A new image centrality descriptor for wrinkle frame detection in WCE videos.

Santi Seguí[1,2], Ekaterina Zaytseva[1,2], Michal Drozdzal[1,2], Carolina Malagelada[3], Fernando Azpiroz[3], Petia Radeva[1,2] and Jordi Vitrià[1,2]
[1] Computer Vision Center (CVC), Universitat Autònoma de Barcelona, Barcelona, Spain
[2] Dept. Matemàtica Aplicada i Anàlisis, Universitat de Barcelona, Barcelona, Spain
[3] Digestive System Research Unit, Hospital Vall dHebron, Barcelona, Spain
santi.segui@ub.edu

Abstract

Small bowel motility dysfunctions are a widespread functional disorder characterized by abdominal pain and altered bowel habits in the absence of specific and unique organic pathology. Current methods of diagnosis are complex and can only be conducted at some highly specialized referral centers. Wireless Video Capsule Endoscopy (WCE) could be an interesting diagnostic alternative that presents excellent clinical advantages, since it is non-invasive and can be conducted by non specialists. The purpose of this work is to present a new method for the detection of wrinkle frames in WCE, a critical characteristic to detect one of the main motility events: contractions. The method goes beyond the use of one of the classical image feature, the Histogram of Oriented Gradients (HoG), and proposes the use of a mid-level image descriptor, centrality. In the case of wrinkle detection in WCE, this descriptor is computed by applying a graph-based centrality measure on histograms of oriented structure tensor image descriptors. We show how to apply this image descriptor to the detection of contractions in WCE videos and that it outperforms previous methods.

1 Introduction

Small bowel motor dysfunctions may cause mild and severe clinical syndromes, such as intestinal pseudo-obstruction, reduced tolerance to feeding and inability to maintain normal body weight. These motility disorders of the small bowel are caused by the involvement of the muscular layer of the intestine or the involvement of neural control system. Mainly, intestinal motility dysfunctions are characterized by an abnormal contractile activity, such as spasm and intestinal paralysis.

The presence or absence of diverse physiological symptoms constitutes the first evidence for the diagnosis of a pathology of the small bowel. Nowadays, the main source of information, and the only one which leads to a conclusive diagnosis of intestinal motility disorders is the one obtained from the result of motility test performed by using manometric devices [6]. However, the application of this technique presents several drawbacks: 1) It is restricted to few referral centers due to the complexity of the technical procedure, 2) it is an invasive technique and 3) it is restricted to the analysis of pressure values, lacking of information about different content, structure, morphology and dynamics of the intestine.

In 2001 Wireless Capsule Endoscopy (WCE) was presented as a new technology that allowed to look at the intestine from inside. From then, several applications based on WCE have been presented to help physicians to obtain a diagnosis for different pathological abnormalities such as bleeding, Crohn disease, polyp and ulcer detection and severe obstruction.

Intestinal contractions, the main event of motility activity, are the result of a muscular stimulation produced by the nerve system. Its role is the mixing and the propulsion of the food we eat. Its importance for the diagnose for motility disorders has been shown [8, 10, 11]. By the means of WCE, intestinal contractions are visualized as a sequence of frames where intestinal lumen appears opened and then is closed and following opened again. The main visual feature to characterize these events is the changing lumen area during a sequence of consecutive frames (see Figure 1).

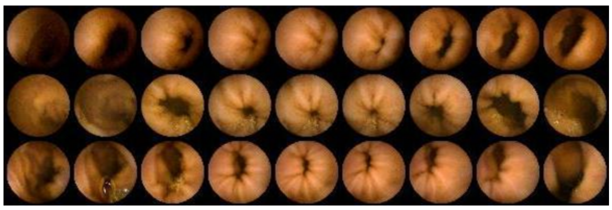

Figure 1. Samples of intestinal contractions. Each row represents different contractions.

Wrinkles are an omnipresent characteristic of contractions and has been mainly studied as a pattern to characterize a subset of intestinal contractions [9, 7]. This pattern is visually observed as a set of folds of intestinal wall, in star-like shape (see Fig. 2). Usually, wrinkle pattern is observed in the central frames on intestinal contractions where strong pressure is produced by the nerve system.

Figure 2. Wrinkle images showing the star-like pattern.

In the recent literature, we find two publications focused on the characterization and the detection of this event. In [9], Vilariño et.al. proposed a method for the detection of tonic contractions [9] based on wrinkle information. The proposed method categorizes wrinkle frames by using general linear radial patterns based on the valleys and ridges of the image. An alternative method was proposed by Spyridonos et.al. in [7]. In this paper, a new image descriptor for categorization of wrinkle images was presented using directional information from the structural tensor matrix. In both works, the characterization of wrinkle frames was achieved by dividing the images into 4 different quadrants and computing the corresponding set of features for each quadrant divided according of lumen center. The weakest point of these approaches is the need to detect the lumen center, because, in most of the cases, it is hard to be successfully accomplished. We have to note, that in wrinkle images the lumen is very small or even appears completely closed because of the wall pressure produced by the nerve system.

Nevertheless, in the literature we can find two papers where wrinkle feature has been successfully used as an indicator in a complex motility diagnose system [4, 5].

In this work we address the problem of detecting wrinkle frames using a new mid-level image descriptor that measures the continuity of certain image features. The image is locally described by the structure tensor or second moment matrix, a matrix derived from the gradients of the image. This image descriptor is used because it summarizes the predominant directions of the gradient in the neighborhood of a pixel. Then, we compute a Histogram of Oriented Features from the structure tensor field in an analogous way to the well known Histogram of Oriented Gradients. The resulting image description, which is composed of a set of orientation histograms defined on image cells, is transformed into a graph where each node represents a cell. All neighboring cells are connected by a link labeled with a value that corresponds to the continuity level between them. That is, if two neighbor cells contain two histograms with compatible histogram bins, then their corresponding link gets a high value. Compatibility between histogram bins is defined in a way that favors the continuity of image structure. Finally, the centrality value of all nodes of the graph is computed and used to build a Support Vector Machine classifier that is able to efficiently detect wrinkle frames in WCE images.

The organization of this paper is as follows: Section 2 presents the proposed image descriptor. Section 3 presents a qualitative and quantitative discussion of our results. Finally, we discuss our contribution and draw some conclusions in Section 4.

2 Method

The proposed method is divided into the following steps: 1) In the first step we detect the predominant gradient directions of an image using the Structural Tensor (ST). This descriptor was previously used in [7] and provides clear information of wall folds (see Fig. 3.a). 2) Using gradient fields from ST we construct a

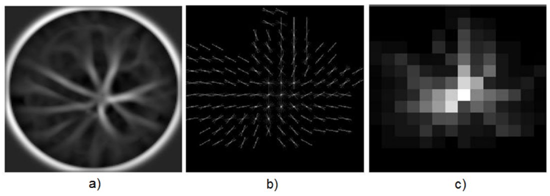

Figure 3. a) First eigenvalues of the Structural Tensor; b) Histogram of oriented features computed from ST; c) Centrality descriptor (darker cell indicate low centrality and lighter cells indicate high centrality).

histogram of oriented features in a similar way as it is done with well known Histogram of Gradients (HOG) [2] (see Fig. 3.b), and 3), we compute the centrality of each cell of the computed histogram of features (see Fig. 3.c). In the following subsections the details of each step are explained.

2.1 Structural Tensor

The structural tensor (ST), also called second-moment matrix, is a matrix derived from the gradient image. This matrix summarizes the predominant directions of the gradients in a neighborhood of a point and their magnitudes. The structural tensor S_w of an image I is computed as follows:

$$S_w(p) = \begin{bmatrix} (I_x(p))^2 & I_x(p)I_y(p) \\ I_x(p)I_y(p) & (I_y(p))^2 \end{bmatrix} \quad (1)$$

where, I_x and I_y are the partial derivatives of the image I with respect to x and y coordinates and $p = (x, y)$ is a point. Following, the eigenvectors of the structural tensor and their associated eigenvalues can be computed at each point of the image. The eigenvector that corresponds to the largest eigenvalue is perpendicular to the wrinkle edge. The second eigenvector is parallel to the edge direction, pointing towards the smoothest image region. We define the magnitude of $S_w(p)$ as the value of the first eigenvalue.

An example that illustrates this procedure is presented in Fig 4. As it can be seen, the wall folds are properly detected and a star-shape is obtained by performing a threshold over the magnitude of the ST.

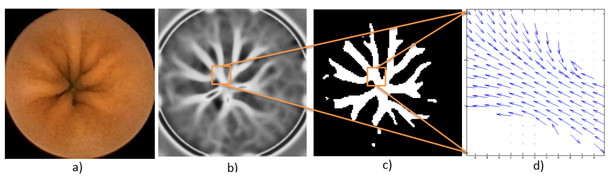

Figure 4. a) Original image b) Magnitude of the first eigenvalue of the structural tensor; c) Binary image after thresholding ST magnitude; d) Orientations of ST of one cell.

2.2 Histogram of Features using Structural Tensor

The classical HoG descriptor is implemented by dividing the image into a set of small connected regions and for each region, or cell, compiling a histogram of gradient directions for the pixels within the cell. HoG descriptor is then built by concatenating the values of the bins of all histograms, getting a high-dimensional vector $H_G = (\mathbf{h}_1, \ldots, \mathbf{h}_m)$ that represents the image, being $\mathbf{h}$ an histogram and m the number of cells. In our method we use the same methodology of classical HoG but using the image field provided by the Structural Tensor, H_{ST}.

2.3 Centrality descriptor

In order to compute a centrality descriptor of the image, we consider the continuity measure presented by Zaytseva et.al in [12]. In this paper the authors propose the computation of a centrality measure on a HoG descriptor, H_G. In our case we propose to compute the centrality of the nodes of a graph G derived from from H_{ST}.

Given the H_{ST} descriptor of an image with m regions, we built a graph G by considering a set of m nodes V, where each node v_j corresponds to the histogram h_j from cell j of H_{ST}, and a set of edges E, where edge $e_{i,k}$ connects the neighboring cells corresponding to the histograms $\mathbf{h}_i$ and $\mathbf{h}_k$ (we have considered a 8-connectivity region).

The cost of an edge $e_{i,k}$ is assigned by taking into account the values of $\mathbf{h}_i$ and $\mathbf{h}_k$. Given a pair of neighbor nodes v_i and v_k, we first consider the predominant orientation α of $\mathbf{h}_i$. Then, the cost of $e_{i,k}$ is computed by adding the value of the bins from $\mathbf{h}_i$ and $\mathbf{h}_k$ that correspond to the following angles: $\alpha - 22.5°$ to $\alpha + 22.5°$. In this way, edges connecting neighboring cells with aligned ST fields will get higher values.

The next step is to calculate the centrality of each node $v_j \in V$ of the graph. To this end, we have considered different node centrality measures on graphs, but the most significant has resulted to be the *betweenness* measure proposed in [3]. The betweeness measure of a node v_j is equal to the number of shortest paths from all vertices to all others that pass through v_j. This value is normalized by dividing through the number of pairs of vertices in graph V.

Formally, this centrality measure can be defined as:

$$C(v_j) = \sum_{s \neq v_j \neq t \in G} \frac{\sigma_{st}(v_j)}{\sigma_{st}} \tag{2}$$

where σ_{st} is the total number of paths from node s to node t and $\sigma_{st}(v_j)$ is the total number of shortest paths from a node s to a node t that pass through v_j. The resulting descriptor is a vector C of size m.

An example that illustrates this procedure is presented in Fig. 3. As can be seen there, the maximum value of the centrality measure is taken in the closed lumen and takes high values where wrinkle pattern ia present.

In order to detect a wrinkle pattern in an image, we propose to use a large sliding window that scans the image and evaluates the presence of a wrinkle pattern with a linear classifier.

3 Results

In order to validate the proposed system a training and a testing set were created using different videos obtained with PillCam SB2 capsule provided by Given Imaging Ltd. The training set consist of 1.000 wrinkle frames and 1.000 non wrinkles frames from 4 videos. The testing set consist of 1500 wrinkles frames and 2500 non wrinkles frames from 5 videos (not considered in the training set). As lumen center is not always located in the center of the image, the lumen center of training wrinkle images was manually labeled.

The models were learnt using a Linear Support Vector Machine (SVM) [1]. SVM parameters where tuned with the training set using a cross validation strategy. Positives samples consists of sub-windows images of winkle samples of 128x128 pixels centered at lumen center, and negative samples, consist of image window of the size 128x128 pixels located at random location of negative samples. The size of the neighborhoods on which the structural tensors were computed was defined with a gaussian function with $sigma = 2$.

The classification on the testing set was done using the sliding window approach. Image score was considered as the maximum score of all tested sub-windows. Table 1 presents the obtained results using three different image descriptors: 1) the standard HoG descriptor computed from image gradients, 2) the raw Structural Tensor descriptor as described in [7] and 3) the proposed centrality C descriptor using ST. As it can be seen in the table, the proposed centrality descriptor outperforms other image descriptor increasing Average Precision of Structural Tensor from 87.78 to 91.91 and Accuracy from 83.83 to 87.76. Additionally, precision/recall curves presented in figure 5 show that centrality descriptor obtains a better compromise between precision and recall.

Figure 6 and 7 shows some qualitative results. In Fig. 7 we observe some False Positive (FP) detections and in Fig. 6 some False Negative (FN) detections. As it can be seen, all FN illustrated in Fig. 7 present very soft folds of intestinal wall and also a completely closed lumen. On the other hand, the obtained FP illustrated in Fig 7 presents several folds of the wall. Moreover, there are two images (marked in red) which are difficult to consider as FP or TP. For instance the first image marked in red is fully covered by intestinal content, however, it presents the classic star-shape observed in the wrinkle frames. It suggest us that the contractions is there although wall visualization is hindered by intestinal content.

4 Conclusions

In this paper we presented a new image descriptor for the classification of wrinkle frames using WCE. The proposed image descriptor is based on the the centrality descriptor which is computed using the histogram

Table 1. Classification performance using 1) standard HoG descriptor (H_G), 2) Histogram of Gradients using Structural Tensor matrix (H_{ST}) and 3) proposed Centrality descriptor C.

	H_G	H_{ST}	Centrality
Average Precision	64.19	87.78	91.91
Accuracy	67.40	83.82	87.76
Sensitivity	54.14	70.75	83.52
Specificity	75.95	92.23	90.17
FAR	24.04	7.76	9.83

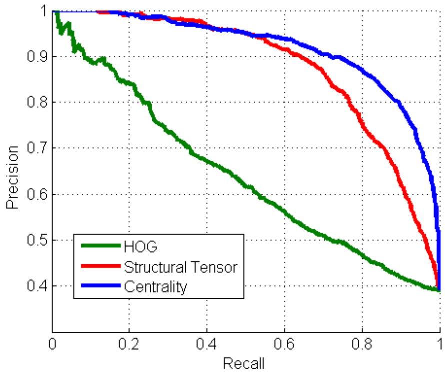

Figure 5. Precision/Recall curves.

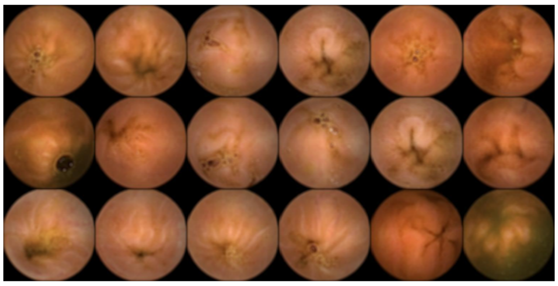

Figure 6. False Negative samples.

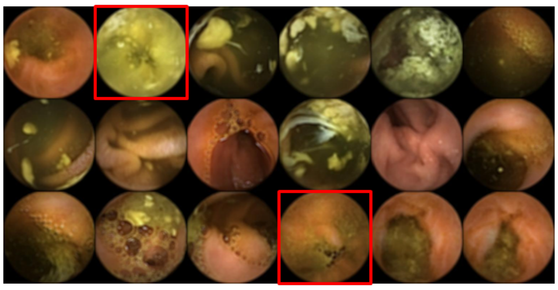

Figure 7. False Positive samples. Images marked in red show unclear label although are considered FP.

of oriented features extracted from the structural tensor field of an image. The validation, carried on a large database, shows that the proposed descriptor successfully detects this particular event of WCE videos, outperforming previous methods.

Acknowledgements

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 TIN2012-38187-C03-01.

References

[1] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011.

[2] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.

[3] L. C. Freeman. A set of measures of centrality based on betweenness. 40, 1977.

[4] C. Malagelada, F. De Iorio, F. Azpiroz, A. Accarino, S. Seguí, P. Radeva, and J.R. Malagelada. New insight into intestinal motor function via noninvasive endoluminal image analysis. *Gastroenterology*, 135(4):1155–1162, 2008.

[5] C. Malagelada, S. Seguí, S. Mendez, M. Drozdzal, J. Vitrià, P. Radeva, J. Santos, A. Accarino, JR Malagelada, F. Azpiroz, et al. Functional gut disorders or disordered gut function? small bowel dysmotility evidenced by an original technique. *Neurogastroenterology & Motility*, 2012.

[6] E M Quigley. Gastric and small intestinal motility in health and disease. *Gastroenterol Clin North Am*, 25(1):113–45, 1996.

[7] Panagiota Spyridonos, Fernando Vilariño, Jordi Vitrià, Fernando Azpiroz, and Petia Radeva. Anisotropic feature extraction from endoluminal images for detection of intestinal contractions. In *MICCAI (2)*, pages 161–168, 2006.

[8] Fernando Vilariño, Panagiota Spyridonos, Fosca De Iorio, Jordi Vitrià, Fernando Azpiroz, and Petia Radeva. Intestinal motility assessment with video capsule endoscopy: Automatic annotation of phasic intestinal contractions. *IEEE Trans. Med. Imaging*, 29(2):246–259, 2010.

[9] Fernando Vilariño, Panagiota Spyridonos, Jordi Vitrià, Carolina Malagelada, and Petia Radeva. Linear radial patterns characterization for automatic detection of tonic intestinal contractions. In *CIARP*, pages 178–187, 2006.

[10] Hai Vu, Tomio Echigo, Ryusuke Sagawa, Keiko Yagi, Masatsugu Shiba, Kazuhide Higuchi, Tetsuo Arakawa, and Yasushi Yagi. Contraction detection in small bowel from an image sequence of wireless capsule endoscopy. In *Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention - Volume Part I*, MICCAI'07, pages 775–783, Berlin, Heidelberg, 2007. Springer-Verlag.

[11] Hai Vu, Tomio Echigo, Ryusuke Sagawa, Keiko Yagi, Masatsugu Shiba, Kazuhide Higuchi, Tetsuo Arakawa, and Yasushi Yagi. Detection of contractions in adaptive transit time of the small bowel from wireless capsule endoscopy videos. *Comp. in Bio. and Med.*, 39(1):16–26, 2009.

[12] Ekaterina Zaytseva, Santi Seguí, and Jordi Vitrià. Sketchable histograms of oriented gradients for object detection. In *CIARP*, pages 374–381, 2012.