CLASSIFICATION OF CAPSULE ENDOSCOPIC IMAGES BASED ON TEXTURE AND NEURO-FUZZY SYSTEMS

V.S. Kodogiannis*, E. Wadge*, and D. Tomtsis

* Mechatronics Group, School of Computer Science,
University of Westminster, London, United Kingdom

V.Kodogiannis@westminster.ac.uk

Abstract: Computerised processing of medical images can ease the search of the representative features in the images. The endoscopic images possess rich information expressed by texture. In this paper schemes have been developed to extract features from the texture spectra in the chromatic and achromatic domains for a selected region of interest from each colour component histogram of images acquired by the new M2A Swallowable Capsule. The implementation of an advanced learning-based neuro-fuzzy scheme and the concept of fusion of multiple classifiers have been also adopted in this paper. The preliminary test results support the feasibility of the proposed methodology.

Introduction

In medical practice, endoscopic diagnosis and other minimally invasive imaging procedures, such as computed tomography, ultrasonography, con-focal microscopy, computed radiography, or magnetic resonance imaging, are now permitting visualisation of previously inaccessible regions of the body. Their objective is to increase the expert's ability in identifying malignant regions and decrease the need for intervention while maintaining the ability for accurate diagnosis. For more than 10 years, flexible video-endoscopes have a widespread use in medicine and guide a variety of diagnostic and therapeutic procedures including colonoscopy, gastroenterology and laparoscopy [1]. Conventional diagnosis of endoscopic images employs visual interpretation of an expert physician. Since the beginning of computer technology, it becomes necessary for visual systems to "understand a scene", that is making its own properties to be outstanding, by enclosing them in a general description of an analysed environment. Computer-assisted image analysis can extract the representative features of the images together with quantitative measurements and thus can ease the task of objective interpretations by a physician expert in endoscopy. A system capable to classify image regions to normal or abnormal will act as a second - more detailed - "eye" by processing the endoscopic video. Endoscopic images possess rich information, which facilitates the abnormality detection by multiple techniques. However, from the literature survey, it has been found that only a few techniques for endoscopic image analysis have been reported and they are still undergoing testing. In addition, most of the techniques were developed on the basis of features in a single domain: chromatic domain or spatial domain. Applying these techniques individually for detecting the disease patterns based on possible incomplete and partial information may lead to inaccurate diagnosis.

Krishnan, *et al.* [2] has been using endoscopic images to define features of the normal and the abnormal colon. New approaches for the characterisation of colon based on a set of quantitative parameters, extracted by the fuzzy processing of colon images, have been used for assisting the colonoscopist in the assessment of the status of patients and were used as inputs to a rule-based decision strategy to find out whether the colon's lumen belongs to either an abnormal or normal category. The analysis of the extracted quantitative parameters was performed using three different neural networks selected for classification of the colon. Endoscopic images contain rich information of texture. Therefore, the additional texture information can provide better results for the image analysis than approaches using merely intensity information. Such information has been used in CoLD (colorectal lesions detector) an innovative detection system to support colorectal cancer diagnosis and detection of pre-cancerous polyps, by processing endoscopy images or video frame sequences acquired during colonoscopy [3]. It utilised second-order statistical features that were calculated on the wavelet transformation of each image to discriminate amongst regions of normal or abnormal tissue. A neural network based on the classic BP learning algorithm performed the classification of the features. CoLD integrated the feature extraction and classification algorithms under a graphical user interface, which allowed both novice and expert users to utilise effectively all system's functions. The detection accuracy of the proposed system has been estimated to be more than 95%.

Intra-operative endoscopy, although used with great success, is more invasive and associated with a higher rate of complications. Though the gastrointestinal (GI) endoscopic procedure has been widely used, doctors must be skilful and experienced to reach deep sites such as the duodenum and small intestine. The cleaning and sterilisation of these devices is still a problem leading to the desire for disposable instruments. In GI tract, great skill and concentration are required for navigating the endoscope because of its flexible structure. Discomfort to the patient and the time required for diagnosis heavily

ISSN: 1727-1983 © 2005 IFMBE

depend on the technical skill of the physician and there is always a possibility of the tip of the endoscope injuring the walls. Standard endoscopic examinations evaluate only short segments of the proximal and distal small bowel and barium follow-through has a low sensitivity and specificity of only 10% for detecting pathologies. Hence, endoscopic examination of the entire small bowel has always been a diagnostic challenge. Limitations of the diagnostic techniques in detection of the lesions located in the small bowel are mainly due to the length of the small intestine, overlying loops and intra-peritoneal location. This caused also the desire for autonomous instruments without the bundles of optical fibres and tubes, which are more than the size of the instrument itself, the reason for the objections of the patients. The use of highly integrated microcircuit in bioelectric data acquisition systems promises new insights into the origin of a large variety of health problems by providing lightweight, low-power, low-cost medical measurement devices. At present, there is only one type of microcapsule which has been introduced recently to improve the health outcome. This first swallowable video-capsule for the gastroenterological diagnosis has been presented by Given Imaging, a company from Israel, and its schematic diagram is illustrated in Fig. 1 [4].

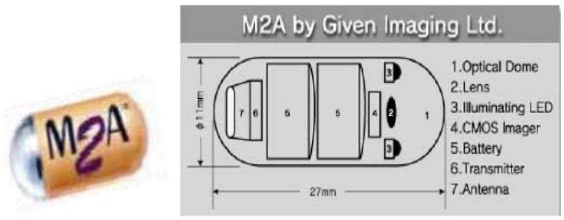

Figure 1: Given Imaging Capsule

The system consists of a small swallowable capsule containing a battery, a camera on a chip, a light source, and a transmitter. The camera-capsule has a one centimetre section and a length of three centimetres so it can be swallowed with some effort. In 24 hours, the capsule is crossing the patient's alimentary canal. For the purpose of this research work, endoscopic images have been obtained using this innovative endoscopic device. They have spatial resolution of 171x151 pixels, a brightness resolution of 256 levels per colour plane (8bits), and consisted of three colour planes (red, green and blue) for a total of 24 bits per pixel. The proposed methodology in this paper is considered in two phases. The first implements the extraction of image features while in the second phase an advanced neural network is implemented / employed to perform the diagnostic task. Texture analysis is one of the most important features used in image processing and pattern recognition. It can give information about the arrangement and spatial properties of fundamental image elements. Many methods have been proposed to extract texture features, e.g. the co-occurrence matrix, and the texture spectrum in the achromatic component of the image. The definition and extraction of quantitative parameters from endoscopic images based on texture information in the chromatic and achromatic domain is been proposed. This information is initially represented by a set of descriptive statistical features calculated on the histogram of the original image. Additionally, in this study an alternative approach of obtaining those quantitative parameters from the texture spectra is proposed both in the chromatic and achromatic domains of the image. The definition of texture spectrum employs the determination of the texture unit (TU) and texture unit number (N_{TU}) values. Texture units characterise the local texture information for a given pixel and its neighbourhood, and the statistics of the entire texture unit over the whole image reveal the global texture aspects. For the diagnostic part, the concept of multiple-classifier scheme has been adopted, where the fusion of the individual outputs was realised using fuzzy integral. An intelligent classifier-scheme based on the adaptive fuzzy logic methodology that utilises a novel defuzzification method, namely area of balance (AOB) has been implemented.

Image features extraction

A major component in analysing images involves data reduction which is accomplished by intelligently modifying the image from the lowest level of pixel data into higher level representations.

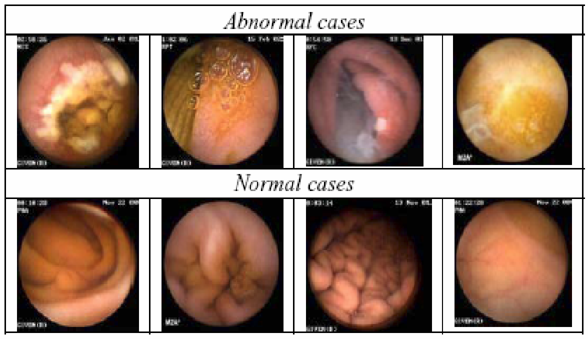

Figure 2: Selected endoscopic images of normal and abnormal cases.

Texture is broadly defined as the rate and direction of change of the chromatic properties of the image, and could be subjectively described as fine, coarse, smooth, random, rippled, and irregular, etc. For this reason, we focused our attention on nine statistical measures (standard deviation, variance, skew, kurtosis, entropy, energy, inverse difference moment, contrast, and covariance) [5]. All texture descriptors are estimated for all planes in both RGB {R (Red), G (Green), B (Blue)} and HSV {H (Hue), S (Saturation), V (Intensity)} spaces, creating a feature vector for each descriptor $D_i=(R_i,G_i,B_i,H_i,S_i,V_i)$. Thus, a total of 54 features (9 statistical measures x 6 image planes) are then estimated. For our experiments, we have used 70 endoscopic images related to abnormal cases and 70 images related to normal ones. Fig. 2 shows samples of selected images acquired using the M2A capsule of normal and abnormal cases. Generally, the statistical measures are estimated on histograms of the original image (1st order statistics). However, the histogram of the original image carries no information regarding relative position of the pixels in the texture. Obviously this can fail to distinguish between textures with similar distributions of grey levels. We therefore have to implement methods which recognise characteristic relative positions of pixels of given intensity levels. An additional scheme is proposed in this study to extract new texture features from the texture spectra in the chromatic and achromatic domains, for a selected region of interest from each colour component histogram of the endoscopic images.

N_{TU} Transformation

The definition of texture spectrum employs the determination of the texture unit (TU) and texture unit number (N_{TU}) values. Texture units characterise the local texture information for a given pixel and its neighbourhood, and the statistics of all the texture units over the whole image reveal the global texture aspects. Given a neighbourhood of $\delta \times \delta$ pixels, which are denoted by a set containing $\delta\times\delta$ elements $P=\{P_0,P_1,\ldots,P_{(\delta\times\delta)-1}\}$, where P_0 represents the chromatic or achromatic (i.e. intensity) value of the central pixel and $P_i\ \{i=1,2,\ldots,(\delta\times\delta)-1\}$ is the chromatic or achromatic value of the neighbouring pixel i, the $TU=\{E_0,E_1,\ldots,E_{(\delta\times\delta)-1}\}$, where $E_i\{i=1,2,\ldots,(\delta\times\delta)-1\}$ is determined as follows:

$$E_i = \begin{cases} 0, & if \quad P_i < P_0 \\ 1, & if \quad P_i = P_0 \\ 2, & if \quad P_i > P_0 \end{cases} \qquad (1)$$

The element E_i occupies the same position as the i^{th} pixel. Each element of the TU has one of three possible values; therefore the combination of all the eight elements results in 6561 possible TU's in total. The texture unit number (N_{TU}) is the label of the texture unit and is defined using the following equation:

$$N_{TU} = \sum_{i=1}^{(\delta\times\delta)-1} E_i \times \delta^{i-1} \qquad (2)$$

Where, in our case, $\delta = 3$. The texture spectrum histogram $(Hist(i))$ is obtained as the frequency distribution of all the texture units, with the abscissa showing the N_{TU} and the ordinate representing its occurrence frequency. The texture spectra of various image components {I (Intensity), R (Red), G (Green), B (Blue), H (Hue), S (Saturation)} are obtained from their texture unit numbers. The statistical features are then estimated on the histograms of the N_{TU} transformations of the chromatic and achromatic planes of the image (R,G,B,H,S,V).

Features evaluation

Recently, the concept of combining multiple classifiers has been actively exploited for developing highly reliable "diagnostic" systems [6]. In this study, six subsystems have been developed, and each of them was associated with the six planes specified in the feature extraction process (*i.e.* R, G, B, H, S, & V). Each subsystem was modelled with an appropriate intelligent learning scheme. In our case, a neuro-fuzzy scheme has been proposed. Such scheme provides a degree of certainty for each classification based on the statistics for each plane. The outputs of each of these networks must then be combined to produce a total output for the system as a whole as can be seen in Fig. 3.

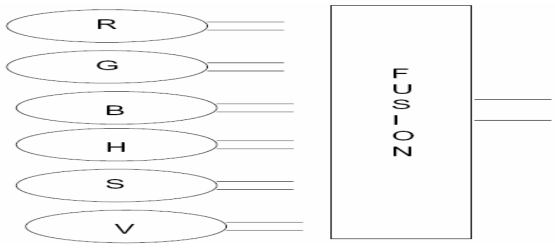

Figure 3: Proposed fusion scheme.

While a usual scheme chooses one best subsystem from amongst the set of candidate subsystems based on a winner-takes-all strategy, the current proposed approach runs all multiple subsystems with an appropriate collective decision strategy. The aim in this study is to incorporate information from each plane/space so that decisions are based on the whole input space. The adopted in this paper methodology was to use the fuzzy integral concept. Fuzzy integral (FI) is a promising method that incorporates information from each space/plane so that decisions are based on the whole input space in the case of multiple classifier schemes. FI combines evidence of a classification with the systems expectation of the importance of that evidence. By treating the classification results a series of disjointed subsets of the input space Sugeno defined the g_λ-fuzzy measure [7].

$$g(A\cup B) = g(A) + g(B) + \lambda g(A)g(B); \quad \lambda \in (-1,\infty) \qquad (3)$$

Where the λ measure can be given by solving the following non-linear equation.

$$\lambda + 1 = \prod_{i=1}^{K}\left(1+\lambda g^i\right) \qquad \lambda > -1 \qquad (4)$$

The $g^i, i \in \{1,\ldots,K\}$ values are fuzzy densities relating to the reliability of each of the K feature networks and satisfy the conditions of fuzzy sets laid out by Sugeno.

Adaptive Fuzzy Logic System (AFLS)

The classification scheme utilised here is an adaptive fuzzy logic system (FLS), *i.e.* a system having adaptive rules. Its structure is the same as a normal FLS but its rules are derived and extracted from given training data. In other words, its parameters can be trained like a neural network approach, but with its structure in a fuzzy logic system structure. Since we have general ideas about the structure and effect of each rule, it is straightforward to effectively initialise each rule. This is a tremendous advantage of AFLS over its NN counterpart. The AFLS is one type of FLS with a singleton fuzzifier and a defuzzifier. The centroid defuzzifier cannot be used because of its computation expense and that it prohibits using the backpropagation training algorithm.

The proposed AFLS consists of a new defuzzification approach, area of balance (AOB). Its feed-forward structure is shown in Fig. 4. The fuzzy basis layer consists of fuzzy basis nodes for each rule. A fuzzy basis node has the following form:

$$\phi_m(\bar{x}) = \frac{\mu_m(\bar{x})}{\sum_{l=1}^{L} \mu_l(\bar{x})} \qquad (5)$$

where $\varphi_m(\bar{x})$ is a fuzzy basis node for rule m and $\mu_m(\bar{x})$ is a membership value of rule m.

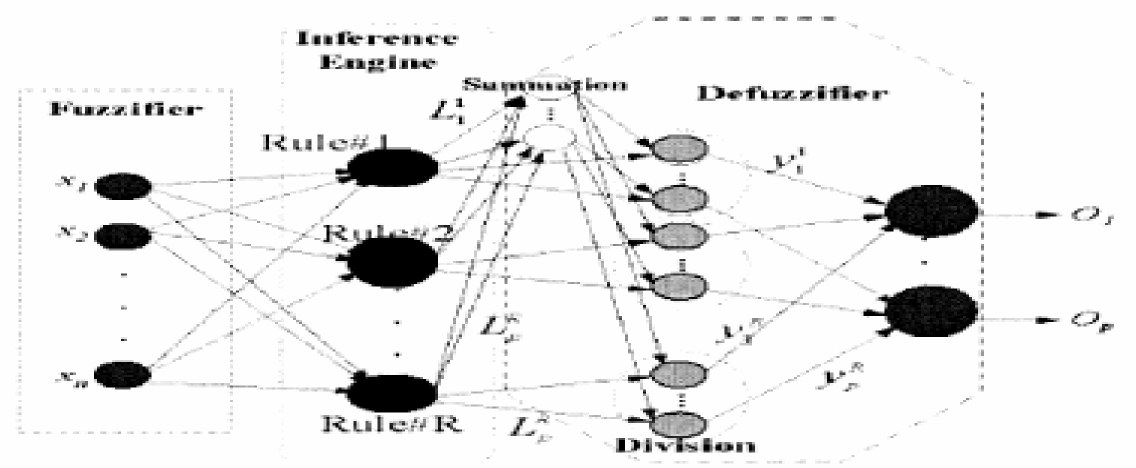

Figure 4: AFLS scheme.

Since we use a product-inference, the fuzzy basis node $\mu_m(\bar{x})$ is in the following form:

$$m(\bar{x}) = \prod_{i=1}^{n} \mu_{F_i^m}(x_i) \qquad (6)$$

where $\mu_{F_i^m}(x_i)$ is a membership value of the ith input of rule m. In our case, a Gaussian shape as a membership function of each input of each rule has been used, hence $\mu_{F_i^m}(x_i)$ will be in the following form:

$$\mu_{F_{ii}^m}(x_i) = \exp\left[-\frac{(x_i - c_i^m)^2}{2(b_i^m)^2}\right] \qquad (7)$$

where c_i^m and b_i^m are the centre and spread parameters, respectively, of the membership function i^{th} input of the m^{th} rule. The most popular defuzzification methods are the centroid of area (COA) and centre average (CA). The former although more accurate than the latter, is well known for its computational cost. Centroid calculation returns the centroid of the area formed by the consequent membership function, the membership value of its rules and the max-min or max- product inference. However, since the COA method provides good performance, its main characteristics, centre of gravity and use of the shape of membership function, will be preserved in the design of the proposed defuzzification approach. The overall output of the system may be the result of fuzzy union or the addition of rule outputs as in Kosko's method. The proposed AFLS uses Kosko's method with product inference [8]. In general form, the calculation of the output, y, will be

$$y_p = \frac{\sum_{m=1}^{M} \mu_m L_p^m y_p^m}{\sum_{m=1}^{M} \mu_m L_p^m} \qquad (8)$$

where y_p : the p^{th} output of the network, μ_m : the membership value of the m^{th} rule, L_p^m: the spread parameter of the membership function in the consequent part of the p^{th} output of the m^{th} rule, y_p^m : the centre of the membership function in the consequent part of the p^{th} output of the m^{th} rule .

ISSN: 1727-1983 © 2005 IFMBE

Results

The proposed approach was evaluated using 140 clinically obtained endoscopic M2A images. For the present analysis, two decision-classes are considered: abnormal and normal. Seventy images (35 abnormal and 35 normal) were used for the training and the remaining ones (35 abnormal and 35 normal) were used for testing.

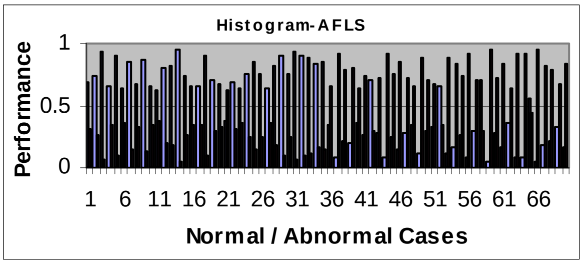

Figure 5: Histogram-based Performance for AFLS

The extraction of quantitative parameters from these endoscopic images is based on texture information. Initially, this information is represented by a set of descriptive statistical features calculated on the histogram of the original image. AFLS network was incorporated into a multiple classifier scheme, where the structure of each individual (for R, G, B, H, S, & V planes) classifier is consisted of 9 input nodes (i.e. nine statistical features) and 2 output nodes. In a second stage, the nine statistical measures for each individual image component are then calculated though the related texture spectra after applying the (N_{TU}) transformation.

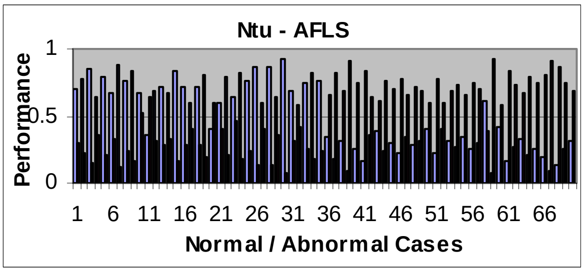

Figure 6: N_{TU}-based Performance for AFLS

The multiple-classifier scheme using the AFLS networks has been trained on the six feature spaces. Each AFLS consisted of 9 input nodes and 2 output nodes. This specific algorithm generally is character-rised by the "quality" of its performance [8]. The soft combination of neural classifiers using FI fusion concept resulted in 92.85% accuracy over the testing dataset (5 mistakes out of 70 testing patterns), demonstrating in this way the efficiency of this scheme in terms of accuracy. Despite the fact that training time was long, it is worth mentioned that results, as shown in Fig. 5, indicate a high confidence levels for each correct classification such as 0.63, while Table 1 presents the performance of individual components.

Table 1: ALFS performance

AFLS Accuracy (70 testing patterns)		
Modules	***Histogram-based***	***N_{TU}-based***
R	92.85% (5 mistakes)	94.28% (4 mistakes)
G	95.71% (3 mistakes)	95.71% (3 mistakes)
B	94.28% (4 mistakes)	97.14% (2 mistakes)
H	94.28% (4 mistakes)	97.14% (2 mistakes)
S	91.42% (6 mistakes)	92.86% (5 mistakes)
V	97.14% (2 mistakes)	97.14% (2 mistakes)
Overall	**92.85% (5 mistakes)**	**95.71% (3 mistakes)**

In the N_{TU}-based extraction process, the texture spectrum of the six components (R, G, B, H, S, V) have been obtained from the texture unit numbers, and the same nine statistical measures have been used in order to extract new features from each textures spectrum. The soft combination of neural classifiers using FI fusion concept resulted in 95.71% accuracy over the testing dataset (3 mistakes out of 70 testing patterns), demonstrating in this way again the efficiency of this scheme in terms of accuracy. More specifically, 2 normal cases as abnormal and one abnormal as normal one provide us a good indication of a "healthy" diagnostic performance. However the level of confidence in this case was slight less than the previous case (i.e. the histogram), that is 0.59 as shown in Fig. 6.

Conclusions

An approach on extracting texture features from endoscopic images using the M2A Given Imaging capsule has been developed. Statistical features based on texture are important features, and were able to distinguish the normal and abnormal status in the selected clinical cases. The multiple classifier approach used in this study with the inclusion of an advanced learning-based algorithm provided encouraging results. Two approaches on extracting statistical features from endoscopic images using the M2A Given Imaging capsule have been developed. In addition to the histogram-based texture spectrum, a new approach of obtaining those quantitative parameters from the texture spectra is proposed both in the chromatic and achromatic domains of the image by calculating the texture unit numbers (N_{TU}) over the histogram spectrum. Future studies will be focused on further development of this "diagnostic" system by incorporating additional features, investigation of algorithms for reduction of input dimensionality as well as the testing of this approach to the IVP-endoscopic capsule which is under development through a European research project. The authors would like to thank European Commission, Division of Information Society Technologies: Components and Subsystems, Applications for their support through the 5th FW "IVP" research project.

References

[1] HAGA Y., ESASHI M. (2004): ‘Biomedical Microsystems for Minimally Invasive Diagnosis and Treatment’, Proceedings of IEEE, Vol. 92, p. 98-114.

[2] KRISHNAN S., WANG P., KUGEAN C., TJOA M (2001): ‘Classification of endoscopic images based on texture and neural network’, Proc 23rd Annual IEEE Int. Conf. in Engineering in Medicine and Biology, Vol. 4, p. 3691-3695.

[3] MAROULIS D.E., IAKOVIDIS D.K., KARKANIS S.A., KARRAS D.A. (2003): ‘CoLD: a versatile detection system for colorectal lesions endoscopy video-frames’, *Computer Methods and Programs in Biomedicine,* **70**, pp. 151–166.

[4] IDDEN G., MERAN G., GLUKHOVSKY A., SWAIN P. (2000): 'Wireless capsule endoscopy', *Nature*, pp. 405-417.

[5] HARALICK R.M. (1979): ‘Statistical and structural approaches to texture’, *IEEE Proc*., **67**, pp. 786- 804.

[6] BOULOUGOURA M., WADGE E., KODOGIANNIS V.S., CHOWDREY H.S. (2004): ‘Intelligent systems for computer-assisted clinical endoscopic image analysis’, 2nd IASTED Int. Conf. on BIOMEDICAL ENGINEERING, Innsbruck, Austria, 2004, p. 405-408.

[7] KUNCHEVA L.I. (2000): ‘Fuzzy Classifier Design’, Physica-Verlag.

[8] KODOGIANNIS, V. (2001): ‘An efficient fuzzy based technique for signal classification’, *Journal of Intelligent & Fuzzy Systems*, **1**, pp. 65-84.