Automatic segmentation of polyps in colonoscopic narrow-band imaging data

Melanie Ganz, Xiaoyun Yang, and Greg Slabaugh

***Abstract*—Colorectal cancer is the third most common type of cancer worldwide. However, this disease can be prevented by detection and removal of precursor adenomatous polyps during optical colonoscopy (OC). During OC, the endoscopist looks for colon polyps. While hyperplastic polyps are benign lesions, adenomatous polyps are likely to become cancerous. Hence it is common practice to remove all identified polyps and send them to subsequent histological analysis. But removal of hyperplastic polyps poses unnecessary risk to patients and incurs unnecessary costs for histological analysis. In this paper, we develop the first part of a novel optical biopsy application based on narrow-band imaging (NBI). A barrier to an automatic system is that polyp classification algorithms require manual segmentations of the polyps, so we automatically segment polyps in colonoscopic NBI data. We propose an algorithm, Shape-UCM, which is an extension of the gPb-OWT-UCM algorithm, a state of the art algorithm for boundary detection and segmentation. Shape-UCM solves the intrinsic scale selection problem of gPb-OWT-UCM by including prior knowledge about the shape of the polyps. Shape-UCM outperforms previous methods with a specificity of 92%, a sensitivity of 71% and an accuracy of 88% for automatic segmentation of a test set of 87 images.**

I. INTRODUCTION

Colorectal cancer is the third most common cancer in men (663,000 cases, 10.0% of the total) and the second most common cancer in women (571,000 cases, 9.4% of the total) worldwide. Colon cancer accounts for 8% of all deaths by cancer, making it the fourth most common cause of death from cancer [1].

However, this disease can be prevented by detection and removal of precursor adenomatous polyps during optical colonoscopy (OC), an endoscopic examination of the colon using a flexible video camera. During OC, the endoscopist looks for anomalous growths, which typically fall into one of two categories: hyperplastic or adenomatous polyps. Hyperplastic polyps are benign lesions that confer little clinical risk of developing into cancer and do not require removal from the colon. In contrast, adenomatous polyps are premalignant tumours that, if left unchecked, are likely to become cancerous and are surgically removed by the endoscopist during a polypectomy.

This decision to remove or leave alone requires expertise beyond that of many endoscopists. Consequently, it is common to remove all identified polyps for subsequent histological analysis [2]. But removal of hyperplastic polyps poses unnecessary risk to patients (as polyp removal carries a risk of colon perforation requiring emergency surgery) and incurs unnecessary costs for histological analysis [3]. This workflow could be significantly improved if there was a way to perform an immediate in vivo biopsy of a polyp during OC [2].

In this project we are interested in developing the first part of a novel optical biopsy application for optical colonoscopy. Several groups have researched optical biopsy based on manual annotations of colon polyps in narrow-band imaging (NBI) [4]–[6]. These optical biopsy algorithms have shown much promise, with preliminary results of 80-90% classification accuracy of hyperplastic and adenomatous polyps from NBI images. For reference, an expert endoscopist has an accuracy of approximately 90% [2].

However, a barrier to a fully automatic optical biopsy system is the limitation that current methods require manual segmentations of the polyps. This task is tedious and unsuitable for practical clinical deployment. In this paper we seek to automatically segment the polyp in the image, thereby facilitating the later automation of an entire optical biopsy system.

A. Related Work

The only papers the authors are aware of that have tackled a similar task of automatically segmenting colon polyps from NBI data are [7], [8] and [9]. All three methods employ quite different strategies.

In [7], Gross et al. make the first attempt at performing polyp segmentation in NBI colonoscopy data. They employ nonlinear diffusion filtering followed by a Canny edge detector to detect the edges of polyps and then use template matching to identify the polyp. The article gives some results on localization, but no details are given regarding the segmentation accuracy.

Breier et al. [8] followed two years later with two different approaches based on active contours and active rays that also incorporate multi-scale processing. They were able to achieve a specificity of 98-100%, with a sensitivity of 3-32% and an accuracy of 45-58% on a set of 184 polyp images, where the polyp was present in the center of the image.

Finally, in [9] Breier et al. improve their segmentation performance by performing a Chan-Vese-segmentation to decrease their specificity to 86%, while increasing their sensitivity and accuracy to 48% and 62%, respectively, on the same data set that was used in [8].

B. Our Contributions

Our contributions in this paper are three-fold:

First, we propose a new method, called Shape-UCM, which is an extension of the gPb-OWT-UCM algorithm, a state of the art algorithm for boundary detection and segmentation. We are the first, to our knowledge, to apply gPb-OWT-UCM to medical imaging. Furthermore, gPb-OWT-UCM has an intrinsic scale selection problem. We have solved this problem by using prior knowledge about the shape of the objects we want to segment. This way we have developed a completely parameter-free algorithm that automatically segments polyps in NBI data. Finally, in our experiments we can show that our algorithm, Shape-UCM, outperforms previous work in this area by a large margin.

C. *Outline of the Paper*

In the following we will introduce the concept of narrow-band imaging (Section II). Then we describe our proposed algorithm (Section III) and perform experiments to evaluate its performance (Section IV). Finally, we discuss our results (Section V) and give a conclusion (Section VI).

II. NARROW-BAND IMAGING

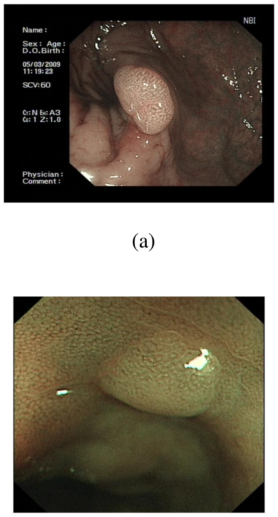
Narrow-band Imaging (NBI) is a recently introduced multi-spectral endoscopic imaging technique that uses several (two or three) discrete bands of light to image tissue. This technique improves visibility of blood vessels, both on the tissue surface (superficial) and below (subepithelial), due to the fact that the depth penetration into the tissue mucosa depends on the wavelengths of light used [10]. NBI became commercially available in 2005 from Olympus. Today endoscopic diagnosis has been rapidly progressing and has now advanced to the area of pathology. The feasibility of NBI has been examined [11] and it is now used for diagnosis of malignant and premalignant gastrointestinal lesions [12]. There has also been a large population trial regarding optical diagnosis of small colorectal polyps based on NBI [2], which showed that the clinical workflow could be significantly improved if there was a way to perform an immediate in vivo biopsy of a polyp during optical colonoscopy. In our case the images are given as screen captures as RGB images. The screen captures also include black background with writing of various sizes. To facilitate automatic segmentation, we will remove this background in our pre-processing procedure. Furthermore, in our dataset the polyps are not always in the center of the image, but can also be situated at the image boundary. In this paper we assume that the endoscopist places the polyp in the center of the image (this is for example the case in the dataset used by [8] and [9]). Hence in our dataset for the cases where the polyp is not in the center, we center it by choosing the segmented region that has the center of the manual annotation in it.
An example of two original images can be seen in Figure 1.

III. AUTOMATIC POLYP SEGMENTATION

In the following we give an overview of the pre-processing and the Shape-UCM algorithm we have devised to perform automatic polyp segmentation from NBI data. A schematic view of the different steps is given in Figure 2.

A. *Pre-processing*

Pre-processing is necessary to remove the irregular background around the image that is due to the images being screen captures and to remove specular reflections that would produce artificial boundaries during the segmentation part of the algorithm.


(a)

(b)

Fig. 1: Two NBI images showing (a) an adenomatous and (b) a hyperplastic polyp.

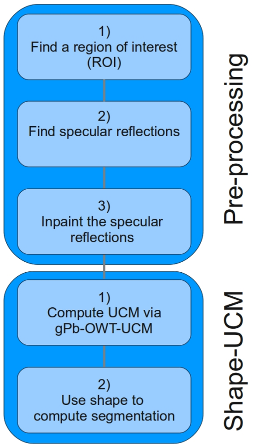


Fig. 2: Overview of our Shape-UCM algorithm for automatic polyp segmentation.

1) Finding a region of interest: First, we have to find a region of interest (ROI). This is done in a two step procedure: based on a simple thresholding, we identify the upper left and the lower right corner of the non-black part of the image and retrieve a rough region of interest. Since the first step sometimes still yields a thin black rim around the image, we additionally extract line segments associated with particular bins for horizontal and vertical lines in a Hough transform [13]. These extracted lines can then be employed to refine the ROI further and remove the leftover thin black rim. The final region of interest is shown in Figure 3 (a) and (b).

2) Removing specular reflections: The second step of the pre-processing is to find the specular reflections. This is done by transforming the image from RGB to HSV (hue, saturation, value) color space and then employing hysteresis thresholding, a bi-threshold procedure typically used for two class object-background pixel segmentation [14]. The hysteresis thresholding consists of three steps. First we threshold once (in the same fashion as described in [4]) using the function

$$f(x) = \begin{cases} \text{true}, & \text{if } (\text{Saturation} < t_1) \cap (\text{Value} > t_2) \\ \text{false}, & \text{otherwise} \end{cases} \tag{1}$$

for each image pixel x in HSV color space. This yields only the high confidence pixels. Next a dilation with a circular structuring element of the size of 5 pixels is performed on the high confidence pixels. Then the same thresholding is applied again with updated parameters t_1 and t_2 yielding a second image after weak thresholding. Finally, we combine the two thresholded images by choosing from among the object pixels selected by the second (weak) thresholding only those pixels connected to pixels in the first (high confidence) thresholding. The thresholds were experimentally determined and are $t_1 = 0.29$ and $t_2 = 0.65$ for the first and $t_1 = 0.22$ and $t_2 = 0.8$ for the second thresholding. The outcome of identifying the

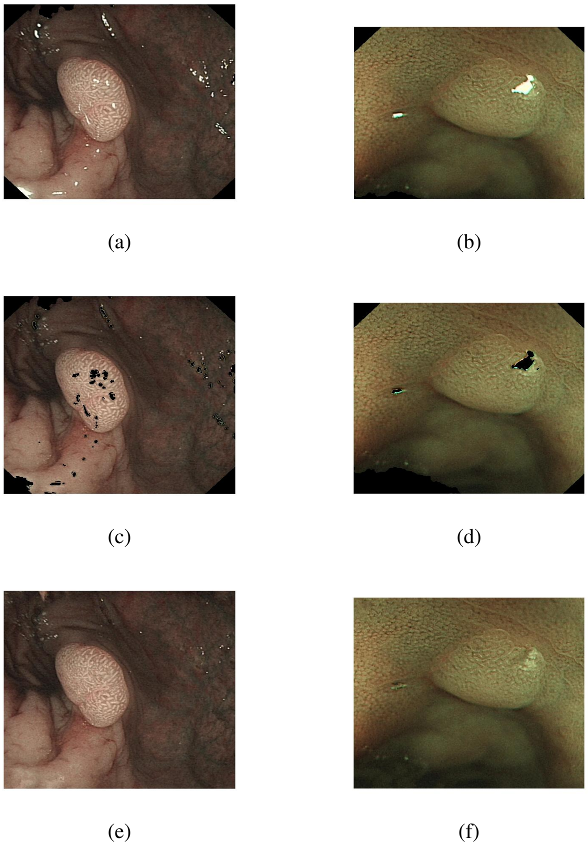

(a) (b) (c) (d) (e) (f)

Fig. 3: Above the three pre-processing steps are shown: The result of finding the region of interest of the example images introduced in Figure 1 can be seen in (a) for an adenomatous and (b) a hyperplastic polyp. After determining the region of interest the specular reflections are identified. The results of identifying the specular reflections are shown in (c) and (d). The last step of the pre-processing consists of inpainting the areas of specular reflections. (e) and (f) display the result of the inpainting.

specular reflections can be seen in Figure 3 (c) and (d).

3) Inpainting: The last step of the pre-processing consist of the application of an exemplar based inpainting method [15], [16] to reconstruct the areas where the specular reflections lead to a loss of image information. We need to reconstruct the specular highlights, since their distinct and bright visual appearance leads to artificial edges and texture that affect the segmentation [17]. Results of the inpainting can be observed in Figure 3 (e) and (f).

B. The Shape-UCM Algorithm

Our algorithm, Shape-UCM, consists of two steps. First we compute a ultrametric contour map (UCM) by using the gPb-OWT-UCM algorithm. Then we employ prior knowledge about the shape of the polyps in order to automatically identify the best segmentation level in the UCM for our problem to segment the polyp surface from the NBI image.

1) Finding the UCM: The first part of the Shape-UCM algorithm we used to segment polyps is described in [18] and in the following we will refer to the method as gPb-OWT-UCM algorithm.

The gPb-OWT-UCM algorithm transforms an image into LAB (lightness and a and b color-opponent dimensions) color space (see Figure 4 (a)-(c)) and a texture representation (see Figure 4 (d)). The texture repreentation is formed by converting the original image to grayscale and convolving it with a set of 17 Gaussian derivative and center-surround filters (details for the filters are given in [18]). The resulting (17-dimensional) vectors are then clustered using K-means and each pixel in the texture representation image is assigned the integer id $\in [1, .., K]$ of the closest cluster center.

Next the gPb-OWT-UCM algorithm finds contours in the image by calculating the gradient magnitude G at each location (x, y) of the image by employing the χ^2 distance of histograms between two regions surrounding the location (x, y). It then combines the multi-scale cues from brightness, colour and textures gradients and forms a multi-scale signal at different orientations θ,

$$mPb(x, y, \theta) = \sum_s \sum_i \alpha_{i,s} G_{i,\sigma(i,s)}(x, y, \theta) \tag{2}$$

where s indexes scales, i indexes feature channels and $G_{i,\sigma(i,s)}(x, y, \theta)$ measures the histogram difference in channel i between two halves of a disk of radius $\sigma(i, s)$. Since $mPb(x, y, \theta)$ is sampled over different orientations, its maximum response over all angles is determined to measure boundary strength

$$mPb(x, y) = \max_\theta \{mPb(x, y, \theta)\}. \tag{3}$$

An example can be seen in Figure 5. Next, gPb-OWT-UCM incorporates global knowledge by employing spectral clustering. The local cues computed by applying oriented gradient operators at every location in the image are combined and yield a global eigenvalue problem. Specifically, the spectral clustering is done by constructing a sparse symmetric affinity matrix W,

$$W_{i,j} = exp\left(-\max_{p \in \overline{ij}} \{mPb(p)\}/\rho\right) \tag{4}$$

where p is any pixel along the line segment $\overline{ij}$ connecting pixels i and j and ρ is a constant. Then the different eigenvectors $\mathbf{v}_k$ of W including their eigenvalues λ_k are used to construct the spectral component of the boundary detector:

$$sPb(x, y, \theta) = \sum_{k=1}^{n} \frac{1}{\sqrt{\lambda_k}} \cdot \nabla_\theta \mathbf{v}_k(x, y) \tag{5}$$

Ultimately the gPb-OWT-UCM algorithm forms a final global probability of boundary (gPb) measure [19] as a weighted sum of local and spectral (global) signals:

$$gPb(x, y, \theta) = \sum_s \sum_i \beta_{i,s} G_{i,\sigma(i,s)}(x, y, \theta) + \gamma \cdot sPb(x, y, \theta). \tag{6}$$

Here, the weights $\beta_{i,s}$ and γ can be learned. We use the same weights as [18].

After the contours are identified, it employs an oriented watershed transform (OWT) to form an over-segmentation

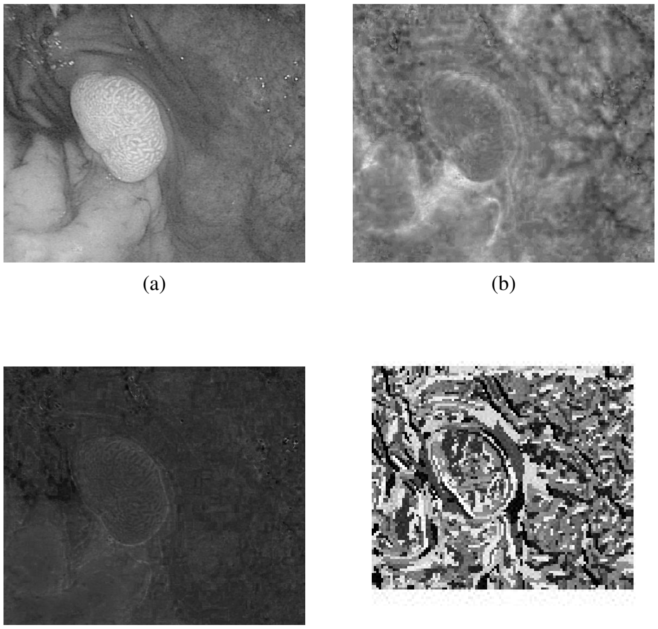


Fig. 4: The LAB color space ((a),(b) and (c)) as well as the texton representation (d) for the image shown in Figure 3 (e). The texton representation (d) results from convolution with 17 derivative filters, followed by clustering. Please see [19] for details.

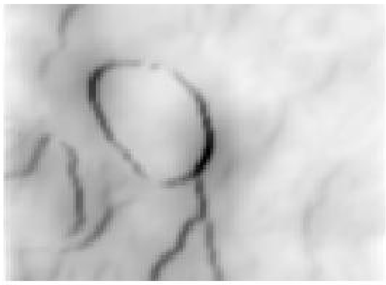


Fig. 5: An example of $mPb(x,y)$ for the image shown in Figure 3 (e). Another example of an $mPb(x,y)$ image can be seen in figure 6 of [19] in the lower right corner.

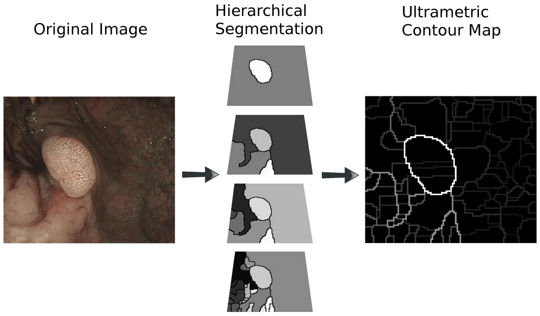
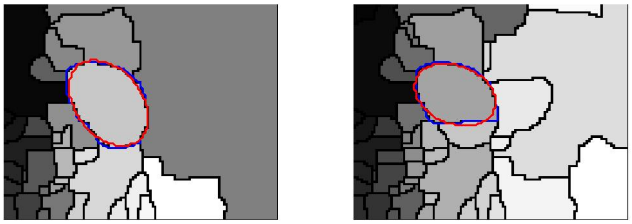


Fig. 6: An ultrametric contour map (UCM) can be constructed by performing boundary detection at different scales and then combining these hierarchical contours into a single image where the boundary strength reflects the level of hierarchy at which the boundary appears.



(a) UCM level = 0.0784 (b) UCM level = 0.0674

(c) UCM level = 0.0537 (d) UCM level = 0.0520

Fig. 7: We can see the gPb-OWT-UCM segmentation output at four different levels (all levels are distributed between [0,1]). The boundaries surrounding the center of the image are shown in blue and to these boundaries we fit an ellipse which is displayed in red. Then the level is chosen at which the ellipse has the highest area overlap with the underlying region.

whose regions determine the highest level of segmentation details; the boundary strength can then be used as an estimate of how likely the contour is [20]. Finally, the gPb-OWT-UCM algorithm constructs a hierarchical region tree in the form of an ultrametric contour map (UCM) [21]. An example of a UCM is shown in Figure 6. There are parameters in the gPb-OWT-UCM algorithm that must be chosen, such as the scales used for the gradient construction in the different channels. For the brightness and texture channels we used a sigma of 5 pixels, whereas a sigma of 10 pixels was used in the case of the color channels.

2) Shape Matching Approach: From the hierarchy of regions given by the UCM, our algorithm automatically determines the level of the region tree at which the partitioning is not an under-, but also not an over-segmentation. We employ a shape matching technique to choose this cutoff point for each image individually and fully automatic. Starting from the gPb-OWT-UCM segmentation output we extract only the boundaries of the central region at each level of the UCM. Then we fit an ellipse using [22] to the points of the boundary to segment the polyp. This kind of ellipse fitting is done for every level of the ultrametric region tree. Finally, we choose the level of the region tree at which the ellipse fits the underlying region best by measuring the area overlap (Eq. 7) at each level. This way we identify the hierarchical level at which the polyp presented is most ellipse-like. An example of ellipse fittings at four different levels of the UCM is given in Figure 7.

C. Comparison to a brute force approach

As already stated from the hierarchy of regions given by the UCM, our algorithm automatically determines the level of the region tree at which the partitioning is not an under-, but also not an over-segmentation. Apart from the Shape-UCM algorithm, we have tested several ways of automatically determining the operating point of the UCM for each image separately:

1) Starting from the top of the UCM, so from an under-segmentation, we progress through the hierarchy of segmentations until there are at least n regions present. A parameter sweep over training data is then used to determine the optimal n.
2) Alternatively, instead of determining the operating point of the UCM by the presence of a certain number of regions, we restrict the maximal size of the regions in the segmentation. Starting from the top of the UCM, we lower the threshold until no single segmented region covers more than n_{area} of the area of the image. Again a parameter sweep over the training data is used to determine the optimal n_{area}.
3) Lastly, instead of requiring a fixed number n of regions to establish the point of optimal segmentation, we find our optimal threshold of the UCM by performing a parameter search over the proportional region number. The proportional region number n_{prop} is given by the optimal region number n_{opt} for each image divided by the maximal region number $n_{\max}$. Here, the maximal region number is the maximum number of regions at the bottom of the UCM.

The results of the parameter sweep for region number n, region size n_{area} and proportional region number n_{prop} can be seen in Figure 8. There we can observe that using n_{prop} for thresholding yields in general higher area overlap, than using n or n_{area}. $n_{\text{prop}} = 0.3$ fetches the best area overlap and we will use it as the threshold for our brute force gPb-OWT-UCM approach in the rest of the paper.

IV. EXPERIMENTS

A. Data

Our data consist of two datasets from two different endoscope systems (Lucera and Exera) that each consist of adenomatous and hyperplastic polyps (see Figure 9). We construct a training set for parameter fitting of the gPb-OWT-UCM algorithm by randomly choosing 40% of each dataset and polyp categories and use the rest for testing. Hence the training set consists of 58 (14 adenomatous and 7 hyperplastic polyps of dataset 1 and 18 adenomatous and 19 hyperplastic polyps of datset 2) and the testing set of 87 images. Further details about the two different datasets are given below.

1) Dataset 1: Dataset 1 consists of 52 images. 35 show adenomatous and 17 hyperplastic polyps. The data was acquired from St. Marks Hospital and Academic Institute as well as Oxford Radcliffe Hospitals. It was provided by Dr. Ana Ignjatovic with the support of Dr. Brian P. Saunders, James E. East, and David Burling. The images were acquired with an Olympus Evis Lucera Spectrum endoscope video system with a CV-260 video processor. This type of endoscope is mostly used in Japan and the UK [23].

2) Dataset 2: Dataset 2 is made up of 46 adenomas and 47 hyperplasts. The second dataset was provided by Douglas K. Rex from Indiana University. These images were taken with an Olympus Evis Exera II endoscope video system with a CV-180 video processor, which is common in Europe and North America [23].

The data was provided to Imperial College, which then cooperated with us.

B. Evaluation Methodology

We consider various region-based measures to evaluate the quality of our segmentation algorithms with regard to manual segmentations and to be able to compare them to previous work.

The area overlap (or Jaccard index) between two regions A and B is defined as

$$\mathcal{J} = \frac{|A \cap B|}{|A \cup B|} = \frac{TP}{FP + TP + FN}, \tag{7}$$

where TP denotes true positives, FP false positives and FN false negatives. It is commonly used to measure segmentation quality in recognition.

To be able to compare our results to previous work on NBI polyp segmentation we also measure the sensitivity given by

$$\mathcal{S}_1 = \frac{TP}{TP + FN} \tag{8}$$

as well as the specificity

$$\mathcal{S}_2 = \frac{TN}{TN + FP}. \tag{9}$$

(TN denote true negatives) and the accuracy

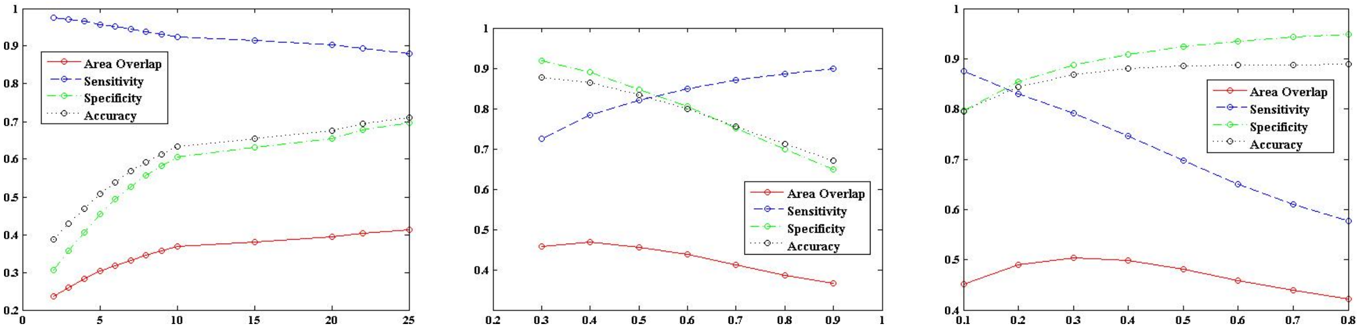
$$\mathcal{A} = \frac{TP + TN}{TP + FP + FN + TN}. \tag{10}$$

C. Inter Observer Variability

To assess the variability occurring between different individuals performing the polyp segmentation, we have acquired three sets of manual annotations for dataset 1. The first was made by an experienced endoscopist (Observer 1) and the other two were made by imaging researchers (Observer 2 and Observer 3), but afterwards checked and corrected by the experienced endoscopist. To check the inter-observer variability we make pairwise comparisons of the manual annotations of the three observers for the adenomatous as well as the hyperplastic polyps. The results can be seen in Table I.

D. Shape-UCM vs. thresholded gPb-OWT-UCM

Finally, we compare our Shape-UCM algorithm directly with the brute force method of Section III-C. In contrast to the gPb-OWT-UCM algorithm, the Shape-UCM algorithm does not need parameter tuning and can be run directly on the training and the testing datasets. The results for both datasets divided in training and testing are displayed in Tables II and III. Examples of the resulting segmentations can be seen in Figures 10 and 11.



(a) n (b) n_{area} (c) n_{prop}

Fig. 8: Results for the gPb-OWT-UCM algorithm parameter sweeps to determine the optimal cutoffs of the UCM. (a) shows the parameter sweep for optimal region number n, while (b) displays it for the maximal region size n_{area} and (c) for the proportional region number n_{prop}. In each case the parameter is fixed at the value which yields the largest area overlap.

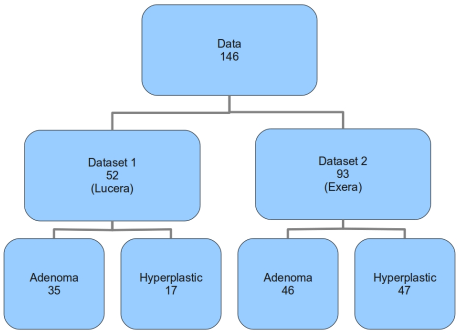


Fig. 9: Overview over our data: Our data consist of two datasets from two different endoscope systems (Lucera and Exera) that each consist of adenomatous and hyperplastic polyps. We construct a training set for parameter fitting of the gPb-OWT-UCM algorithm by randomly choosing 40% of each dataset and category, the rest is used for testing.

TABLE I: Our quantitative inter observer results for (a) the adenoma and (b) the hyperplasts of dataset 1.

Adenoma	Area Overlap	Specificity	Sensitivity	Accuracy
Obs.1 vs. Obs.2	0.83	0.99	0.86	0.97
Obs.1 vs. Obs.3	0.83	0.99	0.89	0.97
Obs.2 vs. Obs.3	0.88	0.98	0.96	0.98

(a)

Hyperplasts	Area Overlap	Specificity	Sensitivity	Accuracy
Obs.1 vs. Obs.2	0.82	0.99	0.85	0.95
Obs.1 vs. Obs.3	0.81	0.97	0.87	0.94
Obs.2 vs. Obs.3	0.89	0.95	0.96	0.97

(b)

TABLE II: Our quantitative results on the training dataset for Shape-UCM (M1) and the gPb-OWT-UCM algorithm (M2).

Method	Area Overlap	Specificity	Sensitivity	Accuracy
M1	0.52 ± 0.28	0.90 ± 0.15	0.78 ± 0.26	0.88 ± 0.14
M2	0.53 ± 0.23	0.95 ± 0.09	0.72 ± 0.25	0.92 ± 0.09

TABLE III: Our quantitative results on the testing dataset for Shape-UCM (M1) and the gPb-OWT-UCM algorithm (M2).

Method	Area Overlap	Specificity	Sensitivity	Accuracy
M1	0.49 ± 0.27	0.92 ± 0.17	0.71 ± 0.29	0.88 ± 0.17
M2	0.44 ± 0.22	0.95 ± 0.08	0.62 ± 0.29	0.90 ± 0.10

V. DISCUSSION

A. *Inter Observer Variability*

Focusing first on the inter-observer variability, we can see in Table I that the inter-observer variability is relatively small considering the variability of the data. Especially, accuracy is extremely high and lies between 94% and 98%. Next, there is no apparent difference in the inter observer variability when annotating adenoma or hyperplasts. This shows that for human observers the task of segmenting an adenoma is equally challenging to segmenting an hyperplast. Furthermore, the experienced observer (Observer 1) agrees less with the two inexperienced observers (Observer 2 and 3) than they agree with each other. This could be the case, because the inexperienced observers make the same errors for challenging polyps whereas the experienced observer annotates more precisely.

B. *Polyp Segmentation*

Next we turn to the automatic segmentation task. The results of the fully automatic segmentation of polyps are shown in Tables II and III for the Shape-UCM and the gPb-OWT-UCM algorithm.

Table II displays the results of both algorithms for the training dataset. There is no statistically significant difference between the two algorithms for training set performance. The area overlap lies at 52-53%, the specificity between and 90%-95%, the sensitivity between 72%-78% and the accuracy between and 88%-92%. While the specificity for the training dataset is only 5-10% lower than that of the manual observers, the sensitivity and area overlap are much lower. Conversely, the accuracy is only about 10% lower.

Table III displays the results for the testing data. In general the results are not very different from the ones for the training data. For the testing data the area overlap is between 44%-49%, the specificity between 92%-95%, the sensitivity

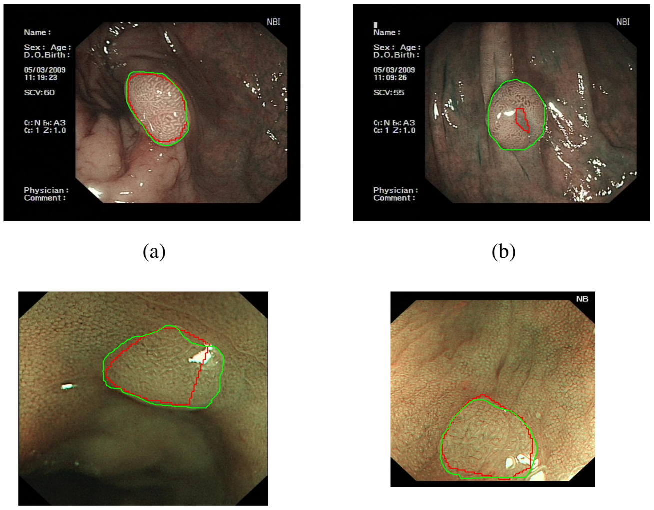

(a) (b) (c) (d)

Fig. 10: Our segmentation results for the gPb-OWT-UCM algorithm [18] for two adenomatous (a) - (b) and two hyperplastic polyps (c) - (d): In green the manual annotation and in red the gPb-OWT-UCM segmentation is shown.

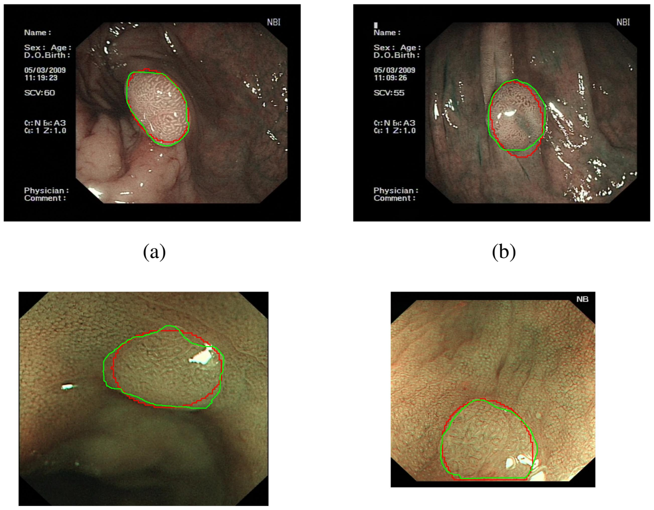
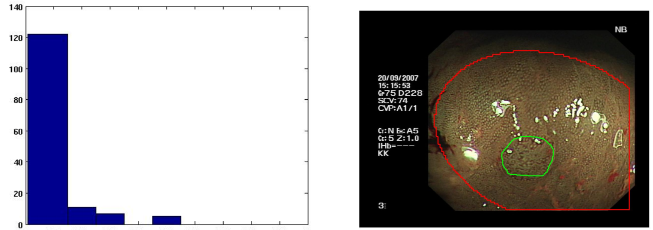

(a) (b) (c) (d)

Fig. 11: Segmentation results for the Shape-UCM algorithm (III-B2) for two adenomatous (a) - (b) and two hyperplastic polyps (c) - (d): In green the manual annotation and in red the Shape-UCM segmentation is shown.

between and 62%-72% and the accuracy between and 88%-90%. But what is interesting for the testing data is that, as expected, the gPb-OWT-UCM algorithm performs worse on the testing then the training data. Especially area overlap and sensitivity drop for the gPb-OWT-UCM algorithm, whereas the Shape-UCM algorithm has a slightly lower sensitivity, but otherwise performs just as well on the test data. Due to the large variations there is again no statistical significant difference between the two algorithms.

Although both algorithms perform equally well and there is no statistically significant difference between them, the big advantage of the Shape-UCM algorithm is its parameter-free nature. Whereas the brute force approach needs a training set to determine the optimal threshold of the UCM, Shape-UCM finds this threshold automatically which is a big advantage.

Furthermore, currently there is a large variation in polyp size in our dataset (see Figure 12(a)). This can lead to erroneous segmentations such as shown in Figure 12(b). If we were able to clinically restrict the size of polyps in the NBI images to e.g. $< 50\%$ of the area of the image, the Shape-UCM approach could be modified to prevent polyp segmentations like in Figure 12(b) that currently achieve only low area overlap. In general, the segmentation results are very good (see e.g. Figure 10 and 11), but have room for improvement due to the gap between the human and automatic performance. Especially, the area overlap and the sensitivity has room for improvement. The reason why automatic segmentation performance is lower is the fact that the task is really challenging due to the biological variability of the data in shape as well as the similarity in color and texture between adenomatous and hyperplastic polyps and the different scale of the polyps in the images resulting from various zoom. A comprehensive atlas showing the large variation in NBI images of polyps can be found in [23].



(a) Polyp coverage % of the image (b)

Fig. 12: In (a) we can see the distribution of polyps in all of our data after size. The size varies between a coverage of 5% and 55%. (b) shows an example of a bad segmentation (red) for a small polyp (green). If our dataset would be more homogeneous in size, Shape-UCM could be refined and errors like this could be avoided.

C. *Comparison To Previous Work*

To provide a fair assessment of our results, we also compare it to the results of previous work in [7]–[9]. In [7] only results regarding the localization are given and the number of polyps detected with a certain area, but no pixel based measurements are given, which makes comparison challenging. Conversely, in [8] and [9] pixel based measurements of the segmentation quality are given. A direct comparison of our results on the testing dataset to the results given in [8] and [9] can be seen in Table IV.

The direct comparison shows the improvement in segmentation. While maintaining 92% specificity, we improve sensitivity from 30% and 48% in [8] and [9] to 71% for the test dataset. Consequently, also the accuracy increases by more than 20% from 59% and 62% in [8] and [9] to 88%.

TABLE IV: Comparison of our results on the test dataset with the results in previous work in [8] and [9] given as mean ± standard deviation.

Method	Specificity	Sensitivity	Accuracy
[8] (184 polyps)	0.98 ± 0.04	0.30 ± 0.26	0.59 ± 0.23
[9] (184 polyps)	0.86 ± 0.02	0.48 ± 0.21	0.62 ± 0.16
Shape-UCM (87 polyps)	0.92 ± 0.17	0.71 ± 0.29	0.88 ± 0.17

VI. CONCLUSION

In this paper we have developed the first part of a novel application for optical colonoscopy based on narrow-band imaging (NBI) by automatically segmenting polyps in colonoscopic NBI data. We developed a new algorithm, called Shape-UCM, which is an extension of the gPb-OWT-UCM algorithm, by including prior knowledge about the shape of the polyps. Hence our Shape-UCM algorithm can automatically determine the optimal segmentation from a selection of hierarchical segmentations and needs no parameter tuning. Our results outperformed previous work on automatic polyp segmentation in NBI data and yielded a specificity of 92%, a sensitivity of 71% and an accuracy of 88% for automatic segmentation in a testing dataset of 87 polyp images from two different endoscope systems.

In conclusion, our automatic segmentation results are promising and in future work we will combine them with a polyp classification procedure to build a fully automatic optical biopsy system. While our results have shown that it is possible to automatically segment polyps in NBI data, there is still a gap to the performance of a human observer. On the one hand, we could try to improve our existing algorithm by using different metrics inside the gPb-OWT-UCM algorithm, such as for example a distance measure from the recently proposed quadratic chi distance family [24]. On the other hand we could try to refine our Shape-UCM approach if the dataset collection and hence size distribution of polyps in the NBI images could be clinically restricted. While the task is left to improve the segmentation of bowel polyps, the segmentation task could also be extended into other organs which are inspected by endoscopy using NBI technology, such as for example the stomach and esophagus.

ACKNOWLEDGEMENT

The authors would like to thank their collaborators from St. Marks Hospital, Ocford Radcliffe Hospital, Indiana University and Imperial College for providing the NBI data. We would especially like to thank Ana Ignatovic for providing expert segmentations.

REFERENCES

[1] J. Ferlay, H. Shin, F. Bray, D. Forman, C. Mathers, and D. Parkin. (24-08-2011) Cancer incidence and mortality worldwide: Iarc cancerbase no. 10. [Online]. Available: http://globocan.iarc.fr

[2] A. Ignjatovic, J. East, N. Suzuki, M. Vance, T. Guenther, and B. Saunders, "Optical diagnosis of small colorectal polyps at routine colonoscopy (detect inspect characterise resect and discard; discard trial): a prospective cohort study," *The Lancet Oncology*, vol. 10, no. 12, pp. 1171–1178, 2009.

[3] W. Kessler, R. Klein, R. Wielage, and D. Rex, "Cost savings of removing diminutive polyps without histologic assessment," *Gastrointestinal Endoscopy*, vol. 67, no. 5, pp. AB105–AB105, 2008.

[4] T. Stehle, R. Auer, S. Gross, A. Behrens, J. Wulff, T. Aach, R. Winograd, C. Trautwein, and J. Tischendorf, "Classification of colon polyps in nbi endoscopy using vascularization features," in *Proceedings of SPIE*, vol. 7260, 2009, pp. 1–12.

[5] A. Varnavas, A. Ignjatovic, A. Bharath, J. East, B. Saunders, J. Ng, and D. Burling, "Classification of colon images using the magnitude of orientation dominance," in *Postgraduate Conference in Biomedical Engineering & Medical Physics*, 2009, p. 13.

[6] T. Tamaki, J. Yoshimuta, T. Takeda, B. Raytchev, K. Kaneda, S. Yoshida, Y. Takemura, and S. Tanaka, "A system for colorectal tumor classification in magnifying endoscopic nbi images," in *Proceedings of the Asian Conference on Computer Vision*. Springer, 2011, pp. 452–463.

[7] S. Gross, M. Kennel, T. Stehle, J. Wulff, J. Tischendorf, C. Trautwein, and T. Aach, "Polyp segmentation in nbi colonoscopy," *Bildverarbeitung fuer die Medizin*, pp. 252–256, 2009.

[8] M. Breier, S. Gross, A. Behrens, T. Stehle, and T. Aach, "Active contours for localizing polyps in colonoscopic nbi image data," in *Proceedings of SPIE*, vol. 7963, 2011, p. 79632M.

[9] M. Breier, S. Gross, and A. Behrens, "Chan-vese-segmentation of polyps in colonoscopic image data," in *Proceedings of the 15th International Student Conference on Electrical Engineering (POSTER)*, 2011.

[10] M. Sambongi, M. Igarashi, T. Obi, M. Yamaguchi, N. Oyama, M. Kobayashi, Y. Sano, S. Yoshida, and K. Gono, "Analysis of spectral reflectance of mucous membrane for endoscopic diagnosis," in *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, 2000, pp. 1026–1029.

[11] W. Song, L. Michel, D. Adler, J. Conway, D. Diehl, F. Farraye, S. Kantsevoy, R. Kwon, P. Mamula, B. Rodriguez *et al.*, "Narrow band imaging and multiband imaging," *Gastrointestinal Endoscopy*, vol. 67, no. 4, pp. 581–589, 2008.

[12] C. Gheorghe, "Narrow-band imaging endoscopy for diagnosis of malignant and premalignant gastrointestinal lesions," *Journal of Gastrointestinal and Liver Diseases*, vol. 15, no. 1, p. 77, 2006.

[13] R. Duda and P. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[14] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 679–698, 1986.

[15] A. Criminisi, P. Perez, and K. Toyama, "Object removal by exemplar-based inpainting," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2. IEEE, 2003, pp. II–721.

[16] S. Bhat. (2004) Object removal by exemplar-based inpainting. [Online]. Available: http://www.cc.gatech.edu/~sooraj/inpainting/

[17] A. Mirko, G. Anarta, A. Stefan, and L. Gerard, "Automatic segmentation and inpainting of specular highlights for endoscopic imaging," *Journal on Image and Video Processing*, vol. 2010, 2010.

[18] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 898–916, 2010.

[19] D. Martin, C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 530–549, 2004.

[20] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "From contours to regions: An empirical evaluation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2009, pp. 2294–2301.

[21] P. Arbelaez, "Boundary extraction in natural images using ultrametric contour maps," in *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop*. IEEE Computer Society, 2006.

[22] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least-squares fitting of ellipses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 476–480, May 1999.

[23] J. Cohen, *Advanced Digestive Endoscopy: Comprehensive Atlas of High Resolution Endoscopy and Narrowband Imaging*. Blackwell Publishing, 2007.

[24] O. Pele and M. Werman, "The quadratic-chi histogram distance family," in *Proceedings of European Conference on Computer Vision*. Springer, 2010, pp. 749–762.