Computer-Assisted Pit-Pattern Classification in Different Wavelet Domains for Supporting Dignity Assessment of Colonic Polyps

Michael Häfner [b], Roland Kwitt [a,*], Andreas Uhl [a],
Friedrich Wrba [c], Alfred Gangl [b], Andreas Vécsei [d]

[a]*Department of Computer Sciences,*
University of Salzburg, 5020 Salzburg, Austria

[b]*Department of Gastroenterology and Hepatology,*
Vienna Medical University, 1090 Vienna, Austria

[c]*Department of Clinical Pathology,*
Vienna Medical University, 1090 Vienna, Austria

[d]*St. Anna Children's Hospital,*
1090 Vienna, Austria

Abstract

In this paper, we show that zoom-endoscopy images can be well classified according to the pit-pattern classification scheme by using texture-analysis methods in different wavelet domains. We base our approach on three different variants of the wavelet transform and propose that the color-channels of the RGB and LAB color model are an important source for computing image features with high discriminative power. Color-channel information is incorporated by either using simple feature vector concatenation and cross-cooccurrence matrices in the wavelet domain. Our experimental results based on k-Nearest Neighbor classification and forward feature selection exemplify the advantages of the different wavelet transforms and show that color-image analysis is superior to grayscale image analysis regarding our medical image classification problem.

Key words: Computer-Assisted Pit-Pattern Classification, Wavelet Transformation, Colorectal Cancer, Color-Texture Analysis

* Corresponding Author; Full-Address: Department of Computer Sciences, University of Salzburg, Jakob-Haringer Strasse 2, 5020 Salzburg, Austria; Tel.: (0043) 662 8044-6310, Fax: (0043) 662 8044-172.

Email addresses: `rkwitt@gmx.at` (Roland Kwitt), `uhl@cosy.sbg.ac.at` (Andreas Uhl).

1 Motivation

Recent statistics of the American Cancer Society reveal that colorectal cancer is the third most common cancer in men and women and the second most common cause of US cancer deaths. Since most colorectal cancers develop from polyps, a regular inspection of the colon is recommended, in order to detect lesions with a malignant potential or early cancer. A common medical procedure to examine the inside of the colon is colonoscopy, which is usually carried out with a conventional video-endoscope. A diagnostic benefit can be achieved by employing so called zoom-endoscopes, which achieve a magnification factor of up to 150 by means of an individually adjustable lens. In combination with dye-spraying to enhance the visual appearance (chromo-endoscopy) of the colon mucosa, zoom-endoscopy can reveal characteristic surface patterns, which can be interpreted by experienced physicians. Commonly used dyes are either methylene-blue, or indigo-carmine, which both lead to a plastic effect. In the research work of (Kudo, 1994), the macroscopic appearance of colorectal polyps is systematically described and results in the so called *pit-pattern* classification scheme.

In this work, we tackle the problem of computer-assisted pit-pattern classification using texture discrimination methods in the wavelet domain. This is mainly motivated by the work of (Kato et al., 2001), where the authors state that assessing the type of mucosal crypt patterns can actually predict the histological findings to a very high accuracy. In particular, we employ three different variants of the discrete wavelet transform, namely the classical, maximally decimated 2-D Discrete Wavelet Transform, the 2-D Stationary Wavelet Transform (á-trous algorithm) and the 2-D Dual-Tree Complex Wavelet Transform, originally proposed by Kingsbury. We further integrate color information into the feature extraction process in order to show the improvements in classification accuracy compared to grayscale image analysis. This is accomplished by either simple concatenation of computed feature vectors from each color-channel or by using an extension of the classical cooccurrence matrices in the wavelet domain.

The paper is structured as follows: in Section 2 we will briefly illustrate the pit-pattern classification scheme from the medical point of view and provide an overview of related work in this research area. In Section 3, we recapitulate the background of the three wavelet transformations we use and outline their differences with regards to feature extraction for image analysis. Further, we discuss the computation of image features, the incorporation of color-channel information and introduce the classification/feature selection setting of our work. Finally, Section 4 presents empirical results and Section 5 concludes the paper with a summary of the main points and an outlook on further research.

2 Pit-Pattern Classification - The Medical Perspective

Polyps of the colon are a frequent finding and are usually divided into metaplastic, adenomatous and malignant. Since the resection of all polyps is rather time-consuming, it is imperative that those polyps which warrant resection can be distinguished. Furthermore, polypectomy[1] of metaplastic lesions is unnecessary and removal of invasive cancer may be hazardous. The classification scheme presented in (Kudo, 1994) divides the mucosal crypt patterns into five types (pit-patterns I-V, see Figure 1), which can be observed using a high-magnification endoscope.

(a) Class I	(b) Class II	(c) Class III-S
(d) Class III-L	(e) Class IV	(f) Class V
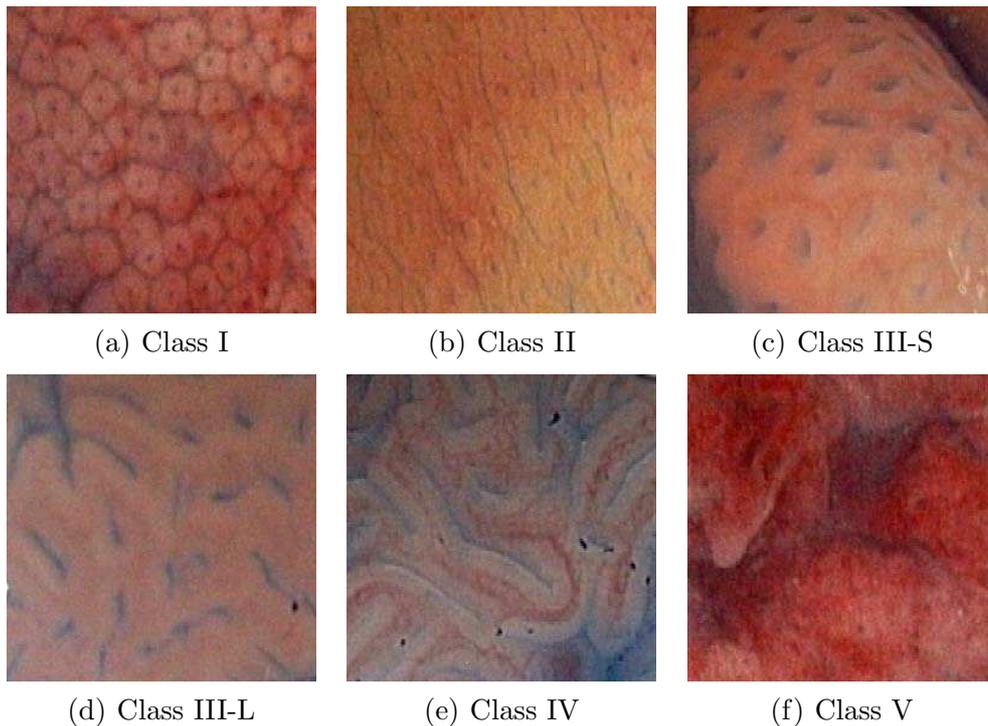
Fig. 1. Example images for the pit-pattern types I-V.

While types I and II are characteristic of benign lesions and represent normal colon mucosa or hyperplastic polyps (non-neoplastic lesions), types III to V represent neoplastic, adenomatous and carcinomatous structures. Our classification problem can be stated as follows: the problem to differentiate pit-patterns I and II from III-L,III-S,IV and V will be denoted as the *two-class* problem (neoplastic vs. non-neoplastic), whereas the more complex and detailed discrimination of all pit-patterns I to V will be denoted as the *six-class* problem. At first sight, the pit-pattern classification scheme seems to be straightforward and easy to be applied. Nevertheless, it needs some experience and exercising to achieve good results. Correct diagnosis very much relies on

---

[1] the process of removing polyps

the experience of the endoscopist as the interpretation of the pit-patterns may be challenging (Hurlstone, 2002).

Regarding the correlation between the mucosal pit-patterns and the histological findings, several studies reported good results, although with quite different diagnostic accuracies. A comparative study by (Kato et al., 2006) shows that the classification accuracy in magnifying colonoscopy ranges from 80.6% to 99.1%. Another extensive study by (Hurlstone et al., 2004) reports error rates of approximately 5%. In (Fu, 2004) the authors claim 95.6% for chromoendoscopy with magnification in contrast to diagnosis using conventional colonoscopy (84.0%) and diagnosis using chromoendoscopy without magnification (89.3%). In addition to that, inter-observer variability of magnification chromoendoscopy has been described at least for Barret's esophagus (Meining, 2004). This inter-observer variability may to lesser degree be also present in the interpretation of pit-patterns of colonic lesions. This work aims at allowing computer-assisted pit-pattern classification in order to enhance the quality of differential diagnosis.

In previous work, we have used several methods from the field of texture-classification to obtain image features with high discriminative power. In (Häfner et al., 2006a) several histogram-based techniques (e.g.: luminance histogram, color-channel histogram) are used to capture the characteristics of the pit-pattern types. A k-Nearest Neighbor (NN) classifier is then used to classify the images based on histogram intersection used as a distance function. The best classification accuracies in the two- (85.6%) and six-class (67.3%) problem are achieved using three-dimensional color histograms.

In (Häfner et al., 2006b) the authors propose to use wavelet-based texture descriptors for feature extraction. More precisely, the classical 2-D DWT and the 2-D wavelet packet decomposition with local discriminant basis (Saito and Coifman, 1994) are used. In the classification stage, the performance of Support Vector Machines (SVM) and k-NN classification is evaluated on each color-channel specific feature set individually. The best Leave-One-Out cross-validation (LOOCV) accuracy is achieved using features computed from the red color-channel of the RGB model with 75% in the two-class and 57% in the six-class case.

A completely different approach is presented in (Häfner et al., 2007a,b), where the authors compute a set of texture-descriptors from the outputs of the Discrete Cosine Transform (DCT) and the discrete Fourier Transform. Features are either computed from non-overlapping pixel blocks in the DCT domain or from adaptively sized rings in the Fourier domain. Concatenation of the feature vectors of each color-channel is then used to incorporate color-information. The authors employ a Bayes normal classifier (Fukunaga, 1990) together with feature subset selection to classify the endoscopy images. The best achieved

LOOCV accuracy in the two-class problem is 97.7% and 86.36% in the six-class problem.

In (Kwitt et al., 2008) the feature extraction step is based on the Gabor Wavelet Transform (Manjunath and Ma, 1996) and the Dual-Tree Complex Wavelet Transform (DT-CWT) (Kingsbury, 1998), which both provide approximate shift-invariance and a directionally-selective frequency partitioning. Color-information from the RGB color-channels is incorporated by feature vector concatenation or by using a parallel multi-classifier. Best reported LOOCV results are 96.28% (two-class) and 81.82% (six-class). However, no feature subset selection was conducted in this work.

Another approach that is completely based on grayscale-image analysis is presented in (Kwitt and Uhl, 2007), where the marginal distributions of wavelet detail subbands computed from the DT-CWT are modeled by two-parameter Weibull distributions. Maximum-likelihood estimates of the distribution parameters are then composed into feature vectors and are used for k-NN classification. Even though this approach is based on the luminance-channel only, the experiments show superior results when compared to the early approaches in (Häfner et al., 2006a,b).

In the work of (Karkanis et al., 2003), so called color wavelet covariance features are computed for a set of color models, including RGB and LAB. The image database of this work consists of endoscopic video frames extracted from video sequences, which were acquired during colonoscopy. The approach is based on computing classical cooccurrence matrices from the detail subbands of wavelet-decomposed color-channels. The feature vectors are then composed by a set of the commonly known Haralick features (Haralick, 1973). Dependencies between the color-channels are incorporated by calculating covariances between equal features of different color-channel subbands. However we have to point out, that the classification problem of (Karkanis et al., 2003) is restricted to the binary case and a conventional video-endoscope is used.

3 Feature Extraction and Classification

In this section we will discuss the different wavelet transforms we use for image representation and introduce the computation of the image features. Our aim is to emphasize the differences of the various wavelet-transform approaches and not to give a thorough treatment of the theoretical details. For further information on the theoretical aspects, take a look at the corresponding references. Further, we outline the incorporation of color-channel information and discuss our classification and feature selection setup.

The three wavelet transforms we discuss are the classical, maximally decimated Discrete Wavelet Transform (DWT) implemented by the Mallat algorithm (Mallat, 1989), the Stationary Wavelet Transform (SWT) implemented by the undecimated à-trous algorithm and the Dual-Tree Complex Wavelet Transform (DT-CWT). All transformations have rather different properties. Whether these properties are either advantages or disadvantages heavily depends on the field of application. Our focus here is on capturing important image information for classification purposes. Thus, our suggestion for the *best suitable* wavelet transform does not necessarily have to coincide with findings in the field of image compression, encryption or watermarking for example.

Probably the most commonly used wavelet transform for texture analysis is the classical 2-D DWT, implemented by the Mallat algorithm (Mallat, 1989), which roughly works as follows: in 1-D, a given input signal is subjected to convolution with two filter-sequences (high- and lowpass filters), followed by decimating the filter outputs by two (subsampling). This convolution-decimation procedure is recursively computed on the lowpass channel. In practice we will work with compactly supported wavelets, where the filters in this case are finite impulse response (FIR) filters. In the language of digital signal processing, this is equivalent to using a two-channel perfect reconstruction (PR) filterbank with Quadrature Mirror Filters (QMF) satisfying certain admissibility conditions, since not every PR filterbank is capable of implementing a DWT (Fliege, 1994). Since we will rely on commonly known wavelet filters (i.e. Daubechies filters) these conditions are of course satisfied in our case. The classical extension of the 1-D DWT to 2-D is straightforward by separate row- and column filtering of the input image, which leads to a decomposition into a set of detail images and one approximation image. The detail images are denoted as the horizontal (LH subband), vertical (HL subband) and diagonal (HH subband) detail subband, where HL signifies lowpass filtering the rows and highpass filtering the columns of an image for example.

Due to the fact that the implementation of the DWT involves downsampling the filter outputs by two, the DWT has zero-redundancy, which is a nice property for many applications (e.g.: compression). However, downsampling introduces aliasing effects, since high-pass content is folded back into low frequencies and leads to severe shift-dependency. This means that shifts of the same input signal might produce completely different coefficients after the transformation. Regarding image analysis purposes, this is definitely a great disadvantage. Just consider the fact, that the process of colonoscopy is subjected to varying physical conditions. We will thus have to deal with shifted images showing the same mucosal pit-patterns. Producing different coefficients in that case might lead to different image features and as a consequence to

wrong classification results.

A second disadvantage arises, when we take a closer look why the 2-D DWT can actually be implemented by separate row and column filtering. This results from the fact that we are using so called tensor-product wavelets, which are constructed by tensoring 1-D wavelet functions. However, since 1-D wavelets are real functions and thus have a two-sided frequency spectrum (positive and negative frequencies), tensoring leads to ambiguities when trying to distinguish image features oriented along $+45°$ or $-45°$ (HH subband). This problem is vividly illustrated in (Selesnick et al., 2005) and is commonly known as the lack of directional selectivity.

A first approach to remedy the problem of shift-dependence of the classical DWT is to use the Stationary Wavelet Transform (Shensa, 1992; Nason and Silverman, 1995) instead. This transform is implemented by the undecimated à-trous (with-holes) algorithm, whose name originates from the fact that it involves spreading the filter-sequences with zeros. The general idea is to completely avoid downsampling the filter outputs, which actually causes the shift-dependency problem of the DWT. Instead, the filters at each decomposition stage of the SWT are upsampled by two (dyadic upsampling). The convolution-decimation procedure now reduces to a simple convolution step, where a given input sequence is convolved with the filter coefficients. The extension to 2-D is again straightforward by separate row and column filtering. However, since the downsampling step is omitted, we end up with a redundancy of $3 \cdot J + 1$ times in 2-D for a decomposition depth of J.

The last wavelet transform transform variant we discuss is the Dual-Tree Complex Wavelet Transform, originally introduced in (Kingsbury, 1998). The DT-CWT is approximately shift-invariant and produces six oriented detail subbands in 2-D. All of these properties come with a very limited redundancy of 2^d in $d \in \mathbb{N}$ dimensions. The basic concept of the DT-CWT in 1-D is, to use two parallel wavelet-transform trees, where the wavelets of each tree form an approximate Hilbert transform pair. With regards to digital filters, the Hilbert-Transform property is equivalent to requiring that the lowpass filter of one tree is a half-sample delayed version of the lowpass filter of the other tree (half-sample delay condition). If this key property is satisfied, the outputs of each tree can be interpreted as the real and imaginary parts of complex wavelet coefficients. From the computational point of view the main advantage of the 2-D DT-CWT is, that it allows an efficient implementation based on four parallel 2-D DWTs. For more information on the theoretical background, see (Selesnick et al., 2005; Selesnick, 2001). The detail subbands containing the real- and imaginary parts of the complex wavelet coefficients are obtained by adding and subtracting the right detail subbands of the four trees. The 2-D DT-CWT leads to six complex detail subbands, which capture image details oriented along $\approx 15°$, $45°$, $75°$, $105°$, $135°$, $160°$. The approxi-

mate shift-invariance together with the the property of directional selectivity make the 2-D DT-CWT a good candidate for capturing texture-information. Remark: if not stated otherwise, the identifiers DWT, SWT and DT-CWT will denote the 2-D variants of the transformations.

3.2   Image Feature Computation

For feature extraction purposes, we consider only the wavelet detail subbands. In general, we cannot directly input the coefficients to a classifier, due to the curse of dimensionality (Bishop, 1995). We have to reduce the dimensionality by computing features from our wavelet detail subbands. Given that $N \times N$ denotes the dimensions of an arbitrary detail subband at decomposition depth m, we search for some mapping $\phi : \mathbb{R}^{N^2} \to \mathbb{R}^d$, with $d \ll N^2$. Each subband $\mathbf{D}_k^m$ is mapped to a feature vector in the d-dimensional feature space $\mathbb{R}^d$, $\mathbf{D}_k^m \mapsto \mathbf{v}_k^m := [v_1, \ldots, v_d]$, where k (i.e. $k \in \{HH, HL, LH\}$ for example) identifies the detail subband.

The features we use in this work are based on two statistics computed from the wavelet detail subbands. The coefficients of $\mathbf{D}_k^m$ will be denoted by $d_k^m(i, j), 1 \leq i, j \leq N$. Note, that in case of the SWT the dimensionality of the subbands is equal to dimensionality of the input image $\mathbf{C}^0$ (the 0 signifies that we work with the original). In case of the DWT and SWT we have $d_k^m(i, j) \in \mathbb{R}$, whereas in case of the DT-CWT we have $d_k^m(i, j) \in \mathbb{C}$, with $k = 1, \ldots, 6$. However, this is no problem since we will use the complex coefficient magnitudes in that case. For the DWT/SWT, the first feature we compute is called the *mean deviation* (Van de Wouwer et al., 1999) λ_k^m, given by

$$\lambda_k^m = \frac{1}{N^2} \sum_{1 \leq i,j \leq N} |d_k^m(i, j) - \overline{d_k^m}|. \tag{3.1}$$

For the DT-CWT we simply compute the mean μ_k^m of the absolute coefficient values, which is used in (Hatipoglu and Mitra, 1999) or (Manjunath and Ma, 1996) for instance. The second feature we employ is the *standard deviation* σ_k^m of the coefficients of a subband at position (m, k), given by

$$\sigma_k^m = \left(\frac{1}{N^2 - 1} \sum_{1 \leq i,j \leq N} \left(d_k^m(i, j) - \overline{d_k^m} \right)^2 \right)^{\frac{1}{2}}. \tag{3.2}$$

Again, in case of the DT-CWT, d_k^m in Eq. (3.2) is replaced by $|d_k^m|$. Based on these features, we can construct a feature vector $\mathbf{v}_i$ for each image $i \in$

$\{1, \ldots, L\}$ by simple concatenation of the subband features. For example, in case of the DWT or SWT we have

$$\mathbf{v}_i := [\lambda_{LH}^1, \lambda_{HL}^1, \lambda_{HH}^1, \sigma_{LH}^1, \sigma_{HL}^1, \sigma_{HH}^1, \ldots,$$
$$\lambda_{LH}^J, \lambda_{HL}^J, \lambda_{HH}^J, \sigma_{LH}^J, \sigma_{HL}^J, \sigma_{HH}^J]. \qquad (3.3)$$

Hence, a J-scale DWT/SWT leads to $J \cdot 6$ dimensional feature vectors, whereas a J-scale DT-CWT (using μ_k^m instead of λ_k^m) leads to $J \cdot 12$ dimensional feature vectors since the number of subbands is doubled. In the next section, we will see how to incorporate color information into the feature extraction process.

3.3 Incorporation of Color-Channel Information

We will introduce two combining methods in this section, which we call *combining by concatenation* and *combining by using cross-cooccurrence matrices*. Both approaches are different in the way color information is incorporated. The first approach is non-integrative, while the latter is integrative in the sense that color information is incorporated directly in the feature extraction step.

3.3.1 Feature Vector Concatenation

We have chosen the very simple way of feature vector concatenation of the color-channel feature vectors to include color information. Without loss of generality, we assume to work with RGB images for a moment and we introduce a superscript p to signify the image plane (i.e. $p \in \{R, G, B\}$ for example). Given a color-channel specific feature vector $\mathbf{v}_i^p$ of image i, we construct a final feature vector from

$$\mathbf{v}_i = [\mathbf{v}_i^R, \mathbf{v}_i^G, \mathbf{v}_i^B]. \qquad (3.4)$$

Since each color-channel is transformed separately without incorporating dependencies, this approach is called non-integrative. Although it is true that we might lose information by neglecting color-channel dependencies, we will see in Section 4 that concatenation is a quite good combining scheme for our problem. Another representative of non-integrative approaches is presented in (Kwitt et al., 2008), where a parallel multi-classifier is used to incorporate color information. However, the results presented there show that concatenation is superior to the parallel multi-classifier in almost all cases.

3.3.2 Color Wavelet Cross-Cooccurrence (CWCC) Matrices

The second combining scheme we propose does not fall into the category of non-integrative approaches. It is based on the extension of classical co-occurrence matrices to cross-cooccurrence matrices (Palm, 2004). Since the cross-cooccurrence matrices were originally defined in the spatial domain, we extend them to the wavelet-domain in order to work with our detail subband images. In contrast to the concatenation scheme, this approach leads to different feature vectors and does not rely on the computation of Eqs. (3.1) and (3.2).

Nevertheless, before we can explain how to compute cross-cooccurrence matrices between wavelet detail subbands, we have to recapitulate the definition of the classical cooccurrence matrix. The cooccurrence matrix $\mathbf{M}_{\mathbf{d}}^p(i,j)$ at position (i,j) of image plane p captures the joint occurrence of intensity values i and j separated by a given displacement vector $\mathbf{d} \in \mathbb{N}^2$. This displacement vector implicitly defines the orientation and the distance of considered pixel pairs (i.e. $\mathbf{d} = (0,1)^T$). Usually, the matrix $\mathbf{M}_{\mathbf{d}}^L$ is computed on the luminance (L) channel of the LUV color model.

The extension to vector images is straightforward. The aim is to capture the joint occurrence of intensity values between different image planes p and p'. This leads to the definition of a matrix $\mathbf{M}_{\mathbf{d}}^{p,p'}(i,j)$, which is equivalent to the formulation of $\mathbf{M}_{\mathbf{d}}^p$ with the only difference that the considered pixel pairs now reside in two different planes. This concept can be easily extended to work with detail subband images in the wavelet domain. We want to capture the joint occurrences of the wavelet detail coefficients of two different image planes (subbands). From that point, these kind of cross-cooccurrence matrices will be denoted as *Wavelet Cross-Cooccurrence (WCC)* matrices. In the most general setting, the matrix $\mathbf{M}_{\mathbf{d},m,k,k'}^{p,p'}(i,j)$ between two arbitrary detail subbands $\mathbf{D}_k^{m,p}$ and $\mathbf{D}_{k'}^{m,p'}$ at decomposition depth m is defined as

$$\mathbf{M}_{\mathbf{d},m,k,k'}^{p,p'}(i,j) = \mathbb{P}(d_k^{m,p}(\mathbf{x}) = i \wedge d_{k'}^{m,p'}(\mathbf{y}) = j | \mathbf{x} - \mathbf{y} = \mathbf{d}). \qquad (3.5)$$

Note, that we have used the abbreviation $\mathbf{x} = (x,y)$ to denote the position of the matrix elements. Eq. (3.5) is the general form of the WCC matrix, since it allows to consider arbitrary subband pairs at scale m. However, in this work we will only consider cases where $k = k'$, which means that we consider those subbands at equal positions in the decomposition structure but in different color channels. Hence we denote this kind of WCC matrices as *Color Wavelet Cross-Cooccurrence (CWCC)* matrices.

In order to make Eq. (3.5) computationally feasible, it is necessary that the coefficients are quantized to integer values $q \in \{1, \ldots, Q\}$, with Q denoting the

quantization factor. Additionally, we point out that we will be using a *zero-displacement* vector $\mathbf{d} = \mathbf{0} = (0,0)^T$, which is equivalent to the computation of a two-dimensional histogram (Palm, 2004). In case of classic cooccurrence matrices (defined in the spatial domain) a zero displacement vector for the computation of $\mathbf{M}_\mathbf{d}^p$ is equivalent to computing an intensity histogram. Using zero-displacement in our setting allows to capture information about the impact of the filtering procedure on different image planes at several scales.

Since the CWCC matrices contain too many elements (at least for reasonable values of Q) to work with them directly as input features, we have to compute a set of features first. We decided to use a selection of the well-known (Haralick, 1973), which were originally proposed in combination with the classical formulation of the cooccurrence matrix. To keep consistency with our notation of the CWCC matrices, we reformulate the four Haralick features we use as follows:

Contrast

$$F_1(\mathbf{D}_k^{m,p}, \mathbf{D}_k^{m,p'}) = -\sum_{i=0}^{Q-1}\sum_{j=0}^{Q-1}(i-j)^2 \mathbf{M}_{\mathbf{d},m,k}^{p,p'}(i,j) \tag{3.6}$$

Correlation

$$F_2(\mathbf{D}_k^{m,p}, \mathbf{D}_k^{m,p'}) = \frac{1}{\sigma_i \sigma_j}\sum_{i=0}^{Q-1}\sum_{j=0}^{Q-1}(i-\mu_i)(j-\mu_j)\mathbf{M}_{\mathbf{d},m,k}^{p,p'}(i,j) \tag{3.7}$$

Homogeneity

$$F_3(\mathbf{D}_k^{m,p}, \mathbf{D}_k^{m,p'}) = \sum_{i=0}^{Q-1}\sum_{j=0}^{Q-1}\frac{\mathbf{M}_{\mathbf{d},m,k}^{p,p'}(i,j)}{1+(i-j)^2} \tag{3.8}$$

Energy

$$F_4(\mathbf{D}_k^{m,p}, \mathbf{D}_k^{m,p'}) = \sum_{i=0}^{Q-1}\sum_{j=0}^{Q-1}\left(\mathbf{M}_{\mathbf{d},m,k}^{p,p'}(i,j)\right)^2 \tag{3.9}$$

In Eq. (3.7) μ_i and σ_i denote the horizontal mean and standard deviation of the quantized detail subbands, whereas μ_j and σ_j denote the vertical mean and standard deviation. Based on these features we can construct feature vectors for each image. However, we make the restriction that our feature vectors will be homogeneous in the sense that they do not contain a mixture of different Haralick features. To give an example, assume that we are working with RGB images and the DWT. The feature vector for image i using Haralick feature F_j is given by

$$\begin{aligned}\mathbf{v}_i = [&F_j(\mathbf{D}_{LH}^{1,R}, \mathbf{D}_{LH}^{1,G}), F_j(\mathbf{D}_{LH}^{1,B}, \mathbf{D}_{LH}^{1,G}), F_j(\mathbf{D}_{lh}^{1,R}, \mathbf{D}_{lh}^{1,B}),\\ &\dots, F_j(\mathbf{D}_{HH}^{J,R}, \mathbf{D}_{HH}^{J,G}), F_j(\mathbf{D}_{HH}^{J,B}, \mathbf{D}_{HH}^{J,G}), F_j(\mathbf{D}_{HH}^{J,R}, \mathbf{D}_{HH}^{J,B})].\end{aligned} \tag{3.10}$$

In case of a J-scale DWT/SWT this feature vector construction approach leads to $J \cdot 9$ dimensional vectors, whereas in case of the DT-CWT we obtain $J \cdot 18$ dimensional feature vectors. Note, that the final feature vector is still constructed by concatenation, but in contrast to the concatenation scheme of Section 3.3.1 each element of $\mathbf{v}_i$ now incorporates information between two image planes.

3.4 Classification and Feature Selection

In this section we discuss the classification and feature selection step of our proposed approach. The basic building elements are a k-NN classifier (Fukunaga, 1990) using the Euclidean formula as a distance function between feature vectors and the procedure of Forward Feature Selection (FFS) (Fukunaga, 1990) to select a subset of features.

Given a set of training examples from K classes $\{\mathbf{x}_i, y_i\} \subset \mathbb{R}^d \times \{1, \ldots, K\}$ with $i = 1, \ldots, L$, a new sample $\mathbf{x} \in \mathbb{R}^d$ is classified into class $c \in \{1, \ldots, K\}$ according to the majority of the labels among its k-nearest neighbors. Since the Euclidean distance is very sensitive to different ranges in the input features, we normalize each feature vector dimension first by using the standard approach of subtracting the mean and dividing by the standard deviation. For simplicity we decided to use a 1-NN classifier for all experiments. On the basis of different feature sets, other work (Häfner et al., 2007a,b; Kwitt and Uhl, 2007; Kwitt et al., 2008) has shown that 1-NN lead to very good results. Since the scope of this paper is on the feature computation side and not on testing classifiers, we limit our discussions to the 1-NN case.

Regarding the total number of computed features, it is questionable that all features contribute discriminative information. Hence, we additionally conduct a feature subset selection step. Since exhaustive search is computationally impossible w.r.t. the dimensionality of our feature space, we use forward feature selection (FFS) as the method of choice. This procedure searches for a (sub)optimal feature subset among all available features using a defined criterion of class separability. Lazy formulated, forward selection starts with the evaluation of all features separately and then successively adds one feature to the best subset in every iteration. As a criterion function for class separability it is perfectly obvious to use a 1-NN classifier together with LOOCV. The LOOCV accuracy (number of correctly classified samples divided by the total sample size) is then used as a quality measure for the feature subsets in every FFS iteration. Since we do not set an upper bound on the number of selected features, we store the LOOCV accuracy for the best subset of an iteration and finally select the very subset, which led to the best result. The number of selected features will be provided for all experiments (see Section 4). Using

this setup, we do not have to evaluate the final 1-NN classifier in a separate step anymore, since both the criterion function of the FFS procedure and the final classifier are the same.

4 Experimental Results

In this section, we present the experimental results of our work. Our image database consists of 484 images, acquired in 2005/2006 at the Department of Gastroenterology and Hepatology (Medical University of Vienna) using a zoom-endoscope (Olympus Evis Exera CF-Q160ZI/L) with a magnification factor of 150. All images were selected by the physician conducting the colonoscopy with the objective to provide images with equal lightning conditions and at approximately the same camera angle. To enhance the visual appearance, dye-spraying with indigo-carmine was applied and biopsies or mucosal resections were taken to obtain a histopathological diagnosis (*our ground truth*). For pit-pattern types I,II and V, biopsies were taken, since these types need not be removed. Lesions of pit-pattern types III-S/III-L and IV have been removed endoscopically. Table 1 lists the number of image samples per class.

I	II	III-L	III-S	IV	V
126	72	62	18	146	60

Table 1
Number of images per pit-pattern class (ground truth).

4.1 Experimental Setup

For all three wavelet transformations we vary the maximum decomposition depth from four to six and evaluate the impact of the decomposition depth on the LOOCV results. When using the DWT or SWT we further vary the wavelet filters to study the effect on the LOOCV rates. In particular, we vary between four Daubechies filters with two to eight taps (abbreviated by db2, db4, db6 and db8). Concerning the choice of wavelet filters for the DT-CWT, we use Kingsbury's (14,14)-tap Q-Shift filters (Kingsbury, 2001) for decomposition levels $j \geq 2$ and the near-symmetric (13,19)-tap filters for the first level. The technical reason for working different filter sets on the first level of the DT-CWT is explained in Selesnick et al. (2005).

In order to have a reference for evaluating the quality of the color-information combining schemes, we first conduct our experiments on the luminance (L) channel of the LUV color model, which is commonly used for grayscale-only

image processing. The feature vectors we compute for this channel are constructed according to Eq. (3.3), since no color-information is available. For the two combining schemes, we decided to evaluate whether different color models lead to different classification results as well. In particular, we compare the RGB to the 1976 CIE L*a*b* (1976) color model (abbreviated by LAB) (Gonzalez and Woods, 2002; Lukac and Plataniotis, 2006). In the RGB model we have one component for each primary color, whereas the LAB color model splits color information into the three components lightness (L) and color-information on a red/green (A component) and yellow-blue (B component) axis. Another interesting difference between RGB and LAB is, that distances (in the Euclidean sense) measured in the LAB coordinate system approximately match distances perceived by the human visual system.

Since we will often obtain very similar LOOCV accuracies for different feature sets, we use McNemar's test (Everitt, 1992) to check if the LOOCV differences obtained on different feature sets are statistically significant at a given significance level α. However, our setup for the McNemar test is slightly different to the standard one (Dieterich, 1998; Zamolotskikh et al., 2006), since we do not want to check for significant differences between two classifiers, but between two feature sets $\mathcal{S}_1, \mathcal{S}_2$. Now, let n_{10} denote the number of correctly classified samples during LOOCV from feature set $\mathcal{S}_1$, which are in turn falsely classified when using $\mathcal{S}_2$. Accordingly, let n_{01} denote the number of samples, correctly classified by using $\mathcal{S}_2$ but wrongly classified on the basis of $\mathcal{S}_1$. Then, the McNemar test statistic T, which is approximately χ_1^2 distributed with 1 degree of freedom is defined as

$$T := \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}} \tag{4.1}$$

The null-hypothesis H_0 of the McNemar test is that both feature sets lead to equal error rates. At a given significance level (we use $\alpha = 0.05$), we can reject the null-hypothesis if our test statistic (4.1) is greater than $\chi_{1,1-\alpha}^2$.

Regarding the notation of the test results, we follow the convention that a '+' sign denotes a significant difference ($T > \chi_{1,1-\alpha}^2$), whereas a '−' sign denotes no significant difference. Since we have two classification problems, almost all McNemar test results will be of the form '±±', where the first ± sign signifies the McNemar test result for the two-class problem and the second ± sign represents the result for the six-class problem.

14

First of all, we evaluate the impact of the wavelet filters on the LOOCV accuracies for DWT and the SWT case. It turns out, that for all possible $\binom{4}{2}$ filter-pairings we cannot reject the the null-hypothesis of equal error rates. This means that the differences in the results obtained on the filter-specific feature sets for each transform and decomposition depth are not statistically significant. Hence, we will omit the filter-specific results in the following explanations and provide per-scale results only for each transformation. These results are the maximum LOOCV accuracies over all wavelet filters. Table 2 displays the LOOCV accuracies for varying decomposition depths. In addition, the numbers in braces show the number of total features selected by FFS procedure.

Problem	**Scales**	DWT	SWT	DT-CWT
2-class	4	84.92 (19)	89.05 (10)	93.60 (41)
	5	84.92 (16)	89.46 (24)	94.63 (21)
	6	84.50 (22)	89.46 (14)	**94.83 (24)**
6-class	4	70.45 (19)	76.65 (10)	81.40 (27)
	5	71.69 (16)	77.69 (14)	83.47 (29)
	6	71.90 (16)	77.69 (29)	**83.88 (65)**

Table 2
LOOCV accuracies (%) for the luminance-channel together with the number of selected features in braces. The best results for both classification problems are marked bold.

Since the intra-transform results for different decomposition depths are very similar (especially in the 2-class problem), we use the McNemar test to check for significant differences. As expected, there is no evidence against the null-hypothesis for almost all $\binom{3}{2}$ considered scale combinations in the two-class case. We observe almost the same situation in the the six-class case, however with the exception that the differences between scales 4 and 6 are significant in some cases.

The next question is, whether we can make a general statement about the impact of the wavelet transformations on the LOOCV rates? From the output of the McNemar test for the inter-transform accuracies (see Table 3) we see, that at least all the differences are statistically significant for every decomposition depth. A look at Figure 2 reveals that the DT-CWT outperforms both other transformations. We can see an almost linear increase in LOOCV accuracy as we switch from DWT to SWT and DT-CWT. This allows the conclusion that the stepwise addition of the properties *shift-invariance* (added by the SWT) and *directional selectivity* (added by the DT-CWT) leads to a direct

Scales	DWT vs. SWT	DWT vs. DT-CWT	SWT vs. DT-DWT
4	++	++	++
5	++	++	++
6	++	++	++

Table 3

McNemar test results for the (L)uminance channel feature vectors.

increase in LOOCV accuracy for both classification problems. We summarize, that for grayscale-only image analysis it is obvious that the DT-CWT is the best choice with regards to our classification problem and setup.

(a) 2-class, Grayscale   (b) 2-class, Grayscale
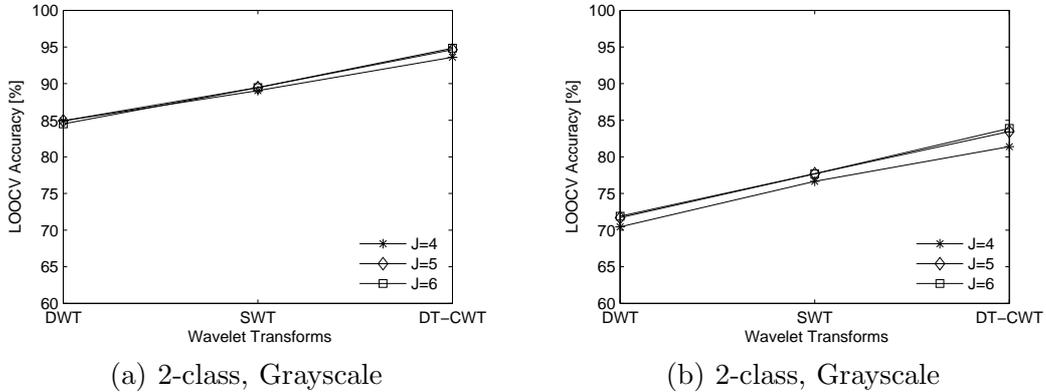
Fig. 2. Visualization of differences in the LOOCV accuracies for the three wavelet transforms over three different decomposition scales.

Concerning the number of selected features, we see that although the search space of the FFS procedure in case of the DT-CWT features is two-times larger than the search space in case of the SWT/DWT features, only 24 are actually selected for the two-class problem. In the six-class problem however, 65 features are selected at six scales for the DT-CWT. This might lead to overtraining issues. In this particular case we would choose in favour of five decomposition scales, since the second-best rate of 83.47% is only slightly worse and the number of features reduces to 29.

The luminance-channel results of this section will serve as reference results for following sections, where we evaluate the quality of the color-information combining schemes. The presented results clearly show that choosing the right image representation for feature computation in grayscale-only image-analysis can lead to an improvement in classification accuracy by $\approx 10\%$.

4.3 Feature Vector Concatenation Results

First, we again check whether the wavelet filters have a significant impact on the LOCCV results. In accordance with the outputs of the McNemar test for the luminance-channel, there is no evidence against the null-hypothesis of equal error rates. However, concerning the impact of the decomposition depth, some differences between depth 4 and 6 are again significant. Table 4 shows the LOOCV results for both the color models and Table 5 provides the number of selected features from FFS.

		DWT		SWT		DT-CWT	
Problem	**Scales**	RGB	LAB	RGB	LAB	RGB	LAB
	4	95.66	96.07	97.52	98.55	98.55	97.52
2-cls.	5	96.28	96.07	96.90	97.93	99.38	97.93
	6	96.69	96.49	98.14	98.55	**99.38**	**99.38**
	4	84.92	88.02	89.05	89.67	90.08	89.26
6-cls.	5	85.74	87.81	88.64	90.50	91.74	91.94
	6	85.33	87.40	89.46	90.70	**93.18**	91.94

Table 4
LOOCV accuracies of the feature vector concatenation approach for the RGB and LAB color model. The best results for both classification problems are marked bold.

		DWT		SWT		DT-CWT	
Problem	**Scales**	RGB	LAB	RGB	LAB	RGB	LAB
	4	30	25	19	31	22	44
2-cls.	5	45	25	42	26	48	17
	6	36	29	25	31	**48**	70
	4	45	26	31	25	62	47
6-cls.	5	47	45	26	44	84	47
	6	60	28	36	41	**61**	71

Table 5
Number of selected features for the feature vector concatenation approach.

From the LOOCV results of Table 4 we notice that, in contrast to Table 2, the inter-transform results become more similar, especially in the two-class problem. However, a minimum LOOCV accuracy of $\approx 96\%$ lets very little room for any transformation to produce significantly better rates. The similarity of the results is reflected in the outcomes of the inter-transform McNemar tests (see Table 6) as well. As the LOOCV accuracies are lifted to a higher level, we

have no evidence against equal error rates for the majority of the transform combinations any longer. The only consistently significant differences for the two-class problem remain between the DWT and DT-CWT results for RGB.

	DWT vs. SWT		DWT vs. DT-CWT		SWT vs. DT-DWT	
Scales	RGB	LAB	RGB	LAB	RGB	LAB
4	−+	+−	++	−−	−−	−−
5	−−	−−	++	−+	++	−−
6	−+	++	++	++	−+	+−

Table 6

Inter-Transform McNemar test results for the feature vector concatenation approach.

The situation is slightly different when considering the six-class problem. When using the highest decomposition depth, most of the differences are significant. We infer, that the additional properties of the SWT and DT-CWT become particularly important for the more complex classification task. The fact that the results move together in general is evident when comparing Figure 2 to Figure 3 as well. Although the trend towards higher LOOCV accuracies persists as we switch from DWT to SWT and DT-CWT, the slope of the graph is less steep than the slope we observed for the luminance-channel results.

By taking a closer look at Table 5, we observe a situation similar to the luminance-channel results. Although the feature space dimensionality is quite high (three times the dimensionality of the luminance-channel feature vectors), only small feature subsets produce the best LOOCV rates. Nevertheless, the good DT-CWT results can be attributed to the fact that the starting set of available features for FFS is bigger than the starting set of the DWT/SWT features and apparently contains more discriminative information.

A question we have not answered at this point is, whether one of the color models leads to better/worse results than the other. From Table 4 we see that the LOOCV rates are actually very similar. Therefore, we conduct an additional McNemar test on the color model specific results. The test outputs given in Table 7 confirm the assumption that neither RGB nor LAB lead to significantly different results, no matter which decomposition depth we choose. We conclude, that the feature-vector concatenation approach is robust against changes in the color model with regards to our experimental setup.

| (a) 2-class, RGB | (b) 2-class, LAB |

| (c) 6-class, RGB | (d) 6-class, LAB |
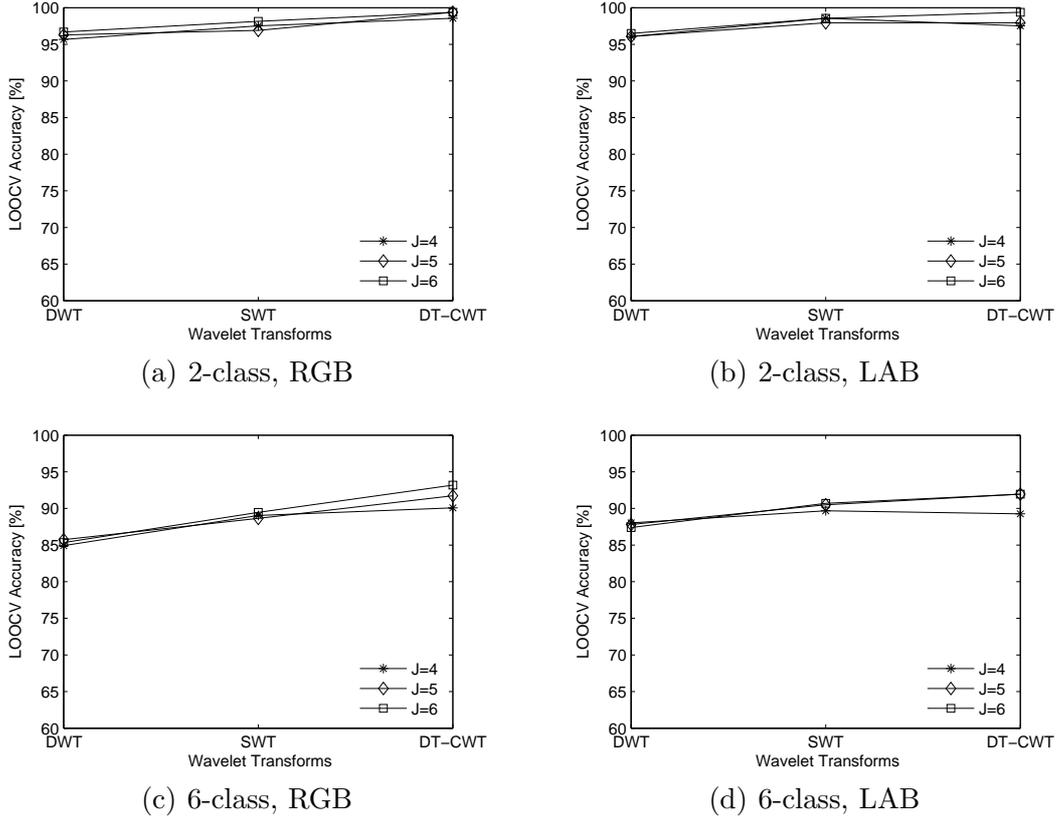
Fig. 3. Visualization the LOOCV accuracies for the feature vector concatenation combining scheme.

Scales	DWT	SWT	DT-CWT
4	––	––	––
5	––	––	––
6	––	––	––

Table 7
McNemar test results between RGB and LAB for the concatenation scheme.

4.4 Color Wavelet Cross-Cooccurrence (CWCC) Results

For the CWCC matrix combining scheme, we fixed the the quantization factor to $Q = 256$. All results presented in this section were obtained by using the *Correlation* feature from Eq. (3.7), which outperformed all other Haralick features in terms of LOOCV accuracy throughout all experiments. Due to space limitations, we omit the other, feature-specific results here.

Again, the filter-specific McNemar tests showed no statistical evidence that a particular filter leads to higher LOOCV accuracies, no matter which classification problem we consider. Furthermore, the outputs of the decomposition

depth specific McNemar tests perfectly fit among the results of the previous sections, since the only significant differences occur between depths 4 and 6. Table 8 shows the LOOCV accuracies for the CWCC matrix approach.

Problem	Scales	DWT		SWT		DT-CWT	
		RGB	LAB	RGB	LAB	RGB	LAB
2-cls.	4	93.18	95.25	92.98	96.28	96.90	95.66
	5	94.63	95.87	94.21	97.31	97.11	96.90
	6	94.63	96.49	95.87	**97.52**	97.11	96.90
6-cls.	4	79.96	83.88	79.34	85.74	83.06	83.05
	5	80.58	85.54	84.50	88.02	84.50	85.33
	6	81.82	86.57	85.95	**88.84**	85.95	85.74

Table 8
Maximum LOOCV accuracies of the CWCC matrix approach for the RGB and LAB color model.

Problem	Scales	DWT		SWT		DT-CWT	
		RGB	LAB	RGB	LAB	RGB	LAB
2-cls.	4	19	24	20	22	54	45
	5	24	36	18	24	41	33
	6	24	30	36	**25**	41	33
6-cls.	4	18	19	15	19	42	38
	5	34	27	21	24	49	65
	6	38	21	23	**29**	67	76

Table 9
Number of selected features for the CWCC matrix approach.

Analog to the concatenation scheme results of the last section, we observe very similar inter-transform accuracies here. For both color models, the results in Table 10 confirm that there is almost no significant difference between the three transforms using RGB and no difference at all using LAB.

A visual examination of the classification results in Figure 4 shows a slight trend towards higher LOOCV accuracies as we switch from DWT to SWT or DT-CWT in the two-class problem. However, the SWT features now produce the best LOOCV rates using the LAB color space. The slopes of the curves are again less steep compared to luminance-channel results. Similar to the concatenation approach, the CWCC matrix scheme somehow compensates the lack of shift-invariance of the DWT and the lack of directional-selectivity of the SWT. Furthermore, the LOOCV rates drop slightly in the six-class

	DWT vs. SWT		DWT vs. DT-CWT		SWT vs. DT-DWT	
Scales	RGB	LAB	RGB	LAB	RGB	LAB
4	$--$	$--$	$+-$	$--$	$+-$	$--$
5	$-+$	$--$	$-+$	$--$	$+-$	$--$
6	$-+$	$--$	$-+$	$--$	$--$	$--$

Table 10
McNemar test results for the CWCC matrices for RGB and LAB.

problem using the DT-CWT features. This is interesting, since the number of features selected from the SWT feature set is lower than the number of features selected by the DT-CWT feature set (see Table 9).

(a) 2-class, RGB

(b) 2-class, LAB

(c) 6-class, RGB

(d) 6-class, LAB
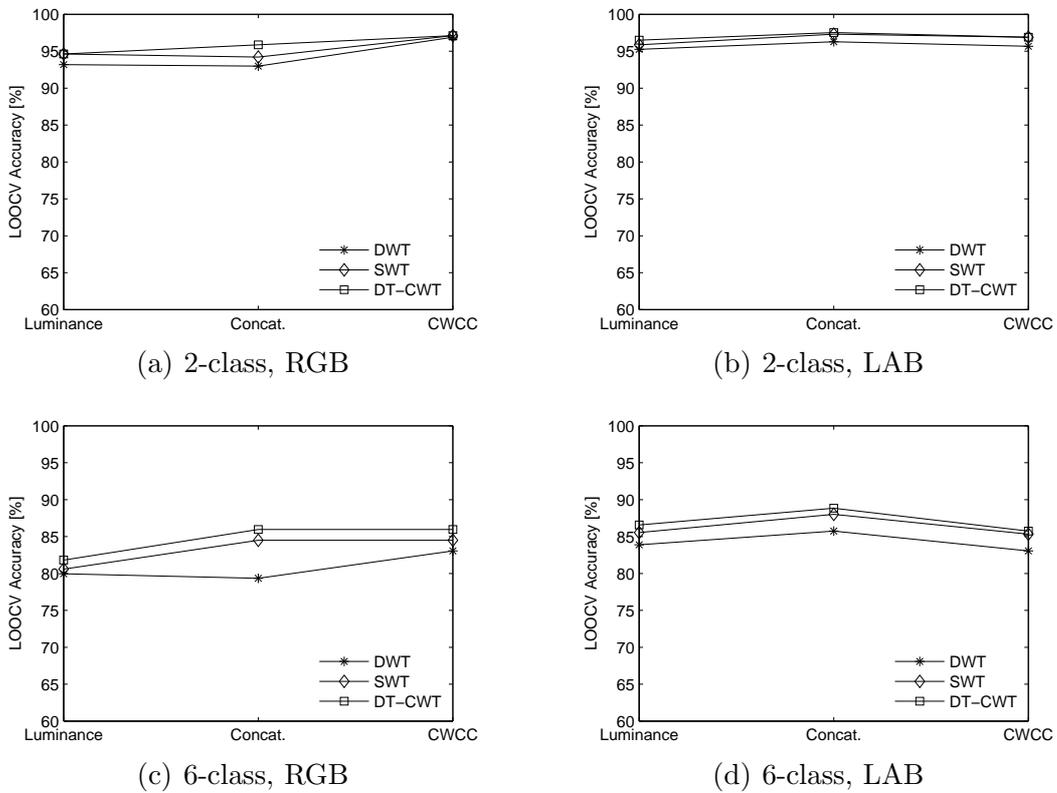
Fig. 4. Visualization of differences in the LOOCV accuracies for the three wavelet transforms over three different decomposition scales for the CWCC matrix combining scheme.

The last question we try to answer is, whether the choice of color model effects the classification results. Table 11 presents the outputs of the color model specific McNemar test, which indicate almost no evidence against equal error rates for the two-class problem, no matter which transformation or decomposition depth we choose. The only exception occurs at decomposition depth four using the SWT features. In the six-class problem, we observe slightly more significant differences in case of the DWT and SWT features. In case

of the DT-CWT features, the situation remains the same as in the two-class case. However, the results are different to the concatenation approach, where we did not observe any effect at all.

Scales	DWT	SWT	DT-CWT
4	$--$	$++$	$--$
5	$-+$	$--$	$--$
6	$-+$	$--$	$--$

Table 11
McNemar test results for the CWCC matrices for RGB and LAB

Taking into account the McNemar test results of Table 11, a comparison of the LOOCV rates for both color models allows the conclusion that LAB is presumably the better choice for the CWCC matrix approach. Furthermore, by referring to Table 10 we emphasize, that the choice of wavelet transform and decomposition depth does not seem to affect the classification rates when using LAB.

4.5 Comparing the Combining Schemes to the Grayscale Results

In this last section we make two special comparisons: First, we compare the results of the two combining schemes to the luminance-channel results of Section 4.2. Second, we compare the combining schemes to each other.

Since all of the maximum LOOCV accuracies were obtained on the highest decomposition depth ($J = 6$), we primarily focus on the results of this level to keep a certain degree of clarity. In Table 12 we provide the LOOCV accuracies for both the RGB and LAB color model. In the context of Table 12 the '$\pm$' signs next to the results of the combining schemes indicate, whether there is a significant difference to the luminance-channel results or not.

As we can see, we have clear evidence against the null-hypothesis of equal error rates in all cases. From the plots in Figure 5 we observe that the combining schemes outperform the luminance-channel results by $\approx 5 - 10\%$ in the two-class case and $\approx 2 - 12\%$ in the six-class case. Hence, we summarize that, independet of the underlying color model, we obtain an improvement in LOOCV accuracy by using the proposed combining schemes.

On the basis of Figure 5 we would further suggest that feature vector concatenation is superior to the CWCC matrix approach. However, since visual inspections can be elusive, we conduct another McNemar test. Table 13 shows the corresponding outputs for both color models. The '$\pm$' signs indicate, whether there is evidence against equal error rates when comparing the concatenation

RGB						
Problem	**Transform**	Lum.	Concat.		CWCC	
2-cls.	DWT	84.50	96.69	+	94.63	+
	SWT	89.46	98.14	+	95.87	+
	DT-CWT	94.83	**99.38**	+	97.11	+
6-cls.	DWT	71.90	85.33	+	81.82	+
	SWT	77.69	89.46	+	85.95	+
	DT-CWT	83.88	**93.18**	+	85.95	+
LAB						
Problem	**Transform**	Lum.	Concat.		CWCC	
2-cls.	DWT	84.50	96.49	+	96.49	+
	SWT	89.46	98.55	+	97.52	+
	DT-CWT	94.83	**99.38**	+	96.90	+
6-cls.	DWT	71.90	87.40	+	86.57	+
	SWT	77.69	90.70	+	88.84	+
	DT-CWT	83.88	**91.94**	+	85.74	+

Table 12

Comparison between the luminance-channel results and the two combining schemes for LAB and RGB.

to the CWCC matrix results.

	DWT		SWT		DT-CWT	
Scales	RGB	LAB	RGB	LAB	RGB	LAB
4	−+	−+	++	++	++	−+
5	−+	−−	−+	−−	++	−+
6	++	−−	−+	−−	++	++

Table 13

McNemar test results for the CWCC matrices in comparision to feature vector concatenation for RGB and LAB.

In the six-class problem the results are consistently different when using the RGB color space. From Figure 5 we can thus safely conclude, that concatenation is superior to the CWCC matrices in that particular case. However, this statement is not true in general for LAB, where only the DT-CWT results are significantly different over all decomposition depths. In the two-class problem we are again in the difficult situation that there is no clear pattern in the results, which would allow to make a general statement.

(a) 2-class, RGB   (b) 2-class, LAB

(c) 6-class, RGB   (d) 6-class, LAB
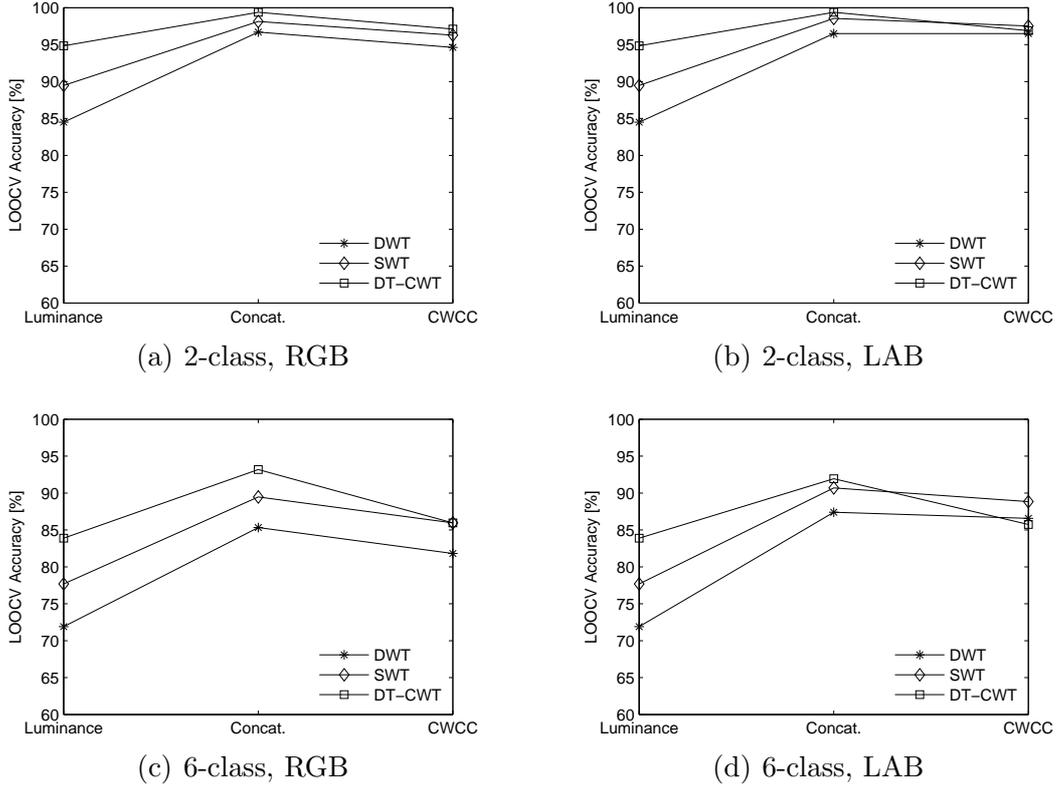
Fig. 5. Visualization of differences in the LOOCV accuracies between the luminance-channel results and the results of both combining schemes for $J = 6$.

5 Conclusion

In this article we focused on computer-assisted pit-pattern classification to support the dignity assessment of colonic polyps. For that purpose, we studied the applicability of texture-analysis methods in different wavelet domains. Our results show that texture-descriptors work pretty well for this medical image classification problem, although the ground truth of our work was obtained by an histophatological analysis. We have particularly addressed to main issues: First, we wanted to know whether different choices of the wavelet transformation for feature computation affect the quality of the classification results. Second, we have tried to incorporate color-channel information into the feature extraction process in order to improve classification accuracy.

The first results based on the luminance-channel of the LUV color model indicate that the lack of shift-invariance and directional selectivity make the classical DWT less suitable for our classification task. We have argued that, due to the varying physical conditions during colonoscopy, these properties are of particular importance. We propose to use the DT-CWT for luminance-channel only image analysis as the wavelet transformation of choice, since we obtained the highest LOOCV accuracies for both classification problems.

In the luminance-channel experiments, we further see two patterns, which persists throughout all other experiments: First, the choice of wavelet filter does not significantly affect the classification results, at least for the selection of filters we presented. Second, in contrast to our expectations, the decomposition depth only slightly affects the LOOCV rates. Significant differences only occur between scales four and six. This can be attributed to an increase in search space dimensionality. However, the number of features selected by FFS does not necessarily increase in proportion to the increase in search space dimensionality.

Regarding the quality of the color-information combining approaches, our experimental results document that both presented schemes are superior to luminance-channel only feature extraction. Especially in the six-class problem we achieve a considerable increase of $\approx 12\%$ in LOOCV accuracy by incorporating color-information. Regarding the differences between the two combining schemes, we summarize that the concatenation scheme performed best for the six-class problem in combination with the DT-CWT. Although the difference between the results of the DWT and DT-CWT are much smaller compared to the luminance-channel experiments, the pattern of an increase in LOOCV accuracy remains. Furthermore, we have seen that the choice of color model does not have an impact on the the concatenation scheme at all with respect to the the presented images features.

Unfortunately, we cannot make such a general statement for the two-class problem. This is due to the already mentioned effect, that the additional color-information somehow compensates the inter-transform differences. As a result, we do not observe any evidence against the null-hypothesis of equal error rates any longer. However, this might also be a side-effect of the already very high LOOCV accuracies in the two-class case. Additionally, more color channel information leads to more features and hence we obtain a higher-dimensional search space for FFS.

Further research on this topic will include the evaluation of other color models and the use of different features for the wavelet detail subbands. Although, the proposed subband features show quite good performance, modeling the marginal distributions of the subband coefficients seems to be promising as well (Kwitt and Uhl, 2007). Furthermore, other classifiers and feature subset selection algorithms might lead to even better results, since we have limited our discussions to the simple 1-NN classifier.

Acknowledgments

References

Bishop, C., 1995. Neural Networks for Pattern Recognition. Oxford University Press.

Dietterich, T., Oct. 1998. Approximate statistical tests for comparing supervised classification learning algorithms. Neural Computation 10 (7), 1895–1923.

Everitt, B., 1992. The Analysis of Contingency Tables, 2nd Edition. Vol. 45 of Monographs on Statistics and Applied Probability. Chapman and Hall/CRC.

Fliege, N., 1994. Multirate Digital Signal Processing. Wiley.

Fu, K., Dec. 2004. Chromoendoscopy using Indigo Carmine Dye Spraying with Magnifying Observation is the most Reliable Method for Differential Diagnosis between Non-Neoplastic and Neoplastic Colorectal Lesions. Endoscopy 36 (12), 1089–1093.

Fukunaga, K., 1990. Introduction to Statistical Pattern Recognition. Morgan Kaufmann.

Gonzalez, R., Woods, R., 2002. Digital Image Processing, 2nd Edition. Prentice-Hall.

Häfner, M., Brunauer, L., Payer, H., Resch, R., Wrba, F., Gangl, A., Vécsei, A., Uhl, A., Jun. 2007a. Pit Pattern Classification of Zoom-Endoscopic Colon Images using DCT and FFT. In: Kokol, P., Podgorelec, V., Micetic-Turk, D., Zorman, M., Verlic, M. (Eds.), Proceedings of the 20th IEEE International Symposium on Computer-Based Medical Systems (CBMS 2007). IEEE Computer Society CPS, Maribor, Slovenia, pp. 159 – 164.

Häfner, M., Brunauer, L., Payer, H., Resch, R., Wrba, F., Gangl, A., Vécsei, A., Uhl, A., Aug. 2007b. Pit pattern classification of zoom-endoscopical colon images using evolved Fourier feature vectors. In: Diamantaras, K., Adali, T., Pitas, I., Larsen, J., Papadimitriou, T., Douglas, S. (Eds.), Proceedings of the 2007 IEEE Machine Learning for Signal Processing Workshop (MLSP'07). IEEE, Thessaloniki, Greece, pp. 99 – 104.

Häfner, M., Kendlbacher, C., Mann, W., Taferl, W., Wrba, F., Gangl, A., Vecsei, A., Uhl, A., Jun. 2006a. Pit Pattern Classification of Zoom-Endoscopic Colon Images Using Histogram Techniques. In: Proceedings of the 7th Nordic Signal Processing Symposium (NORSIG'06). Reykjavik, Iceland, pp. 58–61.

Häfner, M., Liedlgruber, M., Wrba, F., Gangl, A., Vecsei, A., Uhl, A., Jul. 2006b. Pit Pattern Classification of Zoom-Endoscopic Colon Images Using Wavelet Texture Features. In: Proceedings of the 3rd International Conference on Advances in Medical Signal and Image Processing (MEDSIP'06). Glasgow, Scotland.

Haralick, R., Nov. 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Men and Cybernetics 3 (6), 610–621.

Hatipoglu, S., Mitra, S. Kingsbury, N., Aug. 1999. Texture Classification using Dual-Tree Complex Wavelet Transform. In: Proceedings of the 7th International Conference on Image Processing and Its Applications. Brisbane, Australia, pp. 344–347.

Hurlstone, D., Apr. 2002. High-Resolution Magnification Chromoendoscopy: Common Problems Encountered in Pit-Pattern Interpretation and Correct Classification of Flat Colorectal Lesions. American Journal of Gastroenterology 97 (4), 1069–1070.

Hurlstone, D., Cross, S., Adam, I., Shorthouse, A., Brown, S., Sanders, D., Lobo, A., 2004. Efficiacy of High Magnification Chromoscopic Colonoscopy for the Diagnosis of Neoplasia in Flat and Depressed Lesions of the Colorectum: a Prospective Analysis. Gut 53, 284–290.

Karkanis, S., Iakovids, D., Maroulis, D., Sep. 2003. Computer-Aided Tumor Detection in Endoscopic Video Using Color Wavelet Features. IEEE Transactions on Information Technology in Biomedicine 7 (3), 141–152.

Kato, S., Fu, K., Sano, Y., Fujii, T., Saito, Y., Matsuda, T., Koba, I., Yoshida, S., Fujimori, T., Mar. 2006. Magnifying Colonoscopy as a Non-Biopsy Technique for Differential Diagnosis of Non-Neoplastic and Neoplastic Lesions. World of Gastroenterology 12 (9), 1416–1420.

Kato, S., Fujii, T., Koba, I., Sano, Y., Fu, K. I., Parra-Blanco, A., Tajiri, H., Yoshida, S., Rembacken, B., Mar. 2001. Assessment of Colorectal Lesions Using Magnifying Colonoscopy and Mucosal Dye Spraying: Can Significant Lesions Be Distinguished? Endoscopy 33 (3), 306–311.

Kingsbury, N., Aug. 1998. The Dual-Tree Complex Wavelet Transform: A new Technique for Shift-Invariance and Directional Filters. In: Proceedings of the 8th IEEE DSP Workshop. Bryce Canyon, Utah, USA, pp. 9–12.

Kingsbury, N., May 2001. Complex Wavelets for Shift-Invariant Analysis and Filtering of Signals. Journal of Applied Computational Harmonic Analysis 10 (3), 234–253.

Kudo, S., Oct. 1994. Colorectal Tumours and Pit Pattern. Journal of Clinical Pathology 47 (10), 880–885.

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.

Kwitt, R., Uhl, A., Häfner, M., Wrba, F., Gangl, A., Vécsei, A., 2008. Multi-Directional Multi-Resolution Transforms for Zoom-Endoscopy Image Clas-

sification. Vol. 45 of Advances in Soft Computing. Springer.

Lukac, R., Plataniotis, K., 2006. Color Image Processing - Methods and Applications. CRC Press.

Mallat, S., Jul. 1989. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (7), 674–693.

Manjunath, B., Ma, W., Aug. 1996. Texture Features for Browsing and Retrieval of Image Data. IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (8), 837–842.

Meining, A., Feb. 2004. Inter- and Intra-Observer Variability of Magnification Chromoendoscopy for Detecting Specialized Intestinal Metaplasia at the Gastroesophageal Junction. Endoscopy 36 (2), 160–164.

Nason, G., Silverman, B., 1995. The Stationary Wavelet Transform and some Statistical Applications. Lecture Notes in Statistics 103, 281–300.

Palm, C., May 2004. Color Texture Classification by Integrative Cooccurrence Matrices. Pattern Recognition 37 (5), 965–976.

Saito, N., Coifman, R., 1994. Local discriminant bases. Mathematical Imaging: Wavelet Applications in Signal and Image Processing II 2303, 2–14.

Selesnick, I., Jun. 2001. The Design of Hilbert Transform Pairs of Wavelet Bases. Signal Processing Letters 8 (6), 170–173.

Selesnick, I., Baraniuk, R., Kingsbury, N., 2005. The Dual-Tree Complex Wavelet Transform. IEEE Signal Processing Magazine 22 (6), 123–151.

Shensa, M., Oct. 1992. The Discrete Wavelet Transform: Wedding the Á Trous and Mallat Algorithms. IEEE Transactions on Signal Processing 40 (10), 2464–2482.

Van de Wouwer, G., Scheunders, P., Van Dyck, D., Apr. 1999. Statistical texture characterization from discrete wavelet representations. IEEE Transactions on Image Processing 8 (4), 592–598.

Zamolotskikh, A., Delany, S., Cunningham, P., Apr. 2006. A Methodology for Comparing Classifiers that allow the Control of Bias. In: Proceedings of the 2006 ACM symposium on Applied computing (SAC'06). ACM, New York, NY, USA, pp. 582–587.