

Statistical and Structural Wavelet Packet Features for Pit Pattern Classification in Zoom-Endoscopic Colon Images

MICHAEL LIEDLGRUBER
University of Salzburg
Department of Computer Sciences
J.-Haringerstr.2, 5020 Salzburg
AUSTRIA

ANDREAS UHL
University of Salzburg
Department of Computer Sciences
J.-Haringerstr.2, 5020 Salzburg
AUSTRIA

Abstract: We discuss features extracted from a wavelet packet decomposition for image classification. Statistical features computed from wavelet packet coefficients are compared to structural features which are derived from an image dependent wavelet packet decomposition subband structure. Primary application area is the classification of pit pattern structures in zoom-endoscopic colon imagery, while results are also compared to the outcome of a classical texture classification application.

Key-Words: image classification, texture classification, wavelet packets, pit pattern, colon zoom-endoscopy

1 Introduction

Colonoscopy is a medical procedure, which enables a physician to examine the colon's inside appearance. The ability of taking pictures from inside the colon facilitates analysis of images or video sequences with the assistance of computers. This work describes a novel approach to computer-assisted classification of specific tumorous lesions (see [7] for previous work).

To get detailed images, a special endoscope - a *magnifying endoscope* - is used. A magnifying endoscope represents a significant advance in colonoscopic diagnosis as it provides images which are up to 150-fold magnified. This magnification is possible through an individually adjustable lens. Images taken with this type of endoscope are very detailed as they uncover the fine surface structure of the mucosa as well as small lesions. To visually enhance the structure of the mucosa, and therefore the structure of a potentially tumorous lesion, a common procedure is to spray indigo carmine or methylen blue onto the mucosa.

In this work we compare the use of wavelet packet-based structural features to traditional statistical wavelet features for an automated classification of mucosa imagery acquired by a magnifying colonoscope corresponding to different types of lesions. In Section 2, we review the classification of pit patterns of the colonic mucosa. Section 3 describes the classification approach and first gives an overview and extension of the traditional statistical wavelet-based features. In Section 3.2 we describe the alternative structural features used in more detail. Experimental results and configuration details are presented and discussed in Section 4. Section 5 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 nature of lesions at the time of colonoscopy is important.

Diagnosis of tumorous lesions by endoscopy is always based on some sort of staging, which is a method used to evaluate the progress of cancer in a patient and to see to what extent a tumorous lesion has spread to other parts of the body. Staging is also very important for a physician to choose the right treatment of the colorectal cancer according to the respective stage. A recent classification system, based on so-called *pit patterns* of the colonic mucosa, was originally reported by Kudo et al. [4].

As illustrated in figure 1 this classification differentiates between five main types according to the mucosal surface of the colon. The higher the type number the higher is the risk of a lesion to be malignant: It has been suggested that type I and II pattern are characteristic of non-neoplastic lesions, type III and IV are found on adenomatous polyps, and type V are strongly suggestive of invasive carcinoma.

While lesions of type I and II are benign, representing the normal mucosa or hyperplastic tissue, and in fact are nontumorous, lesions of type III to V in contrast represent lesions which are malignant. Lesions of type I and II can be grouped into non-neoplastic lesions, while

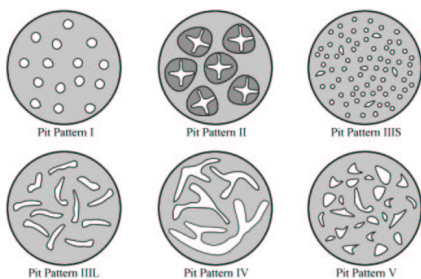


Figure 1: Pit pattern classification according to Kudo et al.

lesions of type III to V can be grouped into neoplastic lesions. Thus a coarser grouping of lesions into two instead of six classes is also possible. Using a magnifying colonoscope together with indigo carmine dye spraying, the mucosal crypt pattern on the surface of colonic lesions can be observed [4]. 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 [5, 10]

Pit type	Characteristics
I	roundish pits which designate a normal mucosa
II	stellate or papillary pits
III S	small roundish or tubular pits, which are smaller than the pits of type I
III L	roundish or tubular pits, which are larger than the pits of type I
IV	branch-like or gyrus-like pits
V	non-structured pits

Table 1: The characteristics of the different pit pattern types.

As depicted in figure 1 pit pattern types I to IV can be characterized fairly well, while type V is a composition of unstructured pits. Table 1 contains a short overview of the main characteristics of the different pit pattern types.

Although at a first glance this classification scheme seems to be straightforward and easy to be applied, it needs some experience and exercising to achieve fairly good results. Correct diagnosis very much relies on the experience of the endoscopist as the interpretation of pit patterns may be challenging [9]. Therefore, a computer-based decision support system would be a valuable help for a physician to provide more reliable on-line diagnosis based on colonoscopy only instead of being forced to wait for the histopathological specimen analysis for custom treatment.

3 Wavelet Packet Classification

If a computer program has to discriminate between different classes of images some sort of classification algorithm has to be applied to the training data during the training phase. During the subsequent classification of an unknown image, the formerly trained classification algorithm is applied to the new, unknown image and tries to classify it correctly. A classification process mainly consist of two parts: the extraction of relevant features from images and the classification based on these features. In this work we rely on the discrete wavelet packet transform (DWP) [13] as a preprocessing stage to extract relevant features.

3.1 Statistical Features

In previous work [7] we have already presented results using two classical feature sets generated from the DWP. The DWP transform domain contains the pyramidal wavelet transform (WT) as a subset the subbands of which are used to extract the corresponding first type of feature vector. Possible features computed are based on the coefficients in the subbands, e.g. the *Energy*, *Logarithm of energy*, *Variance*, *Entropy* or the *l-Norm*.

Local discriminant bases (LDB) [12] is the second type of feature vectors considered in previous work. Contrasting to the previous technique this method is already based on a feature extraction scheme which is highly focused on discrimination between different classes. Here, a wavelet packet basis is constructed which is optimal to discriminate between images of different classes. Once this basis has been identified all training images are decomposed into this basis. The resulting subbands are then used in the subsequent feature extraction step.

In the following we introduce two new ways to extract statistical features from the DWP domain. Both rely on the best-basis algorithm [3] which decomposes a given image into an optimal wavelet packet basis according to a specified cost function (e.g. like *Logarithm of energy*, *Entropy*, L^p -Norm and the *Threshold cost function*). The resulting best basis subband structure usually concentrates the energy of the image in an optimal way.

The Best-basis method (BB) decomposes each image in the training set into an optimal wavelet packet basis with respect to the chosen wavelet family. The resulting subbands are then used to extract features from. Since however the resulting decomposition structures are different among the images, we employ a voting procedure, which assures, that the feature vectors for the different images are based on the same subbands and that the subband ordering within the feature vectors is the same. After all training images are decomposed into their respective best basis subband structure,

we count the occurrence of each subband of a fully decomposed DWP decomposition quadtree in the set of all training images' best basis subband structures. The subbands used to extract features from (also for the images to be subsequently classified) are those with the highest occurrence count.

The Best-basis centroid (BBC) method also decomposes each image in the training set into an optimal wavelet packet basis according to the best basis algorithm. Subsequently, a common decomposition structure is determined, a so-called *centroid*, into which all images are being subsequently decomposed and which is used to extract features from. This centroid is obtained by determining the subband structure which has the smallest average distance to all best-basis decomposition trees of the training image set according to some quadtree distance metric (see next subsection for an example).

3.2 Structural Features

While the features employed in the previous section are based on wavelet coefficients' statistics, this section introduces two feature extraction schemes which use the *decomposition subband structures* resulting from the best basis algorithm as features.

To be able to use the decomposition structures to create feature vectors we need to compare different subband structures and assess their amount of difference. For this purpose, we introduce the concept of distinctive numerical labels of nodes in a decomposition quadtree, so-called *unique node values*. Using the unique node values of a decomposition tree for a given image we are able to create a feature vector. For example, the unique node value u_i for vertex v_i in the decomposition tree can be computed as

$$u_i = \sum_{j=1}^{M-1} 5^j p_t \quad \text{with } p_t \in \{1, 2, 3, 4\}$$

where j is the indicator of the decomposition level in the quadtree, M the level of vertex v_i , and p_t determines the type of the node (whether it is lower left, lower right, upper left, or upper right child of its parent node).

3.2.1 Unique Node Distance (BBS)

The resulting feature vectors cannot be compared using classical metrics due to their different vector length and different positions of corresponding unique node values caused by intermediate vertices. To ensure, that the feature vectors among all images contain the same node positions, for each node present in a tree a zero is inserted into the feature vectors for those images, which do not contain the according node in their decomposition structures. This process of equalizing the fea-

ture vectors in length and in terms of the node positions present is illustrated in figure 2.

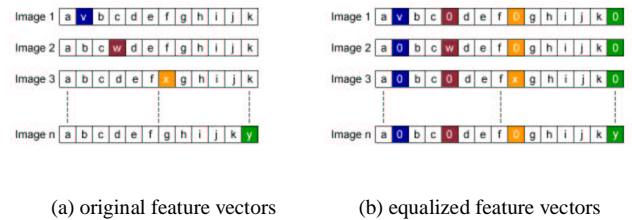


Figure 2: Illustration of the feature vector equalization for unique node values feature vectors.

In figure 2(a) we see feature vectors for some images 1 to n . As we notice, each of these feature vectors contains a feature vector entry (unique node value in this context), which is not present in the other feature vectors (denoted by v , w , x and y in the shaded boxes). The resulting feature vectors after equalization are shown in figure 2(b).

After the equalization, we are able to use the feature vectors in conjunction with any metric and the classifiers mentioned in Section 3.3.

A possible metric used to calculate distances between feature vectors is a weighted version of the Euclidean distance metric

$$D(f, g) = \sum_{i=1}^n \left(\frac{f_i}{w_i^{(f)}} - \frac{g_i}{w_i^{(g)}} \right)^2 \quad (1)$$

where f and g are the feature vectors, $w^{(f)}$ and $w^{(g)}$ are weighting vectors associated with f and g , respectively and n is the length of f and g . The weighting vectors may be derived from the decomposition level of the node or from the cost information of the subbands the feature vectors have been composed from.

3.2.2 Tree Distance (TD)

Just like it is done in the BBS method, all images in the training set are decomposed to their respective best basis decomposition trees according to a chosen wavelet family and cost function. Unlike in BBS, we do not aim at applying a general metric to equalized feature vectors but want to develop a custom tree distance metric.

In order to calculate the distance between two arbitrary quadtrees T_a and T_b , we first create two sets of the unique node values contained in each of the trees \mathbb{U}_a and \mathbb{U}_b . Then the unique node values, which are contained in either the set \mathbb{U}_a or the set \mathbb{U}_b only, are summed up:

$$d(T_a, T_b) = \frac{\sum_N x_i}{y} \quad (2)$$

with

$$x_i \in (\mathbb{U}_a \setminus \mathbb{U}_b) \cup (\mathbb{U}_b \setminus \mathbb{U}_a)$$

where N denotes the number of elements x_i contained only in one of the two sets \mathbb{U}_a or \mathbb{U}_b and y denotes the sum over all elements contained in the set $\mathbb{U}_a \cup \mathbb{U}_b$.

It is obvious, that for equal trees the sets \mathbb{U}_1 and \mathbb{U}_2 contain the same elements and the distance therefore is 0. By dividing the distance by the value y we additionally introduce an upper limit of 1 to the distance.

Another property of this metric is that differences between two quadtrees at nodes near the root node (lower tree levels) contribute more to the distance, while differences at higher tree levels have a much lower impact on the resulting distance. This is due to the fact that the unique node values used are created in a way, which ensures that they get smaller down the tree.

3.3 Classification

The k-NN classifier is one of the simplest classification algorithms. Classification is done by finding the k closest neighbours to an input feature vector \mathbf{x} in the feature space according to some distance metric. The unknown input vector \mathbf{x} is then assigned to the class to which the majority of the k nearest neighbours belongs to.

The SVM classifier, further described in [8, 2], is another, more recent technique for data classification, which has already been used to classify texture using wavelet features successfully [11]. The basic idea behind Support Vector Machines (SVM) is to construct classifying hyperplanes which are optimal for separation of given data.

The Bayes classifier [6] is a probabilistic classifier based on the Bayes theorem. This classifier assigns each unknown image to that class, to which the image belongs most probably or which causes minimal costs in respect to some cost function. This is done by applying the commonly used maximum a posteriori decision rule (MAP decision rule). The MAP decision rule utilizes the Bayes theorem to maximize the a posteriori probability and this way the most probable class for a given (unknown) image is determined.

4 Experiments

4.1 Settings

In our experiments we use 484 images acquired in 2005 and 2006 at the Department of Gastroenterology and Hepatology (Medical University of Vienna) using a zoomcolonoscope (Olympus Evis Exera CF-Q160ZI/L) with a magnification factor set to 150. 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 286 neoplastic cases. The detailed classification results, which are used as ground truth for our experiments, are shown in Table 2.

Pit Type	2 cls.	I		II			
	6 cls.	I	II	IIIS	IIIL	IV	V
Images		126	72	18	62	146	60

Table 2: Number of images per class used in experiments

Using leave-one-out cross-validation, 483 out of 484 images are used as training set. The remaining image is then classified. This process is repeated for each single image.

To be able to compare the classification performance of the methods presented in this paper with other texture databases as well, we carried out additional tests using the Outex image database [1]. Table 3 shows the number of images per class (the respective classes are composed of images of the types Canvas002, Carpet002, Canvas011, Canvas032, Carpet009, and Tile006) used throughout the tests carried out.

Class	1	2	3	4	5	6
2 classes	180	180				
6 classes	180	180	180	180	180	180

Table 3: Details about the Outex images used

For classification, we employ colour channels separately and the Y-channel and show the best results encountered. The same is true for the actual features derived from wavelet subbands (several are tested and only the best result is shown).

4.2 Results

4.2.1 Pit Pattern Images

Table 4 shows the results we obtained using the statistical features described in Section 3 to classify the Pit-Pattern images. In table 5 we see the results of the structural features presented in Section 3.2. All tables display the percentage of correctly classified images for each class and overall.

It is interesting to note that within each statistical technique, k-NN classification delivers the worst results and Bayes classification the best. This is not the case for the structural BBS technique, where SVM clearly gives the best result. When comparing the top results, the statistical features are clearly superior to structural features in the case of 2 classes with 84% correctly classified (WT, Bayes) vs. 73% (BBS, SVM). In the 6

Pit Type	I	II	IIIS	IIIL	IV	V	Total
LOCAL DISCRIMINANT BASES							
k-NN, 2C	66		83			76	
k-NN, 6C	69	42	28	45	57	10	49
SVM, 2C	65		89			79	
SVM, 6C	65	51	0	50	64	48	56
Bayes, 2C	73		86			81	
Bayes, 6C	67	49	0	65	55	55	56
BEST BASIS METHOD							
k-NN, 2C	42		76			62	
k-NN, 6C	52	18	0	42	53	0	38
SVM, 2C	56		81			71	
SVM, 6C	59	43	0	47	53	17	46
Bayes, 2C	71		84			79	
Bayes, 6C	63	29	39	65	43	57	50
BEST BASIS CENTROID METHOD							
k-NN, 2C	70		76			73	
k-NN, 6C	54	35	11	45	42	43	43
SVM, 2C	60		90			78	
SVM, 6C	61	47	11	39	51	38	49
Bayes, 2C	77		87			83	
Bayes, 6C	68	54	6	68	53	62	58
PYRAMIDAL WAVELET TRANSFORM							
k-NN, 2C	56		71			65	
k-NN, 6C	59	32	0	27	47	22	40
SVM, 2C	63		85			76	
SVM, 6C	63	26	0	8	73	30	47
Bayes, 2C	77		88			84	
Bayes, 6C	68	60	6	71	48	65	58

Table 4: Percentage of correctly classified Pit-Pattern images using statistical features.

classes case, the top results are almost identical for statistical and structural features. On average, structural features cannot compete with the statistical ones. However, there are rare configurations where structural features are superior to statistical ones, e.g. comparing BB to BBS using SVM or k-NN classification. Therefore, it is of high importance to choose the right configuration of feature and classifier employed.

Comparing within the statistical features, the BB approach performs clearly worst and LDB is best. BBC provides good results for all classifiers, while WT gives poor results for k-NN classification. Concerning structural features, TD is not competitive at all – this technique delivers the worst results overall.

4.2.2 Outex Images

The results obtained by carrying out our comparative tests using the Outex image database using the statistical features are shown in table 6. Here, we only display the best (LDB) and worst (BB) techniques. In table 7 we see the results obtained using the structural features.

In the 2 classes case, all types of statistical features achieve a classification rate of 100% with all three types of classifiers. LDB exhibits this excellent result also for

Pit Type	I	II	IIIS	IIIL	IV	V	Total
UNIQUE NODE VALUES (BBS)							
k-NN, 2C	47		79			66	
k-NN, 6C	53	31	0	44	52	15	42
SVM, 2C	73		73			73	
SVM, 6C	100	0	0	3	98	0	56
Bayes, 2C	53		75			66	
Bayes, 6C	94	0	0	3	10	22	31
TREE DISTANCE (TD)							
k-NN, 2C	62		46			52	
k-NN, 6C	72	0	0	10	44	0	33

Table 5: Percentage of correctly classified Pit-Pattern images using structural features.

Class	1	2	3	4	5	6	Total
LOCAL DISCRIMINANT BASES							
k-NN, 2C	100	100					100
k-NN, 6C	100	100	100	100	100	100	100
SVM, 2C	100	100					100
SVM, 6C	100	100	100	100	99	100	99
Bayes, 2C	100	100					100
Bayes, 6C	100	100	100	100	100	100	100
BEST BASIS METHOD							
k-NN, 2C	100	100					100
k-NN, 6C	97	92	100	89	87	91	93
SVM, 2C	100	100					100
SVM, 6C	98	98	100	98	94	93	97
Bayes, 2C	100	100					100
Bayes, 6C	100	100	100	100	98	99	99

Table 6: Percentage of correctly classified Outex images using statistical features.

the 6 classes case (except for the SVM classifier where 99% are obtained). The BB method is clearly inferior in the 6 classes case, still resulting in 93% and more correctly classified images. Similar to the pit pattern images, k-NN gives worst results with BB and the Bayes classifier even results in 99% correctness for BB in the 6 classes case.

A very different situation is seen with the structural classifiers. Again, we do not observe the best results for the Bayes classifier, but contrasting to the pit pattern images here k-NN classification provides the best results with 100% correctly classified images in the 2 classes case and 94% in the 6 classes case (which is slightly better than the BB result with k-NN, which is the worst BB result, though). Overall, the results of the structural features are clearly inferior as compared to the statistical features.

The excellent results of the statistical features using the Bayes classifier (for Outex as well as for pit pattern images) suggests the wavelet coefficient statistics distribution closely following the normal distribution whereas obviously this is not the case for the distribution of the structural features' vectors.

Class	1	2	3	4	5	6	Total
UNIQUE NODE VALUES (BBS)							
k-NN, 2C	100	100					100
k-NN, 6C	96	99	99	83	92	97	94
SVM, 2C	91	87					89
SVM, 6C	65	70	58	59	46	27	54
Bayes, 2C	92	87					90
Bayes, 6C	57	63	83	74	50	47	63
TREE DISTANCES (TD)							
k-NN, 2C	100	99					99
k-NN, 6C	91	82	78	53	45	53	67

Table 7: Percentage of correctly classified Outex images using structural features.

5 Conclusion

In this work we compare the use of statistical features and structural features for image classification, both types of features are derived from the wavelet packet domain.

The classification results of both wavelet feature types depend strongly on the classification technique applied, where Bayes classification provides clearly the best results in case of statistical features for both zoom-endoscopic images as well as a texture database. For structural features, the top performing classification technique varies, but Bayes classification is never the best one.

Overall, coefficient distribution based statistical wavelet features turn out to be superior as compared to structural features using the best basis decomposition subband structure itself as feature. However, this is not true for each combination of feature extraction and classification technique employed. Therefore, the selection of feature extraction and classifier needs to be jointly optimized in order to achieve optimal classification accuracy.

Acknowledgements: This work has been partially supported by the Austrian Science Fund FWF, project no. L366-N15 and by the Austrian National Bank “Jubiläumsfonds”, project no. 12514.

References:

- [1] University of Oulu Texture Database. Available online at <http://www.outex.oulu.fi/temp/> (28.11.2005).
- [2] Christopher J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.
- [3] R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2):713–719, 1992.
- [4] Shin-Ei Kudo et al. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointestinal Endoscopy*, 44(1):8–14, July 1996.
- [5] K.-I. Fu et al. Chromoendoscopy using indigo carmine dye spraying with magnifying observation is the most reliable method for differential diagnosis between non-neoplastic and neoplastic colorectal lesions: a prospective study. *Endoscopy*, 36(12):1089–1093, 2004.
- [6] Keinosuke Fukunaga. *Statistical Pattern Recognition*. Morgan Kaufmann, 2nd edition, 1990.
- [7] 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, 2006. paper no. 0038.
- [8] Chih-Wei Hsu and Chih-Jen Lin. A comparison of methods for multi-class support vector machines. *IEEE Transactions on Neural Networks*, 13(2):415–425, 2002.
- [9] 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.
- [10] D.P. Hurlstone et al. Efficacy 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, 2004.
- [11] Kashif Mahmood Rajpoot and Nasir Mahmood Rajpoot. Wavelets and support vector machines for texture classification. In *Proceedings of the 8th IEEE International Multioptic Conference (INMIC'04)*, pages 328–333, 2004.
- [12] Naoki Saito and Ronald R. Coifman. Local discriminant bases and their applications. *J. Mathematical Imaging and Vision*, 5(4):337–358, 1995.
- [13] M.V. Wickerhauser. *Adapted wavelet analysis from theory to software*. A.K. Peters, Wellesley, Mass., 1994.