

A COMPLETE SYSTEM FOR CANDIDATE POLYPS DETECTION IN VIRTUAL COLONOSCOPY

MARCELO FIORI* and PABLO MUSÉ†

Facultad de Ingeniería, Universidad de la República, Uruguay

**mfiori@fing.edu.uy*

†pmuse@fing.edu.uy

GUILLERMO SAPIRO

Duke University, NC, USA

guillermo.sapiro@duke.edu

Received 14 February 2014

Accepted 18 July 2014

Published 12 September 2014

We present a computer-aided detection pipeline for polyp detection in Computer tomographic colonography. The first stage of the pipeline consists of a simple colon segmentation technique that enhances polyps, which is followed by an adaptive-scale candidate polyp delineation, in order to capture the appropriate polyp size. In the last step, candidates are classified based on new texture and geometric features that consider both the information in the candidate polyp location and its immediate surrounding area. The system is tested with ground truth data, including flat and small polyps which are hard to detect even with optical colonoscopy. We achieve 100% sensitivity for polyps larger than 6 mm in size with just 0.9 false positives per case, and 93% sensitivity with 2.8 false positives per case for polyps larger than 3 mm in size.

Keywords: Computed tomographic colonography; computer-aided detection; colonic polyp detection; colon segmentation; curvature motion; differential features.

1. Introduction

Colorectal cancer is nowadays the second leading cause of cancer-related deaths in the United States (only surpassed by lung cancer), and the third cause worldwide.²⁸ The early detection of polyps is fundamental, allowing to reduce mortality rates up to 90%. Polyps can be classified according to their morphology: Pedunculated polyps are attached to the colon wall by a stalk, sessile polyps grow directly from the wall, and flat polyps have less than 3 mm of elevation above the colonic mucosa. Nowadays, optical colonoscopy (OC) is the most used detection method due in part

*Corresponding author.

to its high detection rate. However, this technique is invasive and expensive, making it hard to use in large screening campaigns.

Computed Tomographic Colonography (CTC), or Virtual Colonoscopy (VC), is a promising alternative technique that emerged in the 90's.²⁵ It uses volumetric Computed Tomographic data of the cleansed and air-distended colon. It is less invasive than OC, and much more suitable for screening campaigns once its performance is demonstrated. However, VC is less popular than OC not only because it is a relatively new technique, but also because it is not yet reimbursed by insurance companies. On the other hand, in OC, incomplete studies due to obstructing lesions, colon twists, or anatomical variations are not rare² (5% to 15% of OC examinations) and there is an additional important risk of colon perforation.

Nevertheless, it takes more than 15 min for a trained radiologist to complete a CTC study, and the overall performance of OC is still considered better. In this regard, Computer-Aided Detection (CAD) algorithms can play a key role, assisting the expert to both reduce the procedure time and improve its accuracy.²¹

Flat polyps are of special interest because these are an important source of false negatives in CTC, and although there are different opinions, many authors^{8,27} claim that flat polyps are around 10 times more likely to contain high-grade epithelial dysplasia, while Bond³ states that the major limitation of VC is its current low performance for flat polyps. There are numerous discussions regarding the potential risks of the polyps according to their size. Even though some authors consider that “small” polyps may not represent risk, some gastroenterologists disagree.¹ Summers²⁰ claims that one of the major challenges in the field is in increasing sensitivity for smaller polyps, and Church⁶ states that small adenomas can still be clinically significant and should not be ignored.

The goal of the work presented in this paper is to exploit VC precisely to automatically flag (mark for attention of the expert) colon regions with high probability of being polyps, with special attention to results in challenging small and flat polyps. It is crucial to minimize the false negatives, keeping a reasonable false positives number. We achieve this by an automatic four-steps process that constitutes the entire end-to-end algorithm, from data to candidate polyps flagging.

The proposed system with its four steps is illustrated in Fig. 1. The first step is colon segmentation, which takes as input the computed tomography (CT) volume

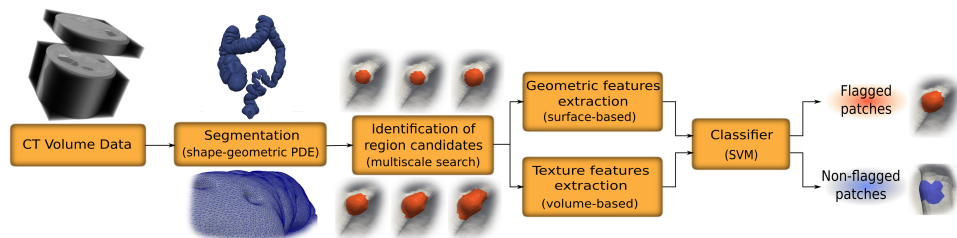


Fig. 1. Basic pipeline of the proposed polyp flagging system.

data, and produces a 3D mesh representing the colon surface. The segmentation technique proposed here is geared toward the subsequent step of polyp detection, and simultaneously segments and prepares the obtained surface for this task. In the second step, from the segmented mesh we perform an adaptive-scale search of candidates in order to capture the appropriate polyp size, obtaining a set of candidate patches. Then, geometrical and textural features are computed for each candidate patch that was identified in the previous stage. The final step consists of a machine learning algorithm that uses the computed features to classify patches as polyps or normal tissue.

The whole algorithm is completely automatic and produces state-of-the-art results. This paper extends our previous conference publications.^{9,11}

The rest of this paper is organized as follows. We address the colon segmentation problem in Sec. 2 and the feature extraction and classification in Sec. 3. In Sec. 4, we describe the classification, and in Sec. 5 we present numerical results. The discussion is presented in Sec. 6 and we conclude in Sec. 7.

1.1. Virtual colonoscopy CAD review

Automatic polyp detection is a very challenging problem, not only because the polyps can have different shapes and sizes, but also because they can be located in very different surroundings. Most of the previous work on CAD of colonic polyps is based on geometric features, some using additional CT image density information, but none of them takes into account the (geometric and texture) information of the tissues *surrounding* the polyp. This local and adaptive differential analysis is part of the contributions of this work.

Early work on CAD methods by Vining *et al.*²⁴ is based on the detection of abnormal wall thickness. Since then, several different approaches were proposed. Most of them have a segmentation step first, and then the classification step itself.

The most common segmentation techniques are based on region growing and thresholding methods (often followed by a smoothing stage), or level set methods, which have the smoothing incorporated as a velocity term in the evolution. Successful examples of these approaches are in Refs. 5, 12, 17, 21 and 29. In general, all interfaces are diffuse due to the partial volume effect, and special care must be taken with the air–fluid–tissue T-junction, since artifacts here generated are a common source of false positives. On the other hand, not much work has been done in comparing the smoothing techniques (or the regularization in the level set method), to see which one is more adapted to polyp detection. The main variations in the detection stage are in the features used and in the classification method. The most discriminant features are the geometric ones, and in particular curvature-based measures have been proved successful, see Refs. 17, 21, 23 and 29. All these techniques based on local geometric computations suffer from a high dependence on the regularity of the polyp shape itself, ignoring how pronounced it is with respect to the surrounding area. Using only geometry is also very sensitive to the accuracy of

the segmentation. Texture features have been also used with promising results in Refs. 19 and 26. To the best of our knowledge, no algorithm reported in the literature can detect small polyps properly. On the other hand, for polyps larger than 6 mm in size, no algorithm can achieve 100% sensitivity with less than one false positive per case.

2. Colon Segmentation

The segmentation of the colon surface, which is critical in particular to compute geometric features, is divided into two parts: a pre-processing stage for dealing with the air–fluid composition of the colon volume, and a second stage that consists on smoothing the pre-processed image and obtaining the final colon surface by thresholding the smoothed volume. The overall procedure here presented is very simple and computationally efficient. More details are available in Ref. 10.

2.1. Classifying CT regions

All the cases from the used database have the same preparation, which includes solid–stool tagging and opacification of luminal fluid (white liquid in Fig. 2). One of the strongest difficulties concerning the segmentation of the colon is the presence of this tagged fluid and its interfaces with air and tissue. Figure 2 shows a CT slice and its pixel values over the highlighted vertical profile. At first sight there are three clearly distinguishable classes: lowest gray levels correspond to air, highest levels to fluid, and the middle gray values correspond to tissue. Nevertheless, there are around six interface voxels between air and fluid whose gray values lie within the normal tissue range. Therefore, a naïve approach, ignoring the physical nature of the tissue and its environment, is not suitable for proper tissue classification and segmentation. Also, with a binary segmentation approach the border will be necessarily bumpy. We propose to compute a function u_0 intended to have homogeneous values in the colon interior and exterior, and a smooth transition between them.

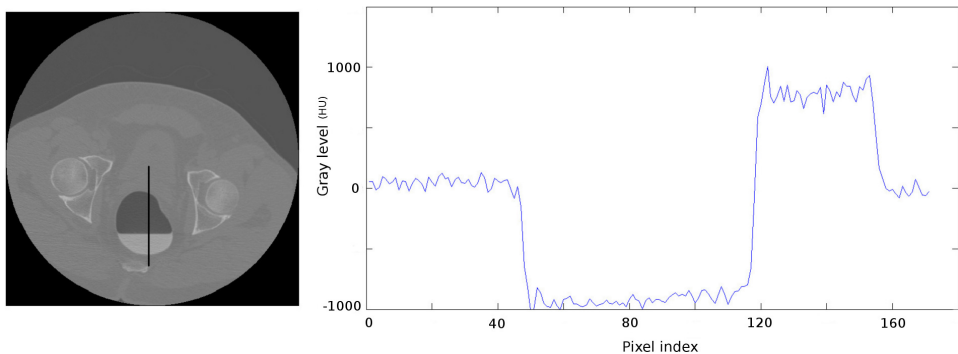


Fig. 2. CT slice and its gray values for air, fluid and normal tissue, along the vertical profile.

In order to do that, it makes sense to assign to each voxel the likelihood of being air, fluid, or air–fluid interface. The air and fluid distributions are estimated using standard kernel density estimation techniques, and these functions are then used to assign the air and tagged fluid likelihood values to the voxels.

Note that this assignment fails on the air–fluid and air–fluid–tissue interfaces. For assigning a value to these voxels, we take advantage of the physics of the problem: The subject is laid horizontally so the interface between the fluid and the air is a plane parallel to the floor. Then, the voxels situated on the interface have a large gradient in the vertical direction.

The implementation of these criteria is as follows. A cubic neighborhood around each voxel \mathbf{x} is considered, and for each “column” that results of fixing the x and y coordinates, the air-likelihoods of the upper voxels and the fluid-likelihoods of the lower voxels are accumulated. The value $IC(\mathbf{x})$ that represents the confidence level of \mathbf{x} being an interface voxel is an increasing function of the accumulated measure. The algorithm below provides a pseudo-code that represents this procedure.

Algorithm 1.

```

for each voxel  $(x,y,z)$  do
    sum=0 for  $i=-1$  to 1 do
        for  $j=-1$  to 1 do
            for  $k=1$  to 2 do
                sum +=  $p((x+i,y+j,z+k)|w_1)$ ;
                sum +=  $p((x+i,y+j,z-k)|w_2)$ ;
            end
        end
    end
     $IC(x,y,z) = \text{sum}/18$ ;
end

```

In order to guarantee that a high value is associated to every interior voxel, we assign to the initial segmentation u_0 the maximum of these three values, namely, the air and fluid likelihoods and the interface confidence level.

After the computation of the initial segmentation u_0 , some spurious (isolated) voxels may have high values (bones for example), so we clean the initial segmentation by keeping the connected components^a containing some chosen voxels used as seeds. These seeds are automatically detected by choosing the voxels with largest values of $IC(\mathbf{x})$, since these high values only occur at the interface between air and fluid. This way, the segmentation step is able to handle the lumen discontinuities problem and to obtain the multiple pieces of the colon that might be disconnected.

The air–fluid–tissue joint may create artifacts in the segmentation. This is a critical point, not only because of the quality of the segmentation, but mainly for the

^aActually, since the initial segmentation u_0 is not binary, a (conservative) threshold of 0.6 is considered to separate the connected components.

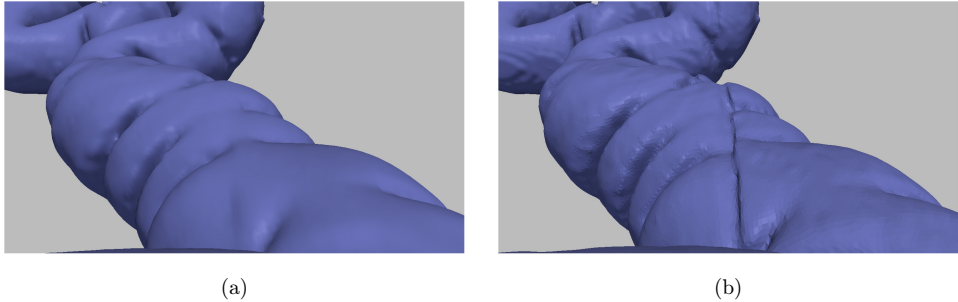


Fig. 3. Comparison of reduced artifacts in our segmentation (a) with a previously tested more standard version (b). The initial segmentation stage, computing the maximum of the air and fluid likelihoods and the interface confidence level, successfully deals with the T-junction problems.

potential of yielding several FPs in the polyps detection step. It is not rare that segmentation algorithms result in “gutter-like” shapes along this interface. Then, if small oscillations occur along the “gutter” (which is expectable due to the CT resolution), artifacts with polyp-like shape are produced, thus degrading the performance of the whole CAD system. We paid particular attention to this issue while designing the segmentation algorithm: the IC computation allows to avoid these artifacts. Figure 3 illustrates the performance of our segmentation method compared to a version of the algorithm without the computation of the IC, clearly presenting some problems along this interface. This helps to detect polyps near the air-fluid interface, and minimizes the FPs resulting from to this artifact.

2.2. Smoothing and colon surface computation

In order to eliminate noise and to obtain a smoother colon surface after the segmentation stage, we proceed to smooth the initial segmentation u_0 . We derive a Partial Differential Equation (PDE) driven smoothing technique that preserves the shape of the polyps, while obtaining a smooth enough surface to reliably compute local geometric features such as curvatures. Of course, the ultimate goal of the method is to simplify and to improve the polyp/nonpolyp classification system. The effectiveness of the proposed approach will be assessed with experiments both qualitatively and quantitatively in Sec. 5, where ROC curves obtained with the proposed PDE and other smoothing alternatives will be compared.

We concentrate on a family of smoothing PDEs of the form

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} = \beta |\nabla u|, \quad u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad (1)$$

where the initial volumetric image u_0 results from the preprocessing described in the previous section. A classical time discretization scheme is used to implement this evolution. After a few iterations of this evolution, the inner colonic wall will be extracted as a suitable iso-level surface of the resulting 3D image $u(\mathbf{x}, T)$.

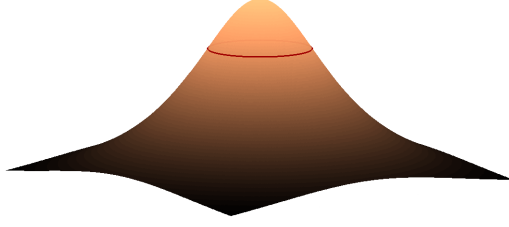


Fig. 4. Polyp with a curve of inflection points (in red), where $\kappa_{\min} = 0$ (color online).

We recall that the Level Set Method¹⁶ states that if $u(\mathbf{x}, t)$ evolves according to (1), then its iso-levels (level sets) satisfy

$$\frac{\partial \mathcal{S}}{\partial t} = \beta \mathcal{N}, \quad (2)$$

where \mathcal{S} is any iso-level surface and \mathcal{N} its unit normal. This geometric view enables to design β to fulfill a set of requirements we will impose to the surface evolution. In particular, we are interested in motions driven by the principal curvatures.

With the mean curvature motion ($\beta = \mathcal{H}$), and the affine motion ($((K^+)^{1/4})$), the polyyps are flattened too fast.¹⁰

As an alternative, a classical motion that appears to be well suited for our problem is the motion by minimal curvature.⁴ Indeed, polyyps have a curve of inflection points all around it, separating its upper and lower sections (see Fig. 4). Along this curve, the minimal curvature is $\kappa_{\min} = 0$, and therefore this part of the polyp does not move (or moves very slowly), so intuitively under this motion the polyyps should persist longer. This PDE already yields very good results in terms of both surface smoothing and polyp enhancement.

We further derive two modifications that lead us to the proposed smoothing PDE, showing qualitative results to support this claim. The first modification is inspired by the exponent $\frac{1}{4}$ of the affine motions in dimension 3:

$$\frac{\partial \mathcal{S}}{\partial t} = \kappa_{\min}^{1/4} \mathcal{N}. \quad (3)$$

Figure 5 shows the result after a few iterations, and Fig. 6 evidences the difference between the motions by κ_{\min} and $\kappa_{\min}^{1/4}$ (in gray and in orange, respectively) with a comparative image. On the polyp protrusion, the orange surface is above the gray surface, while the opposite is observed in the surrounding area. This shows that the evolution by $\kappa_{\min}^{1/4}$ leads to better polyp enhancement.

The second modification is based on the idea of preserving the polyyps qualities that we later use to identify them. A measure of the local shape of a surface is the so-called *shape index* (SI),¹⁵ and the complementary *curvedness* C :

$$\text{SI} := -\frac{2}{\pi} \arctan \left(\frac{\kappa_{\max} + \kappa_{\min}}{\kappa_{\max} - \kappa_{\min}} \right), \quad C := \frac{2}{\pi} \ln \sqrt{\frac{\kappa_{\max}^2 + \kappa_{\min}^2}{2}}.$$

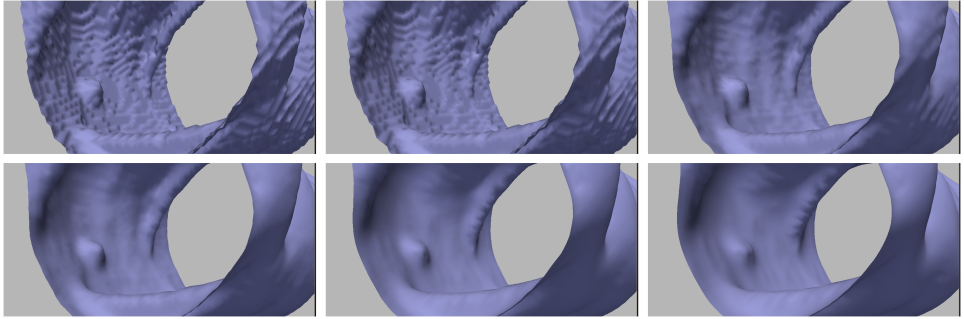


Fig. 5. Evolution by $\kappa_{\min}^{1/4}$: original surface and the result after 2, 8, 15, 30 and 50 iterations.

While the value of SI is scale-invariant and measures the local shape of the surface, the value of C indicates how pronounced it is.

Back to the PDE motion, we now include this information in order to make potential polyps evolve differently than the rest of the colon surface. More precisely, we modify the best motion so far ($\beta = \kappa_{\min}^{1/4}$), in such a way that the resulting motion further enhances the potential polyps. In order to achieve this, we first need to characterize the potentially polyp points, and then modify the deformation function accordingly.

We define a function of the SI that acts as a multiplying factor to the term $\kappa_{\min}^{1/4}$, making the surface evolve slower at the interest points. These function should assign low values to SI near -1 , and values close to unity to other points. A smooth function $g(\text{SI})$ verifying these constraints is $g(\text{SI}) = \frac{1}{\pi} \arctan((\text{SI} - 0.75) \cdot 10) + \frac{1}{2}$.

The final evolution then becomes

$$\frac{\partial \mathcal{S}}{\partial t} = g(\text{SI}) \kappa_{\min}^{1/4} \mathcal{N}. \quad (4)$$

This motion keeps all the advantages of the motion by $\kappa_{\min}^{1/4}$ and in addition, polyps are flattened more slowly, so at the end the obtained surface is smooth and the polyps still stand out.

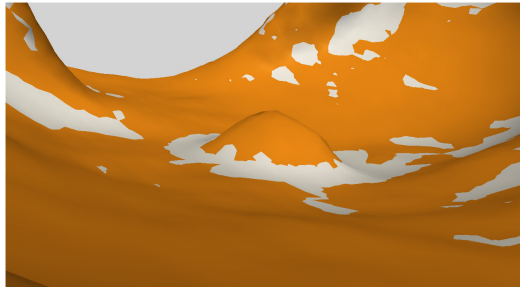


Fig. 6. Comparison between evolutions. Motion by k_{\min} in light gray versus motion by $k_{\min}^{1/4}$ in dark gray. Both surfaces are overlaid, so sections that are not visible are hidden below the other surface (color online).

The number of iterations can be set by choosing the value that maximizes the overall performance of the system, in terms of the free-response ROC curve (FROC). This can be done by trying with several values and keeping the one which maximizes this performance. Alternatively, we can consider a sphere of the size of the CT resolution and compute analytically the number of iterations needed to make it vanish (see Appendix A). The idea behind this procedure is to smooth the surface up to the resolution limit. These two approaches led to the same result, namely 15 iterations, and therefore this is the chosen value for the experiments in this paper.

At this point, after choosing the appropriate diffusion and the number of iterations, we have a smoothed function $u(\mathbf{x}, T)$ whose level sets indicate the region inside of the colon. We then extract the surface of the colon as the iso-value surface of level $\alpha \in [0, 1]$. The choice of the value α can be made by maximizing some criteria, in order to obtain the most contrasted surface in a given sense. This optimization-oriented method was tested, and we observed that in our particular application all the consistent surfaces are very close to each other. Therefore, the computational effort is not justified and we kept the iso-surface $\alpha = 0.7$. Note that this choice can be safely made once for all the data. The result of this stage is then a triangulated surface \mathcal{S} representing the colon wall.

3. Polyp Delineation and Feature Extraction

All the polyp detection methods reported in the literature try to detect or classify the polyps from properties defined only within the candidate region. However, it is important to analyze the spatial context in which the candidate patch is located, not only because different sections of the colon present different characteristics, but also because polyps can be situated over different structures such as folds or plain colonic wall. A good feature including the shape of the neighborhood for example, can help in the discrimination between irregular folds and polyps over folds. In addition, looking for significant differences in the gray level imitates the human-based inspection, which highlights zones that contrast with their vicinity.

In this regard, most of the features described in this section take into account the local information of the area surrounding the candidate patch. This makes the features more robust to the particular local phenomena. The normal tissue of different cases may vary (due to different biological properties of the subjects or to different conditions of the studies), so absolute thresholds in texture features lack meaning; while texture patterns differ from study to study, what does not vary is the fact that polyps have different properties than normal tissue.

3.1. Candidate detection and geometrical features

The starting point for the geometric features described in this section is the segmented surface \mathcal{S} . Let us consider the SI as a function $SI : \mathcal{S} \rightarrow [-1, 1]$, and recall that the polyps have SI values close to -1 . Therefore, it is expected that a region

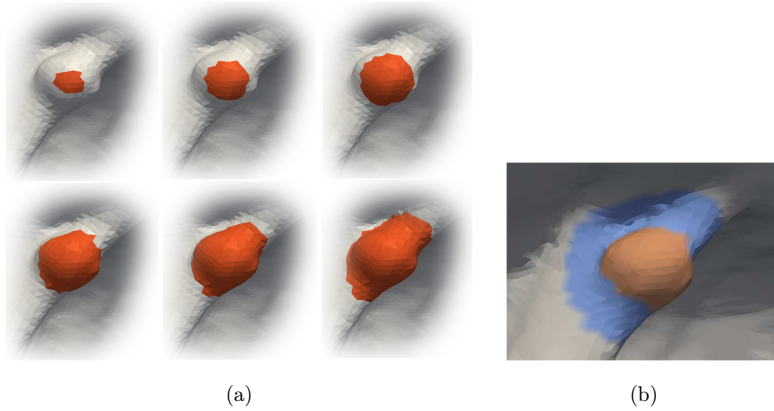


Fig. 7. Patch size selection and ring around polyp (a) Sets $\mathcal{P}_1 \dots \mathcal{P}_n$: Different sizes are tested in order to select the most appropriate patch and (b) Ring (in blue) surrounding a candidate polyp (in orange) (color online).

(patch) of the surface that corresponds to a polyp contains at least one local minimum of this function. The detection of the candidate patches follows an adaptive-scale search: For each local minimum $x_0 \in \mathcal{S}$ of the function SI, several level sets of SI ($\mathcal{P}_1 \dots \mathcal{P}_n$) around x_0 are tested, and the level set \mathcal{P}_i that maximizes the distances between the histograms described below, is the considered candidate patch, which we simply denote by \mathcal{P} (Fig. 7(a)). A total of $n = 7$ level sets are tested, corresponding to the SI values from -0.8 to -0.5 with a 0.05 step. The following description is given for the final chosen patch \mathcal{P} , but the ring and histogram computations are made for all the level sets \mathcal{P}_i in order to select the most appropriate of them.

Given a candidate patch \mathcal{P} , a ring \mathcal{R} around \mathcal{P} is computed, in order to consider geometrical measurements with respect to the area surrounding the patch. The ring is calculated by dilating the patch \mathcal{P} a certain geodesic distance, such that the areas of \mathcal{P} and \mathcal{R} are equal. Figure 7(b) shows a candidate patch (actually a true polyp), and its corresponding ring.

Histograms of the SI values are then computed for the patch \mathcal{P} and the ring \mathcal{R} , and two different distances between them are computed: the L_1 distance and the symmetric Kullback–Leibler divergence. If the patch corresponds to a polyp-like shape then the values of the histogram \mathcal{P} will be concentrated around the -1 extrema, on the other hand, the histogram \mathcal{R} will be inclined to the other extreme in case of a polyp on a normal colon wall (concave), or with tendency to values near -0.5 if the polyp is on a fold. These two features give a measure of the geometric local variation of the candidate patch \mathcal{P} . Although these two distances are the most discriminative features,^b we also consider the following additional ones since they

^bSeveral feature selection techniques confirmed this. For instance, the following methods available in *Weka*¹³: *Information Gain*, *Gain Ratio*, and *Relief-F*. Using the data described below, these techniques sort the features according to their discriminative power.

still help to discriminate some typical false positives:

- The mean value of the SI over the patch \mathcal{P} ,
- The area of the patch, since the target polyps are in a certain range of size.
- The growth rate of the areas at the adaptive-size stage, meaning the ratio between the area of the chosen patch $\mathcal{P} = \mathcal{P}_i$ and the area of the immediately smaller patch \mathcal{P}_{i-1} ; this feature measures how fast the shape of the patch is changing.
- And finally the *shape factor*,

$$SF = \frac{4\pi \cdot \text{Area}}{\text{Perimeter}^2},$$

which measures how efficiently the perimeter is used in order to gain area,^c and it favors circle-like patches (like the polyp patch in Fig. 7(b)), avoiding elongated patches (like the false positives in folds).

Therefore, we have a total of six geometric features.

3.2. Texture features

There is evidence that the gray-level of the CT image and its texture can be very helpful for detecting polyps. This is in particular useful for flat or small polyps, where the geometric information is limited. Some work has been done on the inclusion of texture features (inside the candidate polyps only), in order to reduce false positives.²⁶ According to the reported results, there is still room for improvement. We propose both the use of new texture features and the inclusion of the information on the candidate's surrounding area.

First, for each polyp candidate $\mathcal{P} \subset \mathcal{S}$, a region V_1 is calculated, containing the patch \mathcal{P} and a portion of the inner tissue next to the patch. The region V_1 is obtained by dilating (in 3D) the patch \mathcal{P} towards the inner colon tissue (we discard the air or fluid voxels). A second region V_2 surrounding V_1 is calculated by dilating V_1 . Volume V_2 is intended to contain normal tissue in order to compare it with the polyp candidate tissue. In order to choose how much dilation to perform, we use a technique similar to the one in the previous section: several dilation distances are tested, and we keep the distance that makes the differential features most discriminative.

The features chosen are a subset of the classical Haralick¹⁴ texture features, namely, entropy, energy, contrast, sumMean, and homogeneity. Seven co-occurrence matrices are calculated with the voxels of V_1 , and all the five features are averaged over the seven directions. The analogous computation is made for V_2 , and the differences between the two 3D regions, for each texture feature, are considered. Additionally, the mean gray levels of the voxels in both regions is computed, and their difference is considered as a feature. In this way, six texture features are considered. This approach for computing the texture features, measuring differences with the

^cThe maximum value for the shape factor is 1 and it is achieved only by the circle.

surrounding area, leads to better discrimination than the features computed just for V_1 , as demonstrated next.

4. Classification

Once the the candidates detection has been performed with the adaptive-scale approach, the number of true polyps was much lower than the number of nonpolyps patches, a relation on the order of 500:1, which is a significant problem for the learning stage of the classifier, since most classifiers are designed to maximize the accuracy,⁷ and this measure might not be adequate for imbalanced problems. For instance, if we classify all candidates as “nonpolyps,” we would get an accuracy of 99.8% but without detecting any polyps. Three techniques were considered to overcome this shortcoming.

The MetaCost approach consists of combining several instances of the classifier instead of modifying the proportion of classes in the training data according to the costs. This method does not work with “stable” classifiers (those that produce similar models with slightly different training sets) like support vector machines (SVM).

The Cost Sensitive Learning approach, unlike the MetaCost, tries to balance the classes before the learning stage. The implementation we used from *Weka*¹³ simply takes as input the re-balance parameters (cost matrix) and replicates instances of the minority class. One of the advantages of this approach is that no assumptions are made about the behavior of the classifiers (unlike the previous method) nor the distribution of the data (unlike the next method).

Finally, the Synthetic Minority Over-sampling Technique (SMOTE) is a method to generate artificial instances of the minority class, in order to get a balanced data to learn from. The new artificial instances are created as a convex combination of the existing instances of the minority class.

We tested all these options and the best results were obtained using Cost Sensitive Learning with SVM. The pipeline is then: pre-processing stage as in Sec. 2.1, smoothing according to (4), the 12 features described in Sec. 4, and classification using Cost Sensitive Learning with SVM.

5. Results

A total of 150 patients of the Walter Reed Army Medical Center (WRAMC) database¹⁸ were used to test the proposed CAD algorithm. Most of these patients have two sets of CT images, one for supine and one for prone position. Taking precautions not to train the classifier with, for instance, the prone images set and test it with the supine set (i.e. one cannot use a supine study of a given patient for training, and the corresponding prone study for testing, or vice versa), we can consider the 300 images sets as independent. From now on, we refer to each of these 150 images sets as a *case*. The database contains 134 polyps detected by optical colonoscopy, including 12 flat polyps. Among these 134 polyps, 86 are larger than 6 mm in size, and the other 48 are between 3 and 6 mm in size. Figure 8 shows the distribution of polyps’ sizes in the

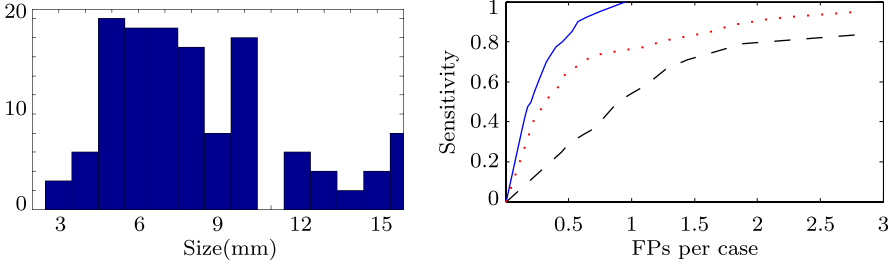


Fig. 8. Histogram of the polyps' sizes (left) and FROC curve (right) of the proposed system for different polyps sizes: larger than 6 mm (solid), smaller than 6 mm (dashed), and all polyps (dotted).

database. The size and shape classification of these polyps was taken from the WRAMC database description. These descriptions were provided by the physicians involved in the OC examinations. Taking these examinations as the ground truth, patches classified as polyp were considered TP if the distance to a ground truth polyp was less than 3 mm.

The evaluation was carried out by splitting the dataset into two halves, training and testing. Under this setting, and classifying with SVM + Cost Sensitive Learning, we obtained the FROC in Fig. 8, which shows the performance for different polyps sizes. Again, the work with such small, as well as flat polyps, is unique to the framework here presented, as will be discussed in Sec. 6.1. About 40% of the polyps were covered by tagging, but the classification results do not vary depending on this fact: the performance is the same for covered and for noncovered polyps.

These values are comparable with state-of-the-art results,^{26,22} but our database includes very small polyps. A more precise comparison of results is not necessarily meaningful, since in general each work considers its own database.

The FROC curve in Fig. 9 compares the performance of the system with the different smoothing methods discussed in Sec. 2 (the rest of the pipeline is

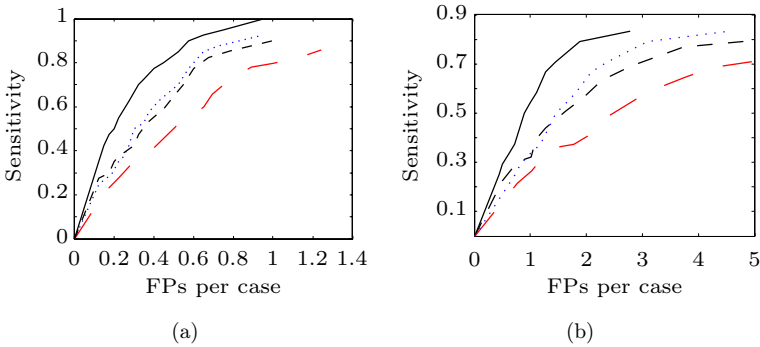


Fig. 9. FROC curve comparing the performances using the different smoothing methods, classifying large polyps (a) and small polyps (b). The curve for the proposed evolution is shown in solid line, the results for the evolution by the mean curvature \mathcal{H} and κ_{\min} are shown in dotted and dashed lines, respectively, and the lower curve is the result when no smoothing is performed.

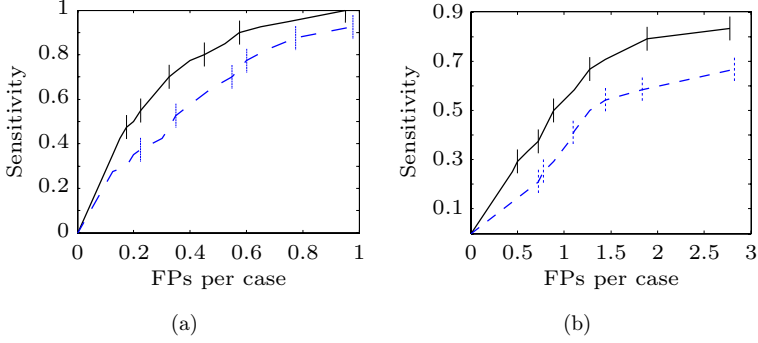


Fig. 10. FROC curve with 95% confidence intervals, comparing the performances with differential (solid) and absolute (dashed) texture features, classifying polyps larger than 6 mm in size (a) and smaller than 6 mm in size (b).

unchanged). The proposed smoothing technique achieves better results than the other discussed methods.

The FROC curve in Fig. 10 shows the comparison between absolute and differential texture features. The classification was performed using all the geometric features and either the absolute texture features (computed just for V_1), or the differential texture features. The results show that, when combined with the differential geometric features, differential texture features are more discriminative than the absolute ones. This fact was also confirmed using several feature selection techniques available in Weka¹³ like *Information Gain*, *Gain Ratio*, and *Relief-F*. All these methods sort the features according to their discriminative power, and in all of them the differential features ranked better than absolute ones.

Finally the FROC curve in Fig. 11(a) compares the results of the different classification approaches. Cost Sensitive, SMOTE, and MetaCost were used as a pre-processing stage for SVM, AdaBoost was used with C4.5 trees. The parameters in

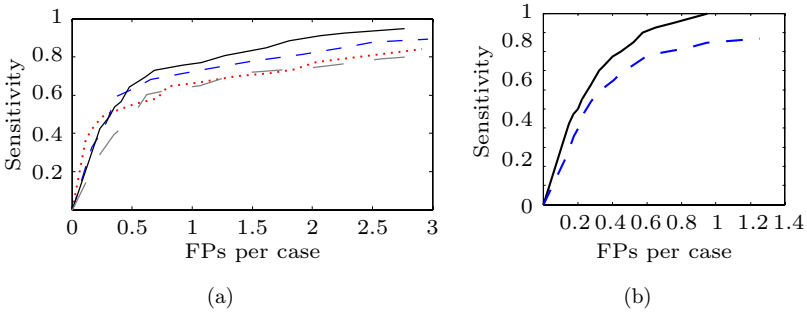


Fig. 11. (a) FROC curve comparing the performances of different classification approaches for all polyps. SVM with cost Sensitive (solid), SVM with SMOTE (dashed), C4.5 trees with AdaBoost (dotted) and plain SVM (long-dashed). (b) FROC curve of the final pipeline (solid) and without the IC computation (see Sec. 2.1).

all classifiers were optimized via cross-validation. Also, the FROC curve in Fig. 11(b) shows the effect of the elimination of the gutter artifact with the computation of IC described in Sec. 2.1.

6. Discussion

6.1. Small, big and flat polyps

It is clear that both the small and flat polyps are much more difficult to detect than the other polyps. What is not clear is if the same kind of algorithm and features are suitable for detecting all the range of polyp types and sizes. We showed that the proposed combination of features, although it might not be optimal for every specific type of lesion, is able to correctly detect all of them.

Both the segmentation and the features considered here, contribute to the good classification results for the whole database. The 93% sensitivity together with the 2.8 FP rate for polyps larger than 3 mm in size is as remarkable as the 0.9 FP rate for polyps beyond 6 mm in size with 100% detection.

6.2. Geometric and texture importance

Although geometrical features are the most discriminative ones (see Table 1), texture features still play a fundamental role in the classification. Adding the texture features to the geometric ones, the sensitivity increases from 88% to 93%, and at the same time the false positives rate decreases by 30%.

Figure 12(a) shows a detected polyp, where geometry is crucial, because the gray-level does not present considerable local variations. This is specially true for polyps

Table 1. Comparison of performance using only geometric versus only texture features.

Polyps > 3 mm	Features		
	All	Geometric	Texture
Sensitivity	93%	88%	68%
FPS p/case	2.8	6.5	19

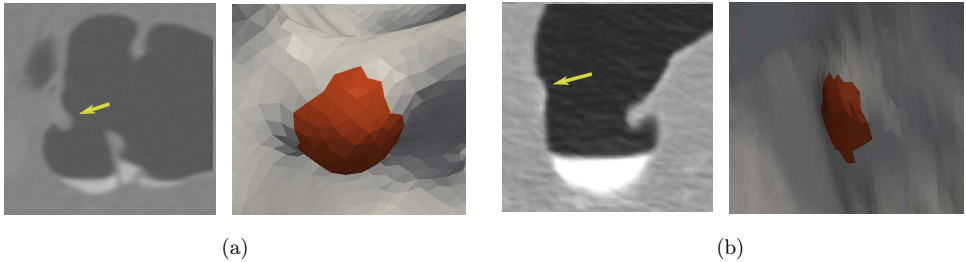


Fig. 12. (a) Polyp with no texture information. (b) Polyp with texture information, but weak geometric information.

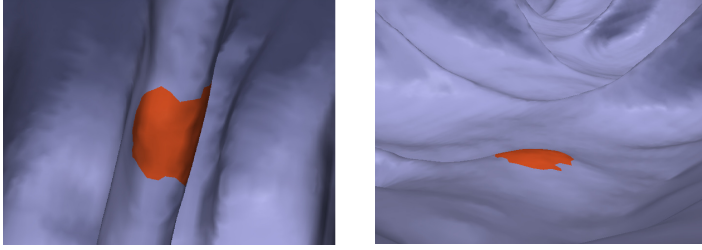


Fig. 13. False positives: Fold and patch similar to polyp.

submerged in tagged material. On the other hand, in the flat polyp of Fig. 12(b), the geometry is weakly discriminative (although the measure considering the ring enhances the detectability), and texture features lead to a correct classification. Texture information is essential also because it is more robust to segmentation errors, as texture features are computed by integrating over the volumetric data.

6.3. Qualitative analysis of false positives

In addition to the number of false positives, it is very important to study how these FP patches look like, since some of them can be quickly ruled out by the expert and some can be avoided by improving some aspects of the segmentation.

About half of the false positives are quite reasonable, in the sense that they are (usually small) sections of the colon that are polyp-like shaped (see Fig. 13), specially

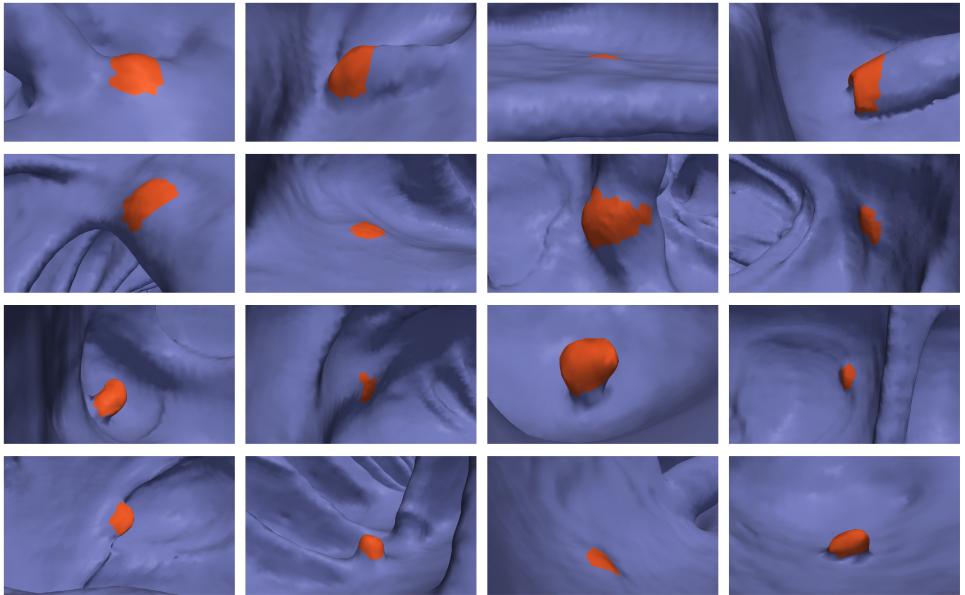


Fig. 14. Examples of false positives according to the available labeled data, with some segmentation errors, parts of the insufflation tube, and some patches with polyp-like shape.

taking into account that we designed the system to also detect small and flat polyps. Among these false positives, 40% are very small patches that may be avoided by incorporating some new features to the classification, or by adding a size threshold if very small polyps are not considered of interest (which was not our case). On the other hand, about 10% of the FPs occurred in fold sections of the wall (Fig. 13) and another 10% occurred in parts of the insufflation tube. All these patches (from folds and the insufflation tube), are easily ruled out by visual inspection. Another 20% of the FPs were caused by colon segmentation errors. About half of these false positives due to segmentation errors are caused by bad quality original CT slices in the region (generally due to the partial volume effect), like the bottom-left example in Fig. 14. Another fraction of FP due to segmentation are protuberances caused by some fluid voxels near to the colon wall, where the gray value is away from standard values. This is the case of the bottom-right example in Fig. 14. An additional representative set of false positives is shown in Fig. 14.

7. Conclusion

We introduced a complete pipeline for a CAD algorithm that flags candidate polyp regions. The segmentation stage is very simple and fast, and its main novelty is the smoothing PDE which enhances the polyps, leading to a better detection. In addition to the incorporation of the Haralick texture features, the main yet simple novelties of the proposed features and classification stages are twofold. First, the surrounding area of candidate polyps are explicitly taken into account. Indeed, the proposed (so-called differential) features are computed by comparing properties in the central and surrounding regions of the polyps. We show that differential features are more discriminative than the absolute ones, as they emphasize local deviations of the geometry and texture over the colon. The other novelty is an adaptive-scale strategy that test regions of different sizes and automatically selects the region that best delineates each candidate polyp. The obtained quantitative results are very promising, detecting 100% of the true-polyps, including small and flat ones, with a low false positives rate. Additional improvement of the segmentation and, in collaborations with radiologists, finding features that are tailored to polyp-like geometries, can further improve these results.

Appendix A. Number of Iterations: Analytical Solving

We want to compute the number of iterations needed to make a certain sphere (of the size of the CT resolution) vanish, according to our proposed PDE. For a sphere, the SI is constant, so the PDE becomes:

$$\frac{\partial S}{\partial t} = \kappa_{\min}^{1/4} \mathcal{N}.$$

Then, as $k_{\min} = 1/r$, the sphere radius satisfies the following differential equation:

$$r'(t) = \frac{-1}{r(t)^{1/4}}.$$

Therefore,

$$\int_0^T r'(t)r(t)^{1/4}dt = -\int_0^T 1dt \Rightarrow \frac{4}{5}(r(T)^{4/5} - r(0)^{4/5}) = -T.$$

As we want to find the value of T so that $r(T) = 0$,

$$T = \frac{4}{5}r(0)^{4/5}.$$

The resolution in the z -direction is 1 mm in our examples, and the time step considered for the numerical method was $t_s = 0.055$. That gives a value N for the number of iterations $N = \lceil 14.54 \rceil = 15$.

References

1. A. Aschoff, A. Ernst, H.-J. Brambs and M. Juchems, CT colonography: An update, *Eur. Radiol.* **18**(3) (2008) 429–437.
2. A. Blachar, M. Graif, A. Kessler and J. Sosna, State-of-the-art CT colonography: Update on technique and performance, *Curr. Colorectal Cancer Rep.* **3**(1) (2007) 49–54.
3. J. H. Bond, Colorectal cancer screening: The potential role of virtual colonoscopy, *J. Gastroenterol.* **37** (2002) 92–96.
4. V. Caselles and C. Sbert, What is the best causal scale space for three-dimensional images? *SIAM J. Appl. Math.* **56**(4) (1996) 1199–1246.
5. D. Chen, R. Fahmi, A. A. Farag, R. L. Falk and G. W. Dryden, Accurate and fast 3D colon segmentation in CT colonography, *ISBI 2009* (2009), pp. 490–493.
6. J. Church, Clinical significance of small colorectal polyps, *Dis. Colon Rectum* **47** (2004) 481–485.
7. M. Di Martino, G. Hernández, M. Fiori and A. Fernández, A new framework for optimal classifier design, *Pattern Recogn.* **46**(8) (2013) 2249–2255.
8. J. Fidler and C. Johnson, Flat polyps of the colon: Accuracy of detection by CT colonography and histologic significance, *Abdom Imaging* **34** (2009) 157–171.
9. M. Fiori, P. Musé, S. Aguirre and G. Sapiro, Automatic colon polyp flagging via geometric and texture features, *Int. Conf. IEEE EMBS*, Vol. 1 (2010) pp. 3170–3173.
10. M. Fiori, P. Musé and G. Sapiro, A complete system for candidate polyps detection in virtual colonoscopy, Available at <http://arxiv.org/abs/1209.6525> (2012).
11. M. Fiori, P. Musé and G. Sapiro, Polyps flagging in virtual colonoscopy, in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* (Springer, 2013), pp. 181–189.
12. M. Franaszek, R. M. Summers, P. Pickhardt and R. Choi, Hybrid segmentation of colon filled with air and opacified fluid for CT colonography, *IEEE Trans. Med. Imaging* **25** (2006) 358–368.
13. M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. Witten, The WEKA data mining software: An update, *AIM SIGKDD Explorations Newsl.* **11**(1) (2009) 10–18.

14. R. M. Haralick, K. Shanmugam and I. Dinstein, Textural features for image classification, *IEEE Trans. Syst., Man and Cybern.*, **3**(6) (1973) 610–621.
 15. J. J. Koenderink, *Solid Shape* (MIT Press, Cambridge, USA, 1990).
 16. S. Osher and J. Sethian, Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations, *J. Comput. Phys.* **79** (1988) 12–49.
 17. D. Paik, C. Beaulieu, G. Rubin, B. Acar, R. Jeffrey, J. Yee, J. Dey and S. Napel, Surface normal overlap: A computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT, *IEEE Trans. Med. Imaging* **23**(6) (2004) 661–675.
 18. P. Pickhardt, J. Choi, I. Hwang, J. Butler, M. Puckett, H. Hildebrandt, R. Wong, P. Nugent, P. Mysliwiec and W. Schindler, Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults, *New Engl. J. Med.* **349**(23) (2003) 2191–2200.
 19. G. Slabaugh, X. Yang, X. Ye, R. Boyes and G. Beddoe, A robust and fast system for CTC computer-aided detection of colorectal lesions, *Algorithms* **3**(1) (2010) 21–43.
 20. R. M. Summers, CT colonography computer-aided detection: Effect on radiologist observers, *MICCAI 2010* (Beijing, China, 2010), pp. 1–4.
 21. P. Sundaram, A. Zomorodian, C. Beaulieu and S. Napel, Colon polyp detection using smoothed shape operators: Preliminary results, *Med. Image Anal.* **12**(2) (2008) 99–119.
 22. K. Suzuki, H. Yoshida, J. Näppi, S. G. Armato and A. H. Dachman, Mixture of expert 3D massive-training ANNs for reduction of multiple types of false positives in CAD for detection of polyps in CT colonography, *Med. Phys.* **35**(2) (2008) 694.
 23. C. van Wijk, V. F. van Ravesteijn, F. M. Vos and L. J. van Vliet, Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow, *IEEE Trans. Med. Imaging* **29**(3) (2010) 688–698.
 24. D. Vining, Y. Ge, D. Ahn and D. Stelts, Virtual colonoscopy with computer-assisted polyps detection, in *Computer-Aided Diagnosis in Medical Imaging* (Elsevier Science, 1999) pp. 445–452.
 25. D. Vining, D. Gelfand, R. Bechtold, E. Scharling, E. Grishaw and R. Shifrin, Technical feasibility of colon imaging with helical CT and virtual reality, *Am. J. Roentgenol.* **162** (1994), 104.
 26. Z. Wang, Z. Liang, L. Li, X. Li, B. Li, J. Anderson and D. Harrington, Reduction of false positives by internal features for polyp detection in CT-based virtual colonoscopy, *Med. Phys.* **32**(12) (2005) 3602–3616.
 27. R. A. Wolber and D. A. Owen, Flat adenomas of the colon, *Human Pathol.* **22**(1) (1991) 70–74.
 28. World Health Organization, Cancer (2011).
 29. H. Yoshida and J. Näppi, Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps, *IEEE Trans. Med. Imaging* **20**(12) (2001) 1261–1274.
-



Marcelo Fiori received his Electrical Engineering degree in 2008, and his M.Sc. degree in 2011 from the Universidad de la República, Uruguay (UdelaR), and he is currently a Ph.D. student. He holds an Assistantship at the Institute of Mathematics at UdelaR. His

main research interests are machine learning, graph matching problems, and sparse linear models, with special focus in signal processing.



Pablo Musé received his Electrical Engineering degree from Universidad de la República, Uruguay, in 1999, his M.Sc. degree in Mathematics, Vision and Learning and his Ph.D. in Applied Mathematics from ENS de Cachan, France, in 2001 and 2004, respectively.

From 2005 to 2006 he was with Cognitech, Inc., Pasadena, CA, USA, where he worked on computer vision and image processing applications. In 2006 and 2007, he was a Postdoctoral Scholar with the Seismological Laboratory, California Institute of Technology, Pasadena, working on remote sensing using optical imaging, radar and GPS networks. Since 2008, he has been with the Division of Electrical Engineering, School of Engineering, Universidad de la República, where he is currently an Associate Professor of signal processing.



Guillermo Sapiro was born in Montevideo, Uruguay, on 3rd April 1966. He received his B.Sc. (summa cum laude), M.Sc., and Ph.D. from the Department of Electrical Engineering at the Technion, Israel Institute of Technology, in 1989, 1991, and 1993, respec-

tively. After postdoctoral research at MIT, Dr. Sapiro became a Member of Technical Staff at the research facilities of HP Labs in Palo Alto, California. He was with the Department of Electrical and Computer Engineering at the University of Minnesota, where he held the position of Distinguished McKnight University Professor and Vincentine Hermes-Luh Chair in Electrical and Computer Engineering. Currently he is the Edmund T. Pratt, Jr. School Professor with Duke University.

G. Sapiro works on theory and applications in computer vision, computer graphics, medical imaging, image analysis, and machine learning. He has authored and co-authored over 300 papers in these areas and has written a book published by Cambridge University Press, in January 2001.

G. Sapiro was awarded the Gutwirth Scholarship for Special Excellence in Graduate Studies in 1991, the Ollendorff Fellowship for Excellence in Vision and Image Understanding Work in 1992, the Rothschild Fellowship for Post-Doctoral Studies in 1993, the Office of Naval Research Young Investigator Award in 1998, the Presidential Early Career Awards for Scientist and Engineers (PECASE) in 1998, the National Science Foundation Career Award in 1999, and the National Security Science and Engineering Faculty Fellowship in 2010. He received the test of time award at ICCV 2011. G. Sapiro is a Fellow of IEEE and SIAM. G. Sapiro is also the founding Editor-in-Chief of the SIAM Journal on Imaging Sciences.