

A Region Segmentation Method for Colonoscopy Images Using a Model of Polyp Appearance

Jorge Bernal, Javier Sánchez, and Fernando Vilariño

Computer Vision Center and Computer Science Department, Campus Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

jbernal@cvc.uab.es

<http://www.cvc.uab.es/>

Abstract. This work aims at the segmentation of colonoscopy images into a minimum number of informative regions. Our method performs in a way such, if a polyp is present in the image, it will be exclusively and totally contained in a single region. This result can be used in later stages to classify regions as polyp-containing candidates. The output of the algorithm also defines which regions can be considered as non-informative. The algorithm starts with a high number of initial regions and merges them taking into account the model of polyp appearance obtained from available data. The results show that our segmentations of polyp regions are more accurate than state-of-the-art methods.

Keywords: Colonoscopy, Polyp Detection, Region Merging, Region Segmentation.

1 Introduction

Colon cancer's survival rate depends on the stage that it is detected on, going from rates higher than 95% in stages I or II to rates lower than 35% in stages IV and V [1], elucidating the importance of detecting it on its early stages. In order to do so, several screening techniques are used. One of the most extended techniques is colonoscopy [2], which consists of introducing a probe with a camera mounted on it through the rectum and the colon. The physician can observe the status of the patient as the colonoscope progresses, and it can even remove polyps during the intervention.

Our global objective is to detect polyps in colonoscopy images. Our processing scheme [3] consists of 3 stages: *region segmentation*, *region description* and *region classification*. This detection scheme can be used in several applications such as: 1) real-time polyp detection, 2) off-line quality assessment of the colonoscopy, and 3) quantitative assessment of the trainee skills in training procedures, just to mention a few.

In this paper we present our *region segmentation* stage in which an input colonoscopy image is segmented into a minimum number of informative regions, one of them containing a polyp, therefore reducing the size of the problem. The term of informative regions is used here as opposite to non-informative

regions, where we are sure no polyp is inside and therefore, there will be no need to continue analyzing them [4]. These results can be used later to classify the informative regions into polyp- vs. non-polyp-containing candidates.

The structure of the paper is as it follows: in Section 2 we present the segmentation algorithm which we will compare our performance results with. In Section 3 we present our segmentation algorithm along with the model of polyp appearance in which it was inspired. In Section 4 we present our experimental setup and show our results. Finally in Section 5 we show the main conclusions that we extract from our work and present some future research lines.

2 Related Work

There are different approaches to polyp detection in colonoscopy video, which can be divided [5] according to the type of feature they are based on, namely shape, color or texture. Some of the include, like us, a *Region Segmentation* stage that also use shape and color cues to guide the process, such as the work of Hwang et al. [6], although they are in a more advanced stage where they classify the segmented regions.

In general segmentation, which is one of the most difficult and critical tasks in computer vision, can be viewed as a perceptual grouping problem in which the image is divided into homogeneous regions, which can represent different features in the images depending on the methodology adopted. Some simple ways of segmentation exist however they prove to be over simplified for semantic grouping of image regions in more complex scenarios, as they are more sensitive to noise and other artifacts [7]. More sophisticated methods of image segmentation can be mainly divided into two different categories: segmentation by fitting and segmentation by clustering [8]. In the former, the problem of segmentation is viewed as an assertion that the pixels in an image conform to a model while, in the latter, the pixels are grouped according to some criteria such as gray level, color or texture. In order to perform efficiently, segmentation by fitting methods need strong gradient differences pertaining to the objects in the images which have to be segmented, which is not our case. Given that we want to segment informative regions containing polyps from clinically uninteresting areas, methods that segment by clustering seem well suited for our scenarios. Because of this, we have chosen two methods from this group to carry out our research:

a) Normalized Cuts: The *normalized cuts* method [9] is a graph theoretic approach for solving the perceptual grouping problem in vision. In *normalized cuts*, all the sets of points lying in the feature space are represented as a weighted, undirected graph. The weight of each arc is assigned using a set of pre-defined criteria. These can be based on the spatial distance among the pixels, their brightness values, etc. Usually the easiest way to perform segmentation in graph theoretic algorithms is to disconnect the edges having small weights usually known as the *minimum cut* [10]. The problem with minimum cuts is that it typically results in over segmentation since the method basically finds local minima. Shi and Malik [9] proposed in 2000 a new approach that aims at extracting the

global impression of an image instead of focusing on its local features. In this approach (*normalized cuts*), the cut between two graphs is normalized by the volumes of the resulting graphs. In this case graphs are constructed by taking each pixel as a node and by defining the edge weight as a the product of feature similarity term and spatial proximity term [9].

b) Watersheds: Watershed transformation [11] is one of the clustering based methods used as a tool for image segmentation. Watersheds operate on intensity gradients to perceive an image as a combination of catchment basins in a hilly area (a hill corresponds to high gradient) simulating the formation of image regions with projected flow of water. After identification of an intensity valley in an image, region growing algorithms are used to combine all the pixels which have similar intensities. This procedure is particularly effective in images where we have strong boundaries of objects which have to be segmented. For images which are rich in texture or where the intensity gradient is not prominent, an over segmentation is usually obtained, making convenient a posterior region merging.

3 Our Proposed Methodology

3.1 A Model of Polyp Appearance

The lighting of the probe gives us hints about how polyps appear in colonoscopy images. As the light falls perpendicularly to the walls of the colon, it creates shadows around the surfaces and, when the light falls into a prominent surface (Figure 1 (a)), it creates a bright spot surrounded by darker areas, which are the shadows, generating edges and valleys in the intensity image (Figure 1 (b)).

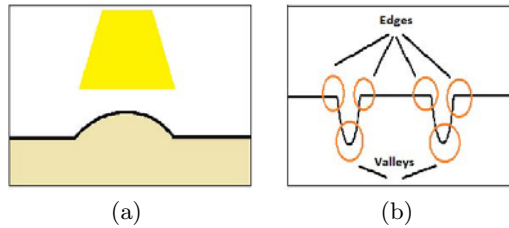


Fig. 1. Simulation of (a) an illuminated prominent surface (b) grey-scale profile

Even considering this model of prominent surface appearance under the lighting conditions of the colonoscope, there are some challenges to be overcome, namely: 1) non-uniform appearance of the polyps (flat or peduncular shapes); 2) the appearance of reflections in the image; and 3) the similarity between tissues inside and outside the polyp, also affected by the lighting conditions. Taking all this into consideration, we base our segmentation method on a model of polyp appearance that we can define as *a prominent shape enclosed in a region with presence of edges and valleys around its frontiers*.

3.2 Algorithm

The segmentation algorithm, which general scheme can be seen in Figure 2, consists of 4 different stages which will be described next.

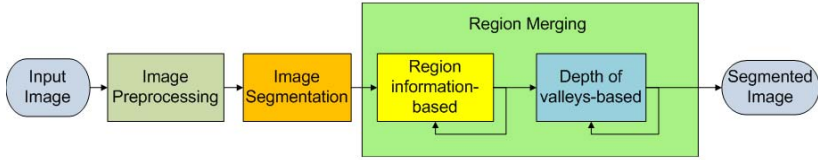


Fig. 2. Scheme of the segmentation algorithm

1. **Image Preprocessing:** Before applying any segmentation algorithm there are some operations that should be done: 1) converting the image into gray-scale, 2) de-interleaving (as our images come from a high definition interleaved video source), 3) correction of the reflections, and 4) inverting the gray-scale image.
2. **Image Segmentation:** We apply watersheds to the gradient image because the boundaries between the regions obtained in such way are closer to the boundaries that separate the different structures that appear in the image (Fig. 3 [4]).

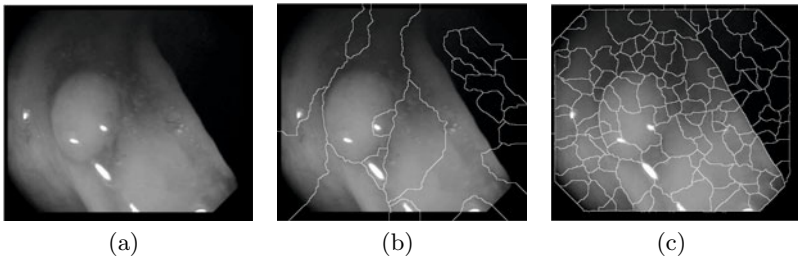


Fig. 3. Use of gradient information: (a) original image (b) basic segmentation (c) Segmentation using gradient information

3. Region Merging:

a) **Region information-based:** We first calculate the neighborhood map of the image and identify the frontier pixels between each pair of regions and then categorize the regions and frontiers, in terms of the amount of information that they contain [4]. For instance, a low information region will have a very high mean (or very low) grey level and very low standard deviation of this grey level. We will only merge regions with the same kind of information separated by weak frontiers. In this case, in order to consider a frontier as weak we propose a frontier weakness measure as defined in Equation 1. This measure combines the information of the the mean gradient of the frontier

pixels and the strength of the frontiers (measured as the frontiers that are kept after applying a two consecutive order increasing median filtering, which helps to remove the regions created by the veins). We merge and categorize regions until their number is stabilized or there are no weak frontiers left.

$$FrontierWeakness = \alpha * gradient + \beta * median \quad (1)$$

b) Depth of valleys-based: We define a *depth of valleys* measure that combines the information of the output of a ridges and valleys detector (see [12] for details) with the information that the morphological gradient provides. This gives information about the depth of the pixels in the valley with higher values for the pixels that constitute the frontier of the region (which will have both high 'valleyness' and gradient values and smaller from the inner pixels, as can be seen in Figure 4). Using this information we can continue merging regions, keeping only those which frontiers are strong in terms of *depth of valleys*. We merge regions until there are no weak frontiers according to the depth of valleys threshold value or when the number of regions is stabilized.

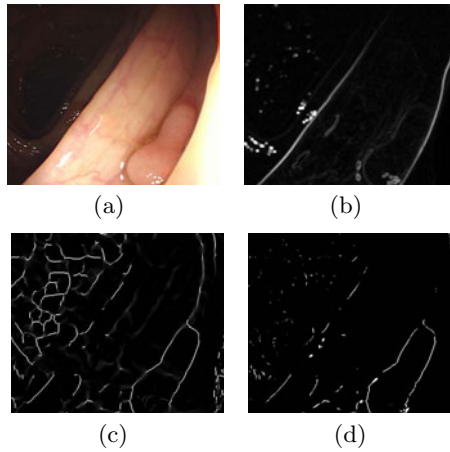


Fig. 4. Creation of the depth of valleys image: (a) Original image (b) Morphological Gradient image (c) Valleys image (d) Depth of valleys image

4 Results

4.1 Experimental Setup

In order to test the performance of our segmentation algorithm we have created a database which consists of 300 different studies of polyp appearance along with their corresponding polyp masks. We will evaluate the performance of our method by using two different measures: Annotated Area Covered (AAC) and Dice Similarity Coefficient (DICE) [7].

Table 1. Comparison between the results obtained by our method and normalized cuts with respect to the value of the depth of valleys

With Borders						
Measure / Method	Ours	NCuts	Ours	NCuts	Ours	NCuts
Threshold Value	0.6		0.7		0.8	
AAC	61.91%	63.66%	70.29%	69.06%	75.79%	70.86%
DICE	55.33%	44.97%	44.6%	37.75%	36.44%	34.01%
Without Borders						
Measure / Method	Ours	NCuts	Ours	NCuts	Ours	NCuts
Threshold Value	0.6		0.7		0.8	
AAC	60.71%	60.2%	70.29%	63.98%	74.32%	64.24%
DICE	55.68%	63.15%	48.01%	61.84%	45.01%	56.73%

Both measures are complementary as the former calculates the amount of annotated polyp area while the latter complements it with the amount of non-polyp information that is kept in the region. We will compare our final segmentation results with the ones obtained using *normalized cuts*. To do so we will set the number of final regions that we have to obtain with *normalized cuts* at the minimum number of final regions that we have obtained with our method that gives us best results in terms of AAC and DICE. To obtain our best results we have run experiments in order to get the combination of parameter values (α , β , depth of valleys threshold) that gives good results for every image.

We also have to consider that our images have black borders around them. Our region segmentation method consider their presence and use the results of non-informative region identification (which borders of the image are part of) to avoid further processing of this areas. In order to test the effect that these borders have in the segmentation results, we have also created a database that eliminates the borders of the images. It has to be mentioned that in order to make the new image suitable to be processed some non-borders parts of the image should also be eliminated which causes the loss of some information. We will compare the results achieved by using these two versions of the images.

4.2 Experimental Results

In Table 1 we show results for polyp region detection comparing the performance of our method, with the performance achieved by *normalized cuts* (with the same number of final regions). Using the whole image we get better results than *normalized cuts* in terms of AAC and DICE. This means that our regions which contain polyps have more polyp information than the ones that *normalized cuts* and it is closer to the real polyp region.

The elimination of the borders in the image, in terms of AAC, has almost no effect for our method but it has more incidence for *normalized cuts*. DICE results improve for both methods by using the image without borders (better as the threshold value increases), although *normalized cuts* results are better. But, as it can be seen in Figure 5, *normalized cuts* non-polyp regions tend

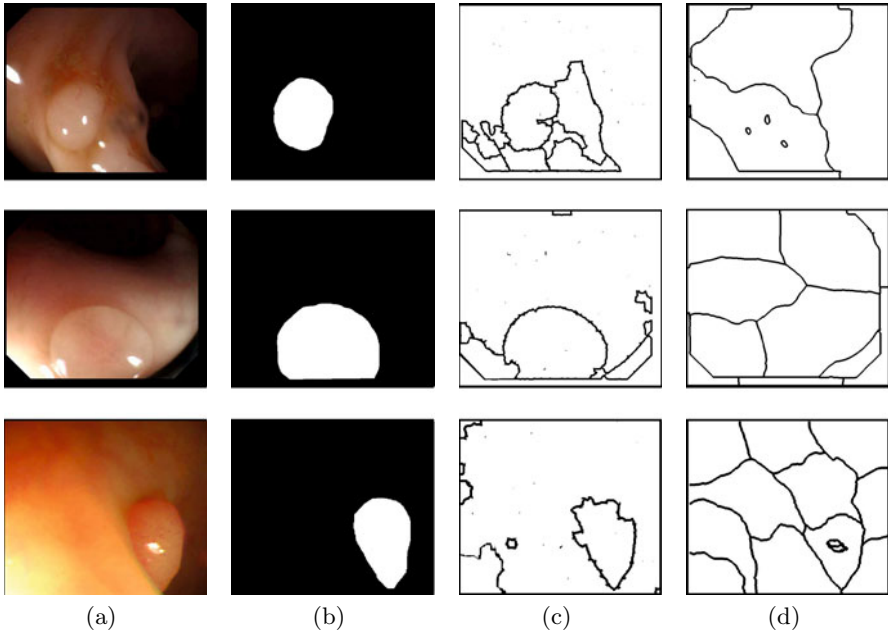


Fig. 5. Comparison of segmentation results: (a) Original images (b) Polyp masks (c) Our method’s output (d) Normalized cuts’ output

to be larger than our non-polyp regions (in our case we know that the larger region corresponds always to the background). In our case, we could apply a size threshold value to discard some of them without having chance of losing the polyp region while in *normalized cuts* this would not happen.

In Figure 5 we can see examples of each method’s output. It can be seen that the images segmented with our method (see Figure 5(c)) fit better the polyp mask (that was segmented by experts). Third row’s results are obtained using the image without borders. We plan to improve our DICE results by merging some small regions that appear inside the polyp region and, after this is achieved, our overall region segmentation results by discarding some of the smallest regions in order to provide as result a very low number of relevant regions.

5 Conclusions and Future Work

In this paper, in the context of region segmentation in colonoscopy images, we present our novel segmentation approach. Our objective is to provide a low number of regions, one of them containing a polyp. Our algorithm also indicates the degree of information of the final regions. Our approach consists of applying a basic segmentation algorithm (such as watersheds) and then applying a region merging algorithm that takes into account our model of polyp appearance. This model states that a polyp is a prominent shape enclosed in a region with presence

of edges and valleys around its frontiers. In order to test the performance of our method we rely on a database of more than 300 studies where the experts have manually segmented the polyps in the images. We compare our method with one state-of-the-art technique (*normalized cuts*) and quantify the accuracy of our segmented regions using two complementary measures. Our method outperforms normalized cuts in the accuracy of polyp region detection and also offers a good non-informative region characterization.

Our future work, in terms of *Region Segmentation*, will be focused on reducing even more the number of final regions and once we have achieved a better performance (in terms of both AAC and DICE) we plan to start with the *region description* and *region classification* stages.

Acknowledgements

The authors would like to thank Farhan Riaz for his helpful comments and suggestions. This work was supported in part by a research grant from Universitat Autònoma de Barcelona 471-01-3/08, by the Spanish Government through the founded project “COLON-QA” (TIN2009-10435) and by research programme Consolider Ingenio 2010: MIPRCV (CSD2007-00018).

References

1. Tresca, A.: The Stages of Colon and Rectal Cancer. New York Times (About.com), p. 1 (2010)
2. Hassinger, J.P., Holubar, S.D., et al.: Effectiveness of a Multimedia-Based Educational Intervention for Improving Colon Cancer Literacy in Screening Colonoscopy Patients. *Diseases of the Colon & Rectum* 53(9), 1301 (2010)
3. Bernal, J., Sánchez, J., Vilariño, F.: Current challenges on polyp detection in colonoscopy videos: From region segmentation to region classification. a pattern recognition-based approach. In: *Proceedings of the 2nd International Workshop on Medical Image Analysis and Description for Diagnosis Systems - MIAD 2011, Rome, Italy (January 2011)* (in press)
4. Bernal, J., Sánchez, J., Vilariño, F.: Reduction of Pattern Search Area in Colonoscopy Images by Merging Non-Informative Regions. In: *Proceedings of the XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica, Madrid, Spain (November 2010)* (in press)
5. Ameling, S., Wirth, S., Paulus, D.: Methods for Polyp Detection in Colonoscopy Videos: A Review. *Inst. für Computervisualistik* (2009)
6. Hwang, S., Oh, J., Tavanapong, W., Wong, J., De Groen, P.: Automatic polyp region segmentation for colonoscopy images using watershed algorithm and ellipse segmentation. *Progress in biomedical optics and imaging* 8(33) (2007)
7. Riaz, F., Ribeiro, M.D., Coimbra, M.T.: Quantitative comparison of segmentation methods for in-body images. In: *Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, EMBC 2009*, pp. 5785–5788 (2009)

8. Forsyth, D.A., Ponce, J.: *Computer vision: a modern approach*. Prentice Hall Professional Technical Reference (2002)
9. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(8), 888–905 (2002)
10. Cheng, Y.: Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(8), 790–799 (2002)
11. Vincent, L., Soille, P.: Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(6), 583–598 (1991)
12. López, A.M., Lumbreras, F., et al.: Evaluation of methods for ridge and valley detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(4), 327–335 (1999)