

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# **Building up the Future of Colonoscopy – A Synergy between Clinicians and Computer Scientists**

---

Jorge Bernal, F. Javier Sánchez, Cristina Rodríguez de Miguel and Gloria Fernández-Esparrach

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/61012>

---

## **Abstract**

Recent advances in endoscopic technology have generated an increasing interest in strengthening the collaboration between clinicians and computers scientist to develop intelligent systems that can provide additional information to clinicians in the different stages of an intervention. The objective of this chapter is to identify clinical drawbacks of colonoscopy in order to define potential areas of collaboration. Once areas are defined, we present the challenges that colonoscopy images present in order computational methods to provide with meaningful output, including those related to image formation and acquisition, as they are proven to have an impact in the performance of an intelligent system. Finally, we also propose how to define validation frameworks in order to assess the performance of a given method, making an special emphasis on how databases should be created and annotated and which metrics should be used to evaluate systems correctly.

**Keywords:** Intelligent systems, Image properties, Validation, Clinical drawbacks, Endoluminal scene description

---

## **1. Introduction**

### **1.1. Motivation**

During the last few years there has been an increasing effort in exploring the use of intelligent systems to assist and provide additional information to clinicians in the different stages of an intervention. In this context, we can find in the literature systems aiming at assisting the clinician in in-vivo diagnosis such as KARDIO proposed in [1], which can automatically analyze electrocardiograms, or methods that provide with data to help in the detection and

diagnosis of breast [2] or prostate cancer [3]. The spread use of Computed Tomography has elicited a new set of methods that help clinicians in intervention planning as exposed in [4]. For instance, we can find systems which allow clinicians to follow the fastest and safest way to target a pulmonary lesion [5], perform laparoscopic surgery [6] or systems such as [7] in the domain of transcatheter aortic valve implantations. However, there is scarce experience with intelligent systems applied to endoscopy where there are only a few methods such as the works presented in [8] in the context of colonoscopy quality assessment which analyzes how clinical procedures have been performed to provide quality scores.

Endoscopic technology has rapidly evolved in the last decade and current equipment allows clinicians to observe the whole endoluminal scene in high definition and, moreover, makes it possible to get different views of the same scene for further analysis by applying automatic techniques of chromoendoscopy [9] as narrow band imaging (NBI) –proposed in [10]–, the Fujinon Intelligent Chromo-Endoscopy (FICE) presented in [11] or Pentax I-scan, which was published in [12]. These advances in endoscopy imaging have generated an increasing interest in strengthening partnerships between clinicians and computer scientists to build applications that can solve some of the challenges that colonoscopy procedures still present nowadays.

It is clear that this potential collaboration between these two domains of knowledge needs from each part to acknowledge the challenges that the analysis of colonoscopy images present related to their area of expertise. Related to this, clinicians need to identify which of the existing drawbacks could be mitigated with the aid of image processing tools and computer scientists must define clearly what can be achieved by means of image processing to provide clinicians with feasible and clinically applicable solutions. Endoscopy imaging analysis present some challenges that are not limited to the ones that the characterization of anatomical structures for detection or diagnosis purposes present; aspects that are rarely covered by existing methods such as image acquisition and formation should be considered as they are proven to have an impact on the output of a given method [13].

Considering this, the focus of this chapter is to present new advances on computer vision methods for colonoscopy and to identify potential clinical issues that may be solved with the aid of computer vision. As it can be observed, this chapter is not written from either a pure clinical or technical point of view but as a way to couple the necessities and challenges of each of the domains in order to build up feasible and clinically applicable systems.

## **2. Introduction to colonoscopy challenges**

### **2.1. A brief history of endoscopy**

The history of endoscopy, as stated in [14], starts in 1805 with P. Bozzini and his attempts to construct a cystoscope (See Table 1). Although this first endoscope was considered as having failed, the principles incorporated in its design - a light source, a reflective surface (lens) and a series of specula (mirrors)- are the basis of current endoscopes. The technical challenges posed since then have been overcome with the collaboration of physicians, engineers, scientists

and optical experts among others. The progress has been slow but constant and initially rigid instruments have been changed by flexible endoscopes; candles and lamps have been replaced by electric filaments and, for vision, single lenses have been supplanted by optic fibers.

Year	Authorship	Development
1805	Philipp Bozzini (Physician)	Design of the first endoscope (Lichtleiter). Illumination is provided by candles.
1825	Pierre Solomon Ségalas (Urologist)	Design of an urethro-cystic speculum that incorporates mirrors for projecting light along the tube.
1827	John D. Fisher (Physician)	Development of a cystoscope. His principal innovation is the inclusion of a double convex lens to amplify the image.
1853	Antonin Jean Desormeaux (Urologist)	Demonstration of the first functional endoscope (cystoscope). Candles are replaced by mixture of alcohol and turpentine for illumination.
1865	Francis Richard Cruise (Urologist)	Improvement of the illumination using camphor and petrol and redesigns the lens and lamp system.
1867	Julius Brück (Dentist)	Design of an unusual instrument that uses a lamp lit by electric current.
1868	Adolf Kussmaul (Surgeon)	Attempt at the creation of the first gastroscopy using a rigid instrument based upon sword swallowers.
1870	Gustav Trouvé (Engineer)	Construction of the first electrical endoscopic instrument with optical system: the polyscope (mostly for laryngeal observations).
1877	Max Nitze (Urologist) Fritz Leiter (Manufacturer)	Development the first effective rigid endoscope that incorporates an optical system and an incandescent platinum wire lamp at the end of the cystoscope.
1880	David Newman (Surgeon)	Incorporation of the Edison incandescent lamp into a cystoscope.
1881	Johann Von Mikulicz (Surgeon) Fritz Leiter (Manufacturer)	Development the first practical and functional esophagoscope.
1894	Howard A. Kelly (Gynecologist)	Introduction the first long (30 cm) rigid rectosigmoidoscope.
1911	Michael Hoffmann (Physician)	Proposal of a solution to the problem of bending light using multiple prisms and lenses and applies this concept to gastroscopy. This is the first attempt to construct a flexible gastroscope.
1911	Hans Elsner (Physician)	Construction of the first rigid gastroscope.
1922	Rudolf Schindler (Gastroenterologist)	Building of the second rigid gastroscope
1930	Heinrich Lamm (Medical student)	Images are successfully transmitted through glass fibers.

Year	Authorship	Development
1932	Rudolf Schindler (Gastroenterologist) Georg Wolf (Manufacturer)	Development of the first semiflexible gastroscope. Schindler is considered the founder of modern endoscopy.
1940	Cameron Surgical Co.	The first flexible gastroscope is made in the USA: the <i>Cameron Schindler Endoscope</i> .
1948	Edward B. Benedict (Surgeon)	Development of the operating gastroscope by incorporating both a biopsy forceps and a suction tube within the gastroscope itself.
1948	Harry Segal (Physician) James Watson (Physician)	Production of a viable endoscopic photographic system.
1952	Tatsuno Uji (Engineer)	Design of a miniature gastrocamera that can be introduced into the stomach.
1957	Basil Hirschowitz (Gastroenterologist) Larry Curtiss (Physicist)	Introduction of the first fiber optic gastroscope.
1960	Machida Endoscope Co. Olympus Optical Co.	Development of the first prototypes of flexible colonoscopes.
1971	William I. Wolff (Surgeon) Hiromi Shinya (Surgeon)	Performance of the first polypectomy with a wire loop snare.
1975	Masahiro Tada (Gastroenterologist)	Description of the first magnifying colonoscope.
1983	Welch Allyn Inc.	Development of an electronic sensor or charge coupled device that is inserted at the tip of the endoscope.
2002	Olympus Co.	HD endoscopes

**Table 1.** Evolution of endoscopy as a result of collaboration of different disciplines

Shortly after having successfully traversed the esophagus and reached the stomach, the assessment of the duodenum, small intestine and colon were the next steps that were progressively addressed and achieved. Other needs were also identified and solved: first, the evolution from diagnostic to operating endoscopes that allowed obtaining biopsies; second, the need of preserving the image of the lesion which was observed. The latter not only reflected clinical needs but also documentation and educational requirements. At that point, several corporations became involved in the development of endoscopic instrumentation and they also designed cameras specifically for endoscopic usage.

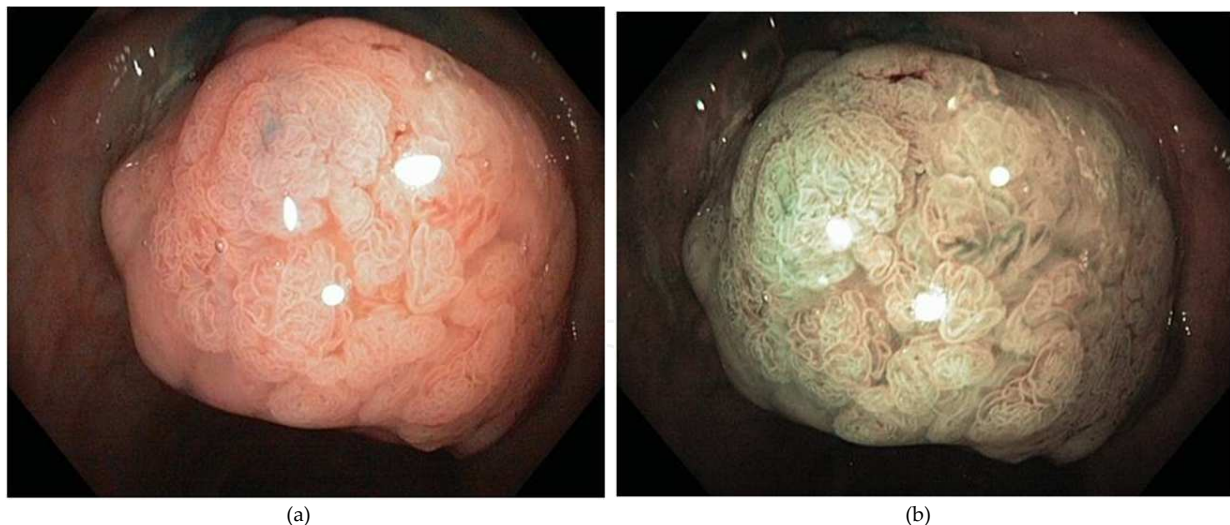
Once the fiber optic endoscope was established as a reality by late 1960s, numerous design modifications were performed with the collaboration of physicians in order to augment the utility of the device and increase its resolution. The decade of 1970 witnessed a series of rapid technological advances where a number of instrumental manufactures including ACMI, Olympus Optical Company and Machida Endoscope Company included a variety of innova-

tions (length, flexibility, channel size...) that improved the performance of the instrument. In 1983 video endoscopy was introduced as the logical consequence of technical advances in microelectronics and all current endoscopes are based on this technology. Video endoscopy allows an easy exploration, instant image acquisition and further storage confirming its utility not only for clinical practice but also for educational purposes.

## 2.2. High definition endoscopy (The quality of image matters)

In the last years, most of the developments in endoscopy have been focused on improving the quality of images, as it is the case of high definition (HD) endoscopes that use a 1080-line television and a high resolution charge coupled device with up to 1.3 million pixels. This allows the acquisition and storage of images with double the resolution of normal television. Other capabilities available in some endoscopes are the following:

- Wide angle: the endoscope has a field of vision of 170° (30% more than the conventional model) that is supposed to improve the detection of lesions hidden behind the folds;
- Electronic zoom: that achieves a  $\times 80$ –100 maximum effect;
- Narrow band imaging (NBI): a modification in the light beam enhances visualization of the network of the mucosa providing contrast and acting as a substitute of chromoendoscopy. This system offers the possibility to switch from conventional white light to blue NBI light alternatively (see Fig.1).



**Figure 1.** Example of a same polyp observed with white light (a) and NBI (b).

HD endoscopes (particularly those with magnification function) facilitate the demonstration of the mucosal architectural and vascular patterns that are altered in dysplastic lesions as it can be observed in Fig.2. With regards to the detection rate of lesions, although it is logical to assume that a higher resolution endoscope could provide better results, the results of several studies [15, 16] do not support this hypothesis.





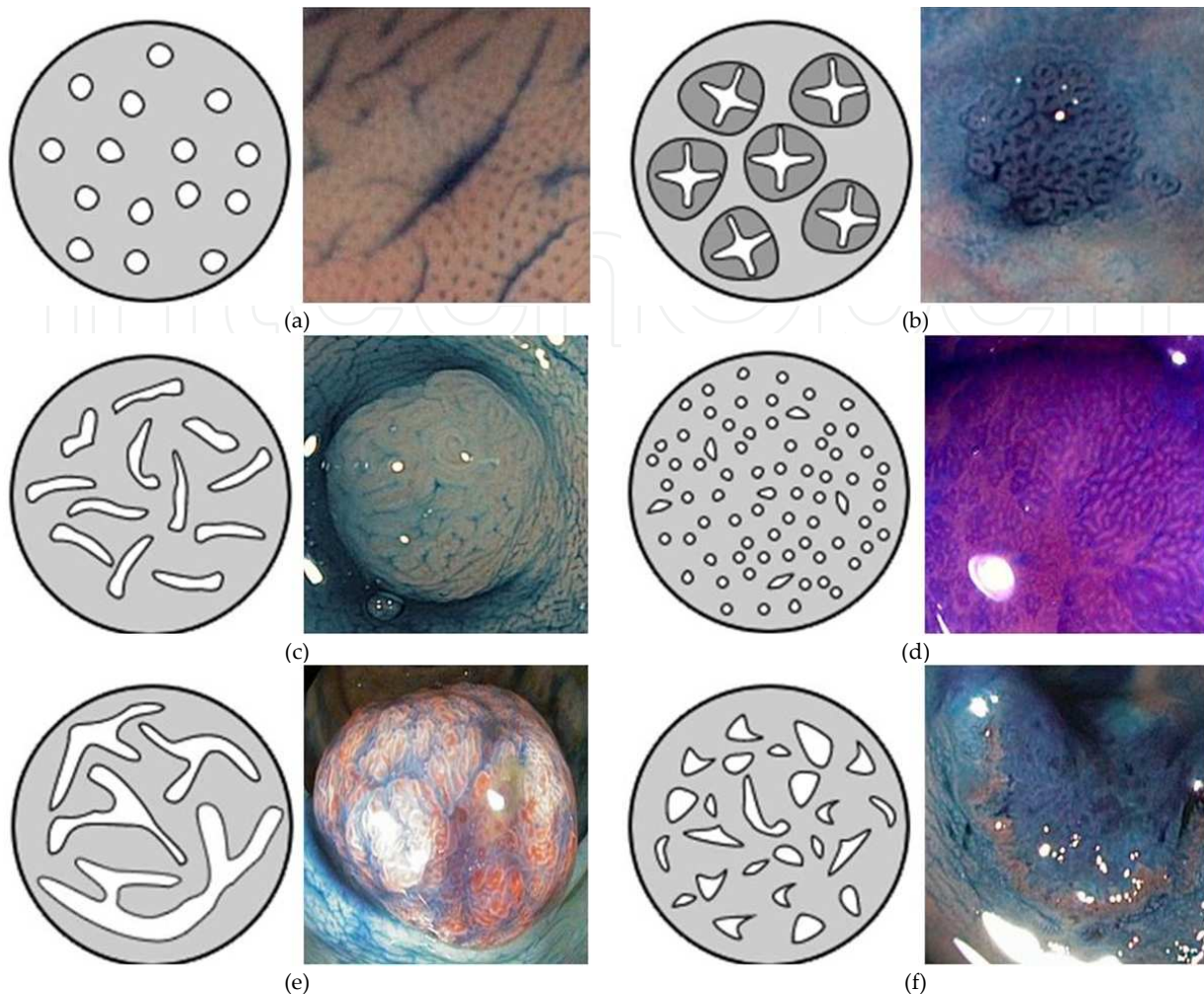
**Figure 2.** Example of a colonoscopy frame observed with conventional endoscope (a) and with high definition endoscope (b).

### 2.3. The problem of colonic polyps

Colorectal cancer (CRC) is a serious health problem in the general population and it is considered that at least two thirds of CRC develop through the adenoma–carcinoma pathway. Consequently, screening with colonoscopy for CRC and its precursor lesion has become an increasingly practice, as shown in [17]. Several actions have been proposed to optimize colonoscopy such as ensuring colon perfect preparation and carrying out a thorough examination of the mucosa which would imply a longer withdrawal inspection time, as indicated in [18].

However, colonoscopy still presents some drawbacks being the most relevant the polyp miss-rate -reported to be as high as 22%- resulting in a lack of total effectiveness [19]. The rate of polyps missed increases significantly in smaller sized polyps (2% for adenomas  $\geq 10$  mm versus 26% for adenomas  $< 5$  mm) and this has a clinical impact, not only because the prevalence of high-grade dysplasia increases with the size as exposed in [20] but because of the risk of having an interval cancer. Interval colorectal cancers are described as cancers occurring after a negative screening test or examination and they are an important indicator of the quality and effectiveness of CRC screening and surveillance, as stated in [21].

The diagnosis of dysplasia has practical consequences on the management of polyps. There is general consensus on removing all polyps detected during colonoscopy but size is a limiting factor for endoscopic polypectomy. Therefore, having a histological diagnostic of presumption is very useful in order to make the decision of performing or not a polypectomy. In this regard, there are several classifications (NICE, Kudo...) that predict the histology of the lesion based on the characteristics of the image. Kudo [22] proposes a gross classification of pit patterns into 7 types: type I and II pit patterns are characteristic of non-neoplastic lesions such as normal mucosa or hyperplastic polyps whereas pattern types IIIS, IIIL, IV, and a subset of VI are intramucosal neoplastic lesions such as adenoma or intramucosal carcinoma and lesions with a type VN pattern and a subset of type VI suggest deep invasive carcinoma (see Fig. 3).



**Figure 3.** Examples of Kudo neoplastic lesion classification: (a) Type I; (b) Type II; (c) Type IIIa; (d) Type IIIb; (e) Type IV and (f) Type V.

As this classification applies for magnification endoscopy, when it is used with conventional endoscopy the results are worse. Contrarily, NICE is an international classification of colorectal tumors on the basis of NBI observation either with or without use of a magnifying endoscope [23]. NICE is a simple categorical classification defining three different types based on three characteristics: (i) lesion color; (ii) micro vascular architecture; and (iii) surface pattern. Type 1 is considered an index for hyperplastic lesions, type 2 an index for adenoma or mucosal/submucosal scanty invasive carcinoma, and type 3 an index for deeply submucosal-invasive carcinoma. The problem with these classifications is that diagnosis derives from a subjective visual analysis and requires specific training and a high degree of experience.

Finally, the precise location of the polyps is another meaningful drawback of colonoscopy, not only when planning a surgery but also during successive colonoscopies. This limitation is especially remarkable in the presence of several polyps. In this case, an exhaustive analysis of the surface and boundaries of the polyp could be very helpful.



## 2.4. Identification of potential collaborative research areas between clinicians and computer scientists

Considering the mentioned drawbacks of colonoscopy, three potential areas in which computer science may play a role have been identified:

- Automatic polyp detection and localization: one of the exposed drawbacks is related to the difficulty on detecting certain types of polyps such as small or flat lesions. Flat polyps can be detected with the support of CT [24, 25] although its detection supposes additional patient radiation and is limited by the size. Detection of small polyps cannot be undertaken with the help of CT as the current available resolution makes it impossible to detect polyps with size smaller than 10 mm as stated in [26], therefore the diagnosis in these cases should only rely on endoscopic exploration.
- Polyp classification: the decision of performing polypectomy is commonly taken by an estimation of the size and histology of the detected lesion. This estimation is commonly made by means of visual observation and therefore incorporates some degree of subjectivity. In this context, a system that can objectively provide an estimation of the size and classification of the polyp could allow taking in-vivo diagnostic decisions and this would optimize the treatment timing.
- Patients lesion follow-up and endoscopy navigation: there is a necessity expressed by some clinicians regarding the recognition of the area that a lesion occupies, which can be useful for two different reasons: 1) for the case of polyps that have not been removed, an univocal recognition of the lesion would allow the study of the evolution of the lesion; 2) an accurate recognition of the marks that clinicians leave to identify the area of the polyp once it is removed would allow the exploration of areas nearby the lesion to search for new pathologies.

## 3. Image processing challenges for the analysis of colonoscopy videos

In order to provide clinicians with meaningful applications, the content of colonoscopy videos and frames must be thoroughly analyzed by computer scientists to search for lesions or indicators defined by clinicians. In this context, the majority of the literature has been focused on developing methods to characterize accurately the different elements of the endoluminal scene, paying special attention to polyps. Although it is clear that anatomical landmarks recognition is essential for application development, the acquisition and generation of high quality images is also crucial for computer vision methods in order to work as they are intended. For instance, the presence of image artifacts has been proven to have an impact in the performance of polyp localization methods, as shown in [13].

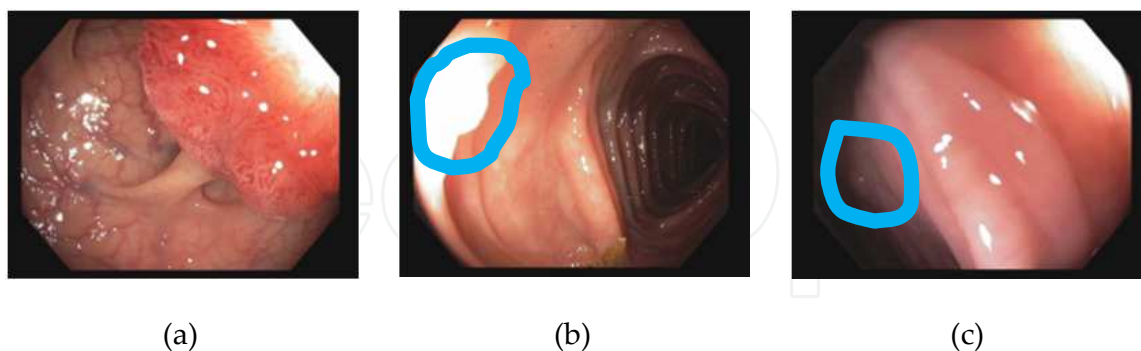
Considering this we present in this section a summary of the most important challenges that a given computer vision method must face in order to provide with efficient support to clinicians. We have divided the challenges in two groups: those related to image acquisition

and formation and those related to the characterization of anatomical structures needed to build up the clinicians' support system.

### 3.1. Identification of endoscopy image particularities with impact in image processing analysis

Videos that endoscopes generate are created following common television standards in a way such they can provide with sufficiently moving image quality while allowing for efficient resource management in case endoscopy images and videos are stored for later inspection. It is important to mention that quality in this case is understood under human's observer point of view but not under computer visions; for instance there are some image processing techniques automatically performed – i.e. sharpening - that may improve how images are observed but, as they modify the original image, they create new elements that affect an automatic analysis by means of computer vision methods. Some of the features that can affect the performance of a computer vision method are listed below and in table 2:

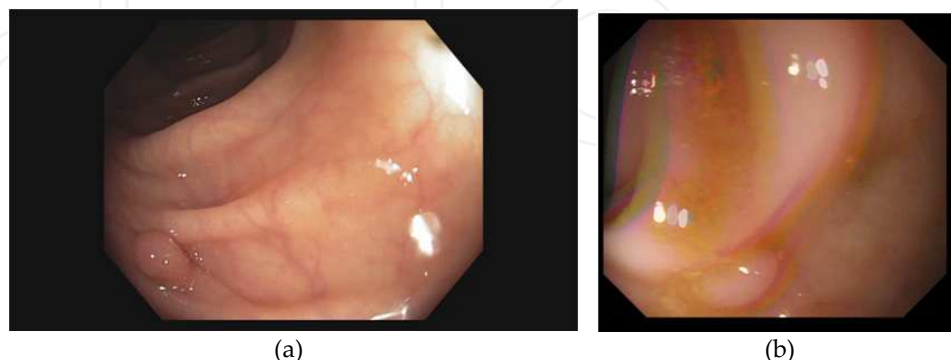
- *Illumination effects:* The way colonoscope illuminates the scene produces an axial illumination which tends to generate *specular highlights* on shiny surfaces such as the mucosa. Mucosa is covered by a thin watery film which generates many specular highlights when it is illuminated in a perpendicular direction to its surface. Specular highlights position will vary with little movements of the colonoscope which will change the angle at which mucosa is illuminated therefore areas of the mucosa affected by specularities will change rapidly. The presence of specular highlights difficult strongly image processing [13] as they appear as very prominent structures which also hinder color and texture information about the surfaces in which they appear. Moreover, axial illumination introduces also an additional side-effect regarding its *lack of uniformity* in the way structures are illuminated: structures closer to the endoscope will appear brighter than others far from the endoscope (see Fig. 4).



**Figure 4.** Examples of illumination effects: (a) specular highlights (b) overexposed polyp and (c) underexposed polyp. Polyps in images b and c are delimited with a blue mask to ease visualization.

- *Sensor acquisition effects:* *Color phantoms* appear due to temporal misalignment of color channels related to some endoscopes that still use monochrome sensors. In this case, color information is generated by illuminating the scene with the three primary colors (red, green and blue) successively. Consequently, three different images are needed to generate a color

image. This process introduces some undesired side-effects associated to camera movement: as we acquire the images in different time instants, specular highlights generated by the light source in each of the three moments will be located in slightly different positions, causing instability in the final color image –Fig. 5(a).

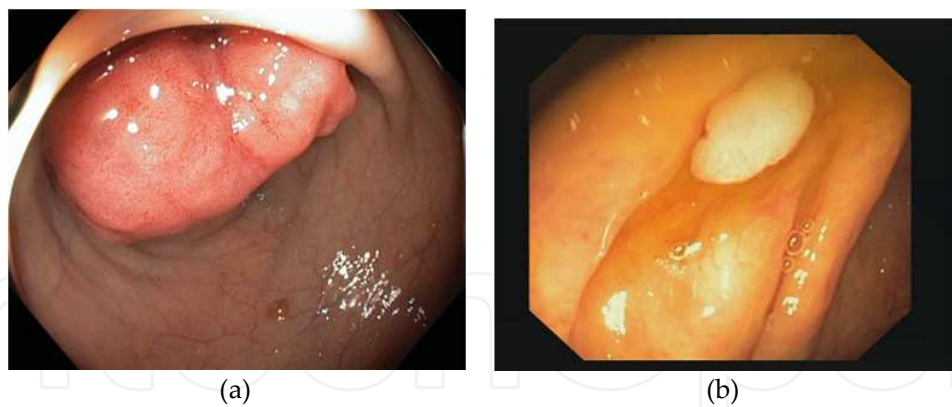


**Figure 5.** Effect of channel misalignment due to monochrome sensors: instability in specular highlights position (a) and apparition of color phantoms (b).

Moreover, as each color channel is acquired in different times, the three components (red, green and blue) will not be exactly aligned if the endoscope moves when the image is acquired. This lack of color channel alignment generates artificial color bands in the contours of the structures –Fig. 5(b) - that appear in the image which limits the performance of any color information-based structure characterization method.

- *Image resolution:* Commercial endoscopes generate videos in formats following television standards (PAL for Europe, NTSC for America and Japan). These formats are meant to generate motion images with enough quality to be observed by the general public but also minimizing the size of the information to be transmitted. By acting this way, videos generated by commercial endoscopes can be played in any standard system (TV, personal computers) without needing format conversion. Moreover, the minimization of the amount of transmitted information allows a reduction of the storage needs which is crucial in clinical settings where the amount of resources dedicated to information storage must be efficiently distributed.

Although the use of standard formats presents clear advantages for visualization and storage purposes, it does not benefit image processing by means of computer vision. Video standards offer images with lower resolution than the one that can be achieved by means of commercial cameras. For instance, NTSC standard provides as output 0.3 Megapixels images, HD standard offers images up to 2 Megapixels and a commercial camera easily exceeds 10 Megapixels [27]. Low resolution images lead to a loss of texture information associated to anatomical structures in the endoluminal scene, which can have an impact on the output of structure classification methods –Fig. 6–.



**Figure 6.** Different colonoscopy images acquired at different resolutions: (a) high resolution image and (b) low resolution image. We can observe greater texture details in the polyp from the highest resolution image.

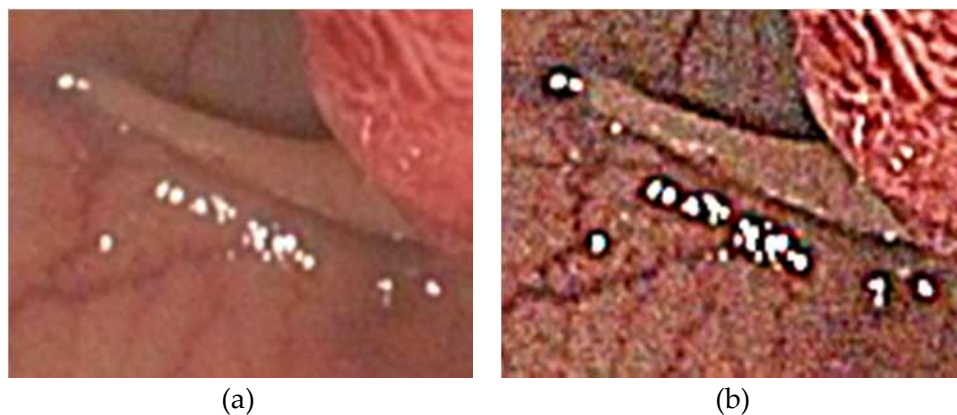
- *Image interlacing:* As it has been mentioned before, from all available video standards those with lowest bandwidth –amount of information that needs to be transmitted–requirements are chosen for use in endoscopy. This reduction in bandwidth is achieved by interlacing image lines, which is performed by acquiring odd and even image lines in different time instants. By this we can double the image refresh rate without increasing the size of the information. This also makes video movement appear smoother and more continuous to the human eye but it has a counterpart that affects posterior image processing. The final image provided by the processor will be a mixture of two different images captured in different time instants: even lines will be from the first capture whereas odd lines will come from the second. As with color channel misalignment, interlacing impact will depend on the amount of endoscope movement between the two acquisitions. For instance, if camera moves horizontally we can observe sawtooth profiles in vertical contours, apart from change of position of specular highlights. We show in Fig. 7 a clear example on how interlacing can affect the quality of the image to be processed by, for instance, the apparition of double and shadowy contours surrounding the elements of the image.



**Figure 7.** Impact of interlacing in image quality: (a) Interlaced image and (b) Separate field of an interlaced image.

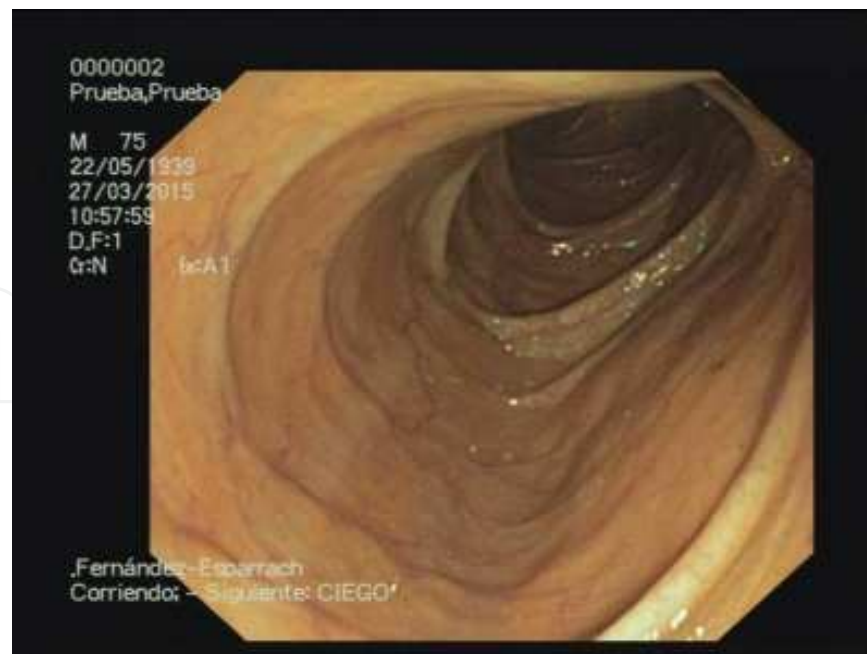


- *Sharpening*: Endoscopes and video processors include functionalities that improve the quality of the image to be visualized by human observers, aiming to simplify the observation of particular structures in the images. One of the most common techniques is sharpening, which describes a subjective perception of sharpness related to edge contrast in an image. By applying this technique, contours that separate different objects in the image can be more clearly identified and consequently structures can be easily separated –Fig. 8 (b)–. This visualization enhancement [28] comes at a cost in terms of image processing as contour enhancement implies a modification of the original image which increases image noise. Sharpening also generate halos around structures that appear in the image such as specular highlights, as observed in Fig. 8 (b).

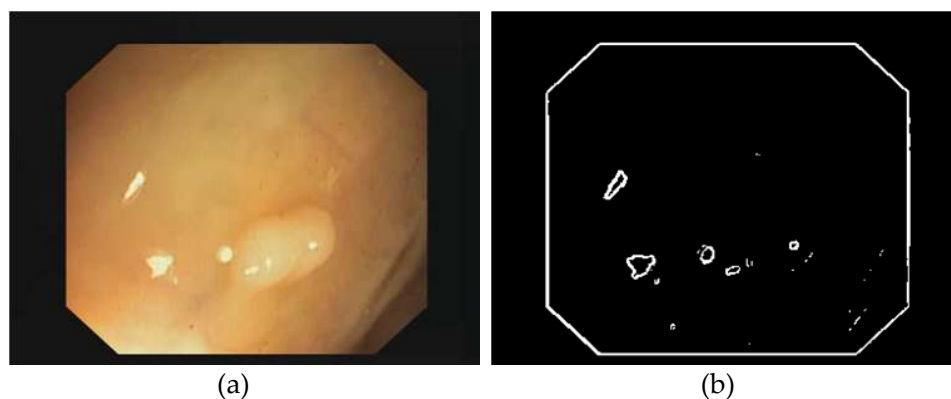


**Figure 8.** Examples of sharpening applied on colonoscopy images: (a) Original image and (b) image with sharpening applied.

- *Information overlay*: Video processors associated to endoscope do not present a specific output dedicated to its connection to a personal computer. Considering this, the image that the clinician is observing will be the same that will be stored for later processing. It is common that some information regarding the procedure such as patient information or procedure date is superimposed to the image provided by the colonoscope, as it can be observed in Fig. 9. The presence of this information precludes its use for research purposes, as this data should be anonymized. Moreover the presence of this information superimposed to the original image may difficult the observation and characterization of structures in the images apart from introducing additional noise and elements (letters, numbers) to the image.
- *Black mask*: Endoscopes automatically add an octagonal or circular black mask surrounding the image acquired by the sensor. This mask covers those regions of the image that are strongly affected by geometric distortions introduced by wide angle optic used in endoscopes. These distortions, similar to fisheye effects present in some cameras, makes structures below the mask appear different to what they are in reality and consequently they should not be analyzed by clinicians. Unfortunately the presence of this black mask affects the performance of image processing methods, as the mask creates strong contours in the separation between the mask and the endoluminal scene, as it can be observed in Fig. 10.



**Figure 9.** Examples of information overlay in colonoscopy images.



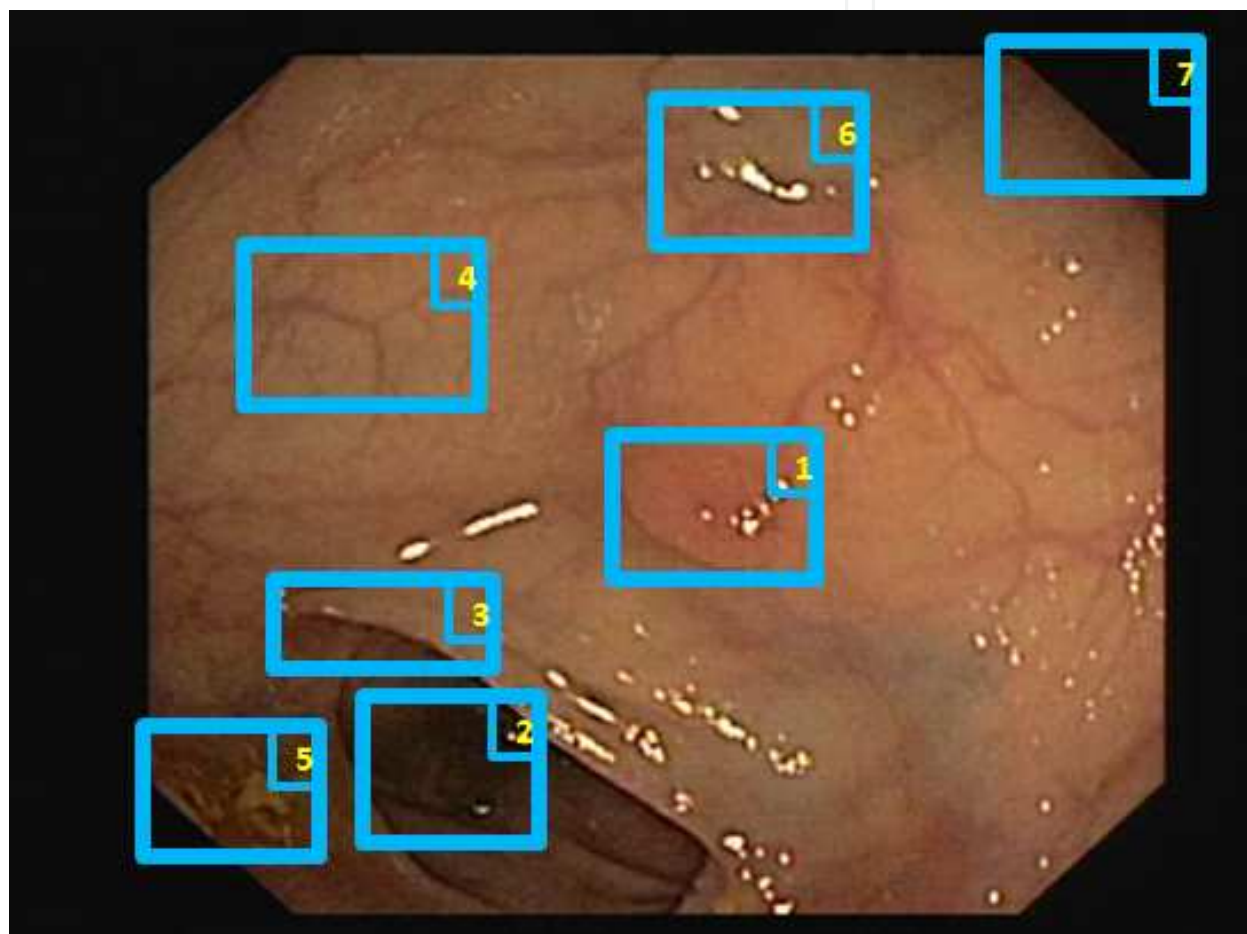
**Figure 10.** Impact of black mask in image processing algorithms. (a) shows the original image whereas (b) shows the output of an edge detection algorithm. Note that mask contours appear as strong as structural elements.

- *Data compression:* Image and video data are commonly compressed in order to save storage space but commonly used formats such as MPEG and JPEG lead to information loss along with the introduction of some artifacts they may difficult fine detail processing in images. In this case the lower the compression, the least impact it will have in further image processing.

### 3.2. Endoluminal scene description challenges

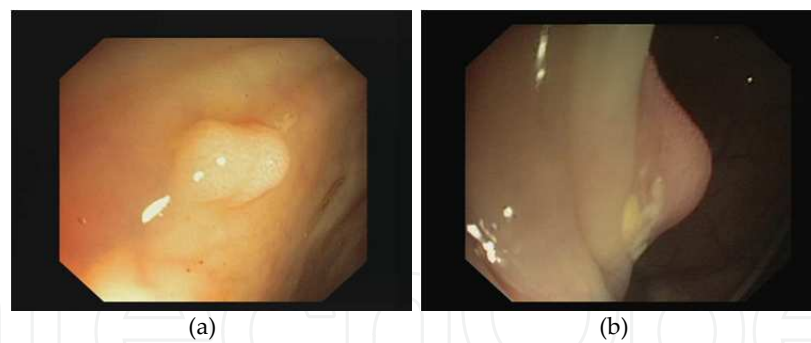
In order to provide with systems that can help clinicians to overcome some of the clinical challenges identified earlier, a description of the elements of the endoluminal scene is needed. We show in Fig.11 an example on how endoluminal scene looks like.

We can make a division of the elements that appear on a given scene into pure anatomical structures (polyps, luminal region, folds, blood vessels or intestinal content) and structures appearing as result of image acquisition and formation processes (specular highlights and black mask). It is clear that a potential intelligent system should focus on the characterization of anatomical structures in order to be clinically useful –being polyps the usual target structure– but, as recent studies demonstrate [29], the consideration of all the elements of the endoluminal scene may result in an improvement of the performance of a given system. Endoluminal structure characterization is not a straightforward task due to three main reasons:



**Figure 11.** Elements of the endoluminal scene: (1) Polyp; (2) Luminal region; (3) Folds; (4) Blood vessels; (5) Intestinal content; (6) Specular highlights and (7) Black mask.

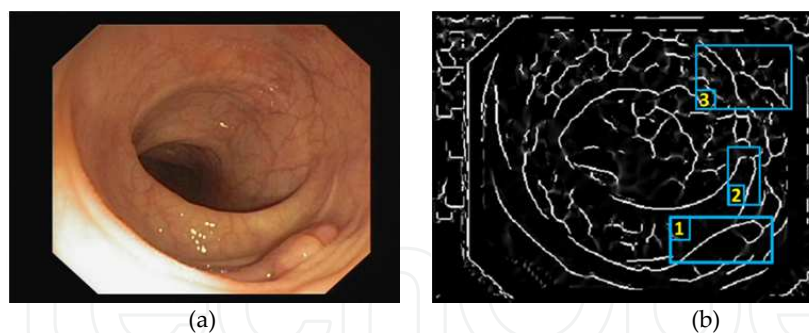
- *Lack of uniform structure appearance:* Anatomical structures appearance differs greatly in different interventions, which may difficult the development of characterization methods that can be widely applicable. For instance, polyp characterization is challenging because there is not an uniform and unique polyp appearance; in fact, polyp appearance depends greatly on the point of view in which it is observed and we can observe different particularities whether we are observing polyps in zenithal or lateral views –see Fig. 12–.



**Figure 12.** Variability in polyp appearance: (1) Zenithal view and (2) Lateral view.

Consequently a definition of a model of appearance for a given structure should consider this great variability in order to be widely applicable and, therefore, search for general features that can be attainable for the majority of the cases.

- *Impact of other elements of the scene on a particular element characterization:* Following with the polyp example, the majority of available works rely on polyp characterization from the identification of polyp boundaries but, in terms of image processing, there is not a big difference in terms of contour appearance between polyps, blood vessels and folds, as the three of them provide with similar response to contour detection operators, as it can be observed in Fig. 13. Considering this, a given intelligent system must consider the impact of all present structures when providing a characterization of a particular one and it will need to find additional cues to differentiate between these structures.



**Figure 13.** Example of similarity of response of different structures to a given operator. Number 1 represents a polyp, number 2 a fold and 3 represents blood vessels.

- *Difficulties on the definition of the structural element:* Another challenge is related to the visual definition of the structure itself, that is, sometimes the definition of the element itself is not clear, which makes it difficult to delimit the structure. For instance, recent studies show a great variability between observers when defining the luminal region –demonstrated in [30], which may have an impact on ground truth creation for assessing the performance of a given intelligent system. This difficult on the definition on the structure can also be applied for other elements such as fecal or intestinal content.



## 4. Equipment setting to favor optimal image processing analysis

We present in this section the optimal settings of clinical equipment to ensure the best possible quality of the images which will be analyzed by the intelligent system.

### 4.1. Endoscopic equipment settings

Chronologically, the first element to be considered is the configuration of both endoscope and video processor in order to obtain the best possible images for further analysis. In this case we propose the following configuration:

- Disable sharpening options, so we can avoid the apparition of artificial information (halos) surrounding image structure contours along with reducing image noise.
- Disable the superimposition of overlay information such as patient or procedure data to obtain a clean view of the endoluminal scene. This also allows a complete anonymization of the information easing its use for research purposes.
- If possible, allow the endoluminal view to occupy the largest portion of the scene without applying any kind of digital zooming operation.
- Configure storage options to obtain data with the minimum possible compression.

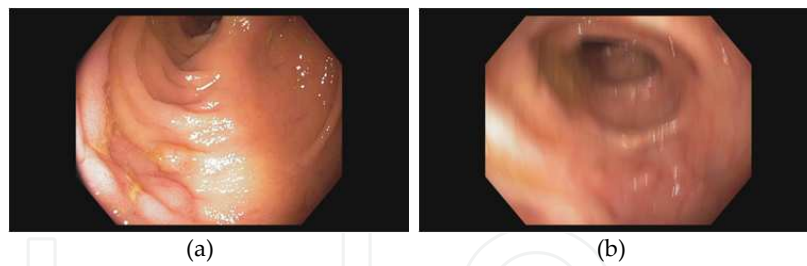
### 4.2. Image storage and anonymization

We have to consider that image or/and video data will be used in research projects from which several research publications will be generated. Access to this image or video data should be granted to other researchers in order to allow an easier comparison of the performance of different methods. Considering this, no information that can allow an identification of either the patient or the clinician should be provided in neither the images or in the metadata associated to them –such as time and date of image capture or endoscopy used-, preventing the association of a given image to a patient, clinician or hospital.

Considering the amount of endoscopic interventions performed in a hospital in a year, images or videos that are stored tend to be compressed. This compression has already been mentioned to have implications for image processing methods so; if possible, the configuration with less possible compression should be chosen.

### 4.3. Endoscopic navigation guidelines

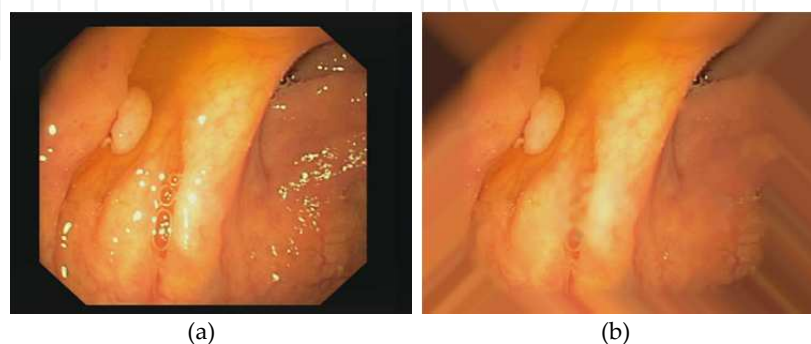
Endoscope movement when images are acquired impacts the quality of the images that are obtained. If there is no scope movement, effects such as interlacing or color phantoms can be almost inexistent -Fig. 14 (a)-. Considering this, we propose still images acquisition to be made being both the scope and the elements of the endoluminal scene static. For the case of video acquisition we suggest slow and smooth endoscope progression through the patient in order to maximize the reduction of movement-related artifacts generation.



**Figure 14.** Difference in image quality related to endoscope movement when acquiring images: (a) still endoscope vs. (b) moving endoscope.

It is clear that even by considering all the suggestions expressed, there will still be a minor movement of the scope between the two time instants in which odd and even lines of the final image are acquired. In order to mitigate the impact of interlacing and to avoid loss of image resolution we propose to make a real-time analysis of the images when they are acquired in order to store only the one which less interlacing impact. This analysis will be made by comparing consecutive frames, where the difference in content between them is so minimal that there is no point on storing them all, considering the small changes that will appear in images extracted from a 30 frames per second video. In case interlacing can still be perceived, its impact can be completely removed by working with one of the two channels of the image [29], although this implies a decrease in final image resolution.

To close this section, we show in Table 2 a summary of the challenges related to image formation and acquisition depicted in Section 3 and our proposal on how to solve/mitigate them. As it can be seen from the table, there are some challenges that cannot be solved by applying specific settings to the devices involved. For instance, those related to image formation are highly device-dependent. In this sense, newer equipment has dedicated sensors for each color channel avoiding the apparition of color phantoms. There are other challenges that must be solved by means of image processing techniques, such as specular highlights. In this sense, the most accepted solution [29] consists of a specular highlight detection followed by a substitution of the pixels in the image belonging to specular highlights by a combination of valid values of neighbor pixels, as it can be observed in Fig. 15. The same operation is applied to mitigate the impact of strong contours created by the black mask.



**Figure 15.** Application of image processing methods to mitigate impact of specular highlights and black mask. (a) Original image and (b) Processed image.

Source	Challenge		Proposed solution
Image formation	Illumination	Specular highlights	Specular highlights correction
		Lack of uniform illumination	Device-dependent
	Sensor acquisition	Color phantoms	Device-dependent
Image acquisition and visualization	Image acquisition and storage	Image resolution	Stabilization of endoscope, interlacing suppression and use of HD endoscopes
		Image interlacing	Interlacing suppression, neighbor frame frames, endoscope stabilization
	Image visualization capabilities enhancement	Sharpening	Disable sharpening
		Presence of patient and procedure information	Disable overlays
		Black mask	Black mask substitution
		Data compression	Use of lossless compression standards.

**Table 2.** Summary of image acquisition and formation challenges along with proposal of solutions

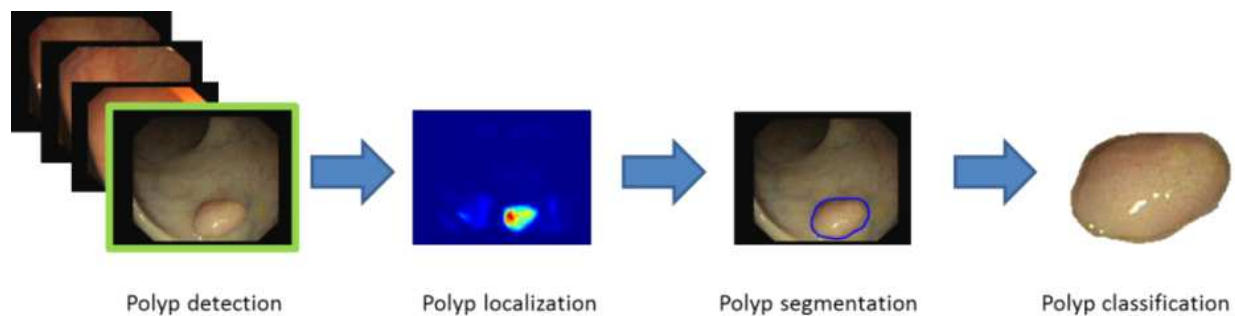
## 5. Current endoluminal scene description methods

We present in this section a review on the most recent works published on the topic of anatomical endoluminal scene elements description.

### 5.1. Polyps

As they are the main focus of colonoscopy explorations, the majority of already existing intelligent systems for colonoscopy deals with polyp characterization. We divide existing systems according to the application they are built for:

- *Polyp detection:* This group of methods aim to decide whether there is a polyp or not in the image. The majority of the works on polyp detection are built on the principle of applying a given feature detector/descriptor to the image in order to guide detection methods. In this sense, we can divide existing approaches in two groups: (a) shape and (b) texture and color-based. The first group aims to detect polyps by observing specific cues on the contours of the polyp –examples of this can be found in works presented in [31-33], or by fitting candidate objects in the image to the most common shapes that polyps present [34]. Regarding the second group, the use of several general descriptors has been proposed, such as wavelets in [35], local binary patterns in [36] or co-occurrence matrices [37]. A method combining MPEG-7 texture and color descriptors was proposed in [38]. One big drawback of descriptor-based methods is that they tend to need of an exhaustive training and they are



**Figure 16.** Example of the output of each polyp characterization group of algorithms.

very sensitive to parameter tuning. Finally the work published in [39] combines shape and texture features to build up a polyp detection method which also considers spatial and temporal adjacency information present in colonoscopy videos.

- *Polyp localization/highlighting:* These methods are focused on highlighting the area of the image more likely to contain a polyp. Considering this, they can be understood as a subgroup of polyp detection method but, in this case, with the objective to establish the area of the image where the polyp is. These methods rely on the definition of a model of polyp appearance and on the exploration of low-level features of the image –in this case, the definition of polyp boundaries in terms of valley information- in order to provide with methods that can be applied in the intervention rooms. Some examples of these methods can be found in the works of Bernal et al [13, 29].
- *Polyp segmentation:* In this case the objective is to delimit the region of the image that the polyp occupies. The majority of available works deal with polyp segmentation in CT images –such as the works depicted in [40,41] –, which can also be useful to provide further features of the polyp such as its size, although considering CT limitations regarding small polyps visibility as mentioned in Section2. Recent works on white light colonoscopy exploit the output from polyp localization methods in order to delimit the final polyp region [42], providing accurate results that could be directly applicable in the intervention room without additional radiation of the patient. Finally there are some recent works [43] that deal with polyp segmentation using narrow-band imaging; preliminary results are promising although its usefulness is restricted to the availability of this imaging modality.
- *Polyp characterization/classification:* The aim of these methodologies concerns lesion characterization according to the content of the polyp region. In this case the objective is to aid clinicians in in-vivo diagnosis and some of the existing works aim to provide automatic lesion labeling using previously-mentioned classifications such as NICE [23] or KUDO [22]. These systems would benefit from an accurate localization and segmentation of the polyp region in order to find features that best discriminate between different polyp types.

As it can be seen from the classification exposed above, a potential intelligent system with applicability in the intervention room could easily use a system from each of the four groups



in order to build up a computer-aided diagnosis tool. We show in Fig. 16 a graphical example of such a system. In a first stage the system will automatically decide which frames contain a polyp and which region of the frame contains the polyp. From this, an accurate segmentation of the polyp region will be obtained in order to extract meaningful features to help in the classification process.

## 5.2. Luminal area

Luminal area is defined as the interior space of a tubular structure, such as the intestine. The detection of the lumen and its position can be crucial in both intervention and post-intervention time.

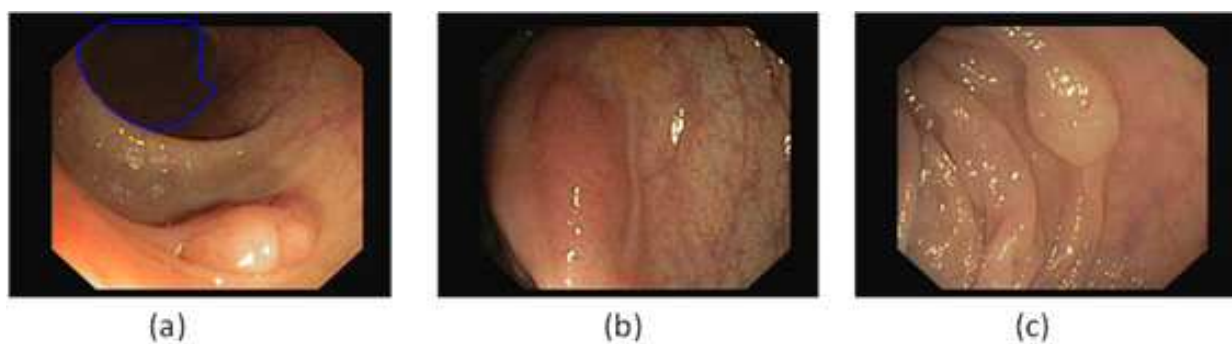
On the one hand, an accurate detection of the lumen region during in-vivo intervention may be useful to discard areas of the image with low visibility –Fig. 17(a) - in order to save computation time for other interesting regions of the image as proposed in [44]. Lumen detection can also be helpful to guide the clinician inside the intestine by pointing out which direction he/she should take to progress. On the other hand, lumen characterization in post-intervention can be used to discard frames for further revision: frames where the proportion of lumen out of the entire image is large can be related to the progression of the colonoscope through the gut but, conversely, frames where the amount of lumen presence is low may potentially indicate areas of the image where the physician has paid more attention. This can be useful to obtain summary videos of the whole procedure. Lumen characterization has been an active topic of research in several endoscopy image modalities such as optical –works of [45] and [46] - and virtual colonoscopy [47]. The main reasoning behind the majority of the luminal region characterization methods is the assumption that lumen is the darkest region of the image and from this seed region growing algorithms are built in order to find lumen boundaries.

## 5.3. Blood vessels

Blood vessels are the part of the circulatory system that transports blood through the body and they can be identified by their tree-like shape with ramifications. The characterization of these branching structures has been reported in domains such as retinal image analysis [48] or palm prints recognition [49]. Blood vessels characterization in colonoscopy images can be useful in two domains: helping in polyp localization and segmentation tasks, as it has been proven in [13, 29, 42], and as key points to be used in potential follow-up methods, as proposed in [50]. Regarding the former, a mitigation of blood vessels related valleys by using contrast properties of blood vessels contours has been proven to be useful to improve polyp localization segmentation, as in some images -Fig. 17(b)- blood vessels can be identified easier than polyp boundaries. Concerning the latter, we could think of a univocal characterization of blood vessels branching patterns using methods such as the one proposed in [51] to recognize a same region during different interventions.

### 5.3.1. Folds

Haustral folds represent folds of mucosa within the colon. They are formed by circumferential contraction of the inner muscular layer of the colon. In the context of intelligent systems for colonoscopy, folds characterization can play a key role in polyp characterization tasks. In this sense, we have to consider that the fold contours appearance in colonoscopy images is very similar to the one of polyps. We can observe in Fig. 17 (c) that folds and polyp contours present similar appearance but different levels of curvature; consequently, an accurate identification of folds could lead to an improvement in polyp characterization tasks. Some recent works build up advances model of polyp appearance to discriminate polyp contours from folds by considering desirable properties of polyp contours such as concavity, completeness or continuity, as proposed in [13].



**Figure 17.** Effect of endoluminal scene structures in polyp characterization: (a) Luminal region (delimited by a blue mask); (b) Blood vessels and (c) Folds.

### 5.4. Fecal content

Apart from the elements that have already been covered, there are more elements that can appear in the endoluminal scene as a result of bad patient preparation. In this sense high presence of intestinal content is considered by clinicians as an indicator to decide whether a procedure has to be repeated or not as no clinician or computer vision method would work with very low quality images. Moreover, there are some cases when the presence of fecal content can affect the output of computer vision methods, as it was shown in [13]. Therefore an accurate identification of fecal content in colonoscopy images could be used to provide automatic indicators of the quality of patients' preparation.

## 6. Building up validation frameworks for intelligent systems

One of the main problems when assessing the performance of the different available intelligent systems for colonoscopy is that the majority of them are tested on private databases, which makes it difficult to observe the differences in performance between them and to extrapolate its functioning in other environments. Moreover, it is very difficult to compare performance

levels of different methods as each of them proposes or uses different evaluation metrics which, for some cases, can be only used with a specific application in mind. Considering this two problems, we present in this section our proposal for a complete validation framework covering from database and ground truth creation to the definition of the metrics to be used to evaluate a given method.

### 6.1. Database creation

In order to validate and assess the performance of a computer vision method, this has to be tested in a set of images covering as many possible cases of study. For instance, if we want our method to be able to characterize polyps from all the types present in Paris classification our database should contain several examples from each of the classes that are defined there. Apart from the original images, a ground truth should also be provided. This ground truth will be used to assess the performance of the method and its configuration will depend on the concrete experiment. Following the same example used before, for polyp localization purposes the ground truth should consist of a binary image where pixels in white should correspond to those pixels which are part of the polyp. If the output of a given method falls in the white pixels of the image, the method will be performing as expected. As it can be seen there are two processes involved when creating databases for intelligent systems validation: the selection of the cases to be included in the database and the creation of the corresponding ground truth.

Regarding the selection of the cases, in order the use of a method can be extended outside research domain, these cases should represent the clinical variability that the clinician can find during interventions. In case we have several types of elements to be characterized, the database should contain as many different examples as possible for all the possible classes. It is important to mention that the more different the examples, the more robust will be our method and the better it will perform once a new case of study is to be analyzed. By doing this if we achieve that a given method offers good performance in our database it will be easy to extrapolate its performance in a potential clinical application.

There is one branch of computer vision known as machine learning which involves method training in a set of images and a posterior testing of this method in a different set of images, once its performance has been optimized in the training stage. Considering this, the size of the database should permit the division in training and testing examples and we should define our database in a way such representative examples of all the possible cases are present both in training and testing databases. The final size of the database should allow extracting statistically significant conclusions. In clinical trials, a variability of less of 10 % is not considered as relevant as stated in [52], being variability calculated as the inverse of the square root of the number of samples  $\sqrt{N}$  in our database. Considering this, the minimum size of the database should be of 100 images.

Once database has been defined, ground truth must be created to validate the performance of the methods. The definition of this ground truth is clearly application dependent: for instance if we are developing a polyp detection method the ground truth may only consist of an excel file indicating for each frame whether there is a polyp or not in the image but for a polyp

segmentation method we would need a binary image representing the structure to be segmented, as it can be seen in Fig. 18.



**Figure 18.** Possible contents of a polyp segmentation database: (a) Original image; (b) Polyp mask; (c) Polyp contour mask and (d) Black mask.

Image-based ground truth are commonly created using image editing software such as Microsoft Paint or Adobe Photoshop, although there is an increasing use of specific tools such as ImageJ [53] which allows the creation of segmentation ground truths by marking a few points in the image. Concerning ground truth creation, it should be created either by clinicians or by experts under clinicians' supervision. Having more than one ground truth per image is recommendable for validation purposes as a way to avoid possible subjectivity in ground truth creation. This allows performing statistical tests and also to assess whether the performance of a given method is within inter-observer variability. If clinical conclusions are meant to be extracted from the performance of intelligent systems, clinical metadata should be provided. For instance, if we want to assess the performance of a polyp classification method, apart from the mask representing where the polyp in the image is, clinicians should provide which is the class of the polyp (i.e., KUDO type I).

Currently there are only, up to our knowledge, three different databases related to colonoscopy image analysis: two of them consisting of still images showing a polyp - CVC-ColonDB and CVC-ClinicDB- and another - ASU-Mayo Clinic polyp database-, which consists of full colonoscopy videos with and without polyps. The first two databases are meant for the validation of model of appearance for polyps to ease polyp localization and segmentation whereas the latter has been developed for the validation of polyp detection algorithms. Currently only CVC-ClinicDB incorporates clinical metadata associated to each polyp, including information regarding polyp size, Paris classification and histological type of polyp. This allows break down of the results according clinical criteria, as exposed in [13]. We introduce the main features of each of the three databases in Table 3.

## 6.2. Performance metrics

The way a given intelligent system method is validated will depend greatly on what this intelligent system is for. The potential application the system is designed for will define both how database and ground truth need to be generated and the metrics used to assess the performance of the method. In this subsection we propose validation protocols for each of the four main types of intelligent systems reported in the literature.



Database	Number of frames/videos	Ground truth content
CVC-ColonDB	380 frames from 15 different sequences with a polyp	For each image with a polyp the following images are provided: 1) original image; 2) polyp mask; 3) non-informative regions mask and 4) polyp contour.
CVC-ClinicDB	612 frames from 29 different sequences with a polyp.	For each image both the original frame along with a mask covering the polyp are provided. For each polyp, clinical metadata associated is provided (size, Paris classification, histological type of polyp after biopsy) [13]
ASU-Mayo Clinic polyp database	Training set: 20 videos (10 with a polyp and 10 without polyps). Testing set: 18 videos	For each frame of the video a binary image is provided. Absence of polyp in the image can be identified by having a completely black associated image. In case of polyp presence, an approximation of polyp region is provided.

Table 3. Summary of available databases for colonoscopy image analysis

- *Polyp Detection:* A given polyp detection method should provide an output whenever a polyp is present in the image and should not provide any output if there is no polyp.

Performance metrics:

Considering this we propose the use of four different concepts (True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN)) which are commonly used in object detection and characterization problems. We present these concepts in Table 4.

Concept	Method	Ground truth
TP	Provides an output	There is a polyp in the image
FP	Provides an output	There is no polyp in the image
TN	Does not provide an output	There is no polyp in the image
FN	Does not provide an output	There is a polyp in the image

Table 4. Explanation of polyp detection metrics

Consequently a good polyp detection method should provide with a high number of TP and TN along the lowest possible number of FN and FP. In order to allow a more clear representation of these results, four different metrics are calculated from TP, FP, TN and FN values:

- Precision, calculated as:  $Prec = \frac{TP}{TP + FP}$ . It represents the fraction of relevant retrieved information. Regarding polyp detection, it represents the percentage of correct alarms (frames where the method provides an output and the image has a polyp). A low precision rate will be interpreted as the system providing a high number of false alarms.

- Recall, calculated as:  $Rec = \frac{TP}{TP + FN}$ . Recall represents the fraction of elements to be retrieved that have been successfully retrieved. In our context, represents the fraction of polyps out of the total that have been correctly detected. Considering this, the highest recall the best the detection method.
- Accuracy, calculated as:  $Acc = \frac{TP + TN}{TP + FP + TN + FN}$ . This measure represents the amount of information that has been correctly labeled. It is useful in cases where positive and negative examples are balanced which is not always the case for polyp detection.
- Specificity, calculated as:  $Spec = \frac{TN}{FP + TN}$ . This represents how good a polyp detection method is when detecting the absence of polyps. A high number of false alarms can be interpreted as the method being less specific regarding polyp presence.

Finally, a polyp detection method will be considered as clinically useful if it can help the clinician to detect the polyp. Considering this and assuming that a given sequence contains a polyp, the following metrics can be defined:

- Reaction time: difference in number of frames between first apparition of the polyp in the sequence and the first frame in which a given method provides detection.
- Dwell time: number of frames with a polyp in which the detection method provides detection.

Considering these two metrics, a comparison can be made between the performance of a given automatic method and clinicians, as it was presented in [13]. This can allow the assessment of the potential of a given method to be included to support clinicians in polyp detection tasks.

#### *Ground truth:*

Ground truth for polyp detection methods validation can consist in either a text file stating which frames contain a polyp or in a binary mask corresponding to each original frame. In this case the binary mask should represent polyp presence and absence (for instance, an all-black image can represent polyp absence).

- *Polyp localization*: Polyp localization methods aim to extend the information provided by polyp detection methods by not only indicating whether there is a polyp in the image or not, but also indicating where the polyp is within the image.

#### *Performance metrics:*

Considering the purpose of localization methods, we cannot use all the four concepts explained before as the use of TN does not make sense in this type of problems as there is always a polyp in the image. In this case several authors [13] propose a more direct performance referred as localization accuracy. Considering that a polyp localization method always provides a potential polyp location, we can define a good localization (GL) whenever the output of the localization method coincides with a polyp. Conversely we define false localization (FL) in the opposite case when the localization proposed by the method falls outside the polyp. Taking this into account, we define localization accuracy as:

$$LAcc = \frac{GL}{GL + FL}$$

In cases where the output of a localization image does not consists of points representing polyp locations but of energy images representing areas with more likelihood of containing a polyp –as it can be seen in Fig. 16- the use of energy concentration metrics seems useful to represent the performance of a method [13]. Considering these two metrics, LAcc and concentration, a good localization method should provide a low number of FL while concentrating the majority of the polyp presence likelihood image inside the polyp mask.

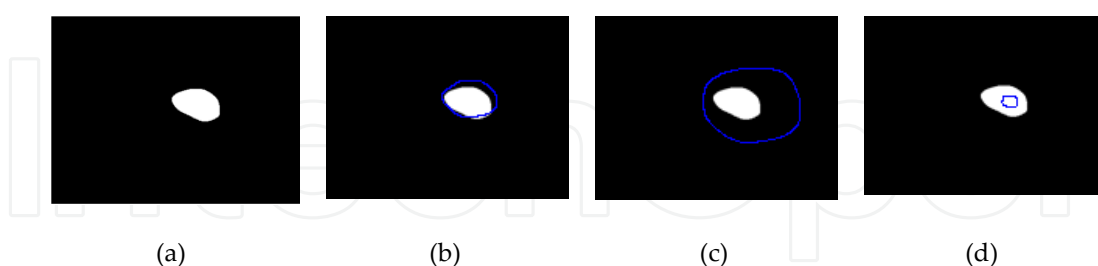
*Ground truth:*

Ground truth for polyp localization should consist of binary masks representing the area of the image that is occupied by the polyp, as it is shown in Figure 18.

- *Polyp segmentation:* An accurate segmentation of the region that contains the polyp can be useful for both lesion recognition tasks as well as for delimiting the area of the image to be used for lesion classification purposes.

*Performance metrics:*

We propose the use of common segmentation metrics such as Precision and Recall, as they were defined for polyp detection. In this case we classify each pixel as TP, FP, TN and FN considering methods' output and the ground truth (i.e. a false positive pixel is defined as a pixel in which our method states it is part of the polyp when it is not). In this context, a good polyp segmentation method should provide higher Precision and Recall results (Fig. 19 (b)); a method providing high Precision with low Recall will provide regions that cannot be used for further polyp characterization as they contain lots of non-polyp information (Fig. 19 (c)). Conversely a method providing with high Recall but low Precision values will be useful for polyp description but will leave a lot of useful polyp content out of posterior analysis (Fig. 19 (d)).



**Figure 19.** Interpretation of segmentations: (a) Original ground truth. Segmentation results with (a) good Precision and Recall values; (c) good Precision but low Recall value and (d) low Precision but good Recall value. Mask representing the output of a given method is represented in blue.

*Ground truth:*

As for the case of polyp localization, ground truth for polyp segmentation should consists of binary masks representing either the area of the image that is occupied by the polyp -Figure 18 (b)- or the contour of the polyp region -Figure 18 (c)-.

- *Polyp classification:* A good polyp classification method should be able to assign the polyp present in the image the same label/class that is attached to the polyp in the ground truth.

*Performance metrics:*

In this case we can have two different types of evaluation, depending on the number of possible classes that we define: if a polyp can only have two different classes we could evaluate our method by checking whether the output of a method coincides or not with the ground truth; in this case for each image we will have a correct (OK) or incorrect classification (NOK). The accuracy of the system will be calculated as

$$Acc = \frac{OK}{OK + NOK}$$

The second type of evaluation is related to multiclass classification; in this case we can also include studies regarding which classes are more easily identified and which classes are mostly confused over each other. In this last case we can use confusion matrices, similar to the ones presented in [54] to represent the output of a given classification method.

*Ground truth:*

Ground truth for polyp classification should consist of a label associated to each frame with a polyp; this label must include the given polyp in any of the possible classes defined in the problem.

## 7. Conclusions

Collaboration between clinicians and computer scientists is crucial for the development of intelligent systems for colonoscopy. Those systems need to be designed to solve real clinical problems if they want to be deployed in clinical environments. Considering this, apart from application development and validation, efforts must be focused on the definition of the aim of the proposed intelligent system.

We have presented in this chapter some of the problems that colonoscopy still present nowadays, being polyp miss-rate the most important of them. Additionally there is a need expressed by clinicians of systems that can allow them to have a first approach to polyp histology, which could be useful to take in-vivo decisions. Considering this we define three possible domains of application of a given intelligent system: polyp detection and localization, polyp classification and development of navigation-assisting and patient follow-up methods.

Once the clinical need is defined, computer scientists must deal with image processing in order to provide with meaningful results. In this context, we have subdivided this problem in two:



image preparation for optimal image processing and endoluminal scene description for intelligent system applications.

Regarding image preparation, one of the main objectives of this chapter was to rise up some concerns about image quality for later processing and clinicians and computer scientists must reach an agreement to obtain images that are useful for both domains. Endoluminal scene description has been proven as a challenging task due to the great variability in structures' appearance throughout different interventions. The majority of bibliographical sources are devoted to polyp characterization, although we have observed an increasing interest in the definition of other elements of the scene, as they have been proven to have an impact in polyp characterization tasks. At this point it is important to mention that there are some aspects that we have not covered in full such as patient preparation although it has a direct consequence on the output of a given intelligent system. In this case we opt to follow the same criteria that clinicians do: if patient preparation is bad neither computer vision nor clinicians would be able to distinguish anything.

The objective of the development of an intelligent system is to take profit of the synergies between clinicians and computer scientists. During the development of a given system, clinicians must provide with data in order to test different methods. We propose in this chapter a validation framework which covers topics such as database and ground truth creation as well as the definition of performance metrics. The proposal of a validation framework including database creation and management along with the definition of standard evaluation metrics can pave the way for a standardized comparison of the performance of intelligent systems which would allow in the future clinicians choose the one that fulfills better their necessities.

The main conclusion that can be extracted from this chapter is that there is indeed room and necessity for the collaboration between these two domains of research. Acknowledging the necessities of each other is meant to play a key role in the development of applicable and deployable intelligent systems for colonoscopy.

## Author details

Jorge Bernal<sup>1\*</sup>, F. Javier Sánchez<sup>1</sup>, Cristina Rodríguez de Miguel<sup>2</sup> and Gloria Fernández-Esparrach<sup>2</sup>

\*Address all correspondence to: [jbernal@cvc.uab.es](mailto:jbernal@cvc.uab.es)

1 Computer Science Department at Universitat Autònoma de Barcelona and Computer Vision Center, Barcelona, Spain

2 Endoscopy Unit, Gastroenterology Department, Hospital Clínic, IDIBAPS, CIBEREHD, University of Barcelona, Barcelona, Spain

## References

- [1] Bratko, I., Mozetič, I., & Lavrač, N. (1990). KARDIO: a study in deep and qualitative knowledge for expert systems. MIT Press. DOI:10.1016/0933-3657(91)90006-W
- [2] Wilschut, J. A., Steyerberg, E. W., van Leerdam, M. E., Lansdorp-Vogelaar, I., Habbe-  
ma, J. D. F., & van Ballegooijen, M. (2011). How much colonoscopy screening should be  
recommended to individuals with various degrees of family history of colorectal cancer?.  
*Cancer*, 117(18), 4166-4174. DOI: 10.1002/cncr.26009
- [3] Viswanath, S., Palumbo, D., Chappelow, J., Patel, P., Bloch, B. N., Rofsky, N et al. (2011,  
March). Empirical evaluation of bias field correction algorithms for computer-aided  
detection of prostate cancer on T2w MRI. In *SPIE Medical Imaging* (pp. 79630V-79630V).  
International Society for Optics and Photonics. DOI:10.1117/12.878813
- [4] Bale, R., & Widmann, G. (2007). Navigated CT-guided interventions. *Minimally Invasive  
Therapy & Allied Technologies*, 16(4), 196-204. DOI:10.1080/13645700701520578
- [5] Schwarz, Y., Greif, J., Becker, H. D., Ernst, A., & Mehta, A. (2006). Real-time electromag-  
netic navigation bronchoscopy to peripheral lung lesions using overlaid CT images: the  
first human study. *CHEST Journal*, 129(4), 988-994. DOI:10.1378/chest.129.4.988
- [6] Fernández-Esparrach, G., Estépar, R. S. J., Guarner-Argente, C., Martínez-Pallí, G.,  
Navarro, R., de Miguel, C. R et al. (2010). The role of a computed tomography-based  
image registered navigation system for natural orifice transluminal endoscopic sur-  
gery: a comparative study in a porcine model. *Endoscopy*, 42(12), 1096. DOI: 10.1055/  
s-0030-1255824
- [7] Grbic, S., Ionasec, R., Mansi, T., Georgescu, B., Vega-Higuera, F., Navab, N., & Comani-  
ciu, D. (2013, April). Advanced intervention planning for Transcatheter Aortic Valve  
Implantations (TAVI) from CT using volumetric models. In *Biomedical Imaging (ISBI),  
2013 IEEE 10th International Symposium on* (pp. 1424-1427). IEEE. DOI: 10.1109/ISBI.  
2013.6556801
- [8] Thomas-Gibson, S., Bassett, P., Suzuki, N., Brown, G. J., Williams, C. B., & Saunders, B.  
P. (2007). Intensive training over 5 days improves colonoscopy skills long-term.  
*Endoscopy*, 39(9), 818-824. DOI: 10.1055/s-2007-966763
- [9] Kiesslich, R., Fritsch, J., Holtmann, M., Koehler, H. H., Stolte, M., Kanzler, S. et al. (2003).  
Methylene blue-aided chromoendoscopy for the detection of intraepithelial neoplasia  
and colon cancer in ulcerative colitis. *Gastroenterology*, 124(4), 880-888. DOI:10.1053/  
gast.2003.50146
- [10] Machida, H., Sano, Y., Hamamoto, Y., Muto, M., Kozu, T., Tajiri, H., & Yoshida, S.  
(2004). Narrow-band imaging in the diagnosis of colorectal mucosal lesions: a pilot  
study. *Endoscopy*, 36(12), 1094-1098. DOI:10.1055/s-2004-826040
- [11] Yoshida, Y., Matsuda, K., Sumiyama, K., Kawahara, Y., Yoshizawa, K., Ishiguro, H.,  
& Tajiri, H. (2013). A randomized crossover open trial of the adenoma miss rate for  
narrow band imaging (NBI) versus flexible spectral imaging color enhancement

- (FICE). *International journal of colorectal disease*, 28(11), 1511-1516. DOI:10.1007/s00384-013-1735-4
- [12] Hoffman, A., Kagel, C., Goetz, M., Tresch, A., Mudter, J., Biesterfeld, S. et al. (2010). Recognition and characterization of small colonic neoplasia with high-definition colonoscopy using i-Scan is as precise as chromoendoscopy. *Digestive and Liver Disease*, 42(1), 45-50. DOI :10.1016/j.dld.2009.04.005
- [13] Bernal, J., Sánchez, F. J., Fernández-Esparrach, G., Gil, D., Rodríguez, C., & Vilarinho, F. (2015). WM-DOVA Maps for Accurate Polyp Highlighting in Colonoscopy: Validation vs. Saliency Maps from Physicians. *Computerized Medical Imaging and Graphics*, 43, pp. 99-111. DOI :10.1016/j.compmedimag.2015.02.007
- [14] Modlin, I. M. (2000). A brief history of endoscopy. *MultiMed*.
- [15] Pellise, M., Fernández-Esparrach, G., Cardenas, A., Sendino, O., Ricart, E., Vaquero et al. (2008). Clinical impact of wide-angle, high-resolution endoscopy in the diagnosis of colorectal neoplasia in a non-selected population: a prospective randomized controlled trial. *Gastrointestinal Endoscopy*, 67(5), AB101. DOI :10.1016/j.gie.2008.03.119
- [16] Rex, D. K., & Helbig, C. C. (2007). High yields of small and flat adenomas with high-definition colonoscopes using either white light or narrow band imaging. *Gastroenterology*, 133(1), pp. 42-47. DOI :10.1053/j.gastro.2007.04.029
- [17] Quintero, E., Castells, A., Bujanda, L., Cubiella, J., Salas, D., Lanas, Á. et al. (2012). Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. *New England Journal of Medicine*, 366(8), pp. 697-706. DOI: 10.1056/NEJMoa1108895
- [18] Barclay, R. L., Vicari, J. J., Doughty, A. S., Johanson, J. F., & Greenlaw, R. L. (2006). Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. *New England Journal of Medicine*, 355(24), pp. 2533-2541. DOI: 10.1056/NEJMoa055498
- [19] van Rijn, J. C., Reitsma, J. B., Stoker, J., Bossuyt, P. M., van Deventer, S. J., & Dekker, E. (2006). Polyp miss rate determined by tandem colonoscopy: a systematic review. *The American journal of gastroenterology*, 101(2), pp. 343-350. DOI:10.1111/j.1572-0241.2006.00390.x
- [20] Bretagne, J. F., Manfredi, S., Piette, C., Hamonic, S., Durand, G., & Riou, F. (2010). Yield of high-grade dysplasia based on polyp size detected at colonoscopy: a series of 2295 examinations following a positive fecal occult blood test in a population-based study. *Diseases of the Colon & Rectum*, 53(3), pp. 339-345. DOI: 10.1007/DCR.0b013e3181c37f9c
- [21] Samadder, N. J., Curtin, K., Tuohy, T. M., Pappas, L., Boucher, K., Provenzale, D. et al. (2014). Characteristics of missed or interval colorectal cancer and patient survival:

- a population-based study. *Gastroenterology*, 146(4), pp. 950-960. DOI:10.1053/j.gastro.2014.01.013
- [22] Kudo, S. E., Wakamura, K., Ikehara, N., Mori, Y., Inoue, H., & Hamatani, S. (2011). Diagnosis of colorectal lesions with a novel endocytoscopic classification—a pilot study. *Endoscopy*, 43(10), 869. DOI: 10.1055/s-0030-1256663
- [23] Hayashi, N., Tanaka, S., Hewett, D. G., Kaltenbach, T. R., Sano, Y., Ponchon, T. et al. (2013). Endoscopic prediction of deep submucosal invasive carcinoma: validation of the narrow-band imaging international colorectal endoscopic (NICE) classification. *Gastrointestinal endoscopy*, 78(4), pp. 625-632. DOI:10.1016/j.gie.2013.04.185
- [24] Fidler, J. L., Johnson, C. D., MacCarty, R. L., Welch, T. J., Hara, A. K., & Harmsen, W. S. (2002). Detection of flat lesions in the colon with CT colonography. *Abdominal imaging*, 27(3), pp. 292-300. DOI: 10.1007/s00261-001-0171-z
- [25] Fidler, J., & Johnson, C. (2009). Flat polyps of the colon: accuracy of detection by CT colonography and histologic significance. *Abdominal imaging*, 34(2), pp. 157-171. DOI: 10.1007/s00261-008-9388-4
- [26] Johnson, C. D., Chen, M. H., Toledano, A. Y., Heiken, J. P., Dachman, A., Kuo, M. D et al. (2008). Accuracy of CT colonography for detection of large adenomas and cancers. *New England Journal of Medicine*, 359(12), pp. 1207-1217. DOI: 10.1056/NEJMc080041
- [27] Poynton, C. (2012). *Digital video and HD: Algorithms and Interfaces*. Elsevier.
- [28] Cambridge in Color: A Learning Community for Photographers. Sharpening: Unsharp Mask. <http://www.cambridgeincolour.com/tutorials/unsharp-mask.htm>. Last visit: 29/03/2015.
- [29] Bernal, J., Sánchez, F. J., & Vilariño, F. (2013, July). Impact of image preprocessing methods on polyp localization in colonoscopy frames. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE* (pp. 7350-7354). IEEE. DOI:10.1109/EMBC.2013.6611256
- [30] Sánchez, C., Bernal, J., Sánchez, F.J., Díez, M., Rosell, A. & Gil, D. (2015). Towards On-line Quantification of Tracheal Stenosis from Videobronchoscopy. *International Journal of Computer Assisted Radiology and Surgery*, 10(6), pp. 935-945. DOI: 10.1007/s11548-015-1196-z
- [31] Krishnan, S. M., Yang, X., Chan, K. L., Kumar, S., & Goh, P. M. Y. (1998). Intestinal abnormality detection from endoscopic images. In *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE* (Vol. 2, pp. 895-898). IEEE. DOI:10.1109/IEMBS.1998.745583
- [32] Zhu, H., Fan, Y., & Liang, Z. (2011). Improved curvature estimation for shape analysis in computer-aided detection of colonic polyps. In *Virtual Colonoscopy and Ab-*



- dominal Imaging. Computational Challenges and Clinical Opportunities (pp. 9-14). Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-25719-3\_2
- [33] Tajbakhsh, N., Gurudu, S. R., & Liang, J. (2014). Automatic Polyp Detection Using Global Geometric Constraints and Local Intensity Variation Patterns. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2014*(pp. 179-187). Springer International Publishing. DOI: 10.1007/978-3-319-10470-6\_23
  - [34] Kang, J., & Doraiswami, R. (2003, May). Real-time image processing system for endoscopic applications. In *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on* (Vol. 3, pp. 1469-1472). IEEE. DOI:10.1109/CCECE.2003.1226181
  - [35] Burrus, C. S., Gopinath, R. A., & Guo, H. (1998). *Introduction to wavelets and wavelet transforms* (Vol. 998). New Jersey: Prentice hall.
  - [36] Bernal, J., Sánchez, F. J., & Vilariño, F. (2010). *Feature Detectors and Feature Descriptors: Where We Are Now*. Technical Report. Computer Vision Center.
  - [37] Ameling, S., Wirth, S., Paulus, D., Lacey, G., & Vilariño, F. (2009). Texture-based polyp detection in colonoscopy. In *Bildverarbeitung für die Medizin 2009*(pp. 346-350). Springer Berlin Heidelberg. DOI: 10.1007/978-3-540-93860-6\_70
  - [38] Coimbra, M. T., & Cunha, J. S. (2006). MPEG-7 visual descriptors—contributions for automated feature extraction in capsule endoscopy. *Circuits and Systems for Video Technology, IEEE Transactions on*, 16(5), 628-637. DOI:10.1109/TCSVT.2006.873158
  - [39] Park, S. Y., Sargent, D., Spofford, I., & Vosburgh, K. G. (2012). A colon video analysis framework for polyp detection. *Biomedical Engineering, IEEE Transactions on*, 59(5), pp. 1408-1418. DOI: 10.1109/TBME.2012.2188397
  - [40] Xu, Y. R., & Zhao, J. (2014). Segmentation of haustral folds and polyps on haustral folds in CT colonography using complementary geodesic distance transformation. *Journal of Shanghai Jiaotong University (Science)*, 19, pp. 513-520. DOI: 10.1007/s12204-014-1534-2
  - [41] Van Wijk, C., Van Ravesteijn, V. F., Vos, F. M., & Van Vliet, L. J. (2010). Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow. *Medical Imaging, IEEE Transactions on*, 29(3), pp. 688-698. DOI: 10.1109/TMI.2009.2031323
  - [42] Bernal, J., Núñez, J. M., Sánchez, F. J., & Vilariño, F. (2014). Polyp Segmentation Method in Colonoscopy Videos by Means of MSA-DOVA Energy Maps Calculation. In *Clinical Image-Based Procedures. Translational Research in Medical Imaging* (pp. 41-49). Springer International Publishing. DOI: 10.1007/978-3-319-13909-8\_6
  - [43] Ganz, M., Yang, X., & Slabaugh, G. (2012). Automatic segmentation of polyps in colonoscopic narrow-band imaging data. *Biomedical Engineering, IEEE Transactions on*, 59(8), pp. 2144-2151. DOI: 10.1109/TBME.2012.2195314

- [44] Bernal, J., Gil, D., Sánchez, C., & Sánchez, F. J. (2014). Discarding Non Informative Regions for Efficient Colonoscopy Image Analysis. In *Computer-Assisted and Robotic Endoscopy* (pp. 1-10). Springer International Publishing. DOI: 10.1007/978-3-319-13410-9\_1
- [45] Arnold, M. An Image Analysis and Machine Learning Approach to Measuring the Quality of Individual Colonoscopy Procedures, PhD Thesis, (2012).
- [46] Tian, H., Srikanthan, T., & Asari, K. V. (2001). Automatic segmentation algorithm for the extraction of lumen region and boundary from endoscopic images. *Medical and Biological Engineering and Computing*, 39(1), pp. 8-14. DOI: 10.1007/BF02345260
- [47] Lu, L., Zhang, D., Li, L., & Zhao, J. (2012). Fully automated colon segmentation for the computation of complete colon centerline in virtual colonoscopy. *Biomedical Engineering, IEEE Transactions on*, 59(4), pp. 996-1004. DOI:10.1109/TBME.2011.2182051
- [48] Azzopardi, G., & Petkov, N. (2013). Automatic detection of vascular bifurcations in segmented retinal images using trainable COSFIRE filters. *Pattern Recognition Letters*, 34(8), pp. 922-933. DOI:10.1016/j.patrec.2012.11.002
- [49] Pudzs, M., Fuksis, R., & Greitans, M. (2013, April). Palmprint image processing with non-halo complex matched filters for forensic data analysis. In *Biometrics and Forensics (IWBF), 2013 International Workshop on* (pp. 1-4). IEEE. DOI:10.1109/IWBF.2013.6547317
- [50] Núñez, J. M., Bernal, J., Ferrer, M., & Vilariño, F. (2014). Impact of Keypoint Detection on Graph-Based Characterization of Blood Vessels in Colonoscopy Videos. In *Computer-Assisted and Robotic Endoscopy* (pp. 22-33). Springer International. DOI: 10.1007/978-3-319-13410-9\_3
- [51] Núñez, J. M., Bernal, J., Sánchez, F. J., & Vilariño, F. (2015). GRowing Algorithm for Intersection Detection (GRAID) in branching patterns. *Machine Vision and Applications*, 26(2-3), pp. 387-400. DOI: 10.1007/s00138-015-0663-4
- [52] Julious, S. A. (2009). *Sample sizes for clinical trials*. CRC Press. DOI: 10.1201/9781584887409
- [53] Abràmoff, M. D., Magalhães, P. J., & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics international*, 11(7), pp. 36-43.
- [54] Chapelle, O., Haffner, P., & Vapnik, V. N. (1999). Support vector machines for histogram-based image classification. *Neural Networks, IEEE Transactions on*, 10(5), pp. 1055-1064. DOI:10.1109/72.788646

