QUALITY ASSESSMENT IN COLONOSCOPY
New Challenges Through Computer Vision-based Systems

Fernando Vilariño
Computer Vision Center and Computer Science Dep. Universitat Autònoma de Barcelona, Spain

Gerard Lacey
Graphics, Vision and Visualisation Group (GV2), School of Computer Science and Statistics, Trinity College Dublin, Ireland

Keywords: Colonoscopy, Quality assessment, Eye-tracker, Computer vision, Polyps, Colon cancer.

Abstract: The assessment of the quality of the colonoscopic interventions arises as a most relevant issue once the number and the availability of these clinical procedures are increased day by day. The use of the latest computer vision-based techniques can provide the physician with both qualitative and, most important, objectively verifiable quantitative indicators of performance. In this paper we present a study in which we propose the automatic analysis of colonoscopy video for the quality assessment of the intervention from different points of view: 1) We propose the characterization of the different parts of the colon in order to obtain metrics of the time used for navigation, portion of gut analyzed, etc. 2) We analyze the image contents in order to automatically characterize the presence of polyps. 3) We use the information obtained by and eye-tracker in order to assess the physician's skills.

1 COLON CANCER AND INTESTINAL SCREENING

1.1 Colon Cancer in Numbers

The main lesions associated to the intestine are: bleeding, lump, ulcer, Crohn disease, and cancer. During the last 20 years, colon cancer has been the second leading cause of cancer deaths in the United States, behind lung cancer, with approximately $60,000$ deaths per year as shown in O'Brien's report (O'Brien et al., 1990), and also analyzed in other studies (U.S. Department of Health and Human Services, 2003). America Cancer Society's 2007 report (American Cancer Society, 2007) provides an extended summary of facts and statistics about colon cancer prevalence and impact in the population. Colorectal cancer is the second leading cause of cancer-related deaths in Singapore and Europe (Ministry of Health Singapore, 1990). Although colon screening has become the main alternative for prevention of colorectal cancer, recent data suggest that there is a significant miss-rate for the detection of even relatively large polyps and cancer (Pabby et al., 2005). For this reason, special efforts have been focused on the development of computer-aided systems for the detection of this type of pathologies. Nowadays, novel lines of research are oriented to widen this perspective by means of the implementation of objective indicators for the assessment of the procedures used in colon screening, since the miss rate of polyps is highly correlated to elements such as the quality of the preparation, the time consumed for the intervention, the amount of intestinal surface screened, the kills of the physician in the manipulation of the endoscope, etc. Moreover, this indicators present a potential value for the training of future endoscopists as reference values to measure abilities in and objective way.

1.2 Screening Technics

Intestinal endoscopy is referred to as the technique for screening the intestinal lumen. For the case of large intestine, this technique receives the name of colonoscopy. Fiberoptic colonoscopy (FOC) is widely accepted as the definitive method for diagnosis of colonic polyps. FOC allows direct visualization of the intestinal surface and affords the possibility of obtaining a in-situ biopsy as well as cauteriza-

Vilariño F. and Lacey G. (2009).
QUALITY ASSESSMENT IN COLONOSCOPY - New Challenges Through Computer Vision-based Systems.
In *Proceedings of the International Conference on Biomedical Electronics and Devices*, pages 320-325
DOI: 10.5220/0001780703200325

tion and clinical intervention such as polyp removal (Winawer et al., 1997). FOC is a minimal invasive surgery (Hunter and Sackier, 1993), (Hulka and Reich, 1994) consisting of the introduction through the anus of a flexible probe which a camera and an illumination device on its tip. The probe consists of a flexible cable which can be controlled by the expert in order to reach every part on the intestinal wall. There exist several reference books regarding clinical colonoscopy for the reader interested in deepening in the specificities of colonoscopy, amongst which Kato's textbook (Kato and Baron, 2003) can provide an insightful introductory view. Several authors have assessed that endoscopic images possess rich information (Nagasako et al., 1998), which facilitates the abnormality detection by multiple techniques (Zheng et al., 2005). The main drawbacks related with FOC can be enumerated as follows (Winawer et al., 1993), (Eddy, 1990): risks of perforation; costs of the intervention; difficulty in visualizing the 100% of the intestinal surface; high number of patients for a reduced number of specialists provide a stress and shrink in the intervention time; difficult visualization due to the intestinal content; preparation needed; imprecise localization of events for a subsequent interventions, and high rate of negative results in inspections looking for polyps (Gokturk and Tomasi, 2001).

In the last years, other modalities for the visualization of the colon have arisen. One of the most relevant is virtual colonoscopy, a technique consisting of the construction of a virtual 3D model of the colon from computed tomography (CT) data (Liang et al., 2004), and for this reason several authors have proposed automatic techniques for the detection of lesions in this modality of images. On the other hand, Wireless Capsule Video Endoscopy (WCVE) (Iddan and Meron, 2000) has devoted particular attention, recently. WCVE (Fireman et al., 2002) consists of a capsule with a camera, a battery and a set of led lamps for illumination attached to it, which is swallowed by the patient, emitting a radio frequency signal which is received and stored in an external device. The result is a video movie which records the trip of the capsule along the intestinal tract with a rate of two frames per second, and that can be easily downloaded into a PC with the camera software installed. It is much less invasive, since the patient simply has to swallow the pill, which will be expelled in the normal cycle through defecation. Moreover, there is no need of hospitalization nor expert support through the process and the patient can lead an ordinary life, since the attached device is recording the video movie emitted by the camera in the capsule (Vilariño, 2006; Vilariño et al., 2009). However, although these new modalities provide new ways of clinical analysis, colonoscopy is the reference technique for clinical intervention in the case of colon screening and colon cancer.

2 COMPUTER VISION FOR QUALITY ASSESSMENT

We state that computer vision-based techniques can be used in order to obtain objective indicators of the quality of the interventions in an automatic way. Our approach provides a framework for the quality assessment which is implemented in the following 3 areas in which we are developing our research, namely: 1) Computer-aided intervention: On-line detection and characterization of potential targets in intervention-time, 2) Post-interventional quality metrics: Automatic computation of the quality measures related to the intervention such as quality of preparation, amount of bowel surface visualized, time measures, etc. 3) Evaluation of skills: Analysis of 3D trajectories of the endoscopes and the screening behavior. For this final point, we propose to use the trajectories together with the information obtained from the tracking of the gaze position of trainees in order to assess their skills. The perspective presented in this paragraph is graphically depicted in Figure 1.

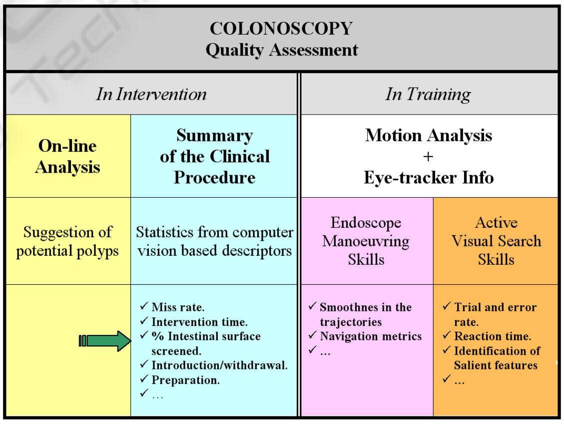

Figure 1: Our framework for automatic quality assessment in colonoscopy.

3 COLON CANCER: CLINICAL CHARACTERIZATION

Colon cancer has devoted wide attention in many studies in order to find proper descriptions and categorization (Rembacken et al., 2000), (Saitoh et al., 2001), (Paris Workshop Participants, 2003). Adenomatous polyps, particularly those larger than 1 cm in diameter, are the most likely precursors of colorectal carcinoma (Gokturk and Tomasi, 2001), (Thoeni and Laufer, 1994). The main features used for cancer and general abnormality characterization are: color, shape and texture.

- **Color.** Color colonoscopic images tend to exhibit the same color features for the same colon status (Kato and Baron, 2003). Malignant tumors are usually inflated and inflamed and this inflammation is usually reddish and more severe in color than the surrounding tissues. Benign tumors exhibit less intense hues. Redness may specify bleeding and black may be treated as deposits due to laxatives. Green may be the presence of faecal materials, which are not clear during the preoperative preparation, and yellow relates to pus formation (Tjoa and Krishnan, 2003).
- **Shape.** Shape is a relevant cue since polyps are associated to rounded or peduncular shapes. Peduncular polyps are relatively easy to visualize during a screening session. Flat polyps present a higher difficulty, and in addition, they are more likely to develop into malignant polyps.
- **Texture.** Texture is known to be an important cue to be evaluated for the discrimination between malignant and benign lesions (Kudo and Kashida, 2000), (Nagata et al., 2000).

The high-level characterization of cancer explained above is usually translated into an image-based feature extraction stage, focused on color, texture or shape cues. Following the feature extraction, a discrimination procedure, based on simple comparisons or more sophisticated machine learning techniques must be applied. In our approach, efficient methods for color, texture and shape characterization must be oriented towards the high speed requirements of on-line procedures. The use of histogram quantization (HQ) (Swain and Ballard, 1991) and the use of different color spaces (Paschos, 2001) in a whole-image level (Hai et al., 2006) or a multi-scale framework (Li et al., 2005) for color; *Gray level co-occurrence matrices* (*GLCM*) (Srivastava et al., 2005), *fractal dimension* (Chaudhuri and Sarkar, 1995), histograms of oriented gradients (HOG) (Won et al., 2002) and wavelets (Karkanis et al., 2001) for texture; and MPEG-7 descriptors and others for shape (Coimbra and Cunha, 2006) should be adapted to run online in order to provide a efficient characterization of our system. This orientation towards the real time performance of discriminant features is one of the most relevant challenges of our current line of research.

Figure 2: Three polyps: flat, peduncular and mixed. Peduncular polyps are prone to their development into malignant cancers.

In recent works we presented preliminary results of discriminative features and classification systems for colon cancer detection and its a posteriori characterization in different types of polyps. Figure 2 shows examples of a) peduncular, b) flat and c) mixed polyps, which have a different degree of clinical relevance, as appearing in our latest contribution (Vilariño et al., 2007). Our next step is, hence, to put these techniques in the horizon of real-time performance. In addition, we argue that the use of mixed strategies (color/texture/shape) for characterization would potentially provide a reacher information than each of them alone, and the study of the selection of appropriate features represents an open field of research.

4 POST-INTERVENTIONAL QUALITY METRICS

We state that computer vision-based techniques can be used in order to provide post-interventional metrics in an automatic way. We define four main lines or research in this area:

1. Automatic estimation of presence of intestinal content for the assessment of the quality of the preparation.
2. Percentage of the intestine visualized.
3. Measure of the intervention time.
4. Automatic detection of the different moments of the intervention (introduction of the endoscope/withdrawal).
5. Characterization of the motion of the endoscope.

All these points can be gathered into a general framework which could be stated in the following way: The definition of a 2D map of the patient's gut in order to have patient-oriented representation of the colon, together with their singular features, lesions, etc. In order to get this wide target, color, textural and motion features must be used together. Color appears as a main cue for intestinal content characterization

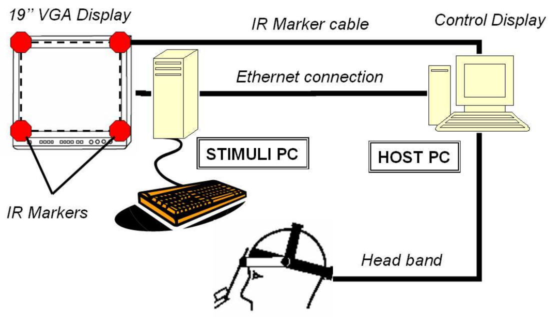

Figure 3: Eye-tracker configuration system.

(Vilariño et al., 2006). Texture can provide a description of the intestinal folds, and motion characterization is essential for the measure of transition times. In a addition, deeper studies must be performed in order to build up a large database of cases which allows us to carry out statistical test with the aim of unveiling the most suitable features and indicators for quality measures.

5 EYE-TRACING INFORMATION FOR SKILL ASSESSMENT

The last line of research we introduce in this paper is related to the assessment of the skills of the trainees in endoscopy training programs. We hold that using both the information of the trajectories of the endoscopes -which can be potentially obtained by means of the analysis of the camera movement- and the tracking of the gaze position -which can be obtained by an eye-tracker device-, we can provide objective indicators for the evaluation of the skills. Moreover, it will be possible to characterize those skills separately: ability in endoscope manipulation vs. visual search.

An eye tracker consists of an adjustable device with a set of cameras pointing towards the user's eyes. The eye tracking procedure involves the recording of the gaze position and the translation of this values into the location on the image where the user is focusing the attention. Figure 3 shows a scheme of the EyeLink II eye-tracker which was used for our experiments. The video signal from the cameras is sent to a host PC in which the eye tracker software is installed, and which communicates to the experiment PC in which the visual stimuli are shown. The pupil position is then translated into gaze position on the stimuli screen at a very high speed (up to one record each four milliseconds).

6 RESULTS

In this position paper, we would like to highlight preliminary results which were obtained in the colon cancer characterization and eye-tracker based analysis. The computation of automatic quality assessment measures is a key part of our ongoing research, and it will be analyzed in more detail in the Discussion section.

The eye-tracker information can be used to label massive data by using the gaze position of the specialist while screening the colonoscopy video off-line (in a post-interventional session). Those areas the specialist is steering the sight towards are associated to visual salient features which are of interest for the colonoscopists. This has two main consequences: On the one hand, we can build up an annotated database of clinical cases in which the position of the observer's gaze determines the position of the cancer screened. On the other hand, we can use the information of the the gaze trajectory to characterize the abilities of the experts in the visual search. In our case, we built up a database consisting of 6 cases, which is freely accessible for the public and from the following link: `hhttp://www.cs.tcd.ie/colon/colon_et_database.zip`.

We used this database to for the experiments associated to the colon cancer detection and characterization. We applied support vector machine classifiers (SVM) (Vapnik, 1995) in order to distinguish between polyp images and random images of 6 different videos. Each image was 500x700 pixels. We manually selected those frames were the polyps were present, and then we trained the classifier with examples of polyps and non-polyps patches of 128x128 pixels by using the gray-level image, achieving a working point of around 80% of both sensitivity and specificity. The visual analysis of these data provided promising results in the clustering of the different types of polyps into three basic types (peduncular, flat and mixed). For a further analysis of these results we refer to the IbPria conference proceedings (Vilariño et al., 2007).

Regarding the analysis of the visual trajectories obtained from the eye-tracker, we carried out a new set of experiments in which 20 full colonoscopy studies were annotated in a two-stage strategy: 1) First, the expert manually selected those sections of the video study in which the expert detected the presence of cancer -this action was performed by clicking and holding a mouse during the visualization of the polyp-. 2) Then, only those parts of the video selected by the specialist in the previous stage were screened with the eye-tracker. We repeated this experiment with different experts and novices in order to look for differences in annotation. The results, which are to be deeply analyzed in a current study to be submitted in a forthcoming publication, point out that experts and trainees show different behaviors in the visual search in terms of reaction time, frequency of saccades, geometry of the 3D trajectories of the gaze position, etc. These preliminary results show statistical significance, and they must be contrasted in terms of inter- and intra-observer variability. This is one of the most relevant drawbacks of this kind of studies since, the time availability of the physicians is a major constraint for large validation studies.

7 DISCUSSION

The preliminary results shown above pave the way to the computation of quality assessment measures. For the case of polyps detection, the main aim is to provide the physician with candidates of polyps to be analyzed during the intervention. Different statistics, such as the number of polyp-candidates analyzed and not analyzed, total time consumed in this analysis, etc., can be pulled out in order to obtain objective indicators. These statistics and indicators must be defined together with the physicians in order to get clinical significance and the appropriate tolerance levels.

For the case of eye-tracker data, statistically significant indicators of the physicians' visualization skills are of relevant importance in a two-fold way: First, we would be able to provide the physician with objective metrics that measure high level skills, such as reaction time, search activity, robustness of the search pattern, etc., together with general indicators such as miss rate of polyps. In addition to the former, we make it possible to decouple the manoeuvering ability from the visual search skills in an objective way. This information, together with the motion analysis, can provide indicators regarding the smoothness of the trajectories that the endoscope performs.

Finally, the automatic estimation of presence of intestinal content, the quality of the preparation, the percentage of the intestine visualized, the measure of the intervention time, and the automatic detection of the introduction and withdrawal stages must be put into a clinical framework of quality assessment. This provides both a study-based assessment and, into a historical archive, a log of the physician's indicators along different interventions. In order to get this done, a full study is being performed currently together with the Royal College of Surgeons of Ireland, funded by an Enterprise Ireland project of the Irish government for the acquisition and annotation of a full database of 100 cases of colonoscopy showing colon cancer in high definition videos. This database comprises several tera bytes of video data and its corresponding annotation both in manual and eye-tracked versions by several specialists with different levels of expertise.

8 CONCLUSIONS

Quality assessment of colonoscopy videos is a relevant issue, since the evaluation of different aspects of the intervention, by providing objective indicators, sets up the foundations for a control and reduction of miss rates in colon cancer detection. We proposed an approach of quality assessment by means of computer vision-based techniques which is underpinned by: 1) The automatic suggestion of potential candidates of colon cancer during the intervention time, 2) The automatic computation of objective quality metrics after the intervention, and 3) The use of eye-tracking information in order to provide metrics to evaluate the skills of the physician both in the visual search and in the endoscope manoeuvering. Preliminary results showed the suitability of such techniques for polyp detection and experts vs. trainees discrimination. The deeper analysis of quality metrics and their correlation is devoted to a further piece of research whose study our team is carrying out currently.

ACKNOWLEDGEMENTS

We would like to acknowledge the collaboration of M.D Steven Patchett and M.D Hugh Mulcahy, from the Royal College of Physicians of Ireland, for their collaboration in the clinical study of this work. This research is supported by Enterprise Ireland contracts PC-2006-038 and CFTD-2006-216.

REFERENCES

American Cancer Society (2007). *Cancer Facts and Figure*. American CancerSociety.

Chaudhuri, B. and Sarkar, N. (1995). Texture segmentation using fractal dimension. *IEEE Trans. Pattern Anal. and Mach. Intel.*, 17:72–77.

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

Eddy, D. (1990). Screening for colorectal cancer. *Annals of Internal Medicine*, 113:373–384.

Fireman, Z., Glukhovsky, A., et al. (2002). Wireless capsule endoscopy. *Israel Medical Association Journal*, 4:717719.

Gokturk, S. and Tomasi, C. o. (2001). A statistical 3-d pattern processing method for computer-aided detection of polyps in ct colonography. *IEEE Transactions on Medical Imaging*, 20(12):1251–1260.

Hai, V., Echigo, T., et al. (2006). Adaptive control of video display for diagnostic assistance by analysis of capsule endoscopic images. In *ICPR*, volume III, pages 980–983.

Hulka, J. and Reich, H. (1994). *Textbook of Laparoscopy*. W.B. Saunders, Phipladelphia, PA, 2 edition.

Hunter, J. and Sackier, J. (1993). *Minimally invasive surgery*. McGraw-Hill, New York.

Iddan, G. and Meron, G. e. a. (2000). Wireless capsule endoscopy. *Nature*, 405:4–7.

Karkanis, S., Iakovidis, D., et al. (2001). Detection of lesions in endoscopic video using textural descriptors on wavelet domain supported by artificial neural network architectures. In *ICIP*, volume II, pages 833–836.

Kato, H. and Baron, J. (2003). *Electronic Videoendoscopy*. Jpapan: Harwood Academic Publisher.

Kudo, S. and Kashida, H. e. a. (2000). Colonoscopic diagnosisi and management of nonpolypod and early colorectal cancer. *World Journal of Surgery*, 24(9):1081–1090.

Li, P., Chan, K., et al. (2005). Learning a multi-size patch-based hybrid kernel machine ensemble for abnormal region detection in colonoscopic images. In *CVPR*, volume II, pages 670–675.

Liang, J., Higgins, W., et al. (2004). Introduction to the special section on virtual endoscopy. *IEEE Transactions on Medical Imaging*, 23(11):1333–1334.

Ministry of Health Singapore (1990). *MOH Annual Report*. Ministry of Health Singapore.

Nagasako, K., Fijimori, Y., et al. (1998). *Atlas of gastroenterologic endoscopy by High-Resolution Video-endoscopy*. Igaku-Shoin Ltd. Tokyo.

Nagata, S., Tanaka, K., et al. (2000). Pit pattern diagnosis of early colorectal carcinoma by magnifying colonoscopy: clinical and histological implications. *International Journal of Oncology*, 16(5):927–934.

O'Brien, M., Winawer, S., et al. (1990). The national polyp study: Patient and polyp characteristics associated with high-grade dysplasia in colorectoral adenomas. *Gastroenterology*, 98:371–379.

Pabby, A., Schoen, R., et al. (2005). Analysis of colorectal cancer occurence during surveillance colonoscopy in the dietary prevention trial. *Gastrointestinal endoscopy*, 61(3):385–391.

Paris Workshop Participants (2003). The Paris endospcopic classification of superficial neoplastic lesions. *Gastrointestinal Endoscopy (Supplement)*, 58(6):3–23.

Paschos, G. (2001). Perceptually uniform color spaces for color texture analysis: empirical evaluation. *IEEE Trans. Image Processing*, 10:932–937.

Rembacken, B., Fuji, T., et al. (2000). Flat and depressed colonic neoplasms: A pospective study of 1000 colonoscopies in the uk. *Lancet*, 17:1157–1167.

Saitoh, Y., Waxman, A., et al. (2001). Prevalence and distinctive biologic features of flat colorectal adenomas in a north american population. *Gastroenterology*, 120:1657–1665.

Srivastava, S., Rodriguez, J., et al. (2005). Analysis of confocal microendoscope images for automatic detection of ovarian cancer. In *ICIP*, volume I, pages 1113–1116.

Swain, M. and Ballard, D. (1991). Color indexing. *Journal of computer vision*, 7(1):11–32.

Thoeni, R. and Laufer, I. (1994). *Textbook of gastrointestinal radiology*, chapter Polyps and Cancer, page 1160. Philadelphia, PA: Sanders.

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

U.S. Department of Health and Human Services (2003). Cdc, the importance of prevention and early detection cdc. http://www.cdc.gov/cancer/colorectal/.

Vapnik, V. (1995). *The nature of statistical learning theory*. Springer.

Vilariño, F. (2006). *A machine learning approach for intestinal motility assessment with capsule endoscopy*. PhD thesis, Dept. Computer Science, Universitat Autonoma de Barcelona, Spain.

Vilariño, F., Lacey, G., et al. (2007). Automatic labeling of colonoscopy video for cancer detection. *LNCS*, 4477:290–297.

Vilariño, F., Spyridonos, P., et al. (2006). Automatic detection of intestinal juices in wireless capsule video endoscopy. In *Pattern Recognition, ICPR*, volume 4, pages 719– 722.

Vilariño, F., Spyridonos, P., et al. (2009). Intestinal motility assessment with video capsule endoscopy: Automatic annotation of phasic intestinal contractions. *IEEE Trans. on Medical Imaging (in press)*.

Winawer, S., Fletcher, R., et al. (1997). Colorectal cancer screening: Clinical guideliness and rationale. *Gastroenterology*, 112:594–642.

Winawer, S., Zauber, A., et al. (1993). Prevention of colorectal cancer by colonoscopy polypectomy. *New England Journal of Medicine*, 329:977–981.

Won, C., Park, D., et al. (2002). Efficient use of MPEG-7 edge histogram descriptor. *ETRI*, 24(1):23–30.

Zheng, M., Krishnan, S., et al. (2005). A fusion-based clinical decision support for disease diagnosis from endoscopic images. *Computers in Biology and Medicine*, 35(3):259–274.