AUTOMATED SYSTEM FOR QUANTIFYING THE LEVEL OF PREPARATION IN COLONOSCOPY

Leticia Angulo-Rodríguez, Xuexin Gao, Dobromir Filip, Christopher N. Andrews and Martin P. Mintchev

***Abstract**: Colonoscopy is the gold standard method for the diagnosis of colorectal cancer (CRC). It detects the first clinical manifestation of CRC, known as polyps.*

One night prior to a colonoscopy procedure, patients are instructed to take laxative agents in order to completely cleanse the colon. This process is called bowel preparation. Contemporary sensitivity of colonoscopy for detecting polyps of a size larger than 10 mm is 98% with the limitation in detection mainly due to poor visualization related to inadequate bowel preparation.

Unfortunately, there is not yet a metric (formally recommended by means of guidelines) for the quantification of bowel preparation. Scales used nowadays are not objective, because generally colonoscopists estimate the level of cleanliness after the conclusion of the colonoscopic test.

This limitation leads to the formalization of the present study, which focuses on the development of a novel cleansing evaluation system for bowel preparation and the assessment of its clinical efficacy. The proposed system consists of a computer-based tool that can automatically measure the quantity of stool and waste matter existing within the patient during a colonoscopy procedure. As these metrics can be obtained automatically, the proposed method can lead to future quality control in daily medical practice. Furthermore, it can be used to create best practice standards for colonoscopy training or as part of medical skill evaluation.

***Keywords**: Colonoscopy; Colon preparation; Efficacy; Quality measurement metrics; Video segmentation*

***ACM Classification Keywords**: A.0 General Literature - Conference proceedings; J.3. Life and Medical Sciences*

Introduction

A. Colonoscopy and colon cancer

Colonoscopy is a procedure that allows real-time visualization of the colon. Additionally, it enables polypectomy (removal of polyps), tissue biopsy and the usage of instruments that can be inserted within the colonoscope [1].

Colonoscopy can assist doctors to diagnose some diseases with a variety of symptoms, including colonic bleeding, abdominal pain, and weight loss [2]. Most importantly, colonoscopy is now considered the “gold standard” for the detection of colorectal cancer (CRC) [3]. Based on the fact that CRC is the second most common cancer in the world causing more than 600 000 deaths annually, colonoscopy has become one of the most commonly performed medical procedures in the United States [4].

Early evaluation by means of colonoscopy is related to the decrease of CRC incidences [5]. This was demonstrated by the National Polyp Study, the objective of which was the long term monitoring of patients with polypectomy. Published results from this study demonstrated a reduction in the incidence of CRC ranging from 76 to 90% [6]. After this publication supported the use of colonoscopy as the major screening method for CRC, the demand for this procedure increased significantly. In 2004 alone, 22 million colonoscopies were performed, resulting in an increase of 8 million compared to the previous year [7].

B. Colon preparation for colonoscopy

Bowel preparation is a very important stage in the colonoscopy procedure, since it cleanses the colon walls from fecal matter for optimal visualization [8]. Poor bowel preparation inhibits the detection of small and large polyps. In such situations clinicians often clean unclear sections of the colon using the water source that is embedded in the colonoscopy system. Such an approach is time consuming and leads to the prolongation of intubation, withdrawal times and sedation, causing an increased risk for the patient. In addition, there is an immediate impact on the associated costs, especially when the examination has to be repeated [9]. Thus, poor preparation is one of the biggest obstacles to colonoscopy effectiveness. Figure 1 demonstrates a clean colon (left) in comparison to colonic segments with poor bowel preparation (middle and right). In the poorly prepared colon, small or even moderately sized polyps can be obscured by stool and turbid fluid.

The American Society for Gastrointestinal Endoscopy (ASGE) and the American College of Gastroenterology (ACG) have recommended that quality of bowel preparation should be registered in the final report after each individual screening [1]. Unfortunately, there is not yet a metric formally recommended by means of guidelines. This is because all three presently introduced scales for bowel preparation measurements are considered to be subjective. They depend purely on the colonoscopist’s appreciation of cleanliness [10].These 3 scales use subjective terms for grading, such as “excellent,” “good, “fair” and “poor” [1]. Although these scales have been a welcome tool for assessing quality of preparation, their subjective nature illustrates the need for a standard quality measurement. These scales are: the Aronchick, the Ottawa and the Boston scales [10].

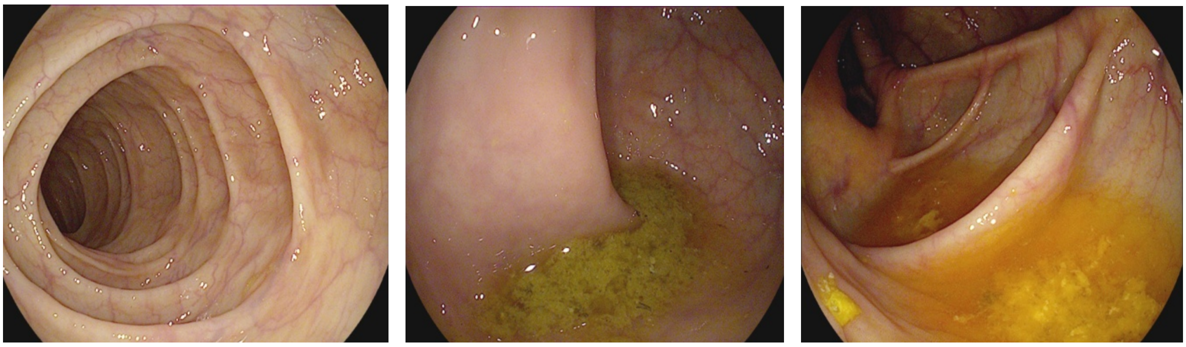

Figure 1. Left: Bowel segment with good preparation. Middle and right: Poorly prepared bowel segments.

C. Quantification scales for bowel preparation

Aronchick scale

The Aronchick Scale evaluates the quality of the colonoscopy according to the percentage of colon surface that is clearly visible [11]. The scale rates bowel preparation as “excellent” when more than 95% of the surface is seen and there is only a small amount of clear liquid. The preparation is “good” if large amounts of clear liquid cover 5% to 25% of the surface allowing the visualization of 90% of the surface. It is “fair” if there is some semisolid stool that could be washed away but still more than 90% of the surface is perceptible. Finally, the bowel preparation is “poor” when semisolid stool cannot be suctioned or washed away and it is possible to see less than 90% of the surface only [11].

Ottawa scale

The Ottawa Bowel Preparation Scale (OBPS) evaluates each segment of the colon (right colon, mid colon and rectosigmoid) regarding the presence of fecal matter. The sum of all segment scores provides a total score for bowel the preparation [10]. This scale calculates cleanliness and fluid volume independently. Cleanliness is graded from 0 to 4 in each of the three segments of the colon. The obtained 3 numbers are added to indicate total cleanliness. On the other hand, fluid quantity is an overall value for the entire colon ranging from 0 to 2 [12]. OBPS assigns a total score of 0 to an excellent preparation where no fluid is on the colon surface. Conversely, it assigns 14 overall points (12 for poor cleanliness and 2 for large fluid quantity) to a poor preparation containing large amount of fluid in each of the three segments [11].

The criteria for cleanliness of each segment in the Ottawa scale are as follows:

♣ Excellent (0): the mucosa is completely visible, there is almost no stool, although there can be fluid, but it should be clear.

♣ Good (1): there is stool and non-transparent fluid, but colon wall is still clearly visible and it does not require washing and suctioning.

♣ Fair (2): there is turbid fluid and stool hindering the visualization of the mucosa. However, mucosal detail becomes visible by suctioning but not washing.

♣ Poor (3): there is stool on the mucosa. By suctioning and washing a fair image of the mucosa can be retrieved.

♣ Inadequate (4): It is not possible to see the colon properly even after washing and suctioning.

Weighing for the presence of fluid is: 0 for small; 1 for moderate; and finally, 2 for large amount of fluid [11].

Boston Scale

The Boston Bowel Preparation Scale (BBPS) is a 10-point scale that assesses bowel preparation with the indication that it has to be done during colonoscope withdrawal, which occurs strictly after the completion of all cleansing maneuvers [8]. In this scale, similarly to the scales discussed above, subjective terms, such as “excellent”, “good”, “fair”, “poor” and “unsatisfactory” are inferred. 4-point (from 0 to 3) scoring system is applied to each of the 3 different segments of the colon. The maximum BBPS score for a perfectly clean colon without any residual liquid is 9, and the minimum BBPS score for an unprepared colon is 0. Conversely, if the procedure has to be aborted due to insufficient preparation, all the segments are assigned a score of 0 [8].

D. Aim of the paper

Despite the fact that colonoscopy is widely practiced nowadays, an established standard for measuring the quality of colonoscopies is still missing. Such system could allow for the continuous improvement of colonoscopy practices. The objective of this study is to propose a novel colon cleansing evaluation system for bowel preparation and to assess its clinical efficacy.

Methods

The core factor to be detected in the process of developing a colon cleansing evaluation system is the level of cleanliness in the colon. In the present study the evaluation of the level of cleanliness was based on the Ottawa scale. The developed algorithm to perform stool detection utilizes color recognition, which is a major approach in feature segmentation [13].

Matlab software (The MathWorks Inc., Natick, Massachusetts) was selected for the implementation of the algorithm because of its friendly interface and the advantage of having an image processing toolbox, in which image recognition by color and shape is facilitated. This algorithm was tested in a set of 13 videos, which were acquired in the McPhail Colon Cancer Screening Center, Foothills Hospital, Calgary, Alberta, Canada. The colonoscopy videos were anonymized videos without any patient and endoscopist information.

Finally, a correlation was performed between the Ottawa 5 point-scoring system and the percentage of stool in the colon retrieved by the proposed software.

A. Color detection

The elements for characterizing an image are color, shape and texture. Of the above mentioned factors, color is the most important feature to segment images [13].

Hue-Saturation-Value color model (HSV) is a method to define color according to its three basic features:

1) Hue
2) Saturation; and
3) Lightness [14].

Hue represents a specific tone of color. Saturation is the estimation of the purity of hue and is related to the intensity of the latter. When a color is completely saturated it excludes any gray from its content. Conversely, when the saturation is low, a color turns entirely into gray. The lightness component determines if the color turns lighter or darker [15].

HSV color space is shown on Figure 2. The values for the hue component range from 0° to 360°, where red is set at 0° and black at 360°. For saturation and lightness, the values range from 0 to 1 [16]. From these 3 components, hue expresses the main characteristic of a color [16]. For instance, in the present study, hue represents the highest contrast between the colon wall and the stool in comparison with the other 2 components. Once the ranges in the HSV color space have been set for a specific target object, only the pixels in the image that are within these limits are extracted.

The spectral characteristics of a camera sensor and its lighting conditions determine the level of color in an image [17]. The use of the HSV color model is suitable for the present study because it is more consistent and efficient than the Red-Green-Blue color space (RGB) while working on color detection, since the hue component remains immune to lighting behavior conditions. This means that the hue histogram, which provides the values for each one of the HSV color space components in an image, remains about the same regardless the change of the illumination level [17].

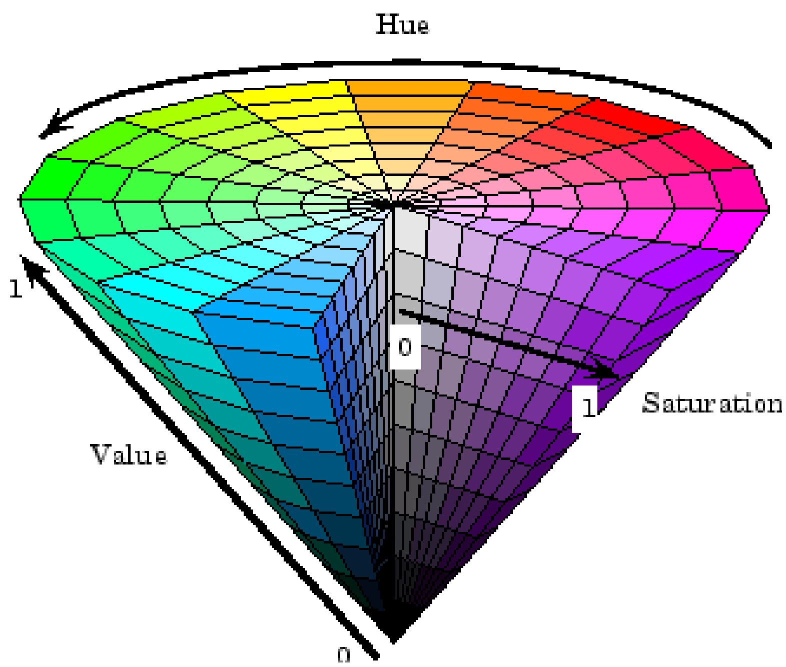

Figure 2. HSV Color Space [17].

B. Post-processing and quantification

The post-processing and quantification algorithm includes the following steps:

1. Increasing the contrast of the RGB image in order to discriminate better between the colon wall and the stool matter on it.

2. Separating the three color components of the HSV color space. As a result of that partition, hue, saturation and lightness components are separately quantified. The target color is defined by limiting the range of these component values as follows: Hmin< H <Hmax, Smin< S <Smax and Lmax< L <Lmin, where min and max are the maximum and minimum values, respectively [18]. The examined color is detected in the image if the pixel color lays within the boundaries of the HSV zone. These values should be previously identified from the HSV histograms. One histogram for each one of the components in the HSV color space was computed.

3. Every small object that has less than a certain number of pixels is removed from the image in order to avoid saturating the screen and allowing only the segmentation of areas that considerably contribute to the percentage that is displayed.

4. A morphological closing is performed on the grayscale image. The ratio of the black pixels, (which represent the pixels of stool) over the total number of pixels in the matrix is calculated. With this method the percentage output of pixels of stool is computed.

5. The perimeter of the detected area is overlaid on the original image to examine the correctness of the detection.

Figure 3 shows the output of the proposed software. At the top, the percentage of stool matter in the image is indicated. The contour around the area with poor bowel preparation is also displayed. The complete procedure is explained as a flow chart in Figure 4.

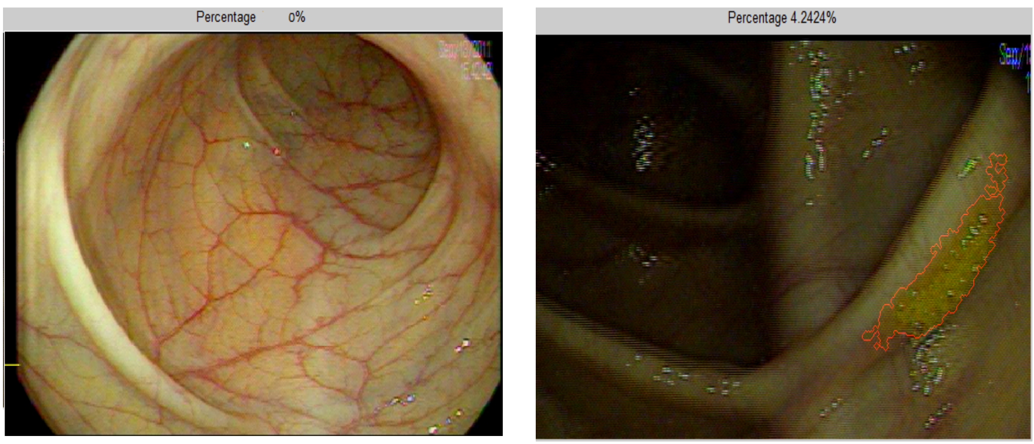

Figure 3. Output of the software. Left: A clean colon where no percentage of stool was recognized. Right: Colon with stool where the target area is highlighted.

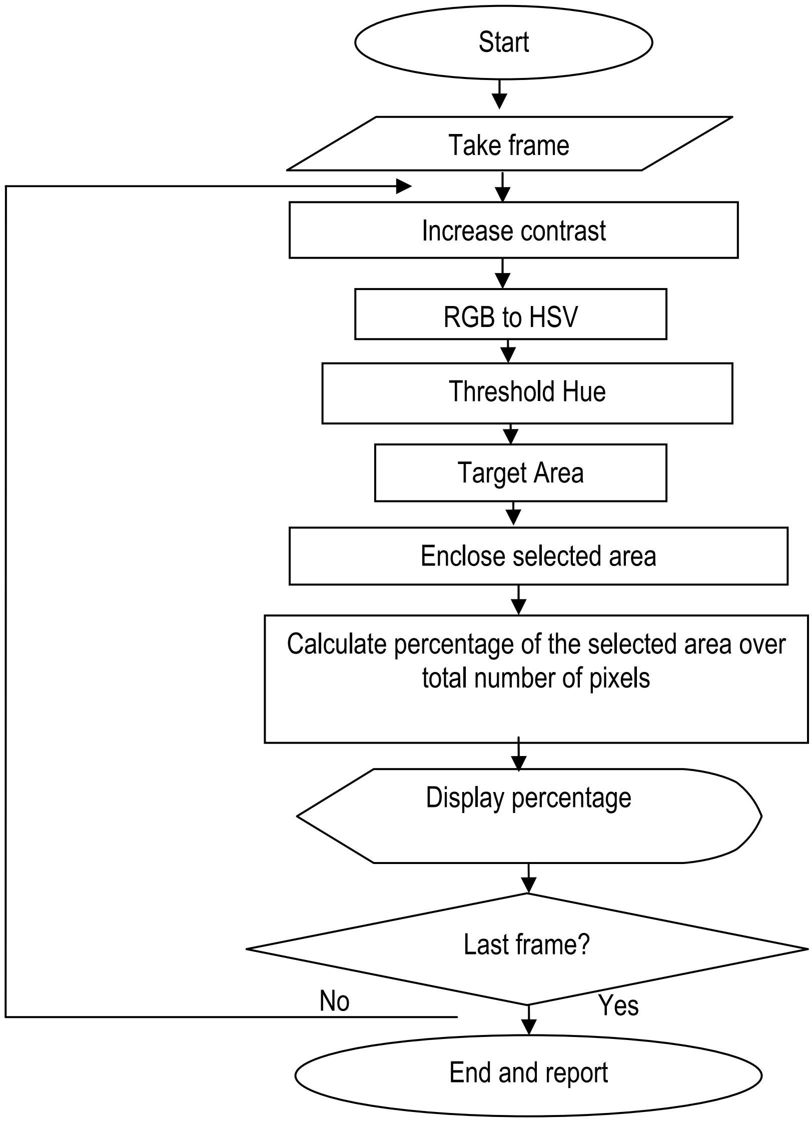

Figure 4. Algorithm for stool detection

C. Ottawa scale integration

Thirteen complete colonoscopy videos have been graded by a professionally licensed academic endoscopist, and each of the 3 segments in the colon (right colon, mid colon and rectosigmoid) was marked according to the Ottawa scale. For the preliminary analysis the value for the quantity of fluid was not taken into consideration, since the fluid recognition stage has not been yet implemented in the proposed software.

All videos were analyzed in Matlab, at 30 frames per second. All of them had different time durations, ranging from 4 to 8 minutes. The percentage of stool in each of the frames was retrieved from each of the colon segments. From the collected data, the mean of the percentages of stool and the standard deviation for each of the three colonic segments in the 13 colonoscopic videos were calculated.

Results

Table 1 lists the results utilized to perform the correlation between the Ottawa scores given for each one of the segments in the colon and the mean value of the percentages of pixels identified as stool over the total pixels in the image for each segment in the 13 videos. The values are plotted in Figure 5, which demonstrates the relationship between the Ottawa scale scores and the percentages retrieved from the proposed system. After performing Pearson correlation analysis [19], the obtained correlation coefficient r was equal to 0.61 which confirmed that there was a statistically significant correlation between the Ottawa scale score given by the endoscopist and the percentage output of the software ($p<0.01$). Alternatively, the coefficient of determination (r^2) was obtained by squaring Pearson correlation coefficient. It calculates the linear relationship strength between two variables [20]. In this study, the coefficient of determination estimated that 37% of the variances of either variable (Ottawa scale scores and percentages of stool as determined by the proposed algorithm) are shared between one another.

Table 1. Comparison between the proposed system results and the corresponding Ottawa scores. The percentages represent the per-video average of the ratios of pixels identified as stool over the total number of pixels per frame. The standard deviation of these percentages per video is also listed.

Segment	Right			Mid			Rectosigmoid		
Video #	Ottawa score	% of stool	Standard Deviation	Ottawa Score	% of stool	Standard Deviation	Ottawa score	% of stool	Standard Deviation
1	1	2.879	3.406	2	22.913	34.456	2	13.743	20.200
2	2	3.004	7.279	3	12.434	21.211	3	37.656	36.917
3	3	27.226	32.404	2	22.271	28.398	2	30.891	38.104
4	3	36.028	29.084	3	36.110	32.423	2	37.289	30.765
5	3	39.683	30.648	1	18.530	26.308	2	20.496	33.054
6	1	2.609	8.174	1	4.209	12.638	1	5.449	15.121
7	1	20.761	17.325	2	5.169	11.220	1	1.420	3.878
8	2	38.260	34.029	0	11.344	20.069	1	3.522	13.671
9	1	18.812	30.531	2	17.530	27.703	1	5.267	18.404
10	2	40.112	26.426	2	20.196	27.842	1	2.231	9.393
11	2	35.074	36.736	1	25.571	32.370	2	33.520	32.678
12	1	12.957	12.521	1	24.967	30.424	1	12.057	20.579
13	0	10.950	11.503	1	19.738	29.246	1	10.450	25.965

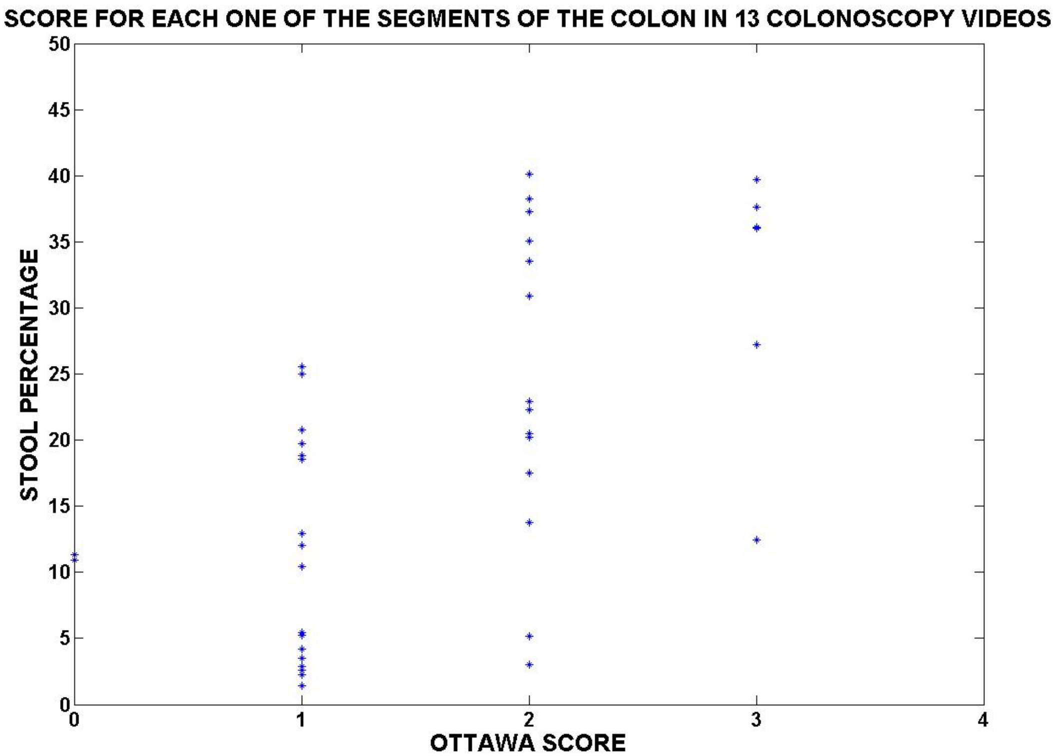

Figure 5. Graphical representation of the statistically significant correlation ($p<0.01$) between Ottawa scale scores and the mean of the stool percentages in each segment of the colon retrieved from the proposed system in 13 videos.

Discussion

The present study describes the development of a system that measures the amount of stool matter in the colon as an approach to quantify the quality of colonoscopy. The main part of the algorithm is color-based detection by hue color space.

In order to set the percentage range that would determine whether the quality of bowel preparation is “Excellent”, “Good”, “Fair”, “Poor” or “Inadequate”, the software should be improved in order to get more consistent results and have less variability in the group of points that correspond to each value in the Ottawa scale. Therefore, the reliability of the software cannot be presently evaluated, and it will be necessary to include some improvements to the code in order to get a greater determination coefficient in order to confidently conclude that the results are close to those in the Ottawa scale.

There are some potential factors that have an impact on the number of target pixels detected. Among them is the type of consumed laxative, since this might determine the color of the stool, ranging from yellow shade color to a green shade color. Furthermore, in order to agree with the Ottawa scale, the amount of fluid within the colon should be detected separately and automatically displayed by the system as a percentage output. Finally, an additional factor that affects the measurement is the quantity of blurry frames that can be identified in the video. These frames hamper the correct detection of stool matter in the colon. Future work should consist of automatically removing unclear image frames, especially when cleaning maneuvers are being conducted. So far, it has been thought that edge analysis could be an appropriate method to discard blurry frames, based on the assumption that the latter would not contain edges.

In summary, our approach for evaluating the quality of bowel preparation accounts for 3 factors:

1. Color of stool
2. Amount of fluid
3. Quantity of blurry images

Since these factors are uncontrollable, it is important to implement a calibration system that can distinguish between different color tonalities of fluid, stool matter and colon walls and can automatically adjust the HSV threshold values accordingly.

Finally, in order to have a general evaluation of colonoscopy procedures, in a future stage the proposed system has to be incorporated with other useful tools that measure additional parameters of importance in colonoscopy, such as the colonoscope withdrawal time and the polyp detection rate. Thus, the proposed software could provide complete quantitative results about colonoscopy quality and would be able to document them in a database. Overall future clinical testing of the developed system should include comparing polyp detection rate in colonoscopy utilizing colon cleanliness evaluation.

The proposed system can lead to future quality control in daily medical practice and can be used to create best practice standards, since colonoscopy quality metrics would be automatically obtained. In addition, the software could also be part of training programs for colonoscopists that would allow for an easier continuing professional development and competence maintenance, as well as for becoming a tool for medical skills evaluation. Finally, it could be pivotal for extracting information from already documented colonoscopy studies to conduct requirement analyses for future improvements of the procedure.

Conclusions

Quality is a very important issue in medical practice. Having guidelines that can help in the establishment of quality standards can significantly improve health care delivery service.

Colonoscopy has become the gold standard for the diagnosis, monitoring and therapeutic treatment of colorectal cancer, yet its continuous improvement is hampered by the lack of quantitative standards for best practices.

In this paper, the development of an objective quantification of bowel preparation stage during colonoscopic procedures is proposed, since the perception of bowel cleanliness may vary from one clinician to another. A detailed revision of the 3 scales currently available was performed for the purpose of establishing quantitative threshold levels for bowel cleanliness. Stool detection algorithm using thresholding in the hue component of HSV space and additional imaging processing procedures was proposed.

Based on the performance of the proposed evaluation system in processing 13 colonoscopic videos, correlation was established between the system and the Ottawa Bowel Preparation Scale.

Bibliography

[1] Y. Hazewinkel and E. Dekker, "Colonoscopy: basic principles and novel techniques", *Nat. Rev. Gastroenterol. Hepatol.,* vol. 8, pp. 554-564, Sep 6, 2011.

[2] National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), "**Colonoscopy**". January, 2010. Retrieved on March 29, 2012. Available: http://digestive.niddk.nhi.gov/ddiseases/pubs/colonoscopy/

[3] American Society of Colonoscopy, "*Colorectal cancer facts & figures 2011-2013*", American Cancer Society, Atlanta, GA, 2011. Available: http://www.cancer.org/Research/CancerFactsFigures/ColorectalCancerFactsFigures/colorectal-cancer-facts-figures-2011-2013-page

[4] World Health Organization, "Cancer", February, 2012. Retrieved on March 29, 2012. Available: http://www.who.int/mediacentre/factsheets/fs297/en/

[5] G. M. Eisen, "Building a better colonoscope?", *Gastrointest. Endosc.,* vol. 68, pp. 711-712, 10, 2008.

[6] R. L. Barclay, J. J. Vicari, A. S. Doughty, J. F. Johanson and R. L. Greenlaw, "Colonoscopic Withdrawal Times and Adenoma Detection during Screening Colonoscopy", *N. Engl. J. Med.,* vol. 355, pp. 2533-2541, 2006.

[7] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward and D. Forman, "Global cancer statistics", *Cancer Journal for Clinicians,* vol. 61, pp. 69-90, Mar, 2011.

[8] E. J. Lai, A. H. Calderwood, G. Doros, O. K. Fix and B. C. Jacobson, "The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research", *Gastrointest. Endosc.,* vol. 69, pp. 620-625, Mar, 2009.

[9] D. K. Rex, J. L. Petrini, T. H. Baron, A. Chak, J. Cohen, S. E. Deal, B. Hoffman, B. C. Jacobson, K. Mergener, B. T. Petersen, M. A. Safdi, D. O. Faigel, I. M. Pike and ASGE/ACG Taskforce on Quality in Endoscopy, "Quality indicators for colonoscopy", *Am. J. Gastroenterol.,* vol. 101, pp. 873-885, Apr, 2006.

[10] S. Landreneau and J. Di Palma, "Update on Preparation for Colonoscopy", *Curr.Gastroenterol. Rep.,* vol. 12, pp. 366-373, 2010.

[11] Y. S. Choi, J. P. Suh, J. K. Kim, I. T. Lee, E. G. Youk, D. S. Lee, S. K. Do and D. H. Lee, "Magnesium citrate with a single dose of sodium phosphate for colonoscopy bowel preparation", *World J. Gastroenterol.,* vol. 17, pp. 242-248, Jan 14, 2011.

[12] A. Rostom and E. Jolicoeur, "Validation of a new scale for the assessment of bowel preparation quality", *Gastrointest.Endosc.,* vol. 59, pp. 482-486, 4, 2004.

[13] M. Abdellatif, "Effect of color pre-processing on color-based object detection", in *SICE Annual Conference,* pp. 1124-1129, 2008.

[14] S. Li and G. Guo, "The application of improved HSV color space model in image processing", in *Future Computer and Communication (ICFCC),* vol.2, pp.10-13, 2010.

[15] R. T. Stevens, *Computer Graphics Dictionary.* Massachusetts, USA: Charles River Media, 2002, p.112.

[16] L. J. Ubong, C.S. Tehand G. W. Ng, "A comparison of RGB and HSI color segmentation in real - time video images: A preliminary study on road sign detection", in *Information Technology. International Symposium on,* 2008, pp. 1-6.

[17] M. Riaz, G. Kang, Y. Kim, S. Pan and J. Park, "Efficient image retrieval using adaptive segmentation of HSV color space", in *Computational Sciences and its Applications. ICCSA '08. International Conference on,* 2008, pp. 491-496.

[18] G. D. Finlayson, "Color in perspective", *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 18, pp. 1034-1038, 1996.

[19] J. D. Jobson, *Applied Multivariate Data Analysis.* Virginia, USA: Springer, 1999.

[20] C. D. McDaniel and R. H. Gates, Marketing Research Essentials. Ohio, USA: South Western College Company, 1998.

Authors Information

L. Angulo-Rodríguez *is with the Department of Biomedical Engineering, University of Calgary, Calgary, AB, T2N 1N4.*

X. Gao *is with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, ABT2N 1N4, Canada.*

D. Filip *is with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada.*

C. N. Andrews *is with the Department of Medicine, University of Calgary, Calgary, AB, T2N 1N4, Canada.*

M. P. Mintchev *is with the Department of Electrical and Computer Engineering, and the Department of Medicine, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada (e-mail: mintchev@enel.ucalgary.ca).*