A Novel Shape Feature Descriptor for the Classification of Polyps in HD Colonoscopy

Michael Häfner[1], Andreas Uhl[2], and Georg Wimmer[2]

[1] St. Elisabeth Hospital, Vienna, Austria
[2] University of Salzburg, Department of Computer Sciences, Salzburg, Austria
{gwimmer,uhl}@cosy.sbg.ac.at

Abstract This work proposes a new method analyzing the shape of connected components (blobs) from segmented images for the classification of colonic polyps. The segmentation algorithm is a novel variation of the fast level lines transform and the resultant blobs are ideal to model the pit pattern structure of the mucosa. The shape of the blobs is described by a mixture of new features (convex hull, skeletonization and perimeter) as well as already proven features (contrast feature). We show that shape features of blobs extracted by segmenting an image are particularly suitable for mucosal texture classification and outperforming commonly used feature extraction methods.
Additionally this work compares and analyzes the influences of image enhancement technologies to the automated classification of the colonic mucosa. In particular, we compare the conventional chromoendoscopy with the computed virtual chromoendoscopy (the i-Scan technology of Pentax). Results imply that computed virtual chromoendoscopy facilitates the discrimination between healthy and abnormal mucosa, whereas conventional chromoendoscopy rather complicates the discrimination.

Keywords: computed virtual chromoendoscopy, chromoendoscopy, i-Scan, texture recognition, HD-endoscopy, segmentation, blob

1 Introduction

Colonic polyps have a rather high prevalence and are known to either develop into cancer or to be precursors of colon cancer. Hence, an early assessment of the malignant potential of such polyps is important as this can lower the mortality rate drastically. As a consequence, a regular colon examination is recommended, especially for people at an age of 50 years and older. The current gold standard for the examination of the colon is colonoscopy, performed by using a colonoscope. Modern endoscopy devices are able to take pictures or videos from inside the colon, allowing to obtain images (or videos) for a computer-assisted analysis with the goal of detecting and diagnosing abnormalities. To enable an easier detection and diagnosis of the extent of a lesion, there are two common image enhancement technologies:

1. Conventional chromoendoscopy (CC) came into clinical use 40 years ago. By staining the mucosa using (indigocarmine) dye spray, it is easier to find and classify polyps.

2. Digital chromoendoscopy is a technique to facilitate “chromoendoscopy without dyes” [1]. The strategies followed by major manufacturers differ in this area:
   - In Narrow band imaging (NBI, Olympus), narrow bandpass filters are placed in front of a conventional white-light source to enhance the detail of certain aspects of the surface of the mucosa.
   - The i-Scan (Pentax) image processing technology [2] is a digital contrast method which consists of combinations of surface enhancement (SE), contrast enhancement (CE) and tone enhancement (TE).
     The FICE system (Fujinon) decomposes images by wavelength and then directly reconstructs images with enhanced mucosal surface contrast.
     Both systems (i-Scan and FICE) apply post-processing to the reflected light and thus are called ”computed virtual chromoendoscopy (CVC)“.

Previous works for the computer assisted staging of colon polyps, which are using endoscopes producing highly detailed images in combination with image enhancement technologies, can be divided in two categories: high-magnification chromoendoscopy ([3]) and high-magnification endoscopy combined with NBI ([4]). In this work we use highly detailed images acquired by a high definition (HD) endoscope without magnification in combination with CC and CVC (the i-Scan technology). To the best of our knowledge, this is the first work for computer assisted colonic polyp classification using HD-endoscopy combined with CVC as well as HD-endoscopy combined with CC. We use three different i-Scan modes:

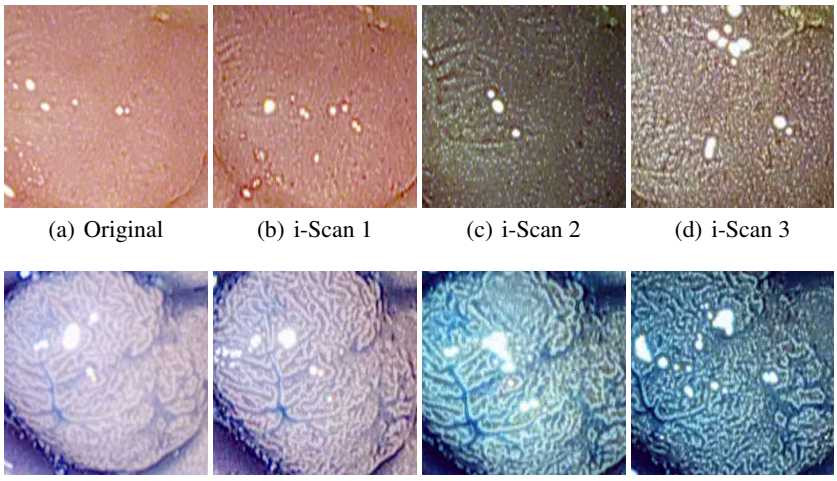
- i-Scan 1 includes SE and CE. This mode enhances surface and pit pattern details, helping to detect dysplastic areas and to accentuate mucosal surfaces.
- i-Scan 2 includes SE, CE and TE. This mode visually enhances boundaries, margins, surface architecture and hard-to-discern polyps.
- i-Scan 3 also includes SE, CE and TE. This mode is similar to i-Scan 2, with increased illumination and emphasis on the visualization of vascular features.

In Fig. 1 we see an image showing an adenomatous polyp without image enhancement technology (a), example images using CVC (b,c,d), an image using CC (e) and images combining CC and CVC by using the i-Scan technology to visually enhance the already stained mocusa (f,g,h).

In this work we will compare classification results with respect to using CVC (i-Scan) or CC. We will also examine the effects of combinations of CVC and CC on the classification results.

For the classification of the images, we propose a new method analyzing the shape of connected components of segmented images.
To find out which image enhancement technologies are most suitable for the computer-aided mucosal texture classification and to compare the results of our proposed method with methods already proven to be successful, we additionally employ a number of well known feature extraction methods for the classification of mucosal texture.

We differentiate between two classes, normal mucosa or hyperplastic polyps (class healthy) and neoplastic, adenomatous or carcinomatous structures (class abnormal) (see



(a) Original (b) i-Scan 1 (c) i-Scan 2 (d) i-Scan 3

(e) CC (f) CC & i-Scan 1 (g) CC & i-Scan 2 (h) CC & i-Scan 3

Figure 1. Images using digital (i-Scan) and/or conventional chromoendoscopy (CC)

Fig. 2 a–d). The various pit pattern types [5] of these two classes are presented in Fig. 2 e–f.

This paper is organized as follows. Section II describes the feature extraction methods, especially our new method based on shapes of connected components (blobs). In section III we describe the experiments and present the results. Section IV presents the conclusion.

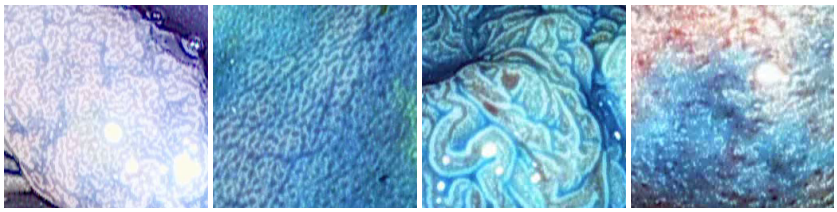
2 Feature extraction

In colonoscopic (and other types of endoscopic) imagery, mucosa texture is usually found at different scales. This is due to varying distance and perspective towards the colon wall and eventually different zoom factors used during an endoscopy session. The differences in scale are for example much higher using HD-endoscopes (especially because of the highly variable distance) than for using high-magnification endoscopes, where the distance of the endoscope to the mucosa is relatively constant. Consequently, in order to design reliable computer-aided mucosa texture classification schemes, the scale invariance of the employed feature sets could be essential.

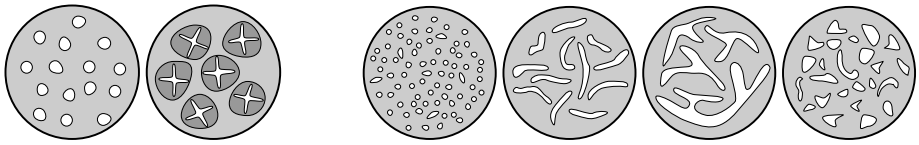
We propose a new scale and rotation invariant feature extraction method denoted as "Segmented Shape Features (SSF)". Similar to the approach of [6], our approach is based on the fast level lines transform (FLLT) [7], which is a fast algorithm decomposing an image I into connected components (blobs S) of upper (X_λ) and lower level set (X^μ):

$$X_\lambda = \{x \in \mathbb{R}^2, I(x) \geq \lambda\}, \qquad X^\mu = \{x \in \mathbb{R}^2, I(x) \leq \mu\}.$$

By grouping the blobs depending on their size, a scale-space representation is created. The family of blobs S is ordered in a tree structure, showing which blob is contained



(a) Healthy (b) Healthy (c) Abnormal (d) Abnormal



(e) Healthy (f) Abnormal

Figure 2. Example images of the two classes (a–d) and the pit pattern types of these two classes (e–f)

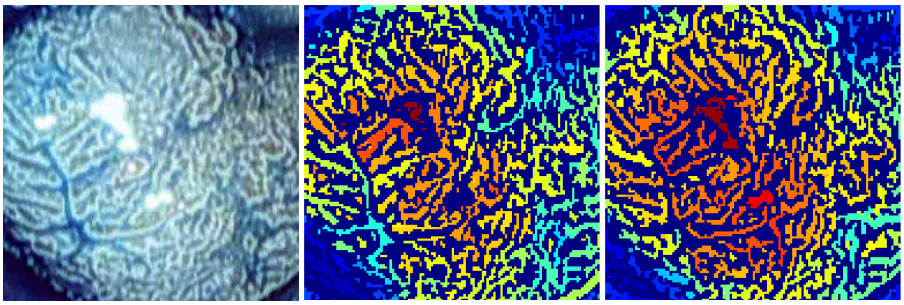
within another.

In case of distinguishing between healthy and abnormal colon mucosa, it is important that the blobs have the right size to represent the typical structures of healthy and abnormal mucosa. It turned out that for too big or too small blobs it is hard to find suitable features for the classification of images. Too small blobs do not contain any discriminative information and for too big blobs it is hard to extract features representing the local mucosal structure (the types of pit patterns). As a solution of this problem we modified (and simplified) the original FLLT algorithm as follows:

Generating dark (bright) blobs R by localized region growing:

1. Scan the image for a not tagged local minimum (maximum for bright blobs) x_0 with gray value g and create a blob R consisting of only x_0 at the first iteration.
2. Find all neighbors N (4-connectivity) of R with $g_N = \min_{x \in N} I(x)$ ($g_N = \max_{x \in N} I(x)$ for bright blobs) and tag them.
3. Two cases are possible:
   - $g \leq g_N$ ($g \geq g_N$ for bright blobs):
     - $R \leftarrow R \cup \{x \in N | I(x) = g_N\}$
     - $g \leftarrow g_N$
     - Return to step 2.
   - $g > g_N$ ($g < g_N$ for bright blobs): Set the gray-levels of the pixels in R to g and go to step 1.

The difference of our simplified FLLT algorithm to that of [7] is that we only record the blobs R when reaching the break condition at step 3 (when $g < g_N$), whereas the original FLLT approach additionally records the R's in step 3 whenever $g < g_N$ (dark blobs) or $g > g_N$ (bright blobs). The original FLLT approach also merges the blobs to bigger



(a) Original (b) Dark blobs (c) Bright blobs

Figure 3. The extracted dark and bright blobs of the original image

blobs resulting in a tree structure of blobs R (with a huge amount of blobs of all possible sizes). For our field of application, it turned out that the blobs R recorded reaching the break condition at the end of step 3 are ideal to distinguish between healthy and abnormal colon mucosa. These dark and bright blobs are well suited (size and shape of the blobs) to model the local pit pattern structure of the mucosa (except for too small blobs –less than 9 pixels– which are not considered for further feature extraction) enabeling the distinction between healthy and affected mucosa . In Fig. 3 we see an example image and the corresponding dark and bright blobs located with our simplified algorithm. The color of a blob in Fig. 3 denotes the averaged brightness of the pixels inside the blob (red is bright, blue is dark).

To extract information about the contrast inside of the blobs we compute the contrast feature CF (used in [6]) as follows: For each pixel x contained in a blob R, a normalized gray value is computed as

$$CF(x) = \frac{I(x) - \text{mean}_{R(x)}(I)}{\sqrt{\text{var}_{R(x)}(I)}}, \tag{1}$$

where $R(x)$ is the blob containing x, $\text{mean}_{R(x)}(I)$ and $\text{var}_{R(x)}(I)$ are the mean and the variance of the image I over R, respectively.
The CF is computed separately for pixels contained in dark and bright blobs, respectively. This results in two contrast feature histograms (CFH), computed by scanning all pixels contained in dark or bright blobs.

Additionally we use three new scale and rotation invariant shape features suitable for mucosal texture classification. The convex hull feature (CH) (see Fig. 4) is showing the proportion of the dilation of a blob R to the density of the blob (the number of pixels of R):

$$CH(R) = \frac{\#\text{ Pixels of Convex Hull}(R)}{\#\text{ Pixels of } R} \tag{2}$$

An example of the convex hull of a blob is shown in Fig. 4. The blob R is shown in gray, the convex hull of R is shown in black and gray.

(a) Convex hull (b) Skeletonization (c) Perimeter

Figure 4. Examples of the blob features

The skeletonization (SK) and the perimeter (PE) feature (see Fig. 4) are both indicating the flatness of a blob R:

$$SK(R) = \frac{\#\text{ Pixels of Skeletonization}(R)}{\sqrt{\#\text{ Pixels of } R}}, \quad PE(R) = \frac{\#\text{ Pixels of Perimeter}(R)}{\sqrt{\#\text{ Pixels of } R}} \tag{3}$$

In Fig. 4, the skeletonizations and the perimeters of the blobs are shown in black and the blobs are shown in gray. The scale invariance of the three shape features is gained by normalizing the shape features (the denominators in equations 2 and 3).
The three shape features are computed separately for dark and bright blobs resulting in 6 shape histograms. The final feature vector of an image consists of the aggregation of the two contrast histograms (25 bins per histogram) and the 6 shape histograms (15 bins per histogram). Each of the four features is able to achieve good results classifying polyps, but the best results are achieved by aggregating the histograms of the four features.

Distances between two feature vectors are measured using the χ^2 statistic which has been frequently used to compare probability distributions (histograms) and is defined by

$$\chi^2(x, y) = \sum_i \frac{(x_i - y_i)^2}{x_i + y_i} \tag{4}$$

Additionally, we employ a number of well known feature extraction methods to compare their results with our SSF method and also to have a higher number of methods resulting in more reliable conclusions with respect to the suitability of the CVC and CC for the automated mucosal texture classification:

DT-CWT [3] is a multi-scale and multi-orientation wavelet transform. The final feature vector of an image consists of the statistical features mean and standard deviation of the absolute values of the subband coefficients (6 decomposition levels $\times$ 6 orientations $\times$ 3 color channels $\times$ 2 features per subband = 216 features per image).

Gabor-Transformation [3] is a multi-scale and multi-orientation wavelet transform. The final feature vector of an image consists of the same statistical features like

in case of the DT-CWT.

LBP [8] is a texture operator which labels the pixels of an image by thresholding the neighborhood (8 neighbors per pixel, radius=1) of each pixel and considers the result as a binary number.

Fractal analysis [9] is a scale invariant method which pre-filters an image using the MR8 filterbank and then computes the local fractal dimensions of the (8) filter outputs followed by building models of the image using the Bag of Visual Words approach.

Multiscale Blob Features (MBF) [10] is a scale and rotation invariant method that produces binary images by thresholding the image by means of blurred versions of the image itself and uses a shape descriptor and the number of connected regions (blobs) as features.

3 Experimental setup and results

Our 8 image databases are acquired by extracting patches of size 256 x 256 from frames of HD-endoscopic (Pentax HiLINE HD+ 90i Colonoscope) videos using CVC (original, i-Scan modes 1–3) with or without CC. The patches are extracted only from regions having histological findings. Table 1 lists the number of images and patients per class and database.

Table 1. Number of images and patients per class with and without CC (staining) and CVC

	No staining				Staining			
i-Scan mode	No CVC	i-Scan 1	i-Scan 2	i-Scan 3	No CVC	i-Scan 1	i-Scan 2	i-Scan 3
$Healthy$								
Number of images	20	16	15	20	20	28	20	17
Number of patients	13	12	11	13	14	17	16	12
$Abnormal$								
Number of images	35	34	31	35	32	36	33	33
Number of patients	29	29	26	29	28	28	27	28
Total nr. of images	55	50	46	55	52	64	53	50

The different numbers of images (and patients) per database are caused by the different length of video sections with or without CC and with using different (or no) i-Scan modes. It is not possible to extract suitable patches for every video section (the videos are quite often blurry, the distance of the endoscope to the mucosa is too high (no details) or too small (blurry), there is too much endoscope movement (blurry) or no regions of interest are in the endoscope's field of view).

For a better comparability of the results, all methods are evaluated using a k-NN classifier. We use this simple classifier since it is adequate for all these methods and

because the focus of this paper lies on feature extraction strategies (especially SSF) and not on classification methods.

The results presented in Table 2 are the mean values of the 20 results of the k-NN classifier using Leave-one-patient-out (LOPO) cross validation with the k-values k=1–20. In that way we avoid the problem of varying results depending on the number of nearest neighbors of the k-NN classifier. The advantage of LOPO compared to leave-one-out cross validation (LOOCV) is the impossibility that the nearest neighbor of an image and the image itself come from the same patient. In this way we avoid overfitting. The row ∅ shows the averaged accuracies over all methods.

Table 2. Accuarcies with and without CC (staining) and CVC

	No staining				Staining			
i-Scan mode	No CVC	i-Scan 1	i-Scan 2	i-Scan 3	No CVC	i-Scan 1	i-Scan 2	i-Scan 3
SSF	69.6	75.6	77.6	73.2	68.1	69.5	**84.2**	78.4
DT-CWT	**80.1**	78.6	71.9	76.4	67.2	72.7	77.3	69.4
Gabor	67.8	71.1	76.5	61.7	65.2	72.7	73.8	67.3
LBP	67.0	72.9	71.3	65.6	64.2	70.9	74.0	66.0
Fractal Analysis	67.5	**80.2**	68.6	68.6	68.9	67.0	66.5	69.8
MBF	67.0	72.9	71.3	65.6	64.2	70.9	74.0	66.0
∅	69.8	**75.2**	72.9	68.5	66.3	70.6	75.0	69.5

As we can see in Table 2 the results of our SSF approach are higher with CVC than without and all in all better with CC than without. The best result is achieved using CC combined with i-Scan mode 2, which is also the best result over all methods. Our SSF approach performs quite competitive compared to the other methods. Only the DT-CWT achieves results comparable to these of the SSF.

Each method achieves its best or worst results at different image enhancement technologies (only SSF and MBF, both using shape features, show similarities). Overall, the best results are achieved using i-Scan modes 1 and 2. CC does not improve the classification rates of most of the methods.

4 Conclusion

With our SSF approach we have shown that shape features of blobs extracted by segmenting an image are particularly suitable for mucosal texture classification.

It also turned out that CVC (especially i-Scan mode 1 and 2) can help to improve classification results, whereas staining doesn't improve the classification results for the automated mucosal texture classification (except for our proposed method SSF using i-Scan mode 2). However, the classification results of the different feature extraction methods are not homogeneous with respect to the 8 image enhancement modes (CVC and CC) and there could be other methods providing contrary results. There is one thing common to all methods, the results using CC alone are below average (considered over all 8 image enhancement modes) for each of the 6 feature extraction methods. These

results are even worse than those without using any image enhancement technology (except for the method "Fractal Analysis"). From this point of view we have to state that CC is not making sense for the automated mucosal texture classification using HD-endoscopy, contrary to high-magnification endoscopes. CVC on the other hand does make sense.

5 Acknowledgments

This work is partially supported by the Austrian Science Fund, TRP Project 206.

References

1. Kiesslich, R.: Advanced imaging in endoscopy. European Gastroenterology & Hepatology Review **5**(1) (October 2009) 22–25
2. Kodashima, S., Fujishiro, M.: Novel image-enhanced endoscopy with i-scan technology. World Journal of Gastroenterology **16**(9) (2010) 1043–1049
3. Häfner, M., Kwitt, R., Uhl, A., Gangl, A., Wrba, F., Vécsei, A.: Feature-extraction from multi-directional multi-resolution image transformations for the classification of zoom-endoscopy images. Pattern Analysis and Applications **12**(4) (December 2009) 407–413
4. Tamaki, T., Yoshimuta, J., Kawakami, M., Raytchev, B., Kaneda, K., Yoshida, S., Yoshito-Takemura, Onji, K., Miyaki, R., Tanaka, S.: Computer-aided colorectal tumor classification in NBI endoscopy using local features. Medical Image Analysis **17**(1) (2013) 78 – 100
5. Kudo, S.E., Hirota, S., Nakajima, T., Hosobe, S., Kusaka, H., Kobayashi, T., Himori, M., Yagyuu, A.: Colorectal tumours and pit pattern. Journal of Clinical Pathology **47** (1994) 880–885
6. Xia, G., Delon, J., Gousseau, Y.: Shape-based Invariant Texture Indexing. International Journal of Computer Vision **88**(3) (2010) 382–403
7. Monasse, P., Guichard, F.: Fast computation of a contrast invariant image representation. IEEE Transactions on Image Processing **9**(5) (2000) 860–872
8. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition **29**(1) (January 1996) 51–59
9. Varma, M., Garg, R.: Locally invariant fractal features for statistical texture classification. In: Proceedings of the IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil. (October 2007)
10. Xu, Q., Chen, Y.Q.: Multiscale blob features for gray scale, rotation and spatial scale invariant texture classification. In: Proceedings of 18th International Conference on Pattern Recognition (ICPR). Volume 4. (September 2006) 29–32