Video Processing Architecture: A Solution for Endoscopic Procedures Results

Isabel Laranjo[1,2], Joel Braga[1,2], Domingos Assunção[1], Carla Rolanda[2,3,4], Luís Lopes[5], Jorge Correia-Pinto[2,3,6], and Victor Alves[1]

[1] CCTC - Computer Science and Technology Center, University of Minho, Braga, Portugal
[2] Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Braga, Portugal
[3] ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal
[4] Department of Gastroenterology, Hospital de Braga, Portugal
[5] Department of Gastroenterology, Santa Luzia Hospital, Viana do Castelo, Portugal
[6] Department of Pediatric Surgery, Hospital de Braga, Portugal
{isabel,valves}@di.uminho.pt,
{joeltelesbraga,dassuncao}@gmail.com,
{crolanda,jcp}@ecsaude.uminho.pt, luis.m.lopes@mac.com

Abstract. In this paper we propose an architecture for processing endoscopic procedures results. The goal is to create a complete system capable of processing any type of endoscopic multimedia results, in order to overcome the most common issues in the endoscopic domain (e.g. video's long-duration, gastroenterologist's possible difficulty to maintain the focus and efficiency during the viewing process, imperfections in images/videos). It was this scenario that led to the conception of the *MIVprocessing* solution, which will address these and other problems, providing an added value to the elaboration of diagnoses. The *MIVprocessing* is composed of five tasks: Video Summarization (elimination of the "non-informative" frames); Pre-Processing (correction/improvement of the frames); Pre-Detection; Segmentation; and Feature Extraction and Classification. The idea is to create a framework that brings together the capabilities of different but at the same time complementary concepts (e.g. image and signal processing, machine learning, computer vision). This conjugation applied to the endoscopic domain provides a set of features capable of improving the gastroenterologist's activities during and after the procedure.

Keywords: Endoscopy, White-light Endoscopy, e-Health, MyEndoscopy, MIVprocessing, Video Summarization, Image and Video Processing.

1 Introduction

Image processing techniques are used in several areas (e.g. Medicine, Archaeology, Biology). Within these areas, the computational tasks of image processing (e.g. contrast enhancement, adjusting intensity levels of gray, fix blurry images), are mainly used to improve the images visual perception by humans [1, 2].

C. Ramos et al. (eds.), *Ambient Intelligence - Software and Applications*, Advances in Intelligent Systems and Computing 291,
DOI: 10.1007/978-3-319-07596-9_13, © Springer International Publishing Switzerland 2014

Another area of application for image processing techniques is the resolution of problems related to the perception of the image by the computer. In this case, the goal is to obtain information contained in the image in a way that the computer can interpret it [2]. This information rarely has some sort of resemblance to the visual features that humans use for interpreting the image content. The coefficients of the Fourier transform, statistical information or multidimensional distance measures, are examples of processing techniques used for computational perception [2].

While humans (e.g. radiologist, gastroenterologist) can learn to interpret patterns in an image the computer solves the problem with processing techniques. Despite significant advances in image processing and in computer vision techniques, the nuances of knowledge extraction from an image still challenges the best algorithms [3].

Nowadays, digestive endoscopy is used to allow the gastroenterologist to locate two types of endoscopic findings (diseases and lesions), and to evaluate a wide variety of endoscopic diagnoses/diseases (e.g. esophageal varices, gastric and duodenal ulcer, benign and malignant tumors) or, in most cases, to ensure that the symptoms are caused by lesions (e.g. polyp - protruding lesions) [4]. The detection and classification of polyps is the prevailing domain of research for the development of computer-aided decision support system. Of all the approaches found in *Liedlgruber et al.* [5], 47 belong to polyps (in particular polyps located in the colon).

This paper is organized as follows: Section 2 describes the problem, in Section 3 we present the related work and in the next Section (4) we present the proposed solution (*MIVprocessing)*. Finally, in Section 5 some conclusions are drawn and it is presented some future work.

2 Problem Definition

With the evolution of technology, the amount of information generated in the healthcare delivery has increased exponentially. In the endoscopic domain each procedure results in several gigabytes of multimedia information. This is related with the fact that sometimes the generated videos are too long. Depending on the type of endoscopic procedure, the videos may present a variable duration (e.g. Capsule Endoscopy ≈ 8 hours, Upper GastroIntestinal (GI) Endoscopy ≈ 25 minutes, Colonoscopy ≈ 40 minutes). The healthcare professionals can subsequently review the video in order to detect potential pathologies with a higher degree of confidence and accuracy, but at the same time, the time consuming reviewing process can make this task quite exhaustive and tedious.

Besides the time it takes to review the endoscopic procedure full-length video, this review is not always made in the best conditions due to technical issues, e.g. poor image quality, blurred image, the presence of specular reflection. These conditions make the presence of signs or symptoms that indicate the presence of pathology harder to detect.

These problems justify the conception of a solution for reducing the file size and duration of the video as well as improve the endoscopic images quality and the ability to assist the healthcare professionals on the detection/classification of endoscopic findings.

3 Related Work

From the capture and image processing systems found in the literature, there are some that stand as the basis for the several variants that have been developed. One of the better known base systems was proposed by *Gonzalez et al.* [2], which comprises a step for image/video acquiring, followed by the pre-processing, segmentation, feature representation and extraction, recognition and interpretation and, finally, the presentation of results. At each step, the information will be compared with a knowledge base.

Liedlgruber et al. presents, in its review work on endoscopic imaging [5], the main steps involved in computer-aided decision support system. Beyond the generic steps it includes a step for feature extraction and another for feature post processing. The classification step uses machine learning techniques.

Some recent studies have showed the potential of a Computer-Aided Diagnosis (CAD) system for the detection and classification of pathologies in endoscopic videos, providing the healthcare professional valuable technical support. There are CAD systems capable of detecting and/or classifying certain pathologies (e.g. polyps [6, 7], tumors [8], cancer [9], ulcers [6, 10], bleeding [11, 12]). Based on these studies a generic CAD system should follow a sequence of tasks: acquisition of image/video, low-level processing, segmentation and classification.

Other studies focuses on improving the quality of the image (e.g. elimination of specular reflection [13]) or detecting “non-informative” frames [14]. In these examples there are two common tasks: segmentation and classification. In 2007, *Lau et al.* [15] proposed a system for processing capsule endoscopic videos based on a framework that enables the analysis of multiple image features. This framework includes contrast enhancement tools, chamfer matching, threshold analysis, similarity between images, as well as color analysis tools.

All the studies referenced in this section implement some processing tasks that improve their own systems capabilities. However, there is no system integrating a complete set of tasks working together.

4 Proposed Architecture

The *MIVprocessing* was conceptualized from the idea of designing a processing system capable of grouping several tasks into a single framework, helping more effectively and accurately the healthcare professional at any time he is viewing, analyzing and/or interpreting the videos/images. The *MIVprocessing* is integrated in the *MyEndoscopy* system developed by *Laranjo et al.* [16]. The *MyEndoscopy* is a web-based application system, which allows the acquisition, processing, archiving and diffusion of endoscopic procedure results [16].

In Fig. 1 is presented a simple workflow describing the moments that occurs in a gastroenterology medical appointment in the healthcare institution that results in an endoscopy procedure (Moment 2). The *MIVprocessing* solution can be seamlessly integrated in the current workflow by performing some additional processing tasks, which can help the gastroenterologist's activities during and after the procedure.

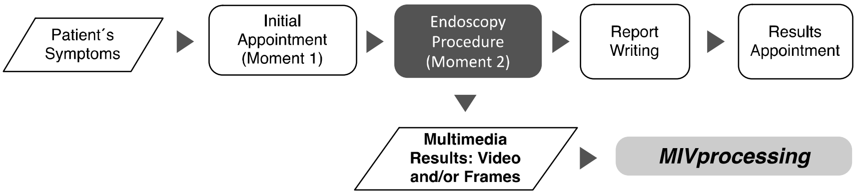


Fig. 1. Workflow of a gastroenterology medical appointment

The *MIVprocessing* results from the merger of two types of systems (**Supportive System** and **Computer-Aided Diagnosis (CAD) System**) and the tasks running on each of them (**Fig. 2**). The Supportive system consists of tasks that will support the CAD system. This means that these tasks are the basis of processing to be applied to videos, in order to make the detection and classification of abnormal pathologies more efficient and accurate in CAD system. The Supportive system is composed of three tasks (Video Summarization; Pre-Processing; and Pre-Detection), while the CAD system is composed of two tasks (Segmentation; Feature Extraction and Classification).

The *MIVprocessing* is integrated in *MIVbox*, which is a device for the acquisition, processing and storage of the endoscopic results.

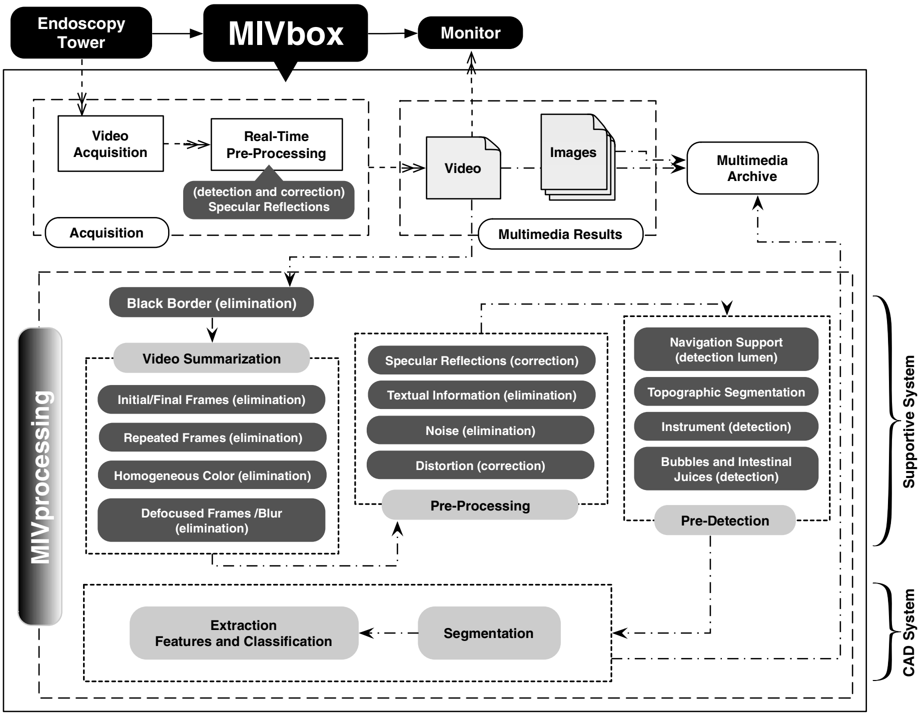
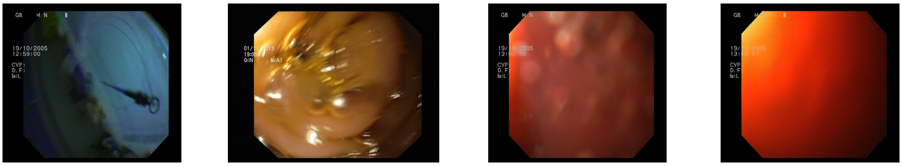


Fig. 2. MIVprocessing's architecture

Fig. 2 details the *MIVprocessing* architecture and also illustrates the steps that occur since video acquisition until archiving, not forgetting the visualization and processing steps. During the video acquisition and visualization (this process is based on the system described in [17]), *MIVprocessing* performs the detection and correction of specular reflections in real-time. The focus of this solution is in the white-light endoscopy imaging technique.

There is a preliminary task that deals with the elimination of the black borders, since these borders can interfere with the video processing, due to the contrast with the region of interest.

The main task, **video summarization**, aims to decrease the duration and file size of the full-length video, eliminating the "non-informative" frames (e.g. initial/final - captured outside the GI tract (Fig. 3.a), blurred (Fig. 3.b), homogeneous color (Fig. 3.c, Fig. 3.d), repeated) for easier visualization in the future by the gastroenterologist.

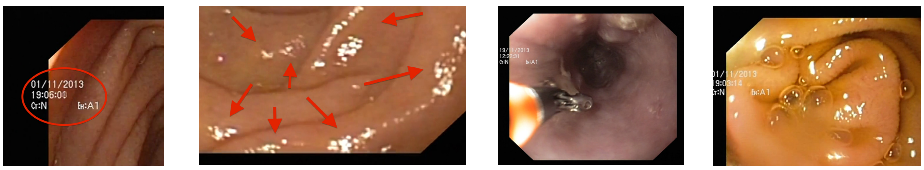


(a) (b) (c) (d)

Fig. 3. "Non-informative" frames: (a) initial, (b) defocused and (c, d) homogeneous color

Even after removing the frames considered not relevant, the reduced-length video still contains some information that, due to its characteristics, is considered unnecessary or even detrimental to the diagnosis. Some examples of this can be the noise present in the images, regions with textual information (Fig. 4.a) or regions with specular reflection (brightness) (Fig. 4.b). The goal of the main task, **pre-processing** (or low-level processing) is exactly to remove the information that is not necessary or even detrimental for diagnosis purposes. Some of the routines included are: cutting regions of interest, image rotation, scaling, color conversion, contrast enhancement or distortion correction.

The **pre-detection** is the last main task belonging to the Supportive system. Its routines are: navigation support, topographic segmentation, instrument detection ((Fig. 4.c), bubbles (Fig. 4.d) and intestinal juices detection and elimination. Detection of



(a) (b) (c) (d)

Fig. 4. Regions with (a) textual information, (b) specular reflection, and frames with (c) instrument and (d) bubbles

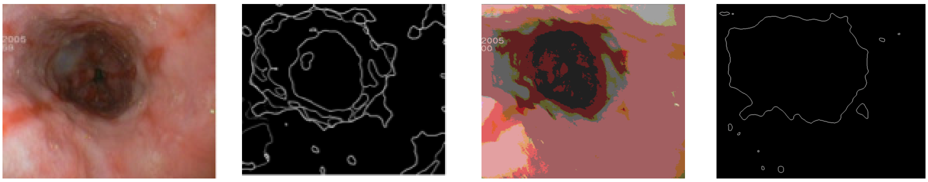
the lumen (in navigation support) allows the extraction of important parameters for processing algorithms, e.g. determining the center and the border, which allows knowing the orientation of the GI tract. On the topographic segmentation occurs the video division into segments, specifically in different parts of the GI tract.

The purpose of the CAD system is to address the following issues:

- The correct detection of a given pathology;
- The exact location, along the GI tract, of the pathology found;
- Correct classification of the endoscopic findings.

The CAD system tasks aim to perform video segmentation, visual features extraction and its classification into semantic information. It also provides a reliable and accurate help at the time of diagnosis, never leading to the adulteration of the original information contained in the video. It is important to make clear that none of these tasks are intended to replace the gastroenterologist in the elaboration of diagnoses, since they are only support tools to help the diagnostic.

Segmentation is the task where the input data is divided into its constituents. The main role of segmentation is to distinguish the image background from the objects delimitation. The segmentation techniques used in this task are thresholding (Fig. 5.b), segmentation based on regions (Fig. 5.c), segmentation based on contours (Fig. 5.d) and classification by pixels.



(a) (b) (c) (d)

Fig. 5. An example of image segmentation (a) original, (b) segmented - adaptive thresholding, (c) segmentation based on regions and (d) segmentation based on contours

At last, the CAD system has its **feature extraction and classification** task, which in a first moment selects and extracts the most important features (e.g. shape, texture, color) present in each frame and then, in a second moment, assigns a semantic designation to the endoscopic finding. The methodologies used in this task can be divided in two categories: Signal processing (processing information in the frequency domain, e.g. Fourier transforms, wavelet transforms) and Artificial intelligence (uses machine learning tools, e.g. Support Vector Machines (SVM), Bayes classifier).

As noted in [18], the clinical procedures and terminologies must be represented in expressive models that allow the storage of all relevant and necessary information. To ensure lexical compatibility of extracted and classified information, a standardized endoscopic vocabulary, the Minimal Standard Terminology (MST), is used [19]. The choice fell on the MST, in some part, because it is a terminology directed at Digestive

Endoscopy, but also for being complete and detailed (the terms described vary depending on the body part that is being examined and depending on the endoscopic technique), and easy to interpret (well organized and structured).

5 Conclusions and Future Work

In this paper, we gave an overview of the proposed solution (*MIVprocessing*). This is a solution that can be used with any type of endoscopic procedure – e.g. Capsule Endoscopy, Upper GI Endoscopy, Colonoscopy (with minor adjustments on the algorithms of feature extraction and semantic classification). In the case of Capsule Endoscopy, there is a slight difference once there is no detection step and elimination of specular reflection, in real-time.

MIVprocessing allows solving the problems described in section two with the following contributions:

- **Reduction in the duration and file size of the full-length video:** with the goal of creating a reduced-length video with only the informative frames;
- **Medical imaging quality improvement:** processing and/or elimination of detrimental information to the elaboration of the diagnosis;
- **Ability to assist the gastroenterologist in the detection/classification of endoscopic findings:** automatic detection of pathologies; defining the exact location of the detected pathology; and correctly classify the endoscopic finding.

The set of functionalities that are integrated in this solution makes it an asset to the quality increase in the provision of healthcare. Currently, the *MIVprocessing* is in the testing phase. Being integrated in *MyEndoscopy* it will provide this system with an extra set of functionalities, making it an optimal solution to the endoscopic domain.

From the studies referenced in the related work section, there is no study proposing an integrated system with a complete set of tasks working together. *MIVprocessing* has an architecture that integrates several pieces proposed in the literature and builds a complete endoscopic processing system.

Future work will be focused on fine tuning the proposed solution and evaluating it not only at the technical level, but also at the functional level by using end users like gastroenterologists and researchers. The assessment of the medical experts' opinion is also an important step towards optimizing the task of classification of endoscopic pathologies, since these experts are the ones that can identify the most relevant features for correct identification of pathologies.

Although the *MIVprocessing* has been designed for the white-light endoscopy imaging technique it can be extended and adapted to other techniques (e.g. Narrow Band Imaging (NBI), Confocal Laser Endomicroscopy (CLE)).

Acknowledgments. This work is funded by ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT - *Fundação para a Ciência e a Tecnologia* (Portuguese Foundation for Science and Technology) within project

FCOMP-01-0202-FEDER- 013853 and by National Funds through the FCT - *Fundação para a Ciência e a Tecnologia* (Portuguese Foundation for Science and Technology) within project PEst-OE/EEI/UI0752/2014.

References

1. Moeslund, T.B.: Introduction to Video and Image Processing: Building Real Systems and Applications. Springer, London (2012)
2. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Pearson Education International (2009)
3. Bui, A.A.T., Taira, R.K., Kangarloo, H.: Introduction - What is Medical Imaging Informatics? In: Bui, A.A.T., Taira, R.K. (eds.) Medical Imaging Informatics, pp. 3–14. Springer US (2010)
4. Schiller, K.F.R., Warren, B.F., Hunt, R.H.: Atlas of Gastrointestinal Endoscopy and Related Pathology. Wiley-Blackwell (2002)
5. Liedlgruber, M., Uhl, A.: Computer-Aided Decision Support Systems for Endoscopy in the Gastrointestinal Tract: a Review. IEEE Rev. Biomed. Eng. 4, 73–88 (2011)
6. Karargyris, A., Bourbakis, N.: Detection of Small Bowel Polyps and Ulcers in Wireless Capsule Endoscopy Videos. IEEE Trans. Biomed. Eng. 58, 2777–2786 (2011)
7. Alexandre, L.A., Casteleiro, J.M., Nobreinst, N.: Polyp Detection in Endoscopic Video Using SVMs. In: Kok, J.N., Koronacki, J., Lopez de Mantaras, R., Matwin, S., Mladenič, D., Skowron, A. (eds.) PKDD 2007. LNCS (LNAI), vol. 4702, pp. 358–365. Springer, Heidelberg (2007)
8. Barbosa, D.C., Roupar, D.B., Ramos, J.C., Tavares, A.C., Lima, C.S.: Automatic Small Bowel Tumor Diagnosis by using Multi-scale Wavelet-based Analysis in Wireless Capsule Endoscopy Images. Biomed. Eng. Online. 11, 1–17 (2012)
9. Iakovidis, D.K., Maroulis, D.E., Karkanis, S.: a: An Intelligent System for Automatic Detection of Gastrointestinal Adenomas in Video Endoscopy. Comput. Biol. Med. 36, 1084–1103 (2006)
10. Chen, Y., Lee, J.: Ulcer Detection in Wireless Capsule Endoscopy Video. In: Proceedings of the 20th ACM International Conference on Multimedia, pp. 1181–1184. ACM (2012)
11. Pan, G., Yan, G., Qiu, X., Cui, J.: Bleeding Detection in Wireless Capsule Endoscopy Based on Probabilistic Neural Network. J. Med. Syst. 35, 1477–1484 (2011)
12. Li, B., Meng, M.Q.-H.: Computer-Aided Detection of Bleeding Regions for Capsule Endoscopy Images. IEEE Trans. Biomed. Eng. 56, 1032–1039 (2009)
13. Stehle, T.: Removal of Specular Reflections in Endoscopic Images Removal of Specular Reflections in Endoscopic Images. Acta Polytech. J. Adv. Eng. 46, 32–36 (2006)
14. Bashar, M.K., Kitasaka, T., Suenaga, Y., Mekada, Y., Mori, K.: Automatic Detection of Informative Frames from Wireless Capsule Endoscopy Images. Med. Image Anal. 14, 449–470 (2010)
15. Lau, P.Y., Correia, P.L.: Analyzing Gastrointestinal Tissue Images using Multiple Features. In: Proceedings of the International Conference on Telecommunications, pp. 435–438 (2007)
16. Laranjo, I., Braga, J., Assunção, D., Silva, A., Rolanda, C., Lopes, L., Correia-Pinto, J., Alves, V.: Web-Based Solution for Acquisition, Processing, Archiving and Diffusion of Endoscopy Studies. In: Omatu, S., Neves, J., Rodriguez, J.M.C., Paz Santana, J.F., Gonzalez, S.R. (eds.) Distrib. Computing & Artificial Intelligence. AISC, vol. 217, pp. 317–324. Springer, Heidelberg (2013)

17. Braga, J., Laranjo, I., Assunção, D., Rolanda, C., Lopes, L., Correia-Pinto, J., Alves, V.: Endoscopic Imaging Results: Web based Solution with Video Diffusion. Procedia Technol. 9, 1123–1131 (2013)
18. Oliveira, T., Novais, P., Neves, J.: Guideline Formalization and Knowledge Representation for Clinical Decision Support. Adv. Distrib. Comput. Artif. Intell. J. 1, 1–12 (2012)
19. Aabakken, L., Rembacken, B., LeMoine, O., Kuznetsov, K., Rey, J.-F., Rösch, T., Eisen, G., Cotton, P., Fujino, M.: Minimal Standard Terminology for Gastrointestinal Endoscopy (MST 3.0) (2009)