Available online at www.sciencedirect.com
SCIENCE DIRECT®
Measurement 36 (2004) 297–313
Measurement
www.elsevier.com/locate/measurement

Texture multichannel measurements for cancer precursors' identification using support vector machines

D.K. Iakovidis [a], D.E. Maroulis [a,*], S.A. Karkanis [b], P. Papageorgas [a], M. Tzivras [c]

[a] *Real Time Systems and Image Analysis Group, Department of Informatics and Telecommunications, University of Athens, Panepistimiopolis, Illisia, 15784 Athens, Greece*
[b] *Department of Informatics and Computer Technology, Technological Educational Institute of Lamia, Lamia 35100, Greece*
[c] *Gastroenterology Section, Department of Pathophysiology, Medical School, University of Athens, Laiko General Hospital, 17 St. Thomas Str., 115 27, Athens, Greece*

Available online 18 October 2004

Abstract

Colorectal cancer is one of the leading types of cancer in the developed countries. Epidemiological studies have shown that the risk of developing colorectal cancer can be significantly reduced through early detection and removal of cancer precursor lesions. We propose a novel framework for the automated identification of colon cancer precursors based on the processing of color video frames acquired during endoscopy. The spectral information of the three color channels forming the endoscopic frames is used for the description of the colonic mucosa. The suitability of different color models for this application is investigated. The textural properties of the colonic mucosal surface are measured using second order statistical descriptors on the wavelet transform of the multichannel video signals. A new reduced set of measures based on the inter-channel covariance of the features has proven to provide high discrimination of image regions corresponding to normal and abnormal tissue. The proposed framework was tested using a Support Vector Machine classifier on different video frame sets presenting adenomatous polyps and the average sensitivity and specificity was estimated to reach 94% and 95.7%, respectively.

Keywords: Wavelets; Support vector machines; Color texture analysis; Endoscopy; Cancer; Medical application

1. Introduction

Cancer of the colon and rectum is one of the commonest forms of malignancy in developed countries, and the incidence appears to be rising

* Corresponding author. Tel.: +30 210 7275307; fax: +30 210 7275333.
E-mail address: rtsimage@di.uoa.gr (D.E. Maroulis).

0263-2241/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.measurement.2004.09.010

[1]. The National Cooperative Colon Polyp study clearly showed that the identification and removal of adenomatous colon polyps significantly reduce the risk of developing colorectal cancer [2]. Polyps are visible tissue masses protruding from the mucosal surface. They are characterized by their color, the appearance of their mucosal surface, the presence of ulcers, their bleeding tendency, and above all the presence of pedunculus (pedunculated or non-pedunculated). Their size varies from barely visible transparent protrusions to penduculated lesions with a diameter of 3–5 cm or non-pedunculated lesions with a diameter of 10–20 cm. Although there are many histopathologic types of colonic polyps, approximately 75% of them are adenomatous [3]. There are several screening procedures that are used for the detection of colonic polyps, including barium enema, computed tomography or magnetic resonance imaging of the colon, virtual endoscopy, rectosigmoidoscopy and finally endoscopy, which is the most accurate screening procedure for patients with clinical or laboratory suspicion of colonic polyposis [4].

During the endoscopic examination of the colon it is possible for some polyps to be discounted by the physicians and become malignant in the following years. This fact imposes the need of a reliable system that would be capable of supporting the detection of polyps. Such a system could increase the physician's ability to accurately locate early stage polyps, and could contribute to a reduction in the duration of the endoscopic procedure, which is in most cases painful for the patients. Moreover, a consequent cost reduction of the operation would be feasible, since more patients would be able to be examined in less time.

Computer-aided diagnosis is a useful tool for improving accuracy for several diagnostic tasks [5]. In many of these tasks, the measurement of color and textural characteristics of medical images have shown to provide important information for the automatic detection of lesions, such as lesions of the liver [6,7], prostate [8], brain [9], breast [10,11], heart [12], skin [14–16], cervix [13] and colon [17–22]. As it has been proposed by Kudo [23,24] changes in the cellular pattern (pit pattern) of the colon lining can be used for qualitative and quantitative diagnosis of colonic lesions. This textural information can be exploited for the automated identification of polyps.

In this paper, we propose a novel framework for the identification of colorectal polyps using wavelet domain measurements of the textural content of color video endoscopy frames. This framework includes: (a) acquisition and preprocessing of the endoscopic color video frames, (b) calculation of textural image measures and, (c) classification of these measures into classes that correspond to image regions depicting healthy or suspicious for abnormality tissue.

The textural measures are calculated on the 2-dimensional Discrete Wavelet Transform (2D-DWT) of each color channel of the video frames. 2D-DWT is a multiresolution technique, which intends to transform images to a representation in which both spatial and frequency information is present. It conforms to the way the human visual system processes images [25], and its application in texture classification problems has resulted in higher success rates [20–22,26]. The utilization of the textural information existing in or between the different color channels of an image have been a point of interest only in the recent years of research, where the widely available computational power has increased. Significant improvement of texture classification results has been reported using color instead of grayscale images [27–29].

The classification task has been assigned to a Support Vector Machine (SVM). SVMs are supervised machine learning algorithms that are based on statistical learning theory [30]. Their remarkably robust performance, even with sparse and noisy data and their ability to resist overfitting and the "curse of dimensionality", makes them attractive for a number of real world problems [31]. They have been successfully applied to many classification tasks such as face recognition [32], handwritten digit recognition [33], defect detection [34], detection of bacilli in medical images [35], and classification and validation of cancer tissue samples using microarray expression data [36].

The rest of this paper is organized as follows. In Section 2, we provide information on the acquisition setup and the spectral content of the endoscopic video frames, in Section 3 we review the

color models that have been adopted for texture and medical image analysis, and in Section 4 we present a novel scheme for the analysis of the colonoscopic images as applied on the wavelet domain of our multichannel measurements. Section 5, outlines the principles of the SVM models, and the experimental results of our study are presented in Section 6. Finally, the conclusions of our study are summarized in Section 7.

2. Endoscopic imaging

Endoscopes usually consist of a fiber-optic tube attached to a viewing device, and are used for the exploration and biopsy in such areas as the gastrointestinal system and the bronchi of the lungs. Endoscopes employ miniature cameras and tiny surgical implements that allow exploration and minimally invasive endoscopic surgery through small incisions. Such surgery is much less traumatic to the patients than traditional open surgery. The endoscopy oriented to the visual examination of the colonic mucosa is also known as colonoscopy. Standard colonoscopes are quite small having an average length of around 185 cm and are only 12–13 mm in diameter. The instrument head is connected to a variety of auxiliary devices via a separate cable, such as a suction box, an external cold light source and a water feed tank which is used for intestine cleaning. The recording of the colonoscopy procedure is commonly performed by a standard video recording system.

The color images of the colon or the rectum that are acquired by an endoscope provide important information for diagnosing various kinds of rectal and colon diseases. An endoscopic camera usually consists of three light sensors, symbolized as $i = R, G, B$. Each of these sensors is characterized by its spectral response function $S_i(\lambda)$, which indicates the sensitivity of the sensor to different wavelengths λ. The spectral response functions of these sensors correspond to overlapping intervals of the visible spectrum that are centered on red (R), green (G) and blue (B) wavelengths. This overlap introduces a correlation between the RGB components. The color reproduction characteristics of the endoscope depend on several factors, including (a) the spectral radiant distribution of the illuminant $E(\lambda)$, which specifies the energy emitted by the light source of the endoscope at each wavelength λ, (b) the spectral sensitivity $S_i(\lambda)$ of the sensors $i = R$, G, B, which specifies its sensitivity to the light energy at each wavelength λ, (c) the spectral transmittance of the imaging lenses $L(\lambda)$ and (d) the spectral reflectance of the mucosal surface $O(\lambda)$. When the light travels from the source, is reflected by the mucosal surface and finally arrives at the sensor through the lens, its spectrum is modified and the spectral characteristics of these stages are multiplied. The spectrum of a light beam, arriving at each sensor, is weighted by the response function $S_i(\lambda)$ of the sensor, since each sensor captures a particular interval of the signal's wavelength. The intensity value $V_i(x, y)$ of the sensor response to this beam, at a specific point with coordinates (x, y), is calculated by convolving the incident spectrum with the bandpass filter $S_i(\lambda)$:

$$V_i(x, y) = \int_w E(\lambda) S_i(\lambda) L(\lambda) O(\lambda, x, y) \mathrm{d}\lambda$$

w is the range of wavelengths for which the sensors have non-zero sensitivity (visible spectrum) [37].

Multispectral imaging of different objects has shown that more than three components could be necessary for an accurate reproduction of their spectra. For example, three components are needed for the reproduction of oil paintings' spectra [37]. Under this framework, Principal Component Analysis (PCA) on a number of normal rectal membrane reflectance spectra from different patients at Kyoto National Hospital [37], showed that 99.7% of the spectra could be expressed by only three principal components, while the introduction of more principal components could only contribute to an increase of 0.3%. Thus, the reflectance spectra of the rectal membrane can be adequately estimated from the RGB output channels of the electronic endoscope without significant loss of valuable information needed for medical diagnosis. The analog output signals of the endoscope are then led to the composite input of a standard video recording system and the video frames are digitized so that they can be processed by a personal computer system (Fig. 1).

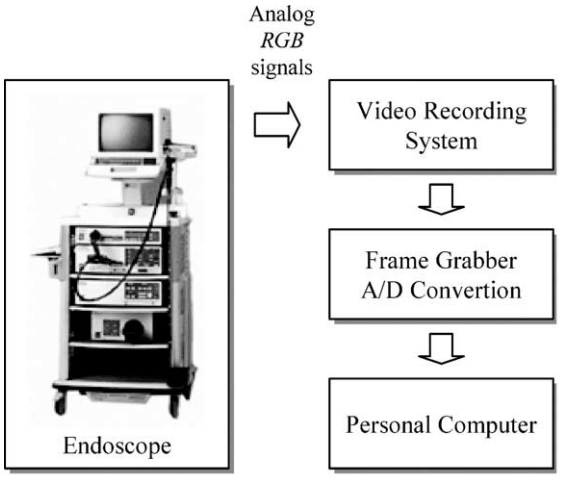

Fig. 1. Acquisition of colonoscopic video frames.

3. Color models for medical diagnosis

The *RGB* color model is a common representation of color used in digital imaging. The three primary components that it uses to represent color (Red, Green and Blue) correspond to different regions of the visible spectrum. In many medical decision support applications, the direct use of the *RGB* model has been considered inadequate to represent accurately the clinical and pathological characteristics of the examined tissue [14,38,39]. Moreover, it has been reported that *RGB* endoscopic images are not good enough for the diagnosis of a disease in early stages [37,40]. Major disadvantages of the *RGB* representation advocating to its low performance are (a) the high correlation among the *RGB* channels and (b) the representation of colors does not conform to the perceptual mechanisms of the human brain [41,42].

Different color models have been proposed to overcome these drawbacks. The acquired *RGB* signals can undergo a pre-processing stage in which they will be transformed linearly or non-linearly into a color model that could enhance the significant image characteristics for the identification of the cancer pre-cursors. The normalized *RGB*, known as *rgb* (see Appendix A), has proven to be independent of viewpoint, surface orientation, illumination direction and illumination intensity, assuming dichromatic reflection and white illumination [43]. It has been employed for automatic lip reading [44] and face detection applications [45–49]. A major drawback of *rgb* is that it provides poor segmentation results when intensity is low [50].

Best segmentation results for the analysis of skin lesions [14] and for the characterization of color textures [29] have been achieved with the use of the Karhunen–Loeve (K–L) color model [51] which is derived from the application of PCA on digital images. The K–L color model is formed by the eigenvector of the correlation matrix of an *RGB* image, which remains approximately the same for a large set of natural color images [29,51]. The images are transformed into an orthogonal basis in which the axes are statistically uncorrelated. In that sense, the information present in *RGB* space is decorrelated. Practically, it can be derived as a linear transformation of the *RGB* coordinates (see Appendix A).

Phenomenal color models attempt to classify colors in relation to how they are perceived by the human brain. In general, these color models mainly incorporate Hue, Saturation and Brightness as classifying descriptors, and they are more intuitive in manipulating color [52]. A representative example is *HSV* (Hue, Saturation and brightness Value). *HSV* have been successfully employed in different applications of skin segmentation [53,54] and the measurement of plaque from intra-oral video frames [55]. It has led to higher classification performance than *RGB* in both noisy and noise-free conditions for color texture analysis [27], but on the other hand Palm [56] showed that *HSV* performs equivalently to *RGB* for color texture classification using different measures.

Perceptually uniform color models, as $L^*a^*b^*$, $L^*u^*v^*$ and their derivatives [28,42,57], have been proposed to describe color closer to the way humans perceive color, in the sense that the perceived color differences are measured by Euclidean distances. They have been applied successfully in many cases, such as general color image segmentation [58], the retrieval of color patterns using textural features [59], the analysis of skin lesions [14], the segmentation of human flesh [60] and color clustering in medical image database [61]. A reason for inhibiting these spaces from being

widely used in image processing tasks is their noise-sensitivity due to the nonlinear transformations involved [62]. In addition, their non-linearity poses a computational complexity problem [52]. $L^*a^*b^*$, $L^*u^*v^*$ and HSV are calculated as a non-linear transformations of the RGB model (see Appendix A).

4. Color texture analysis of the colonoscopic video frames

Texture is a fundamental characteristic of the digital images and it usually reflects the composition and structure of the pictured objects. Measurement of these textural characteristics can provide significant information for the discrimination of these objects. For a number of years, texture analysis mainly focused on the use of gray-level image information [63–65] and only a few efforts have been presented exploiting color image information, for instance in [27–29,66]. Color and texture is the discriminating information that is usually used by gastroenterologists to differentiate normal from abnormal tissue of the colonic mucosa [23,24]. On this basis, a system that would be capable of identifying colorectal polyps should automatically incorporate processing algorithms for the analysis of texture in color images.

The polyps that the system should be capable of identifying can have different sizes. The image resolution cannot be defined so that to cover the whole range of their sizes. This leads to deal with the classification of normal/abnormal regions in different resolutions in order to exploit the information from the intermediate scales for the final decision. Multiresolution analysis is performed by applying the 2-dimensional Discrete Wavelet Transform (2D-DWT) on the colonoscopic video frames.

4.1. Discrete wavelet transform

In 1-dimensional DWT (1D-DWT) of a signal, two functions mutually orthonormal are initially adopted: the scaling function φ and the mother wavelet ψ. Other wavelets are then produced by translations of the scaling function φ and dilations by the mother wavelet ψ, according to the equations [67]:

$$\begin{aligned}&\varphi_{j_0,k}(t) = 2^{j_0/2}\varphi(2^{j_0}t-k),\\ &\psi_{j,k}(t) = 2^{j/2}\psi(2^{j}t-k), \quad j = j_0, j_0+1,\ldots,\ k\in \boldsymbol{Z}\end{aligned} \tag{1}$$

for some $j_0 \in \boldsymbol{Z}$, where $\boldsymbol{Z}$ is the set of integers. The scaling function φ defines a kernel function and the mother wavelet ψ results in an oscillation of the input signal. Families of scaling functions can act as suitable bases for $\boldsymbol{L}^2(\boldsymbol{R})$ or for alternative spaces. The structure of a wavelet basis is deterministic in location and frequency due to translation and dilation respectively. A function $f \in \boldsymbol{L}^2(\boldsymbol{R})$ can be represented in a wavelet series, using a given basis, as [67]

$$f(t) = \sum_k c_{j_0,k}\varphi_{j_0,k}(t) + \sum_{j=1}^{j_0}\sum_{k\in Z} w_{j,k}\psi_{j,k}(t) \tag{2}$$

where j is the scale of the transform, j_0 is the "coarsest scale", $c_{j_0,k} = \langle f, \varphi_{j_0,k}\rangle$ and $w_{j,k} = \langle f, \psi_{j,k}\rangle$ are the wavelet coefficients and $\langle\cdot,\cdot\rangle$ is the standard $\boldsymbol{L}^2$ inner product of two functions

$$\langle f_1, f_2\rangle = \int_R f_1(t) f_2(t)\mathrm{d}t \tag{3}$$

The first term of Eq. (2) corresponds to a low resolution signal $L_{j_0} = \{c_{j_0}, k\}$, $k \in \boldsymbol{Z}$ that can be obtained by lowpass filtering. The coefficients $D_j = \{w_{j,k}\}$, $k \in \boldsymbol{Z}$, $1 \leqslant j \leqslant j_0$, constitute the detail signal at scale j, that can be obtained by highpass filtering. Together L_{j_0} and D_j are known as the *wavelet representation of depth* j_0 of the signal f. The expansion of DWT for 2D-signals, is straightforward if we consider that the wavelet transform is separable [67]. The 2D-DWT of a 2D-signal can be calculated by first applying 1D-DWT on its rows and then apply the 1D-DWT on the columns of the resulted 2D-signal. More precisely a separable filterbank is applied to the original 2D-signal L_0 according to the following recursive equations:

$$\left.\begin{aligned} L_{j_0} &= [H_x * (H_y * L_{j_0-1})_{\downarrow 2,1}]_{\downarrow 1,2}\\ D_{j1} &= [H_x * (G_y * L_{j_0-1})_{\downarrow 2,1}]_{\downarrow 1,2}\\ D_{j2} &= [G_x * (H_y * L_{j_0-1})_{\downarrow 2,1}]_{\downarrow 1,2}\\ D_{j3} &= [G_x * (G_y * stL_{j_0-1})_{\downarrow 2,1}]_{\downarrow 1,2}\end{aligned}\right\} \tag{4}$$

where $k \in \boldsymbol{Z}$, $1 \leqslant j \leqslant j_0$, $j_0 \in \boldsymbol{Z}$, $\downarrow 2,1$ and $\downarrow 1,2$ denote the sub-sampling along the rows and columns respectively, * is the convolution operator, H is the lowpass filter and G is the highpass filter. As in the 1D-DWT the coefficients $\{L_{j_0}, D_{j1}, D_{j2}, D_{j3}\}$, $1 \leqslant j \leqslant j_0$ are known as the *wavelet representation of depth* j_0 of a 2D-signal L_0.

4.2. Second-order statistical measures

An efficient method to measure the textural information of a single-channel video frame digitized in a number of intensity levels, is the gray-level cooccurrence method [68,69]. Cooccurrence matrices encode the gray-level spatial dependence based on the estimation of the second-order joint conditional probability density function $f(i,j,d,a)$, which is computed by counting all pairs of pixels at distance d having gray levels i and j at a given direction a. The angular displacement usually is included in the range of the values $\{0, \pi/4, \pi/2, 3\pi/4\}$. Among the 14 statistical measures, originally proposed by Haralick [63,70], that are derived from each cooccurrence matrix we have considered four:

$$\mathrm{ASM} = \sum_{i}^{N_g} \sum_{j}^{N_g} p(i,j)^2$$

Energy-Angular Second Moment (5)

$$\mathrm{COR} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i \cdot j) p(i,j) - \mu x \mu y}{\sigma_x \sigma_y} \quad \text{Correlation} \tag{6}$$

$$\mathrm{IDM} = \sum_{i}^{N_g} \sum_{j}^{N_g} \frac{1}{1 + (i-j)} p(i,j)$$

Inverse Difference Moment (7)

$$\mathrm{ENT} = -\sum_{i}^{N_g} \sum_{j}^{N_g} p(i,j) \log(p(i,j)) \quad \text{Entropy} \tag{8}$$

where $p(i,j)$ is the ijth entry of the normalized cooccurrence matrix, N_g is the number of gray levels of the image, μ_x, μ_y, σ_x, and σ_y are the means and standard deviations of the marginal probability $p_x(i)$ obtained by summing the rows of matrix $p(i,j)$. ASM is a measure of the homogeneity of an image. COR is a measure of linear intensity dependence between the pixels at the specified positions relative to each other. IDM is a measure of lack of variability of the intensity levels within the image and ENT measures the randomness of the intensity distribution within the image [70]. Studies on grayscale texture and colonoscopic images have shown that these four measures provide high discrimination accuracy, which can be only marginally increased by adding more measures in the feature vector [18–22,71].

4.3. Wavelet domain multichannel second-order statistical measures

The technique we propose for the analysis of the colonoscopic video frames, is based on the covariance estimation of the second-order statistical measures on the DWT of each channel of the video frame. Different covariance measures have been proposed in the literature for color texture analysis [27–29,72], but in most of them the 1st order statistical information of an image has been taken into account. It is described in the following four steps:

Step 1

Let I be the original multichannel signal composed by the separate channels C_i, $i = 1,2,\ldots,c$. The models used in this framework to describe image color, impose a maximum number of channels $c = 3$. Each of these channels is raster scanned with a fixed size sliding square window.

Step 2

On each window a K-level 2D-DWT ($j_0 = K$) is applied according to the equations of wavelet decomposition Eq. (4). This transform results in a new representation of the original window, which consists of

$$B = 3K + 1 \tag{9}$$

sub-windows corresponding to the different wavelet bands L_{j_0}, D_{j1}, D_{j2}, D_{j3}, $1 \leqslant j \leqslant j_0$ as illustrated in Fig. 2. In this figure each band is denoted as $B_b(k)$, where $b = 0,1,2,3$ for $k = K$ and $b = 1,2,3$

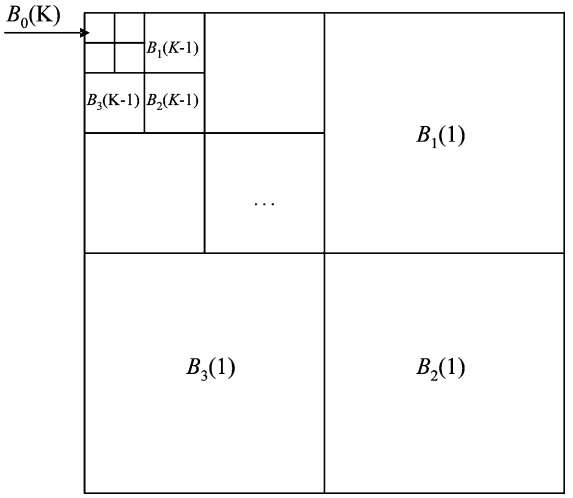

Fig. 2. Scheme of the sub-images resulted from K-level DWT of an image.

for $k < K$. These bands correspond to $B_0(k) = L_K, B_b(k) = D_{Kb}, b = 1,2,3$ for $k = K$, and $B_b\ (k) = D_{jb},\ \ 1 \leqslant j < j_0,\ \ b = 1,2,3,$ for $k < K$, respectively.

Step 3

The cooccurrence measures described in the previous paragraph are calculated over each subwindow $B_b(k)$, $b = 1,2,3$, $k = 1,2,\ldots,K$. The resulted set of measures corresponds to different channels and wavelet bands

$$F_{C_i}^{B_b(k)}(a) \quad i = 1,2,\ldots,c, \quad b = 1,2,3, \quad k = 1,2,\ldots K, \tag{10}$$

where $F \in \{\mathrm{ASM}, \mathrm{COR}, \mathrm{IDM}, \mathrm{ENT}\}$ and a corresponds to the angle considered in the calculation of the cooccurrence matrices, $a \in \{0, \pi/4, \pi/2, 3\pi/4\}$. A one-level wavelet decomposition of a color window leads to 144 measures (16 cooccurrence measures $\times$ 3 wavelet bands $\times$ 3 color channels), which comprise a 144-dimensional feature space.

Step 4

A significant reduction of the feature space dimension is achieved by considering the covariances of these features between the different channels C_i, $i = 1,2,\ldots,c$. For $c = 3$, we define *Color Wavelet Covariance* (CWC) of a measure $F \in \{\mathrm{ASM}, \mathrm{COR}, \mathrm{IDM}, \mathrm{ENT}\}$ Eqs. (5)–(8) at wavelet band $B_b(k)$, $b = 1,2,3$, $k = 1,2,\ldots,K$, between two color channels C_l and C_m as

$$\mathrm{CWC}_F^{B_b(k)}(l,m) = \mathrm{Cov}\left(F_{C_l}^{B_b(k)}, F_{C_m}^{B_b(k)}\right)$$

$$\mathrm{CWC}_F^{B_b(k)}(l,m) = \frac{\sum_a \left(\left(F_{C_l}^{B_b(k)}(a) - \overline{F}_{C_l}^{B_b(k)}\right) \cdot \left(F_{C_m}^{B_b(k)}(a) - \overline{F}_{C_m}^{B_b(k)}\right)\right)}{n_a} \tag{11}$$

where $l \leqslant m$, $\overline{F}$ is the mean value of F over the different angles a, and n_a is the number of angles considered for the calculation of the cooccurrence matrices. For $a \in \{0, \pi/4, \pi/2, 3\pi/4\}$, $n_a = 4$. Using these measures, the 144-dimensional feature space is reduced to 72 features, and consequently the complexity of the classification task that follows will be reduced.

Fig. 3, outlines the complete framework for the automated identification of colon cancer

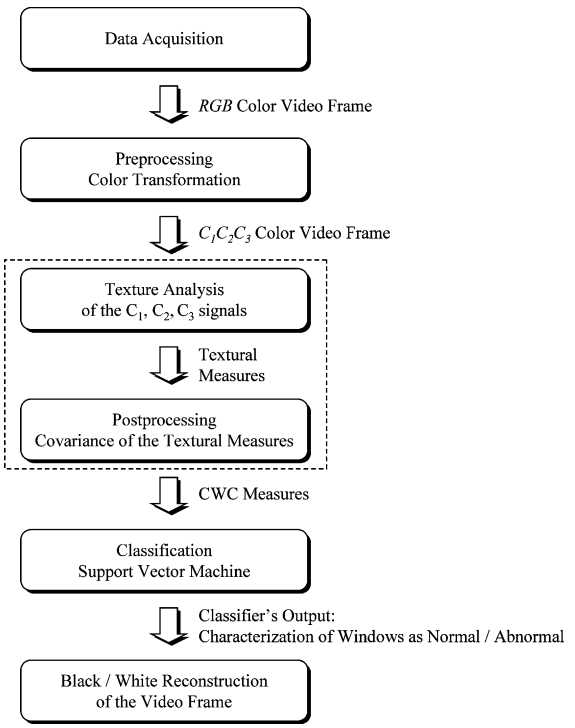

Fig. 3. Framework for the identification of normal/abnormal tissue in colonoscopic video frames.

precursors. First, a video frame is acquired as a three-channel signal with values corresponding to the *RGB* color model. These values are then preprocessed and transformed to a different color model $C_1C_2C_3$, which enhances the properties that characterize the normal/abnormal regions of the colonic mucosal surface. The textural information existent in each of the C_1, C_2, and C_3 channels of a video frame is measured in the wavelet domain and the CWC measures are calculated. The resulting measures form a set of feature vectors that feed an SVM classifier. The classification result is finally depicted as an artificially generated frame constructed by overlapping windows corresponding to the characterized regions of the original video frame. The windows that are classified as "normal" or "abnormal" are depicted in black and white color respectively. The same procedure is repeated for the rest of the colonoscopic video frames.

5. Support vector machines

Given a finite number of training data the learning machine should be able to learn any training set without error. Suppose we are given a number of observations (l) each one consisting of a vector x_i in $\boldsymbol{R}^n$, $i = 1,2,\ldots,l$ and the corresponding truth value y_i in $\{-1,1\}$. An unknown distribution $P(x,y)$ exists, according to which the data are produced. A learning machine should be able to learn the mapping $x_i \to y_i$. Different functions can be defined as $x_i \to f(x,\alpha)$ where α is an adjustable parameter. Choosing various values for α we can produce a different "trained machine". Any such machine is deterministic. The error that is produced by the machine during testing is called *expected risk* and can be described as

$$R(\alpha) = \int \frac{1}{2l} \mid y - f(x,\alpha) \mid \mathrm{d}P(x,y) \tag{12}$$

The mean error on the training set is called empirical risk R_{emp} and is defined as

$$R_{\mathrm{emp}}(\alpha) = \frac{1}{2l} \sum_{i=1}^{l} \mid y - f(x,\alpha) \mid \tag{13}$$

R_{emp} is a fixed number for each choice of α and training set $\{x_i, y_i\}$. Vapnik [30], proposed the following bound of risk:

$$R(\alpha) \leqslant R_{\mathrm{emp}}(\alpha) + \sqrt{\left(\frac{h(\log(2l/h)+1) - \log(\eta/4)}{l}\right)} \tag{14}$$

where h is the VC (Vapnik Chervonenkis) dimension representing the capability of the machine to learn. Three important properties of the *risk bound* can be mentioned

1. Risk bound is independent of $P(x,y)$.
2. $R(\alpha)$ cannot be computed.
3. If we know h we can compute the Risk bound.

So we are looking for a learning machine that produces the lowest upper bound on the actual risk.

VC dimension is a property of the functions $\{f(\alpha)\}$ and is defined as the maximum number of training points that can be shattered by $f(\alpha)$. In the case that the data are in $\boldsymbol{R}^2$, the VC dimension is three. In general when we are dealing with data from $\boldsymbol{R}^n$, the VC dimension is $n + 1$ [73].

The linear machines on separable data is the simplest case but as it comes from the study of non-linear machines on non-separable data, such general cases result to a similar quadratic programming problem. Supposing that we have a hyperplane

$$H : w \cdot x + b = 0$$

where w is normal to the hyperplane and $|b|/\|w\|$ is the distance from origin, which separates positive from negative samples. The support vector algorithm for the linearly separable case looks for *separating hyperplane* with the *largest margin*. According to this, we have the following constraints for the data:

$$(x_i \cdot w + b) \geqslant +1 \quad \text{for } y_i = +1, \tag{15}$$

$$(x_i \cdot w + b) \leqslant -1 \quad \text{for } y_i = -1 \tag{16}$$

and finally

$$y_i(x_i \cdot w + b) - 1 \geqslant 0 \quad \forall i \tag{17}$$

The two hyperplanes produced by Eqs. (15) and (16) are parallel and the maximum margin between them is $\|w\|^2$, subject to constraints Eq. (17). The *support vectors* are those training points that satisfy Eq. (17) and whose removal would change the solution found.

Support vectors are critical elements for the training set because

1. They lie close to the decision boundary.
2. If all the other training samples were removed and training was repeated, the same separating hyperplane would be found.

For the non-separable data, the above described algorithm can be applied under the constraints of Eqs. (15) and (16) by introducing additionally and when it is necessary a further cost function to the primal objective function by introducing positive slack variables ξ_i, $i = 1, 2, \ldots, l$, in the constraints [74,75]:

$$(x_i.w + b) \geqslant +1 - \xi_i \quad \text{for } y_i = +1 \tag{18}$$

$$(x_i \cdot w + b) \leqslant -1 + \xi_i \quad \text{for } y_i = -1 \tag{19}$$

$$\xi_i \geqslant 0 \quad \forall i \tag{20}$$

where ξ_i must exceed unity in training error cases and $\sum \xi_i$ forms an upper bound on the number of training error.

The above can be generalized for the non-linear case [76,77] according to which training data are produced as dot products x_i, x_j. We are looking for a mapping ϕ:

$$\phi : R^d \mapsto H$$

where H is a Euclidean (or in general a Hilbert) space of infinite dimension. The training algorithm depends on the dot products defined on H, i.e. $\phi(x_i)$. $\phi(x_j)$. If there was a *kernel function* K

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \tag{21}$$

we could use K for training even if we do not know function ϕ. The problem is still a linear separation problem but in a different space. It is easy to find *kernel functions* such that the training algorithm and solution are independent of the dimension of H. Such kernels must hold Mercer's condition [74] which tells us whether or not a perspective kernel is a dot product in some space.

Kernel functions used in pattern recognition problems are

Linear	$K(x, y) = x \cdot y$
Polynomial	$K(x, y) = (x \cdot y + 1)^p$
Gaussian radial basis functions	$K(x, y) = \mathrm{e}^{-\|x-y\|^2/2\sigma 2}$
Sigmoidal neural network	$K(x, y) = \tanh(\kappa x \cdot y - \delta)$

where p, σ, κ, δ are constants.

6. Experimental study

6.1. Implementation

Sixty colonoscopic video frame sequences of closely captured early stage adenomatous polyps, corresponding to different patients, were used in our experiments, according to the physicians' recommendations. The colonoscopic videos were provided by the Gastroenterology Section, Department of Pathophysiology, Medical School, University of Athens and partially by the Section of Minimal Invasive Surgery, University of Tübingen. A representative sample of the video sequences is illustrated in Fig. 4. Each video frame is raster scanned with 2401 overlapping windows of 128×128 pixels size and eight pixels overlap. The window size is relatively large compared to the frame size, in order to obtain adequate statistical population for the calculation of the second-order statistical measures. The relatively small overlap (16 times smaller than the window size) has been chosen to increase the accuracy in the determination of the abnormal regions. The co-occurrence matrix dimension was limited to 64. The use of larger cooccurrence matrices leads to insignificant improvement of the results and disproportional increase of the overall computational complexity.

The colonoscopic video frames were acquired with an Olympus CF-100HL endoscope, and digitized at 512×512 and 24 bit color depth (8 bit per channel). A Pentium IV 1.4 GHz was used for the

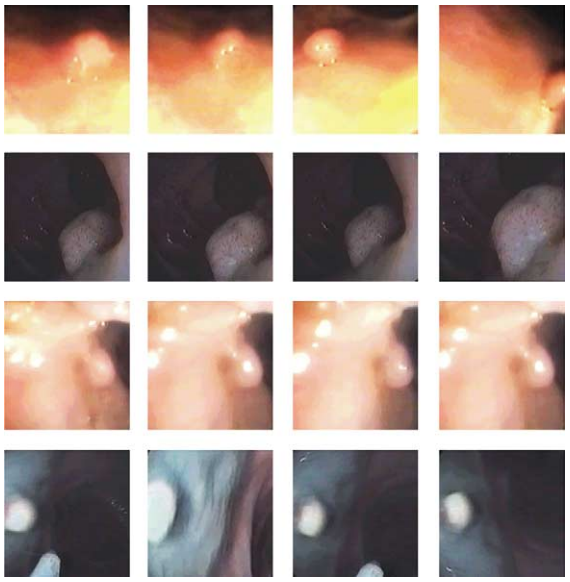

Fig. 4. Four representative cases of adenomatous polyps.

processing task (Fig. 1). The software implementation of the proposed framework was written completely in C++ under Microsoft Windows 98/NT operating systems. It uses Intel performance library functions [78], which provide high performance for complex computational tasks on Intel microprocessors. The modules implementing the SVM classifier are based on functions of the LIBSVM library [79]. The high processing speed provided by the hardware and the efficient software implementation was essential because multiple frames had to be processed in a reasonable time for a large amount of experiments.

6.2. Classification performance criteria

In many medical decision support applications that involve classification of negative (normal) and positive (abnormal) patterns, the data sets are highly imbalanced. The patterns corresponding to healthy instances are more than those corresponding to pathological. In the video frames of our experiments, the proportion of abnormal to normal patterns reaches 10% on average. *Accuracy*, i.e the ratio of correctly classified patterns to the total number of patterns in a data set, is a classic metric for the evaluation of the classification performance in machine learning. However, it assumes that the class distribution is unknown, unchanging and that the costs of a false positive and a false negative are equal. These assumptions make it unreliable when imbalanced data sets are involved [80]. In such cases, the correct classification of a minority class pattern is usually much more important than the correct classification of a pattern belonging to the majority class, especially when the classification task is oriented to the identification of pathological situations. To formulate more reliable criteria of the performance of pattern recognition systems, statisticians use the confusion matrix (Table 1). Its columns and rows correspond to the classifier's outcome and the actual characterization of the patterns respectively: a is the number of correctly classified negative patterns (true negatives), b is the number of negative patterns that have been classified as positive (false positives), c is the number of positive patterns that have been classified as negative (false negatives) and d is the number of correctly classified positive patterns (true positives).

If a, b, c and d values are known, the accuracy (A), the true positive (TP) and false positive (FP) rates are formulated as follows:

$$A = \frac{a+d}{a+b+c+d} \cdot 100(\%)$$

$$\mathrm{TP} = \frac{d}{c+d} \cdot 100(\%)$$

$$\mathrm{FP} = \frac{b}{a+b} \cdot 100(\%)$$

The tradeoff between TP and (100 − FP) rates can be used as a reliable criterion for the evaluation of classification performance for imbalanced data sets [80–82], overcoming the disadvantages of using accuracy. TP and (100 − *FP*) rates are also

Table 1
Confusion matrix

	Resulted	
	Negative	Positive
Actual		
Negative	*a*	*b*
Positive	*c*	*d*

known as *sensitivity* (SN) and *specificity* (SP), respectively. *Sensitivity* is the accuracy among positive patterns, while *specificity* is the accuracy among negative 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.

6.3. Results

The objective of this experimental study is to evaluate the performance of the proposed framework and determine the most suitable color model for the identification of colon cancer precursors. The system is initially trained and then is able to identify lesions of similar clinical and histological characteristics. In the following experiments the SVM classifier was trained and consequently used for the recognition of the polyps. Training was performed on 2401 patterns of a video frame, and tested on 7203 patterns of the following frames. Different training frames were used in order to measure the generalization performance of the proposed framework.

Tests were performed for both grayscale and color images, using a Support Vector Machine with radial basis function kernel since such kernels have led to better classification results compared to other kernel mapping techniques [73]. The color models tested were the *RGB*, *rgb*, K–L, *HSV*, *L*a*b** and *L*u*v**, respectively as previously described. The results of these tests are illustrated in Fig. 5, in terms of average sensitivity and specificity of the corresponding video frame set. The classification performance is considered high when the plot points are located closer to the upper left corner of the diagram SN vs. (100 − SP).

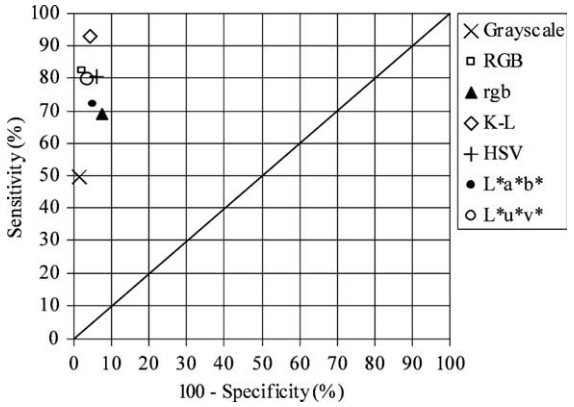

Fig. 5. Average sensitivity and specificity using grayscale and color models.

From the experimentation, we concluded that the highest average (SN, SP) reaching (94%,

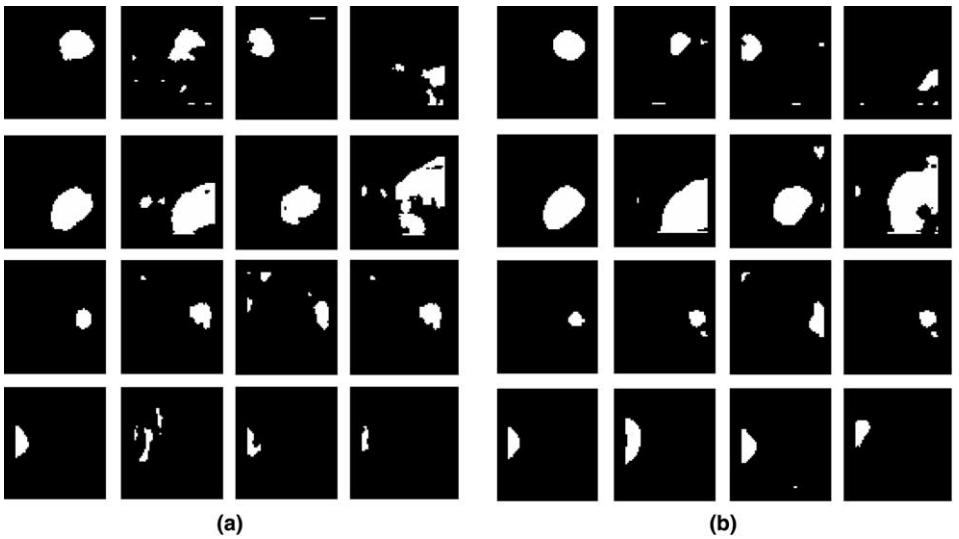

Fig. 6. Classification results corresponding to Fig. 4, (a) using *RGB* and (b) using K–L color models.

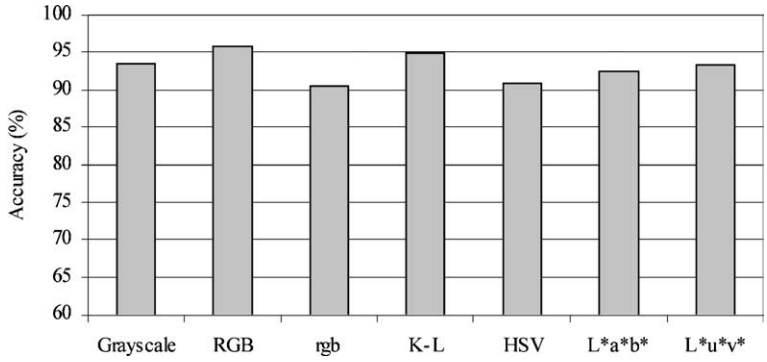

Fig. 7. Average accuracy, using grayscale and color models.

95.7%), was achieved using the K–L color model. The average performance using *RGB* was also high (83.1%, 97.7%) but almost equivalent to the perceptual *HSV* (81.6%, 93.9%) and *L*u*v** (80.7%, 96.3%). *L*a*b**(73.3%, 95.1%) did not perform as well as the other perceptually uniform model *L*u*v**. Worse results (70.1%, 92.4%) were achieved using the normalized *RGB* model, possibly due to the low intensity levels appearing in some frames. The use of grayscale video frames (single 8 bit intensity channel) resulted in low sensitivity but high specificity (50.6%, 98.6%).

Comparing the reconstructed frames after the application of the SVM algorithm for the recognition of polyps, the use of K–L color transformation leads to more accurate results than the corresponding *RGB* spaces. The original images used are shown in Fig. 4 and the reconstructed ones are illustrated in Fig. 6a and Fig. 6b, respectively. K–L model results in more accurate identification of lesions, having well defined borders and less misleading false classified regions.

The average accuracy is high for all models, Fig. 7, exceeding 90%. However its distribution over the different color models is not in agreement with the results provided using sensitivity and specificity. For instance, the use of K–L color model gives high accuracy in the recognition of polyps 94.9%

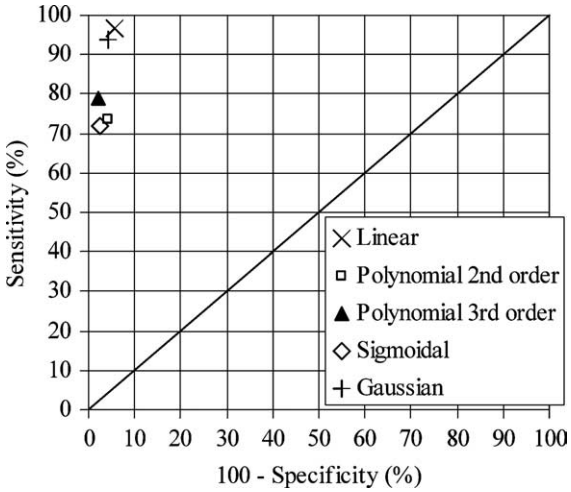

Fig. 8. Average sensitivity and specificity using different SVM kernels.

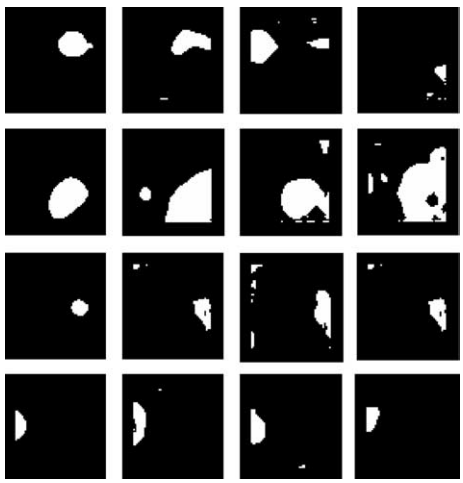

Fig. 9. Classification results corresponding to Fig. 4, using K–L color model and linear kernel SVM.

whereas the use of *RGB* leads to a marginally higher accuracy than K–L, reaching 95.7%.

The performance in polyp's identification was studied along with the use of different SVM kernel functions. Linear, second and third order polynomial, Gaussian and sigmoidal kernels were tested in order to determine the most suitable for the K–L CWC set of measures. As it is illustrated in Fig. 8, SVMs with Linear or Gaussian kernel function produce contiguous points closer to the upper left corner of the diagram. However, the use of linear kernel results in higher average sensitivity but lower specificity. This has the effect of many small white regions (false positives) to appear in the reconstructed images (Fig. 9). These can possibly be misleading indications for the interpretation of the colonoscopic findings. Using Gaussian kernel functions in the SVM algorithm results in more preferable reconstructed images, since they contain less false positives due to higher specificity than the linear kernel. The requirements of the application, as these have been agreed on with the medical experts, suggest a higher specificity to sensitivity ratio.

7. Conclusions

We have proposed a novel framework for the identification of colon cancer precursors using multichannel (color) video endoscopy measurements and SVMs. We presented a number of observations based on different measurements of various experimentations with real data. An approach like this can be used to assist colon endoscopists to accurately locate early stage adenomatous polyps and thus decrease the probability of missinterpratations during the clinical examination. The *R*, *G* and *B* color channels of the endoscopic imaging systems can adequately describe the colonic mucosa, but textural alterations of the colonic mucosal surface (pit patterns) can be used to differentiate normal from abnormal tissue by the endoscopists. The textural features proposed were based on the covariance of the second-order statistical measures estimated on the wavelet domain between the color channels and have been efficiently used for the discrimination between normal/abnormal tissue.

Different color models have been applied to enhance the discriminative properties of the normal/abnormal mucosal surface. The highest generalization performance was achieved using the K–L model. This leads to the conclusion that the orthogonality of the color model used is of high importance for the characterization of the colonic mucosa texture in color video frames. The classification results were evaluated in terms of sensitivity and specificity. Among popular SVM kernels, the Gaussian was found to perform better for the pattern classification problem discussed. The average sensitivity and specificity estimated on 60 video frame sequences of real endoscopy data was 94% and 95.7%, respectively. The identification of adenomatous polyps of the colon can be performed accurately by measuring and processing the second order textural information of color endoscopic video frames in the wavelet domain.

Acknowledgment

This work was partially funded by National and Kapodistrian University of Athens, Special Account of Research Grants. We would like to acknowledge Dr. M.O. Schurr M.D., University of Tübingen Section of Minimal Invasive Surgery, for provision of a part of endoscopic videos used in our study and his contribution in the evaluation of the results.

Appendix A. In this appendix we present the formulas for the color transformations used in this paper [29,42,51,57].

(a) The coordinates of the *rgb* (normalized *RGB*) model are calculated by dividing each of the *RGB* coordinates by the summation of Red, Green and Blue

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

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

where $r + g + b = 1$.

(b) The conversion procedure from *RGB* to *HSV* includes the following steps:

(i) Normalization of the *RGB* values to [0, 1]

$$M = \max(R, G, B)$$
$$m = \min(R, G, B)$$
$$r = (M - R)/(M - m)$$
$$g = (M - G)/(M - m)$$
$$b = (M - B)/(M - m)$$

(ii) Calculate Value $V \in [0, 1]$

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

(iii) Calculate Saturation $S \in [0, 180]$

$$\text{if } (M = 0) \quad \text{then } S = 0 \quad \text{and } H = 180$$
$$\text{if } (M \neq 0) \quad \text{then } S = (M - m)/M$$

(iv) Calculate Hue $H \in [0, 360]$

$$\text{if } (R = M) \quad \text{then } H = 60(b - g)$$
$$\text{if } (G = M) \quad \text{then } H = 60(2 + r - b)$$
$$\text{if } (B = M) \quad \text{then } H = 60(4 + g - r)$$
$$\text{if } (H \geqslant 360) \quad \text{then } H = H - 360$$
$$\text{if } (H < 0) \quad \text{then } H = H + 360$$

(c) Both $L^*a^*b^*$ and $L^*u^*v^*$ models are based on the CIE-*XYZ* standard color model, which is estimated as a linear transform of the *RGB* coordinates as follows:

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

The lightness L^*, and chroma coordinates a^*, b^* are calculated as

$$L^* = \begin{cases} 116 \cdot \left(\frac{Y}{Y_n}\right)^{1/3} - 16, & Y/Y_n > 0.008856 \\ 903.3 \cdot \left(\frac{Y}{Y_n}\right), & Y/Y_n \leqslant 0.008856 \end{cases}$$

$$a^* = 500[f(X/X_n) - f(Y/Y_n)]$$
$$b^* = 200[f(Y/Y_n) - f(Z/Z_n)]$$

where $f(t)$ is estimated as

$$f(t) = \begin{cases} t^{1/3}, & t > 0.008856 \\ 7.787 \cdot t + 16/116, & t \leqslant 0.008856 \end{cases}$$

where X, Y, Z refer to the color considered, X_n, Y_n, Z_n refer to a suitably chosen reference white (in this case CIE standard illuminant D_{65} obtained by setting $R = G = B = 100$ in *RGB* to *XYZ* transformation), and $t \in \{X/X_n, Y/Y_n, Z/Z_n\}$.

(d) The $L^*u^*v^*$ model's coordinates are calculated by the following equations

$$L^* = 116 \cdot \left(\frac{Y}{Y_n}\right)^{1/3} - 16$$
$$u^* = 13[L^*(u' - u_n)]$$
$$v^* = 13[L^*(v' - v_n)]$$

where

$$u' = [4X/(X + 15Y + 3Z)]$$
$$v' = [9Y/(X + 15Y + 3Z)]$$

and X, Y, Z refer to the color considered, X_n, Y_n, Z_n, v_n, u_n refer to a suitably chosen reference white (in this case CIE standard illuminant D_{65} obtained by setting $R = G = B = 100$ in *RGB* to *XYZ* transformation) and v_n, u_n are calculated by the same equations as u', v'.

(e) The K–L color model's coordinates (K_1, K_2 and K_3) can be calculated as a linear transformation of *RGB*:

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

References

[1] S. Parker, T. Tong, S. Bolden, P. Wingo, Cancer statistics 1997, CA Cancer Journal for Clinicians 47 (1997) 5–27.

[2] M.J. O'Brien, S.J. Winawer, A.G. Zauber et al., The national polyp study: patient and polyp characteristics associated with high grade dysplasia in colorectal adenomas, Gastroenterology 98 (1990) 371–379.

[3] S.H. Itzkowitz, Y.S. Kim, Sleisinger and Fordtran's Gastrointestinal and Liver Disease, 6th ed., vol. 2, WB Saunders Company; 1998, Philadelphia, PA.

[4] D. Rex, R. Weddle, D. Pound, K. O'Connor, R. Hawes, R. Dittus, J. Lappas, L. Lumeng, Flexible sigmoidoscopy plus air contrast barium enema versus colonoscopy for suspected lower gastrointestinal bleeding, Gastroenterology 98 (1990) 855–861.

[5] D. West, V. West, Model selection for a medical diagnostic decision support system: a breast cancer detection case, Artificial Intelligence in Medicine 20 (2000) 183–204.

[6] C.M. Wu, Y.C. Chen, Multi-threshold dimension vector for texture analysis and its application to liver tissue classification, Pattern Recognition 26 (1) (1993) 137–144.
[7] H. Sujana, S. Swarnamani, S. Suresh, Artificial neural networks for the classification of liver lesions by image texture parameters, Ultrasound Medicine in Biology 22 (9) (1996) 1177–1181.
[8] S.B. Premkumar, A.G. Houston, D.E. Pitts, R.J. Babaian, R.B. Evans, Statistical description of prostate cancer from ultrasound imagery, Journal of Ultrasound in Medicine 11 (3) (1992) 576–584.
[9] F. Lachmann, C. Barillot, Brain tissue classification from MRI data by means of texture analysisProceedings of the of Medical Imaging: VI Image Processing, vol. 1652, SPIE Press, Newport Beach CA, 1992, pp. 72–83.
[10] S. Bovis, S. Singh, Detection of masses in mammograms using texture measuresProceedings of the 15th International Conference on Pattern Recognition, Barcelona, vol. 2, IEEE Press, New York, 2000, pp. 267–270.
[11] N.R. Mudigonda, R.M. Rangayyan, J.E.L. Desautels, Gradient and texture analysis for the classification of mammographic masses, IEEE Transactions of the Medical Imaging 19 (10) (2000) 1032–1043.
[12] C. Fortin, W. Ohley, Automatic segmentation of cardiac images: texture mapping, in: Proceedings of the IEEE 17th Annual Northeast Bioengineering Conference, Hartford, CT, 1991.
[13] Q. Ji, J. Engel, E. Craine, Texture analysis for classification of cervix lesions, IEEE Transactions of the on Medical Imaging 19 (11) (2000) 1144–1149.
[14] S. Fischer, P. Schmid, J. Guillod, Analysis of skin lesions with pigmented networks, in: Proceedings of the International ICIP Conference on Image Processing, Lausanne, Switzerland, , vol. 1, 1996, pp. 323–326.
[15] H. Ganster, A. Pinz, R. Rohrer, E. Wildling, M. Binder, H. Kittler, Automated melanoma recognition, IEEE Transactions of the Medical Imaging 20 (3) (2001) 233–239.
[16] T.W. Pope, W.L. Williams, S.B. Wilkinson, M.A. Gordon, Comparisons in Biomed Research 29 (1996) 429–437.
[17] A.N. Esgiar, R.N.G. Naguib, B.S. Sharif, M.K. Bennett, A. Murray, Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa, IEEE Transactions of the Information Technology in Biomedicine 2 (1998) 197–203.
[18] S. Karkanis, G.D. Magoulas, N. Theofanous, Image recognition and neuronal networks: intelligent systems for the improvement of imaging information, Minimal Invasive Therapy and Allied Technologies 9 (2000) 225–230.
[19] S.A. Karkanis, G.D. Magoulas, D.K. Iakovidis, D.A. Karras, D.E. Maroulis, Evaluation of textural feature extraction schemes for neural network-based interpretation of regions in medical images, in: Proceedings of the International ICIP Conference on Image Processing, Thessaloniki, Greece, 2001, pp. 281–284.
[20] S.A. Karkanis, D.K. Iakovidis, D.A. Karras, D.E. Maroulis, Detection of lesions in endoscopic video using textural descriptors on wavelet domain supported by artificial neural network architectures, in: Proceedings of the International ICIP Conference on Image Processing, Thessaloniki, Greece, 2001 pp. 833–836.
[21] S.A. Karkanis, G.D. Magoulas, D.K. Iakovidis, D.E. Maroulis, N. Theofanous, Tumor recognition in endoscopic video images, in: Proceedings of the 26th EUROMICRO Conference, Maastricht, Netherlands, 2000, pp. 423–429.
[22] S. Karkanis, G.D. Magoulas, M. Grigoriadou, M. Schurr, Detecting abnormalities in colonoscopic images by textural description and neural networks, in: Proceedings of the ACAI Advance Course in Artificial Intelligence, Chania, Greece, 1999, Workshop on Machine Learning in Medical Applications, pp. 59–62.
[23] S. Nagata, S. Tanaka, K. Haruma, M. Yoshihara, K. Sumii, G. Kajiyama, F. Shimamoto, Pit pattern diagnosis of early colorectal carcinoma by magnifying colonoscopy: clinical and histological implications, International Journal of Oncolology 16 (2000) 927–934.
[24] S. Kudo, H. Kashida, T. Tamura, E. Kogure, Y. Imai, H. Yamano, A.R. Hart, Colonoscopic diagnosis and management of nonpolypoid early colorectal cancer, World Journal of Surgery 24 (2000) 1081–1090.
[25] J.G. Daugman, Computational Neuroscience, MIT Press, Cambridge, 1990.
[26] T. Randen, J.H. Husoy, Filtering for texture classification: a comparative study, IEEE Transactions of the Pattern Analysis and Machine Intelligence 21 (4) (1999) 291–310.
[27] G. Paschos, Chromatic correlation features for texture recognition, Pattern Recognition Letters 19 (1998) 643–650.
[28] G. Paschos, Perceptually uniform color spaces for color texture analysis: an empirical evaluation, IEEE Transactions of the Image Processing 10 (6) (2001) 932–937.
[29] G. Van de Wouwer, P. Scheunders, S. Livens, D. Van Dyck, Wavelet correlation signatures for color texture characterization, Pattern Recognition 32 (1999) 443–451.
[30] V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, 1995.
[31] S. Haykin, Neural Networks: A Comprehensive Foundation, second ed., Prentice-Hall, Englewood Cliffs, NJ, 1999.
[32] E. Osuna, R. Freund, F. Girosi, Training support vector machines: an application to face detection, Proceedings of the Computer Vision and Pattern Recognition (1997) 130–136.
[33] Y. LeCun, L.D. Jackel, L. Bottou, A. Brunot, C. Cortes, J.S. Denker, H. Drucker, I. Guinon, U.A. Muller, E. Sackinger, P. Simard, V. Vapnik, Comparison of learning algorithms for handwritten digit recognition, in: F. Fogelman, P. Gallinari (Eds.), Proceedings of the International Conference on Artificial Neural Networks, 1995, pp. 53–60.
[34] D.A. Karras, S.A. Karkanis, D. Iakovidis, D.E. Maroulis, B.G. Mertzios, Support vector machines for improved defect detection using novel multidimensional wavelet feature extraction involving vector quantization and PCA techniques, in: NIMIA Adnvanced Study Institute on

Neural Networks for Instrumentation, Measurement and Related Industrial Applications, Crema, Italy, 2001, pp. 139–144.
[35] K. Veropoulos, N. Cristianini, C. Campbell, The application of support vector machines to medical decision support: a case study, in: Proceedings of the ACAI Advanced Course in Artificial Intelligence, Chania, Greece, Workshop W10, 1999, pp. 17–21.
[36] T.S. Furey, N. Christianini, N. Duffy, D.W. Bednarski, M. Schummer, D. Haussler, Support vector machine classification and validation of cancer tissue samples using microarray expression data, Bioinformatics 16 (10) (2000) 906–914.
[37] Y. Miyake, N. Tsumura, M. Takeya, R. Inagawa, in: H. Tanaka, Y. Miyake, M. Nishibori, D. Mukhopadhyay (Eds.), Digital Color Images in Biomedicine, Japan 2001.
[38] C. Chiu, A novel approach based on computerized image analysis for traditional Chinese Medical diagnosis of the tongue, Computer Methods and Programs in Biomedicine 61 (2000) 77–89.
[39] G. Hamarneh, A. Chodorowski, T. Gustavsson, Active contour models: application to oral lesion detection in color images, in: IEEE Conference On Systems Man, and Cybernetics, vol. 4, 2000, pp. 2458–2463.
[40] T. Shiobara, H. Haneishi, Y. Miyake, Improved color reproduction of electronic endoscopes, Journal of Imaging Science and Technology 40 (1996) 494–501.
[41] D.G. Chamberlin, Colour: Its Measurement Computation and Application, Heyden, 1980.
[42] G. Wyszecki, W.S. Styles, Color Science: Concepts and Methods, Quantitative Data and Formulae, John Wiley & Sons, New York, 1982.
[43] T. Gevers, W.M. Smeulders, Pic-To-Seek: combining color and shape invariant features for image retrieval, IEEE Transactions of the Image Processing 9 (1) (2000) 102–119.
[44] P. Duchnowski, M. Hunke, D. Busching, M. Meier, A. Waibel, Toward moment-invariant automatic lip-reading and speech recognition, in: Proceedings of the ICASSP, vol. 1, 1995, pp.109–112.
[45] H.P. Graf, E. Cosatto, D. Gibbon, M. Kocheisen, E. Petajan, Multi-modal system for locating heads and faces, in: Proceedings of the 2nd International Conference on Automatic Face and Gesture Recoginition, Killington, Vermont, 1996, pp. 88–93.
[46] S.H. Kim, N.K. Kim, S.C. Ahn, H.G. Kim, Object oriented face detection using range and color information, in: Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998, pp. 76–81.
[47] B. Schiele, A. Waibel, Gaze tracking based on face-color, in: Proceedings of the International Workshop on Automatic Face and Gesture Recognition, Zurich, 1995, pp. 344–349.
[48] Q.B. Sun, W.M. Huang, J. K. Wu, Face detection based on color and local symmetry information, in: Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998, pp. 130–135.
[49] J. Yang, A. Waibel, Tracking human faces in real time, Technical Report CMU-CS-95-210, C.M.U., 1995.
[50] T. Gevers, A.W.M. Smeulders, H. Stockman, Photometric invariant region detection, in: Proceedings of the 9th British Machine Vision Conference, University of Southampton, UK, 14–17 September, 1998.
[51] Y. Ohta, T. Kanade, T. Sakai, Color information for region segmentation computer graphics and image processing (13) (1980) 222–241.
[52] A.R. Smith, Integrated spatial and feature image systems: retrieval, compression and analysis, Ph.D. dissertation, Columbia University, New York, 1997.
[53] K. Sobottka, I. Pitas, Extraction of facial regions and features using color and shape information, in: Proceedings of the ICIP 96, vol. 3, 1996, pp. 483–486.
[54] C.H. Lee, J.S. Kim, K.H. Park, Automatic human face location in a complex background using motion and color information, Pattern Recognition 29 (1996) 1877–1889.
[55] S. Grosskopf, P. Neugebauer, H. Schumann, Plaque measurement from intra-oral video frames, in: Proceedings of the IADMFR/CMI'97 Advances in Maxillofacial Imaging, Louisville, Kentucky, 1997, pp. 89–94.
[56] C. Palm, D. Keysers, K. Spitzer, Gabor Filtering of complex hue/saturation images for color texture classification, in: Joint Conference on Information Sciences—International Conference on Computer Vision, Pattern Recognition, and Image Processing, Atlantic City, NJ, USA, vol. 2, February 2000, pp. 45–49.
[57] J. Keith, Video Demystified, second ed., Hightext Interactive, 1996.
[58] L. Shafarenko, H. Petrou, J. Kittler, Histogram-based segmentation in a perceptually uniform color space, IEEE Transactions of the Image Processing 7 (1998) 1354–1358.
[59] A. Mojsilovic, J. Kovacevic, J. Hu, R.J. Safranek, K. Ganapathy, Retrieval of color patterns based on perceptual dimensions of texture and human similarity rules, in: Proceedings of the Human Vision and Electronic Imaging, vol. 3644, SPIE, San Jose, 1999, pp. 441–452.
[60] R.P. Schumeyer, K.E. Barner, Color-based classifier for region identification in video, in: Proceedings of the VCIP, vol. 3309, SPIE, pp.189–200.
[61] C.H. Li, Regularized color clustering in medical image database, IEEE Transactions of the Medical Imaging 19 (2000) 1150–1155.
[62] K.Y. Song, J. Kittler, M. Petrou, Defect detection in random color textures, Imaging and Visual Computation 14 (9) (1996) 667–684.
[63] R.M. Haralick, Statistical and structural approaches to texture, in: Proceedings of the IEEE, vol. 67, 1979, pp. 786–804.
[64] L. Van Gool, P. Dewaele, A. Osterlinck, Texture Analysis Computer Vision Graphics Image Processing 29 (12) (1983) 336–357.
[65] T.R. Reed, J.M.H. Hans Du Buf, A review of recent texture segmentation and feature extraction techniques, Computer Vision Graphics Image Process: Image Understanding 57 (3) (1993) 359–372.

[66] M. Mirmehdi, M. Petrou, Segmentation of color textures, IEEE Transactions of the Pattern Analysis and Machine Intelligence 22 (2) (2000) 142–159.
[67] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transactions of the on Pattern Analysis and Machine Intelligence 11 (1989) 674–693.
[68] B. Sahiner, H.P. Chan, N. Petrick, D. Wei, M.A. Helvie, D.D. Adler, M.M. Goodsitt, Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images, IEEE Transactions of the Medical Imaging 15 (1996) 598–610.
[69] Q.A. Holmes, D.R. Nuesch, R.A. Shuchman, Texture analysis and real-time classification of sea-ice types using digital SAR data, IEEE Transactions of the Geoscience Remote Sensing 22 (1984) 113–120.
[70] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Transactions of the Systems, Man and Cybernetics 3 (1973) 610–621.
[71] A. Latif-Amet, A. Ertuzun, A. Ercil, An efficient method for texture defect detection: sub-band domain cooccurrence matrices, Image and Vision Computing 18 (2000) 543–553.
[72] N. Vandenbroucke, L. Macaire, J.G. Postaire, Unsupervised color texture feature extraction and selection for soccer image segmentation, in: Proceedings of the ICIP'2000, IEEE International Conference on Image Processing, Vancouver Canada, vol. 2, 2000, pp. 800–803.
[73] C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, Kluwer Academic Publishers, Boston, 1998.
[74] V. Vapnik, Estimation of Dependencies based on Empirical Data, Springer Verlag, New York, 1982.
[75] C. Cortes, V. Vapnik, Support vector networks, Machine Learning 20 (1995) 273–297.
[76] A. Aizerman, B. Braverman, L. Rozoner, Theoretical foundations of the potential function method in pattern recognition learning, Automation and Remote Control 25 (1964) 821–837.
[77] B. Boser, I. Guyon, V. Vapnik, A training algorithm for optimal margin classifiers, in: ACM 5th Annual Workshop on Computational Learning Theory, Pittsburgh, 1992, pp. 144–152.
[78] Intel Performance Libraries (WWW page), http://developer.intel.com/software/products/, Accessed on January 2002.
[79] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines. Software Available from http://www/csie.ntu.edu.tw/~cjlin/libsvm Accessed: Jan. 2002.
[80] M. Kubat, S. Matwin, Addressing the curse of imbalanced training sets: one-sided selection, in: Proceedings of the 14th International Conference Machine Learning, Morgan Kaufmann, Los Altos, CA, 1997, pp. 179–186.
[81] J.A. Swets, R.M. Dawes, J. Monahan, Psychological science can improve diagnostic decisions, Psychological Science in the Public Interest 1 (2000) 1–26.
[82] B. Mac Namee, P. Cunningham, S. Byrne, O.I. Corrigan, The problem of bias in training data in regression problems in medical decision support, AI in Medicine 25 (1) (2002) 51–70.