

# Computer-Aided Tumor Detection in Endoscopic Video Using Color Wavelet Features

Stavros A. Karkanis, *Member, IEEE*, Dimitris K. Iakovidis, Dimitris E. Maroulis, *Member, IEEE*,  
Dimitris A. Karras, *Associate Member, IEEE*, and M. Tzivras

**Abstract**—We present an approach to the detection of tumors in colonoscopic video. It is based on a new color feature extraction scheme to represent the different regions in the frame sequence. This scheme is built on the wavelet decomposition. The features named as color wavelet covariance (CWC) are based on the covariances of second-order textural measures and an optimum subset of them is proposed after the application of a selection algorithm. The proposed approach is supported by a linear discriminant analysis (LDA) procedure for the characterization of the image regions along the video frames. The whole methodology has been applied on real data sets of color colonoscopic videos. The performance in the detection of abnormal colonic regions corresponding to adenomatous polyps has been estimated high, reaching 97% specificity and 90% sensitivity.

**Index Terms**—Color texture, computer aided colonoscopy, image analysis, medical imaging, polyp detection, wavelet features.

## I. INTRODUCTION

COLORECTAL cancer is the second leading cause of cancer-related deaths in the United States [1], [2]. More than 130 000 people are diagnosed with colon cancer each year and about 55 000 people die from the disease annually. Colon cancer can be prevented and cured through early detection, so early diagnosis is of critical importance role for patient's survival. Screening is the current and most suitable prevention method for an early detection and removal of colorectal polyps. If such polyps remain in the colon, they can possibly grow into malignant lesions. Colonoscopy is an accurate screening technique for detecting polyps of all sizes, which also allows for biopsy of lesions and resection of most polyps [3]. The colonic mucosal surface is granular and demarcated into small areas called nonspecific grooves. Changes in the cellular pat-

tern (pit pattern) of the colon lining might be the very earliest sign of polyps. Pit patterns can be used for a qualitative and quantitative diagnosis of lesions. These textural alterations of the colonic mucosal surface can also be used for the automatic detection of colorectal lesions [4]–[6].

The scope of this work is the location of regions suspicious for malignancy in video colonoscopy, regions that require more thorough examination by medical experts for a second evaluation. Tumor detection schemes using textural information have been proposed for various tissues such as liver [7], prostate [8], breast [9], brain [10], cervix [11], and cardiac [12]. Automated classification and identification of colonic carcinoma using microscopic images and involving texture analysis compared with geometric features based on statistical analysis has been proposed by Esgiar *et al.* [13], [14]. The use of endoscopic video frames for the identification of adenomatous polyps involving a novel wavelet based color texture analysis scheme is a topic that it has not been reported in the literature to the best of our knowledge. In the proposed approach the video frame sequences are transformed in scale and frequency by using the wavelet transform since it has been observed that the textural information is localized in the middle frequencies and lower scales of the original signal [15]. Statistical color wavelet features have been encountered in this texture analysis scheme, for the discrimination of normal and abnormal (i.e., tumor) regions. The construction of the texture feature space follows the multiresolution approach on the color domain. The resulted space is found to be discriminant. A linear classification scheme was used to label image regions with a low error rate. The novel proposed color wavelet textural features are favorably compared to the rival approach of wavelet correlation signatures [16].

The proposed detection scheme involves a) a novel feature extraction technique based on a discrete wavelet decomposition applied on different color spaces and b) statistical analysis of the wavelet coefficients associated with the color bands. The wavelet features are based on second-order textural information estimated on the domain of the discrete wavelet decomposition of each color band of a video frame. In this paper, the textural characteristics estimated on the color discrete wavelet frame transform and in the sequel processed by using correlational analysis, give valuable information about the set of features that produce the most discriminant subspaces for normal/abnormal tissue regions. The proposed scheme was tested on real data sets of color colonoscopic videos provided by the Gastroenterology Section, Department of Pathophysiology, Medical School, University of Athens, Greece, and partially by the Section for Minimal Invasive Surgery, University of Tübingen, Germany. The

Manuscript received March 6, 2002. This work was supported in part by the Special Account of Research Grants, National and Kapodestrian University of Athens.

S. A. Karkanis was with Realtime Systems and Image Analysis Group, Department of Informatics and Telecommunications, University of Athens. He is now with the Department of Informatics and Computer Technology, Technological Educational Institute of Lamia, Lamia 35100, Greece (e-mail: sk@teilam.gr).

D. K. Iakovidis and D. E. Maroulis are with Realtime Systems and Image Analysis Group, Department of Informatics and Telecommunications, University of Athens, 15784 Athens, Greece (e-mail: rtsimage@di.uoa.gr; rtsimage@di.uoa.gr).

D. A. Karras is with Hellenic Aerospace Industry, Schematari, Greece (e-mail: dkarras@haicorp.com).

M. Tzivras is with Gastroenterology Section, Department of Pathophysiology, Medical School, University of Athens, Athens 11527, Greece (e-mail: dkarras@haicorp.com).

Digital Object Identifier 10.1109/TITB.2003.813794

video sequences used for evaluation were selected to contain relatively small polyps, as physicians suggested. The sequences were evaluated by endoscopy experts and compared with the corresponding histological results, proving the accuracy of the proposed methodology (this evaluation procedure with respect to the histological data led to specificity ranging from 86% to 98% and sensitivity ranging from 79% to 96.5%).

The rest of the paper is organized as follows. Medical information on colorectal polyps is provided in Section II. In Section III, the fundamental properties of color and texture analysis involved along with the proposed methodology are presented. Section IV describes the evaluation approach and the results obtained from the extensive experimentation. Finally, discussion of the results as well as the conclusions of this study is presented in Sections V and VI, respectively.

## II. MEDICAL BACKGROUND

A polyp is defined as any visible tissue mass protruding from the mucosal surface. Polyps are characterized according to their color, appearance of their mucosal surface, presence of ulcers, their bleeding tendency, and above all the presence of pedunculus (pedunculated or nonpedunculated). Their size varies from barely visible transparent protrusions to pedunculated lesions with a diameter of 3 to 5 cm. Although there are many histopathologic types of polyps, the majority of them are adenomatous. Approximately 75% of the colonic polyps are adenomatous [17]. Adenomatous polyps are neoplasms that result from disordered cell proliferation, differentiation, and apoptosis [18]. The evolution of an adenomatous polyp to cancer is the result of a multistep process that involves many molecular and genetic mechanisms including activation of oncogenes and suppression of tumor genes [19]. The real prevalence of colonic polyps in the general population is not known. Polyps may be found in the colon of 30%–50% of people older than 55 years old, while colonoscopy surveys showed a lower incidence, at the level of 30% [20]. Today, the international consensus for the treatment of polyposis dictates removal of all polyps, regardless of the location, size or other characteristics, in order to prevent a possible development to cancer. Colonoscopy remains the best available procedure to detect polyps, with many advantages such as the ability to have simultaneous tissue biopsy or polypectomy [3]. A competitive new generation technique used for the detection of colorectal polyps is virtual colonoscopy based on computer tomography (CT) or magnetic resonance (MR) data. This technique utilizes specialized imaging software that allows for a three-dimensional visualization of the colon and the rectum by combining multiple volumetric tomographic data [21]–[24]. It has the advantage that it does not discomfort the patients as the standard colonoscopy, but it is not so accurate for the detection of small lesions and it can not easily discriminate polyps among retained stool or thickened folds because they can mimic their shape and density and does not allow for tissue biopsy or polypectomy [23], [24].

Important research on the automated detection of polyps on virtual colonoscopy data has been reported in the recent literature. Most of this research was concentrated on the use of geo-

metric features for the discrimination of polyps from normal colonic regions [24]–[27].

## III. COLOR TEXTURE ANALYSIS

Color texture analysis is based on the combined information from both color and texture fields of the image. Texture processing was mainly focused on the use of gray-level image information for a number of years [28], [29]. Pioneering studies exploiting the combination of both color and texture information, have been presented by Caelli and Raye [31], Sharkanski *et al.* [32], and Kondepudy *et al.* [33]. More recent studies involving color texture analysis, include the calculation of chromaticity moments [34], a perceptual approach for the segmentation of color textures [35], Gabor filtering of complex hue/saturation images [36], moving average modeling [37] and color and texture fusion by combining color and multiresolution simultaneous autoregressive models [38]. Drimborean and Whelan [39] performed experiments using grayscale and color features based on discrete cosine transform, Gabor and cooccurrence matrices in different color spaces. The results of this study led to the conclusion that the introduction of color information, especially by calculating grayscale texture features on the different color channels, significantly improves color texture classification. Other approaches that have taken into account the correlation of texture measures between the different color channels, have shown that color texture information can also be found in the way color channels are related to each other. Under this framework Paschos [40] proposed a set of discriminative and robust chromatic correlation features using directional histograms, Van de Wouwer *et al.* [41] achieved high classification results using correlation signatures calculated on the wavelet coefficients of the different color channels of the images and Vandenbroucke *et al.* [42] exploited the correlation of first-order statistical features among the different color channels for unsupervised soccer image segmentation. In this work, we propose the covariance of second-order statistical features in the wavelet domain for the characterization of colonic polyps.

### A. Color Spaces

Color is a property of the brain and not of the outside world [43]. The nervous system, instead of analyzing colors, uses the information of the external environment, namely the reflectance of different wavelengths of light and transforms this information into colors [44]. The use of the red-green-blue (RGB) space is very common in image and video-processing research, dictated primarily by the availability of such data as they are produced by most color image-capturing devices. Drawbacks in the use of RGB in computer vision applications are: the high correlation among RGB channels for natural images [45], the representation of RGB is not very close to the way humans perceive colors [46] and it is not perceptually uniform [47].

In RGB space, each color is represented as a triple (R, G, B), where R, G, and B represent red, green, and blue signals corresponding to different wavelengths of the visible spectrum. Assuming dichromatic reflection and white illumination, a color transform that is independent of the viewpoint, surface orientation, illumination direction, and illumination intensity, has been

proposed by Gevers [48]. A first-order invariant instantiation of this transform, which is also more robust to noise comparing to other invariant instantiations, has proven to be the normalized RGB space (Appendix I-A). Normalized *rgb* has been used for automatic lip reading [49] and other face detection applications [52].

A variety of other color spaces are used in different applications. The international committee on colorimetry, Commission Internationale de l'Eclairage (CIE), established the XYZ color space as standard, based on the response curves of the eyes and statistics that were performed on human observers [47]. Normalizing the XYZ (Appendix I-B), the occurring *xyY* space, has been proven to be noise robust for texture recognition using chromatic correlation features [40]. All of the above color spaces have the advantage of isolating the luminance component *Y* from the two-chrominance components [53].

The Karhunen–Loeve (K–L) transformation applied on images, has been proved to be best for color texture characterization as reported by Van De Wouwer *et al.* [41], and for the analysis of skin lesions [54]. K-L transform is formed by the eigenvector of the correlation matrix of an image, which remains approximately the same for a large set of natural color images [41], [55]. It transforms an image to an orthogonal basis in which the axes are statistically uncorrelated. In that sense, the information presented in RGB space is decorrelated. Practically, it that can be produced as a linear transformation of the RGB coordinates (Appendix I-C).

Perceptual uniformity has been considered to form color spaces that describe color similarity to the way humans perceive color. Generally, a system is perceptually uniform if a small perturbation to a component value is approximately equally perceptible across the range of that value [47]. CIE-Lab is a perceptually uniform color space that has proved to perform better than *RGB* for color texture analysis, but not in the presence of noise [53]. It has been applied for several color texture classification tasks such as: the retrieval of color patterns using textural features [56], analysis of skin lesions [54] and segmentation of human flesh [57], with a performance that has been considered high. The coordinates of CIE-Lab as a function of *R, G, B* are given in Appendix I-D.

Another, approximately perceptually uniform color space is defined in terms of hue, saturation and value (HSV), a phenomenal color space [58]. Phenomenal color spaces attempt to classify colors in relation to how they are perceived and interpreted by the human brain and they are more “intuitive” in manipulating color. HSV has led to higher classification performance than CIE-Lab and RGB in both noisy and noise-free conditions for color texture analysis [53]. On the other hand, Palm *et al.* [36] showed that HSV performs equivalently to *RGB* for color texture classification using different features. Another common alternative similar to HSV is hue, lightness, saturation (HLS) space [46], [59]. HLS has been applied to represent the color of the tongue for medical diagnosis [60].

### B. Second-Order Statistics on the Wavelet Domain as Grayscale Textural Features

As it has already been noted, the size of the lesions to be detected using the proposed framework varies. The image resolution cannot be defined so as to cover the majority of the lesions sizes. It will be useful to face the problem in a way that detects

the information in different resolutions by exploiting the intermediate scales for the final decision. Multiresolution analysis of an image can be achieved by using the discrete wavelet transform.

Texture is the discriminating information that differentiates normal from abnormal lesions [4]–[6]. Since texture is essentially a multiscale phenomenon, multiresolution approaches such as wavelets perform well for texture analysis. A characterization of texture is usually based on the local information that appears within a neighborhood distribution of the gray levels. The proposed methodology focuses on a single scale in order to extract the relevant information. Recent studies have come to the conclusion that a spatial/frequency representation, which preserves both global and local information, is adequate for the characterization of texture. The wavelet transform offers a tool for spatial/frequency representation by decomposing the original images to the corresponding scales. When decomposition level decreases in the spatial domain, it increases in the frequency domain providing zooming capabilities and local characterization of the image. Since the low-frequency image produced by the transformation does not contain major texture information and the most significant information of a texture often appears in the middle-frequency channels, we choose to use discrete wavelet transform (DWT) for the decomposition of the frequency domain of the image [61], [16], [63]. Wavelet frame representation of the image offers a representation of the frequency domain. Such representations have been proposed because they have greater robustness in the presence of noise, can be sparser, and can have greater flexibility in representing the structure of the input data. The dimensionality and the representation of input is not a unique combination of basis vectors. The two-dimensional (2-D) DWT transformation is implemented by applying a separable filterbank to the image [64].

This filtering procedure convolves the image with a lowpass *H* and bandpass filter *G*, which produces a low-resolution image  $L_n$  at scale *n* and the detail images  $D_1, \dots, D_{3n}$  at scale *n*. The repetition of this filtering procedure ( $k = 1 \dots n$ ) results in a decomposition of the image at several scales. The final set consisting of the low resolution image  $L_n$  and all the detailed images  $D_1, \dots, D_{3n}$  along the scale is the multiscale representation of the image at a specific depth defined by the total number of scales. This filtering procedure can be described by the following recursive equations [16]:

$$\begin{aligned} L_k(b_i, b_j) &= [H_x * (H_y * L_{k-1})_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \\ D_{3k-2}(b_i, b_j) &= [H_x * (G_y * L_{k-1})_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \\ D_{3k-1}(b_i, b_j) &= [G_x * (H_y * L_{k-1})_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \\ D_{3k}(b_i, b_j) &= [G_x * (G_y * L_{k-1})_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \end{aligned} \quad (1)$$

where the arrow ( $\downarrow$ ) denotes the subsampling procedure, the asterisk ( $*$ ) is the convolution operator, and *H* and *G* are the two filters for all  $k = 1 \dots n$ .

The cooccurrence matrices approach has been considered in this work for the description of a statistical model of the texture encoded within the decomposed subimages. It captures second-order gray-level information, which is mostly related to the human perception and discrimination of textures [65]. For a coarse texture these matrices tend to have higher values

near the main diagonal whereas for a fine texture the values are scattered. The cooccurrence matrices encode the gray level spatial dependence based on the estimation of the second-order joint-conditional probability density function  $f(i, j, d, \theta)$ , which is computed by counting all pairs of pixels at distance  $d$  having gray levels  $i$  and  $j$  at a given direction  $\theta$ . The angular displacement is usually included in the range of the values  $\{0, \pi/4, \pi/2, 3\pi/4\}$ . Among the 14 statistical measures, originally proposed by Haralick [28], [66], that derive from each cooccurrence matrix we consider only four. Namely, angular second moment, correlation, inverse difference moment and entropy

$$F_1 = \sum_i^{N_g} \sum_j^{N_g} p(i, j)^2 \quad (2)$$

$$F_2 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i \cdot j) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (3)$$

$$F_3 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i - j)} p(i, j) \quad (4)$$

$$F_4 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log(p(i, j)) \quad (5)$$

where  $p(i, j)$  is the  $ij$ th entry of the normalized cooccurrence matrix,  $N_g$  is the number of gray levels of the image, and  $\mu_x, \mu_y, \sigma_x$ , and  $\sigma_y$  are the means and standard deviations of the marginal probability  $p_x(i)$  obtained by summing up the rows of matrix  $p(i, j)$ . These measures provide high discrimination accuracy which can be only marginally increased by adding more measures in the feature vector [67].

In addition to the features (2)–(6), Esgiar *et al.* proposed the use of contrast [14] or the use of both contrast (also known as difference moment) and dissimilarity [13], for microscopic image analysis of colonic tissue

$$F_5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 p(i, j) \quad (6)$$

$$F_6 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |i - j| p(i, j). \quad (7)$$

In Section IV, we experimentally show that the use of these features do not provide additional textural information that is significant for the analysis of the macroscopic video images used in our application.

### C. Second-Order Color Wavelet Covariance (CWC) Features

The proposed approach is based on the extraction of color textural features. These features are estimated over the second-order statistical representation of the wavelet transform of the color image. Since each feature represents a different property of the examined region, we consider as valuable information the covariance among the different statistical values between the color channels of the examined region.

According to the definition of texture, it is mainly related to the distribution of the intensities [28], [61]. It is then expected that similar textures will have close statistical distributions and

consequently they should appear to have similar feature values of the features. This similarity property of the selected features can be described by measuring the variance in pairs of them. By using the covariance between two features, we can have a measure of their “tendency” to vary together. The texture covariance has been proposed in the literature [29] as a measure that is used directly on the image intensities or among the color intensities of the examined region. Our method uses the covariance in order to rank the changes in the statistical distribution of the intensities between the examined regions in the different color channels. By noticing the way the features of examined texture regions covary it will be an easy task to decide if they belong to the same texture class since in similar textures we expect measures to covary.

By considering the original image  $I$ , we obtain its color transformation. Color transformations result in three decomposed color channels

$$I^i, \quad i = 1, 2, 3. \quad (8)$$

A three-level discrete wavelet frame transformation is consequently applied on each color channel ( $I^i$ ). This transformation results in a new representation of the original image, according to the corresponding equations of wavelet decomposition (1). This decomposition procedure produces a low-resolution image  $L_n^i$  at scale  $n$  and the detail images  $D_l^i$  and  $l = 1, 2, 3, \dots, 3n$ . In our case, we have

$$I^i \equiv \{L_n^i, D_l^i\}, \quad i = 1, 2, 3, \quad l = 1, 2, 3, \dots, 3n \quad (9)$$

where  $n$  is the decomposition level.

Since the textural information is better presented in the middle wavelet detailed channels, we consider the second level detailed coefficients. Thus, the image representation that is finally considered is the one consisting of the detail images produced from (9) for the values  $l = 4, 5, 6$ . This results in a set of nine different subimages

$$\{D_l^i\}, \quad i = 1, 2, 3, \quad l = 4, 5, 6. \quad (10)$$

For the extraction of the second-order statistical textural information, we use cooccurrence matrices calculated over the above nine different subimages. These matrices reflect the spatial interrelations between the intensities within the wavelet decomposition level. The cooccurrence matrices are estimated in four different directions of intensities' relation,  $0^\circ, 45^\circ, 90^\circ$ , and  $135^\circ$ , resulting to 36 matrices

$$C_a(D_l^i), \quad i = 1, 2, 3, \quad l = 4, 5, 6, \quad \alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ. \quad (11)$$

Finally the four statistical measures, namely angular second moment, correlation, inverse difference moment, and entropy are estimated for each matrix resulting in 144 wavelet features

$$F_m(C_a(D_l^i)), \quad i = 1, 2, 3, \quad l = 4, 5, 6 \\ \alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ, \quad m = 1, 2, 3, 4 \quad (12)$$

where  $m$  is the respective statistical measure.

In the proposed scheme, we consider as a textural measure  $CWC_m^l(i, j)$  the covariance of the same statistical measure be-

tween color channels  $i, j$  at wavelet band  $l$  which is defined according to the following equations:

$$\begin{aligned} & \text{Cov} \left[ F_m \left( C_a \left( D_l^i \right) \right), F_m \left( C_a \left( D_l^j \right) \right) \right] \\ &= \sum_a \left[ F_m \left( C_a \left( D_l^i \right) \right) - E \left( F_m \left( C_a \left( D_l^i \right) \right) \right) \right] \\ & \quad \times \left[ F_m \left( C_a \left( D_l^j \right) \right) - E \left( F_m \left( C_a \left( D_l^j \right) \right) \right) \right] \\ & \text{CWC}_m^l(i, j) \\ &= \begin{cases} \text{Cov} \left[ F_m \left( C_a \left( D_l^i \right) \right), F_m \left( C_a \left( D_l^j \right) \right) \right], & i < j \\ \text{Var} \left[ F_m \left( C_a \left( D_l^i \right) \right) \right], & i = j \end{cases} \end{aligned} \quad (13)$$

where  $\alpha$  represents the different angles for the cooccurrence matrices  $\alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ .

Since the covariance (13) relates pairs of features, the proposed set of features is a set of 72 components. The 36 of them are the variances as they relate features of the same color channel and the rest 36 represent features of different color channels estimated by the corresponding covariance values. We call this set of the 72 components color wavelet covariance features, the CWC feature vector.

The extraction of the CWC vector can be described in the following steps.

- The original color image (video frame) is decomposed into three separate color bands.
- Each band is scanned across with fixed size sliding square window.
- Each window is then transformed according to a three-level 2-D discrete wavelet transform by using decomposition functions that follow the properties of the wavelet frames. The detail coefficients of the middle decomposition level are considered for further processing. This step results to a set of nine subimages.
- The cooccurrence matrices, for each image of the previous step, are estimated into four directions, producing 36 matrices that are a second-order statistical representation of the original image.
- Four statistical measures (angular second moment, entropy, inverse difference moment, and correlation) are calculated for each matrix, resulting in a set of 144 components. Each of the measure carries different information about the texture.
- Covariance values of pairs of the estimated features (e) constitute the 72-dimensional CWC feature vector, to be used for the classification of the image regions (windows).

#### IV. EXPERIMENTS AND RESULTS

The experimental study of this paper outlines the series of the conducted experiments and the obtained results in order to evaluate the proposed novel feature-extraction methodology, along with its associated parameters in the problem of tumor detection using color colonoscopic video sequences.

##### A. Data Acquisition and Processing

The colonoscopic data used in the following experiments was acquired from different patients with an Olympus CF-100 HL

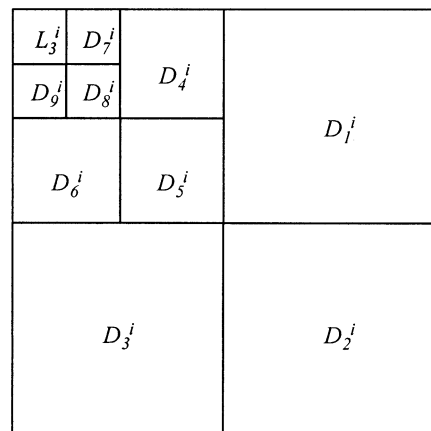


Fig. 1. Three-level wavelet decomposition scheme of the original image for color channel  $i$ .

TABLE I  
HISTOLOGICAL CHARACTERIZATION OF THE AVAILABLE DATASET

Sample	Absolute Population	Percentage (%)	Percentage (%)
Patients	66		
Polyps	95	100	
Hyperplastic Adenomas	35	36.8	
	60	63.2	100
Benign	27		45
Malignant	33		55

endoscope. The major interest for the tumor detection problem, as the experts have suggested it, has led us to the use of video frames mainly of small size adenomatous polyps. Since they are not easily detectable, they are more common and more likely to become malignant compared to the hyperplastic polyps [17].

Sixty-six patients having relatively small polyps were examined within a period of eight months. The results of the histological evaluation of these polyps are presented in Table I [62]. The mean diameter of the adenomatous polyps was estimated to be  $5.66 \pm 0.63$  mm. A total number of 60 video sequences corresponding to the different adenomas with a duration ranging between 5 to 10 s, were used for the evaluation of the proposed methodology. The video frame sequences were recorded during the clinical examination of the patients and then digitized by using a commercial RGB-color frame grabber at a rate of 25 frames per second, a resolution of  $1 \text{ K} \times 1 \text{ K}$  pixels and 24 bits per pixel color depth (eight bits for each color channel). Each one of these video frame sequences, selected by the physician as indicative cases, illustrates small size lesions of interest at different position, scale and lighting conditions (Fig. 2). In the experiments outlined in the following section training and test set of frames have been considered. The training set comprised of 180 frame images (up to three frames per video sequence) shown by the experts group. The selection of the frame images to be incorporated in the training set has been very carefully performed by the experts group in order to minimize the bias introduced in the training procedure. Each one of the five members of the expert group has independently selected 200 image frames as representative of the image frames encountered in the normal subject of the video-sequences as well as 400 image frames as

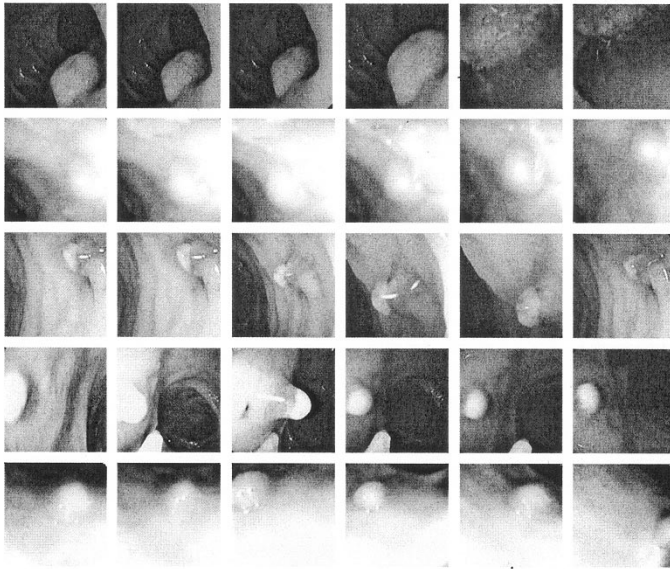


Fig. 2. Representative sample of the dataset.

representative of the different types of polyps encountered in the total of 66 patient corresponding video sequences. In the sequel, considering the different sets of frame images obtained by each expert, an automated statistical analysis has been performed in order to select a final set of training frame images achieving a very high interrater agreement (Spearman's correlation coefficient,  $r = 0.94$ ,  $P < 0.001$ ). Such an inter-rater agreement could be safely considered to lead to the construction of a training set of image frames with reduced bias introduced in the evaluation procedure [71]–[73]. The test set used for evaluation of the recognition performance was comprised of 1200 (up to 20 frames per video sequence) randomly selected frames from the video sequences. In order to improve the reliability of our experimentation we have chosen nonoverlapping training and test sets.

A 1.4 GHz Pentium IV processor-based workstation with 512 Mb RAM was used for the video processing and the execution of the previously described algorithms. A special purpose software suite implementing these algorithms was developed on Microsoft Visual C++, and many modules incorporate calls to Intel Performance Library functions [71], which provides optimum performance for Intel Pentium processors.

### B. Experiments

The experimental procedure generated a large volume of results that can be classified into the following five categories:

- 1) Benefits of the second-order statistics on the wavelet domain of grayscale endoscopic video frames.
- 2) Comparison of the second-order CWC features with color correlation signatures.
- 3) Determination of the most suitable color space transformation, which enhances the textural properties of the colonic mucosal surface and increases detection accuracy.
- 4) Selection of the least correlated second-order CWC features for the tumor detection problem.

- 5) Effect of the size of the training set to the generalization performance of the proposed methodology.

The whole experimentation procedure was based on two major criteria. The first criterion is related to the classification task and the second is related to the evaluation of system's performance.

The classification task is based on stepwise linear discriminant analysis (LDA). It is a simple model involving a minimum set of parameters and has been used in medical decision support tasks providing increased sensitivity [75], [13]. The utilization of a more complex classifier would increase the number of parameters associated with the evaluation of the proposed feature set. The details of the application of LDA include the use of Fisher's function coefficients and computation of the prior probabilities from group sizes,  $P = 0.0001$  and  $F$  statistic for insertion/remove variable has been set at 3.87. As in many medical applications, the data sets consisting of normal and abnormal regions are highly unbalanced [78]. In our experimentation the proportion of abnormal to normal patterns for each of the available frames is about 10% average. Instead of measuring the accuracy, i.e., the rate of successfully recognized patterns, more reliable measures for the evaluation of the classification performance can be achieved by using the sensitivity (true positive rate) and the specificity (100 minus false positive rate) measures [76], [77]. These two measures can be calculated by the following formulas:

$$\text{Sensitivity} = \frac{d}{c+d} \cdot 100(\%) \quad (14)$$

$$\text{Specificity} = \left(100 - \frac{b}{a+b} \cdot 100\right) (\%) \quad (15)$$

where  $a$  is the number of the true negative patterns,  $b$  is the number of the false positive patterns,  $c$  is the number of the false negative patterns, and  $d$  is the number of the true positive patterns. The classification performance is high when both sensitivity and specificity are high, in a way that their tradeoff favors true positive or false positive rate depending on the application. In the following paragraphs, we summarize the results on the above five categories.

1) *Second-Order Statistics on the Wavelet Domain of Grayscale Endoscopic Video Frames*: Primarily, the color video frames were transformed to eight-bit intensity maps and the optimal window size for the detection of polyps was investigated. Each frame is raster scanned by a sliding window, with a step of eight pixels in order to ensure detailed scanning since the regions that possibly contain lesions are expected to be small. A three-level wavelet frame transform was applied on each window. The size of the cooccurrence matrix was set at  $64 \times 64$ , since the classification performance does not improve significantly for larger sizes. According to the second-order statistics on the wavelet domain methodology and (2)–(7), the total number of the gray level features used is 72 (six cooccurrence measures  $\times$  3 wavelet bands  $\times$  4 directions). This feature set was analyzed by using Pearson correlational analysis [79], which can be used as a classifier-independent feature selection method, by discarding the features with absolute correlation exceeding a given threshold [80]. The analysis showed that contrast (6) and dissimilarity were highly correlated to the

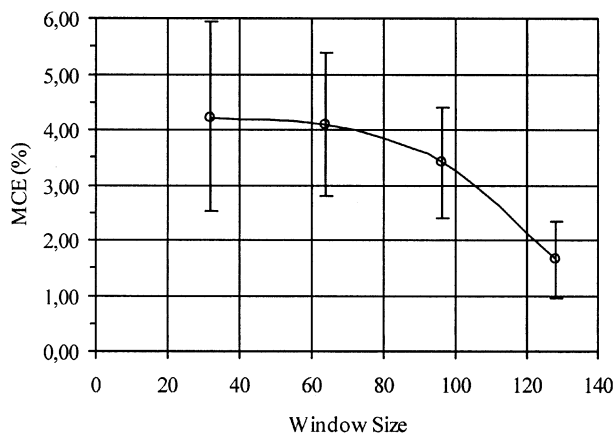


Fig. 3. MCE with respect to the window size.

inverse difference moment (4) exceeding 90% in all wavelet bands. Similarly, high correlation exceeding 85% was also observed between contrast and entropy. The features (2)–(5) were selected as the least correlated in all wavelet bands, having a correlation less than 71% on average. The resulted feature set consists of a total of 48 features (four cooccurrence measures  $\times$  3 wavelet bands  $\times$  4 directions).

Different window sizes, including  $32 \times 32$ ,  $64 \times 64$ ,  $96 \times 96$ , and  $128 \times 128$ , were tested to determine which one provides the lowest mean classification error rate (MCE), estimated on the whole population of the available video frames. The resulted MCE for the various window sizes, is illustrated in Fig. 3. MCEs are depicted on the vertical and the window sizes on the horizontal axis respectively. The error bars correspond to the uncertainty estimated in terms of standard deviation. The lowest average MCE and uncertainty ( $1.65 \pm 0.7\%$ ) correspond to a window size of  $128 \times 128$ , which is the one chosen for the experimentation. The smaller the window size set, the higher the MCE and the uncertainty achieved. This indicates that a large population of pixels is required to characterize tumor regions using second-order statistical features on the wavelet domain. Such result can be justified if we consider the tumor dimensions within the given images, which in most cases reach or exceed  $128 \times 128$  pixels.

The proposed approach was also tested by omitting the wavelet transform. The second-order statistical features (2)–(5) were calculated directly from the intensity values of the corresponding windows, and the average classification performance was estimated  $97 \pm 1\%$  and  $78 \pm 2\%$  in terms of specificity and sensitivity, respectively. As it is illustrated in Fig. 4, the introduction of wavelets increases significantly both specificity (white column) and sensitivity (gray column) at  $99.3 \pm 0.3\%$  and  $87.5 \pm 0.8\%$ , respectively.

2) *Second-Order CWC Features Versus Color Correlation Signatures*: We compare the proposed second-order CWC features with the corresponding correlation signatures proposed by Van de Wouwer [16], on the *RGB* color space. The latter feature extraction scheme uses the correlation of the wavelet coefficients of the different color bands, while our method involves covariance of textural features on the color wavelet domain. The comparison was held by using the complete set of frames. Fig. 5, illustrates the results of the comparison between the CWC fea-

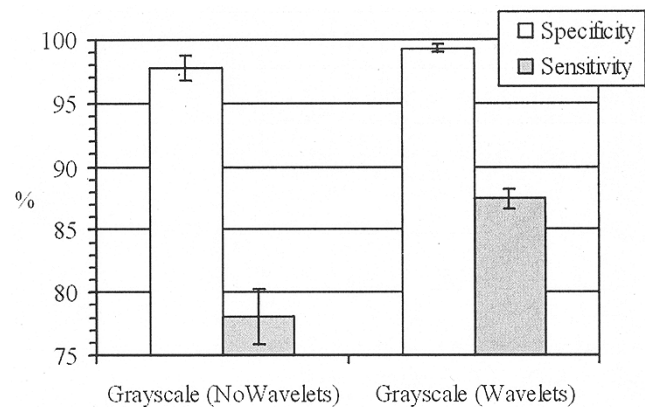


Fig. 4. Comparative results for grayscale video frames.

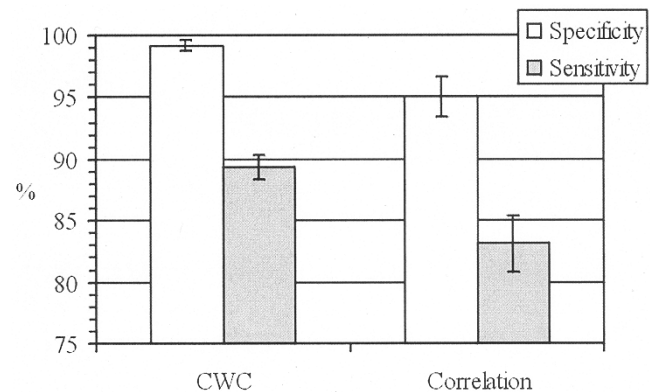


Fig. 5. Comparative results between CWC and correlation signatures.

tures and the color correlation signatures. The specificity and the sensitivity achieved using the CWC features reached  $99.2 \pm 0.4\%$  and  $89 \pm 1\%$ , respectively, while the color correlation signatures led to a  $95 \pm 2\%$  specificity and an  $83 \pm 2\%$  sensitivity. The higher average classification performance of the CWC features and the nonoverlapping uncertainty estimates, show that CWC features are more appropriate for the characterization of the tumor regions.

These results also show that the *RGB* CWC features provide improved results compared to the grayscale wavelet domain features (Fig. 4) in terms of sensitivity. Thus, we expect that color contribute to additional information for tumor detection.

3) *Optimal Color Space for CWC Textural Features*: The color video frame sequences were transformed into different color spaces in order to determine the transformation contributing to the highest classification performance. These results are illustrated in Fig. 6.

The specificity is high in all cases and the small perturbations that are present fall within the uncertainty range. The variations of sensitivity are significant, which means that in this case sensitivity should be the selection criterion of the optimal color space for the discrimination of colorectal polyps and healthy tissue. In the following specificity and sensitivity estimates are given in parentheses in the form of (specificity, sensitivity) for each case. Normalized *RGB*, *rgb* ( $98.6 \pm 0.3, 78.5 \pm 0.8\%$ ) and *xyY* ( $99.0 \pm 0.4, 86 \pm 1$ ), resulted the lowest overall sensitivity. *HSV* ( $98.9 \pm 0.5, 91.5 \pm 0.8\%$ ), *HLS* ( $99.2 \pm 0.4, 90.6 \pm 0.7\%$ ) and the perceptually uniform

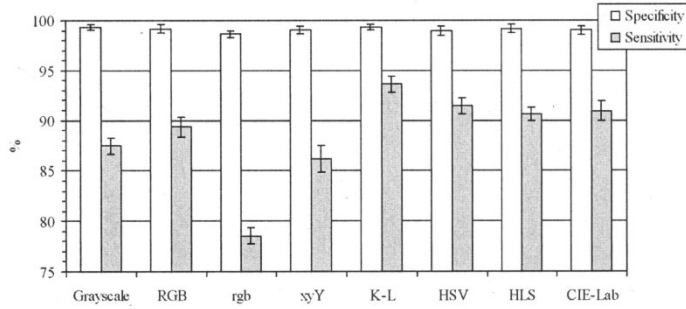


Fig. 6. Comparative results for CWC features in various color spaces.

TABLE II  
HIGHLY CORRELATED CWC FEATURES. THE NOTATION IS ACCORDING TO (13)

CWC Features		Correlation (%)
Correlated Pairs		
$CWC_3^l(2, 2)$	$CWC_3^l(2, 3)$	95-96.5
$CWC_3^l(3, 3)$	$CWC_3^l(1, 2)$	95-96.5
$CWC_3^l(2, 2)$	$CWC_3^l(1, 1)$	90-95
$CWC_3^l(2, 2)$	$CWC_3^l(1, 2)$	90-95
$CWC_4^l(3, 3)$	$CWC_4^l(1, 2)$	80-90
$CWC_4^l(2, 2)$	$CWC_4^l(1, 1)$	80-90
$CWC_4^l(2, 2)$	$CWC_4^l(2, 3)$	80-90
$CWC_4^l(1, 1)$	$CWC_4^l(2, 3)$	80-90
$CWC_4^l(3, 3)$	$CWC_4^l(1, 2)$	80-90
$CWC_1^l(2, 2)$	$CWC_1^l(1, 1)$	80-90
$CWC_1^l(2, 2)$	$CWC_1^l(2, 3)$	80-90
$CWC_2^l(2, 2)$	$CWC_2^l(1, 2)$	80-90

CIE-Lab ( $99.0 \pm 0.4\%$ ,  $90.8 \pm 1\%$ ) color spaces give higher sensitivity than RGB ( $99.2 \pm 0.4\%$ ,  $89 \pm 1\%$ ). The highest overall accuracy was achieved using the *K-L* space, reaching  $99.3 \pm 0.3\%$  specificity and  $93.6 \pm 0.8\%$  sensitivity, which means that its inherent characteristics enhance the covariance of the textural properties of the colonic mucosal surface between color bands. It should be noted at this point that the achieved rates in specificity and sensitivity have been judged as significantly compared with the literature.

4) *Selection of Least Correlated Second-Order CWC Features for Tumor Detection*: Highly correlated features often lead to the degradation of the overall classification performance. It is also worth noting that our major aim is to determine a set of features that produces separable subspaces and do not select the optimal feature subset for the current classifier. In order to investigate the correlation between the features of the *K-L* CWC feature set, we have performed Pearson correlational analysis. The application of correlational analysis showed that the maximum correlation reached 96.5% between the inverse difference moment CWC features listed in Table II in all wavelet bands. Furthermore the angular second moment, entropy and correlation CWC features which are also listed in Table II, are correlated by 80–90% in wavelet bands  $l = 4, 5, 6$ .

Since the left column of Table II consists only of feature variances, it can be concluded that higher correlation is observed between variance and covariance features and not between the different covariance features.

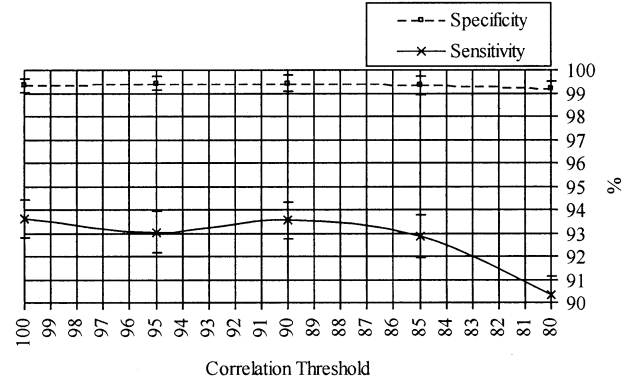


Fig. 7. Selection on CWC features using different correlation thresholds.

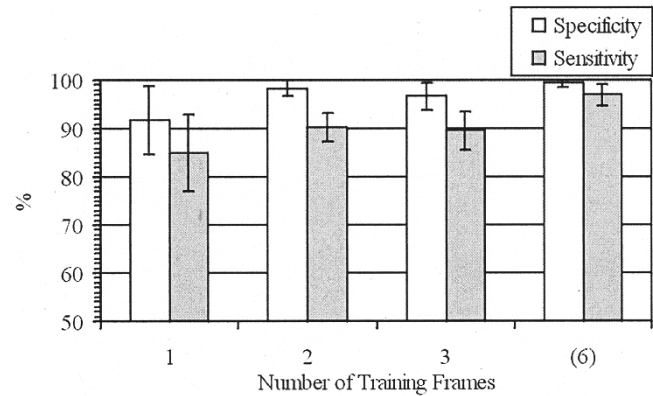


Fig. 8. Effect of the training set to the Generalization performance of the proposed method.

For different correlation thresholds, namely 95%, 90%, 85%, and 80%, discarding the variance features that exceed the threshold, we considered the different feature spaces of different dimensions produced. The classification results illustrated in Fig. 7 show that the sensitivity decreases as the correlation threshold falls below 90%. Discarding all variance features leads to approximately 3% reduction of sensitivity, which means that the textural information contained in variances is not negligible. The fact that the sensitivity at 90% correlation threshold falls within the uncertainty range of the complete set of features (100% correlation threshold), suggests that the first four features (Table II) can be omitted, leading to the reduction of the feature space dimension by nine features, without any harmful implication in the resulted overall sensitivity.

5) *Indicative Experiment on the Generalization Performance of the Proposed Methodology*: From each of the video frame sequences, a set of frames was selected to train/test the linear classifier and have an additional estimation of its generalization performance. Three tests were performed using a different number of frames for training and evaluation. The test results, in terms of average specificity and sensitivity are presented in Fig. 8. The last category of this diagram corresponds to the control case where a set of 20 frames was used for both training and testing.

This diagram shows that two training frames provide slightly better generalization performance ( $97 \pm 2\%$  specificity and  $90 \pm 3\%$  sensitivity), because the average specificity and sensitivity are higher and the corresponding uncertainty ranges are shorter than in training with one frame. Fig. 9, illustrates the recon-



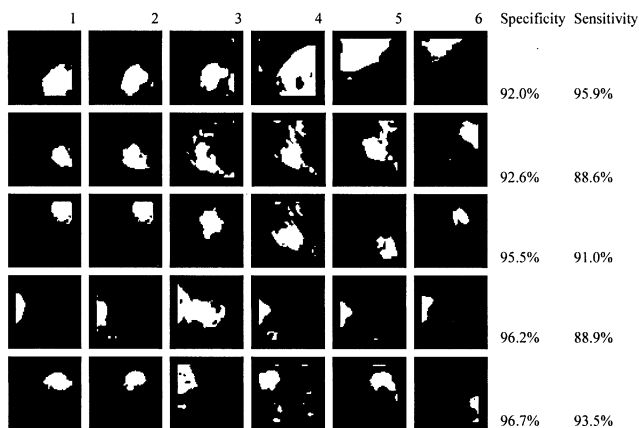


Fig. 9. Classification results for the five different colonoscopic video sequences illustrated in Fig. 2. Average sensitivity and specificity over each set is presented on the left of each row.

structed video frame sequences that correspond to the representative sample of frames of Fig. 2, as they were produced by the classifier's output, by using the first two frames of each set for training and the rest four for testing. White areas correspond to abnormal and black areas to normal regions of the mucosal surface.

## V. DISCUSSION

The results in this paper showed that the use of CWC features led to rather high specificity (14) and sensitivity (15) values, ( $99.3 \pm 0.3\%$  and  $93.6 \pm 0.8\%$ , respectively) estimated on the classified image regions. The validation of the results was experts based since highly experienced physicians (see acknowledgment) reviewed the data in comparison with the histological findings. This gold standard allows us to know if the detected polyps are true polyps or not. Expert endoscopists have defined manually on the original video sequences all the image regions that correspond to polyps and normal tissue. This was done by creating artificial black and white images in which the two possible classes are indicated. These images were used as reference images for the evaluation procedure. Comparing the result of the classifier, which in our case is the discriminant analysis methodology with the characterization of the examined region in the reference image we validated the sensitivity and specificity values. It is worth noting at this point that the expert endoscopists did not utilize any preprocessed image data for review, but they relied only on their reading experience and histological findings. This evaluation procedure is commonly used to similar computer-aided systems for the detection of colorectal polyps [24].

The use of discrete wavelet frame transform contributed to a significant increase of the classification performance by a factor of 2.4% and 12.2% to the values of specificity and sensitivity.

The contribution of the color textural information involved led to additional increase to the value of sensitivity without sacrificing the specificity measure. This increase estimated in a percentage of 7.8% compared to the results obtained using grayscale images. Comparing different color spaces for the detection of tumors we have shown that  $K-L$  color space resulted to the best classification performance with an increase of 5.9% in sensitivity compared to the RGB space.

The proposed methodology can be easily applied in clinical routine. The hardware used for the experiments is a low-cost personal computer with standard configuration. By using such equipment the time performance reaches 1.6 min per selected color video, as the ones used in this work. The algorithm is fully parallelizable and thus it can be executed in parallel on different image regions. The use of special hardware will drastically speed-up the performance of the final system, by a factor that depends on the number of processing elements (PEs) involved. For example, a parallel architecture, which uses 100 of PEs, will accelerate the system by a factor of 100 times and the 1.6 min is estimated to be less than 20 ms. Such implementation could be used during the colonoscopy to increase the physician's capability to detect polyps faster, and thus reduce the duration of the examination, which is rather uncomfortable for the patients. Our group today is working to the direction of the development of such a high-performance embedded system.

## VI. CONCLUSION

We have presented a novel methodology for the extraction of color image features that utilize the covariances of the second-order statistical measures calculated over the wavelet frame transformation of different color bands. It has been applied on the detection of colorectal polyps in colonoscopic video frame sequences, and it has been found that the feature subspaces corresponding to normal and abnormal tissue are highly discriminant. Classification was performed using step-wise LDA and the results of the experimental study have led to the following conclusions.

- 1) The use of second-order statistical features on the wavelet domain results in higher classification accuracy in terms of specificity and sensitivity.
- 2) The proposed CWC features perform significantly better than correlation signatures for tumor detection.
- 3)  $K-L$  was found to be the most suitable color space for the detection of colorectal polyps using CWC features, resulting to a  $99.3 \pm 0.3\%$  and  $93.6 \pm 0.8\%$  specificity and sensitivity, respectively.
- 4) The majority of the proposed CWC features show low correlation, as this has been reached according to the correlational analysis performed.
- 5) The reconstructed images using classifiers output verified that the polyps were well located.

Future extension of this work will be to determine a more robust classification scheme. The overall system could be enhanced under a classifier fusion scheme for the identification of different types of colorectal polyps.

## APPENDIX I

### COLOR TRANSFORMATIONS [47], [59]

#### A. RGB to rgb (Normalized RGB)

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B}$$

where  $r + g + b = 1$ .

### B. RGB to K-L

$$\begin{pmatrix} K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0.333 & 0.333 & 0.333 \\ 0.500 & 0.000 & -0.500 \\ -0.500 & 1.000 & -0.500 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

where  $K_1$ ,  $K_2$ , and  $K_3$  the coordinates of the  $K$ - $L$  color space.

### C. RGB to XYZ

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.412 & 0.357 & 0.180 \\ 0.212 & 0.715 & 0.072 \\ 0.019 & 0.119 & 0.950 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}.$$

### D. RGB to xyY

$$x = X/(X + Y + Z), \quad y = Y/(X + Y + Z).$$

### E. RGB to CIE-Lab

If  $(Y/Y_n > 0.008 856)$  then  $L = 116 (Y/Y_n)^{1/3} - 16$

else  $L = 903.3(Y/Y_n)$

$$a = 500(f(X/X_n) - f(Y/Y_n))$$

$$b = 200(f(Y/Y_n) - f(Z/Z_n))$$

If  $(t > 0.008 856)$  then  $f(t) = t^{1/3}$

else  $f(t) = 7.787t + 16/116$

where  $X_n, Y_n, Z_n$  correspond to the coordinates of a reference white as defined by CIE standard illuminant  $D_{65}$  and are obtained by setting  $R = G = B = 100$  in  $RGB$  to  $XYZ$  transformation, and  $t \in \{X/X_n, Y/Y_n, Z/Z_n\}$ .

### F. RGB to HSV

1)  $RGB$  coordinates are normalized to  $[0, 1]$

$$M = \max(R, G, B)$$

$$m = \min(R, G, B)$$

$$r = (M - R)/(M - m)$$

$$g = (M - G)/(M - m)$$

$$b = (M - B)/(M - m).$$

2) Value  $V \in [0, 1]$

$$V = \max(R, G, B).$$

3) Saturation  $S \in [0, 180]$

if  $(M = 0)$  then  $S = 0$  and  $H = 180$

if  $(M \neq 0)$  then  $S = (M - m)/M$ .

4) Hue  $H \in [0, 360]$

if  $(R = M)$  then  $H = 60 * (b - g)$

if  $(G = M)$  then  $H = 60 * (2 + r - b)$

if  $(B = M)$  then  $H = 60 * (4 + g - r)$

if  $(H >= 360)$  then  $H = H - 360$

if  $(H < 0)$  then  $H = H + 360$ .

### G. RGB to HLS

The normalized  $RGB$  values (Step 1) and hue (Step 4) are calculated in the same way as in the  $RGB$  to  $HSV$  conversion algorithm. Lightness and saturation are calculated as follows.

1) Lightness  $L \in [0, 1]$

$$L = (M + m)/2.$$

4) Saturation  $S \in [0, 1]$

if  $(M = m)$  then  $S = 0$  and  $H = 180$

if  $(I \leq 0.5)$  then  $S = (M - m)/(M + m)$

if  $(I > 0.5)$  then  $S = (M - m)/(2 - M - m)$ .

### ACKNOWLEDGMENT

The authors would like to thank Dr. M. O. Schurr, MD, Section for Minimal Invasive Surgery, University of Tübingen, Germany, for the provision of a part of the endoscopic videos used in our study and his contribution to the evaluation of the results.

### REFERENCES

- [1] *Cancer Facts and Figures*, 2000.
- [2] S. Parker, T. Tong, S. Bolden, and P. Wingo, "Cancer statistics 1997," *CA Cancer J. Clinicians*, vol. 47, pp. 5-27, 1997.
- [3] D. Rex, R. Weddle, D. Pound, K. O'Connor, R. Hawes, R. Dittus, J. Lappas, and L. Lumeng, "Flexible sigmoidoscopy plus air contrast barium enema versus colonoscopy for suspected lower gastrointestinal bleeding," *Gastroenterology*, vol. 98, pp. 855-861, 1990.
- [4] S. Nagata, S. Tanaka, K. Haruma, M. Yoshihara, K. Sumii, G. Kajiyama, and F. Shimamoto, "Pit pattern diagnosis of early colorectal carcinoma by magnifying colonoscopy: Clinical and histological implications," *Int. J. Oncol.*, vol. 16, pp. 927-934, 2000.
- [5] S. Kudo, H. Kashida, T. Tamura, E. Kogure, Y. Imai, H. Yamano, and A. R. Hart, "Colonoscopic diagnosis and management of nonpolypoid early colorectal cancer," *World J. Surgery*, vol. 24, pp. 1081-1090, 2000.
- [6] S. Kudo, S. Tamura, T. Nakajima, H. Yamano, H. Kusaka, and H. Watanabe, "Diagnosis of colorectal tumorous lesions by magnifying endoscopy," *Gastrointestinal Endoscopy*, vol. 44, pp. 8-14, Jan. 1996.
- [7] H. Sujana, S. Swarnamani, and S. Suresh, "Artificial neural networks for the classification of liver lesions by image texture parameters," *Ultrasound Med. Biol.*, vol. 22, pp. 1177-1181, Sept. 1996.
- [8] A. G. Houston and S. B. Premkumar, "Statistical interpretation of texture for medical applications," presented at the Biomedical Image Processing and Three Dimensional Microscopy, San Jose, CA, 1991.
- [9] C. Enderwick and E. Micheli-Tzanakou, "Classification of mammographic tissue using shape and texture features," *Proc. 19th Annu. Int. Conf. IEEE Engineering Medicine Biology Society*, pp. 810-813, 1997.
- [10] F. Lachmann and C. Barillot, "Brain tissue classification from MRI data by means of texture analysis," in *Proc. Medical Imaging VI: Image Processing*, vol. 1652, Newport Beach, CA, 1992, pp. 72-83.
- [11] Q. Ji, J. Engel, and E. Craine, "Texture analysis for classification of cervix lesions," *IEEE Trans. Med. Imag.*, vol. 19, pp. 1144-1149, Nov. 2000.
- [12] C. Fortin and W. Ohley, "Automatic segmentation of cardiac images: Texture mapping," *Proc. IEEE 17th Annu. Northeast Bioeng. Conf.*, 1991.
- [13] A. N. Esgiar, R. N. G. Naguib, B. S. Sharif, M. K. Bennett, and A. Murray, "Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa," *IEEE Trans. Inform. Technol. Biomed.*, vol. 2, pp. 197-203, Mar. 1998.
- [14] —, "Automated feature extraction and identification of colon carcinoma," *Anal. Quant. Cytology Histology*, vol. 20, pp. 297-301, 1998.
- [15] K. W. Abyoto, S. J. Wirdjosoedirdjo, and T. Watanabe, "Unsupervised texture segmentation using multiresolution analysis for feature extraction," *J. Tokyo Univ. Inform. Sci.*, vol. 2, pp. 49-61, Jan. 1998.
- [16] G. Van de Wouwer, P. Scheunders, and D. Van Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Trans. Image Processing*, vol. 8, pp. 592-598, Apr. 1999.
- [17] S. H. Itzkowitz and Y. S. Kim, *Sleisinger & Fordtran's Gastrointestinal and Liver Disease*, 6th ed. Philadelphia, PA: Saunders, 1998, vol. 2.

- [18] E. R. Fearon and B. Vogelstein, "A genetic model for colorectal tumorigenesis," *Cell*, vol. 61, pp. 759–767, 1990.
- [19] B. Vogelstein, E. R. Fearon, and S. Hamilton, "Genetic alterations during colorectal-tumor development," *N. Eng. J. Med.*, vol. 319, pp. 525–532, 1998.
- [20] A. R. Williams, B. A. W. Balasooriya, and D. W. Day, "Polyps and cancer of the large bowel: A necroscopy study in Liverpool," *GUT*, vol. 23, pp. 835–842, 1982.
- [21] C. D. Johnson and A. H. Dachman, "CT colonography: The next colon screening examination," *Radiology*, vol. 216, pp. 331–341, 2000.
- [22] B. Saar, J. T. Heverhagen, T. Obst, L. D. Berthold, I. Kopp, K. J. Klose, and H. Wagner, "Magnetic resonance colonography and virtual magnetic resonance colonoscopy with the I.0-T system: A feasibility study," *Invest. Radiol.*, vol. 35, pp. 521–526, Sept. 2000.
- [23] Virtual Colonoscopy, M. Macari. (2002, June). [Online]. Available: [www.nysge.org/Postgrad\\_1999/Macari.htm](http://www.nysge.org/Postgrad_1999/Macari.htm)
- [24] S. B. Gokturk, C. Tomasi, R. B. Acar, C. F. Beaulieu, D. S. Paik, R. B. Jeffrey Jr., J. Yee, and S. Napel, "A statistical 3-D pattern processing method for computer-aided detection of polyps in CT colonography," *IEEE Trans. Med. Imag.*, vol. 20, pp. 1251–1260, Dec. 2001.
- [25] D. S. Paik, C. F. Beaulieu, R. B. Jeffrey, C. A. Karadi Jr., and S. Napel, "Detection of polyps in CT colonography: A comparison of a computer-aided detection algorithm to 3D visualization methods," presented at the Radiological Society of North America 85th Scientific Sessions, Chicago, IL, Nov. 1999.
- [26] H. Yoshida, Y. Masutani, P. M. MacEaney, K. Doi, Y. Kim, and A. H. Dachman, "Detection of colonic polyps in CT colonography based on geometric features," *Radiology*, vol. 217, pp. 582–582, Nov. 2000.
- [27] J. Nappi and H. Yoshida, "Automated detection of polyps with CT colonography: Evaluation of volumetric features for reduction of false positive findings," *Acad. Radiol.*, vol. 9, no. 4, pp. 386–397, 2002.
- [28] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, pp. 786–804, 1979.
- [29] *The Handbook of Pattern Recognition and Computer Vision*, 2nd ed., C. H. Chen, L. F. Pau, and P. S. P. Wang, Eds., World Scientific, Singapore, 1998, pp. 207–248.
- [30] T. R. Reed and J. M. H. Du Buf, "A review of recent texture segmentation and feature extraction techniques," *Computer Vision Graphics Image Process. Image Understanding*, vol. 57, pp. 359–372, Mar. 1993.
- [31] T. Caelli and D. Reye, "On the classification of image regions by color texture and shape," *Pattern Recog.*, vol. 26, pp. 461–470, Apr. 1993.
- [32] J. Scharcanski, J. H. Hovis, and H. C. Shen, "Representing the color aspects of texture images," *Pattern Recog. Lett.*, vol. 15, pp. 191–197, 1994.
- [33] R. Kondepudy and G. Healey, "Modeling and identifying 3-D color textures," in *Proc. Int. Conf. Comput. Vision Pattern Recognition*, 1993, pp. 577–582.
- [34] G. Paschos, "Fast color texture recognition using chromaticity moments," *Pattern Recog. Lett.*, vol. 21, pp. 847–841, 2000.
- [35] M. Mirmehdi and M. Petrou, "Segmentation of color textures," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 142–159, Feb. 2000.
- [36] C. Palm, D. Keysers, and K. Spitzer, "Gabor filtering of complex hue/saturation images for color texture classification," in *Proc. Joint Conf. Information Sciences Int. Conf. Comput. Vision, Pattern Recognition, and Image Processing*, vol. 2, 2000, pp. 45–49.
- [37] K. B. Eom, "Segmentation of monochrome and color textures using moving average modeling approach," *Image Vision Comput.*, vol. 17, pp. 233–244, 1999.
- [38] M. P. Dubuisson-Jolly and A. Gupta, "Color and texture fusion: Application to aerial image segmentation and GIS updating," *Image Vision Comput.*, vol. 18, pp. 823–832, 2000.
- [39] A. Drimbarean and P. F. Whelan, "Experiments in color texture analysis," *Pattern Recognition Lett.*, vol. 22, pp. 1161–1167, 2001.
- [40] G. Paschos, "Chromatic correlation features for texture recognition," *Pattern Recognition Lett.*, vol. 19, pp. 643–650, 1998.
- [41] G. Van de Wouwer, P. Scheunders, S. Livens, and D. Van Dyck, "Wavelet correlation signatures for color texture characterization," *Pattern Recognition*, vol. 32, pp. 443–451, 1999.
- [42] N. Vandenbroucke, L. Macaire, and J.-G. Postaire, "Unsupervised color texture feature extraction and selection for soccer image segmentation," *Proc. IEEE Int. Conf. Image Processing*, vol. 2, pp. 800–803, 2000.
- [43] E. Thompson, A. Palacios, and F. J. Varela, "Ways of coloring: Comparative color vision as a case study for cognitive science," *Behavioral Brain Sci.*, vol. 15, pp. 1–74, Jan. 1992.
- [44] S. Zeki, "Color coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colors," *Neuroscience*, vol. 9, pp. 741–765, 1983.
- [45] C. H. Li, "Regularized color clustering in medical image database," *IEEE Trans. Med. Imag.*, vol. 19, pp. 1150–1155, Nov. 2000.
- [46] D. G. Chamberlin, *Color: Its Measurement, Computation and Application*: Hayden, 1980, ch. New York.
- [47] G. Wyszecki and W. S. Styles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed. New York: Wiley, 1982.
- [48] T. Gevers and W. M. Smeulders, "PicToSeek: Combining color and shape invariant features for image retrieval," *IEEE Trans. Image Processing*, vol. 9, pp. 102–119, Jan. 2000.
- [49] P. Duchnowski, M. Hunke, D. Busching, M. Meier, and A. Waibel, "Toward moment-invariant automatic lip-reading and speech recognition," in *Proc. ICASSP*, vol. 1, 1995, pp. 109–112.
- [50] S. H. Kim, N. K. Kim, S. C. Ahn, and H. G. Kim, "Object oriented face detection using range and color information," in *Proc. 3rd Int. Conf. Automatic Face Gesture Recognition*, 1998, pp. 76–81.
- [51] Q. B. Sun, W. M. Huang, and J. K. Wu, "Face detection based on color and local symmetry information," in *Proc. 3rd Int. Conf. Automatic Face Gesture Recognition*, 1998, pp. 130–135.
- [52] J. Yang and A. Waibel, "Tracking Human Faces in Real Time," Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-CS-95-210, 1995.
- [53] G. Paschos, "Perceptually uniform color spaces for color texture analysis: An empirical evaluation," *IEEE Trans. Image Processing*, vol. 10, pp. 932–937, June 2001.
- [54] S. Fischer, P. Schmid, and J. Guilloid, "Analysis of skin lesions with pigmented networks," in *Proc. Int. Conf. Image Processing*, vol. 1, 1996, pp. 323–326.
- [55] Y. Ohta, T. Kanade, and T. Sakai, "Color information for region segmentation," in *Proc. Comput. Graphics Image Processing*, vol. 13, 1980, pp. 222–241.
- [56] A. Mojsilovic, J. Kovacevic, J. Hu, R. J. Safranek, and K. Ganapathy, "Retrieval of color patterns based on perceptual dimensions of texture and human similarity rules," in *Proc. SPIE Human Vision and Electronic Imaging*, vol. 3644, 1999, pp. 441–452.
- [57] R. P. Schumeyer and K. E. Barner, "Color-based classifier for region identification in video," in *Proc. SPIE Visual Communications Image Processing*, vol. 3309, pp. 189–200.
- [58] A. R. Smith, "Integrated Spatial and Feature Image Systems: Retrieval, Compression and Analysis," Ph.D. dissertation, Columbia Univ., New York, 1997.
- [59] J. Keith, *Video Demystified*, 2nd ed., 1996. Hightext Interactive.
- [60] C. Chiu, "A novel approach based on computerized image analysis for traditional Chinese medical diagnosis of the tongue," *Computer Methods and Programs in Biomedicine*, vol. 61, pp. 77–89, 2000.
- [61] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Processing*, vol. 4, pp. 1549–1560, 1995.
- [62] A. Archimandritis, M. Tjivras, P. Davaris, D. Kalogeras, M. Chronaki, S. Bougas, and A. Fertakis, "Colonic polyps found on flexible sigmoidoscopy: A retrospective study," *J. Clin. Gastroenterol.*, vol. 17, pp. 87–89, Jan. 1993.
- [63] S. Liapis, N. Alvertos, and G. Tziritas, "Unsupervised texture segmentation using discrete wavelet frames," in *Proc. European Signal Processing Conf.*, 1998, pp. 1341–1344.
- [64] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, 1989.
- [65] B. Julesz, "Texton gradients: The texton theory revisited," *Biol. Cybern.*, vol. 54, pp. 245–251, 1986.
- [66] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 610–621, 1973.
- [67] S. Karkanis, G. D. Magoulas, and N. Theofanous, "Image recognition and neuronal networks: Intelligent systems for the improvement of imaging information," *Minimal Invasive Therapy and Allied Technologies*, vol. 9, pp. 225–230, 2000.
- [68] S. A. Karkanis, G. D. Magoulas, D. K. Iakovidis, D. A. Karras, and D. E. Maroulis, "Evaluation of textural feature extraction schemes for neural network-based interpretation of regions in medical images," *Proc. IEEE ICIP*, pp. 281–284, 2001.
- [69] S. A. Karkanis, D. K. Iakovidis, D. A. Karras, and D. E. Maroulis, "Detection of lesions in endoscopic video using textural descriptors on wavelet domain supported by artificial neural network architectures," *Proc. IEEE ICIP*, pp. 833–836, 2001.
- [70] 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*, vol. 70, pp. 151–186, Feb. 2003.

- [71] P. Armitage and G. Berry, *Statistical Methods in Medical Research*, 2nd ed. Oxford, U.K.: Blackwell, 1987.
- [72] M. G. Kendall, *Rank Correlation Methods*, 4th ed, London, U.K.: Griffin, 1970.
- [73] G. Semler *et al.*, "Test-retest reliability of a standardised psychiatric interview (DIS/CIDI)," *European Archives of Psychiatry Neurology Science*, vol. 236, pp. 214–222, 1987.
- [74] Performance Libraries (2002, Jan.). [Online]. Available: <http://developer.intel.com/software/products/>
- [75] D. West and V. West, "Model selection for a medical diagnostic decision support system: A breast cancer detection case," *Artificial Intelligence in Medicine*, pp. 183–204, 2000.
- [76] M. Kubat and S. Matwin, "Addressing the curse of imbalanced training sets: One-sided selection," in *Proc. 14th Int. Conf. Machine Learning*, 1997, pp. 179–186.
- [77] J. A. Swets, R. M. Dawes, and J. Monahan, "Psychological science can improve diagnostic decisions," *Psych. Sci. Public Interest*, vol. 1, pp. 1–26, 2000.
- [78] B. Mac Namee, P. Cunningham, S. Byrne, and O. I. Corrigan, "The problem of bias in training data in regression problems in medical decision support," *AI Med.*, vol. 24, Jan. 2002.
- [79] J. F. Kenney and E. S. Keeping, "Linear regression and correlation," in *Mathematics of Statistics*, 3rd ed. Princeton, NJ: Van Nostrand, 1962, pp. 252–285.
- [80] G. C. Looney, *Pattern Recognition Using Neural Networks: Theory and Algorithms for Engineers and Scientists*. New York: Oxford Univ. Press, 1997, pp. 316–319.

**Stavros A. Karkanis** (M'89) received the B.Sc. degree in mathematics in April 1986 and the Ph.D. degree in December 1995, both from the Department of Informatics and Telecommunications, University of Athens, Athens, Greece.

He has worked in the field of image processing and especially in texture recognition from various academic and industry positions. Until 2003, he was Research Assistant with the Department of Informatics and Telecommunications, University of Athens, working with the Real-time Systems and Image Analysis Group. Currently, he is Associate Professor in the Department of Informatics at the Technological Educational Institute of Lamia. His main interests include texture recognition, wavelet transform for texture, pattern recognition for image processing applications, and statistical learning methodologies for classification.

**Dimitris K. Iakovidis** received the B.Sc. degree in physics from the University of Athens, Athens, Greece, in November 1997 and the M.Sc. degree in electronic automation in April 2001. Currently, he is pursuing the Ph.D. degree in data acquisition systems and image analysis.

**Dimitris E. Maroulis** (M'02) received the B.Sc. degree in physics, the M.Sc. degree in radioelectricity, the M.Sc. in electronic automation, and the Ph.D. degree in informatics, all from University of Athens, Athens, Greece, in 1973, 1977, 1980, and 1990, respectively.

In 1979, he was appointed Assistant in the Department of Physics, in 1991, he was elected Lecturer and in 1994, he was elected Assistant Professor, in the Department of Informatics of the University of Athens. He is currently teaching and conducting research activities, including projects with European Community. His main areas of activity include data acquisition systems, real-time systems, signal processing, and digital communications.

**Dimitris A. Karras** (A'86) received the Diploma and M.Sc. degrees in electrical and electronic engineering from the National Technical University of Athens, Athens, Greece in 1985 and the Ph.D. degree in electrical engineering, from the National Technical University of Athens, in 1995, with honors. He received the Diploma degree in mathematics from the University of Athens in 1999.

From 1989 to 1995, he was with the National Centre of Scientific Research as a collaborating researcher. From 1995 to 2000, he served as a Visiting Professor in the Department Informatics, University of Ioannina, Ioannina, Greece, and in the Department Business Administration, University of Piraeus, Piraeus, Greece. From 2000 to 2003, he was with the Hellenic Aerospace Industry as Technical Manager in Telecommunications projects as well as with the Department of Informatics, Hellenic Open University, as a Visiting Professor. In addition, he is collaborating with the University of Hertfordshire, U.K., as an External Professor. He has published more than 25 research Journal papers in various areas of pattern recognition, image/signal processing and neural networks and more than 60 research papers in international scientific conferences. His research interests span the fields of pattern recognition and neural networks, multidimensional digital signal processing, image processing and analysis, communications and security, as well as parallel algorithms and fast processing.

Dr. Karras received a research grant from DKFZ, Germany, in 1995. He is listed in International Who's Who and has served as session chair/international program committee member in various International Scientific Conferences and as a referee in various research journals/conferences. He is a Member of ACM and INNS as well as a member of the Technical Chamber of Greece.

**M. Tzivras** received the B.Sc. degree in medicine in 1968 and he was named a Certified Specialist in internal medicine in 1977. He received the Ph.D. degree in 1979 from the University of Athens and he was named a Certified Specialist in gastroenterology in 1980.

From 1981 to 1982, he was Senior Registrar and in 1982, he was appointed Senior Lecturer, Department of Pathophysiology, Medical School, University of Athens, Athens, Greece. Since 1983, he has been in charge of the Gastroenterology Clinic of this department. In 1988, he was appointed Assistant and in 1999, he was appointed Associate Professor of Medicine, Medical School, University of Athens. He is currently in charge of the Gastroenterology and Endoscopy Section of the Department of Pathophysiology.