SHAPE AND SIZE ADAPTED LOCAL FRACTAL DIMENSION FOR THE CLASSIFICATION OF POLYPS IN HD COLONOSCOPY

A. Uhl, G. Wimmer *

University of Salzburg
Department of Computer Sciences
Jakob Haringerstrasse 2, 5020 Salzburg, Austria

Michael Häfner

St. Elisabeth Hospital
Vienna, Austria

ABSTRACT

This work proposes a new method for computing the local fractal dimension for the classification of colonic polyps. First an image is segmented by an algorithm based on the idea of the watershed transform. The resultant connected components (blobs) show the local mucosal structure at local minima and maxima in the image and model the pit pattern structure of the mucosa. The local fractal dimension is computed using two different filter masks, an anisotropic Gaussian filter mask and an elliptic binary filter mask, which are especially adapted to the shapes and sizes of the blobs. By specifically fitting shapes and sizes of the filter masks for each blob, our feature is scale, orientation and viewpoint invariant. The proposed method outperforms other methods commonly used for mucosal texture classification.

Index Terms— Local fractal dimension, scale invariance, colonoscopy, colonic polyps

1. INTRODUCTION

Colonic polyps have a rather high prevalence and are known to either develop into cancer or to be precursors of colon cancer. 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.

In this work we use highly detailed images acquired by a high definition (HD) endoscope without magnification in combination with the i-Scan technology and conventional chromoscopy (staining the mucosa). In particular we use the i-Scan 2 mode, which includes surface enhancement, contrast enhancement and tone enhancement. This mode visually enhances boundaries, margins, surface architecture.

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. Consequently, in order to design reliable computer-aided mucosa texture classification schemes, the scale and viewpoint invariance of the employed feature sets could be essential.

Promising texture descriptors that are invariant to scale and viewpoint conditions and have demonstrated good results on real world datasets are [1, 2, 3, 4, 5]. Over the last years, fractal and multifractal geometries were applied extensively in many signal analysis applications like texture analysis [2, 3, 6, 7] and segmentation [8], including medical applications [9, 10]. Two of these methods are based on the computation of the local fractal dimension [2, 3], a viewpoint and scale invariant feature analyzing the 'non-uniformity" of the local intensity distribution by filtering the image or transformations of the image (MR8 filter responses [3] or directional derivatives (e.g. gradient image) [2]) with disk shaped binary filter masks.

In this work we propose a novel texture descriptor denoted as "Blob-Adapted Local Fractal Dimension" (BFD), which is based on computing a local fractal dimension like feature. The differences between BFD and previous released descriptors based on computing the local fractal dimension are:

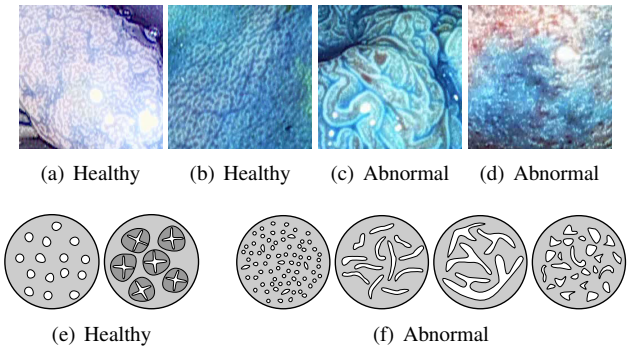
- We use ellipsoidal binary filters and Gaussian filters, whose shapes and sizes are adapted to the local texture regions. The other descriptors use disk shaped binary filter masks with predetermined radii.
- We compute the local fractal dimension only for interest points determined by the segmentation of the image instead of computing it for each pixel of an image respectively transformation of an image.

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 Fig. 1

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

a–d). The various pit pattern types [11] of these two classes are presented in Fig. 1 e–f.



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

(e) Healthy (f) Abnormal

Fig. 1. Example images of the two classes (a–d) and the pit pattern types of these two classes (e–f)

This paper is organized as follows. Section II describes the feature extraction methods, especially our new method based on the local fractal dimension. In section III we describe the experiments and present the results. Section IV presents the conclusion.

2. FEATURE EXTRACTION

The proposed feature extraction method BFD consists of a segmentation step followed by computing the local fractal dimension in a locally adapted manner.

2.1. Segmentation

In our segmentation approach, we want to extract the shapes of the pits and peaks inside of mucosal (gray scale) images. With pits we denote local areas of the mucosal image with lower gray values as the surrounding area and with peaks we mean local areas with higher gray values as the surrounding area. Our segmentation algorithm is a slight modification of the algorithm used in [12]:

Generating pit (peak) blobs R by localized region growing:

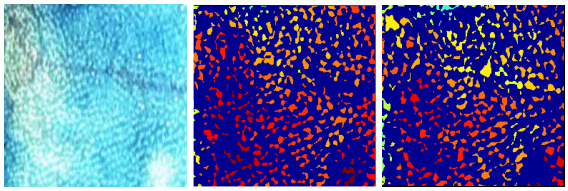
1. Scan the image I for a untagged local minimum (maximum for peak blobs) x_0 with gray value g and create a blob R consisting of only x_0 at the first iteration.
2. Find all neighbors N (4-connectivity) of R with $g_N = \min_{x \in N} I(x)$
   ($g_N = \max_{x \in N} I(x)$ for peak blobs) and tag them.
3. Two cases are possible:
   - $g \leq g_N$ ($g \geq g_N$ for peak 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 peak blobs):
     - # (Pixels of R)$\geq$8: use R for the BFD.
     - # (Pixels of R)$<$8: do not use R any further.

We only use these blobs for further feature extraction (computing the local fractal dimension) with $N \geq 8$ pixels. In this way we ensure that only these blobs are used that represent a distinct pit or peak in the mucosal texture structure and exclude these blobs that evolve of minima or maxima that are caused by noise. That means that our segmentation algorithm also acts as interest point detector.

The only difference between our segmentation algorithm and the one proposed in [12] is, that the algorithm in [12] sets in step 3, point two ($g > g_N$) the gray values of the image I to g_N for all pixels being part of the blob R. By resetting the gray values of the pixels of a blob, an already tagged blob can be fused with another blob resulting in bigger blobs. In our algorithm it is impossible that blobs are merged, each blob represents the surrounding area of a local minima (pit blob) or maximum (peak blob). This is important for the subsequent local fractal dimension computation in Section 2.2.

The idea behind this segmentation approach is that the two classes Healthy and Abnormal have different typical pit pattern types (see Figure 1). By filling up the pits and peaks of a mucosal image, the resultant blobs represents the shapes of local structures of the image including the different types of pit pattern. In that way the shape of the blobs contain information that enables an distinction between healthy and abnormal mucosa (see [12]).

The idea behind our segmentation algorithm is similar to the watershed segmentation by immersion [13, 14], but the results are completely different (see Figure 2.



(a) Image (b) Peak blobs (c) Pit blobs

Fig. 2. The extracted peak and pit blobs of the image

2.2. Feature Extraction

Let μ be a finite Borel regular measure on $\mathbb{R}^2$. For $x \in \mathbb{R}^2$, denote $B(x,r)$ as the closed disk with center x and radius $r > 0$. $\mu(B(x,r))$ is considered as an exponential function of r, i.e. $\mu(B(x,r)) = c\, r^{D(x)}$, where $D(x)$ is the density function and c is some constant. The local fractal dimension [2] of x is defined as

$$D(x) = \lim_{r \to 0} \frac{\log \mu(B(x,r))}{\log r}. \quad (1)$$

The local fractal dimension measures the "non-uniformity" of the intensity distribution in the region neighboring the considered point.

The local fractal dimension D is invariant under the bi-Lipschitz map, which includes view-point changes and non-rigid deformations of texture surface as well as local affine illumination changes [2].

In this work we propose a feature derived from the local fractal dimension, which uses ellipsoidal binary filters and anisotropic, ellipsoidal Gaussian filters fitted to the shape and size of blobs instead of disks with preassigned radii. To simplify matters, we also denote our feature as local fractal dimension.

In the following we describe the derivation of the elliptic binary filter masks and the ellipsoidal Gaussian filter masks.

Let us consider an image I, and a blob b extracted from I (it doesn't matter if it is a pit or peak blob) consisting of N Pixels with coordinates $(x_i, y_i),\ i \in \{1, ..., N\}$ in I and center of mass $(\overline{x}, \overline{y})$ $(\overline{x} = 1/N \times \sum_i x_i)$.

For the two integer values p, q, the normalized two-dimensional $(p+q)$th order central moments $\mu_{p,q}$ of the blob b are defined as follows:

$$\mu_{p,q} = \frac{\sum_{i=1}^{N}(x_i-\overline{x})^p(y_i-\overline{y})^q}{N^{\frac{p+q+2}{2}}}.$$

$$\left(\mu_{2,0} = \frac{\sum_{i=1}^{N}(x_i-\overline{x})^2}{N^2}, \qquad \mu_{1,1} = \frac{\sum_{i=1}^{N}(x_i-\overline{x})(y_i-\overline{y})}{N^2}\right)$$

From the inertia matrix $C = \begin{pmatrix}\mu_{2,0} & \mu_{1,1}\\ \mu_{1,1} & \mu_{0,2}\end{pmatrix}$, we compute the eigenvalues λ_1 and λ_2 (with $\lambda_1 > \lambda_2$) and the corresponding eigenvectors v_1 and v_2. Roughly speaking, the eigenvectors contain the informations of the orientation of the blob b in the image and the eigenvalues contain the information about the dilation of the blob in the eigenvectors direction.

The angle of v_1 in the coordinate system is computed by the arctangent function: $\alpha = -atan_2(y_{v_1}, x_{v_1})$, where y_{v_1} is the y-coordinate and x_{v_1} is the x-coordinate of the eigenvector v_1.The proportion of the eigenvalues is computed as follows: $p_\lambda = \sqrt{\lambda_1/\lambda_2}$.

Now, the blob can be approximated by an ellipse with mayor axis v_1 and minor axis v_2, whereat the proportion of the mayor axis length to the minor axis length is p_λ.

The two eigenvectors span a coordinate system, which is, compared to the Cartesian coordinate system, twisted by the angle α . The coordinate (x, y) of the Cartesian coordinate system can be transformed to the coordinate (x', y') of the eigenvectors coordinate system:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

A binary elliptic $f \times f$ filter approximating the shape of the blob b consisting of N pixels can be computed as follows:

$$E_b^t(x,y) = E_{p_\lambda,N,}^t(x,y) = \begin{cases} 1 \text{ if } \sqrt{x'^2 + (y'p_\alpha)^2} < \sqrt{\frac{N}{\pi}}t, \\ 0 \text{ otherwise.} \end{cases}$$

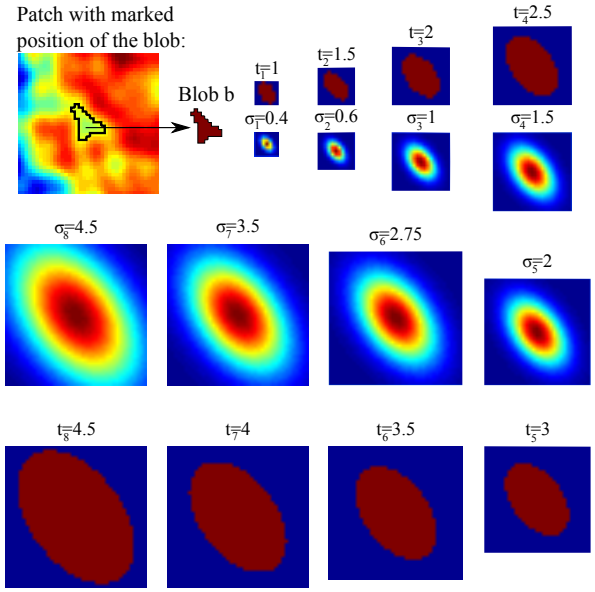


Fig. 3. A patch containing a pit blob b in his center and the binary elliptic filter masks $E_b^{t_i}$ and elliptic Gaussian filter masks $G_b^{\sigma_i}$ adapted to the Blob with parameters t_i respectively λ_i.

The term $\sqrt{\frac{N}{\pi}}$ and the threshold t control the size of the ellipse. The term $\sqrt{\frac{N}{\pi}}$ is derived from the computation of the area A of a disk with radius r: $A = r^2 * \pi \Leftrightarrow r = \sqrt{\frac{A}{\pi}}$.

The x and y coordinates of the filter are integers and range from $-(f-1)/2$ to $+(f-1)/2$ (f has to be odd-numbered).

In Figure 3 we see a blob b and the corresponding binary elliptic filters $E_b^{t_i}$ with the thresholds $t_i,\ i \in \{1, \ldots, 8\}$.

The elliptic Gaussian filter mask $G_{p_\lambda,\alpha,N,}^{\sigma}$ with size $f \times f$ is computed in a similar way as the binary elliptic filters:

$$GF_b^\sigma(x,y) = GF_{p_\lambda,N}^\sigma(x,y) = e^{-\left(\frac{x'^2+(y'p_\lambda)^2}{2\left(\sigma\sqrt{\frac{N}{\pi}}\right)^2}\right)},$$

Finally, the Gaussian filter has to be normalized in two steps:

$$G1_b^\sigma(x,y) = \frac{GF_b^\sigma(x,y)}{\sum GF_b^\sigma},\ G_b^\sigma(x,y) = G1_b^\sigma(x,y) - \frac{\sum G1_b^\sigma}{f^2}.$$

In Figure 3 we see a blob b and the corresponding elliptic Gaussian filters $G_b^{\sigma_i}$ with the standard deviations $\sigma i,\ i \in \{1, \ldots, 8\}$.

So for each blob in the image we generate especially fitted binary elliptic filters E and elliptic Gaussian filters G and with these filters we now compute the fractal dimension:

For a given Blob b with center position $(\overline{x}, \overline{y})$ in the image I and the corresponding filters G_b^σ (E_b^t analogous) with filter size $f \times f$, μ is defined as follows:

$$\mu(G_b^{\sigma_i}) = \sum_{x=-\frac{f-1}{2}}^{\frac{f-1}{2}} \sum_{y=-\frac{f-1}{2}}^{\frac{f-1}{2}} I(\overline{x}-x, \overline{y}-y) \quad G_b^{\sigma_i}(x,y)$$

Since $\mu(G_b^{\sigma_i})$ can be negative, $\log \mu(G_b^{\sigma_i})$ would be complex. We solve this problem by adding a value to the 8 $\mu(G_b^{\sigma_i})$'s ($i \in \{1, \ldots, 8\}$), so that the smallest $\mu(G_b^{\sigma_i})$ over all i's is one: $\mu(G_b^{\sigma_i}) = \mu(G_b^{\sigma_i}) + 1 - \min_{i \in \{1,\ldots 8\}} \mu(G_b^{\sigma_i})$.

We define the two local fractal dimensions for a Blob b as:

$$D_E(b) = \lim_{t \to 0} \frac{\log \mu(E_b^{t_i})}{\log i}, D_G(b) = \lim_{\sigma \to 0} \frac{\log \mu(G_b^{\sigma_i})}{\log i}, \quad (2)$$

where σ_i and t_i are strictly monotonic increasing.

We practically compute the local fractal dimension D_E (D_G) for each blob by linear fitting the slope of the line in a scaling plot of $\log \mu(E_b^{t_i})$ ($\log \mu(G_b^{\sigma_i})$) against $\log i$ with $i \in \{1, \ldots, 8\}$. σ_i and t_i have the values as shown in Figure 3. That means we locally filter the image I at each position where a blob b exists with filter masks that are fitted to the shape and size of the blobs and analyze the degree of increase or decrease of the filter responses for increasing filter sizes.

The local fractal dimension measures the "non-uniformity" of the intensity distribution in the region and neighboring region of a blob. Beginning with the center region of a pit or peak, it analyzes the changes in the intensity distribution with expanding region. In that way it analyzes the changing intensity distribution from the inside to the outside of a pit or peak in an image. Since size and shape of the filters are adapted to the blob representing the pit or peak, our version of the local fractal dimension is even more invariant to varying scales, orientations and viewpoint conditions as the original fractal dimension using disks with fixed radii [2].

The final feature vector of an image consists of the histograms (15 bins) of the local fractal dimensions D_E as well as D_G, separately computed for pit and peak blobs (60 features per feature vector).

Distances between two feature vectors are measured using the χ^2 statistic.

2.3. Other Methods

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

Segmented shape features (SSF) [12] uses a similar segmentation algorithm as in our work, with only one difference that is pointed out in Section 2.1. The feature vector of an image consists of the histograms of three features describing the shape of the blobs and of one histogram of a feature describing the contrast inside of the blobs.

DT-CWT [15] 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).

Methods	BFD	SSF	DTCWT	LBP	FA
Accuracies	**84.6**	79.6	73.1	72.8	82.3

Table 1. Accuracies of the methods in %

LBP [16] 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 (FA) [3] is a scale invariant method that 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.

3. EXPERIMENTAL SETUP AND RESULTS

Our image databases is acquired by extracting patches of size 256 x 256 from frames of HD-endoscopic (Pentax HiLINE HD+ 90i Colonoscope) videos. The patches are extracted only from regions having histological findings. The database consists of 94 patches, 32 patches of class Healthy (from 23 patients) and 62 patches of class Abnormal (from 52 patients). Before segmentation and feature extraction, the images are Gaussian blurred with $\sigma = 2$.

For a better comparability of the results, all methods are evaluated using a k-NN classifier.

The results presented in Table 1 are the mean values of the 10 results of the k-NN classifier using Leave-one-patient-out (LOPO) cross validation with the k-values k=1–10. 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. As we see in Table 1 the results of our proposed method are better than the results of the other methods.

4. CONCLUSION

With our BFD approach we have shown that the local fractal dimension, adapted to the shape of local texture structures, is particularly suitable for mucosal texture classification.

It was already shown in [12], that the shape of the blobs, generated by local region growing starting with local extrema, is a good indicator to discriminate between healthy and abnormal mucosa.

In this work we showed that analyzing the changes in the intensity distribution of an expanding region centered at a pit or peak of an image, whose shape and and size is adapted to the blob representing the pit or peak, seems to be an even better indicator for polyp classification.

5. REFERENCES

[1] C. Schmid S. Lazebnik and J. Ponce, "A sparse texture representation using local affine region," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1265–1278, 2005.

[2] Y. Xu, H. Ji, and C. Fermüller, "Viewpoint invariant texture description using fractal analysis," *International Journal of Computer Vision*, vol. 83, no. 1, pp. 85–100, 2009.

[3] M. Varma and R. Garg, "Locally invariant fractal features for statistical texture classification," in *Proceedings of the IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil*, Oct. 2007, pp. 1–8.

[4] G.S. Xia, J. Delon, and Y. Gousseau, "Shape-based Invariant Texture Indexing," *International Journal of Computer Vision*, vol. 88, no. 3, pp. 382–403, 2010.

[5] M. Mellor, Byung-Woo Hong, and M. Brady, "Locally rotation, contrast, and scale invariant descriptors for texture analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 1, pp. 52–61, Jan 2008.

[6] Sebastien Deguy, Christophe Debain, Albert Benassi, Universite Blaise Pascal, Pole Physique, and Complexe Scientifique Des Cezeaux, "Classification of texture images using multi-scale statistical estimators of fractal parameters," in *British Machine Vision Conference*, 2000, pp. 192–201.

[7] S. Peleg, Joseph Naor, Ralph Hartley, and D. Avnir, "Multiple resolution texture analysis and classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 4, pp. 518–523, July 1984.

[8] L.M. Kaplan, "Extended fractal analysis for texture classification and segmentation," *IEEE Transactions on Image Processing*, vol. 8, no. 11, pp. 1572–1585, Nov 1999.

[9] R. Lopes and N. Betrouni, "Fractal and multifractal analysis: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 634–649, 2009.

[10] A. Uhl, A. Vécsei, and G. Wimmer, "Fractal analysis for the viewpoint invariant classification of celiac disease," in *Proceedings of the 7th International Symposium on Image and Signal Processing (ISPA 2011)*, Dubrovnik, Croatia, Sept. 2011, pp. 727 –732.

[11] S.-E. Kudo, S. Hirota, T. Nakajima, S. Hosobe, H. Kusaka, T. Kobayashi, M. Himori, and A. Yagyuu, "Colorectal tumours and pit pattern," *Journal of Clinical Pathology*, vol. 47, pp. 880–885, 1994.

[12] A. Häfner, A. Uhl, and G. Wimmer, "A novel shape feature descriptor for the classification of polyps in HD colonoscopy," in *Medical Computer Vision. Large Data in Medical Imaging (Proceedings of the 3rd Intern. MICCAI - MCV Workshop 2013)*, 2014, vol. 8331 of *Springer LNCS*, pp. 205–213.

[13] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, Jun 1991.

[14] Jos B.T.M. Roerdink and Arnold Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies," *Fundamenta Informaticae*, vol. 41, no. 1,2, pp. 187–228, 2000.

[15] M. Häfner, R. Kwitt, A. Uhl, A. Gangl, F. Wrba, and A. Vécsei, "Feature-extraction from multi-directional multi-resolution image transformations for the classification of zoom-endoscopy images," *Pattern Analysis and Applications*, vol. 12, no. 4, pp. 407–413, Dec. 2009.

[16] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, January 1996.