Multi-Directional Multi-Resolution Transforms for Zoom-Endoscopy Image Classification

Roland Kwitt and Andreas Uhl

Department of Computer Sciences[†]
University of Salzburg
rkwitt,uhl@cosy.sbg.ac.at

Summary. In this paper, we evaluate the discriminative power of image features, extracted from subbands of the Gabor Wavelet Transform and the Dual-Tree Complex Wavelet Transform for the classification of zoom-endoscopy images. Further, we incorporate color channel information into the classification process and show, that this leads to superior classification results, compared to luminance-channel based image processing.

1 Introduction

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 et al. [9], the macroscopic appearance of colorectal polyps is systematically described and results in the so called *pit-pattern* classification scheme.

In this work, we evaluate the discriminative power of features extracted from subbands of the Gabor Wavelet Transform and Kingsbury's Dual-Tree Complex Wavelet Transform, to allow computer-assisted pit-pattern classification. Existing approaches in this research area include histogram- and 2-D DWT-based methods for pit-pattern classification [2, 3] or classic video-endoscopy image classification by statistical second-order measures [10, 5].

The paper is structured as follows: in Section 2 we briefly illustrate the pit-pattern classification scheme. Section 3 outlines the proposed feature extraction step and the classification process. Finally, Section 4 presents our experimental results and Section 5 concludes the paper with an summary of our work.

2 Pit-Pattern Classification

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[3] of metaplastic lesions is unnecessary and removal of invasive cancer may be hazardous. The classification scheme of Kudo et al. divides the mucosal crypt patterns into five groups (pit-patterns I-V, see Figure 1). While types I and II are characteristic of benign lesions and represent normal colon mucosa or hyperplastic polyps, types III-V represent neoplastic, adenomatous and carcinomatous structures. Our classification problem can be stated as follows: the problem to differentiate pit-types I and II from III-V will be denoted as the *two-class* problem, whereas the more detailed discrimination of pit-types I-V will be denoted as the *six-class* problem.

(a) Pit-Types   (b) Pit-Types I-V (left to right)
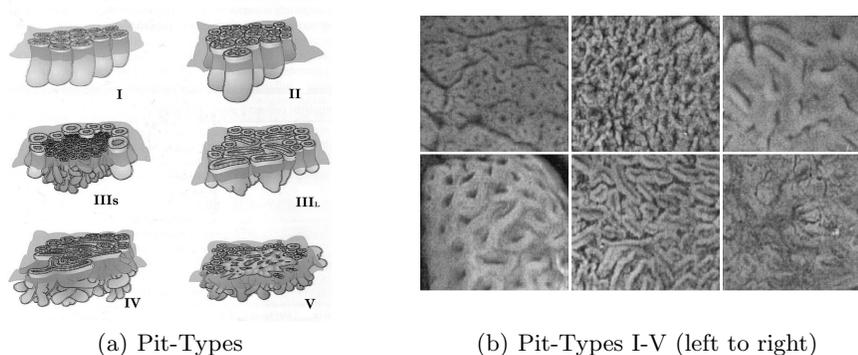
Fig. 1. Pit-Pattern classification according to Kudo et al. [9]

By comparing the prediction of dignity with the histological findings, recent studies [8] have shown, that zoom-endoscopy performs significantly better than the conventional video-endoscopy method, at least in the two-class problem. Although the classification based on Kudo's results seems to be straightforward at first glance, the interpretation of the specific pit-patterns can be challenging [4]. Furthermore, a high-degree of inter-observer variability, which has been reported at least for diagnosing specialized intestinal metaplasia in Barret's esophagus [11], can be a disturbing factor. Therefore, our work intends to find a set of image features with enough discriminative power to enable computer-assisted pit-pattern classification.

3 Feature Extraction and Classification

Each image of our database is decomposed by the Gabor Wavelet Transform [1] and Kingsbury's Dual-Tree Complex Wavelet Transform (DT-CWT) [6]. Both transforms aim to remedy one commonly known shortcoming of the maximally decimated 2-D tensor-product DWT: the inability to capture directional information, which results from the fact that the wavelet filters are separable and real. With respect to our colonoscopy images, this particular shortcoming is important, since the images exhibit structures that are directed at various orientations (see Figure 1(b)). The mother wavelet in the Gabor Wavelet transform is defined by

---

[3] the process of removing polyps

$$\psi(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi jWx\right]. \tag{1}$$

The Gabor wavelets for a given number of orientations N and decomposition scales S are obtained through the generating function

$$\psi_{mn}(x, y) = a^{-m}\psi(\hat{x}, \hat{y}), \quad a > 1, m, n \in \mathbb{N} \tag{2}$$

with $m = 1, \ldots, S - 1$ and scale factor a. The coordinates $\hat{x}, \hat{y}$ are given by

$$\hat{x} = a^{-m}(x\cos\theta + y\sin\theta) \text{ and } \hat{y} = a^{-m}(-x\sin\theta + y\cos\theta). \tag{3}$$

Here, $\theta = n\pi/N$ $(n = 1, \ldots, N)$ denotes the rotation angle. To obtain a frequency partitioning, where the half-peak magnitudes of the filters touch each other (see Figure 3), the filter parameters $\sigma_u = 1/2\pi\sigma_x, \sigma_v = 1/2\pi\sigma_y$ in the frequency domain are related to each other according to [1]. The final decision is, how to choose the upper U_u and lower U_l center frequencies of the filters, which, together with the number of decomposition scales S, determine the scaling factor a.

The second transformation we employ here, the DT-CWT, provides an efficient realization of a complex wavelet transform, which is approximately shift-invariant and produces approximately analytic wavelets, oriented at a fixed number of six directions ($\approx$ 15°, 45°, 75°, 105°, 135°, 165°). Furthermore, the DT-CWT is only four-times redundant in 2-D and allows an efficient implementation based on parallel filter bank trees. The key concept is, that the filters in each tree are designed so that the real wavelets they generate from an approximate Hilbert transform pair and the resulting complex wavelet is thus approximately analytic. The basic structure of the DT-CWT in 1-D is shown in Figure 2, where a different set of low- $(h_0(n), g_0(n))$ and high-pass $(h_1(n), g_1(n))$ filters is used in the first stage. This is a crucial requirement of the DT-CWT, because otherwise the first stages would not be approximately analytic [13]. One tree of the DT-CWT gives the real part of the complex wavelet coefficients, while the other gives the imaginary part. The extension to two dimensions is straightforward and requires four separable wavelet transforms in parallel.
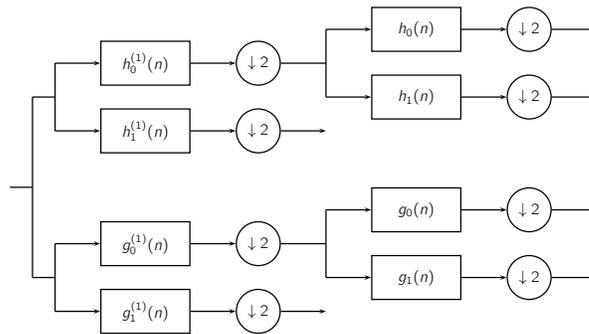
Fig. 2. DT-CWT analysis FB (1-D)

Our feature extraction step is based on the assumption, that the marginal distributions of the transform coefficients $\{x_{in}\}_{1 \leq n \leq L_i}$, $x_{in} \in \mathbb{C}$ of subband i characterize the various pit-types. Here, L_i denotes the number of coefficients of subband i. We use the classic mean μ_i and standard deviation σ_i of the magnitude of the transform coefficients, given by

$$\mu_i = \frac{1}{L_i} \sum_{j=1}^{L_i} |x_{ij}| \quad \sigma_i = \frac{1}{L_i} \left(\sum_{j=1}^{L_i} (|x_{ij}| - \mu_i)^2 \right)^{\frac{1}{2}} \tag{4}$$

to constitute a general feature vector $\mathbf{v} = [\mu_1, \sigma_1, \ldots, \mu_B, \sigma_B]^T$ for a given image, with B denoting the total number of subbands. The same image features have already been successfully used in [1, 12] for texture image retrieval.

To incorporate color information into the classification process, we first decompose each image into its red (R), green (G) and blue (B) color channel and transform each channel separately. Thus, we obtain three features vectors $\mathbf{v}_R, \mathbf{v}_G$ and $\mathbf{v}_B$ for each image of our database. To show, that color channel information can improve classification results, we also compute a feature vector $\mathbf{v}_L$ for the luminance (L) channel.

A question, that has not been answered yet is, how to choose the parameters of the Gabor Wavelet Transform and the DT-CWT. Regarding the Gabor Wavelets, we have to choose U_l and U_u, as well as the number of decomposition scales S and orientations N. In [1], the authors found that four scales and six orientations, with $(U_l, U_u) = (0.05, 0.4)$ was optimal for their problem. This led to a scaling factor of $a = 2$, where the center frequencies are one octave apart. In our studies, we adopt the number of orientations and scales, but we vary the upper and lower center frequencies between $0.02 \leq U_l \leq 0.08$ and $0.1 \leq U_u \leq 0.4$ for each color channel. Our results show, that this is a reasonable step, since it is doubtful that the same parameter setting is optimal for all channels. In case of the DT-CWT the number of orientations is fixed to six and we choose a decomposition depth of four scales. Thus, we obtain a total of $B = 24$ subbands for each color channel for both transforms. For that reason, the dimensionality of the resulting feature vectors is $|\mathbf{v}_R| = |\mathbf{v}_G| = |\mathbf{v}_B| = 48$.

Since we either get three feature vectors for the RGB image or one feature vector for the luminance channel, the question arises, how to combine the color information to obtain the best classification results. A first intuitive approach is to *concatenate* the feature vectors of the color channels for each image into one big feature vector $\mathbf{v} = [\mathbf{v}_R \mathbf{v}_G \mathbf{v}_B]^T$ with $3 \cdot 48 = 144$ dimensions.

Another approach to tackle this problem is, to use a so called *multi-classifier*, with three single classifiers operating on the distinct feature sets. The outputs of the classifiers are then combined by using the common combining rules, introduced by Kittler et al.[7]. In multi-classifier terminology, the resulting classifier is a so called *parallel multi-classifier*. We will use k-NN classifiers as building blocks for our experiments. To obtain a reasonable error prediction, we employ leave-one-out cross-validation (LOO-CV). The LOO-CV results will be given by the classification *accuracy*, which is defined as the number of correctly classified samples (true positives) divided by the total sample size.

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. To enhance visual appearance, dye-spraying with indigo-carmine was applied and biopsies or mucosal resections were taken to obtain a histopathological diagnosis. 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/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 image samples per pit-pattern class (ground truth)

Before actually transforming the images, we perform two image-quality enhancing preprocessing steps: histogram equalization using the CLAHE [14] (contrast limited histogram equalization) algorithm and blurring with a 3×3 Gaussian kernel ($\sigma = 0.5$). The LOO-CV results for the red, green and blue color channel are given in Table 2. The numbers represent the classification accuracy of the LOO-CV.

Pit-Type	I	II	III-L	III-S	IV	V	Total
Gabor Wavelet Transform							
2-cls., red	85.86			90.21			88.43
2-cls., blue	85.35			88.81			**87.40**
2-cls., green	83.33			89.86			87.19
6-cls., red	73.81	61.11	61.29	38.89	69.18	63.33	66.32
6-cls., blue	73.02	61.11	61.29	38.89	68.49	60.00	65.50
6-cls., green	65.08	58.33	79.03	50.00	72.60	60.00	**66.94**
Dual-Tree Complex Wavelet Transform							
2-cls., red	90.91			95.10			**93.39**
2-cls., blue	84.34			87.76			86.36
2-cls., green	86.87			89.51			**88.43**
6-cls., red	80.95	73.61	70.97	72.22	79.45	75.00	**77.07**
6-cls., blue	74.60	63.89	69.35	44.44	69.86	71.67	**69.42**
6-cls., green	65.87	62.50	61.29	55.56	65.07	63.33	63.84

Table 2. Comparison of the two- and six-class LOO-CV results for both transforms, using a k-NN classifier with $k = 1$. The top results for each color channel are marked bold.

We can see, that for both transforms the features obtained from the red color channel lead to the best results in the two-class problem, whereas the features from the green color channel are most discriminative in the six-class problem. Regarding the parameter k for the k-NN classifier, we ran several tests with $1 \leq k \leq 20$, with $k = 1$ leading to the best results in almost all cases. Table 3 lists the best parameter settings (U_l, U_u) for the Gabor Wavelet transform, together with the resulting scale factor a. We point out, that the optimal parameters for the color channels are in no case equal to the parameters in [1] $(U_l, U_u) = (0.05, 0.4)$ and additionally differ for the two classification problems.

	2-class				**6-class**			
	Gray	Red	Blue	Green	Gray	Red	Blue	Green
U_l	0.05	0.05	0.07	0.08	0.05	0.06	0.07	0.05
U_u	0.40	0.50	0.30	0.40	0.30	0.40	0.20	0.30
$\rightarrow a$	2.00	2.15	1.62	1.71	1.81	1.88	1.42	1.81

Table 3. Optimal parameter settings for the Gabor Wavelet Transform

.

To illustrate the difference between the Gabor Wavelet Transform and the DT-CWT regarding the frequency partitioning, Fig. 3 shows the half-peak (Gabor Wavelets) and 70% peak magnitude (DT-CWT) of the filter responses in the fre-

quency domain for a selection of filters from both transforms. The parameters of the Gabor Wavelet Transform are the optimal parameters for the red color channel in the two-class problem.
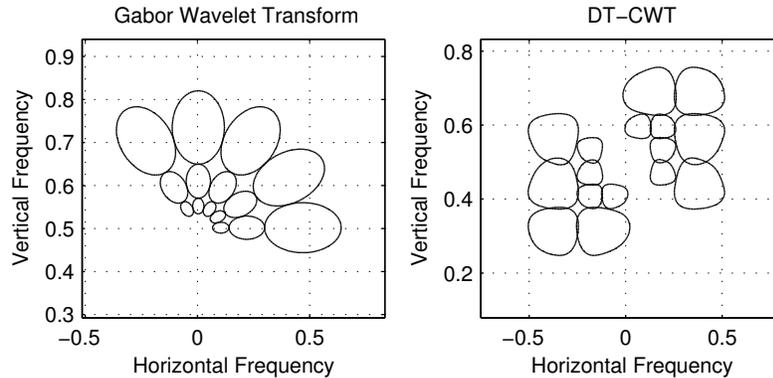
Fig. 3. Frequency partitioning for the Gabor Wavelet Transform and the DT-CWT

Regarding the combination of the information from the different color channels of the RGB color model, Table 4 lists the LOO-CV results for both transforms and combination methods, together with the LOO-CV results of the luminance (gray) channel features. The concatenation scheme is abbreviated by *concat.* and the multi-classifier is abbreviated by *MC*.

Pit-Type	I	II	III-L	III-S	IV	V	Total
Gabor Wavelet Transform							
2-cls., gray	87.37		89.86				88.44
2-cls., concat.	88.89		92.66				91.12
2-cls., MC	89.90		95.45				93.18
6-cls., gray	73.63	63.89	62.90	66.67	74.66	56.67	69.98
6-cls., concat.	77.78	66.67	74.19	50.00	71.23	66.67	71.28
6-cls., MC	82.54	68.06	75.81	38.89	84.93	63.33	76.26
Dual-Tree Complex Wavelet Transform							
2-cls., gray	87.88		91.61				90.08
2-cls., concat.	92.63		98.60				**96.28**
2-cls., MC	93.43		97.90				96.07
6-cls., gray	73.02	68.06	66.13	55.56	69.86	70.00	69.42
6-cls., concat.	85.71	69.44	88.71	83.33	82.88	78.33	**81.82**
6-cls., MC	84.92	68.06	82.26	66.67	84.25	76.67	80.17

Table 4. Comparison of the LOO-CV results for the luminance channel and the combined color channel information, using simple concatenation and a 1-NN multi-classifier with the *product* combining rule. The top results for both classification problems are marked bold.

For both transforms, the so called *product rule* [7] performs best for the two- and six-class problem using the multi-classifier. Although, the overall best LOO-CV results are obtained from the DT-CWT features using the simple concatenation scheme. Due to the fact, that no combination scheme is optimal in all cases, we follow that there is no common rule for combining color channel information for our classification problem. However, both combining schemes are at least superior to the luminance channel features alone.

5 Conclusion

In this work, we presented an approach to computer-assisted pit-pattern classification by means of two multi-directional multi-resolution transforms. Compared to the top results in [3] (75% in the two-class problem and 57% in the six-class problem), which are based on the same image database, we have come to the conclusion that the directional selectivity of both transforms leads to improved classification results and outperforms feature extraction based on the maximally decimated 2-D DWT. In addition, color channel information leads to even better results and is superior to luminance-channel based image processing, at least for our two classification problems. This is consistent with the results in [2], where information from three-dimensional color histograms led to the top classification results. Regarding the choice, which transform to choose, we favour the DT-CWT, since it is less computationally expansive than the Gabor Wavelet transform and leads to slightly better classification results.

References

1. W.Y. Ma B.S. Manjunath. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842, August 1996.
2. M. Häfner, Ch. Kendlbacher, W. Mann, W. Taferl, F. Wrba, A. Gangl, A. Vécsei, and A. Uhl. Pit pattern classification of zoom-endoscopic colon images using histogram techniques. In Johannes R. Sveinsson, editor, *Proceedings of the 7th Nordic Signal Processing Symposium (NORSIG 2006)*, pages 58–61, Reykavik, Iceland, June 2006. IEEE.
3. M. Häfner, M. Liedlgruber, F. Wrba, A. Gangl, A. Vécsei, and A. Uhl. Pit pattern classification of zoom-endoscopic colon images using wavelet texture features. In W. Sandham, D. Hamilton, and C. James, editors, *Proceedings of the International Conference on Advances in Medical Signal and Image Processing (MEDSIP 2006)*, Glasgow, Scotland, UK, July 2006. paper no. 0038.
4. D.P. Hurlstone. High-resolution magnification chromoendoscopy: Common problems encountered in "pit pattern" interpretation and correct classification of flat colorectal lesions. *American Journal of Gastroenterology*, 97:1069–1070, 2002.
5. Stavros A. Karkanis. Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE Transactions on Information Technology in Biomedicine*, 7(3):141–152, September 2003.
6. Nick G. Kingsbury. The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters. In *Proceedings of the IEEE Digital Signal Processing Workshop, DSP '98*, pages 9–12, Bryce Canyon, USA, August 1998.
7. J. Kittler, M. Hatef, R.P.W. Duin, and J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, March 1998.
8. Kazuo Konishi, Kazuhiro Kaneko, Toshinori Kurahashi, Taikan Yamamoto, Miki Kushima, Akira Kanda, Hisao Tajiri, and Keiji Mitamura. A comparison of magnifying and nonmagnifying colonoscopy for diagnosis of colorectal polyps: a prospective. *Gastrointestinal Endoscopy*, 57:48–53, 2003.
9. S. Kudo, S. Hirota, T. Nakajima, S. Hosobe, H. Kusaka, T. Kobayashi, M. Himori, and A. Yagyuu. Colorectal tumorous and pit pattern. *Journal of Clinical Pathology*, 47:880–885, 1994.
10. D. E. Maroulis, D. K. Iakovidis, S. A. Karkanis, and D. A. Karras. CoLD: a versatile detection system for colorectal lesions in endoscopy video-frames. *Computer Methods and Programs in Biomedicine*, 70(2):151–66, February 2003.
11. A. Meining, T. Rösch, R. Kiesslich, M. Muders, F. Sax, and W. Heldwein. Inter- and intra-observer variability of magnification chromoendoscopy for detecting specialized intestinal metaplasis at the gastroesophageal junction. *Endoscopy*, 36:160–164, 2004.

12. P. Rivaz and N. Kingsbury. Complex wavelet features for fast texture image retrieval. In *Proceedings of the IEEE International Conference on Image Processing (ICIP'99)*, pages 109–113, Kobe, Japan, 1999.
13. I. W. Selesnick, R. G. Baraniuk, and N. Kingsbury. The dual-tree complex wavelet transform - a coherent framework for multiscale signal and image processing. *IEEE Signal Processing Magazine*, 22(6):123–151, November 2005.
14. K. Zuiderveld. Contrast limited adaptive histogram equalization. In Paul S. Heckbert, editor, *Graphics Gems IV*, pages 474–485. Morgan Kaufmann, 1994.