Adaptable image cuts for motility inspection using WCE

Michal Drozdzal[1,2], Santi Seguí[1,2], Jordi Vitrià[1,2], Carolina Malagelada[3], Fernando Azpiroz[3], Petia Radeva[1,2]

[1]*Dept. Matemàtica Aplicada i Anàlisis, Universitat de Barcelona, Barcelona, Spain*
[2]*Computer Vision Center (CVC), Universitat Autònoma de Barcelona, Barcelona, Spain*
[3]*Digestive System Research Unit, Hospital Vall d'Hebron, Barcelona, Spain*

Abstract

The Wireless Capsule Endoscopy (WCE) technology allows the visualization of the whole small intestine tract. Since the capsule is freely moving, mainly by the means of peristalsis, the data acquired during the study gives a lot of information about the intestinal motility. However, due to: 1) huge amount of frames, 2) complex intestinal scene appearance and 3) intestinal dynamics that make difficult the visualization of the small intestine physiological phenomena, the analysis of the WCE data requires computer-aided systems to speed up the analysis. In this paper, we propose an efficient algorithm for building a novel representation of the WCE video data, optimal for motility analysis and inspection. The algorithm transforms the 3D video data into 2D longitudinal view by choosing the most informative, from the intestinal motility point of view, part of each frame. This step maximizes the lumen visibility in its longitudinal extension. The task of finding "the best longitudinal view" has been defined as a cost function optimization problem which global minimum is obtained by using Dynamic Programming. Validation on both synthetic data and WCE data shows that the adaptive longitudinal view is a good alternative to the traditional motility analysis done by video analysis. The proposed novel data representation a new, holistic insight into the small intestine motility, allowing to easily define and analyze motility events that are difficult to spot by analyzing WCE video. Moreover, the visual inspection of small intestine motility is 4 times faster then by means of video skimming of the WCE.

Preprint submitted to Computerized Medical Imaging and Graphics September 21, 2012

Keywords: Intestinal motility visualization, Wireless Capsule Endoscopy, Dynamic Programming

1. Introduction

Small bowel (also called small intestine) is a part of Gastro-Intestinal (GI) tract connecting stomach with large intestine. The length of the small intestine in an adult human is variable and depending on the conditions can measure from 3 to 8 meters. The main function of the small bowel is the digestion and absorbtion of nutrients and minerals found in the food. In order to do so, the small intestine pushes the food through by means of a physiological mechanism called motility. In general, the intestinal motility can be divided into two categories [1]: 1) Peristalsis - synchronized movement of the intestinal wall responsible for moving the food in one direction. 2) Independent contractions - unsynchronized movement of the intestinal wall where the muscles squeeze more or less independently of each other; this has the effect of mixing the contents but not moving them up or down.

The accurate assessment of small bowel motility constitutes one of the most relevant issues in gastroenterology. Intestinal motility dysfunction appears when the organ losses its ability to coordinate muscular activity, manifesting the abnormal contractile activity (e.g. spasms or intestinal paralysis) [2]. In a broad sense, any alteration in the transit of foods and secretions into the small intestine tube may be considered a motility disorder[1] [2].

Currently, the main source of information, and the only one which leads to a diagnosis of small intestine motility disorders is manometry [3]. The diagnosis is based on the change of intestinal wall pressure in a fixed part of the small intestine. However, this technique has several drawbacks: 1) it is highly invasive causing discomfort to the patient; 2) it does not offer the visualization of the small intestine; 3) only a portion of the small bowel can be evaluated and 4) the performance of this test is limited to referral centers around the world due to its complexity and the difficulty in the interpretation of the results.

A significant technical breakthrough in GI tract analysis was achieved in 2000 when Wireless Capsule Endoscopy (WCE) was presented [4]. From the

[1]It is important to note that the motility dysfunctions are just one part of all digestive problems (like: bleeding, polyps or tumors)

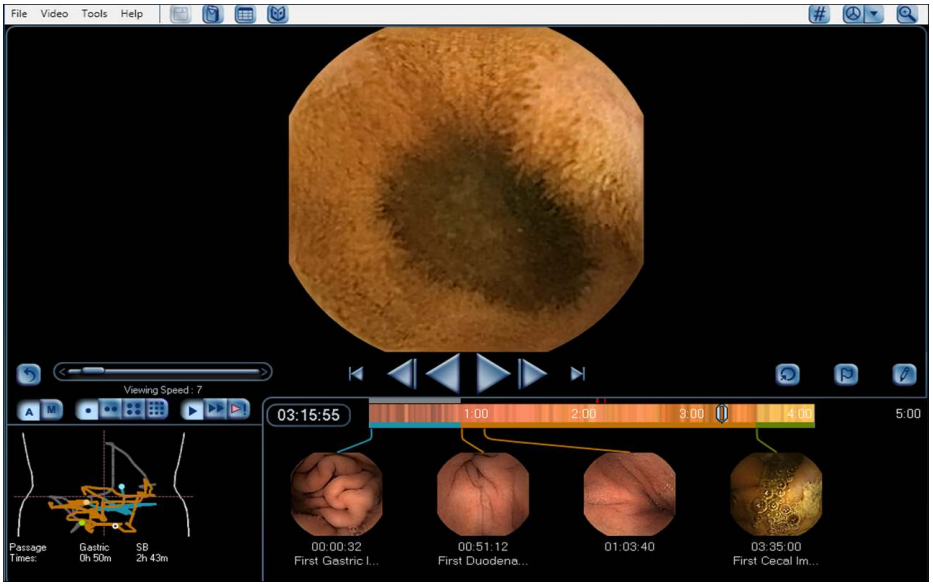


Figure 1: Rapid Reader interface by GivenImaging [5]. The interface presents: (up) video filed presenting frame view, (bottom-left) an approximation of the capsule position inside the GI tract and (bottom-right) color bar representing the whole video, each element in the stripe represents a mean color intensity in one-minute frame period.

first moment, the medical community received WCE with high enthusiasm since it allowed to visualize the intestine content, shape and function. Thus, the device has expanded the diagnostic capabilities of small bowel based on the view of the whole organ. It contains a camera and a full electronic set which allows the radio frequency emission of a video movie in real time, acquiring 2 frames per second. The video, showing the whole trip of the pill along the GI tract, is stored into an external hard drive which is carried by the patient. The videos can have a duration from 1h up to 8h, what means that the capsule captures a total of 7200 up to 60000 images. The video can be visualized by using the software provided by the capsule manufacture. The software interface for the analysis of WCE data is shown in Fig. 1, where three types of information are presented: 1) video filed presenting frame view (Fig. 1 - top), 2) an approximation of the capsule position inside the GI tract (Fig. 1 - bottom left) and 3) a color bar representing the whole video, where each column in the stripe represents a mean color intensity in one-minute frame period (Fig. 1 - bottom right).

The WCE videos contains very valious information about small intestine

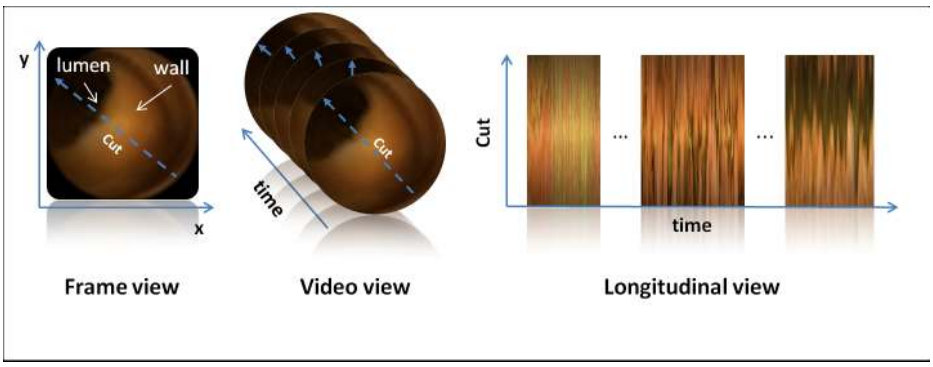
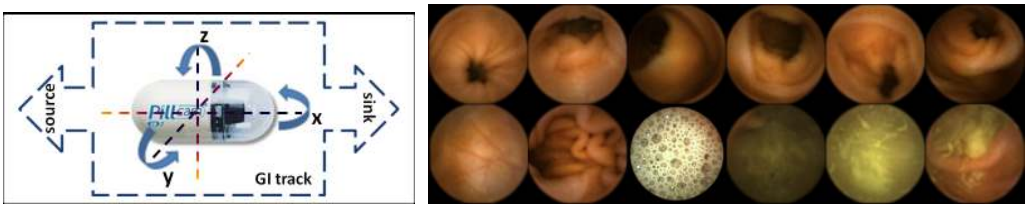


Figure 2: Different views that can be obtained from the WCE video, from left to right: 1) single frame view, 2) video view and 3) longitudinal view of adaptive cuts.



(a) (b)

Figure 3: a) Scheme representing possible camera movements. b) Different frames acquired by the WCE capsule, first row presents different lumen position due to camera/intestine movement, second row presents (from left to right) two examples of intestinal wall and four examples of intestinal content.

and its motility. Using WCE signal, two types of view for data analysis can be constructed: 1) frame view and 2) longitudinal view. The frame view gives information about a slice of intestine (e.g. lumen/wall appearance, see Fig. 2(left)). While the longitudinal view, usually perpendicular to the intestinal tube, shows the desired segment of the video (e.g. lumen/wall change in time - motility, see Fig. 2(right)). One way of obtaining longitudinal view is by "cutting"[2] a line of pixels in the sequence of frames. Note that these cuts do not have to be fixed at the same angle/position, they can adapt to the lumen position (e. g. compensating in a way the camera rotation).

[2]Throughout the paper we will refer with the word "cut" to a straight line of pixels that passes through the center of the frame (diameter)

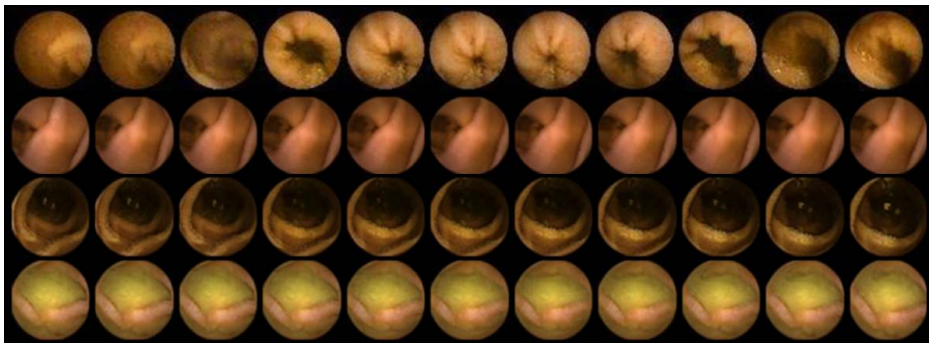


Figure 4: Some examples of intestinal events. From the top to the bottom: 1) contraction - movement of intestinal wall/lumen, 2) static - paralyzed intestine, 3) tunnel - paralyzed intestine with open lumen and 4) turbid - lumen/wall occluded by intestinal content.

Longitudinal view represents substantial benefit in analyzing the intestinal motility by using WCE, the following characteristics should be considered:

- **Complex appearance of intestinal events:** The camera is freely moving inside the intestine (upwards, backwards and with respect to all three rotational axis, (see scheme represented on Fig. 3(a)). The capsule is traveling through the small intestine (from the beginning of small intestine - source, to the end of small intestine - sink) by means of: a) the small intestine motor activity and b) the gravity. These are the only two factors that control the capsule movement, velocity and direction. Generally, the capsule moves forward into the sink, but it is also possible that, for some period of time, the capsule travels backward. Hence, it is very difficult (or even impossible) to determine the exact capsule position or orientation. The image from WCE can show the lumen, whole or only a part of it, or the intestinal wall. Moreover, often the field of view of the WCE is partially or completely occluded by intestinal content (intestinal juices and food in digestion, see Fig. 3(b)).

- **Complex interpretation of intestinal events:** The intestinal movement can be characterized into four categories of events: 1) "contraction" - movement of intestinal wall/lumen, 2) "static" - paralyzed intestine with closed lumen, 3) "tunnel" - paralyzed intestine with open lumen and 4) "turbid" - lumen/wall occluded by intestinal content.

Fig. 4 presents some examples of intestinal events seen by the WCE capsule. As it can be observed, the contraction pattern is visible as the open-closed-open lumen pattern. When no movement of intestinal wall (and lumen) is visible, the intestine is not moving and the capsule is "paralyzed" inside the intestine. When the intestine is paralyzed and the lumen is opened then a tunnel sequence is seen. Although the importance of these events in motility disorders diagnosis has been shown ([6], [7]), the exact medical importance of each of the event and the relations between them is still an open issue.

- **Large number of images:** Having available significant amount of data allows detailed description of physiological events. Meanwhile, a huge amount of the data (up to 60000 frames) requires a long time (up to several hours) for video visualization and for diagnosing a study by the physician.

The problem of manual inspection of the high amount of intestinal images has attracted the attention of researchers from medical imaging community. The main effort has been put on the lesions detection. From the variety of the work dealing with the problem of lesions detection and characterization, we can distinguish, to list just a few: 1) polyp detection [8], [9], 2) bleeding [10], [11], [12], 3) abnormal regions (tumors) [13], 4) ulcers [14] and 5) general pathologies [15], [16], [17].

In the area of intestinal motility some work on detection and characterization of specific events of intestinal motility has been done, such as intestinal contractions detection [18], [19] and [20]. Besides, in [6] and [7] the authors have approached the evaluation of motility from WCE images. Those two papers are an extension of works on specific events characterization making the analysis of the small intestine motility more complex. Both methods, first, extract several motility descriptors (each descriptor representing a specific intestinal event) from WCE videos, and second, combine the extracted characteristics and draw a conclusion on small intestine motility. However, as both methods use a multidimensional and non linear classifiers, the interpretation of the results by the physicians is complicated. It is hard to get an intuition about what is happening inside the small intestine.

Regarding the problem of WCE video visualization techniques, the researchers have focused their efforts on video compaction, resulting in eliminating/compacting similar frames [21], [22], [23] and/or by applying variable

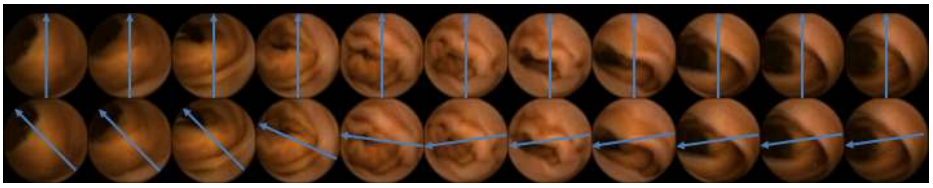


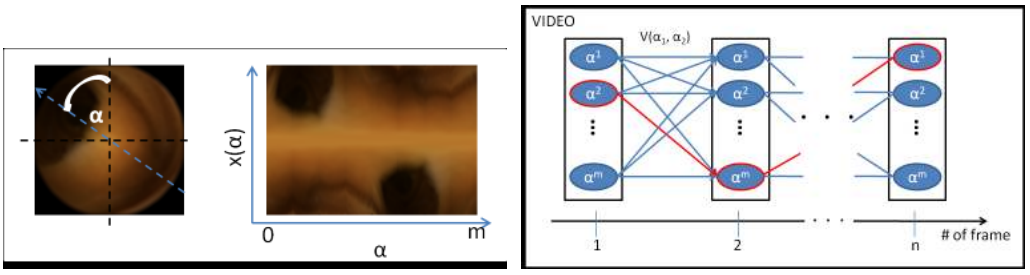
Figure 5: Example of fixed cuts (first row) and adaptable cuts (second row).

sampling rate at acquisition step [24]. Video compaction, by elimination of the frames of the video in which the capsule is paralyzed, permits to reduce the experts skimming process reducing the time needed for video visualization. However, only one work addresses the intestinal motility [24]. In this work, the proposed video visualization removes the central part of the frame, focusing on the wall visualization. This leads that the presence of the lumen information is not considered in the video visualization.

To the best of our knowledge, no other works have been proposed to improve the visualization of the motility and at the moment there does not exist any efficient and compact visualization technique suitable for small intestine motility analysis using the WCE videos.

In this paper, a novel technique, called Adaptive Cut (Acut) longitudinal view, for visualization of WCE videos is presented. The paper addresses the problem of choosing the optimal sequence of cuts through the consecutive frames of the video in order to optimally represent the intestinal motility. Due to the free movement of the capsule inside the intestine, the lumen is not always presented in the center of the frame. Therefore, a straight-forward approach to cut through a fixed angle (e.g. the vertical line) of the pixels of the image will lead to losing the motility information (for an example of a fixed cut see Fig. 5(top)). The proposed methodology is based on an optimization problem that maximizes the probability of passing through the lumen in order to preserve motility information in the new visualization scheme (for an example of adaptive cut see Fig. 5(bottom)). The main advantages of the proposed method are:

- It reduces the information provided by the WCE video, transforming 3D video signal into 2D image and thus permits to evaluate, by an expert, the intestinal motility, in a quick view.
- It preserves the motility information by applying adaptive cuts that maintain, where possible, the lumen/wall information.



(a) (b)

Figure 6: a) An illustration of the cut angle α_i. Left ith image, right all possible cuts through the original image. b) A graph illustration of Dynamic Programming problem in WCE video.

The paper is divided as follows: In Section 2, the algorithm for adaptive image cuts (Acut) based on dynamic programing is presented. Section 3 shows experimental results and validation. Finally, Section 4 concludes the paper.

2. Adaptive cuts for longitudinal view

Let us consider a longitudinal view mosaic that is obtained by cutting a stripe of pixels from consecutive frames of a video. In this section, the algorithm for adaptive longitudinal cuts is presented. The word adaptive, in this context, means that the algorithm, for each video, searches for an optimal path through all video frames, considering for each frame a set of all possible cuts.

The WCE video can be seen as a chain of n frames. Each frame i has m possible cuts $\alpha_i \in \mathbf{\Omega} = \{1, \cdots, m\}$, where the angle α_i denotes the angle between the vertical line passing through the center of the frame and the line representing the cut (see Fig. 6(a)). Moreover, let us define the cost of passing from cut α_i in frame i to cut α_{i+1} in frame $i+1$ as $V(\alpha_i, \alpha_{i+1})$.

The problem of adaptive longitudinal view construction can be seen as an optimization problem, where two constrains are introduced: 1) the lumen visibility and 2) the smoothness of the view. The first term ensures that the cut passes through the lumen and thus ensures its visualization in the longitudinal view, while the smoothness term is important to avoid sudden changes between consecutive frames and thus maintain the interpretability of the view. This task can be reformulated as a problem of finding an optimal

path (presented in red in the graph on Fig. 6(b)). The cost of a candidate solution $(\alpha_1, \cdots, \alpha_n)$ to the problem can be defined according to the following cost equation [25]:

$$E(\alpha_1, \cdots, \alpha_n) = \sum_{i=1}^{n} D(\alpha_i) + \sum_{i=2}^{n} V(\alpha_{i-1}, \alpha_i). \quad (1)$$

The terms $D(\alpha_i)$ are used to ensure that the cut in the ith image passes through the lumen, while the $V(\alpha_{i-1}, \alpha_i)$ ensures that the change of angles α_{i-1} and α_i is smooth (this term captures the cost of change between two consecutive frames, $i-1$ and i). The best solution is the one that passes through all video frames and has the minimal cost.

Because of the size of the WCE video (n up to 60000 frames), we propose to use Dynamic Programming (DP) in order to find the minimum of the function described by Eq. (1) and to obtain the angles for the cuts in the adaptive longitudinal view of the WCE video.

2.1. Dynamic Programming

Dynamic Programming is broadly used in discrete optimization problems. The DP finds the global optimum to the given problem. The basic idea of DP is to decompose a problem into a set of subproblems, where the original solution can be quickly solved and the subproblems can be efficiently solved in a recursive way [25]. Hence, the difference with respect to the classical recursive methods is the memoization (storing the solutions to already solved subproblems).

Let the table $B(\alpha_i)$ denote the cost of the best assignment of angle cuts to the elements from 1 to i, with the constraint that ith element has label α_i. The size of the table B is $n \times m$, where m is the cardinality of the set $\mathbf{\Omega}$ and n the number of frames. The table $B(\alpha_i)$ can be filled in increasing i by using the following recursive equations:

$$B(\alpha_1) = D(\alpha_1),$$
$$B(\alpha_i) = D(\alpha_i) + min_{\alpha_{i-1}}(B(\alpha_{i-1}) + V(\alpha_{i-1}, \alpha_i)) \quad (2)$$

The subproblems are defined as: for the first frame, $B(\alpha_1)$ is as the cost of assigning the angle $D(\alpha_1)$ to the first frame. For every other frame, $B(\alpha_i)$ is the cost of assigning the angle $D(\alpha_i)$ plus the minimal transition cost from $i-1$th to the ith frame $min_{\alpha_{i-1}}(B(\alpha_{i-1}) + V(\alpha_{i-1}, \alpha_i))$.

In order to avoid recalculating the solutions to sub-problems, a matrix T is being filled in while calculating the tables $B(\alpha_i)$. The matrix T stores optimal solutions to sub-problems (thanks to this book-keeping, each sub-problem is calculated only once). Each row of the matrix has a size of m and stores the "best way" to get to the ith solution from the $i-1$th solution. Each time a new value is added to the $B(\alpha_i)$, the matrix T is updated according to the rule:

$$T(\alpha_i) = \underset{\alpha_{i-1}}{\operatorname{argmin}}(B(\alpha_{i-1}) + V(\alpha_{i-1}, \alpha_i)) \tag{3}$$

As a result, the matrix T stores the indices of nodes through which the algorithm should pass in order to get the optimal solution to the problem. Finally, the overall solution is tracked back $\alpha_{i-1} = T(\alpha_i)$ starting at $i = n$. And as a result the sequence of optimal cuts $(\alpha_1, \cdots, \alpha_n)$ through all frames in the video is obtained. This sequence of optimal cuts can be seen as the path of minimal cost through the frames in the video.

2.2. Lumen visibility

The lumen in the WCE image is seen as a dark blob often surrounded by the intestinal wall. The image cut that passes through the intestinal lumen can be characterized in terms of mean and variance of the color intensity. In order to ensure the lumen visibility, the algorithm is looking for the cut of high variance and low mean value (mean and variance are calculated using the pixels that compose the cut). High variance σ^2 assures that the cut preserves maximal information of the frame maintaining, where possible, the lumen/wall information. Low mean value μ assures that the cut passes through the dark area of the image. Note that the dark area of the image presents a lumen with high probability.

Let $x(\alpha_i)$ denote the vector of the pixels from the image cut localized in the angle α_i and passing through the center of the image (see Fig. 6(a)), the lumen visibility cost D can be defined as follows:

$$D(\alpha_i) = 1/(\sigma(x(\alpha_i)) + 1) + \mu(x(\alpha_i)) \tag{4}$$

where it is assumed that the values of the vector $x(\alpha_i)$ are in the range $[0, 1]$.

2.3. Smoothness

In order to assure the longitudinal view smoothness, a term that controls the changes between the angles of the consecutive frames is introduced.

The smoothness is restricted by two factors: 1) angle change $V'(\alpha_{i-1}, \alpha_i) = 180^\circ - |180^\circ - |\alpha_{i-1} - \alpha_i||$ and 2) similarity between two consecutive cuts $V''(\alpha_{i-1}, \alpha_i) = \|x(\alpha_{i-1}) - x(\alpha_i)\|_2$. The final smoothness term V is defined as follows:

$$V(\alpha_{i-1}, \alpha_i) = \beta(V'(\alpha_{i-1}, \alpha_i)/\gamma_1)^2 + (1-\beta)(V''(\alpha_{i-1}, \alpha_i)/\gamma_2)^2 \qquad (5)$$

where quadratic terms in V' and V'' are introduced in order to penalize the sudden changes, γ_1, γ_1 normalization terms, and $\beta \in [0, 1]$ is a parameter controlling the weight between change of angles and similarity of cuts in consecutive frames.

2.4. Computational issues

Let m denote the number of possible cuts and n denote the number of frames for each frame. At each iteration the algorithm calculates: 1) m means of pixels in cut; 2) m variances of pixels in cut; 3) m^2 angle differences between cuts in consecutive frames and 4) m^2 similarities between cuts in consecutive frames. So, the computational complexity of the algorithm is $O(m^2 n)$.

We have implemented the algorithm in Matlab and run the code on 2.6GHz Intel Xenon machine with 16 GB of memory. The running time for one video of 9679 frames was of 1468 seconds. The memory used: 1) cost matrix storing 90 cuts for every frame of double precision 6.8 MB and 2) vector with 1 uint8 index for each frame 76 kB.

3. Results

We tested our algorithm on synthetic data and on WCE data. The WCE data has been obtained using the SB2 capsule endoscopy camera developed by Given Imaging, Ltd., Israel [5]. All videos have been conducted at Digestive Diseases Department, Hospital General "Vall d'Hebron" in Barcelona, Spain.

During the validation three types of cuts for longitudinal view are tested:

1. ($Acut$) - The adaptive cuts described in section 2,
2. ($Acut^-$) - The modification of the proposed algorithm by removing the smoothing term V. In this way the Eq. (1) transforms to $E(\alpha_1, \cdots, \alpha_n) = \sum_{i=1}^{n} D(\alpha_i)$. This is done to test the influence of the smoothness term in the energy function,
3. ($Fcut$) Longitudinal view with fixed cut.

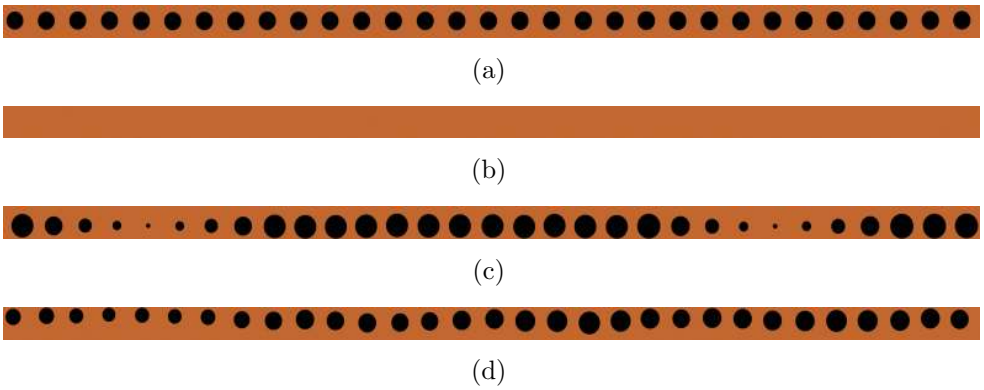


Figure 7: Some examples of the synthetically created intestinal events. From the top: a) tunnel, b) static, c) contractions and d) undefined movement.

3.1. Synthetic data

In this experiment, a synthetic video of 40.000 frames has been created with frame rate of 2 *fps*. On a uniform background, a blob has been placed. The blob position has been changed on consecutive frames depending on the intestinal event. The following intestinal events have been used to create the synthetic video: {*tunnel*, *static*, *contraction*, *undefined movement*}. In order to make the video more realistic, the events order and duration have been defined with random number generator.

The following definitions of the specific events have been used to generate the sequence:

- *Tunnel* - defined as a sequence of paralyzed intestine with open lumen. Lumen size is defined as highly constant (± 2 pixels difference in diameter between two consecutive frames) and highly opened (larger than 70 pixel in diameter). An example of the tunnel sequence is presented on Fig. 7(a).

- *Static* - defined as a sequence of paralyzed intestine with closed lumen (no lumen observed in the frame). An example of the static sequence is presented on Fig. 7(b).

- *Contraction* - defined as a sequence of frames with the presence of intestinal contraction defined as symmetric pattern of open-close-open lumen with duration of 9 frames and a fixed frequency of 6 contractions

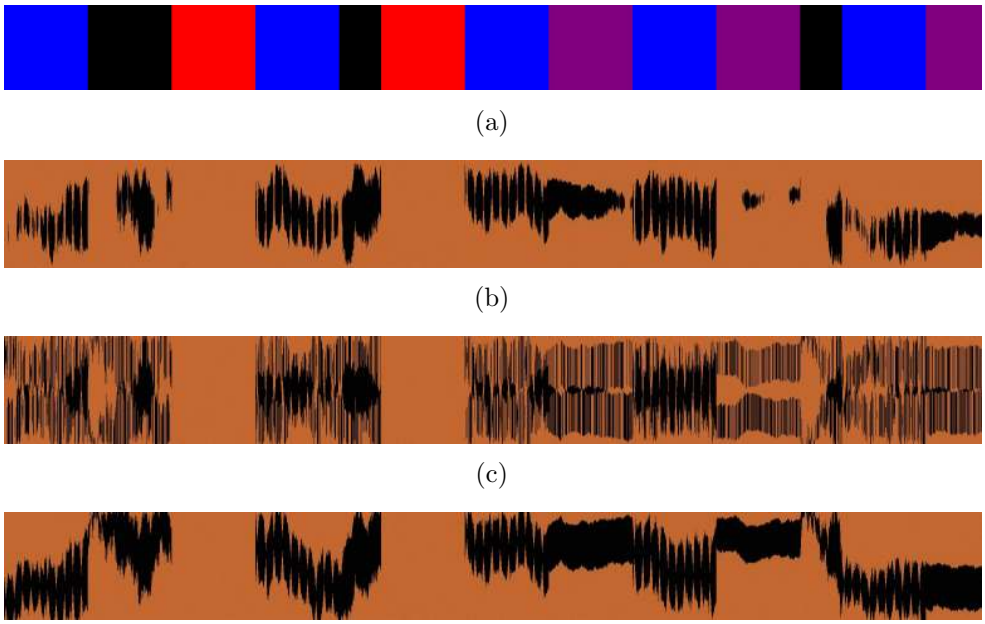


Figure 8: An example of three longitudinal views generated from the synthetic data. From the top: a) ground truth (blue - contraction sequence, black - undefined movement sequence, red - static sequence, purple - tunnel sequence), b) $Fcut$, c) $Acut^{-}$ and d) $Acut$.

$Fcut$	$Acut^{-}$	$Acut$
73%	85%	92%

Table 1: Table presenting the video segmentation score for different cuts. The number represents the segmentation accuracy.

per minute. The lumen size of the central frame of the intestinal contraction has been defined as 10% of the initial lumen size. An example of the contraction sequence is presented on Fig. 7(c).

- *Undefined movement* - defined as an irregular movement of the lumen size ($\pm$30 pixels variation between consecutive frames) and capsule ($\pm$30 pixels variation between consecutive frames). An example of the undefined movement sequence is presented on Fig. 7(d).

Some examples of the longitudinal views that are obtained using different cuts are presented in Fig. 8. Note that, the $Fcut$ losses the blob information and leads to miss-interpretation of the events (e. g. tunnel sequence, Fig. 8(a) vs. Fig. 8(b)). Applying the adaptive cut without the smoothing term $Acut^{-}$ leads to the good lumen detection, but the view is uninterpretable (here, only the static sequence can be visually detected). The adaptive cut $Acut$ presents well both lumen information and view smoothness.

The synthetic video has been presented to an expert asking him/her to recognize and to mark the beginning and the end of the intestinal event sequences. The results were evaluated using the Jaccard index (1 means perfect video segmentation and 0 means no coincidence between visually detected sequences and ground through).

The overall results are presented in Tab. 1. As it can be seen, the longitudinal view obtained by using $Acut$ method achieves 92% and outperforms $Fcut$ and $Acut^{-}$ methods (73% and 85%, accordingly). Analyzing the confusion matrixes presented on Fig. 9, the following conclusions can be drown:

- In $Fcut$, the tunnel sequences are frequently confused with static sequences, this happens when the blob is placed out of the cutting plane (see Fig. 9(a)).

- Applying the $Acut^{-}$ improves the lumen detection reducing the confusion between tunnel and static sequences. The contraction detection

	Contractions	Undefined	Static	Tunnel
Contractions	65,59%	4,59%	0,29%	1,57%
Undefined	6,54%	53,61%	0,16%	4,63%
Static	6,72%	3,56%	63,37%	15,14%
Tunnel	1,15%	9,84%	0,25%	44,12%

(a)

	Contractions	Undefined	Static	Tunnel
Contractions	64,44%	2,35%	0,18%	0,16%
Undefined	18,13%	57,14%	0,20%	3,00%
Static	0,45%	0,11%	97,40%	0,22%
Tunnel	0,56%	5,62%	0,21%	82,48%

(b)

	Contractions	Undefined	Static	Tunnel
Contractions	86,24%	5,41%	0,19%	0,15%
Undefined	1,82%	69,40%	0,18%	3,67%
Static	0,45%	0,11%	97,40%	0,22%
Tunnel	0,65%	5,31%	0,22%	82,74%

(c)

Figure 9: Confusions matrixes presenting Jaccard index obtained using the synthetic data. a) $Fcut$, b) $Acut^{-}$ and c) $Acut$.

rate is still small due to lack of the smoothness, sequences are often confused with undefined sequences (see Fig. 9(b)).

- In $Acut$, the biggest confusion is caused by the undefined sequence. The expert has a problem in distinguishing between the undefined and the contraction sequence and between the undefined and tunnel sequence. Still note that the confusion is small, less than 6% (see Fig. 9(c)).

3.2. Blob detection

In this part of the validation, we evaluate the blob detection using different cuts. Here, the correct detection means that the cut passes through the intestinal lumen. The lumen has been manually segmented in 24740 frames from WCE video. As expected, applying the smoothing term reduces blob detection rate by 7% resulting in overall score of 87% of the blob detected that is a good score in comparison to fixed cut that losses a lumen one every two frames. The results are presented in Tab. 2.

3.3. Real data

As a first step of the validation on real data, the qualitative inspection of the videos obtained with the Acut and the Fcut is done. Fig. 10 shows

	$Acut$	$Acut^{-}$	$Fcut$
detection rate	87%	94%	55%

Table 2: Table presenting results on lumen detection using different image cuts. Numbers represent the detection rate.

several examples of applying the Acut and the Fcut to the WCE videos. The difference between cuts is clearly visible when analyzing the lumen. In detail, the difference (Acut vs. Fcut):

- Fig. 10(a) vs. Fig. 10(b) - the lumen is better followed by the Acut, where the reconstruction of the tunnel is clearly visible, while in the case of Fcut it can be observed that the intestinal wall is present in some parts of the intestinal tunnel.
- Fig. 10(c) vs. Fig. 10(d) - Acut presents the lumen in the whole segment, while Fcut presents the lumen only in some parts. The presence of the lumen in the longitudinal view facilitates the interpretation of the motility information.
- Fig. 10(e) vs. Fig. 10(f) - the larger radius of tunnel is visible in Acut, while Fcut shows the small radius. Moreover, the Acut follows the radius while it is changing the position in the WCE video.
- Fig. 10(g) vs. Fig. 10(h) - Acut presents well the open-close-open lumen pattern, while in Acut this pattern is not clearly visible.

In the next part of the validation, the *Acut* longitudinal view (Fig. 11) and the original WCE video (using Rapid Viewer interface see Fig. 1) were presented to an expert asking him/her to mark the beginning and the end of the following sequences: {*tunnel*, *static*, *contraction*, *turbid* and *undefined movement* }. Table 3 points out the main visual aspects of different segments seen in longitudinal view.

Some examples of sequences seen on longitudinal view can be seen in Fig. 12:

- Fig. 12(a) and Fig. 12(b) show static sequences that are seen as homogeneous images of intestinal wall (tissue),

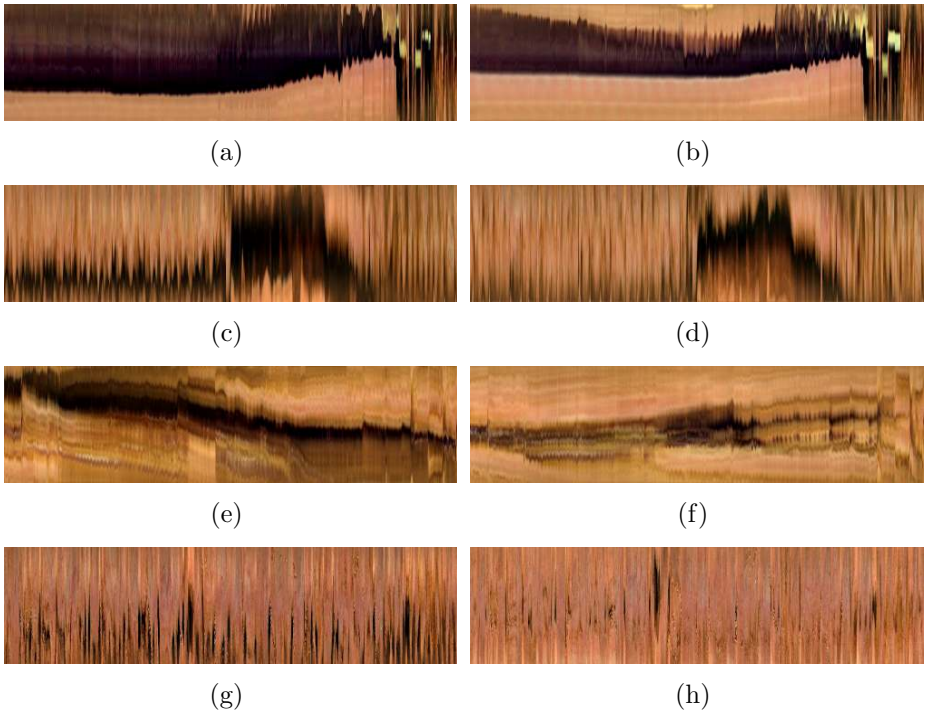


Figure 10: Visual comparison of different cuts (a, c, e, g) - *Acut*, (b, d, f, h) - *Fcut*.

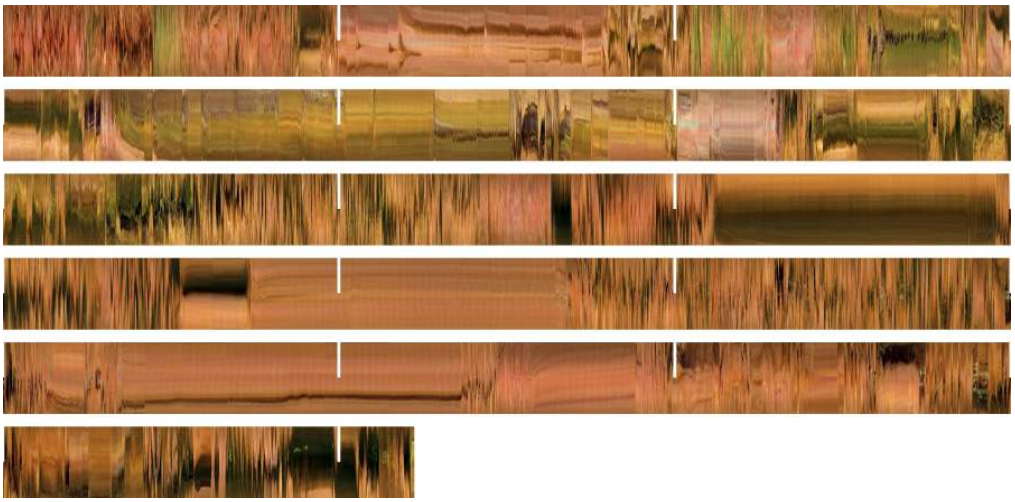


Figure 11: An example of longitudinal view presenting whole small intestine from duodenum to cecum. Each white stripe marks 10 minutes of the video duration.

	Contraction	Static	Tunnel	Turbid
Lumen presence	Yes	No or small lumen	Yes	Can be occluded
Intestine tissue presence	Yes	Yes	Yes	Can be occluded
Colors seen	Orange (tissue) and dark brown/black (lumen)	Orange (tissue)	Orange (tissue) and dark brown/black (lumen)	Light green to dark brown (turbid), and optionally orange (tissue)
Lumen/tissue movement	Yes	No	No or small changes	Can be occluded
Open-close-open lumen pattern	Yes (visible as "stripes")	No	No	No

Table 3: Table pointing out the main visual aspects of different sequences seen on longitudinal view.

- Fig. 12(c) and Fig. 12(d) show tunnel sequences where both intestinal lumen and wall are static and clearly visible,
- Fig. 12(e) and Fig. 12(f) show turbid sequences where the green color of intestinal content occluding lumen/wall can be seen,
- Fig. 12(g) and Fig. 12(h) show contraction sequences with the periodical changes in the visible lumen size.

It is interesting to see that the turbid mixing contractions on Fig. 12(e). Also, note the variety in contraction rhythm in, Fig. 12(g) and Fig. 12(h).

In the first step, the expert labeled[3] five types of event sequences in a video view. In the second part, the expert labeled the sequences using longitudinal view (*Acut*). The similarity between different annotations in video view and in longitudinal view (*Acut*) is 81%. The time needed by an expert for visual inspection of the data of duration of 164 minutes was 80 minutes in video display vs. 18 minutes in the longitudinal view. The longitudinal view reduces the inspection time by 62 minutes that means the reduction of 444%.

For the obtained annotations, the confusion matrix presenting the overlapping between two annotations, was calculated. Analyzing the confusion matrix presented on Fig. 13 the following conclusions can be drawn:

[3]The experts knew how the Acut works, and knew that the task is being timed.

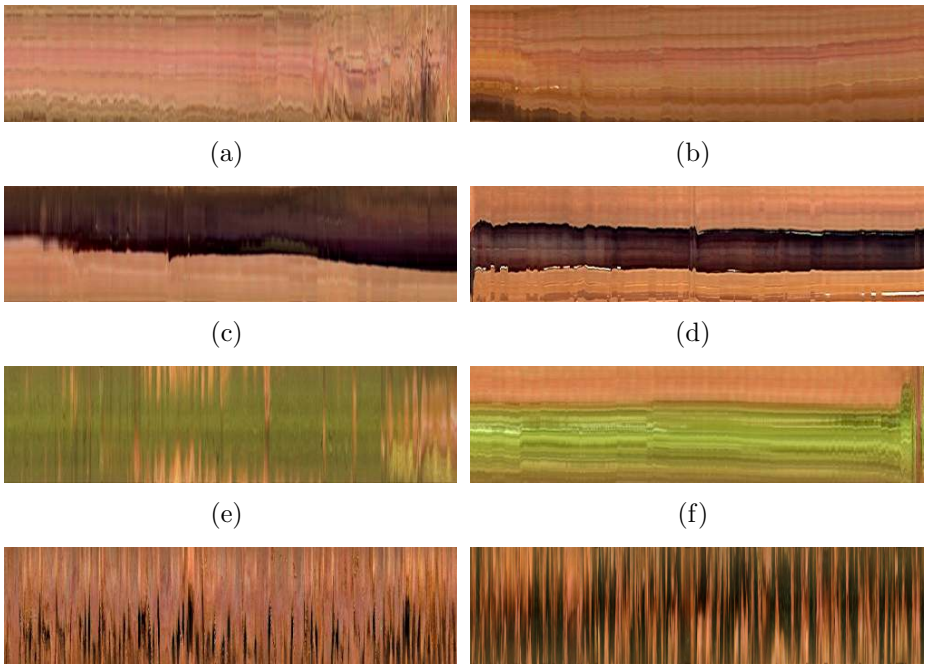


Figure 12: Examples of longitudinal views presenting different intestinal events (a - b) static, (c - d) tunnel, (e - f) turbid and (g - h) contractions.

LONG. VIEW \ VIDEO VIEW	Turbid	Tunnel	Static	Contractions	Undefined
Turbid	69,70%	0,00%	0,00%	3,49%	0,00%
Tunnel	7,14%	50,00%	1,67%	1,30%	3,70%
Static	2,70%	1,75%	67,27%	4,81%	1,72%
Contractions	2,17%	1,33%	4,76%	73,97%	2,67%
Undefined	0,00%	4,17%	5,36%	0,00%	52,94%

Figure 13: Confusion matrix for annotation comparison between video based annotations and longitudinal view based annotations. The numbers represent the Jaccard index.

- Turbid sequences are coinciding quite well between both annotations. The biggest confusion 7% is caused by the tunnel sequence in the longitudinal view. This is due to the opaque turbid that is difficult to see on the longitudinal view while, it only slightly hinders the lumen/wall (in this case hinders a tunnel sequence).

- The 50% coincidence in tunnel between the annotations in the longitudinal view and the video view is caused by: 1) opaque turbid that is difficult to detect in longitudinal view and 2) the camera rotation together with small intestine wall oscillations make it difficult to spot the tunnel sequence in video view.

- Static sequences coincide well in both annotations with the score of 67%. The biggest confusion is between static and undefined movement sequence.

- Contractions have the highest score of coincidence of all the annotated events 74%.

To check the intra-user variability, another user labeled the events in the longitudinal view image. The obtained kappa score was: 0.80, with 0.84 observed agreement and 0.21 random agreement. Analyzing the confusions/errors of the annotations, the following observations can be done:

- Turbid vs. contractions - one expert in the presence of turbid and contractions preferred to annotate turbid, the other annotated contractions.

- Undefined - this label is used whenever an expert is not sure about the considered event. In many cases, it is used in borderline situations. When analyzing the annotations, it can be seen that most frequent confusion is between contractions and turbid.

4. Conclusions

In this paper, a fast and efficient algorithm for construction of adaptive longitudinal view for motility analysis has been presented. It allows compact display and fast inspection of motility data acquired with WCE. To the best of our knowledge, no previous longitudinal visualization technique has been tested for motility analysis. The algorithm adapts the frame cut angle to the lumen position by the minimization of a cost function. This problem formulation permits to apply the Dynamic Programming framework to efficiently find the global minimum to the proposed cost function. Experimental results on both: 1) synthetic data and 2) WCE data show that the proposed algorithm preserves well lumen/wall separation allowing to inspect the intestine motility in details. The annotations obtained by using adaptive longitudinal view coincide well with the ones obtained by using video view. Moreover the time needed for visual inspection is four time faster in longitudinal view than in video view.

This visualization technique offers a new, holistic view into small intestine motility. It opens new investigation lines in the intestinal motility analysis, allowing to see the events that are difficult to spot by using traditional WCE video analysis. It would be interesting to analyze the different contractile rhythms and sequence combinations (e. g. one-minute-cycle, where each minute the intestine changes the state from paralyzed intestine to contractile movement).

The proposed method has some limitations that are worth mentioning. First of all, the cut passes always through the central point of the frame, this set-up reduces the space of possible cuts omitting the cuts that are not passing through the center. As a result, the obtained solution does not have to be the solution of the maximal lumen region. Second limitation is that the first term of proposed energy function (based on basic statistics of the cut) may not infer lumen regions, it only assures that the cut passes through dark part of the frame as it is assumed that the lumen is a dark blob. Examples of the frames where this assumption might fail are: frames with dark food content or water bubbles. This fact is not critical since usually frames with food content or water bubbles use to have a significant coverage and thus any cut has the same clinical interest during inspection and analysis.

The proposed algorithm could be further improved by placing the lumen position in the center of the longitudinal view. This is not an easy task mainly because of the free capsule movement inside the intestine, and due to

the fact that in a lot of frames the lumen is not visible. Placing the lumen in the center of the view could be important step to obtain 3D reconstruction and display of small intestine that could be of high interest to detect lesions and achieve image-guided interventions.

Acknowledgments

This work was supported in part by a research grant from Given Imaging Ltd., Yoqneam Israel, as well as by the MICINN Grants TIN2009-14404-C02 and CONSOLIDER-INGENIO 2010 (CSD2007-00018).

References

[1] Whitehead, W.E.. Gastrointestinal motility disorders of the small intestine, large intestine, rectum, and pelvic floor. International Foundation for Functional Gastrointestinal Disorders 2001;IFFGD Fact Sheet No. 162.

[2] Medscape. 2009. URL `http://emedicine.medscape.com/article/179937-overview`.

[3] Quigley, E.M.. Gastric and small intestinal motility in health and disease. Gastroenterology Clinics of North America 1996;25:113–145.

[4] Iddan, G., Meron, G., et al. Wireless capsule endoscopy. Nature 2000;405:417.

[5] Given imaging, ltd. 2011. URL `http://www.givenimaging.com/`.

[6] Malagelada, C., De Iorio, F., Azpiroz, F., Accarino, A., Seguí, S., Radeva, P., et al. New insight into intestinal motor function via noninvasive endoluminal image analysis. Gastroenterology 2008;135(4):1155–62.

[7] Malagelada, C., De lorio, F., Seguí, S., Mendez, S., Drozdzal, M., Vitria, J., et al. Functional gut disorders or disordered gut function? small bowel dysmotility evidenced by an original technique. Neurogastroenterology and Motility 2012;24(3):223–e105.

[8] Figueiredo, I.N., Prasath, S., Tsai, Y.H.R., Figueiredo, P.N.. Automatic detection and segmentation of colonic polyps in wireless capsule images. In: ICES REPORT 10-36, The Institute for Computational Engineering and Sciences, The University of Texas at Austin. 2010,.

[9] Zhao, Q., Meng, M.H.. Polyp detection in wireless capsule endoscopy images using novel color texture features. In: Intelligent Control and Automation (WCICA), 2011 9th World Congress on. 2011, p. 948 –952.

[10] Jung, Y., Kim, Y., et al. Active blood detection in a high resolution capsule endoscopy using color spectrum transformation. In: International Conference on BioMedical Engineering and Informatics. 2008, p. 859–862.

[11] Hwang, S., Oh, J., et al. Blood detection in wireless capsule endoscopy using expectation maximization clustering. In: Proceedings of the SPIE'06; vol. 6144. 2006, p. 577–587.

[12] Shah, S., Lee, J., Celebi, M.. Classification of bleeding images in wireless capsule endoscopy using hsi color domain and region segmentation. In: Proceedings of the ASEE'07. 2007,.

[13] Karkanis, S.A., Iakovidis, D.K., Maroulis, D.E., Karras, D.A., Member, A., Tzivras, M.. Computer-aided tumor detection in endoscopic video using color wavelet features. IEEE Transactions on Information Technology in Biomedicine 2003;7:141–152.

[14] Li, B., Meng, M.Q.H.. Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments. Computers in Biology and Medicine 2009;39:141–147.

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

[16] Ciaccio, E.J., Tennyson, C.A., Lewis, S.K., Krishnareddy, S., Bhagat, G., Green, P.H.R.. Distinguishing patients with celiac disease by quantitative analysis of videocapsule endoscopy images. Computer Methods and Programs in Biomedicine 2010;100:39–48.

[17] Bejakovic, S., Kumar, R., Dassopoulos, T., Mullin, G., Hager, G.. Analysis of Crohn's disease lesions in capsule endoscopy images. In: Proceedings of the 2009 IEEE International Conference on Robotics and Automation. ICRA'09; 2009,.

[18] Vu, H., Echigo, T., et al. Contraction detection in small bowel from an image sequence of wireless capsule endoscopy. In: Proceedings of MICCAI'07; vol. 1. 2007, p. 775–783.

[19] Vu, H., Echigo, T., Sagawa, R., Yagi, K., Shiba, M., Higuchi, K., et al. Detection of contractions in adaptive transit time of the small bowel from wireless capsule endoscopy videos. Computers in Biology and Medicine 2009;39:16–26.

[20] Vilarino, F., Spyridonos, P., Deiorio, F., Vitria, J., Azpiroz, F., Radeva, P.. Intestinal motility assessment with video capsule endoscopy: automatic annotation of phasic intestinal contractions. IEEE Transactions on Medical Imaging 2010;29(2):246–59.

[21] Vu, H., Sagawa, R., Yagi, Y., Echigo, T., Shiba, M., Higuchi, K., et al. Evaluating the control of the adaptive display rate for video capsule endoscopy diagnosis. In: Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics. 2009,.

[22] Hai, V., Echigo, T., et al., . Adaptive control of video display for diagnostic assistance by analysis of capsule endoscopic images. In: Proceedings of the ICPR'06; vol. III. 2006, p. 980–983.

[23] Yagi, Y., Vu, H., et al. A diagnosis support system for capsule endoscopy. Inflammopharmacology 2007;5(2):78–83.

[24] Szczypinski, P.M., Sriram, R.D., Sriram, P.V., Reddy, D.N.. A model of deformable rings for interpretation of wireless capsule endoscopic videos. Medical Image Analysis 2009;13(2):312 – 324. Includes Special Section on Functional Imaging and Modelling of the Heart.

[25] Felzenszwalb, P., Zabih, R.. Dynamic programming and graph algorithms in computer vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 2011;33:721–740.