

# Color Treatment in Endoscopic Image Classification using Multi-scale Local Color Vector Patterns

M. Häfner<sup>a</sup>, M. Liedlgruber<sup>b,\*</sup>, A. Uhl<sup>b</sup>, A. Vécsei<sup>c</sup>, F. Wrba<sup>d</sup>

<sup>a</sup>*St. Elisabeth Hospital  
Department for Internal Medicine  
Vienna, Austria*

<sup>b</sup>*University of Salzburg  
Department of Computer Sciences  
5020 Salzburg, Austria*

<sup>c</sup>*St. Anna Children's Hospital  
Endoscopy Unit  
Vienna, Austria*

<sup>d</sup>*Medical University of Vienna  
Department of Clinical Pathology  
Vienna, Austria*

---

## Abstract

In this work we propose a novel method to describe local texture properties within color images with the aim of automated classification of endoscopic images. In contrast to comparable Local Binary Patterns operator approaches, where the respective texture operator is almost always applied to each color channel separately, we construct a color vector field from an image. Based on this field the proposed operator computes the similarity between neighboring pixels. The resulting image descriptor is a compact 1D-histogram which we use for a classification using the k-nearest neighbors classifier.

To show the usability of this operator we use it to classify magnification-endoscopic images according to the pit pattern classification scheme. Apart from that, we also show that compared to previously proposed operators we are not only able to get competitive classification results in our application scenario, but that the proposed operator is also able to outperform the other methods either in terms of speed, feature compactness, or both.

*Keywords:* Colonoscopy, colon cancer, local binary patterns, multi-scale, color, classification

---

\*Corresponding author

*Email addresses:* mliedl@cosy.sbg.ac.at (M. Liedlgruber), uhl@cosy.sbg.ac.at (A. Uhl)

## 1. Introduction

Today, the third most common malignant disease in western countries is colon cancer. Therefore, a regular colon examination is recommended, especially for people at an age of 50 years and older. Currently the gold standard for colon examination is colonoscopy, which is performed by using a colonoscope. Modern colonoscopes are able to take pictures from inside the colon which allows to obtain images for a computer-assisted analysis with the goal of detecting tumorous lesions. To get highly detailed images a magnifying endoscope can be used (Bruno, 2003). Such an endoscope represents a significant advance in colonoscopy as it provides images which are up to 150-fold magnified, thus uncovering the fine surface structure of the mucosa as well as small lesions.

There already exists some previous work devoted to an automated cancer staging employing different colonoscopic imaging modalities or videos. For classic white-light endoscopy several studies have demonstrated that computer-based image analysis is capable of detecting colorectal polyps (Karkanis, 2003; Maroulis et al., 2003) in endoscopic video frames to a certain extent and to perform a first assessment of the malignant potential of these polyps (Krishnan et al., 1998; Karkanis et al., 2001; Iakovidis et al., 2005, 2006). Narrow-band-imaging (NBI) has been shown to facilitate discrimination between neoplastic and non-neoplastic polyps relying on features of the observed microvasculature to some extent (Gross et al., 2009; Stehle et al., 2009; Tischendorf et al., 2010). Confocal laser endomicroscopic images have also been used to differentiate lesions into the categories neoplastic and benign (André et al., 2009a,b) using a dense variant of the bag-of-visual-words features.

While diagnostic accuracies of automated staging techniques employing the imaging modalities described so far range between 70% and 94%, classification accuracies ranging from 95% up to 99% depending on employed features and classification schemes (see e.g. Kwitt and Uhl (2007); Häfner et al. (2009a,b, 2010)) have been achieved using high-magnification chromo-colonoscopy (where contrast enhancement is achieved by actual staining during colonoscopy).

We specifically target at the classification of the latter type of imagery based on the pit pattern scheme. In our previous work in this field (Kwitt and Uhl, 2007), for example, we used the the Dual-Tree Complex Wavelet Transform (DT-CWT). While the DT-CWT allows to analyze content within an image in an approximately shift-invariant manner and is directionally selective, additional computational burden is introduced due to the necessity of applying the wavelet transform to an image multiple times.

Another work based on Fourier-features has been presented in (Häfner et al., 2010). In this work features are extracted from Fourier transformed images. To be able to focus on dominant information within the images the features are extracted from a certain number of frequency bands with configurable bandwidths only. Since however the set of optimal frequency bands is not known at beforehand (neither the ideal number of frequency bands nor the respective bandwidths) different feature selection techniques are compared to an evolutionary algorithm. While we achieve rather high classification accuracies using

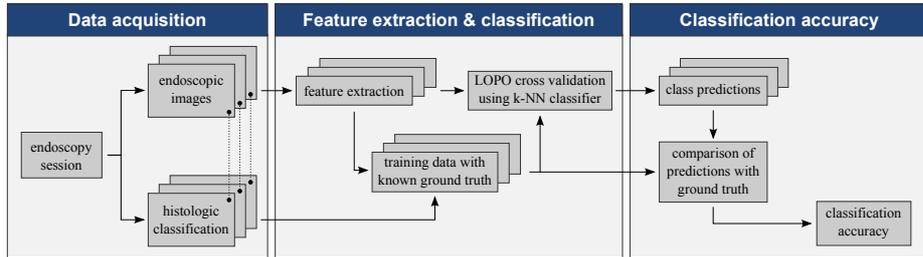


Figure 1: Overview of our system for endoscopic image classification.

the evolutionary algorithm, the respective computational demand is very high.

An ensemble classification system, combining a set of different previously developed methods for classification, has been applied in (Häfner et al., 2009c). Compared to the single methods, the ensemble classifier delivered a superior performance regarding the classification accuracy. However, to be able to build a fast ensemble classifier system it is also important that the single methods have a low computational demand, in terms of feature extraction as well as for the classification of an image. In addition to these aspects with respect to computational performance issues, a high error diversity of the techniques combined in ensemble classifiers is required to significantly enhance results as compared to the single classifiers' results. Besides the overall aim to develop techniques which are computationally efficient to be computed, we also focus on developing methods suited for inclusion in ensemble classifiers which rely on significantly different feature sets compared to earlier work (Kwitt and Uhl, 2007; Häfner et al., 2010).

In this work we propose a novel color texture operator, called the Local Color Vector Patterns operator (LCVP). This operator introduces a novel way to deal with color in Local Binary Patterns (LBP) based texture descriptors. In addition, LCVP offers a high compactness in terms of the features generated, leading to a fast classification. Another advantage of the compact texture descriptors is the small amount of time required to generate and classify them. We use the LCVP operator for an automated classification of visual data acquired by a magnifying colonoscope corresponding to different types of pit patterns. An overview of our system for endoscopic image classification is shown in Fig. 1. After the acquisition of endoscopic images and collecting the respective histologic classification, features are extracted using LCVP and classified using the k-NN classifier (based on the leave-one-patient-out cross-validation protocol). In the final step, the class predictions of the classifier are compared against the ground truth, providing the classification accuracy of the system.

In Section 2 we review the classification of pit patterns of the colonic mucosa. Section 3 summarizes related work on using color-based LBP methods. The Multi-scale Block LBP operator, which is the basis for this work, is reviewed in Section 4. The proposed color texture operator is introduced in Section 5, followed by the classification of the resulting features in Section 6. Experimental

results and configuration details of the classification system proposed are given in Section 7, where we also compare the proposed scheme to other ways of dealing with color in LBP-based schemes. Section 8 concludes the paper.

## 2. Pit Pattern Classification

Polyps of the colon are a frequent finding and are usually divided into metaplastic, adenomatous, and malignant. As resection of all polyps is time-consuming, it is imperative that those polyps which warrant endoscopic resection can be distinguished: polypectomy of metaplastic lesions is unnecessary and removal of invasive cancer may be hazardous. For these reasons, assessing the malignant potential of lesions at the time of colonoscopy is important as this would allow to perform targeted biopsy. While such systems are still not standard-of-care, the aim of developing such automated polyp classification systems is to avoid random and, probably, unnecessary biopsies. Hence, such systems could potentially help to save time, lower the cost for a colonoscopy procedure, and reduce the risk of complications during the procedure.

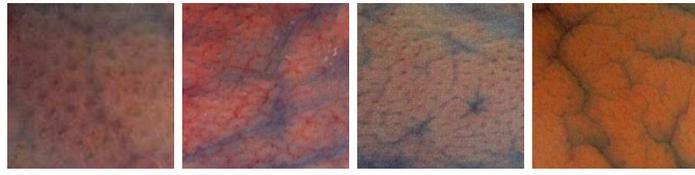
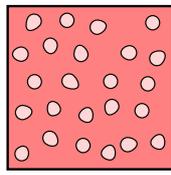
The most commonly used classification system to distinguish between non-neoplastic and neoplastic lesions in the colon is the pit pattern classification, originally reported by Kudo et al. (1994). This system allows to differentiate between normal mucosa, hyperplastic lesions (non-neoplastic), adenomas (a pre-malignant condition), and malignant cancer based on the visual pattern of the mucosal surface. Thus this classification scheme is a convenient tool to decide which lesions need not, which should, and which most likely can not be removed endoscopically. The mucosal pattern as seen after dye staining and by using magnification endoscopy shows a high agreement with the histopathologic diagnosis. Due to the visual nature of this classification it is also a convenient choice for an automated image classification.

In this classification scheme exist five main types according to the mucosal surface of the colon, as illustrated in Figure 2. Type III is divided into types III-S and III-L, designating the size of the pit structure. It has been suggested that type I and II pattern are characteristic of non-neoplastic lesions (benign and non-tumorous), type III and IV are found on adenomatous polyps, and type V are strongly suggestive of invasive carcinoma, thus highly indicative for cancer (Kudo et al., 1996).

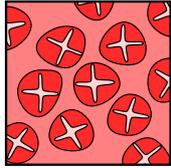
Furthermore lesions of type I and II can be grouped into non-neoplastic lesions and types III to V can be grouped into neoplastic lesions. This allows a grouping of lesions into two classes, which is more relevant in clinical practice as indicated in a study by Kato et al. (2006).

Using a magnifying colonoscope together with indigo carmine dye spraying, the mucosal crypt pattern on the surface of colonic lesions can be observed (Kudo et al., 1996). Several studies found a good correlation between the mucosal pit pattern and the histological findings, where especially techniques using magnifying colonoscopes led to excellent results (Kato et al., 2006).

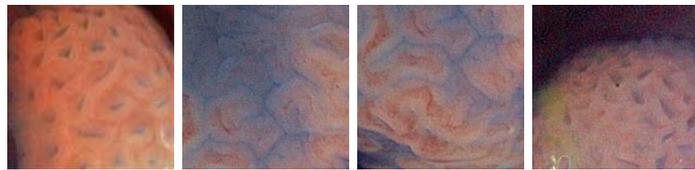
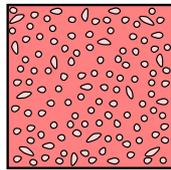
From Figure 2 we notice that pit pattern types I to IV can be characterized fairly well, whereas type V is a composition of unstructured pits. At a first



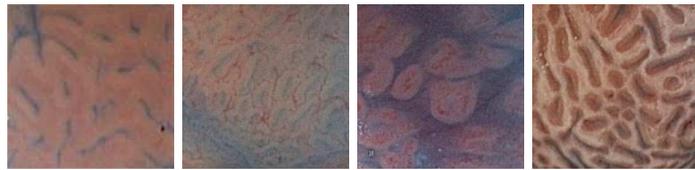
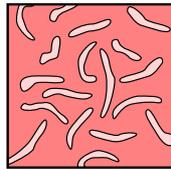
(a) Pit pattern type I



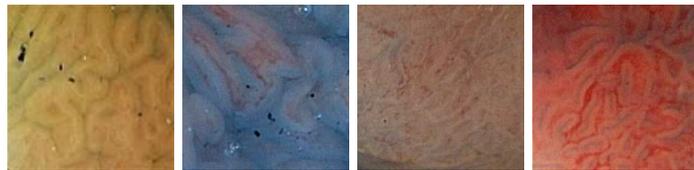
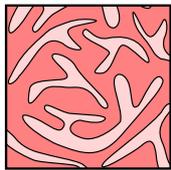
(b) Pit pattern type II



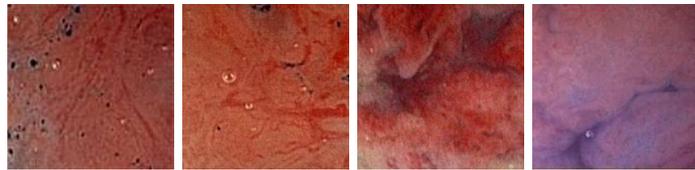
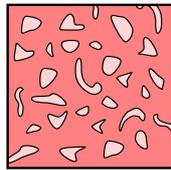
(c) Pit pattern type III S



(d) Pit pattern type III L



(e) Pit pattern type IV



(f) Pit pattern type V

Figure 2: Schematic illustration of the pit pattern classification along with example images for each pit pattern type.

glance this classification scheme seems to be straightforward and easy to be applied. But it needs some exercising to achieve fairly good results (Hurlstone, 2002). Apart from that, similar to the reported inter-observer variability of NBI-based colonoscopy ( $\kappa \approx 0.57$  (Rastogi et al., 2009),  $\kappa \approx 0.63$  (Chang et al., 2009),  $\kappa \approx 0.69$  (Raghavendra et al., 2010)) inter-observer variability of magnification chromo-endoscopy in the interpretation of pit patterns of colonic lesions has been described ( $\kappa \approx 0.56$  (Zanoni et al., 2007),  $\kappa \approx 0.64$  (Chang et al., 2009)). This work aims at allowing computer-assisted pit pattern classification in order to enhance the quality of differential diagnosis.

The topical staining used in chromo-endoscopy has the effect of visually enhancing mucosal crypt patterns or vascular features. Depending on the color dye applied, enhanced features appear in different colors such as red, violet, or blue (Song et al., 2007). Hence, it is important – for the medical expert as well as for an automated polyp classification system – to have color images at hand.

### 3. Related Work on Color in LBP Operators

The LBP operator, introduced by Ojala et al. (1996), has become a very popular technique to describe the local properties of texture. This operator and extensions to it have been used already in many different areas of research such as for example medical image processing, face recognition, and image retrieval – just to mention a few. However, the LBP operator is usually either applied to grayscale images, or, in case of color images, the most common way to deal with color is to combine color channels by concatenating LBP histograms from color channels, which have been processed separately using LBP (e.g. Mäenpää and Pietikäinen (2004)).

Another, more specific color-based LBP operator is the Opponent Color LBP operator (OCLBP) (Mäenpää et al., 2002) which exploits the inter-channel relationships between color channels. For this purpose the LBP transform is computed for all possible pairs of color channels, thresholding the neighbors from one color channels to a pixel from a second channel. Assuming that the histogram is not quantized, this results in nine different combinations, each yielding a histogram containing 256 bins. Concatenating these histograms results in a total number of 2304 bins. This LBP variant has been used for example in (Ameling et al., 2009) for texture-based polyp detection in endoscopic videos and in (Pietikäinen et al., 2002) for color texture classification.

The work presented in (Mäenpää and Pietikäinen, 2004) is an extensive study, evaluating the performance of color-based LBP (concatenation of LBP histograms originating from different color channels) compared to color histograms and grayscale features in the context of image retrieval. In addition to the traditional techniques, color histogram concatenation and the OCLBP operator are evaluated. Similar experiments and comparisons can be found in (Mäenpää et al., 2002) and (Pietikäinen et al., 2002).

In (Connah and Finlayson, 2006) joint LBP histograms have been proposed. In this work the LBP operator is again applied to each color channel separately. But instead of concatenating 1D histograms three-dimensional joint histograms

Reference	$N_C$	Histogram type	$N_B$
Mäenpää et al. (2002)	3	Inter-channel histograms	2304
Mäenpää and Pietikäinen (2004)	3	Concatenated histograms	768
Connah and Finlayson (2006)	3	Joint histograms	205379
Häfner et al. (2009b)	2	Joint histograms	65536

Table 1: A comparison of LBP-based methods using color information ( $N_C$  and  $N_B$  denote the number of color channels considered and the number of histogram bins used, respectively).

are created to capture relationships between different color channels. By using uniform patterns (Mäenpää et al., 2000), only binary patterns containing either zero or two bit transitions between 0 and 1 (and vice versa) are considered for the resulting histograms, resulting in a reduced number of 58 binary patterns. The remaining binary patterns are accumulated in an additional bin, hence a total of 59 bins is used. For three-dimensional histograms the final histograms used as features for classification contain  $59^3 = 205379$  histogram bins. These features have been used for example in (Mackiewicz et al., 2008) to segment endoscopic videos. Recent work aiming at the classification of magnification-endoscopic images is also based on joint LBP histograms (Häfner et al., 2009b). However, we used two color channels only to lower the dimensionality of the features. To compensate for the eventual loss in discriminative information, we considered all possible patterns instead of the restricted set of uniform patterns which results in  $256^2 = 65536$  histogram bins.

Table 1 provides a brief overview of the different methods found in literature, which incorporate color information into LBP-based systems.

#### 4. Basis for the Local Color Vector Patterns Operator

Since the original LBP operator (Ojala et al., 1996) is very sensitive to noise and our images suffer from noise, we base our work on the Multi-scale Block LBP operator (MB-LBP), which has been proposed recently (Liao et al., 2007). To achieve a higher robustness the comparison between pixel values in LBP is replaced by a comparison of pixel block intensity averages, which are computed for each pixel (hence, the blocks overlap). This operator is applied to each color channel  $C_a$  of an image  $I$  separately ( $a = 1, 2, 3$ ).

This process is shown in Fig. 3(a). The average intensity of the center block, which has a size of  $K \times K$  pixels and is centered at the pixel position denoted by the white box, is compared against the intensity averages of the neighboring blocks (the center positions of these blocks are denoted by the dark gray boxes). The blue boxes surrounding the dark gray boxes denote the subwindows of pixels used to compute the respective intensity averages (in this example the blocks used are of size  $3 \times 3$ , hence  $K = 3$ ).

This is equivalent to a convolution of  $C_a$  with a square averaging filter kernel of size  $K \times K$ . The result of the convolution is denoted by  $C_a^*$ . The intensity of a pixel is compared against the intensities of the neighboring pixels, as depicted in Fig. 3(b). In this case the neighboring pixels are not necessarily the

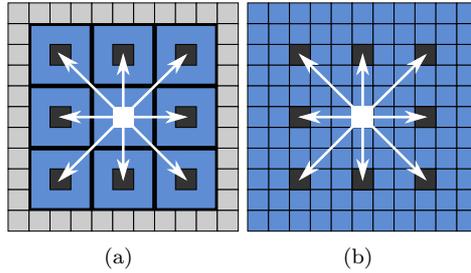


Figure 3: LBP code generation for one pixel position (a) using block averaging and (b) based on a pre-processed image (using an averaging filter).

pixels adjacent to the center pixel. The reason for this is that the position of the neighboring pixels must coincide with the block centers shown in Fig. 3(a) since these pixels now contain the intensity values for the respective neighboring blocks. As a consequence the distance of the neighboring pixels used for intensity comparisons depends on the size of the averaging filter kernel used. The averaging used in this work is based on a whole-point symmetric extension which allows us to compute LBP values near and on the border of an image as well. For  $K = 1$  the averaging has no effect and MB-LBP reduces to LBP.

The thresholding result within a pre-processed color channel  $C_a^*$  for a pixel located at  $(i, j)$  is computed as

$$T_{n,i,j} = \begin{cases} 1, & \text{if } C_a^*(x_n, y_n) \geq C_a^*(i, j) \\ 0, & \text{otherwise} \end{cases}, \quad (1)$$

where  $x_n$  and  $y_n$  denote the position of the  $n$ -th neighboring pixel (these are ordered in a clock-wise fashion, starting with the top left neighbor). Using  $T_{n,i,j}$ , the LBP number can be formulated as

$$LBP_{i,j} = \sum_{n=1}^N 2^{n-1} T_{n,i,j}, \quad (2)$$

with  $N = 8$  since we are using a 8-neighborhood. The computation of the LBP number for an example pixel neighborhood is illustrated in Fig. 4.

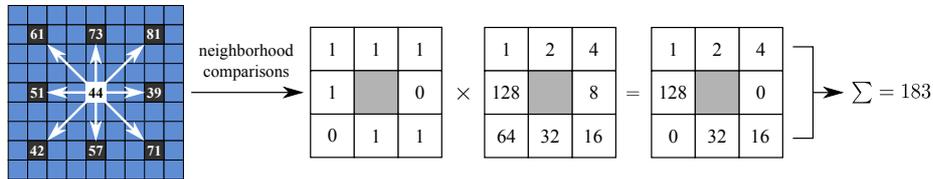


Figure 4: An example pixel neighborhood within some color channel along with the respective computation steps to obtain the final LBP number (for the center pixel denoted by the white box).

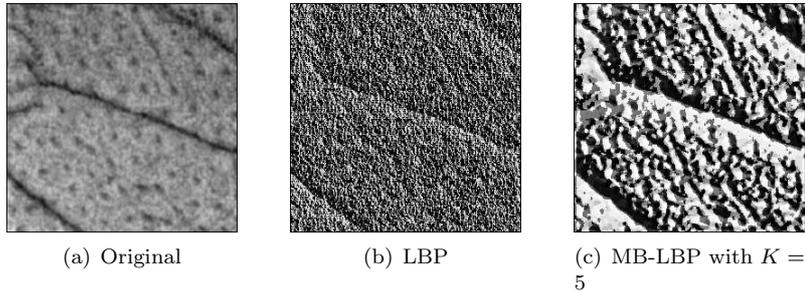


Figure 5: Differences between different variants of LBP applied to the red channel of an example endoscopic image.

As shown in Fig. 5 the MB-LBP operator significantly reduces the influence of noise. Choosing a small value for  $K$  preserves more detail but the operator gets more sensitive to noise, whereas higher values for  $K$  lower the sensitivity to noise but also result in loss of detail. Hence, adjusting the parameter  $K$  allows to find a trade-off between noise and detail loss. In addition, multiple resolutions of an image can be created by using different values for  $K$ . Based on this fact, the multi-scale LCVP operator in Section 5.4 is introduced.

## 5. The Local Color Vector Patterns Operator (LCVP)

In (Häfner et al., 2009b) we combined MB-LBP with joint color LBP histograms to classify endoscopic images and exploit inter-channel relationships between different color channels (we refer to this method as joint color multi-scale block LBP (JC-MB-LBP) from now on). While the results of this method are superior compared to many of our previous approaches it has two major drawbacks. First, as already pointed out in Section 3, we used a combination of two color channels resulting in histograms containing  $256^2$  bins. The result was a rather poor performance in terms of classification speed. Secondly, the choice of the color channels used is somewhat arbitrary. Therefore, we conducted experiments with all possible combinations of two channels to find the combination performing best in terms of the classification accuracy in (Häfner et al., 2009b) (equivalent to a manual feature selection). This results in an additional computational burden. But this is necessary since the different possible combinations might yield considerably different classification accuracies. While we could have used all color channels available this would have resulted in histograms containing  $256^3$  bins which would have ended up in an even more time consuming classification.

The computational burden of using OCLBP histograms (see Section 3) for classification is definitely lower compared to JC-MB-LBP. But things get worse as soon as multiple image scales come into play when introducing the multi-scale variant of OCLBP. This is quite obvious since the time for histogram

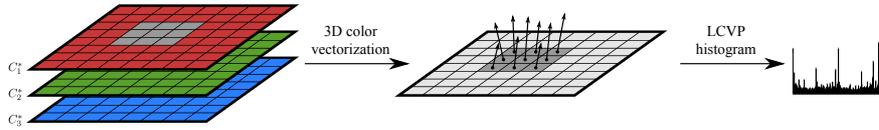


Figure 6: Schematic overview of the proposed operator. Each pixel of an (already averaged) input image is treated as a 3D vector, resulting in a color vector field. Then the similarity between pixels is computed using a similarity function. Based on the resulting LCVP numbers we finally create a 1D histogram (the dark gray part within the slices denotes a sample window along with the respective color vectors).

comparison during classification is then approximately multiplied by the number of resolution stages used.

### 5.1. Motivation for the LCVP Operator

Since the JC-MB-LBP operator yielded very promising results already in (Häfner et al., 2009b) we introduce the LCVP operator with the aim of coping with the main limitations of JC-MB-LBP: the computational demand in terms of the classification and the fact that not all color information available is used. Hence, our goal is to develop a texture descriptor which, compared to JC-MB-LBP, utilizes all color information available, while being more compact and computable in a smaller amount of time.

It is out of question that the primary goal in medical image classification systems must be a high diagnostic accuracy. However, considering the fact that an endoscopic procedure should allow a real-time diagnosis (to enable the examiner to set an appropriate reaction like taking a biopsy or similar) it is also important to create algorithms steering into that direction, and thus to develop fast and efficient feature extraction and classification methods.

### 5.2. Threshold-Based LCVP Operator

The main difference between LCVP and other local color texture operators similar to or derived from LBP is the way we process pixels. As described before, most work published so far using LBP or some variant of LBP in the context of color processing is based on computing the LBP transform considering pixel intensities within single color channels only. Thus, each color channel is processed separately. Although OCLBP at least incorporates the relationship between pairs of channels there exists no LBP variant using the full color information of a pixel at once.

LCVP by contrast treats each pixel as a 3D color vector, as illustrated in Figure 6. This means that, in contrast to LBP, we are not able to use a simple thresholding of intensity values to get the binary sequence making up the LCVP number for a pixel. Our solution to this problem is to use functions measuring the similarity between color vectors. Along with a threshold such a function again produces a binary output. Possible choices for such similarity measures

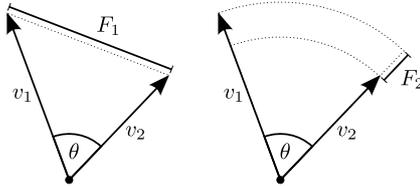


Figure 7: Illustration of the similarity measures used throughout this work (between two vectors  $v_1$  and  $v_2$ ).

are (Smolka and Venetsanopoulos, 2007)

$$F_1(v_1, v_2) = \|v_1 - v_2\| \quad (3)$$

$$F_2(v_1, v_2) = \left| \|v_1\| - \|v_2\| \right| \quad (4)$$

where  $\|\cdot\|$  denotes the length of a vector in Euclidean space. The measures  $F_1$  and  $F_2$  are illustrated in Figure 7.  $F_1$  corresponds to the magnitude of the difference vector between two color vectors  $v_1$  and  $v_2$ , approximating the perceptual difference between two colors.  $F_2$  is the absolute difference between the lengths of  $v_1$  and  $v_2$ .

Figure 8 shows different 2D-color planes which are either transformed to the HSV color space or to the LAB color space. As we notice from this figure, in case of RGB images the length of a color vector influences the brightness (value) only, while the color tone (hue) is only dependent of the color vector direction. As a consequence,  $F_2$  approximately corresponds to the brightness difference between two color vectors. However, when converting an image to the LAB color space the length of a color vector reflects color changes ( $a^*$  and  $b^*$ ) as well as brightness changes ( $L^*$ ).

Similar to MB-LBP we first apply an averaging filter to each color channel  $C_a$  of an image  $I$  to obtain the pre-processed image  $I^*$ . In analogy to Equ. (1), the thresholding for a color vector located at  $(i, j)$  within  $I^*$  can be expressed as

$$T_{n,i,j}^{(p)} = \begin{cases} 1, & \text{if } F_p(I^*(x_n, y_n), I^*(i, j)) < t \\ 0, & \text{otherwise} \end{cases}, \quad (5)$$

where  $x_n$  and  $y_n$  denote the position of the  $n$ -th neighboring color vector (ordered in a clock-wise fashion, starting at the top left color vector),  $F_p(\cdot, \cdot)$  denotes a similarity measure as defined in Equ. (3) and (4), and  $t$  is a threshold allowing to control the strictness of the operator with  $t > 0$ .

Using  $T_{n,i,j}^{(p)}$  the final LCVP number based on color vectors can be formulated as

$$LCVP_{i,j}^{(p)} = \sum_{n=1}^N 2^{n-1} T_{n,i,j}^{(p)}. \quad (6)$$

Similar to MB-LBP, the parameter  $K$  we used to pre-process our images (averaging filter), allows us to find a trade-off between the influence of noise and loss of information. In addition, by adjusting  $K$  we are able to investigate an

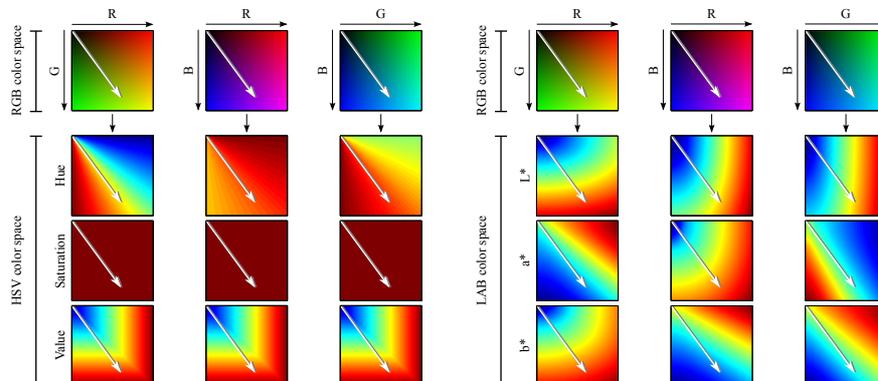


Figure 8: Illustration of color changes along a certain color vector in RGB color space and LAB color space.

image at different scales, which is the basis for the multi-scale LCVF operator introduced in Section 5.4.

The color texture descriptor for an image is obtained by creating a 1D histogram with 256 bins based on the LCVF values for the respective image.

### 5.3. Avoiding Thresholds

Although the LCVF operator as described above already uses the entire color information available and results in a very compact descriptor, we still have the problem that we do not know the optimal choice for the threshold value  $t$  in Equ. (5).

If we assume the color vectors having component values between 0 and 255, we quickly realize that the number of possible choices for  $t$  is rather big (between 0 and  $256\sqrt{3}$  in case of  $F_1$  and  $F_2$ ).

A simple way to obviate the need for a threshold is to compute the aggregated similarity between a color vector and its neighbors. For this purpose the similarity from one color vector to all other color vectors belonging to the neighborhood is computed and summed up. The motivation behind using the aggregated measure is to obtain a single value for each color vector within the neighborhood which reflects the overall similarity of a vector to the remaining vectors belonging to the neighborhood.

As shown in Fig. 9 this computation is performed for each color vector position  $(x_l, y_l)$  within the neighborhood and can be expressed more formally as

$$A^{(p)}(x_l, y_l) = \sum_{k=1}^9 (1 - \delta_{kl}) F_p(I^*(x_l, y_l), I^*(x_k, y_k)), \quad (7)$$

where  $(x_k, y_k)$  denotes the position of the  $k$ -th color vector within the neighborhood with  $l = 1, \dots, 9$  (numbered row-wise, starting from the top left color

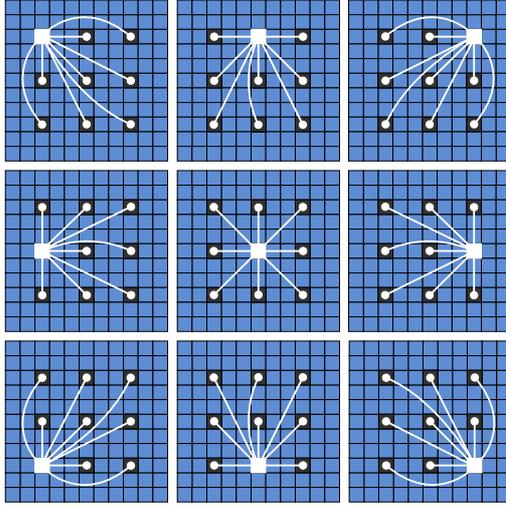


Figure 9: Computation of the aggregated similarity between one color vector at  $(x_l, y_l)$  (shown by a white box) and all color vectors  $(x_k, y_k)$  within the neighborhood. This computation is repeated for each color vector  $l = 1, \dots, 9$  within the neighborhood.

vector).  $\delta_{kl}$  denotes the Kronecker delta defined as

$$\delta_{kl} = \begin{cases} 1, & \text{if } k = l \\ 0, & \text{otherwise} \end{cases} . \quad (8)$$

Using Equ. (7) we can rewrite Equ. (5) as

$$T_{n,i,j}^{(p)} = \begin{cases} 1, & \text{if } A^{(p)}(x_n, y_n) \geq A^{(p)}(i, j) \\ 0, & \text{otherwise} \end{cases} , \quad (9)$$

where  $x_n$  and  $y_n$  denote the position of the  $n$ -th neighboring color vector (ordered in a clock-wise fashion, starting at the top left color vector). Based on this equation we use Equ. (6) to compute the LCVP numbers and, subsequently, the histogram for the LCVP transformed image.

Based on Equ. (7) we define the aggregated similarity measures  $A^{(1)}$  and  $A^{(2)}$  based on the similarity measures  $F_1$  and  $F_2$ . While these measures do not depend on a threshold value, we must note that the computational demand is higher compared to the threshold-based versions. This is due to the additional computations introduced by Equ. (7).

#### 5.4. Multi-scale Extension to LCVP (MS-LCVP)

As already pointed out above, different values for the averaging filter kernel width  $K$  allow us to apply LCVP at different scales. To get a descriptor considering multiple scales of an image we extract 1D histograms with different

choices for  $K$  and concatenate the resulting histograms. Hence, the final LCVP histogram  $H$  for an image is obtained by

$$H = H_{K_1} \oplus \dots \oplus H_{K_S}, \quad (10)$$

where  $H_{K_s}$  denotes the histogram obtained by using LCVP with a filter kernel width of  $K_s$ ,  $S$  is the number of scales used, and  $\oplus$  denotes the histogram concatenation. Starting with a kernel width of  $K_1$  for the finest scale, the subsequent kernel widths  $K_s$  for coarser scales are computed as

$$K_s = K_1 + 2 \cdot (s - 1) \quad \text{with} \quad 1 < s \leq S \quad (11)$$

where  $s$  denotes the scale the kernel width should be computed for. This corresponds to increasing the filter kernel width between two successive scales by 2.

## 6. Classification

To estimate the classification accuracy of our system we use leave-one-patient-out cross-validation (LOPO-CV). In this setup one image out of the database is considered as an unknown image. The remaining images are used to train the classifier (omitting those images which originate from the same patient as the image left out). The class of the unknown image is then predicted by the system. These steps (training and prediction) are repeated for each image, yielding an estimate of the overall classification accuracy.

To actually classify an unknown image (not contained in the training set) we use the k-nearest-neighbor classifier (k-NN). This rather weak classifier has been chosen to emphasize more on quantifying the discriminative power of the features used. The distance between two histograms  $H_1$  and  $H_2$  is computed by using the discrete Bhattacharyya distance metric. We chose this metric since, compared to the histogram intersection, it already delivered slightly higher results in earlier work (Häfner et al., 2009b). This distance metric is defined as

$$D(H_1, H_2) = \sqrt{1 - \sum_{i=1}^B \sqrt{H_{1,i} H_{2,i}}}, \quad (12)$$

where  $B$  denotes the number of bins within the histograms, and  $H_{1,i}$  and  $H_{2,i}$  denote the values of the  $i$ -th bin of  $H_1$  and  $H_2$ , respectively. Prior to computing the distance between two histograms each histogram is normalized such that the bins sum up to 1.

## 7. Experiments

### 7.1. Experimental Setup

The image database used throughout our experiments consists of 716 color images of size of  $256 \times 256$  acquired between the years 2005 and 2009 at the Department of Gastroenterology and Hepatology (Medical University of Vienna)

	Non-neoplastic	Neoplastic	Total
# of images	198	518	716
# of patients	14	32	46

Table 2: Number of images and patients per class.

using a zoom-colonoscope (Olympus Evis Exera CF-Q160ZI/L) with a magnification factor of 150. To obtain the images 40 patients underwent colonoscopy.

Lesions found during colonoscopy have been examined after application of dye-spraying with indigocarmine, as routinely performed in colonoscopy. Biopsies or mucosal resection have been performed in order to get a histopathological diagnosis. Biopsies have been taken from type I, II, and type V lesions, as those lesions need not to be removed or cannot be removed endoscopically. Type III and IV lesions have been removed endoscopically. Out of all acquired images, histopathological classification resulted in 198 non-neoplastic and 518 neoplastic cases. The detailed classification results, which are used as ground truth for our experiments, are shown in Table 2. Since some patients appear in more than one class, the total sum of patients shown in Table 2 is slightly higher than 40.

Since the choice of the  $k$ -value for the  $k$ -NN classifier may have an impact on the classification rates, we chose a set of values (leading to the best overall classification rates in most cases) and carried out the classification of all images for each of these values. We then computed the mean classification rates over the overall classification rates achieved using the different choices for  $k$ . In addition we computed the respective deviations of the results from these mean values (maximum absolute difference between the mean rates and each experimental result for the different choices for  $k$ ). The set of  $k$ -values used is  $\{3, 5, 7, 9\}$ . The reason for choosing odd values only is to avoid draws during the classification. While such draws could be resolved by assigning an unknown image to the class with the highest a priori probability this would lead to biased results due to the rather high number of neoplastic images in our image database.

In order to be able to compare the classification accuracy and runtime performance of the LCVP operator to other color-based LBP operators, we also conducted experiments using the OCLBP operator and JC-MB-LBP operator. To allow a comparison of the classification results, we use the feature extraction and classification methodologies, as described in the previous sections, for these operators too (single scale and multi-scale). To make a comparison in terms of runtime performance feasible, all operators have been implemented and compared using Java 1.6. The runtime measurements have been carried out on a machine equipped with an Intel Core i7 CPU at 2.66 GHz (single-threaded), running Linux.

We also conducted all experiments in two different color spaces, namely the RGB color space and the CIELAB color space (Gonzalez and Woods, 2002). The RGB color space has been chosen since the endoscopic images have been supplied to us in this color format. The CIELAB color space, on the other hand, is an interesting alternative to the RGB color space since it corresponds

more to human perception. Apart from that, the relative perceptual distance between two colors in the CIELAB space can be approximated by computing the Euclidean distance between the according color vectors (Jain, 1989), which perfectly corresponds to our similarity measure  $F_1$  (and therefore also to  $A^{(1)}$ ). To investigate the overall importance of color information we also conducted experiments using JC-MB-LBP with grayscale images (1D-histograms instead of 2D-histograms).

The single scale results for the different methods presented in the next section have been obtained using a filter kernel width of  $K = 7$ . While we also conducted experiments with other values, this choice almost always yielded the best overall classification results (in case of all methods investigated). It must be noted, however, that this value is of course image resolution dependent and most probably will differ for images with lower or higher resolutions.

In case of multi-scale experiments  $K_1$  has been set to 5. Using 3 scales and Equ. (11) we end up with the filter kernel widths 5, 7, and 9 for the different scales. Hence, besides using the scale used in the single scale experiments we also use one finer and one coarser scale. This set of filter widths has been used throughout all multi-scale experiments.

It must be noted that we also apply an averaging filter to the images in case of the OCLBP operator. This way we are able to introduce different scales for the OCLBP operator the same way we do it in case of LCVP and JC-MB-LBP.

Since the JC-MB-LBP operator uses two color channels only, we conducted experiments for all different pairs of color channels. Concerning the average classification accuracy computed over different choices for  $k$ , only the channel combination yielding the highest overall classification for the respective  $k$ -value is taken into consideration. In case of the RGB experiments this was always the combination RB (red and blue channel), while in case of the CIELAB color space either the color channel combination  $L^*a^*$  or  $L^*b^*$  yielded the highest classification accuracies.

In order to be able to assess whether two different methods produce statistically significant differences in the results obtained we employ McNemar's test (Everitt, 1977). For two methods  $M_1$  and  $M_2$  this test statistic keeps track of the number of images which are misclassified by method  $M_1$  but classified correctly by method  $M_2$  (denoted by  $n_{01}$ ) and vice versa (denoted by  $n_{10}$ ). The test statistic, which is approximately Chi Square distributed (with one degree of freedom), is then computed as

$$T = \frac{(|n_{01} - n_{10}| - 0.5)^2}{n_{01} + n_{10}}. \quad (13)$$

From  $T$  the  $p$ -value can be computed as

$$p = 1 - F_{\chi_1^2}(T) \quad (14)$$

where  $F_{\chi_1^2}$  denotes the cumulative distribution function of the Chi Square distribution with one degree of freedom. The null-hypothesis  $H_0$  for McNemar's test is that the outcomes of  $M_1$  and  $M_2$  lead to equal error rates. Given a fixed

Single scale					
Method	Accuracy	Specificity	Sensitivity	$\mathbf{A}^{(1)}$	$\mathbf{A}^{(2)}$
LCVP $A^{(1)}$	75.9 $\pm$ 0.3	33.3 $\pm$ 3.0	92.2 $\pm$ 0.7		
LCVP $A^{(2)}$	74.2 $\pm$ 0.6	25.5 $\pm$ 4.3	92.8 $\pm$ 1.3		
OCLBP	70.3 $\pm$ 1.8	33.5 $\pm$ 4.9	84.4 $\pm$ 3.7	✓ (+)	✓ (+)
<b>JC-MB-LBP</b>	<b>82.4 <math>\pm</math> 0.8</b>	<b>52.4 <math>\pm</math> 4.2</b>	<b>93.9 <math>\pm</math> 1.2</b>	✓ (-)	✓ (-)
JC-MB-LBP (gray)	75.4 $\pm$ 1.1	27.8 $\pm$ 7.1	93.6 $\pm$ 3.9		
Multi-scale					
Method	Accuracy	Specificity	Sensitivity	$\mathbf{A}^{(1)}$	$\mathbf{A}^{(2)}$
LCVP $A^{(1)}$	79.1 $\pm$ 0.5	45.5 $\pm$ 2.0	91.9 $\pm$ 0.7		
LCVP $A^{(2)}$	79.6 $\pm$ 0.8	40.4 $\pm$ 5.1	94.5 $\pm$ 1.1		
OCLBP	70.0 $\pm$ 1.1	31.9 $\pm$ 6.9	84.6 $\pm$ 3.3	✓ (+)	✓ (+)
<b>JC-MB-LBP</b>	<b>82.7 <math>\pm</math> 0.8</b>	<b>55.1 <math>\pm</math> 4.5</b>	<b>93.2 <math>\pm</math> 0.8</b>		
JC-MB-LBP (gray)	74.6 $\pm$ 0.7	22.0 $\pm$ 9.3	94.7 $\pm$ 2.6	✓ (+)	✓ (+)

Table 3: Comparison of the average classification results for the RGB color space (the methods yielding the highest mean overall classification rates are shown in bold).

significance level  $\alpha$ , there is evidence that the methods  $M_1$  and  $M_2$  produce significantly different results if  $p < \alpha$ . As a consequence we can reject the null-hypothesis  $H_0$ . Throughout this work we chose a significance level of  $\alpha = 0.05$ . This implies that, if  $M_1$  and  $M_2$  are significantly different, there is a confidence level of 95% that the differences between the outcomes of the methods are not caused by random variation.

## 7.2. Classification Results

Tables 3 and 4 show the average classification results we achieved with the different methods (in case of the RGB and the CIELAB color space, respectively). The results represent the mean classification rates obtained by carrying out the experiments with different choices for the number of neighbors for the k-NN classifier. To get a picture of the result variations between the different choices for  $k$ , the tables also contain the observed deviations from the mean rates.

Since we also aim at answering the question whether the LCVP operator generates results which are significantly different compared to the other LBP variants tested, we carry out McNemar’s test for method pairs only which include either LCVP  $A^{(1)}$  or LCVP  $A^{(2)}$ . Throughout Tables 3 and 4 a check mark (✓) in column  $\mathbf{A}^{(1)}$  ( $\mathbf{A}^{(2)}$ ) shows that the observed result differences between the method in the respective table row and LCVP  $A^{(1)}$  (LCVP  $A^{(2)}$ ) are statistically significant according to McNemar’s test. If one of the LCVP variants delivers significantly higher (lower) classification results compared to the method in the respective row this is indicated by ‘+’ (‘-’). To obtain the results for the significance results we fixed the parameters for the methods to values which on average resulted in the best classification results ( $k = 7$  for the k-NN classifier, an averaging filter kernel width of  $K = 7$ ,  $K_1 = 5$  in case of multi-scale

Single scale					
Method	Accuracy	Specificity	Sensitivity	A <sup>(1)</sup>	A <sup>(2)</sup>
LCVP A <sup>(1)</sup>	76.9 ± 0.2	37.0 ± 1.4	92.1 ± 0.4		
LCVP A <sup>(2)</sup>	78.7 ± 2.2	35.9 ± 1.5	95.1 ± 3.0		
OCLBP	71.5 ± 0.7	27.0 ± 4.8	88.5 ± 2.7	✓ (+)	✓ (+)
<b>JC-MB-LBP</b>	<b>85.3 ± 0.7</b>	<b>76.6 ± 9.0</b>	<b>88.6 ± 2.5</b>	✓ (-)	✓ (-)
JC-MB-LBP (gray)	75.4 ± 1.1	27.8 ± 7.1	93.6 ± 3.9		✓ (+)
Multi-scale					
Method	Accuracy	Specificity	Sensitivity	A <sup>(1)</sup>	A <sup>(2)</sup>
LCVP A <sup>(1)</sup>	79.5 ± 0.5	46.3 ± 1.4	92.2 ± 0.7		✓ (+)
LCVP A <sup>(2)</sup>	85.0 ± 1.9	57.6 ± 2.0	95.5 ± 2.2	✓ (-)	
OCLBP	71.7 ± 0.8	28.4 ± 3.9	88.2 ± 2.5	✓ (+)	✓ (+)
<b>JC-MB-LBP</b>	<b>85.1 ± 0.2</b>	<b>65.0 ± 3.2</b>	<b>92.8 ± 1.4</b>	✓ (-)	
JC-MB-LBP (gray)	74.6 ± 0.7	22.0 ± 9.3	94.7 ± 2.6	✓ (+)	✓ (+)

Table 4: Comparison of the average classification results for the LAB color space (the methods yielding the highest mean overall classification rates are shown in bold).

experiments, and for the method JC-MB-LBP the color channel combinations RB and L\*b\* in case of RGB and CIELAB experiments, respectively).

### 7.2.1. LCVP and LBP Variants

From Tables 3 and 4 we notice that in most cases there are no significant differences between the overall classification rates obtained by LCVP A<sup>(1)</sup> and LCVP A<sup>(2)</sup>. Only in the case of multi-scale experiments carried out in the CIELAB color space LCVP A<sup>(2)</sup> delivers significantly higher overall classification rates (approximately 5.5 % higher compared to LCVP A<sup>(1)</sup>). Comparison of single scale and multi-scale overall classification accuracies reveals that the LCVP operator benefits from multiple scales (up to 6.3% higher compared to single scale experiments – in case of CIELAB) as well as from switching to the CIELAB color space (approximately 5.4% higher compared to the results from the RGB experiments – in case of multiple scales). In all cases the classification performance gain is higher for LCVP A<sup>(2)</sup> compared to LCVP A<sup>(1)</sup>. The reason for LCVP A<sup>(2)</sup> delivering higher results in case of the CIELAB color space compared to the RGB color space stems from the fact, that as already indicated earlier, LCVP A<sup>(2)</sup> is better suited for CIELAB. While measuring the approximate brightness difference between two color vectors in case of RGB, in case of CIELAB LCVP A<sup>(2)</sup> also measures color differences. We therefore recommend using the multi-scale extension with the CIELAB color space due to its superiority compared to the single scale operator and the RGB results.

Compared to the OCLBP operator the LCVP operator is always able to achieve significantly higher classification rates. In contrast to the LCVP operator, the OCLBP operator always delivers roughly equal classification rates, no matter whether we switch from single scale to multiple scales or from the RGB color space to CIELAB color space.

In case of the JC-MB-LBP operator the picture is quite different. When

using a single scale only the color version of JC-MB-LBP always delivers significantly higher classification rates compared to both LCVP variants (up to 8.4% in case of the CIELAB experiments). However, when using multiple scales the superiority of JC-MB-LBP vanishes. As a result, in case of RGB experiments JC-MB-LBP delivers only an insignificantly higher overall classification rate compared to both LCVP variants. In case of CIELAB experiments at least LCVP  $A^{(2)}$  is able to deliver an approximately equal overall classification rate compared to JC-MB-LBP (again, the classification rate obtained with JC-MB-LBP is only insignificantly higher compared to LCVP  $A^{(2)}$ ). When applied to grayscale images, the JC-MB-LBP operator delivers either equal or significantly lower classification rates compared to LCVP. Similar to the OCLBP operator, the JC-MB-LBP is not able to benefit from multiple scales or from switching from the RGB colorspace to the CIELAB color space. Although LCVP  $A^{(2)}$  approximately reflects brightness differences when applied in the RGB color space, the operator delivers higher classification accuracies compared to JC-MB-LBP when applied to grayscale images. This stems from the fact that the comparison between neighboring pixels is slightly more complex in case of LCVP  $A^{(2)}$ .

Considering the rather low number of images available for our experiments (see Table 2) we may also run into the problem of overfitting in case of operators relying on high-dimensional features (high number of bins used in our histograms). The situation gets even worse since we use the LOPO-CV which reduces the number of training samples available for each image by the total number of images available for the respective patient. As a consequence we might run into a problem commonly known as “curse of dimensionality” (Bellman, 1961). This problem is caused by the fact that high-dimensional feature vectors are more likely to be sparsely distributed in the feature space if the pool of training images is not sufficiently large. Hence, especially the operators OCLBP and JC-MBLBP, both using more than 2000 bins (even in the single scale case), may suffer from the “curse of dimensionality”. This indicates that the results we achieved could get even better in practical use due to the higher number of training samples which would then be available.

From Tables 3 and 4 we also notice that – compared to JC-MB-LBP based on color – the LCVP operator yields a rather low specificity. Experiments with a balanced training set have shown that this most probably is not due to the imbalance in the training set used. Hence, it is quite possible that the LCVP features extracted are not distinctive enough to capture the properties of the non-neoplastic images as good as JC-MB-LBP.

### 7.2.2. *Previously Developed Approaches*

To be able to compare the results of previously developed approaches (color-extended version of DT-CWT-based method in (Kwitt and Uhl, 2007) and Fourier-based in (Häfner et al., 2010)) to the results of the LCVP operator, we repeated the respective experiments on the image database used throughout this work. Compared to alternative feature extraction approaches developed earlier (color-extended version of DT-CWT-based method in (Kwitt and Uhl,

2007) and Fourier-based in (Häfner et al., 2010)) the overall classification rates yielded by color-based LBP schemes are highly competitive.

While the Fourier-based method always outperforms all LBP and LCVP variants (the best result is obtained in case of CIELAB), at least in case of LCVP  $A^{(2)}$  applied in the CIELAB color space the classification rates are only insignificantly higher ( $85.0 \pm 1.9\%$  for LCVP compared to  $87.4\%$  in case of the Fourier features). It must be noted, however, that the Fourier-based method proposed in (Häfner et al., 2010) uses a genetic algorithm for feature selection in order to choose the most suitable ring filters (due to the vast amount of such filters which may potentially be used). As a consequence, one explanation for the superiority of this method is the selection of optimal features, which however results in overfitting. Such an optimization is neither used in case of the LBP variants nor in the case of LCVP.

The overall classification rate of the DT-CWT-based method is significantly higher in case of the RGB color space ( $86.7\%$  compared to  $79.6 \pm 0.8\%$  in case of LCVP  $A^{(2)}$  with multiple scales). According to McNemar’s test, however, the result of the DT-CWT method applied in the RGB color space is not significantly different from the result achieved with LCVP  $A^{(2)}$  using multiple scales in the CIELAB color space ( $85.0 \pm 1.9\%$ ). In the CIELAB color space the DT-CWT method yields a significantly lower classification rate ( $84.4\%$  compared to  $85.0 \pm 1.9\%$  in case of LCVP  $A^{(2)}$  with multiple scales).

We notice that, compared to the results obtained in (Kwitt and Uhl, 2007) and (Häfner et al., 2010), the results presented for these methods in this work are considerably lower. The result discrepancies are due to the different kinds of cross-validation protocols used. While in (Kwitt and Uhl, 2007) and (Häfner et al., 2010) the leave-one-out cross-validation (LOO-CV) has been used, this work employs LOPO-CV, which is more restrictive (the training set must not contain features from the patient from whom an image is currently classified). This results in the limitation that the training set available for each image is smaller. In addition, LOPO-CV also inhibits the use of more than one image from the same polyp. Thus, the images contained within the training set are less likely to be similar to the image to be classified. This makes it more likely that neighbors from wrong image classes are selected as nearest neighbors during the classification, resulting in more classification errors and, therefore, lowered classification accuracies.

### 7.3. Performance Analysis

Table 5 shows the results of our performance measurements (times are given in milliseconds). The column  $T_O$  shows the time needed to compute the respective operator on a single image including the time consumed by the histogram generation.  $T_C$  represents the time needed to classify a single image using the  $k$ -NN classifier with  $k = 1$  and the Bhattacharyya distance. Carrying out the timing measurements for  $k = 1$  only can be justified by the fact that the choice of this value has no noticeable influence on the time needed for classification. This is due to the fact that, no matter which  $k$ -value we choose, we have to

Single scale					
Method	Bins	$T_O$	$T_C$	$T_T$	SF
LCVP $A^{(1)}$	256	166	6	172	2.1
LCVP $A^{(2)}$	256	75	7	82	1.0
OCLBP	2304	80	54	134	1.6
JC-MB-LBP	65536	38	426	464	5.7
Multi-scale					
Method	Bins	$T_O$	$T_C$	$T_T$	SF
LCVP $A^{(1)}$	768	505	26	531	2.1
LCVP $A^{(2)}$	768	232	27	259	1.0
OCLBP	6912	254	158	412	1.6
JC-MB-LBP	196608	126	1824	1950	7.5

Table 5: Result of the performance analysis of the different methods (time measurements are given in milliseconds).

compute the distances from an unknown sample to all training samples (which is the most time consuming part of the classification).

To obtain the values  $T_O$  and  $T_C$  we carried out the respective computations for all images in our database and divided the resulting times by the number of images.  $T_T = T_O + T_C$  is the total time. In case of the CIELAB experiments the color space conversion takes place as a pre-processing step. Since it therefore does not matter which color space we base our experiments on, we present time measurements for the RGB color space only. The column  $SF$  roughly indicates how much higher the total computational demand is for each method, compared to the fastest method listed in Table 5, which is  $A^{(2)}$  (determined separately for single scale experiments and multi-scale experiments).

As we can easily see from the number of bins needed by each method (see Table 5), the LCVP operator is superior in terms of compactness. In case of the single scale tests we obtain histograms containing 256 bins only, while OCLBP results in 9 such histograms, resulting in a total of 2304 bins. The JC-MB-LBP method contains even more bins since in this case we compute 2D joint-color histograms.

Hence, it is no surprise that the proposed operator is also the fastest one in terms of the classification for a single image. But it is also obvious that in case of LCVP the time needed for classification does not affect  $T_T$  very much since  $T_O$  is significantly higher in these cases. In case of JC-MB-LBP the situation is the other way round. While the time needed for the computation of the JC-MB-LBP data and the construction of the histograms ( $T_O$ ) is considerably lower compared to the other methods, the time needed for classification is significantly higher. Altogether this makes the JC-MB-LBP operator the slowest one.

In case of the multi-scale experiments a similar behavior can be observed. As indicated in Section 3, in this case the total computational demand for each method roughly multiplies by 3. Regarding  $T_T$ , LCVP  $A^{(1)}$  and  $A^{(2)}$  are considerably faster compared to JC-MB-LBP. OCLBP is slower than LCVP  $A^{(2)}$  but slightly faster than LCVP  $A^{(1)}$ .

From the definitions of  $F_1$  and  $F_2$  (see Equ. (3) and (4)) one would expect that  $T_O$  should be lower for LCVP  $A^{(1)}$  compared to LCVP  $A^{(2)}$ . However, this is not the case, as we notice from Table 5. The reason for this behavior is that in case of  $F_1$  the computation of  $\|v_1 - v_2\|$  (coming along with a rather slow square root) has to be carried out for each possible pair of pixel positions within a neighborhood (see Equ. (7)). In case of  $F_2$  the needed color vector lengths are precomputed, hence, Equ. (7) reduces to the computation of the absolute values of a couple of subtraction results. Considering the huge number of different color vector pairs which must be taken into account in case of  $F_1$  (approximately 500000 for our images of size  $256 \times 256$ ),  $A^{(2)}$  will always be faster compared to  $A^{(1)}$  since in case of  $F_2$  we need to compute one square root only per color vector (65536 in case of our images).

If we assume an endoscopy video to be captured at a frame rate of 25 frames per second a real-time application demands processing times of at most 40 milliseconds for a single frame. From the timing results in Table 5 we notice that this requirement is currently not met by any of the operators. Nevertheless, the significant speedup between JC-MB-LBP and LCVP  $A^{(2)}$  represents a step into the direction of a real-time application – especially when considering that both methods perform equally well in case of CIELAB multi-scale experiments.

In addition, LCVP can potentially be further optimized in terms of the runtime performance by employing hardware acceleration (e.g. GPU-based computing (Sanders and Kandrot, 2010)).

## 8. Conclusion

In this work we proposed a novel color texture operator which is based on a noise-robust LBP variant. Treating the color components of pixels as color vectors and introducing suitable similarity measures resulted in a compact descriptor for an image, which nevertheless incorporates all color information available.

Applied to the classification of chromo-colonoscopy imagery, we have shown that the inclusion of color information in LBP-based classification improves staging. This is supported by the fact that the grayscale variant of JC-MB-LBP is mostly outperformed by the color-based LBP operators (except for OCLBP, which consistently delivers lower classification rates compared to all other methods).

Compared to OCLBP, the different variants of the LCVP operator are always significantly superior, in terms of speed as well as in terms of the overall classification accuracy. But we notice that the JC-MB-LBP operator in many cases delivers higher classification accuracies as compared to LCVP. However, the multi-scale extension of the LCVP operator in the CIELAB color space either outperforms the JC-MB-LBP operator or at least performs equally well (with no statistically significant differences). In addition, comparing the computational demand between LCVP and all other LBP variants reveals that LCVP is up to 7.5 times faster compared to the other methods. This can be attributed to the compactness of the features used in case of LCVP.

From Table 5 we may further deduce that LCVP results may be improved by employing a stronger classifier at a moderate increase of overall classification time: The time spent in classification ( $T_C$ ) covers only a small share of the total time ( $T_T$ ) for LCVP, while for JC-MB-LBP  $T_T$  is already dominated by the  $T_C$  share. Hence, employment of a more costly classification will dramatically increase  $T_T$  in case of JC-MB-LBP, but will only marginally affect LCVP.

The promising results and the rather low computational complexity of the LCVP operator make it an alternative to JC-MB-LBP. Moreover, LCVP is also a sensible candidate for inclusion into a classifier ensemble.

One strength of the multi-scale LCVP operator is the fact that it needs only two parameters (the averaging filter kernel width for the finest scale and the number of scales to be used). Apart from that, since the operator is not tailored to a specific type of imagery we also conclude that the operator is generic and may also be applicable to other application scenarios.

## 9. Acknowledgments

This work is partially funded by the Austrian Science Fund (FWF) under Project No. L366-N15 and by the Austrian National Bank “Jubiläumsfonds” Project No. 12514.

## References

- Ameling, S., Wirth, S., Paulus, D., Lacey, G., Vilarino, F., 2009. Texture-based polyp detection in colonoscopy, in: *Bildverarbeitung für die Medizin 2009*. Springer Berlin Heidelberg. number 15 in *Informatik aktuell*, pp. 346–350.
- André, B., Vercauteren, T., Perchant, A., Buchner, A.M., Wallace, M.B., Ayache, N., 2009a. Endoscopic image retrieval and classification using invariant visual features, in: *In Proceedings of the 6th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'09)*, Boston, Massachusetts, USA. pp. 346–349.
- André, B., Vercauteren, T., Perchant, A., Wallace, M.B., Buchner, A.M., Ayache, N., 2009b. Introducing space and time in local feature-based endoscopic image retrieval, in: *Proceedings of the MICCAI 2009 Workshop - Medical Content-based Retrieval for Clinical Decision (MCBR-CDS'09)*, London, UK. pp. 18–30.
- Bellman, R., 1961. *Adaptive Control Processes: A Guided Tour*. Princeton University Press.
- Bruno, M.J., 2003. Magnification endoscopy, high resolution endoscopy, and chromoscopy; towards a better optical diagnosis. *Gut* 52, 7–11.

- Chang, C.C., Hsieh, C.R., Lou, H.Y., Fang, C.L., Tiong, C., Wang, J.J., Wei, I.V., Wu, S.C., Chen, J.N., Wang, Y.H., 2009. Comparative study of conventional colonoscopy, magnifying chromoendoscopy, and magnifying narrow-band imaging systems in the differential diagnosis of small colonic polyps between trainee and experienced endoscopist. *Int J Colorectal Dis* 24, 1413–1419.
- Connah, D., Finlayson, G.D., 2006. Using local binary pattern operators for colour constant image indexing, in: *Proceedings of the 3rd European Conference on Colour in Graphics, Imaging and Vision (CGIV'06)*, Leeds, UK.
- Everitt, B., 1977. *The Analysis of Contingency Tables*. Chapman and Hall.
- Gonzalez, R.C., Woods, R.E., 2002. *Digital Image Processing*. Prentice Hall. 2nd edition.
- Gross, S., Stehle, T., Behrens, A., Auer, R., Aach, T., Winograd, R., Trautwein, C., Tischendorf, J., 2009. A comparison of blood vessel features and local binary patterns for colorectal polyp classification, in: *Proceedings of Medical Imaging 2008: Computer-Aided Diagnosis*, Orlando, USA.
- Häfner, M., Brunauer, L., Payer, H., Resch, R., Gangl, A., Uhl, A., Vécsei, A., Wrba, F., 2010. Computer-aided classification of zoom-endoscopic images using fourier filters. *IEEE Trans Inform Tech Biomed* 14, 958–970.
- Häfner, M., Gangl, A., Kwitt, R., Uhl, A., Vécsei, A., Wrba, F., 2009a. Improving pit-pattern classification of endoscopy images by a combination of experts, in: *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'09)*, London, UK. pp. 247–254.
- Häfner, M., Gangl, A., Liedlgruber, M., Uhl, A., Vécsei, A., Wrba, F., 2009b. Pit pattern classification using extended local binary patterns, in: *Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine (ITAB'09)*, Larnaca, Cyprus.
- Häfner, M., Gangl, A., Liedlgruber, M., Uhl, A., Vécsei, A., Wrba, F., 2009c. Pit pattern classification using multichannel features and multiclassification, in: T.P. Exarchos, A. Papadopoulos, D.F. (Ed.), *Handbook of Research on Advanced Techniques in Diagnostic Imaging and Biomedical Applications*. IGI Global, Hershey, PA, USA, pp. 335–350.
- Hurlstone, D.P., 2002. High-resolution magnification chromoendoscopy: Common problems encountered in “pit pattern” interpretation and correct classification of flat colorectal lesions. *Am J Gastroenterol* 97, 1069–1070.
- Iakovidis, D.K., Maroulis, D.E., Karkanis, S.A., 2006. An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy. *Comput Biol Med* 36, 1084–1103.

- Iakovidis, D.K., Maroulis, D.E., Karkanis, S.A., Brokos, A., 2005. A comparative study of texture features for the discrimination of gastric polyps in endoscopic video, in: Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05), Dublin, Ireland. pp. 575–580.
- Jain, A., 1989. Fundamentals of Digital Image Processing. Prentice Hall.
- Karkanis, S., 2003. Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE Trans Inform Tech Biomed* 7, 141–152.
- Karkanis, S.A., Iakovidis, D., Karras, D., Maroulis, D., 2001. Detection of lesions in endoscopic video using textural descriptors on wavelet domain supported by artificial neural network architectures, in: Proceedings of the IEEE International Conference in Image Processing (ICIP'01), Thessaloniki, Greece. pp. 833–836.
- Kato, S., Fu, K.I., Sano, Y., Fujii, T., Saito, Y., Matsuda, T., Koba, I., Yoshida, S., Fujimori, T., 2006. Magnifying colonoscopy as a non-biopsy technique for differential diagnosis of non-neoplastic and neoplastic lesions. *World J Gastroenterol* 12, 1416–1420.
- Krishnan, S.M., Yang, X., L.Chan, K., Kumar, S., Goh, P.M.Y., 1998. Intestinal abnormality detection from endoscopic images, in: Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS'98), Hong Kong, China. pp. 895–898.
- Kudo, S., Hirota, S., Nakajima, T., Hosobe, S., Kusaka, H., Kobayashi, T., Himori, M., Yagyuu, A., 1994. Colorectal tumours and pit pattern. *J Clin Pathol* 47, 880–885.
- Kudo, S., Tamura, S., Nakajima, T., Yamano, H., Kusaka, H., Watanabe, H., 1996. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointest Endosc* 44, 8–14.
- Kwitt, R., Uhl, A., 2007. Modeling the marginal distributions of complex wavelet coefficient magnitudes for the classification of zoom-endoscopy images, in: Proceedings of the IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA '07), Rio de Janeiro, Brasil. pp. 1–8.
- Liao, S., Zhu, X., Lei, Z., Zhang, L., Li, S., 2007. Learning multi-scale block local binary patterns for face recognition, in: Advances in Biometrics. Springer, pp. 828–837.
- Mackiewicz, M., Berens, J., Fisher, M., 2008. Wireless capsule endoscopy color video segmentation. *IEEE Trans Med Imag* 27, 1769–1781.
- Mäenpää, T., Ojala, T., Pietikäinen, M., Soriano, M., 2000. Robust texture classification by subsets of local binary patterns, in: Proceedings of the 15th International Conference on Pattern Recognition, (ICPR'00), Barcelona, Spain. pp. 935–938.

- Mäenpää, T., Pietikäinen, M., 2004. Classification with color and texture: jointly or separately? *Pattern Recogn* 37, 1629–1640.
- Mäenpää, T., Pietikäinen, M., Viertola, J., 2002. Separating color and pattern information for color texture discrimination, in: *Proceedings of the 16th International Conference on Pattern Recognition, (ICPR'02)*, Quebec City, Canada. pp. 668–671.
- Maroulis, D.E., Iakovidis, D.K., Karkanis, S.A., Karras, D.A., 2003. CoLD: a versatile detection system for colorectal lesions in endoscopy video-frames. *Comput Meth Programs Biomed* 70, 151–166.
- Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on feature distributions. *Pattern Recogn* 29, 51–59.
- Pietikäinen, M., Mäenpää, T., Viertola, J., 2002. Color texture classification with color histograms and local binary patterns, in: *Proceedings of the 2nd International Workshop on Texture Analysis and Synthesis, Copenhagen, Denmark*. pp. 109–112.
- Raghavendra, M., Hewett, D.G., Rex, D.K., 2010. Differentiating adenomas from hyperplastic colorectal polyps: narrow-band imaging can be learned in 20 minutes. *Gastrointest Endosc* 72, 572 – 576.
- Rastogi, A., Pondugula, K., Bansal, A., Wani, S., Keighley, J., Sugar, J., Callahan, P., Sharma, P., 2009. Recognition of surface mucosal and vascular patterns of colon polyps by using narrow-band imaging: interobserver and intraobserver agreement and prediction of polyp histology. *Gastrointest Endosc* 69, 716–722.
- Sanders, J., Kandrot, E., 2010. *CUDA by Example: An Introduction to General-Purpose GPU Programming*. Addison-Wesley Longman, Amsterdam.
- Smolka, B., Venetsanopoulos, A.N., 2007. Noise reduction and edge detection in color images, in: Lukac, R., Plataniotis, K.N. (Eds.), *Color Image Processing: Methods and Applications*. CRC Press. Image Processing Series, pp. 75–102.
- Song, L.M.W.K., Adler, D.G., Chand, B., Conway, J.D., Croffie, J.M.B., DiSario, J.A., Mishkin, D.S., Shah, R.J., Somogyi, L., Tierney, W.M., Petersen, B.T., 2007. Chromoendoscopy. *Gastrointest Endosc* 66, 639–649.
- Stehle, T., Auer, R., Gross, S., Behrens, A., Wulff, J., Aach, T., Winograd, R., Trautwein, C., Tischendorf, J.J.W., 2009. Classification of colon polyps in NBI endoscopy using vascularization features, in: *Medical Imaging 2009: Computer-Aided Diagnosis, SPIE, Orlando, USA*.
- Tischendorf, J.J.W., Gross, S., Winograd, R., Hecker, H., Auer, R., Behrens, A., Trautwein, C., Aach, T., Stehle, T., 2010. Computer-aided classification of colorectal polyps based on vascular patterns: a pilot study. *Endoscopy* 42, 203–207.

Zanoni, E., Cutait, R., Averbach, M., et al., 2007. Magnifying colonoscopy: interobserver agreement in the assessment of colonic pit patterns and its correlation with histopathological findings. *Int J Colorectal Dis* 22, 1383–1388.

**Michael Häfner** was born on December 9, 1968 in Vienna, Austria. He studied Medicine in Vienna and Perugia, Italy and graduated on January 23, 1995. He started working at the Department of Gastroenterology and Hepatology of the Medical University of Vienna in 1994, where he also was trained in Internal Medicine. He is specialist in Gastroenterology and Hepatology with his main focus on interventional endoscopy. Michael Häfner is currently secretary of the Austrian Society of Gastroenterology and Hepatology and organizer of several courses on interventional endoscopy, including an annual postgraduate course on endoscopy in Vienna. He acts as tutor at various national and international hands-on training courses. Recently he spent half a year in Malaysia to start up an endoscopy unit in Kuala Lumpur. In September 2009 he became chief physician of the Department of Internal Medicine at the St. Elisabeth Hospital, Vienna, Austria.

**Michael Liedlgruber** received his Master's degree in Computer Sciences from the University of Salzburg (Austria) in May 2006. His Master's thesis has already been devoted to the classification of colonic polyps using wavelets. Currently he is working on his PhD thesis at the Department of Computer Sciences (University of Salzburg) in the same field of research. His main research interests include medical image processing and pattern recognition.

**Andreas Uhl** is an associate professor at the Department of Computer Sciences (University of Salzburg) where he heads the Multimedia Processing and Security Lab. His research interests include image and video processing and compression, wavelets, media security, medical imaging, biometrics, and number-theoretical numerics.

**Andreas Vécsei** graduated (Dr. med. univ.) at the University of Vienna. His post-graduate training in Austria and Kerala (India) made him receive his license to practice general pediatrics. After having been instructed in pediatric gastroenterology according to the syllabus of GPGE (an association of German-speaking pediatric gastroenterologists) at the Medical University of Vienna and Munich, he received his certificate in 2006. Since 2006 he is in charge of the Endoscopy Unit of the St. Anna Children's Hospital, Vienna. His current research interests include *Helicobacter pylori* infection, celiac disease and computer-assisted evaluation of endoscopic images including the assessment of villous atrophy on endoscopic examination of the small intestine.

**Friedrich Wrba** is professor of Pathology at the Department of Clinical Pathology of the University of Vienna. After training as a general practitioner in various Austrian hospitals he started his special training in clinical Pathology, which he finished successfully at the Department of Clinical Pathology at the University of Vienna in the year 1987. In 1990–1992 he spent 2.5 years as a Visiting Research Fellow at the Imperial Cancer Research Fund Laboratories in London, England, (Walter Bodmer's laboratory). Since 1991 he is lecturer and Professor of Pathology, and staff member of the department. His areas of special interest are gastroenterology, hepatology, and molecular pathology.