journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Texture and color based image segmentation and pathology detection in capsule endoscopy videos

Piotr Szczypiński[a,*], Artur Klepaczko[a], Marek Pazurek[b], Piotr Daniel[b]

[a] Technical University of Lodz, Institute of Electronics, Medical Electronics Division, 90-924 Lodz, ul. Wolczanska 211/215, Poland
[b] Medical University of Lodz, Faculty of Medicine, Department of Digestive Tract Diseases, 90-153 Lodz, ul. Kopcinskiego 22, Poland

ARTICLE INFO

Article history:
Received 19 April 2012
Received in revised form
12 September 2012
Accepted 13 September 2012

Keywords:
Capsule endoscopy
Feature selection
Texture analysis
Medical image analysis

ABSTRACT

This paper presents an in-depth study of several approaches to exploratory analysis of wireless capsule endoscopy images (WCE). It is demonstrated that versatile texture and color based descriptors of image regions corresponding to various anomalies of the gastrointestinal tract allows their accurate detection of pathologies in a sequence of WCE frames. Moreover, through classification of single pixels described by texture features of their neighborhood, the images can be segmented into homogeneous areas well matched to the image content. For both, detection and segmentation tasks the same procedure is applied which consists of features calculation, relevant feature subset selection and classification stages. This general three-stage framework is realized using various recognition strategies. In particular, the performance of the developed Vector Supported Convex Hull classification algorithm is compared against Support Vector Machines run in configuration with two different feature selection methods.

1. Introduction

Wireless capsule endoscopy (WCE) has gained a firm position in the catalogue of examination tools available in modern gastroenterology [1–3]. It is widely accepted by patients who consider it less traumatic than traditional gastroscopy or colonoscopy. Secondly, it provides the means to optically and non-invasively visualize interior of the small intestine. Moreover, capsule endoscopy proved useful in diagnosing sources of occult bleeding or causes of gastroesophageal reflux and abdominal pain such as Lesniewski–Crohn's disease, erosive lesions or peptic ulcers [4–8]. On the other hand however, clinicians claim that usage of WCE technique is a time consuming task – the video capture itself takes about 8 h. Then, even if the recorded sequence is played back with larger frame rate than the recording tempo, the diagnostic session occupies at least 2 h for a trained physician and as such it requires high level of concentration. These disadvantages motivate efforts leading to automation of WCE video inspection and they are addressed also in this work.

One can distinguish three principal research directions that aim at solving the problem of computer-aided WCE diagnosis. One approach focuses on segmentation of WCE video sequence into integral parts of the gastrointestinal tract [9–12]. This enables narrowing the preview of the recorded sequence only to an interesting part (e.g. small intestine) and neglect the other pieces (like oesophagus, stomach or colon).

Another direction of investigations concerns the issue of movement speed of a capsule inside the digestive system [13]. This kind of information can be useful when describing intestinal peristalsis or detecting occlusions. In [14] a special case of deformable model (the model of deformable rings) is introduced and the capsule speed estimation is based on the analysis of contents variation between subsequent WCE frames. This in turn allows to smoothly adjust the playback

* *Corresponding author*. Tel.: +48 42 6312642; fax: +48 42 6362238.
E-mail address: pms@p.lodz.pl (P. Szczypiński).
0169-2607/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cmpb.2012.09.004

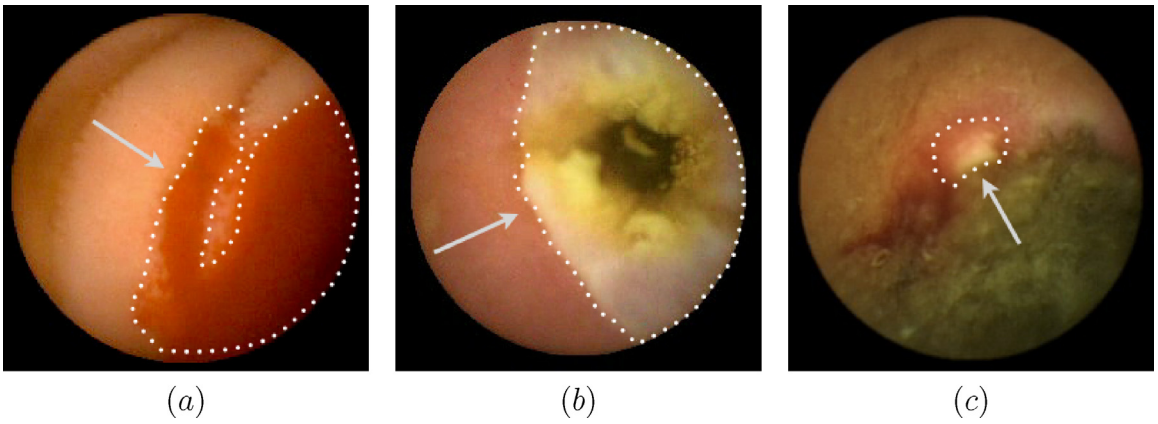

Fig. 1 – Sample WCE images showing abnormality structures: (a) bleeding, (b) excessive and (c) focal ulcerations.

ratio depending on the state of a capsule – if it moves fast and there are sudden changes in succeeding images, the video is slowed down accordingly, otherwise the video is played back at higher frame-rates.

Eventually, in a series of studies the task was to analyze each frame of a given video sequence and detect images that contain certain pathological alterations in internal body organs [15–24]. This is performed with the use of a variety of pattern recognition and image processing techniques. The latter aim at generating proper image characteristics, often taking into account color information derived from various color models. These characteristics form the basis for classifying image regions as representing healthy tissue or abnormal structures. Some examples of this approach include detection of ulcers, bleedings, petechiae and tumors with the use of color and texture analysis and various classification algorithms such as minimum distance classifier, k-nearest-neighbor or support vector machines.

This paper falls into the third category of research directions outlined above. Similarly to other approaches we explore local texture characteristics of a WCE frame describing its pixels by a set of texture and color-dependent features. In particular we aim at designing a tool capable of efficiently detecting bleedings and different forms of ulcerations (especially focal and excessive ones – see Fig. 1 for sample images). There are however a few properties of our strategy which make it a unique concept. Firstly, we analyze WCE examinations at two viewpoints realizing two independent goals: pathology detection and image segmentation. For detection, images are divided into overlapping circular regions of interest covering whole field of view (see Fig. 2). These regions are then described by feature vectors and classified into one of abnormality categories. In the segmentation task images are analyzed pixel-wise, and thus every image element obtains its own vector of features. Per-pixel classification allows distinguishing fine-grained segments of images giving the opportunity to further process delineated regions. This post-processing may include, e.g. shape analysis or quantification of pathological alterations leading to potentially better understanding of a given disease case.

Secondly, no specific texture or color space model is a priori preferred in the proposed method. Instead, every pixel is ascribed a set of up to 2600 features which fully characterize its proximal neighborhood and which are derived from versatile texture models (such as histogram, grey level co-occurrence matrix, run-length matrix, image gradient, autoregressive model and Gabor wavelets) calculated. All these features are calculated for various color components – RGB, YUV, YIQ, HSB, CIE L*a*b, and CIE XYZ. Additionally, Y-relative normalized chrominance components are also examined in order to make texture features independent from the non-uniform lighting conditions. Furthermore, features vary between one another by the size and shape of their local neighborhoods taken into account during calculations. This approach assumes that in such a huge-dimensional feature space, there exist a few significant descriptors which enable accurate tissue classification. In order to identify these most relevant features, we boost the classifier training stage with the attribute selection procedure. The final procedures of pathology detection and WCE video frame segmentation are performed only for the selected features which reduces computational effort.

The problem of feature selection – being itself a challenging task [25–30] – is tackled in our study with special attention.

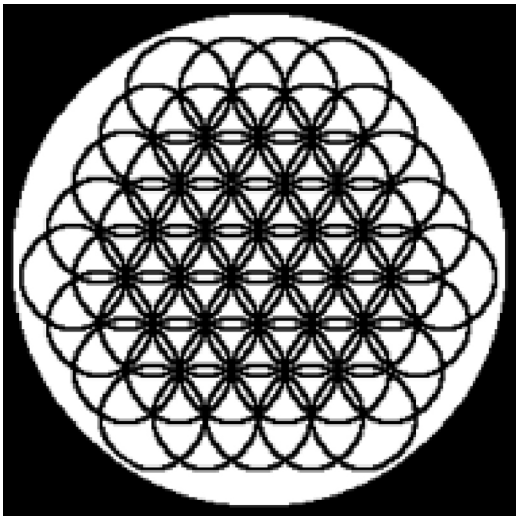

Fig. 2 – Layout of the circular regions of interest used in detection.

In the medical diagnosis context, there appears an additional requirement regarding features discriminative capabilities. It refers to two types of possible classifier errors – false positive (when a normal case is classified as unhealthy) and false negative (the opposite condition). Naturally, the consequences of the second type errors are more risky for a patient. Thus, we developed a novel feature selection algorithm (called hereafter the vector-supported convex hull method, VSCH) designed to minimize the false negative rate, i.e. to maximize sensitivity of the proposed WCE computer-aided diagnosis tool.

The above considerations are discussed more closely in subsequent sections of this paper as follows. Section 2 recalls texture analysis and color space transformation techniques with the focus on methods which have found application in biomedical image processing. A concise review of feature selection algorithms is presented in Section 3. The designed VSCH method is also described therein. In Section 4, we demonstrate the results of experiments which validate the proposed approach. Finally, Section 5 concludes.

2. Feature computation

An image feature is a numerical indicator somehow characterizing properties of the image or its fragment. Therefore, a feature computation is a procedure which examines the image to establish a value of such the indicator. There are various algorithms for computation of features that focus on characterization of various aspects of image appearance, such as brightness, contrast, color or texture. Usually to describe an image appearance comprehensively, an ordered assortment of several features is used – called a feature vector. Further steps of image examination may involve analysis of such vectors distributions within a vector space. Therefore, the application of image feature computation algorithms followed by the methods of pattern recognition is a way of indirect image analysis.

The goals of image analysis range from image recognition, through detection or localization of specific details within the image to the image segmentation. In result and as mentioned above, the features are computed to characterize the image as a whole, to characterize some selected image region or to indicate or detect specific image property of local nature. For that reason, features can be computed within the area of the whole image, within the region of interest defined to cover some known and specific fragment of the image, or within immediate neighborhood around some particular image location. The latter allows creation of a feature map, which is the feature function in the domain of image coordinates. The map is usually presented as a two-dimensional grey-level representation of a feature value computed at every location of the input image. The maps are used in image segmentation.

2.1. *Texture based analysis*

An image texture perceived by humans is a visualization of complex patterns composed of spatially organized, repeated sub-patterns, which have a characteristic, somewhat uniform appearance. The texture may demonstrate specific brightness, color size, roughness, directivity, randomness, smoothness, granulation, etc. and thus, may carry substantial information about the structure of physical objects. In medical images it describes internal structure of human tissues or organs. Consequently, textural image analysis is an important issue also in medical image processing and understanding.

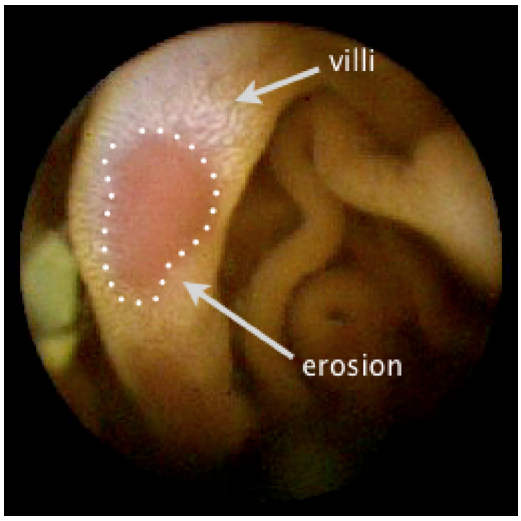

Fig. 3 – Texture patterns discriminating surfaces of erosion and intenstinal villi.

Texture analysis plays important role in interpretation of capsule endoscopy images. Detection of some abnormal changes involves investigation of texture properties. Example of this is healthy small intestine mucosa with protruding the intestinal villi, which create specific pattern within image. Such pattern do not show up in case of abnormalities, such as erosions, ulcers or bleedings (Fig. 3).

Humans usually assess texture only qualitatively, while often quantitative texture analysis is required. To perform such quantification the texture features, mathematically defined texture properties, have to be computed. There is a number of various algorithms to perform such computations. However, as far as wireless endoscope video is concerned, most of publications deal with algorithms defined by the MPEG-7 standard only [11,12,18,19,31–33]. Despite the algorithms of the standard were not particularly designed to characterize medical images, the features computed according to the standard were utilized in pathology detection, topological segmentation or even motion analysis.

There are several problems behind applying MPEG-7 algorithms in medical image analysis. First, they were selected for multimedia content description – in case of images to describe overall content of the image as such, not allowing characterization of local image properties or arbitrary shaped regions of interest. Second, the implementation [34] of the MPEG-7 algorithms can compute image descriptors (features or groups of features) within rectangular regions at least 64×64 (or even 128×128 pixels in case of homogeneous texture descriptor). These relatively big regions are not appropriate for characterization of local image properties.

In our approach to wireless endoscope image analysis the algorithms implemented in MaZda software [35] are used. The program can compute feature maps or feature vectors within

regions of interest of arbitrary, user-defined shape and size. It implements statistical, model-based and image transform algorithms. It calculates statistics of image brightness histogram, image gradient magnitude, grey-level co-occurrence matrix, run-length matrix, parameters of autoregressive model and energies of image signal within frequency bands obtained by means of Haar wavelet transform. In total, there are 279 various features which can be computed. The features are computed for monochromatic images. The utility of the software was demonstrated in numerous applications including the analysis of medical images [35,36].

2.2. Color components and color analysis

It is evident, that in addition to texture, the image color carries information essential for correct interpretation of the image content. For example, in endoscopic images, the color of fibrotic ulcer is white or light-grey, bleedings are red and normal mucosa looks pink. The usual approach to make use of color information is an analysis of color components of color spaces. Most often the analysis is performed on RGB color space, which is a natural representation for displays of computer systems. However, all the three components jointly carry information on image brightness (luminance) and are highly correlated. The analysis of other color spaces, such as YIQ or YUV, which separate color information (chrominance) from luminance, may be more appropriate. Fig. 4 shows an example color image of focal fibrotic ulcer surrounded by erythema and its color components in YIQ color space. Apparently, the larger values of Q component noticeably indicate the pathology.

It is not obvious in advance what color components of which color spaces are the most useful to expose particular kind of abnormality. Thus, we exploit a variety of components of several color spaces such as RGB, YUV, YIQ, HSB, CIE XYZ, CIE L*a*b*. In addition, we examine color component ratios such as U/Y, V/Y, I/Y and Q/Y, in which the impact of illumination intensity may be reduced even further. Therefore, the total number of color components explored in our study is 21.

The MPEG-7 standard defines several descriptors of image color: dominant color descriptor (DCD), scalable color descriptor (SCD), color layout descriptor (CLD) and color structure descriptor (CSD). However, in computation of the DCD the color space has to be arbitrarily chosen beforehand, and as mentioned above the right choice of the color space is not obvious. Moreover, the SCD is computed by means of HSB space only. The CLD is essentially a miniature of the whole image and the CSD is an image histogram in HMMD space, they both extensively describe image, however, can hardly be utilized in image segmentation or classification of small regions.

The natural way to produce exhaustive characterization of the image is to combine color and texture analysis. The MaZda software can be used to compute texture features for particular color component. If the features computed for different color components are joined, it creates up to 5859 features vector (279 texture features in every of 21 color components). After initial experiments with WCE images, we decided to preselect most promising color components and texture description models. Eventually, in the analyses described in Section 4 input feature space was narrowed to 2494 and 2600 attributes in the case of detection and segmentation respectively.

Analysis of vector distributions in such huge-dimensional space is extremely difficult. On the other hand, it can be supposed, the only small subset of all the features may be sufficient for detection of images showing abnormal state of mucosa. Therefore, selection of such features and dimensionality reduction of vector space becomes necessary.

3. Feature selection and classification

Abundance of WCE image descriptors obtained in the feature extraction procedure involves the need to select the most significant attributes, helping to discriminate gastrointestinal pathology classes. In the data mining domain there exist two general approaches to the feature selection problem [37,27]. In the *filter* strategy an attribute saliency is estimated individually according to some evaluation measure. The main advantage of these methods is their low computational complexity as they avoid time-consuming feature space exploration. However, it is often observed that discriminative power of a feature emerges only when it is accompanied by some other attributes. Thus, a subset of individually best features may occur inferior to the actual best possible solution.

In the alternative *wrapper* approach discriminative power of attribute collections is evaluated based on their impact on performance of a given data classification algorithm. Efficiency of wrapper-type feature selection largely depends on the strategy of constructing subsequent feature subsets – candidates for final selection. Especially for high dimension patterns and computationally-demanding classifier training techniques, testing all possible feature combinations is intractable. In that case a search algorithm should intelligently decide which feature subspaces should be examined in order to identify a near-optimal solution without performing a complete search.

Taking into account very high dimension of the proposed texture- and color-based WCE image characteristics and potentially complex structures of pathology class boundaries, three processing pathways have to be examined: (1) filter-type selection method followed by an arbitrarily complex classifier, (2) wrapper-type selection with an arbitrary classifier and suboptimal search strategy and (3) wrapper-type selection with a computationally-efficient classifier and globally optimal search method. In the following we describe our choices for each of the options. Note, that for options (1) and (2) we picked the same classifier – the Support Vector Machines (SVM) algorithm – so that efficiency of both feature selection approaches in application to WCE image processing can be compared in a credible manner. We recall the overall concept of SVM and motivate its choice in Section 3.1. Then, we provide the description of the one-way ANOVA statistics – our chosen filter-type selection method (Section 3.2) – and the SFFS feature space exploration strategy as an example of the wrapper-type technique (Section 3.3). Since for both ANOVA and SFFS selection methods, the ultimate classifier is SVM, we will refer to the entire selection/classification process as SVM/ANOVA and SVM/SFFS respectively. In Section 3.4 we outline the designed VSCH algorithm which performs well with the complete feature search method in lower dimensions (i.e. up to 3D). Larger dimensionalities can be still explored in the

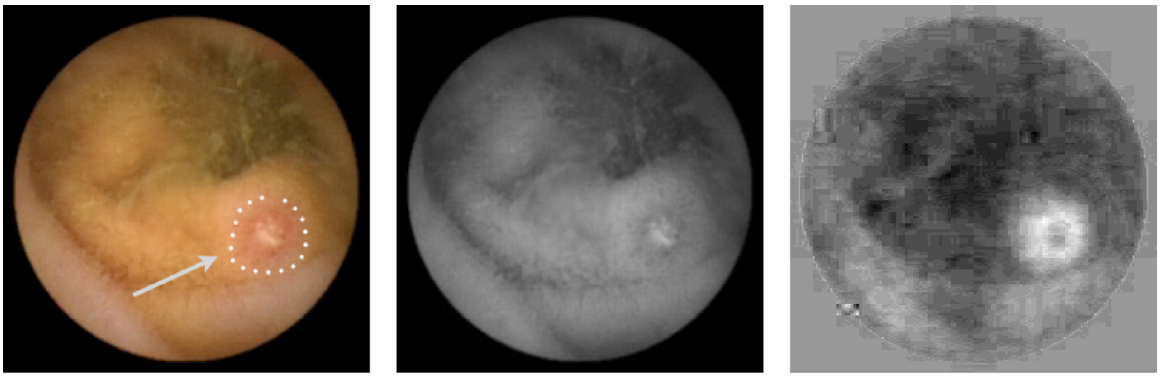

Fig. 4 – Focal ulcer on the small intestine wall: (a) original color image, (b) grey-scale intensity image, (c) Q component image of the YIQ color space model with exposed ring-shaped erythema region around ulceration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

exhaustive manner but on the feature spaces restricted to the best results found so far.

3.1. *Support vector machines*

The choice of a classification algorithm used to process WCE feature vectors during the wrapper-type feature selection method must reflect its capabilities to: (1) properly classify both linearly and non-linearly separable classes; (2) promote features which ensure desired properties of classification results, such as small false negative error rate, e.g. (3) converge quickly and scale well with the size of a training data set in order to reduce computational burden of the selection procedure; (4) reproduce accurate results across different data samples so that saliency of selected attributes remains stable.

It must be noted, that the first requirement holds for both feature selectors because even in the case of the filter-type method which is not linked to any classifier chosen a priori, the ultimate feature set may provide better separability if non-linear boundaries between classes are assumed allowable. Moreover, requirements (2) and (3) directly drive the feature search and thus determine its time-efficiency and define significance of the final reduced attribute space. The last requirement implies that performance of a classifier does not rely on any data-dependent configuration parameters.

Therefore, in our experiments we used the Support Vector Machine (SVM) [38] classifier due to its robustness in the following aspects: the quality of generalization ensured by the optimal margin constructed during training stage; ease of training and availability of the off-the-shelf implementations; extendibility to non-linear cases through the use of kernel functions and inherent resistance to statistical outliers. The SVM itself is a linear classification algorithm which aims at constructing a decision hyperplane defined as

$$y(\mathbf{x}) = b + \sum_{\alpha_i \neq 0} \alpha_i y_i \mathbf{x}_i \cdot \mathbf{x}, \tag{1}$$

where parameters α_i and b together with the support vectors $\mathbf{x}_i$ determine location and orientation of the separating hyperplane. The learning procedure involves solving a constrained quadratic optimization problem which leads to determination of α_i coefficients, whose values are non-zero only for those vectors in a training sample which lie closest (on either side) to a decision boundary. It must be noted, that SVM algorithm constructs a hyperplane which defines the largest margin between different data vectors classes in a particular feature subspace. In this aspect, the further described VSCH method behaves similarly to SVM.

Extending SVM to non-linearly separable cases involves replacement of the dot product $(\cdot)$ in (1) by the kernel function which corresponds to the dot product of data vectors non-linearly transformed into higher dimensional feature space. In this research the radial basis function kernel was chosen:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2). \tag{2}$$

The choice of the RBF kernel was due to the fact that it supports spherical geometry of class boundaries. Since images of a particular pathology class are similar, also in terms of computed features, we expect the feature vectors to form a compact cluster. On the other hand images of reference class vary in appearance, thus the feature values will span. Therefore, the distribution of feature vectors is scattered and may surround the cluster.

Unfortunately, as with any other kernel the RBF function needs to be adjusted to the specific properties of a given data set. The value of γ coefficient cannot be determined automatically and several trials must be made before a trained classifier gains its discriminative power. In principle the larger γ is, the better accuracy on a training set is observed. On the other hand, there appears a risk of overfitting when the value of γ becomes too large. These impediments of course violate restriction (4) mentioned above. However, due to other advantages of SVM, the ultimate value of parameter γ was fine-tuned in the course of experiments.

1. $t = 0$;
2. **repeat**
   //forward selection
3. $\hat{\delta} = \arg\max_{\delta \in \mathbf{\Delta}_t} J(\mathbf{\Xi}_t \cup \{\delta\})$;
4. $\mathbf{\Xi}_{t+1} = \mathbf{\Xi}_t \cup \{\hat{\delta}\}$;
   $\mathbf{\Delta}_{t+1} = \mathbf{\Delta}_t \setminus \{\hat{\delta}\}$;
5. $t = t + 1$;
   // backward selection
6. $\hat{\xi} = \arg\max_{\xi \in \mathbf{\Xi}} J(\mathbf{\Xi}_t \setminus \{\xi\})$;
7. **if** $J(\mathbf{\Xi}_t \setminus \{\hat{\xi}\}) > J(\mathbf{\Xi}_{t-1})$ **then**
8. $\mathbf{\Xi}_{t-1} = \mathbf{\Xi}_t \setminus \{\hat{\xi}\}$;
   $\mathbf{\Delta}_{t-1} = \mathbf{\Delta}_t \cup \{\hat{\xi}\}$;
9. $t = t - 1$;
10. **goto** step 6.
11. **endif**
12. **until** $t = d + r$

Fig. 5 – Pseudo-code of the SFFS algorithm.

3.2. *Analysis of variance*

The one-way analysis of variance, widely known as ANOVA [39], tests whether differences between the means of two or more groups of a data sample are significant or accidental. In the feature selection setting, the term groups refers to data vector classes or categories. Then, the goal of ANOVA is to recognize those attributes which provide the strongest evidence against the null hypothesis of there being no statistically important difference between classes. The procedure involves calculation of the *F*-statistics given by

$$F = \frac{\mathrm{MS}_{\mathrm{effect}}}{\mathrm{MS}_{\mathrm{error}}} \tag{3}$$

where $\mathrm{MS}_{\mathrm{effect}}$ and $\mathrm{MS}_{\mathrm{error}}$ estimate the between- and within-class variances respectively. The *F* statistics promotes features where data vectors belonging to different classes form compact and well separated clusters.

3.3. *Sequential floating forward selection*

Basically, strategies of constructing candidate feature subsets in the wrapper approach fall into four categories: optimal (exhaustive or branch-and-bound), sequential, randomized and hybrid, the latter being a mixture of sequential and randomized techniques. Genetic and hybrid genetic algorithms or simulated annealing [28] though being robust to local minima of the objective function (classification accuracy) suffer from relatively slow convergence rate, especially near optimum. Exhaustive search is computationally intractable for very high feature space dimension and time-consuming classification algorithm, while branch-and-bound strategies require monotonicity of the objective function. Thus, sequential search methods constitute viable alternative providing trade-off between flexibility, compute time and optimality of the final solution.

Among various sequential strategies described in the literature, the sequential floating forward search (SFFS) gained dominance in the field. It proves particularly effective in tackling with the nesting problem [26,40]. The SFFS algorithm is outlined in Fig. 5.

On initialization all features are gathered in insignificant feature subset $\mathbf{\Delta}$ whereas relevant subset $\mathbf{\Xi} = \emptyset$. The search mechanism consists of consecutive forward and backward selection steps. In the forward direction, relevant features are added to $\mathbf{\Xi}$, one at a time. The goal of this step is to identify next feature which either ensures best improvement or lowest degradation of *J*. On the other hand, feature removal (backward step) is executed only if it improves optimality criterion. The algorithm creates a collection of subsets $1 \le t \le d$, each containing the best subset of size *t*. Note that the search continues after assumed target dimensionality *d* is reached. The number of extra steps is controlled by the parameter *r* which ensures that the truly best feature subset of size *d* is found. Note also, that the algorithm formulated as in Fig. 5 assumes, that the objective function *J* is to be maximized which takes place in the case of classification accuracy. If *J* corresponds to classification error, the equations in step 3, 6 and 7 must be altered accordingly.

3.4. *Vector supported convex hull method*

As we explained earlier we assume specific distributions of vectors. Feature vectors related with a particular abnormality form a compact cluster surrounded by scattered feature vectors representing normal tissue, intestine content or other artifact structures. Such distributions cannot be separated by means of linear discriminant analysis. On the other hand, application of nonlinear classifiers such as SVMs turns out to be time consuming. Therefore, there is a need for efficient algorithm which defines nonlinear classification rules and quickly assess separability criterion applicable for selection of vector subspace. To solve the problem we developed the Vector Supported Convex Hull (VSCH) algorithm. The overall concept of VSCH was initially proposed in [21], where only some preliminary results were presented. Below we introduce this method in its mature form and after substantial refinement. Also, we have designed its efficient implementation in the C programming language, so that its credible and thorough validation has become plausible.

The VSCH is essentially a discriminant analysis method of supervised learning for reduction of vectors dimensionality and for data classification. Since only a limited number of features carry relevant information needed for discrimination and other descriptors are redundant for classification, the VSCH aims at finding a subset of descriptors, which present best discrimination ability to separate two classes of vectors. Moreover, the VSCH produces a classification rule to separate the two classes.

The input data consist of two sets (classes), of feature vectors in an n-dimensional space (Fig. 6a). The first class represents images of a selected pathology and the second one represents the images without pathology. We search for a k-dimensional subspace ($k<n$) such that vectors of the set number one form a cluster surrounded by vectors of the set number two. For each subspace the convex hull on the vectors belonging to the first class is computed. Then we count vectors of the second class enclosed by the convex hull (Fig. 6b). A subspace with the smallest number Q_1 of such vectors is selected.

If there is more than one subspace with the same smallest Q_1 then we use the second principle of selection. The convex hull is maximally scaled up around its centroid, not to include any additional vectors of the second class (Fig. 6c). The scaling factor a is computed – we select a subspace with the factor of highest value.

Based on the above criteria, the single value index Q of class separability can be defined. The index is a sum of Q_1 and a reciprocal of a, i.e.

$$Q = Q_1 + a^{-1}. \tag{4}$$

The first component is an integer value, whereas a second one is a fraction (a is always grater then on, thus, its reciprocal is positive and less then 1). Based on the Q index, the subspace with the smallest value of the index is selected.

The VSCH algorithm searches through all the 1D, 2D up to ND feature subsets, where N is a parameter to choose. After a subset is selected, the algorithm can also produce rules for classification. The rules are given by a set of inequalities defining the convex hull which was scaled up by a square root of a. In many medical applications it is crucial not to overlook any indications of pathology. Such indications usually are later verified by medical experts and may be rejected. If they are mistakenly rejected by an automatic method, an expert may never notice the pathology. Therefore, it is important to find methods characterized by a minimal false negative error. The VSCH method reveals a property which is particularly useful in biomedical image analysis. The method produces classification rules, such that for the training set vectors, the false negative error is equal to zero.

There are several algorithms for convex hull computation. In our implementation we adapt a *qhull* library implementation [41].

4. Experiments

In order to confirm utility of the above discussed feature computation and selection methods for analysis of the wireless endoscope images, two kinds of experiments were carried out. The goal of the first kind of experiments was to evaluate the discriminative power of texture and color features generated by MaZda software as well as to compare VSCH feature selection and classification algorithm against SVM/SFFS and SVM/ANOVA methods. The second experimental framework demonstrates feasibility of endoscopic image segmentation by means of color and texture feature maps. In these analyses we also proceed with feature computation and feature selection stages. However, the selected features provide here a means of image segmentation rather than pathology detection. The segmentation is a result of individual pixels classification utilizing either VSCH or SVM classifiers. The details of both experiments, the data, the way of feature computation, lists of the selected features, classification and segmentation results, are given in the following subsections.

4.1. *Material*

In the experimental part of this study we analyzed a set of over 50 capsule endoscopy videos. The videos were supplemented with medical reports, including image lists of gastrointestinal tract specific landmarks and pathological abnormalities – so-called findings.

Our efforts were focused on three types of anomalies occurring in the gastrointestinal tract: bleedings, focal ulcers and extensive fibrotic ulcers. These anomalies are characteristic to various diseases, including Crohn's–Leśniewski disease – often lethal, chronic, nonspecific inflammation of the small intestine. The etiology of the disease is not known and the effects of treatment are dependent on its early diagnosis. Therefore, the development of techniques for detecting these lesions is important and reasonable.

In the 50 videos used in the experiment, there were found: 5 sequences with 38 example images of bleeding, 4 with 36 images of focal ulcers and 3 containing 39 images showing excessive fibrotic ulcerations. Moreover, for comparison, more than 500 diverse images without visible lesions were selected. Gastroenterologists participating in the experiment confirmed the presence of specific lesions or lack of such changes in the

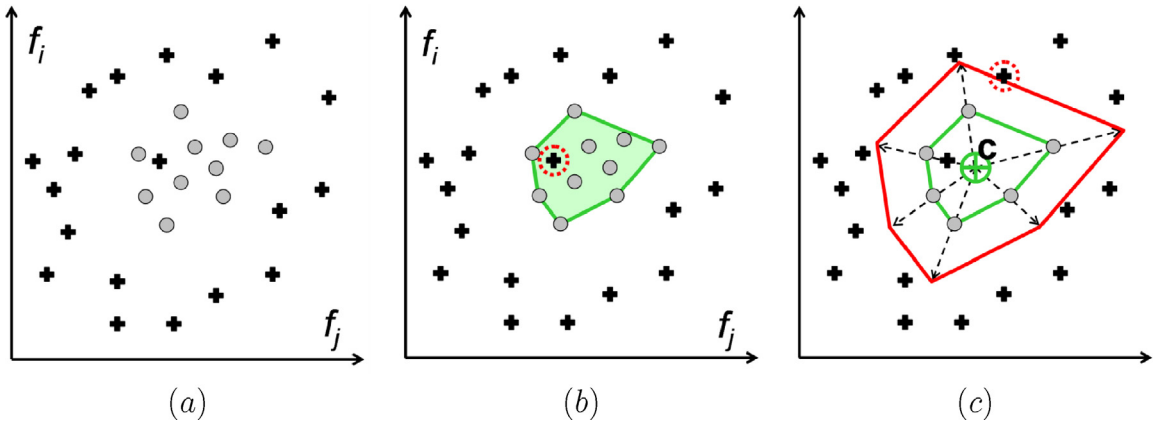

Fig. 6 – Illustration of convex hull method in 2D space ($k = 2$).

selected images, and furthermore, the abnormal regions were outlined.

It should be noted, images of the reference class, without any visible abnormality, presented not only the walls of gastrointestinal system, but also contents of intestines such as partially digested food, bubbles, froth and other artifacts obscuring the view. Thus, the class comprised a mixture of images varied in terms of color, texture and shape of objects presented within.

4.2. *Abnormality detection*

The first experiment was designed to evaluate the discriminative power of the texture and color features in detection of gastrointestinal tract pathologies. In this experiment every feature vector characterizing local image properties was computed within circular region of interest of 25 pixel radius. In the case of bleedings and excessive ulcerations, the regions of interest were semi-randomly placed within the regions indicated by gastroenterologists as abnormal. In contrast to bleedings and excessive ulcers, areas of focal ulcers are small, usually smaller or comparable in size with area of the circular region of interest. Therefore, in these cases the circular region of interest was placed to comprise abnormal area or to comprise at least a substantial portion of it. In addition, the feature vectors were computed within images of the reference class – with no visible abnormalities. In every image of reference class 48 regions of interest were placed evenly within the field of view, and for every such region the feature vector was computed.

The number of calculated vectors exceeded number of 230 per each pathology. There were also 24,000 vectors of the reference class. For the need of discriminant analysis, four sets of labeled vectors were assembled. The set of the reference class contained 1024 randomly selected vectors, the three sets of vectors representing abnormal tissue contained 128 vectors each. Then, three training sets were assembled, each for separate pathology type, by combining reference class vectors with the appropriate abnormality samples. The proportion of class counts in the training sets is adjusted such that cardinality of reference class surpasses the number of lesion class vectors in relation 8:1. In a real full-length video the ratio of normal to abnormal images is several orders larger. However, if this ratio were to be preserved in a training set, classifiers, such as SVM, would suffer from the problem of unbalanced class sizes [42]. The errors made on less frequent class would exert only a negligible impact on overall classification quality, misleading training procedure. Thus we purposely enlarged the percentage of lesion category vectors in training sets to give the SVM algorithm a chance to properly construct classification rule. On the other hand, the portion of reference vectors still remains remarkably larger in order to – at least to some extent – simulate real conditions.

Validation of the discriminant analysis was performed using separate testing sets, composed of 100 randomly selected vectors of each class. The sampling procedure was accomplished in a way ensuring mutual exclusiveness of training and testing sets although for each pathology both kinds of sets could have contained vectors derived from the same WCE sequence. However, at the end of this section we describe an experiment, where new sequences are used for validation and effectiveness of the proposed approach is verified using images not included in classifiers training stage.

The training data sets were examined by the feature selection and classification algorithms, VSCH, SVM/SFFS and SVM/ANOVA. In VSCH, to extract 2D or 3D feature subspace, the feature space exploration was performed using exhaustive search. The extraction of 4D and 5D subspace was performed on a space of preselected, top-ranked attributes, which demonstrated the best discriminative power in preceding explorations. The SFFS selection was launched with target dimensionality set to 5 (value of the ripple parameter equaled 2) and thus it sequentially produced best 2D, 3D, 4D and 5D feature subspaces. In the case of ANOVA, all features were ordered relative to their respective F-statistics measure and then the desired number of attributes from the top of the list were taken for further analysis.

Relevance of a feature subsets obtained by one of the three selection methods was assessed by the accuracy of a corresponding classifier. Note that in the case of SVM algorithm several trails with different values of the gamma coefficient (cf. Eq. (2)) were executed in order to achieve best attainable results for a particular data set. The classification specificity and sensitivity indicators were computed for the

Table 1 – Sensitivity and specificity ratios obtained for the bleeding class.

Exploration strategy	Evaluation ratio	Feature space dimension			
		2D	3D	4D	5D
Training set results					
SVM/	Sensitivity	0.959	0.962	0.977	0.977
SFFS	Specificity	0.986	0.997	0.997	0.998
SVM/	Sensitivity	0.970	0.970	0.970	0.970
ANOVA	Specificity	0.997	0.999	0.999	1.000
VSCH	Sensitivity	1.000	1.000	1.000	1.000
	Specificity	1.000	1.000	1.000	1.000
Testing set results					
SVM/	Sensitivity	0.867	0.980	0.990	0.990
SFFS	Specificity	1.000	1.000	1.000	1.000
SVM/	Sensitivity	0.878	0.970	0.970	0.970
ANOVA	Specificity	1.000	1.000	1.000	1.000
VSCH	Sensitivity	1.000	0.980	0.880	0.880
	Specificity	0.990	1.000	1.000	1.000

Table 2 – Sensitivity and specificity ratios for the focal ulceration class.

Exploration strategy	Evaluation ratio	Feature space dimension			
		2D	3D	4D	5D
Training set results					
SVM/	Sensitivity	0.773	0.826	0.826	0.833
SFFS	Specificity	0.984	0.991	0.996	0.997
SVM/	Sensitivity	0.000	0.561	0.652	0.674
ANOVA	Specificity	1.000	0.965	0.975	0.981
VSCH	Sensitivity	1.000	1.000	1.000	1.000
	Specificity	0.889	0.959	0.989	0.999
Testing set results					
SVM/	Sensitivity	0.794	0.804	0.794	0.794
SFFS	Specificity	0.963	0.963	0.963	0.953
SVM/	Sensitivity	0.000	0.454	0.557	0.557
ANOVA	Specificity	1.000	0.963	0.972	0.972
VSCH	Sensitivity	0.970	0.830	0.720	0.670
	Specificity	0.660	0.940	0.970	0.980

corresponding subspaces of training and testing sets, they are presented in Tables 1–3.

Firstly, it is evident, that the selected texture- and color-based features enable accurate discrimination between abnormal and reference vectors. However, the accuracy level varies depending on the feature selection and classification method and also on the pathology type. The best recognition rates for all exploration approaches are attained for the bleeding class. In the case of training data sets, all methods reach the level of minimum 95% for sensitivity and specificity ratios and all tested feature space cardinalities. In particular, the VSCH algorithm exhibits outstanding performance

Table 3 – Sensitivity and specificity ratios for the excessive ulceration class.

Exploration strategy	Evaluation ratio	Feature space dimension			
		2D	3D	4D	5D
Training set results					
SVM/	Sensitivity	0.553	0.682	0.826	0.856
SFFS	Specificity	0.991	0.994	0.992	0.996
SVM/	Sensitivity	0.417	0.932	1.000	1.000
ANOVA	Specificity	0.995	1.000	1.000	1.000
VSCH	Sensitivity	1.000	1.000	1.000	1.000
	Specificity	0.933	0.989	0.999	1.000
Testing set results					
SVM/	Sensitivity	0.429	0.592	0.663	0.643
SFFS	Specificity	0.990	1.000	1.000	1.000
SVM/	Sensitivity	0.286	0.020	0.000	0.000
ANOVA	Specificity	0.990	1.000	1.000	1.000
VSCH	Sensitivity	1.000	0.880	0.700	0.630
	Specificity	0.720	0.940	0.990	0.990

with 100% scores on the both evaluation criteria. Comparing results achieved in the training stage by ANOVA and SFFS selection methods suggests primacy of the former strategy. However, the testing part reveals opposite tendency. Clearly, the SFFS procedure identifies features which allow construction of a classifier possessing high generalization properties. Moreover, the sensitivity ratios gained for the test sets by the SVM/SFFS and SVM/ANOVA methods are only slightly worse from the training results. On the contrary, performance of VSCH remarkably degrades with the feature space dimension, falling below the acceptable level of 90% for the 5D case. However, if all tested dimensionalities are taken into account, still the VSCH methods outperforms the others, as it is the only method to correctly detect all true positive vectors for the 2D case.

This trend in sensitivity scores – observable also for the other two pathology classes – can be explained by the way, in which the VSCH feature selection procedure is carried on. After examining all possible 3D feature combinations, the attribute triplets are ranked according to their corresponding Q factor value. Full search in 4D is restricted to the top of that rank. Similarly, exploration in 5D is based on the results of 4D selection. Thus, some of the potentially good 4D and 5D feature combinations are never tested and consequently the SFFS selection method appears as a better option in this respect.

Analysis of the results obtained for focal and excessive ulceration classes confirms supremacy of the VSCH method as it again lead in both cases to almost perfect sensitivity ratios for the testing sets (97% and 100% respectively) in 2D. The scores achieved by SVM/SFFS and SVM/ANOVA are far from satisfactory. In particular, classification based on features identified by ANOVA completes with no true positive recognitions at all. This unwelcome effect is directly caused by the structure of the training sets with strong bias towards the reference class. As a result, the classifier error committed on the pathology class has lower impact on the total error and in turn on the overall evaluation of the feature subset. On the other hand, the proposed VSCH algorithm always constructs a convex hull surrounding all true positive cases causing sensitivity ratio to be equal to 100% for any input training set. This property of VSCH ensures its robustness to unbalanced proportions between cardinalities of different-class vector collections so that no clinically important WCE images are omitted during diagnosis.

Regarding specificity ratios, all methods achieve good results (at least 94% for the testing sets) except for the VSCH in the case of ulcers and the 2D space. It is apparent here, that good sensitivity was achieved at the cost of large number of false positive recognitions. Moreover, it can be noted, that in general the preferable dimensionality of target feature space is 3D, as it provides acceptable trade-off between specificity and sensitivity scores for VSCH and SVM/SFFS methods. The latter algorithm performs better in 4D in the case of ulcers, but the benefit from increasing dimension is only minor, not equaling computational effort required by additional texture feature calculation.

It is also interesting to analyze scatter plots of data instances in feature subspaces selected by different methods. Figs. 7–9 depict the example plots obtained for the bleeding class in 3D. It can be seen that for the ANOVA method feature vectors are located mostly along one distinct direction. This distribution, observed also for the other two pathology classes, shows that ANOVA tends to select mutually correlated features which are more expressive (in terms of incorporated data set variance) rather than discriminative in meaning. On the other hand, SVM/SFFS and VSCH methods identify features which convey more diversed information and hence allow for better, possibly non-linear, separation of classes.

Another point of investigation is the ability of particular color spaces to emphasize lesion structures in WCE images and thus facilitate their automatic recognition. To answer this question we performed additional series of experiments, where feature selection was restricted to RGB, HSB and YUV spaces separately. Tables 4 and 5 present the obtained sensitivity and specificity ratios for different number of dimensions, as well as lists of channels of selected features. The presented results were calculated for two example pathology classes – bleeding and excessive ulcerations – and using the VSCH method. For convenience, these tables also recall the corresponding outcomes of experiments where all color spaces were simultaneously taken into account (Tables 1 and 3 respectively).

It can be observed that different color spaces allow achieving various levels of classification accuracy. In particular, the scores obtained for RGB channels are slightly better than for HSB and YUV spaces. It performs best for 2-dimensional cases for both tested pathology classes ensuring relatively high specificity and sensitivity indicators. However, the isolated color spaces – especially for higher dimensions – lead to more classification errors if compared to color spaces mixture. Note that in the latter case, the selected features comprise of descriptors derived from various color models, potentially complementary to one another. This supports the postulated approach which avoids prioritization of any color model, leaving the choice of the best color channels to feature selection procedure.

Eventually, it must be noted that all the above described results were computed only for randomly chosen sample of WCE video frames. However, it is necessary to assess the proposed approach performance in relation to entire sequence, typically composed of over 50,000 endoscopic images. For that purpose, we designed a computer program dedicated to WCE video analysis whose goal is to detect frames suspected to include pathology structures. We focused again on two lesion classes – bleedings and excessive ulcerations whose automated identification – according to achieved experimental results – is mostly reliable and feasible in practice. During program operation, a frame is first divided into 48 circular regions of interest (laid out as shown in Fig. 2) and each ROI is submitted to texture analysis routine. Then an image is denoted as a suspected one, if any of the implemented classifiers (trained with the use of 3D VSCH method) marks at least one ROI as positive, i.e. including pathology.

Using the designed program, we examined two new WCE data sets, whose frames were not employed for classifiers training stage. One of the investigated sequences contained a series of images with visible ulcerations, however differing from previously analyzed excessive structures. Although this new type of lesion also occupies large areas of intestine walls, it additionally contains small reddish spots resampling

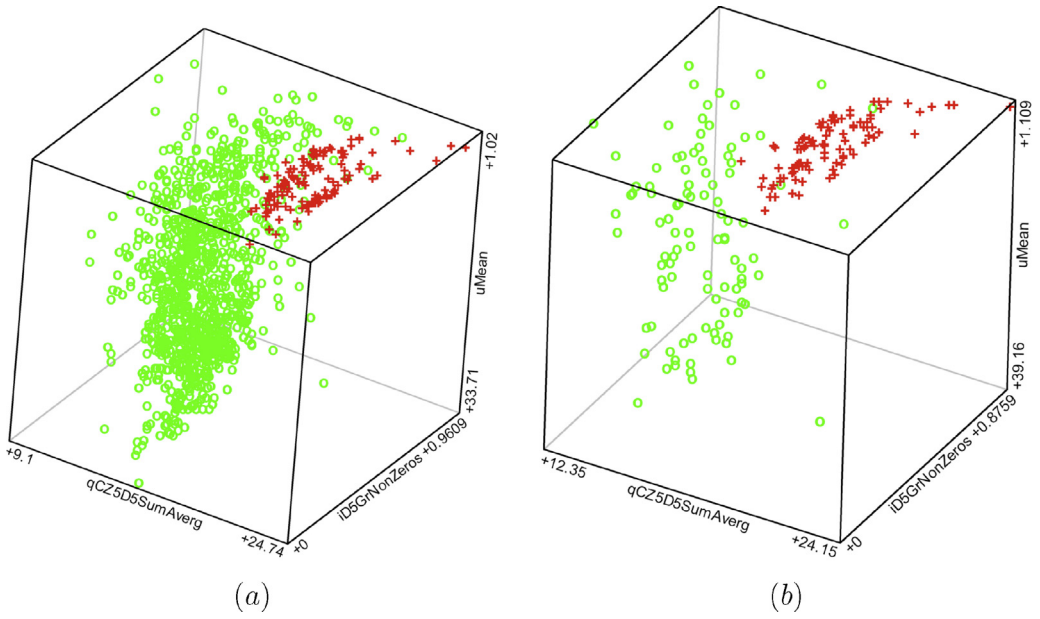

Fig. 7 – Scatter plots of the bleeding (denoted as red '+') and reference (green 'o') classes in the reduced feature subsets as selected by VSCH algorithm: (a) training and (b) test data vectors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

bleedings or petechiae (Fig. 10). In the second analyzed sequence, medical experts did not identify any pathology structure and thus it can be assumed normal.

Since in the analyzed movies there were no pathology classes exactly matching those considered in the training phase, overall classification performance was evaluated disregarding the actual lesion category. Classification result was assumed correct if an existing pathology structure was detected by any of the two implemented classifiers. Thus, the analysis of the first sequence containing new type of ulceration resulted in sensitivity and specificity ratios equal to 0.72 and 0.88 respectively. These rates are lower than the scores given in Tables 1 and 3, however the pattern of new pathology is significantly different both from bleeding and excessive ulceration. If this factor is taken into account, the designed system can still be seen as preserving practical value even if dedicated pathology structures do not appear in a given WCE examination. This conclusion is further supported by the results of the second movie analysis. In this case, in the sequence of 57,061 frames, only 17 (0.03%) cases of excessive ulcerations and 5127 (9.0%) of bleedings were found. These numbers reduce to 0 and 1429 (3.5%) false positive errors if

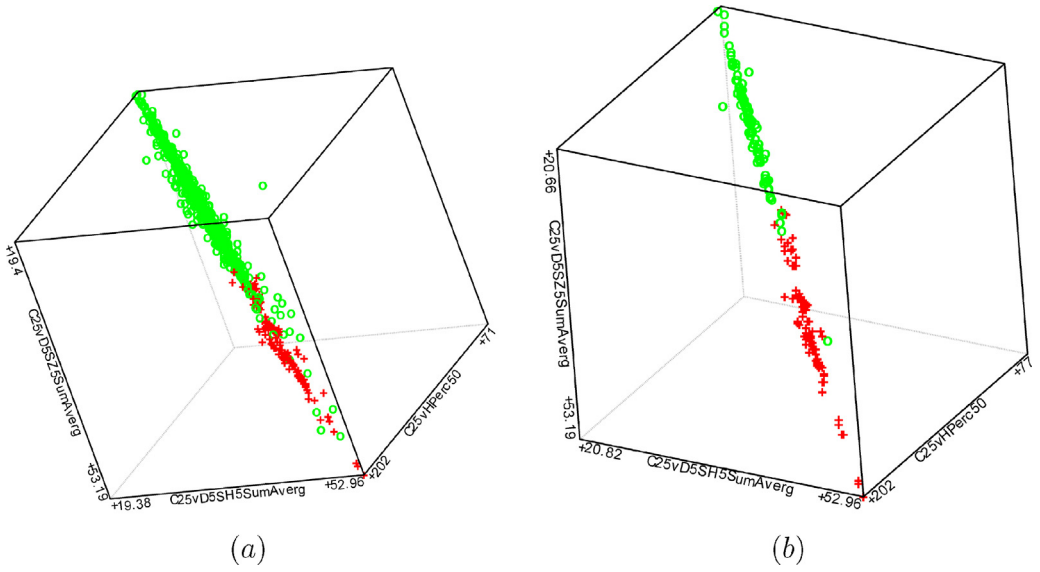

Fig. 8 – Scatter plots of the bleeding (denoted as red '+') and reference (green 'o') classes in the reduced feature subsets as selected by ANOVA algorithm: (a) training and (b) test data vectors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

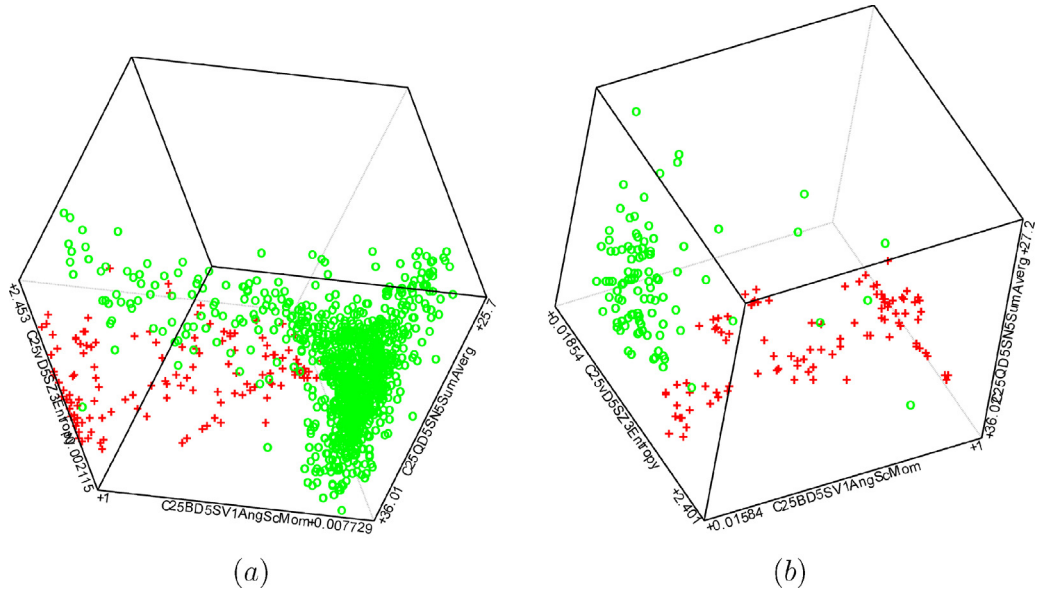

Fig. 9 – Scatter plots of the bleeding (denoted as red '+') and reference (green 'o') classes in the reduced feature subsets as selected by a–b) SVM/SFFS algorithm: (a) training and (b) test data vectors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

the analysis of entire video is limited to 40,785 frames embracing duodenum and small intestine, i.e. the fragments usually examined in capsule endoscopy.

4.3. *Image segmentation*

At the first glance, calculation of maps of the features selected in the previous section and then, using these maps as input to the segmentation procedure seems to be the right solution to the problem. Unfortunately, it is not. The problem is, the feature vectors of the reference class were computed only within images that do not contain lesions. As a result, the selected features correctly discriminate images belonging to different groups, but they are not really useful to distinguish between regions of normal and abnormal state within the same image.

To select features useful for image segmentation, the feature selection process should also take into account vectors computed for normal state regions appearing within images showing abnormalities. Therefore, in order to select features capable of image segmentation, the set of reference vectors was augmented with additional entities computed in such the regions. Moreover, the original feature space was extended on attributes calculated in only direct proximity of a given pixel. Thus, apart from features determined for circular ROI of 25-pixel-long radius, all data vectors contained components corresponding to 3×3-pixel squares centered at the point of interest. These additional attributes ensure even more local

Table 4 – Sensitivity and specificity ratios obtained in different color spaces for the bleeding class (testing sets results).

Dimensionality	Color space	Sensitivity	Specificity	Selected color channels
2D	RGB	0.965	0.979	G, R
	YUV	0.965	0.954	Y, V
	HSB	0.929	0.954	H, B
	All	1.000	0.990	Q, U/Y
3D	RGB	0.912	0.996	G, R, B
	YUV	0.912	1.000	Y, V, U
	HSB	0.859	1.000	H, B, B
	All	0.980	1.000	Q/Y, U/Y, I/Y
4D	RGB	0.912	1.000	R, G, B, G
	YUV	0.877	1.000	Y, V, U, U
	HSB	0.859	1.000	B, H, B, B
	All	0.880	1.000	H, V/Y, U/Y, V/Y
5D	RGB	0.859	1.000	G, B, G, G, R
	YUV	0.807	0.966	Y, V, U, U, V
	HSB	0.877	1.000	B, B, H, S, S
	All	0.880	1.000	H, U/Y, V/Y, U/Y, U/Y

Table 5 – Sensitivity and specificity ratios obtained in different color spaces for the excessive ulceration class (testing sets results).

Dimensionality	Color space	Sensitivity	Specificity	Selected color channels
2D	RGB	0.965	0.730	G, G
	YUV	0.965	0.663	Y, V
	HSB	0.929	0.630	B, S
	All	1.000	0.720	Z, H
3D	RGB	0.830	0.966	G, B, B
	YUV	0.877	0.912	Y, V, V
	HSB	0.859	0.966	B, H, H
	All	0.880	0.940	G, U/Y, U/Y
4D	RGB	0.651	0.990	B, B, G, G
	YUV	0.630	0.966	V, V, Y, Y
	HSB	0.592	0.966	B, H, H, H
	All	0.700	0.990	G, U/Y, U/Y, G
5D	RGB	0.592	0.990	B, B, G, G, G
	YUV	0.553	0.996	Y, Y, Y, U, U
	HSB	0.553	0.993	B, B, H, H, S
	All	0.630	0.990	G, L, Y, V/Y, Y

characteristics of pixels giving opportunity for fine-grained segment delineation. The selection of features for image segmentation was performed the same way as described in the previous section. Three different methods (VSCH, SVM/SFFS and ANOVA) were applied for selection, which yielded three self-sufficient groups of features and thus three different classifiers.

Image segmentation involves qualification of image pixels into different classes. Pixels classified as belonging to a lesion class should designate a region that coincides with the region of the actual lesion. Other pixels should fall into areas without visible pathological changes. Thus, to allow for classification of individual pixels, using the selected features, the values of these features must be computed around each individual pixel to form feature maps.

Fig. 10 – Example of large-area ulcer with red bleeding-like structures. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

For the purpose of the experiment 16 images of bleeding, 16 images of focal ulcers, 16 images of excessive fibrotic ulcers and 64 images of reference class were randomly chosen. The images were segmented by means of the three classifiers. To validate the results, the obtained regions of suspected lesions were compared with the regions depicted manually by a gastroenterologist. For the need of quantitative validation, classification precision and recall ratios as well as Jaccard index were computed for each pair of regions being compared, as defined below:

$$\text{Precision} = \frac{A \cap M}{A}, \tag{5}$$

$$\text{Recall} = \frac{A \cap M}{M}, \tag{6}$$

$$\text{Jaccard index} = \frac{A \cap M}{A \cup M}, \tag{7}$$

where A and M denote areas of automatically and manually delineated segments.

The average values of the three measures are listed in Table 6. Example images of lesions and results of their segmentation are shown in Fig. 11. As it can be seen, precise delineation of pathology regions is a challenging task, which is particularly apparent in the case of small structures such as focal ulcers. The highest scores on Jaccard coefficient – which estimates the accuracy of fit between manually and automatically found regions – do not exceed the value of 0.62 (VSCH method, excessive ulcerations), while the lowest results fall into the range 0.2–0.3 (focal ulcers). On the other hand, the recall ratios are relatively high, especially for the proposed VSCH approach. These results ensure, that approximately 90% of the region of interest area will be properly recognized. The SVM/SFFS and SVM/ANOVA methods perform worse in this respect, although the prior strategy gives similar recall values for the bleeding class.

Comparing the results achieved on the precision ratio, it emerges that none of the tested methods gains supremacy over the others. SVM/ANOVA and SVM/SFFS occurred to be

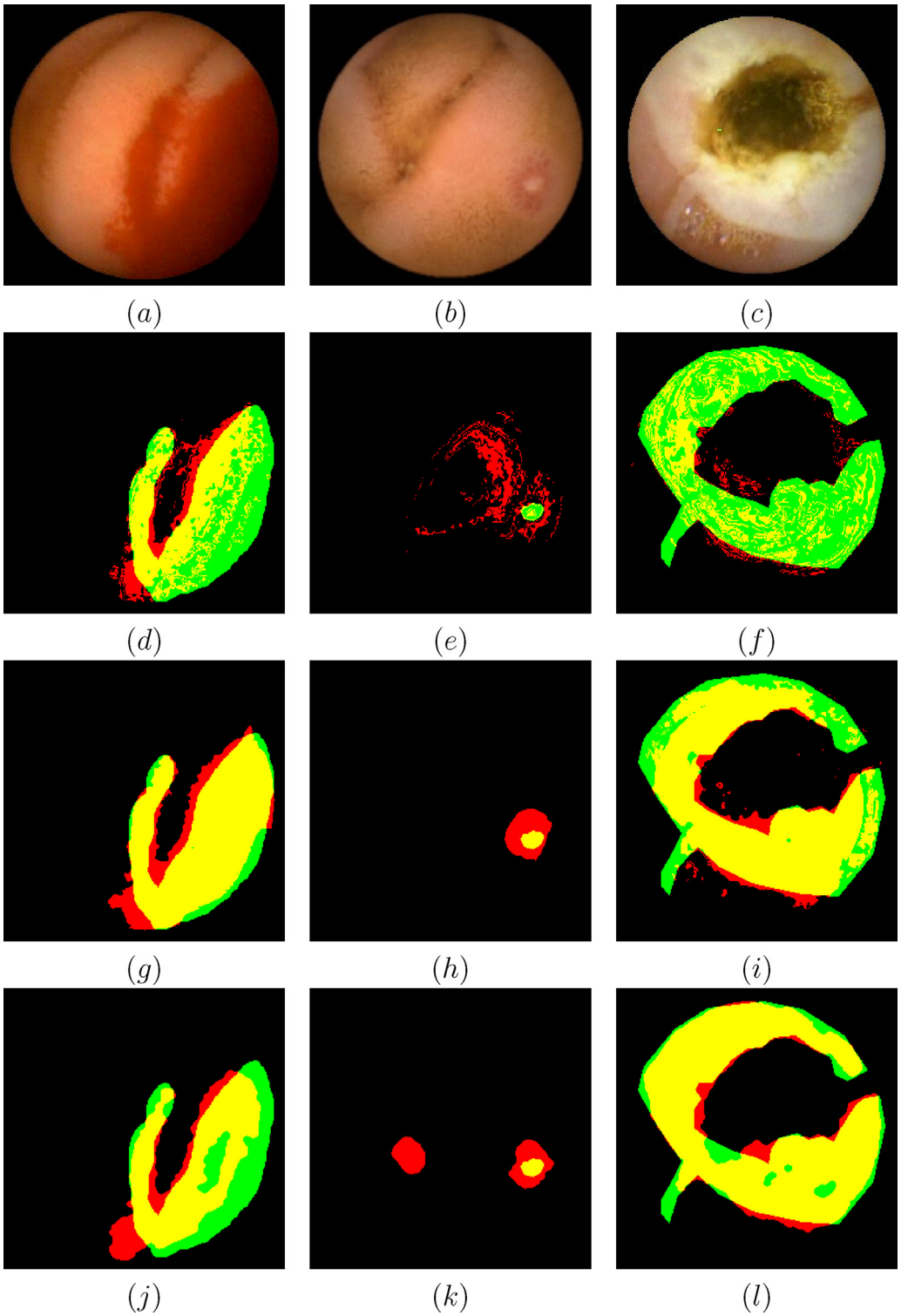

Fig. 11 – Example images of (a) bleeding, (b) focal, (c) excessive ulcerations and their corresponding segmentation results obtained by (d–f) SVM/ANOVA, (g–i) SVM/SFFS, and (j–l) VSCH algorithms. Yellow areas denote common regions found by manual delineation (green outside) and automatic segmentation methods (red outside). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

the best in the case of excessive ulcerations reaching the levels of 0.87 and 0.92 respectively. However, both strategies classify remarkably large amounts of pixels into pathology segment in the case of bleedings (mean precision values close to 0.45) and focal ulcerations (around 0.3). On the other hand, VSCH method exhibits stable behavior and irrespective of the abnormality class gives precision value approximately equal to 0.7. This score is better than the results obtained by the

Table 6 – Quantitative evaluation of the segmentation results.

Algorithm	Evaluation ratio	Bleeding	Focal ulcer	Excessive ulcer
SVM/ ANOVA	Recall	0.693	0.443	0.303
	Precision	0.448	0.287	0.869
	Jaccard index	0.377	0.194	0.285
SVM/ SFFS	Recall	0.869	0.769	0.580
	Precision	0.495	0.305	0.918
	Jaccard	0.462	0.287	0.546
VSCH	Recall	0.875	0.932	0.917
	Precision	0.732	0.696	0.669
	Jaccard index	0.487	0.235	0.621

other two methods for ulcers, however still unsatisfactory. This phenomenon – similar to effects observed in the detection experiment – is caused by the intrinsic property of the VSCH algorithm, which aims at minimizing false negative recognitions even if it increases false positive errors.

5. Summary

The obtained experimental results show great potential of the employed methods in detection and segmentation of anomolies in the WCE images. Especially the SVM/SFFS strategy and the introduced VSCH algorithm reach high accuracy levels, although the latter outperforms the other in all three tested gastrointestinal pathology classes. The best detection efficiency has been achieved for the bleeding category, where VSCH scores over 98% in sensitivity and specificity ratios. Similarily, VSCH occured most accurate in identifying segments representing various abnormality structures.

It is also apparent that usage of non-linear classification rules and extensive feature selection based on the wrapper model is necessary to enable reliable diagnosis. The worst results for any abnormality class are obtained for the SVM/ANOVA method. It is caused by the intrinsic linearity of the ANOVA test and by the fact, that a feature relevance is assessed in isolation from the other attributes. Therefore, ANOVA tends to select features which are correlated with each other and whose common discriminative power is low in comparison with the subsets selected by the wrapper methods.

It can be concluded that exploratory analysis based on texture and color image information and using combined feature selection and classification algorithms helps improving WCE interpretation, making it reliable and cost-effective. Eventually, due to automation of the employed image processing procedures, the diagnosis will also lead to reproducible and more objective results.

On the other hand, if we take into account the results obtained in the analysis of the entire video sequences, over one thousand of false positives is not acceptable in the clinical application as it would undermine the clinician's trust in the system. Therefore, it is necessary to improve the detection algorithm, with the aim of increasing specificity measure. For this reason, we plan to carry out further work on integrating information on color and texture with shape parameters and movement characteristics. Such an approach stems from the observation that many structures are properly recognized by gastroenterologists only if they are observed in motion, i.e. in the range of several frames. This would enable distinction between objects of specific shape, whose position alters independently from the intestine wall and those which shift coherently with the background. In the former case, it can be suspected, that the observed object is an artifact, such as some food residue, while in the latter case it is more likely to be a tissue malformation. Thus, it can be presumed that motion and shape analysis of WCE images will reduce the number of false positive detections.

The described procedure of image color and texture feature computation as well as VSCH feature selection algorithm is implemented in MaZda software [35]. The algorithm for detection of bleedings and ulcers, being the result of presented studies is implemented in WCE Player program. Both programs are available for download from http://www.eletel.p.lodz.pl/programy/mazda/ and http://wce.naviton.pl/.

Conflicts of interest statement

I declare that there are no conflicts of interest related to the above manuscript.

REFERENCES

[1] E. Scapa, H. Jacob, S. Lewkowicz, M. Migdal, D. Gat, A. Gluckhovski, N. Gutmann, Z. Fireman, Initial experience of wireless-capsule endoscopy for evaluating occult gastrointestinal bleeding and suspected small bowel pathologyclinical use of wireless-capsule video endoscope, The American Journal of Gastroenterology 97 (2002) 2776–2779.
[2] B. Lewis, P. Swain, Capsule endoscopy in the evaluation of patients with suspected small intestinal bleeding: results of a pilot study, Gastrointestinal Endoscopy 56 (2002) 349–353.
[3] D. Adler, C. Goustout, Wireless capsule endoscopy, Hosp Physician 39 (2003) 14–22.
[4] P. Swain, Wireless capsule endoscopy and Crohn's disease, GUT 54 (2005) 323–326.
[5] F. Wartel, V. Maunoury, P. Bulois, S. Papadopoulos, B. Filoche, J.F. Colombel, Smallbowel ulcerations at wireless capsule endoscopy: go the whole way, GUT 56 (2007) 1132.
[6] M.M. Chait, Gastroesophageal reflux disease: important considerations for the older patients, World Journal of Gastrointestinal Endoscopy 2 (2010) 388–396.
[7] H.P. Lukashok, Carlos Robles-Jara, Carlos Robles-Medranda, Multiple Intestinal Erosions as a Result of Hemorrhage due to Parasites: Case Reports and Review of the Literature,

Diagnostic and Therapeutic Endoscopy 2011 (2011), http://dx.doi.org/10.1155/2011/340869, Article ID 340869, 3 pages.
[8] M.K. Goenka, S. Majumder, S. Kumar, P.K. Sethy, U. Goenka, Single center experience of capsule endoscopy in patients with obscure gastrointestinal bleeding, World Journal of Gastroenterology 17 (2011) 774–778.
[9] J. Berens, M.W. Mackiewicz, G.D. Bell, Stomach, intestine, and colon tissue discriminators for wireless capsule endoscopy images, in: Medical Imaging 2005: Image Processing, San Diego, CA, USA, 2005, pp. 283–290.
[10] M. Coimbra, P. Campos, J. Cunha, Topographic segmentation and transit time estimation for endoscopic capsule exams, in: IEEE ICASSP, volume 2, Toulouse-France, pp. 1164–1167.
[10] M. Coimbra, P. Campos, J. Cunha, Topographic segmentation and transit time estimation for endoscopic capsule exams, in: IEEE ICASSP, volume 2, Toulouse-France, 2006, pp. 1164–1167.
[12] M. Mackiewicz, J. Berens, M. Fisher, Wireless capsule endoscopy color video segmentation, IEEE Transactions on Medical Imaging 27 (2008) 1769–1781.
[13] F. Vilarino, L.I. Kuncheva, P. Radeva, Roc curves and video analysis optimization in intestinal capsule endoscopy, Pattern Recognition Letters 27 (2006) 875–881.
[14] P.M. Szczypiński, R.D. Sriram, P.V. Sriram, D.N. Reddy, A model of deformable rings for interpretation of wireless capsule endoscopic videos, Medical Image Analysis 13 (2009) 312–324.
[15] N. Bourbakis, Detecting abnormal patterns in wce images, in: Proc. 5th IEEE Symposium on Bioninformatics and Bioengineering, 2005.
[16] V. Kodogiannis, M. Boulougoura, J. Lygouras, I. Petrounias, A neuro-fuzzy-based system for detecting abnormal patterns in wireless-capsule endoscopic images, Neurocomputing 70 (2007) 704–717.
[17] D.J.C. Barbosa, J. Ramos, C.S. Lima, Detection of small bowel tumors in capsule endoscopy frames using texture analysis based on the discrete wavelet transform, in: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008, pp. 3012–3015.
[18] B. Li, M.-H. Meng, Ulcer recognition in capsule endoscopy images by texture features, in: Proceedings of the 7th World Congress on Intelligent Control and Automation, China, 2008, pp. 234–239.
[19] M.W. Mackiewicz, M. Fisher, C. Jamieson, Bleeding detection in wireless capsule endoscopy using adaptive colour histogram model and support vector classification, in: Medical Imaging 2008: Image Processing, San Diego, CA, USA, 2008, p. 69140R-12.
[20] B. Li, M. Meng, Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments, Computers in Biology and Medicine 39 (2009) 141–147.
[21] P. Szczypiński, A. Klepaczko, Convex hull-based feature selection in application to classification of wireless capsule endoscopic images, in: J. Blanc-Talon, W. Philips, D. Popescu, P. Scheunders (Eds.), Advanced Concepts for Intelligent Vision Systems, volume 5807 of Lecture Notes in Computer Science, Springer, Berlin/Heidelberg, 2009, pp. 664–675.
[22] S. Hwang, M. Celebi, Polyp detection in wireless capsule endoscopy videos based on image segmentation and geometric feature, in: Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 678–681.
[23] V. Charisis, L. Hadjileontiadis, C. Liastos, C. Mavrogiannis, G. Sergiadis, Capsule endoscopy image analysis using texture information from various colour models, Computer Methods and Programs in Biomedicine (2011), 10.1016/j.cmpb.2011.10.004.
[24] D. Barbosa, D. Roupar, J. Ramos, A. Tavares, C. Lima, Automatic small bowel tumor diagnosis by using multi-scale wavelet-based analysis in wireless capsule endoscopy images, Biomedical engineering online 11 (2012) 3.
[25] W. Siedlecki, J. Sklansky, On automatic feature selection, International Journal of Pattern Recognition and Artificial Intelligence 2 (1988) 197–220.
[26] P. Pudil, J. Novovicova, J. Kittler, Floating search methods in feature selection, Pattern Recognition Letters 15 (1994) 1119–1125.
[27] R. Kohavi, G. John, Wrappers for feature subset selection, Artificial Intelligence 97 (1997) 273–324.
[28] I. Oh, J. Lee, B. Moon, Hybrid genetic algorithms for feature selection, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (2004) 1424–1437.
[29] J. Dy, C. Brodley, Feature selection for unsupervised learning, Journal of Machine Learning Research 5 (2004) 845–889.
[30] A. Klepaczko, A. Materka, Clustering-stability based feature selection for unsupervised texture classification, Machine Graphics and Vision 18 (2009) 125–141.
[31] M. Coimbra, P. Campos, J.P.S. Cunha, Extracting clinical information from endoscopic capsule exams using mpeg-7 visual descriptors, IEE Seminar Digests 2005 (2005) 105–110.
[32] M. Coimbra, J. Cunha, Mpeg-7 visual descriptors–contributions for automated feature extraction in capsule endoscopy, IEEE Transactions on Circuits and Systems for Video Technology 16 (2006) 628–637.
[33] B. Giritharan, X. Yuan, J. Liu, B. Buckles, J.-H. Oh, S.J. Tang, Bleeding detection from capsule endoscopy videos, in: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2008, pp. 4780–4783.
[34] J.M. Martí nez, Mpeg-7 Overview, 2004, http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm
[35] P. Szczypinski, M. Strzelecki, A. Materka, A. Klepaczko, Mazda—a software package for image texture analysis, Computer Methods and Programs in Biomedicine 94 (2009) 66–76.
[36] A. Klepaczko, M. Kociński, A. Materka, Quantitative description of 3D vascularity images: texture-based approach and its verification thhrough cluster analysis, Pattern Analysis and Applications 14 (2011) 415–424.
[37] A.L. Blum, P. Langley, Selection of relevant features and examples in machine learning, Artificial Intelligence 97 (1997) 245–271.
[38] V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.
[39] D.A. Freedman, Statistical Models: Theory and Practice, Cambridge University Press, 2005.
[40] P. Pudil, P. Somol, Current feature selection techniques in statistical pattern recognition, in: M. Kurzynski, E. Puchala, M. Wozniak, A. Zolnierek (Eds.), Computer Recognition Systems. Advances in Soft Computing, Springer, 2005.
[41] C. Barber, D. Dobkin, H. Huhdanpaa, The Quickhull algorithm for convex hulls, ACM Transactions on Mathematical Software 22 (1996) 469–483.
[42] I.H. Witten, E. Frank, Data Mining. Practical Machine Learning Tools and Techniques, Morgan Kaufmann, San Francisco, 2005.