that will not be passed to posterior stages. We have already used the term of non-informative regions (as we proposed in Bernal, J., Sánchez, J., Vilariño, F. (2010)) but we can also take profit of some of the works presented in this chapter: for instance, we can use some results for lumen segmentation in order to help our description of what is on the scene. Segmentation results could be potentially improved by applying a size threshold to discard some of the smallest regions. It would not have an impact in terms of polyp detection if the hypothesis of a minimum size for detection of polyps is considered. The DICE results could be also potentially improved by merging some small regions that appear inside the polyp region.

Our efforts now are focused on developing more the approach of depth of valleys. Our idea is that polyps appear in the image surrounded by intensity valleys and contours where the light falls perpendicularly on them. We complement the information that a valley detector provides with the morphological gradient in order to achieve a method that enhances both valley and contour information, because in certain type of views (in general, in lateral views) we do not have a whole valley surrounding the polyp. However, non-connected edges are available in these cases, and once we have this information, that can be complete or incomplete, our next steps will be focused on finding out the points which are inside or outside the polyp. To achieve this goal, we will use the depth of valleys image as the seed for an object identification algorithm. This algorithm is based on the fact that if a point is interior to an object, it will be surrounded by boundaries (in this case, high values in the depth of valleys image) in many directions. Conversely, if a point is exterior to an object, it will be surrounded only in a few directions. This concept can be better understood by following the representation shown in Figure 12.

Finally, this method needs an accurate and robust boundary detector, although these contours can be partially incomplete. This method could also be adapted to the idea of ellipse-fitting (Hwang, S. and Oh, J.H. and Tavanapong, W. et al. (2007)) in order to fit points interior to objects to those which would be interior to an ellipse. The development of a method that identifies accurately which points are interior and which are exterior to objects may lead to an efficient pruning of the regions obtained from the region segmentation stage.

5. Conclusions and prospective

As mentioned in our introduction, we believe that there is a great potential to use intelligent systems in colonoscopy. Throughout this chapter we have shown several methods that, starting from a description of what is on the scene, can be used to add in valuable information to the output of the colonoscopy intervention.

We have shown how the quality of colonoscopy videos can be improved, by eliminating some of the artifacts that can alter the performance of computer vision-based systems, such specular highlights or color channel misalignment. We have also shown that, by a general analysis of the image and the motion of the camera, it is possible to automatically annotate the different parts of the intervention for video indexation and retrieval.

There is a vast majority of methods in the literature devoted to lesions and cancer detection. Several approaches have been carried out, which involve from the direct analysis of the shapes of the objects that appear on the image to an analysis of the texture of the tissues. We argued that these methods (included ours) should be tested on a benchmark database with a solid common ground truth, which is not available yet, in order to have solid inference about their actual performance. Despite this, the contributions presented show that the particular results could be potentially used to aid the diagnosis of the physician, both during and post-intervention.

We have proposed methods that could be used to assess the physician skills in training programs. By analyzing training results we can improve our systems or, as one method presented show, we can use the physicians' reaction of what they see in colonoscopy videos to learn about what can be relevant in each colonoscopy frame. This provides a baseline to understand how the attention models are working in intervention time, and which and why the regions of interest call the physician's attention.

In the near future, we can think of colonoscopy interventions where intelligent systems add key information in real-time and allow objective metrics for the assessment of the quality

(a) (b) (c)

Fig. 12. a) Graphical example for the definition of points exterior and interior to objects. Blue lines indicate directions where the point is surrounded by a strong boundary and red lines indicate the direction where the point is not surrounded by a strong boundary. b) Original image, and c) intuitive approximation for the performance of the method in the previous image.

of the intervention. Whether it is an alert to the physician by highlighting some part of the image where the intelligent system observes some evidence of lesion or cancer presence, or by automatically storing a patient-specific report of the patient's colon, intelligent systems are called to be a relevant tool in the future of colonoscopy.

6. Acknowledgements

This work was supported in part by a research grant from Universitat Autonòma de Barcelona 471-01-3/08, by the Spanish Government through the founded project "COLON-QA" (TIN2009-10435) and by research programme Consolider Ingenio 2010: MIPRCV (CSD2007-00018).

7. References

Ameling, S., Wirth, S., Paulus, D., Lacey, G. & Vilariño, F. (2009). Texture-based polyp detection in colonoscopy, *Bildverarbeitung für die Medizin 2009* pp. 346–350.

Arnold, M., Ameling, S., Ghosh, A. & Lacey, G. (2011). Quality Improvement of Endoscopy Videos, *Proceedings of the 8th IASTED International Conference on Biomedical Engineering*, Insbruck, Austria.

Arnold, M., Ghosh, A., Lacey, G., Patchett, S. & Mulcahy, H. (2009). Indistinct frame detection in colonoscopy videos, *2009 13th International Machine Vision and Image Processing Conference*, pp. 47–52.

Bernal, J. and Vilariño, F. and Sánchez, J. (2010). Feature detectors and feature descriptors: Where we are now, *Technical Report 154*, Computer Vision Center & Computer Science Department Universitat Autònoma de Barcelona.

Bernal, J. and Sánchez, J. and Vilariño, F. (2011). A region segmentation method for colonoscopy images using a model of polyp appearance, *Lecture Notes in Computer Science. Pattern Recognition and Image Analysis. Proceedings of the 5th Iberian Conference, IbPRIA 2011*, pp. 134–142.

Bernal, J., Sánchez, J., Vilariño, F. (2010). Reduction of Pattern Search Area in Colonoscopy Images by Merging Non-Informative Regions, *Proceedings of the XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, Madrid, Spain.

Bernal, J., Sánchez, J., Vilariño, F. (2011). Current challenges on polyp detection in colonoscopy videos: From region segmentation to region classification. a pattern recognition-based approach, *Proceedings of the 2nd International Workshop on Medical Image Analysis and Description for Diagnosis Systems - MIAD 2011*, Rome, Italy.

Bevilacqua, V., Cortellino, M., Piccinni, M., Scarpa, A., Taurino, D., Mastronardi, G., Moschetta, M. & Angelelli, G. (2009). Image processing framework for virtual colonoscopy, *Emerging Intelligent Computing Technology and Applications* pp. 965–974.

Bratko, I., Mozetič, I. & Lavrač, N. (1990). KARDIO: A study in deep and qualitative knowledge for expert systems.

Cao, Y., Liu, D., Tavanapong, W., Wong, J., Oh, J. & de Groen, P. (2007). Computer-aided detection of diagnostic and therapeutic operations in colonoscopy videos, *Biomedical Engineering, IEEE Transactions on* 54(7): 1268–1279.

Coimbra, M. & Cunha, J. (2006). MPEG-7 visual descriptorsï£¡contributions for automated feature extraction in capsule endoscopy, *Circuits and Systems for Video Technology, IEEE Transactions on* 16(5): 628–637.

Dahyot, R., Vilariño, F. & Lacey, G. (2008). Improving the quality of color colonoscopy videos, *Journal on Image and Video Processing* 2008: 1–7.

Dhandra, B., Hegadi, R., Hangarge, M. & Malemath, V. (2006). Analysis of abnormality in endoscopic images using combined hsi color space and watershed segmentation, *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, Vol. 4, pp. 695–698.

Gunduz-Demir, C., Kandemir, M., Tosun, A. & Sokmensuer, C. (2010). Automatic segmentation of colon glands using object-graphs, *Medical image analysis* 14(1): 1–12.

Hassinger, J.P., Holubar, S.D. et al. (2010). Effectiveness of a Multimedia-Based Educational Intervention for Improving Colon Cancer Literacy in Screening Colonoscopy Patients, *Diseases of the Colon & Rectum* 53(9): 1301.

Hwang, S. and Oh, J.H. and Tavanapong, W. et al. (2007). Polyp detection in colonoscopy video using elliptical shape feature, *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, Vol. 2.

Hwang, S., Oh, J., Lee, J., Cao, Y., Tavanapong, W., Liu, D., Wong, J. & de Groen, P. (2005). Automatic measurement of quality metrics for colonoscopy videos, *Proceedings of the 13th annual ACM international conference on Multimedia*, pp. 912–921.

ICT Programme Commitee (2011). Fp7. ict - information and communication technologies. work programme 2011-12, Community Research and Development Information Service.

Kang, J. and Doraiswami, R. (2003). Real-time image processing system for endoscopic applications, *IEEE Canadian Conference on Electrical and Computer Engineering*, Vol. 3, pp. 1469 – 1472 vol.3.

Karkanis, S., Iakovidis, D., Maroulis, D., Karras, D. & Tzivras, M. (2003). Computer-aided tumor detection in endoscopic video using color wavelet features, *Information Technology in Biomedicine, IEEE Transactions on* 7(3): 141–152.

Krishnan, S., Yang, X., Chan, K., Kumar, S. & Goh, P. (1998). Intestinal abnormality detection from endoscopic images, *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE*, Vol. 2, pp. 895–898.

Krishnan, S.M. and Goh, P.M.Y. (1997). Quantitative parametrization of colonoscopic images by applying fuzzy technique, *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, Vol. 3, pp. 1121 –1123 vol.3.

Li, P., Chan, K. & Krishnan, S. (2005). Learning a multi-size patch-based hybrid kernel machine ensemble for abnormal region detection in colonoscopic images.

Liu, D., Cao, Y., Kim, K., Stanek, S., Doungratanaex-Chai, B., Lin, K., Tavanapong, W., Wong, J., Oh, J. & de Groen, P. (2007). Arthemis: Annotation software in an integrated capturing and analysis system for colonoscopy, *Computer methods and programs in biomedicine* 88(2): 152–163.

López, A.M., Lumbreras, F. et al. (1999). Evaluation of methods for ridge and valley detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(4): 327–335.

Malik, A. (2010). *End of insertion detection in colonoscopy videos*, PhD thesis, University of North Texas.

Maroulis, D., Iakovidis, D., Karkanis, S. & Karras, D. (2003). CoLD: a versatile detection system for colorectal lesions in endoscopy video-frames, *Computer Methods and Programs in Biomedicine* 70(2): 151–166.

Morán, S., Torrella, E., Esteban, D., Baños, M., García, A., Ono, A., Pérez, C., Parra, P., Cruzado, Q., Pérez, R. et al. (2009). Colonoscopy quality assessment., *Revista española de enfermedades digestivas: Órgano oficial de la Sociedad Española de Patología Digestiva* 101(2): 107.

National Cancer Institute (2011). Image Archive Resources. URL: *http://tinyurl.com/63n4qo4*

Nunes, C., Mendonça, T., Amorim, P., Ferreira, D. & Antunes, L. (2005). Comparison of Neural Networks, Fuzzy and Stochastic Prediction Models for return of consciousness after general anesthesia, *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, pp. 4827–4832.

Oh, J., Hwang, S., Cao, Y., Tavanapong, W., Liu, D., Wong, J. & de Groen, P. (2009). Measuring objective quality of colonoscopy, *Biomedical Engineering, IEEE Transactions on* 56(9): 2190–2196.

Oh, J., Hwang, S., Tavanapong, W., de Groen, P. & Wong, J. (2003). Blurry-frame detection and shot segmentation in colonoscopy videos, *Proceedings of SPIE*, Vol. 5307, p. 531.

Segnan, N., Patnick, J. & von Karsa, L. (2011). *European guidelines for quality assurance in colorectal cancer screening and diagnosis*, Luxembourg: Publications Office of the European Union.

Tian, H., Srikanthan, T. & Vijayan Asari, K. (2001). Automatic segmentation algorithm for the extraction of lumen region and boundary from endoscopic images, *Medical and Biological Engineering and Computing* 39(1): 8–14.

Tjoa, M. & Krishnan, S. (2003). Feature extraction for the analysis of colon status from the endoscopic images, *BioMedical Engineering OnLine* 2(9): 1–17.

Tjoa, M., Krishnan, S. & Doraiswami, R. (2002). Automated diagnosis for segmentation of colonoscopic images using chromatic features, *Electrical and Computer Engineering, 2002. IEEE CCECE 2002. Canadian Conference on*, Vol. 2, pp. 1177–1180.

Tresca, A. (2010). The Stages of Colon and Rectal Cancer, *New York Times (About.com)* p. 1. URL: *http://tinyurl.com/3y4acut*

van Wijk, C., van Ravesteijn, V., Vos, F. & van Vliet, L. (2010). Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow, *Medical Imaging, IEEE Transactions on* 29(3): 688–698.

Vilariño, F. (2006). *A Machine Learning Approach for Intestinal Motility Assessment with Capsule Endoscopy*, PhD thesis, Universitat Autònoma de Barcelona and Computer Vision Center.

Vilariño, F., Ameling, S., Lacey, G., Ghosh, A., Patchett, S. & Mulcahy, H. (2009). Eye tracking search patterns in expert and trainee colonoscopoist: A novel method of assessing endoscopic competency?, *Book of abstracts from the Digestive Disease Week*, Chicago, America.

Vilariño, F., Lacey, G., Zhou, J., Mulcahy, H. & Patchett, S. (2007). Automatic labeling of colonoscopy video for cancer detection, *Pattern Recognition and Image Analysis* pp. 290–297.

Vincent, L. , Soille, P. (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations, *IEEE transactions on pattern analysis and machine intelligence* 13(6): 583–598.

Viswanath, S., Palumbo, D., Chappelow, J., Patel, P., Bloch, B., Rofsky, N., Lenkinski, R., Genega, E. & Madabhushi, A. (2011). Empirical evaluation of bias field correction algorithms for computer-aided detection of prostate cancer on T2w MRI (Proceedings Paper).

Wei, J., Chan, H., Zhou, C., Wu, Y., Sahiner, B., Hadjiiski, L., Roubidoux, M. & Helvie, M. (2011). Computer-aided detection of breast masses: Four-view strategy for screening mammography, *Medical Physics* 38: 1867.

Zhang, D., Zhao, J., Lu, L., Li, L. & Wang, Z. (2010). Virtual eversion and rotation of colon based on outer surface centerline, *Medical physics* 37: 5518.

Zhu, H., Fan, Y. & Liang, Z. (2010). Improved Curvature Estimation for Shape Analysis in Computer-Aided Detection of Colonic Polyps, *Beijing, China* p. 19.