Capview: computer-aided diagnosis for endoscopic capsule

João Paulo Silva Cunha[1], Miguel Coimbra[2] and José Soares[3]

[1]IEETA/Dep. Electrónica, Telecomunicações e Informática, Universidade de Aveiro;

[2]IT-Porto/Faculdade de Ciências, Universidade do Porto;

[3]Serviço de Gastroenterologia, Hospital Geral de Stº. António, Porto

Resumo – **A endoscopia por cápsula é uma técnica muito recente que coloca diversos problemas novos à engenharia biomédica. Por um lado, trata-se dum exame de vídeo clínico de longa duração (até 8h) que necessita de ser revisto na sua totalidade para se chegar ao diagnóstico. Por outro lado, a necessidade de localizar eventos nas diferentes partes do tracto gastrointestinal e as características desse vídeo, como luminosidade, presença de artefactos na imagem, etc., são complexas o que torna difícil a sua interpretação. Apresentamos neste artigo a ferramenta computacional "*CapView*" tem vindo a agregar um conjunto de desenvolvimentos para auxílio ao diagnóstico neste tipo de exames de vídeo clínico.**

Abstract – **Endoscopic capsule is a rather recent technique that is raising new challenges to biomedical engineers. On one hand, it results on a long clinical video (up to 8h) that needs to be reviewed to reach a diagnosis. On the other hand, the location of events in the gastrointestinal tract and its image characteristics, such as luminance, artefact presence, etc. makes its analysis complex and difficult its interpretation. In this paper we present the "CapView" software tool that is being developed to support diagnosis in this type of clinical video exams.**

I. INTRODUCTION

The endoscopic capsule is the first autonomous micro-device to explore the human inner body of wide clinical application. This 11 by 30 mm device includes a camera, a light source, RF transmitter and batteries (see Fig. 1). It is ingested by the patient and films the whole gastrointestinal tract, reaching places where conventional endoscopy is not capable of. The full system consists of the capsule itself, an external receiving antenna and a portable hard drive worn by the patient. According to its distributor (Given Imaging, Israel), "over 230,000 patients worldwide have experienced the advantages of painless and effective PillCam™ Capsule Endoscopy" [1] The literature has shown its clinical importance [2,3,4].

Currently, one of the main setbacks of this new technology is the long duration of the exam analysis task. A specialized doctor needs to view around 60,000 images such as the ones in Fig. 2, looking for both abnormal situations (events) such as blood or ulcers, and defined topographic marks of the gastrointestinal tract (e.g.

Fig. 1. The wireless endoscopic capsule (1 - Optical dome; 2 - Lens holder; 3 – Lens; 4 – Illuminating LEDs; 5 – CMOS imager; 6 – Battery; 7 – ASIC transmitter; 8 – Antenna).

pylorum, ileo-cecal valve). This process, when performed by a trained specialist, can take about 2 hours. Ravens [4] comments that "the time a doctor needs to analyze the exam may be the most costly part of the procedure". There is a pressing need for automatic tools that can reduce these long annotation times. Even small time savings can be vital when multiplied by the 230,000 exams that, according to Given Imaging [1], have been performed worldwide.
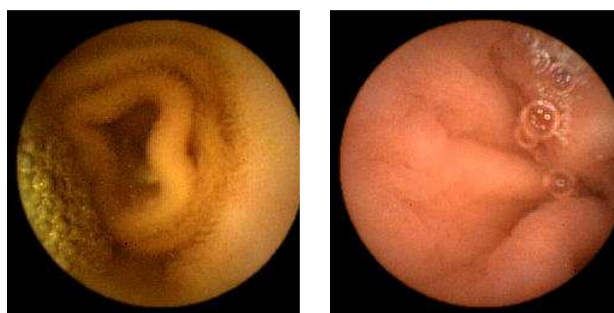

Fig. 2 – Examples of endoscopic capsule images

In the present paper we will describe our efforts in the R&D of computer aided tools that MDs can use to improve their diagnosis and reduce the revision time needed to perform diagnosis in capsule endoscopy.

II. THE CAPVIEW ANNOTATION SOFTWARE

The CapView annotation software resulted from the collaboration between biomedical engineers from the

IEETA/University of Aveiro (www.ieeta.pt/sias) and medical specialists working in capsule endoscopy at Santo António General Hospital, Porto, Portugal (www.hgsa.pt). The most vital contribution of CapView is providing a intuitive user friendly data management facilities for the large number of exams that have been accumulated in just a few years of capsule endoscopy (Fig. 3). When annotating an exam, a doctor labels video images with specific comments. The ones that are deemed important are added to the medical report and small videos of the event are stored.
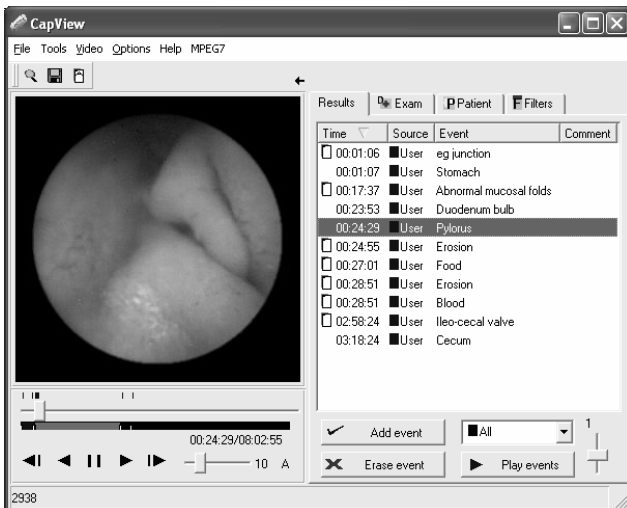

Fig. 3. CapView exam annotation software developed by our research group and currently used for routine work at two hospitals and a private clinic. Over 350 exams were already analyzed and annotated with this software tool.

The clinical experience at Santo António General Hospital (HGSA) shows that each doctor tends to use different annotation nomenclatures making database search a non-trivial task. An important contribution of *CapView* is to provide a normalized annotation language for creating medical reports. Annotated events are selected from lists of pre-defined events defined by the World Organization for Digestive Endoscopy [5] that are divided into the following topics: location, lumen, contents, diffuse alterations of the mucosa, flat lesions, protruding lesions, excavated lesions, and others. By forcing the labelling of topological marks, each event is associated with a location (e.g. blood event in the stomach). This greatly improves the efficiency of the reviewing process of the capsule video, reducing the time MDs need to reach a diagnosis and output a detailed report.

The CapView is also a platform where we can "plug-in" new algorithms to further improve the review procedure. One such example is the "topographic segmentation" algorithm recently published at the IEEE Trans. On Medical Imaging [6]. This algorithm automates this task for the medical specialist, associating each event with its topographic location and automatically extracting capsule transit times. Later, this information is inserted directly into the template-driven automated report creation functionality of the *CapView* software.

A. Topographic Segmentation algorithm

Our topographic segmentation system can be divided into three stages as shown in Fig. 4. We have a training stage where support vector machine (SVM) models and Bayesian distributions are obtained, a classification stage where each image is labeled as belonging to one of four topographic zones, and a final stage where results are fitted to a four section model representing the gastrointestinal tract.
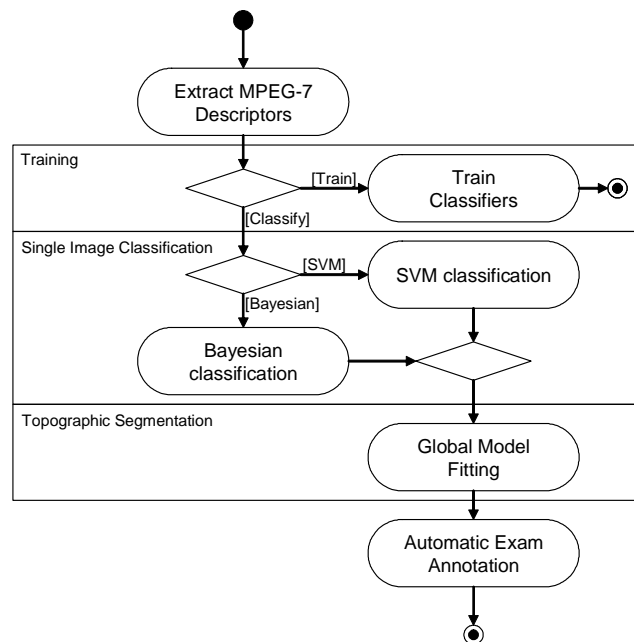

Fig. 4. Structure of the topographic segmentation system (using UML diagram notation [7]).

The dataset used for this study (see Table I) consists of 60 full endoscopic capsule exams of patients suffering from different small intestine pathologies (e.g. obscure bleeding, celiac disease). This amounts to a total of 3.6 million images, averaging 60,000 frames per exam. Each exam was manually annotated by a senior clinical specialist, who marked the temporal position of the eso-gastric junction, pylorus and ileo-cecal valve.

All data was anonymized and integrated onto the *IEETA-CapDB* database: an annotated endoscopic capsule database that aims to be a collection of real exams classified by medical experts, to be used by the computer vision research and development community (more info available at www.capview.org). For maximization of the relevance of our results, all 60 exams were used for testing and training. In order to avoid over-fitting of results, the well known 3-fold cross validation method was used. This consists in dividing the test set into sub-groups of 20 exams and using two for training and one for testing. Final results are obtained by averaging results of the three possible permutations. This technique avoids classifying data used in training and maximizes the usefulness of the available annotated data.

Nr. Patients	Total frames	Avg. frames per exam	Image Resolution
60	3.6 million	60,000	256x256
Avg. Frames Zone 1	Avg. Frames Zone 2	Avg. Frames Zone 3	Avg. Frames Zone 4
86	3974	31331	20004

Table I. Dataset details: Zone 1 – entrance; Zone 2 – stomach; Zone 3 – small intestine; Zone 4 – large intestine).

All results have been obtained on a Pentium4 3,2GHz with 1GB RAM. The well-known SVM-light software package [8,9] was used for SVM training. Classification results were obtained by a SVM-light module integrated into *CapView*.

To reduce development time, we've decided to use MPEG-7 visual descriptors as our low-level features. These constitute a set of well known and studied features with available free source code that we could use. We were especially interested in color and texture features (please see Manjunath [10] for definitions), and used the reference software freely available at the website of the Institute for Integrated Systems of TU Munich.

(www.lis.ei.tum.de/research/bv/topics/mmdb/e_mpeg7.ht ml).

Instead of testing all descriptors, we've selected the most useful ones using our own previous published relevance studies [11]. Following this, two were used for the experiments: *Scalable Colour* (SC) and *Homogenous Texture* (HT). They not only obtained best results but have a complementary nature evaluating both colour and texture.

Before looking at the median segmentation errors themselves, we can inspect an example of the behaviour of *SC Poly* (uses polynomial kernels) classifiers (Zones 1-4) near the esogastric junction of a real exam (transition from Zone 1 - entrance to Zone 2 - stomach). This will hopefully illustrate why the topographic segmentation method presented performs so well using such a simple global fitting stage. Fig. 5 shows the correct segmentation of this topographic junction: dark colour represents the distance of each image to the Zone 1 *SC Poly* SVM hyperplane, and light colour the distance to Zone 2. The vertical line shows the manually annotated transition. Although these distances have strong fluctuations, they usually produce good classifications (negative distances mean negative classification results) and there is a clear transition near the annotated boundary. For a detailed description of this method we refer to Burges' tutorial [12]
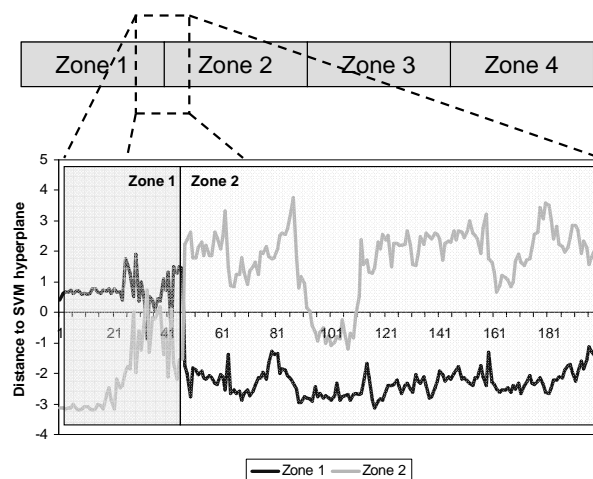

Fig. 5. Example of the behavior of Zone 1 and Zone 2 SC Poly classifiers near the esogastric junction. The horizontal axis shows temporal positions as image numbers

The SVM segmentation performance using selected classifiers can be seen on Table II. These include not only the best classifiers from the previous section but also comparative examples such as the simpler *SC_Lin* (uses linear kernels) and *L1_SC* (with "radial-based" functions) algorithms. SC with polynomial kernels performance obtained clearly the best segmentation performance as indicated in table II.

|		Median Error (images)			
		$E(z_{12})$	$E(z_{23})$	$E(z_{34})$	Total
SVM	SC Lin	60	326	2689	3021
	SC Poly	2	211	1070	1285
	SC Rbf	9	704	1902	2615
	HT Rbf	15	647	1815	2477
	SC+HT Rbf	11	295	1322	1630

Table II. Median errors of the topographic segmentation: E(12)-esogastric junction error, E(23)-pylorus error, E(34)-ileo-cecal valve error.

This tool is now integrated into the *CapView* annotation software (www.capview.org) developed by our group and currently used by two hospitals and a private clinic, responsible for over 250 capsule exams per year. When using this tool, topographic marks are automatically added to the events panel (red circles in Fig.6-a) and zones color-coded in the topographic bar (Fig.6-c). All this information is automatically included in the report generating functionalities of *CapView*, labeling the location of each manually annotated event (Fig.6-b) and automatic estimation and annotation of transit times (example of a real medical report: Fig.6-d).
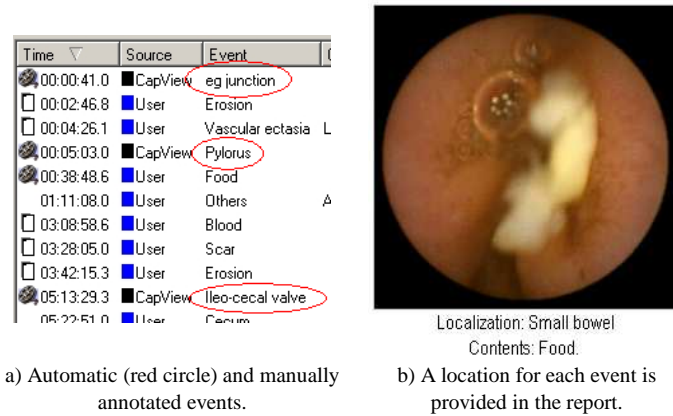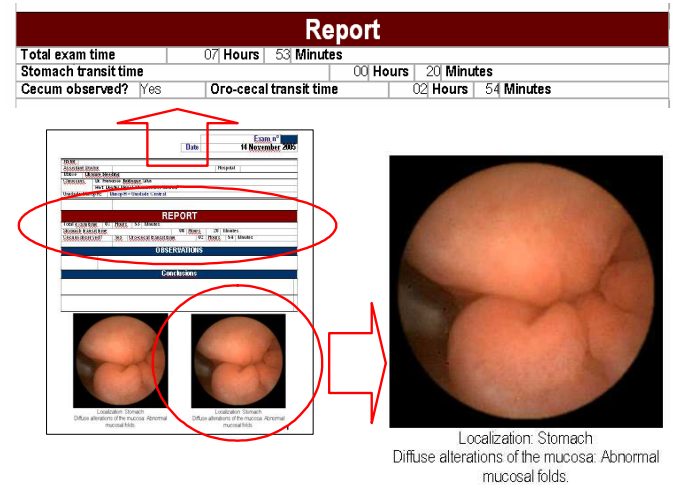
a) Automatic (red circle) and manually annotated events.

b) A location for each event is provided in the report.
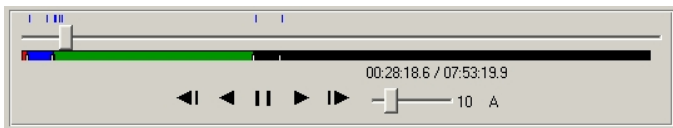
c) Topographic bar showing a different color per region (red – entrance, blue – stomach, green – small intestine, black – large intestine).

d) Example of automatically annotated report fields in a real medical report.

Fig. 6. Integration of the topographic segmentation tool into the *CapView* capsule exam annotation software.

III. CONCLUSIONS

The development of computer-aided tools to support diagnosis in endoscopic capsule has been a challenging task for the past few years. The *CapView* software tool has evolved from a simple video review tool to a sophisticated platform where we can "plug-in" different algorithms to improve the usability of the tool in the clinical workflow. *CapView* is now capable of supporting different parts of this workflow, from patient global search to automated report generation and segmentation tools.

Although we have made great progress and are now capable of providing a much efficient environment for the clinical workflow in endoscopic capsule exam analysis, we still have many improvements that can be done. Event

detection algorithms, such as polyps or other malignant or pre-malignant structures of the gastrointestinal tract, are our next target. These can also be "plugged-in" *CapView* to further improve its capacity to help clinicians in the difficult and long procedure of analysing endoscopic capsule videos.

REFERENCES

[1]   Given Imaging Home Page - www.givenimaging.com.

[2]   G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless Capsule Endoscopy", in *Nature*, pp. 405-417, (2000)..

[3]   W.A. Qureshi, "Current and future applications of the capsule camera", in *Nature*, vol.3, pp. 447-450, (2004).

[4]   A.F. Ravens, C.P. Swain, "The wireless capsule: new light in the darkness", in *Digestive Diseases*, vol. 20, pp. 127-133, (2002).

[5]   M. Crespi, M. Delvaux, M. Schapiro, C. Venables, and F. Zwiebel, "Minimal standard terminology v2.0", *Working party report, European Society of Gastrointestinal Endoscopy (ESGE)*. Available at: http://www.omed.org/omed2004/minimal.html

[6]   J. P. S. Cunha, M. Coimbra, P. Campos, and J. Soares, "Automated Topographic Segmentation and Transit Time Estimation in Endoscopic Capsule Exams," in *IEEE Trans. on Medical Imaging*, accepted for publication, 2007.

[7]   G. Booch, J. Rumbaugh, and I. Jacobson, *The Unified Modeling Language User's Guide*. Addison-Wesley, 1999.

[8]   SVM-Light support vector machine - http://www.cs.cornell.edu/People/ tj/svm_light/

[9]   T. Joachims, "Making large-Scale SVM Learning Practical", in Advances in Kernel Methods - Support Vector Learning, B. Schölkopf and C. Burges and A. Smola (ed.), MIT-Press, 1999.

[10]   B.S. Manjunath, J.R. Ohm, V.V. Vasudevan, and A. Yamada, "Color and texture descriptors", in *Special Issue on MPEG-7, Trans. Circ. Syst. for Video Tech.*, vol. 11/6, pp. 703-715, (2001).

[11]   M. Coimbra, and J.P. Silva Cunha "MPEG-7 Visual Descriptors – Contributions for Automated Feature Extraction in Capsule Endoscopy", in *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16/5, 2006, pp. 628-637.

[12]   C.J. Burges, "A tutorial on support vector machines for pattern recognition", in *Knowl. Disc. Data Mining*, vol.2, no.2, 1998, pp.1-43.