34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Segmentation of small bowel tumor tissue in Capsule Endoscopy Images by using the MAP algorithm

Pedro Vieira, Jaime Ramos, Daniel Barbosa, Dalila Roupar, Carlos Silva, Higino Correia and Carlos S. Lima, *Member, IEEE*

Abstract— **State of the art algorithms for diagnosis of the small bowel by using capsule endoscopic images usually rely on the processing of the whole frame, hence no segmentation is usually required. However, some specific applications such as three-dimensional reconstruction of the digestive wall, detection of small substructures such as polyps and ulcers or training of young medical staff require robust segmentation. Current state of the art algorithms for robust segmentation are mainly based on Markov Random Fields (MRF) requiring prohibitive computational resources not compatible with applications that generate a great amount of data as is the case of capsule endoscopy. However context information given by MRF is not the only way to improve robustness. Alternatives could come from a more effective use of the color information. This paper proposes a Maximum A Posteriori (MAP) based approach for lesion segmentation based on pixel intensities read simultaneously in the three color channels. Usually tumor regions are characterized by higher intensity than normal regions, where the intensity can be measured as the vectorial sum of the 3 color channels. The exception occurs when the capsule is positioned perpendicularly and too close to the small bowel wall. In this case a hipper intense tissue region appears at the middle of the image, which in case of being normal tissue, will be segmented as tumor tissue. This paper also proposes a Maximum Likelihood (ML) based approach to deal with this situation. Experimental results show that tumor segmentation becomes more effective in the HSV than in the RGB color space where diagonal covariance matrices have similar effectiveness than full covariance matrices.**

I. Introduction

State of the art systems for tumor detection in the small bowel usually do not consider the Region of Interest (ROI) paradigm therefore frames are processed usually as a whole [1], [2]. Therefore image segmentation is not a major concern for this type of systems. However applications such as training of medical staff, three dimensional reconstruction of the digestive wall and detection of small structures such as polyps and ulcers rely on robust segmentation algorithms.

Karargyris and Bourbakis [3] use segmentation to extract regions where similarity between two frames is used for frame interpolation which goal is three-dimensional reconstruction of the digestive wall. The same authors [4] use segmentation with the purpose of detection of polyps and ulcers in the small bowel. They improved the current paradigm mainly based on color and texture characterization to a new paradigm that also includes anatomical characteristics. For this purpose a robust segmentation procedure is needed since the anatomy makes part of the classification process.

The purpose of this paper is to present a robust segmentation algorithm capable of accurately detect small bowel tumors with applications at the medical training level. The algorithm can also be extended to other pathologies and/or applications such as them described in [3] and [4]. Lesion and normal tissues usually differ in texture and most part of current automatic lesion detection systems are therefore based on texture measures [1], [2]. However as texture is usually characterized by a significant group of pixels while segmentation is a pixel by pixel (or a very small group of pixels) classification procedure, segmentation by using texture measures is not obvious to conceive.

Looking at a lesion we can see that lesion tissue has usually white spots or increased lightning partially due to a higher reflection coefficient when compared with normal tissue. This suggests that a segmentation process based on intensity can be adequate to separate normal from abnormal tissue. In this paper we suggest a lightning measure obtained from the vectorial sum of the three color channels, since the segmentation will be based on simultaneous observations in the three channel color, hence three dimensional vectors are observed. At this point it is convenient to mention that pixel intensity alone is not a reliable way to classify normal from abnormal frames. In fact both normal and abnormal frames have different intensity pieces of tissue. However, according to our observations, it is a fact that in abnormal frames lesion tissue has higher intensities than normal tissue. Regarding normal frames a segmented image is also obtained, however the ML procedure classifies these frames as normal frames.

State of the art approaches for image segmentation are usually classified in three groups; boundary-based, region-based and integration of both boundary and region based.

Boundary-based approaches are known as Snakes [5] or sometimes as parametric deformable models. Gradient Vector Flow (GVF) [6] is currently seen as a Snake that combines almost all the merits of previous snakes with a large capture range by using a regularization term. Corresponding to parametric deformable models are geometric deformable models or level sets, where the evolution is in the form of curve [7]. The main drawback sometimes associated with boundary based methods is that they are sensitive to noise and easy to be trapped in local minimum in noisy images [8].

Region-based techniques segment the image/volume into regions/subvolumes. These can be further classified into

This work was partly funded by FEDER grants through the Operational Competitiveness Program - COMPETE and by national grants through FCT - Foundation for Science and Technology, with the Project: FCOMP-01-0124-FEDER-022674.

Pedro Vieira, Carlos Silva, Higino Correia and Carlos S. Lima are with Industrial Electronics Department of University of Minho, Campus de Azurem, 4800-058 Guimaraes Portugal. e-mail: carlos.lima@dei.uminho.pt.

Jaime Ramos is with the Gastroenterology Department, Hospital dos Capuchos, Portugal.

Daniel J. Barbosa and Dalila Roupar are with the Catholic University of Leuven, Belgium.

many categories however in most cases a conditional probability in the Gaussian form represents the possibility of the observed data given its actual value. One of the most promising methods is Markov Random Field (MRF) [9][10] where the prior probability models the spatial constraint in the neighborhood. Region-based approaches in general and MRF in particular are well-known for their robustness, but they are computationally expensive and therefore not adequate for the current application where a great amount of images are generated in each capsule exam. To alleviate computational load we propose to disregard context (neighborhood) information and proceed with the Maximum A Priori (MAP) approach, however trying to improve robustness by a more effective use of color information.

The main idea of the proposed algorithm is to segment the image into two regions and taking the most intense as the ROI. However sometimes segmentation failures arose. One of the most common cases of segmentation failures occur when the camera is located perpendicular and too close to the wall intestine. In this case a high intense region located at the center of the frame is selected as lesion. We propose dealing with this situation by using a double segmentation procedure, into two and three classes. The right model can be chosen by using Maximum Likelihood (ML) estimation. In the cases where a three Gaussian model is more likely than a two Gaussian model the most intense class is discarded and the ROI is assumed to be the second most intense region. The proposed algorithm was tested in RGB and HSV color spaces by using full and diagonal covariance matrices. The main conclusion is that HSV color space is more effective than RGB color space. In RGB color space full covariance matrices are more effective than diagonal covariance matrices showing the correlation between color channels in RGB. HSV performance is similar when using diagonal and full covariance matrices showing that the channels are not correlated.

II. METODOLOGY

A. Statistical Model

Regarding region-based image segmentation the most used statistical model is the Gaussian Mixture Model (GMM). The number of mixture components is usually the number of expected classes. In the present case the goal is to distinguish normal and abnormal tissues, hence a two Gaussian model seems to be appropriate. Looking at figure 1 we can see that the lesion tissue has higher intensity than normal tissue, so intensity can perhaps be a discriminating factor between normal and abnormal tissue.

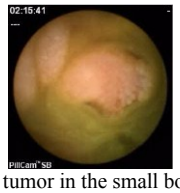

Figure 1. Image of a tumor in the small bowel taken from CE.

Figure 2 shows a frame where the camera was too close to the intestine wall which originates a super-intense region that will be classified as tumor region. To overcome this difficulty, a double segmentation process is required and a ML procedure can be used to choose the most likely number of classes presented in the frame including one only class which means frame containing only normal tissue.

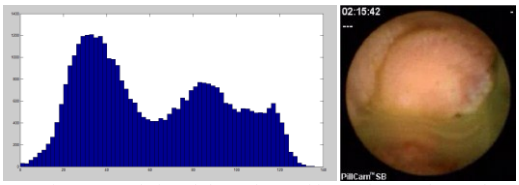

Figure 2. Histogram (left) of the values of intensity on the B channel of an image with a tumor (right).

The current GMM is a multivariate GMM, where observations are three dimensional vectors since the three channels of color are used [11]. RGB and HSV color spaces are used in the ambit of this paper with diagonal and full covariance matrices.

B. Bayesian classification

Bayes rule, Eq. 1, gives the posterior probability of a class, given feature measurements [12].

$$P\left(C_i|x_j\right) = \frac{p\left(x_j|C_i\right).P(C_i)}{p(x_j)}. \qquad (1)$$

In this equation j refers to a different position in the image, i to a different class, $p(x_j|C_i)$ represents the conditional probability density function of the measure x_j for the class C_i; $P(C_i)$ is the *a priori* probability of class C_i; and $p(x_j)$ the probability density function of x_j also known as the evidence and serves only as a scale factor, which makes Eq. 1 a true probability [12]. It can usually be ignored when comparing the result of Eq. 1 for all classes. Eq. 1 without $p(x_j)$ is the maximum a posteriori (MAP) estimate of Eq. 1, which can be computed, assuming only two classes as:

If $p(x_j|C_1).P(C_1) > p(x_j|C_2).P(C_2)$, then x_j belongs to C_1, otherwise x_j belongs to C_2.

Since these models are not known for small bowel tissues they are required to be estimated. This estimation procedure is usually achieved by using the Expectation-Maximization (EM) Algorithm.

C. Estimation of probability density functions

To estimate $p(x_j|C_i)$ the well-known EM algorithm is used. The main idea of the EM algorithm is to iteratively find the most appropriate parameters of the GMM model according to the ML criterion. The likelihood of an observation vector regarding the GMM is given by:

$$p\left(x_j|\varphi\right) = \sum_{i=1}^{k} P(C_i).p(x_j \mid C_i, \varphi_i). \qquad (2)$$

In Eq. 2 φ_i is the vector containing the parameters of the distribution of class C_i [13]. The likelihood of the whole image can be computed as:

$$p\left(X|\varphi\right) = \prod_{j=1}^{n} p(x_j \mid \varphi). \qquad (3)$$

The X represents the set of all the samples ($x_0, x_1, \ldots, x_n$). The log-likelihood of the whole image is given by:

$$l(\varphi) = \log p(X|\varphi) = \log \prod_{j=1}^{n} p(x_j|\varphi) = \sum_{j=1}^{n} \log p(x_j|\varphi) = \sum_{j=1}^{n} \log \sum_{i=1}^{k} P(C_i).p(x_j|C_i,\varphi_i) \quad (4)$$

Maximization of this log-likelihood function requires the application of the gradient operator [13], which results in:

$$\nabla_{\varphi_k} l(\varphi) = \sum_{j=1}^{n} \frac{1}{p(x_j|\varphi)}.\nabla_{\varphi_k}[\sum_{i=1}^{k} P(C_i).p(x_j|C_i,\varphi_i)] = \sum_{j=1}^{n} \frac{1}{p(x_j|\varphi)} \nabla_{\varphi_k}[P(C_k).p(x_j|C_k,\varphi_k)] \quad (5)$$

The log-likelihood function can be written only in terms of *a priori* and *a posteriori* probabilities and probability density function of each class and vector of data:

$$\nabla_{\varphi_k} l(\varphi) = \sum_{j=1}^{n} P(C_k|x_j,\varphi_k).\frac{\nabla_{\varphi_k}[P(C_k).p(x_j|C_k,\varphi_k)]}{p(x_j|C_k,\varphi_k).P(C_k)} = \sum_{j=1}^{n} P(C_k|x_j,\varphi_k) \nabla_{\varphi_k} \log[p(x_j|C_k,\varphi_k).P(C_k)] \quad (6)$$

Maximizing the likelihood of the data requires equation the gradient to zero. Knowing that the probability density function of a multivariate normal distribution is given by:

$$p(x_j|C_k,\varphi_k) = \frac{1}{(2\pi_k)^{\frac{D}{2}}.|\Sigma_k|^{\frac{1}{2}}}.\exp(-\frac{1}{2}(x_j-\mu_k)^T \Sigma_k^{-1}(x_j-\mu_k)) \quad (7)$$

where D is the dimension of the distribution (in this case, D=3) [14]. The update of the distribution parameters (mean, covariance matrix and *a priori* probability) [13] are given by:

$$\hat{\pi} = \frac{1}{n}\sum_{j=1}^{n} P(C_k|x_j,\varphi_k)$$

$$\hat{\mu} = \frac{\sum_{j=1}^{n} P(C_k|x_j,\varphi_k) x_j}{\sum_{j=1}^{n} P(C_k|x_j,\varphi_k)}$$

$$\hat{\Sigma} = \frac{\sum_{j=1}^{n} P(C_k|x_j,\varphi_k)(x_j-\mu)(x_j-\mu)^T}{\sum_{j=1}^{n} P(C_k|x_j,\varphi_k)} \quad (8)$$

Algorithmically we have an iterative process that maximizes the likelihood of the data. Initial estimates are however required and can be given by the K-means algorithm, which divides the pixels into k clusters in such a way that each pixel will be in the cluster with the nearest mean of the pixel value.

The EM has two main steps:

- **E-step** (expectation): according to previous estimated parameters, the likelihood of each sample for each cluster is computed.
- **M-step** (maximization): each pixel is associated with the cluster for the which the likelihood is higher. When this is done for all samples, new estimates of the model parameters are calculated.

After the M-step, the algorithm is repeated until convergence is achieved. Convergence is achieved when the content of each cluster doesn't change in consecutive iterations [15].

E. Color Space

There are many color spaces that can be used to represent images. All have some advantages and disadvantages and are better for some application. The most common are RGB and HSV. RGB is probably the most used color space, and it's an additive model. A huge disadvantage presented with the use of this color space is its high correlation between the different components (Red, Green and Blue). HSV (Hue, Saturation and Value) is another color space used many times, and it's a cylindrical-coordinate representation of the RGB model. It was designed to better approximate the human perceive and interpretation of color [16] and therefore usually the preferred for machine vision applications.

III. EXPERIMENTAL RESULTS

A PC with a 2.20 GHz i7-2670QM processor, with 6 GB of RAM was used with MATLAB to run the proposed algorithm. The algorithm proposed had a convergence time of approximately 1 second for each frame. The first segmentation procedure divides the set of pixels into 3 clusters, and the second one divides the frame into 2 clusters. The log-likelihood under each distribution was calculated in order to see which segmentation is best suited for each image. The algorithm was ran using RGB and HSV color spaces by considering both cases of random variables identically distributed in both dependent and independent cases.

Figures 3, 4 and 5 show three frames with tumors to provide visual results of this segmentation method. In each figure it can be seen the original image on the top, in the middle row the segmentation in RGB color space and on the bottom row the segmentation in HSV color space. From left to right it can be seen segmentation into three clusters considering the space color as dependent random variables, and independent; and segmentation in two clusters with similar considerations regarding statistical dependence. These results show that more segmentation errors appear in RGB than in HSV color spaces. Comparing RGB and HSV based segmentations, it can be seen that in general, the segmentation is better in the HSV color space.

Figure 5 shows that frames with high intensity normal regions are better characterized by a three component GMM. Figures 3 and 4 show that when a tumor is the brightest area of the image the best distribution is the GMM with two components. These results can be confirmed by comparing the log-likelihood of each segmentation.

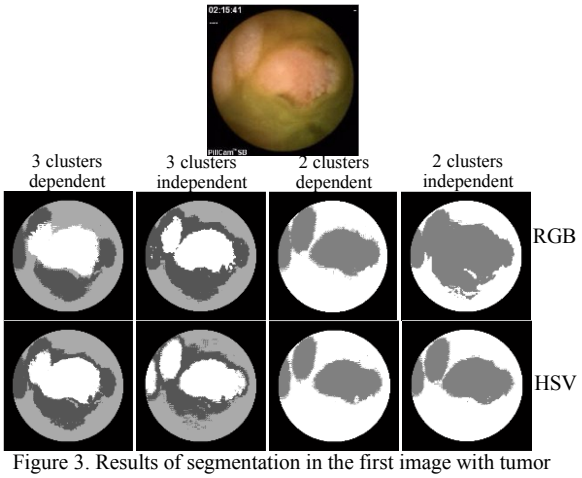

Figure 3. Results of segmentation in the first image with tumor tissue in the small bowel taken from CE.

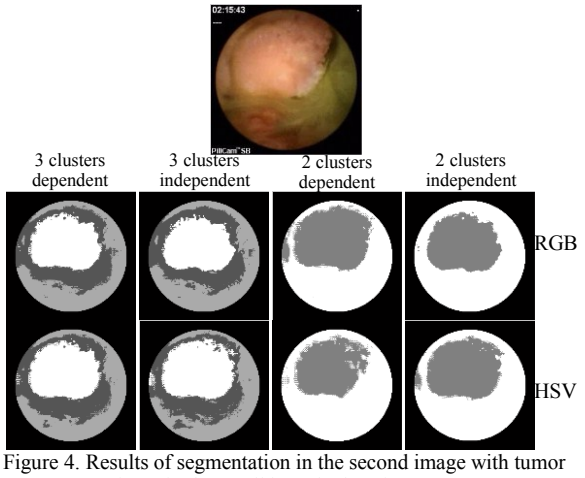

Figure 4. Results of segmentation in the second image with tumor tissue in the small bowel taken from CE.

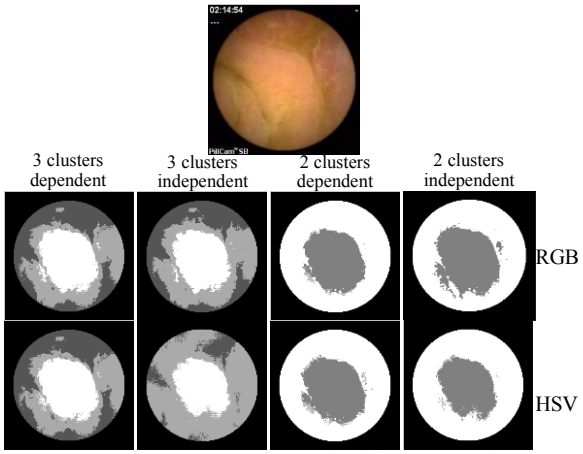

Figure 5. Results of segmentation in the third image with tumor tissue in the small bowel taken from CE.

IV. Conclusion

Current algorithms for processing the small bowel images taken from capsule endoscopy usually do not use the ROI paradigm since the image is processed as a whole. However selected applications such as training of medical staff among others need robustly segmented images. Robustness regarding grey level images is usually obtained by using the MRF concept, which demands a high degree of computational load and hence not adequate for applications generating a great amount of images. This paper proposes changing context information with color information to improve segmentation robustness and was successfully applied in the detection of small bowel abnormal tissue. In this regard HSV color space seems to be more useful than the RGB color space. The algorithm needs to be tested in other pathologies in order to measure the generalization degree. Applications such as three dimensional reconstruction of the wall intestine and detection of small structures such as polyps and ulcers rely heavily on the segmentation process.

References

[1] V. S. Kodogiannis, M. Boulougourab, E. Wadge and J.N. Lygourasc, "The usage of soft-computing methodologies in interpreting capsule endoscopy," *Engineering Applications of Artificial Intelligence,* vol. 20, pp. 539–553, 2007.

[2] D. Barbosa, J. H. Correia, J. Ramos and C. S. Lima, Texture Classification of Images from Endoscopic Capsule by Using MLP and SVM – A Comparative Approach" Proceedings of the World Congress on Medical Physics and Biomedical Engineering 2009; Munich 7-12 September 2009, Vol. 5 pp. 271-274.

[3] Alexandros Karargyris and Nikolaos Bourbakis (2011). Three-dimensional reconstruction of the digestive wall in capsule endoscopy videos using elastic video interpolation. In IEEE transactions on Medical Imaging Volume 30, Number 4, 957-971.

[4] Alexandros Karargyris and Nikolaos Bourbakis (2011). Detection of small bowel polyps and ulcers in wireless capsule endoscopy videos. In IEEE transactions on Biomedical Engineering Volume 58, Number 10, 2777-2786.

[5] Michael Kass, Andrew Witkin and Demetri Terzopoulos (1987). Snake: Active contour models. In Proc. First Int. Conf. Computer Vision, London pp. 259-268.

[6] Chenyang Xu and Jerry L. Prince (1998). Generalized gradient vector flow external forces for active contours. Signal Processing 71, 131-139.

[7] Ravikanth Malladi, James A. Sethian and Baba C. Vemuri (1996). A fast level set based algorithm for topology-independent shape modeling. Journal of mathematical Imaging & Vision, vol. 6, no 2/3 pp. 269-290.

[8] Amit Chakraborty (1996). Feature and module integration for image segmentation. Ph. D. thesis, Yale Univ.

[9] Kapur T. (1999). Model based three dimensional medical image segmentation. Ph. D. thesis, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

[10] Koen Van Leemput, Frederik Maes, Dirk Vandermeulen, Alan Colchester and Paul Suetens. (2001). Automated segmentation of multiple sclerosis lesions by model outlier detection. IEEE transactions on medical imaging, vol. 20, nº8, August 2001.

[11] J. Marin, K. Mengersen and C. Robert, "Bayesian Modelling and Inference on Mixtures of Distributions," in *Handbook of Statistics 25*.

[12] David J.C. MacKay. (2003). Information Theory, Inference, and Learning Algorithms. Cambridge University Press.

[13] J. Bilmes, A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models, Berkeley: U.C. Berkeley, 1998.

[14] A. Gut, An Intermediate Course in Probability, 2009.

[15] S. Borman, The Expectation Maximization Algorithm: A short tutorial, 2004.

[16] M. Tkalcic and J. F. Tasic, "Colour spaces: Perceptual, historical and applicational background," in Proc. IEEE Region 8 Comput. as Tool. EUROCON, Sep. 22–24, 2003, vol. 1, pp. 304–308.