

Handbook of Research on Advanced Techniques in Diagnostic Imaging and Biomedical Applications

Themis P. Exarchos
University of Ioannina, Greece

Athanasios Papadopoulos
University of Ioannina, Greece

Dimitrios I. Fotiadis
University of Ioannina, Greece

Director of Editorial Content: Kristin Klinger
Senior Managing Editor: Jamie Snavely
Managing Editor: Jeff Ash
Assistant Managing Editor: Carole Coulson
Typesetter: Kim Barger
Cover Design: Lisa Tosheff
Printed at: Yurchak Printing Inc.

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue, Suite 200
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: <http://www.igi-global.com/reference>

and in the United Kingdom by
Information Science Reference (an imprint of IGI Global)
3 Henrietta Street
Covent Garden
London WC2E 8LU
Tel: 44 20 7240 0856
Fax: 44 20 7379 0609
Web site: <http://www.eurospanbookstore.com>

Copyright © 2009 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher.

Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Handbook of research on advanced techniques in diagnostic imaging and biomedical applications / Themis P. Exarchos, Athanasios Papadopoulos, and Dimitrios I. Fotiadis, editors.

p. ; cm.

Includes bibliographical references and index.

Summary: "This book includes state-of-the-art methodologies that introduce biomedical imaging in decision support systems and their applications in clinical practice"--Provided by publisher.

ISBN 978-1-60566-314-2 (hardcover)

1. Imaging systems in medicine--Data processing--Handbooks, manuals, etc. 2. Diagnostic imaging--Digital techniques--Handbooks, manuals, etc. I. Exarchos, Themis P., 1980- II. Papadopoulos, Athanasios, 1971- III. Fotiadis, Dimitrios Ioannou.

[DNLM: 1. Diagnostic Imaging--methods--Handbooks, manuals, etc. 2. Diagnostic Imaging--trends. 3. Image Processing, Computer-Assisted--methods. WN 180 H2365 2009]

R857.O6H367 2009

616.07'54--dc22

2008050005

British Cataloguing in Publication Data

A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

Chapter XXII

Pit Pattern Classification Using Multichannel Features and Multiclassification

Michael Häfner

Medical University of Vienna, Austria

Alfred Gangl

Medical University of Vienna, Austria

Michael Liedlgruber

Salzburg University, Austria

A. Uhl

Salzburg University, Austria

Andreas Vécsei

St. Anna Children's Hospital, Austria

Friedrich Wrba

Medical University of Vienna, Austria

ABSTRACT

Wavelet-, Fourier-, and spatial domain-based texture classification methods have been used successfully for classifying zoom-endoscopic colon images according to the pit pattern classification scheme. Regarding the wavelet-based methods, statistical features based on the wavelet coefficients as well as structural features based on the wavelet packet decomposition structures of the images have been used. In the case of the Fourier-based method, statistical features based on the Fourier-coefficients in ring filter domains are computed. In the spatial domain, histogram-based techniques are used. After reviewing the various methods employed we start by extracting the feature vectors for the methods from one color channel only. To enhance the classification results the methods are then extended to utilize multichannel features obtained from all three color channels of the respective color model used. Finally, these methods are combined into one multiclassifier to stabilize classification results across the image classes.

INTRODUCTION

Today, the third most common malignant disease in western countries is colon cancer. For that reason a regular examination of the colon is recommended, especially for people at an age of 50 years and older. *Colonoscopy* is currently the best test available to identify colon cancer.

Colonoscopy is a medical procedure which allows a physician to investigate the inside of the colon. This is done by using a colonoscope, a flexible instrument equipped with a CCD chip for visualisation of the organ and controlled by the physician. In case a lesion is detected, tissue samples can be taken and relevant lesions can be removed, avoiding thus surgery.

Modern colonoscopies allow the acquisition of digital images and video sequences from inside the colon during the colonoscopy. This makes it easier for the physician to review the results from a colonoscopy and to document the growth and spreading of an eventually tumorous lesion. To obtain images which are as detailed as possible a magnifying colonoscope is used. This type of colonoscope provides images which are up to 150-fold magnified and thus are very detailed as they uncover the fine surface structure of the mucosa as well as small lesions.

A common procedure to visually enhance the structure of the mucosa is to spray indigo carmine or methylen blue onto the mucosa. While dyeing with indigo carmine causes a plastic appearance of the mucosa, dyeing with methylen blue helps to highlight the boundary of a lesion. Cresyl violet is often used to actually stain the margins of the pit structures, which is also referred to as staining.

In this work we document the good performance of several texture classification techniques to perform an automated classification of pit pattern images acquired by a magnifying colonoscope. Based on these methods, we show the benefit of using features based on three color channels. Finally, we present one possible way to combine several methods and classifiers to build a multiclassifier.

Note that the developed techniques are not meant to replace the physicians' diagnosis but are designed to act as a decision support system for the human operator during colonoscopy – here a reliable and immediate diagnosis is a significant advantage since a second colonoscopy required in many cases can be avoided as there is no need to wait for the histological classification of eventually extracted biopsies.

PIT PATTERN CLASSIFICATION

Polyps of the colon are a frequent finding and are usually divided into metaplastic, adenomatous, and malignant. As resection of all polyps is time-consuming, it is imperative that those polyps which warrant endoscopic resection can be distinguished: polypectomy of metaplastic lesions is unnecessary and removal of invasive cancer may be hazardous. For these reasons, assessing the malignant potential of lesions at the time of colonoscopy is important.

To be able to differentiate between the different types of lesions a classification method is needed.

The most commonly used classification system for distinguishing between non-neoplastic and neoplastic lesions in the colon is the pit pattern classification originally reported by Kudo, Hirota et al. (1994) and Kudo, Tamura et al. (1996).

This system allows a differentiation between normal mucosa, hyperplastic lesions (non-neoplastic), adenomas (a pre-malignant condition), and malignant cancer based on the visual pattern of the mucosal surface. Hence, this classification scheme is a convenient tool to decide which lesions need not, which should, and which most likely can't be removed endoscopically. The mucosal pattern as seen after dye staining and by using magnification endoscopy shows a high agreement with the histopathologic diagnosis. Furthermore, due to the fact that this method is based on the histopathologic (and therefore visual) structure of

the mucosa, it is a convenient choice for a classification using image processing methods.

As illustrated in Figure 1, this classification method differentiates between the five main types I to V according to the mucosal surface of the colon. Type III is divided into two sub-types, III-S and III-L, designating the size of the pit structure. The higher the number of the pit type is, the higher is the risk that the lesion under investigation is malignant.

It has been suggested that pattern of type I and II are characteristic of non-neoplastic lesions, type III and IV are found on adenomatous polyps, and type V are strongly suggestive of invasive carcinoma.

Lesions of type I and II are benign, representing the normal mucosa or hyperplastic tissue, and in fact are non-tumorous. Lesions of type III and IV in contrast represent lesions which are neoplastic. Type V lesions usually are highly indicative for cancer. 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 (Kudo et al., 1996). Several studies found a good correlation between the mucosal pit pattern and the histological findings, where especially techniques using magnifying colonoscopes led to excellent results (Hurlstone et al., 2004).

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 (Hurlstone, 2002; Tung, Wu, & Su, 2001) – here, an automated decision support system for the physician conducting the colonoscopy would improve the situation. To illustrate this, Figure 2 contains images out of the training set used throughout this work.

FEATURE EXTRACTION AND CLASSIFICATION

The task of automated image classification consists of two major parts: the extraction of relevant features from images and the classification based on these features. These parts are outlined in the following section.

Wavelet-Based Methods

Previous work has already shown that wavelet-based methods can be used successfully for the classification of colon cancer. In (Karkanis et al., 2001) frames of an endoscopic video are transformed to the wavelet domain using the discrete pyramidal *wavelet transform*. Based on the resulting wavelet coefficients, second order statistics are computed from co-occurrence matrices for the wavelet subbands. These statistical values are used as input for an artificial neural network. An implementation along with results is documented in (Maroulis et al., 2003).

The approaches described in (Karkanis et al., 1999; Karkanis et al., 2000) are very similar, but instead of using all subbands, only the subband with the highest variance in the coefficient histogram is used to obtain features.

In (Häfner et al., 2006a; Liedlgruber & Uhl, 2007) different wavelet-based methods in conjunction with different classifiers have been used successfully for *pit pattern classification*. Six distinct methods have been investigated to obtain features based on the wavelet transform. The types of features include statistical features as well as structural features.

Statistical Features

In previous work (Häfner et al., 2006a) we have already presented results using two classical feature sets generated from the discrete wavelet packets transform (DWP). The DWP transform domain contains the pyramidal wavelet transform

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

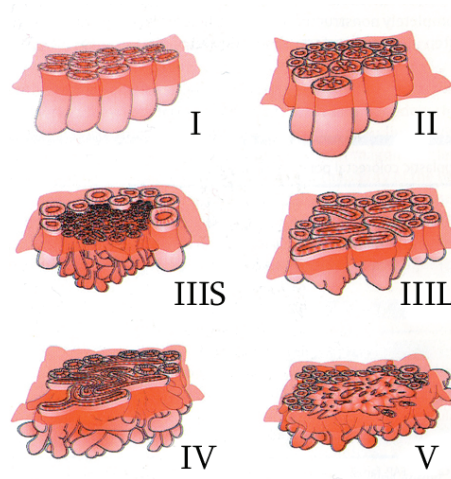
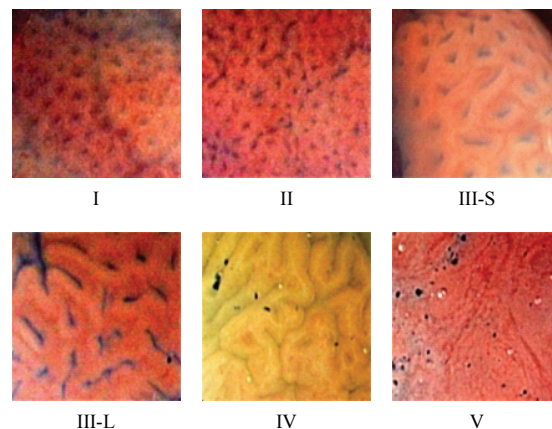


Table 1. The characteristics of the different pit pattern types.

Pit type	Characteristics
I	roundish pits, which designate a normal mucosa
II	stellar 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

Figure 2. Images taken with a colonoscope showing the different types of pit pattern.



(WPC) as a subset of the subbands 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_1 -Norm).

The Local discriminant bases algorithm (LDB) (Saito & Coifman, 1995; Saito, 1997; Rajpoot, 2003) is used to generate the second type of feature vectors considered in previous work. Contrasting to the previous technique this method is already highly focused on discrimination between different classes. Here, a wavelet packet basis optimal for discrimination between images of different classes is constructed. 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.

Wavelet packets can be used in two additional ways to extract statistical features from the DWP domain (Liedlgruber & Uhl, 2007). Both rely on the best-basis algorithm (Coifman & Wickerhauser, 1992) 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 subband structure usually concentrates the energy of the image in an optimal way.

The Best-basis method (BB) decomposes each training image into an optimal wavelet packet basis with respect to the chosen wavelet family. The resulting subbands are 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 structures, 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 (BBCB) method also decomposes each training image into an optimal wavelet packet basis according to the best-basis algorithm. Subsequently, a common decomposition structure – a so-called centroid – is determined, 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 images according to some quadtree distance metric.

Structural Features

In contrast to the feature extraction methods presented in the previous section, the methods presented in this section rely on the best-basis subband structures.

In the best-basis structural method (BBS) we use two different ways to create a feature vector (Liedlgruber & Uhl, 2007). The first method creates a feature vector for a given image, which contains the so-called unique node values of the respective decomposition quadtree. These unique node values uniquely identify each possible node in a quadtree. 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. Having obtained these feature vectors, the euclidean distance is used to calculate the similarity between two images.

The second feature extraction method uses the decomposition trees directly as features. Using a quadtree distance metric the distance between two images can then be calculated. These distances are subsequently used to classify an unknown image using the k -NN classifier (see below).

Fourier-Based Approach

Contrasting to the feature extraction methods presented above, also the discrete Fourier transform (DFT) can be used for feature extraction resulting in excellent classification results (Häfner et al., 2007). This approach is based on the *Fast Fourier Transform* (FFT), which is an efficient algorithm to compute the DFT. Using the 2D-FFT, all training images are transformed to the Fourier domain. Based on the resulting Fourier coefficients, the power spectrum is computed for each image. Several non-overlapping rings of different starting offsets and widths are chosen from the power spectrum (ring filters, band-pass filters). These rings are then used to extract features such as the mean and the standard deviation of the coefficients' magnitudes contained in each of the rings. This yields a feature vector consisting of n entries for n distinct rings in the filter. Figure 3 shows two examples of ring filters.

In (Häfner et al., 2007) the entire information contained in the RGB color model has been chosen to create feature vectors from. Therefore a single feature vector is created from each of the three color channels (R, G and B). These feature vectors are then concatenated to form one large feature vector. Then the Bayes classifier (see below) is used for classification.

One major problem is the huge number of possible ring configurations to choose from. As a consequence, in (Häfner et al., 2007) a genetic algorithm is proposed to perform a search for optimal ring configurations. We follow this strategy but use all types of classifiers considered in this work (see below).

Spatial Domain Approach

Pit pattern classification can also be performed in the spatial domain using a selection of different *histograms* (Häfner et al., 2006b).

The experiments in (Häfner et al., 2006b) have been conducted using 1D, 2D and 3D histograms.

The classical intensity 1D histograms are created for all channels of the RGB color model. The 2D histograms (co-occurrence histograms) are created for the luminance channel of the YUV color model. The 3D histograms are created for all channels of the RGB color model concurrently per definition.

For the classification process the k-NN classifier (see next section) is used. The distance measure used to compute the distance between two normalized histograms H_a and H_b for images I_a and I_b , respectively, is based on the so-called histogram intersection. For 1D histograms the distance is defined as

$$d(H_a, H_b) = 1 - \sum_{i=1}^n \min(H_a[i], H_b[i]) \quad (1)$$

where n is the number of bins in each histogram. The equivalents for the 2D and 3D case are

$$d(H_a, H_b) = 1 - \sum_{i=1}^n \sum_{j=1}^n \min(H_a[i, j], H_b[i, j]) \quad (2)$$

and

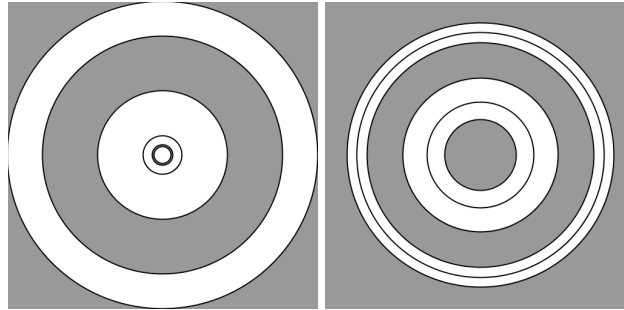
$$d(H_a, H_b) = 1 - \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \min(H_a[i, j, k], H_b[i, j, k]) \quad (3)$$

The experiments in (Häfner et al., 2006b) consider the RGB color model and the luminance channel of the YUV color model only. In this work we examine the information of all color channels of the YUV and HSV color model too.

Classification

In (Häfner et al., 2006a) we employed two distinct classifiers, namely the *k-Nearest Neighbors classifier* (k-NN) and *Support Vector Machines* (SVM). Using the k-NN classifier classification is done by finding the k closest neighbors of an input feature vector \mathbf{x} in the feature space according to some distance metric (e.g. euclidean). The unknown sample \mathbf{x} is then assigned

Figure 3. Two examples of possible ring configurations



to the dominating class among the k nearest neighbors.

The SVM classifier, presented in more detail in (Chang & Lin, 2001; Hsu, Chang, & Lin; Burges, 1998), is another, recently developed classifier, which has already been successfully used to classify texture using wavelet features (Rajpoot & Rajpoot, 2004). The basic idea behind SVM is to construct classifying hyperplanes, which are optimal for separation of given data. These hyperplanes divide the feature space into two distinct classes. For the 6-classes case a voting procedure, combining the two classes classification steps, is employed.

In addition to these two classifiers, we also use the *Bayes classifier*, which is explained in more detail in (Fukunaga, 1990). This is a probabilistic classifier based on the Bayes theorem. During classification each unknown image is assigned to that class, to which the image belongs to most probably or which causes minimal costs with respect to some cost function.

MULTICHANNEL FEATURES AND THE MULTICLASSIFIER

As already mentioned above, the experiments presented in (Häfner et al., 2007) have been carried out by using all color channels for feature extraction. Since this improved the results, hope

is raised that this will also apply to the wavelet-based methods. Therefore we extended the wavelet-based methods using non-structural features to use all color channels.

Multiclassifier for Two Classes

Although the methods presented above deliver a very promising classification accuracy already, we now describe how these methods have been combined in the 2-classes case to improve the accuracy.

The *multiclassifier* mainly relies on two pieces of information: method ranking and method reliability. The ranking expresses how accurately a method classifies images mostly misclassified. Based on a list containing the x most misclassified images for class c the rank for each method is updated for all x images. The rank for the most accurate method for the given image is incremented by N , which denotes the total number of methods combined. The rank for the second-best method is incremented by $N-1$, and so on. Finally, the ranking $R_{m,c}$ for each method m and class c is normalized and transformed to lie between -1 and 1 . This computation is repeated for each image class c to get the ranking information for all classes.

The method reliability is telling us how much we can rely on the classification result of a specific method. The computation of the reliability A_m for

a method m is based on the Bayesian a posteriori probability:

$$A_m = \left(\sum_{i=1}^C N_i \frac{b_i r_i}{(1-b_i)(1-r_i) + b_i r_i} \right)^{1/2} \frac{N}{2} - 1 \quad (4)$$

where C is the number of classes used, N_i is the number of images in class i , N is the total number of images, b_i is the a priori probability for class i and r_i is the classification rate for class i . Multiplying the inner part of the sum by N_i produces a weighted measure, which accounts for the unbalanced training set. Finally, the value is normalized and transformed to lie between -1 and 1 .

Two allow controlling the strength of influence the ranking and the reliability have on the result of the multiclassifier, these values are remapped by

$$V_f(x) = \text{sign}(x) |x|^{\ln(f)/\ln(0.5)} \quad (5)$$

where f is the parameter controlling the shape of the remapping function. The effect of choosing different values for f is depicted in Figure 4. Obviously $f=0.5$ corresponds to a linear mapping. The resulting image class c_i for an image i is calculated by

$$c_i = \sum_{j=1}^M D_{i,j} V_f(A_j) V_f(R_{j,p}) \quad (6)$$

where M is the number of methods combined, $D_{i,j}$ is the remapped value of the previously assigned class p for image i by method j (-1 for class 1 and 1 for class 2), A_j is the reliability of method j and $R_{j,p}$ is the ranking for method j and class p . The resulting class is then

$$\text{resultclass} = \begin{cases} 1, & \text{if } c_i < 0 \\ 2, & \text{else} \end{cases} \quad (7)$$

Multiclassifier for Six Classes

Due to the binary nature of the *multiclassifier* introduced by Equation (6), the 6-classes case

needs to be handled slightly different. The computation of the ranking and the reliability remain almost the same, just the transformation to the range between -1 and 1 has been omitted. Additionally the reliability is computed for each class which reflects the reliability of a method for a specific class.

The final classification result is obtained by a weighted majority voting, based on value x_c which is calculated for each image i and class c as follows:

$$x_c = \sum_{j=1}^M \alpha_{i,j} V_f(A_{j,c}) V_f(R_{j,c}) \quad (8)$$

where $A_{j,c}$ is the reliability of method j for class c , $R_{j,c}$ is the ranking for method j and $\alpha_{i,j}$ is

$$\alpha_{i,j} = \begin{cases} 1, & \text{if } D_{i,j} = c \\ 0, & \text{else} \end{cases} \quad (9)$$

The final class C for the unknown image is obtained by

$$C = \arg \max_c x_c \quad (10)$$

Methods which perform poor for a class c , are filtered out on a per-class basis. Thus only the X_c most reliable methods are considered for class c . Furthermore, to reduce the worsening effect of unreliable methods, we apply a threshold function to the reliability of a class c . The result is a modified version of Equation (8):

$$x_c = \sum_{j=1}^M \alpha_{i,j} \beta_{j,c,k} B_t(A_{j,c}) V_f(R_{j,c}) \quad (11)$$

where $\beta_{j,c,k}$ is set to 1 if method j is among the k most reliable methods for class c . Otherwise $\beta_{j,c,k}$ is set to 0 and therefore method j is ignored. B_t is the threshold function using threshold t .

To consider the low sample count for class III-S, different threshold values are used for this class and all other classes (empirical values 0.1 and 0.4 , respectively). This is necessary since the

reliability is based on a priori knowledge, which leads to very low reliabilities for class III-S - even if a method performs equally well or better than for another class.

EXPERIMENTS AND RESULTS

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 Evris Exera CF-Q160ZI/L) with a magnification factor set to 150. Lesions found during colonoscopy have been examined after application of dye-spraying with indigo carmine 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.

Due to the rather limited set of images available for our experiments, we use leave-one-out cross-validation. Thus, 483 out of 484 images are used as training set. The remaining image is then classified. This process is repeated for each image.

Results

Multichannel Features

Figure 5 shows the differences between the overall classification results (percentage of correctly classified images) we obtain using features from single color channels and multichannel features in the 2-classes case using the RGB color model. Figure 6 depicts the same comparison, but for the 6-classes case. In Tables 5 and 6 the respective results are presented in a more detailed fashion (especially with respect to results for each single class).

As we can see from Figure 5, regarding the 2-classes case, using multichannel features im-

Figure 4. The effect of different choices for f in the value remapping function

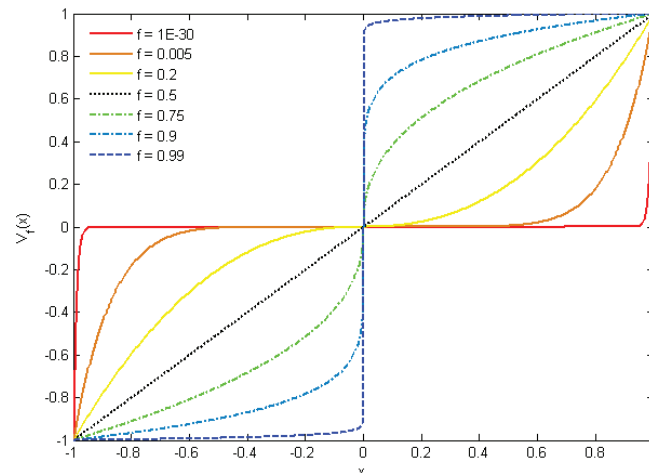
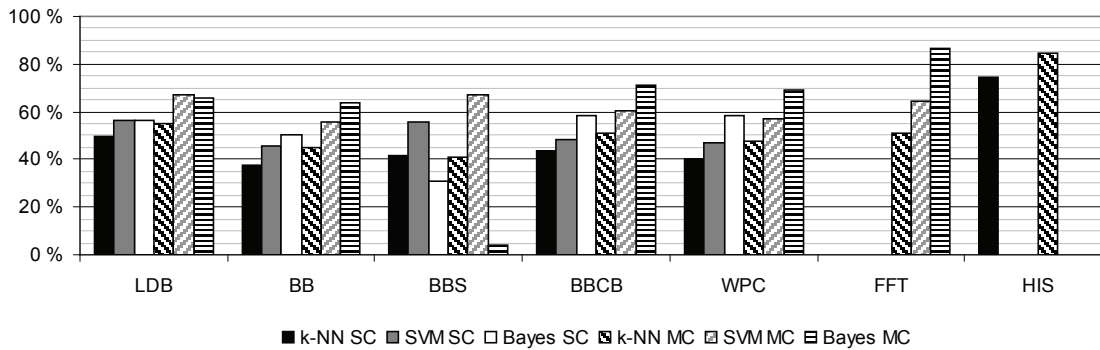


Table 2. Number of images per class used in our experiments

Pit Pattern	I	II	III-S	III-L	IV	V
# of Images	126	72	18	62	146	60

Figure 5. Comparison of the different methods using single channel (SC) and multichannel (MC) features in the 2-classes case



proves the results for nearly all methods. The result improvements lie between 5 % and 17 % for the BB method with the k-NN classifier and BBCB with the Bayes classifier, respectively. Only the result for the BBS method used with the Bayes classifier drops by 4 %. Concerning classifiers, the Bayes classifier gives the best results, followed by SVM and k-NN (except for the BBS method which performs best using SVM).

Regarding the 6-classes case, the improvements are very similar, as shown in Figure 6. Using multichannel features the results are improved by values between 5 % and 13 % for the LDB method with the k-NN classifier and the BB and BBCB methods using the Bayes classifier, respectively. Again the results drop considerably using the BBS method.

While the result drops by 1 % only when using BBS in conjunction with the k-NN classifier, the result decreases by 27 % for the combination of BBS and the Bayes classifier. Again, the Bayes classifier provides the best overall results but again

BBS behaves significantly different compared to the rest of the feature extraction methods.

The best classification results given in Tables 3 to 6 have been found by large scale experimentation testing a significant amount of different parameter settings for each technique considered (e.g. by using different feature vector lengths, different color channels or k-values when using the k-NN classifier).

During our experiments it turned out that regarding the single channel tests most of the best results have been achieved using the red channel of the RGB color model. The other color channel often yielding good results is the luminance channel of the YUV color model. Regarding the multichannel tests the best results have always been achieved using the RGB color model.

Apart from that, it has been observed that some of the feature vector lengths in the multichannel case are considerably higher compared to their single channel counterparts. In the 2-classes case the feature vector lengths vary between 3 and 100, and 3 and 261, in the single channel and

Figure 6. Comparison of the different methods using single channel (SC) and multichannel (MC) features in the 6-classes case

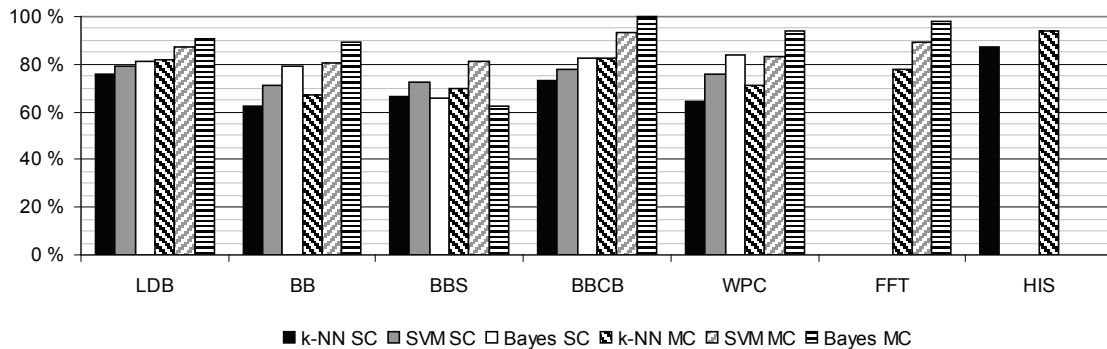


Table 3. Percentage of correctly classified pit pattern images using single channel features in the 2-classes case

	I - II	III-S - V	Total
LOCAL DISCRIMINANT BASES			
k-NN	66	83	76
SVM	65	89	79
Bayes	73	86	81
BEST BASIS METHOD			
k-NN	42	76	62
SVM	56	81	71
Bayes	71	84	79
STRUCTURAL BEST BASIS METHOD			
k-NN	47	79	66
SVM	73	73	73
Bayes	53	75	66
BEST BASIS CENTROID METHOD			
k-NN	70	76	73
SVM	60	90	78
Bayes	77	87	83
PYRAMIDAL WAVELET TRANSFORM			
k-NN	56	71	65
SVM	63	85	76
Bayes	77	88	84
HISTOGRAM METHOD			
k-NN	83	90	87

Table 4. Percentage of correctly classified pit pattern images using multichannel features in the 2-classes case

	I - II	III-S - V	Total
LOCAL DISCRIMINANT BASES			
k-NN	73	88	82
SVM	78	93	87
Bayes	89	92	91
BEST BASIS METHOD			
k-NN	37	88	67
SVM	67	90	81
Bayes	84	92	89
STRUCTURAL BEST BASIS METHOD			
k-NN	62	75	70
SVM	54	100	81
Bayes	44	74	62
BEST BASIS CENTROID METHOD			
k-NN	82	84	83
SVM	92	95	94
Bayes	100	100	100
PYRAMIDAL WAVELET TRANSFORM			
k-NN	56	81	71
SVM	77	87	83
Bayes	91	95	94
HISTOGRAM METHOD			
k-NN	94	93	94
FOURIER METHOD			
k-NN	74	81	78
SVM	85	92	89
Bayes	97	98	98

multichannel case, respectively. In the 6-classes case we observe lengths between 7 and 91, and 3 and 150, when using single channel features and multichannel features, respectively. This resulted in notably higher computational demand.

The k-values for the k-NN classifier used to get the best results in the 2-classes case are rather low compared to the number of samples in each class (between 1 and 11, and 1 and 29, in the single channel and multichannel case,

Table 5. Percentage of correctly classified pit pattern images using single channel features in the 6-classes case

	I	II	III-S	III-L	IV	V	Total
LOCAL DISCRIMINANT BASES							
k-NN	69	42	28	45	57	10	49
SVM	65	51	0	50	64	48	56
Bayes	67	49	0	65	55	55	56
BEST BASIS METHOD							
k-NN	52	18	0	42	53	0	38
SVM	59	43	0	47	53	17	46
Bayes	63	29	39	65	43	57	50
STRUCTURAL BEST BASIS METHOD							
k-NN	53	31	0	44	52	15	42
SVM	100	0	0	3	98	0	56
Bayes	94	0	0	3	10	22	31
BEST BASIS CENTROID METHOD							
k-NN	54	35	11	45	42	43	43
SVM	61	47	11	39	51	38	49
Bayes	68	54	6	68	53	62	58
PYRAMIDAL WAVELET TRANSFORM							
k-NN	59	32	0	27	47	22	40
SVM	63	26	0	8	73	30	47
Bayes	68	60	6	71	48	65	58
HISTOGRAM METHOD							
k-NN	71	61	67	81	84	67	74

respectively). Regarding the 6-classes case, the values are rather high compared to the number of samples of class III-S (between 1 and 10, and 1 and 22, in the single channel and multichannel case, respectively).

In the case of the wavelet-based tests we tried different statistical features of which no dominating one has been observed. For structural features, in most cases using the unique node IDs delivers the best results. This is due to the fact that the underlying implementation is not able to use quadrees as features for the SVM and Bayes classifier.

Figures 7 and 8 show the ring configurations used for the Fourier method in the 2-classes and 6-classes case, respectively. From these figures it seems that mostly low frequency components contain important coefficients for the classifica-

Table 6. Percentage of correctly classified pit pattern images using multichannel features in the 6-classes case

	I	II	III-S	III-L	IV	V	Total
LOCAL DISCRIMINANT BASES							
k-NN	61	38	6	58	73	33	55
SVM	71	51	11	76	76	63	67
Bayes	75	60	6	52	77	60	66
BEST BASIS METHOD							
k-NN	55	26	0	45	67	8	45
SVM	94	6	6	2	100	0	56
Bayes	68	51	0	65	71	68	63
STRUCTURAL BEST BASIS METHOD							
k-NN	54	4	0	6	67	42	41
SVM	92	92	0	8	95	0	67
Bayes	0	0	100	0	0	0	4
BEST BASIS CENTROID METHOD							
k-NN	56	43	17	56	53	53	51
SVM	60	51	0	53	76	60	61
Bayes	75	51	0	65	79	95	71
PYRAMIDAL WAVELET TRANSFORM							
k-NN	64	44	6	48	53	17	48
SVM	53	51	6	61	69	52	57
Bayes	84	56	0	47	90	47	69
HISTOGRAM METHOD							
k-NN	81	82	83	87	89	80	84
FOURIER METHOD							
k-NN	53	47	11	52	59	43	51
SVM	70	53	0	61	77	60	64
Bayes	87	81	28	89	94	88	86

tion regarding the 2-classes case as well as the 6-classes case.

Multiclassifier

Since the multichannel features outperformed the single channel features as pointed out above, the multiclassifier has been tested using the multichannel feature based methods only.

From Table 7 we see that in the 2-classes case the multiclassifier outperforms all methods except the Fourier method and the BBCB method. Both superior methods are using the Bayes classifier in this case and reach 98 % and 100 %, respectively, compared to 98 % overall classification accuracy reached by the multiclassifier.

Regarding the 6-classes case, the multiclassifier delivers a considerably better overall classification result of 94 % compared to the single

Figure 7. The Fourier filters in the 2-classes case for the R, G and B channel (from left to right)

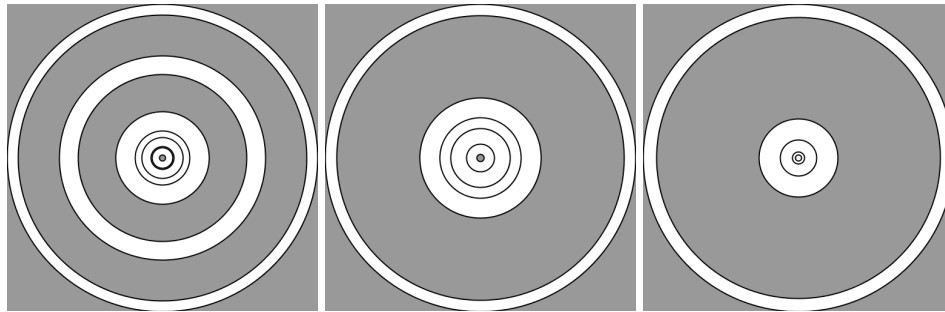
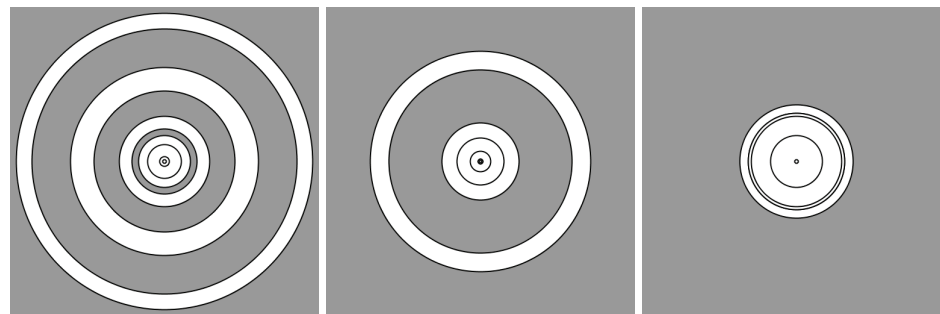


Figure 8. The Fourier filters used in the 6-classes case for the R, G and B channel (from left to right)



methods, which reach an overall classification result between 4 % and 86 % for the BBS method and the Fourier method, respectively, both using the Bayes classifier.

DISCUSSION AND FUTURE TRENDS

To highlight the clinical relevance of the results obtained throughout this work, a comparative meta-study of Kato, Fu et al. (2006) should be mentioned. This study showed, that regarding the 2-classes case, the overall (human) classification accuracy of magnifying colonoscopy based on the pit pattern scheme varies between approximately 80 % and 99 %. The multiclassifier presented in this work delivers a classification accuracy of 98

% in the 2-classes case which ranges among the top-values achieved by physicians documented in this study. The Fourier method in the multichannel case is able to deliver almost the same accuracy as well. This clearly shows that the discussed approach can be a valuable decision support technique in clinical usage.

One possible way to further improve the classification results would be to map features into another, better suitable feature space using Linear discriminant analysis or Principal component analysis. Apart from that, additional preprocessing to the images may be applied. Up to now no particular preprocessing has been performed except for the histogram method, where a Gaussian blur has been used. Although this has been tried with the rest of the methods too, there was no significant gain in classification performance.

Table 7. Overall classification results in percent using the multiclassifier based on methods using multichannel features

Pit Type	I	II	III-S	III-L	IV	V	Total
2 classes	94		100				98
6 classes	100	94	83	81	97	92	94

CONCLUSION

In this work we show that automated pit pattern classification is feasible using general purpose texture classification techniques. However, optimal results are achieved after careful selection and optimization of classification parameters only, except for the histogram-based techniques which deliver satisfying results in an almost ad-hoc manner. It has turned out that classification results can be enhanced when considering all the color information stored in an image.

Apart from that we showed one possible way to combine different methods into one multiclassifier. The proposed multiclassifier is able to yield significantly better results in the 6-classes case as compared to the single classification techniques (while in the 2-classes case a slight degradation of the best results is found). The classification accuracy observed for the best standalone multichannel techniques and the multiclassifier already qualify the approach as an interesting choice for a decision support system for clinical usage.

ACKNOWLEDGMENT

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

REFERENCES

- Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2), 121-167.
- Chang, C., & Lin, C. (2001). LIBSVM: A Library for Support Vector Machines. Retrieved April 1, 2005, from <http://www.csie.ntu.edu.tw/~cjlin/>
- Coifman, R. R., & Wickerhauser, M. V. (1992). Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2), 713-719.
- Fukunaga, K. (1990). *Statistical Pattern Recognition*. Morgan Kaufmann.
- Hsu, C., Chang, C., & Lin, C. A Practical Guide to Support Vector Classification.
- Häfner, M., Liedlgruber, M., Wrba, F., Gangl, A., Vécsei, A., & Uhl, A. (2006a). Pit pattern classification of zoom-endoscopic colon images using wavelet texture features. *Proceedings of the International Conference on Advances in Medical Signal and Image Processing (MEDSIP 2006)*.
- Häfner, M., Kendlbacher, C., Mann, W., Taferl, W., Wrba, F., Gangl, A., Vécsei, A., & Uhl, A. (2006b). Pit pattern classification of zoom-endoscopic colon images using histogram techniques. *Proceedings of the 7th Nordic Signal Processing Symposium (NORSIG 2006)*, 58-61.
- Häfner, M., Brunauer, L., Payer, H., Resch, R., Wrba, F., Gangl, A., et al. (2007). Pit pattern classification of zoom-endoscopic colon images

using evolved Fourier feature vectors *Proceedings of the 2007 IEEE Machine Learning for Signal Processing Workshop (MLSP'07)*, 99 - 104.

Hsu, C., & Lin, C. (2002). A Comparison of Methods for Multi-Class Support Vector Machines. *IEEE Transactions on Neural Networks*, 13(2), 415-425.

Hurlstone, D. P. (2002). 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.

Hurlstone, D. P., Cross, S. S., Adam, I., Shorthouse, A. J., Brown, S., Sanders, D. S., & Lobo, A. J. (2004). 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.

Karkanis, S., Iakovidis, D., Maroulis, D., Theofanis, N., & Magoulas, G. (2000). Tumor Recognition in Endoscopic Video Images using Artificial Neural Network Architectures. *Proceedings of the 26th EUROMICRO Conference (EUROMICRO'00)*, 423-429.

Karkanis, S., Iakovidis, D., Karras, D., & Maroulis, D. (2001). Detection of Lesions in Endoscopic Video using Textural Descriptors on Wavelet Domain supported by Artificial Neural Network Architectures. *Proceedings of the International Conference on Image Processing (ICIP'01)*, 833-836.

Karkanis, S., Magoulas, G., Grigoriadou, M., & Schurr, M. (1999). Detecting Abnormalities in Colonoscopic Images by Textural Description and Neural Networks. *Proceedings of the Workshop on Machine Learning in Medical Applications, Advance Course in Artificial Intelligence (ACAI'99)*, 59-62.

Kato, S., Fu, K., Sano, Y., Fujii, T., Saito, Y., Matsuda, T., et al. (2006). Magnifying colonoscopy as a non-biopsy technique for differential diagnosis of non-neoplastic and neoplastic lesions. *World journal of gastroenterology : WJG*, 12(9), 1416-20.

Kudo, S., Hirota, S., Nakajima, T., Hosobe, S., Kusaka, & H., Kobayashi, T. (1994). Colorectal tumours and pit pattern. *Journal of Clinical Pathology*, 47(10), 880-885.

Kudo, S., Tamura, S., Nakajima, T., Yamano, H., Kusaka, H., & Watanabe, H. (1996). Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointestinal endoscopy*, 44(1), 8-14.

Liedlgruber, M., & Uhl, A. (2007). Statistical and structural wavelet packet features for Pit pattern classification in zoom-endoscopic colon images (P. Dondon, V. Mladenov, S. Impedovo, & S. Cepisca, Hrsg.). *Proceedings of the 7th WSEAS International Conference on Wavelet Analysis & Multirate Systems (WAMUS'07)*, 147-152.

Maroulis, D., Iakovidis, D., Karkanis, S., & Karras, D. (2003). CoLD: a versatile detection system for colorectal lesions in endoscopy video-frames. *Computer methods and programs in biomedicine*, 70(2), 151-66.

Rajpoot, N. (2003). Local Discriminant Wavelet Packet Basis for Texture Classification. *Proceedings of the International Society for Optical Engineering SPIE Wavelets: Applications in Signal and Image Processing X*, 774-783.

Rajpoot, K., & Rajpoot, N. (2004). Wavelets and support vector machines for texture classification. *Multitopic Conference, 2004. Proceedings of INMIC 2004. 8th International*, 328-333.

Saito, N. (1997). Classification of Geophysical Acoustic Waveforms and Extraction of Geological Information Using Time-Frequency Atoms. *1996 Proceedings of the Computing Section of the American Statistical Association*, 322-327.

Saito, N., & Coifman, R. R. (1995). Local Discriminant Bases and their Applications. *J. Mathematical Imaging and Vision*, 5(4), 337-358.

Tung, S., Wu C-S, & Su, M. (2001). Magnifying Colonoscopy in Differentiating Neoplastic From Nonneoplastic Colorectal Lesions. *American Journal of Gastroenterology*, 96, 2628-2632.

KEY TERMS

Classification Feature: A numerical or syntactical value used to describe an observed property of an object (e.g., size, color, shape, ...).

Classification Feature Vector: A collection of classification features describing the properties of an object.

Classifier: An algorithm to assign unknown object samples to their respective classes. The decision is made according to the classification feature vectors describing the object in question.

Colonoscope: A flexible, lighted instrument used to examine the inside of the colon.

Colonoscopy: A medical procedure during which a physician is examining the colon for polyps using a colonoscope.

Color Histogram: A graphical representation of a distribution of colors within an image. The data contained in a histogram is obtained by counting the occurrence of each possible color of the respective color model within the image.

Fourier Transform: An algorithm used to decompose a signal (e.g., an image) into its frequency components and to compute the frequency spectrum for a given signal.

Wavelet Transform: A transform used to decompose a signal into its frequency components, similar to the Fourier transform. But the time-frequency resolution of the wavelet transform can be adjusted since basis functions with compact support are used, in contrast to the Fourier transform, where sine and cosines are used as basis functions.