June 2001 Technical Report: KUL/ESAT/PSI/0104

Computer Aided Diagnosis for Virtual Colonography by Geometric Model Fitting

Gabriel Kiss,
Johan Van Cleynenbreugel,
Maarten Thomeer, Guy Marchal,
Paul Suetens

Katholieke Universiteit Leuven - Center for Processing Speech

PSI

Kasteelpark Arenberg 10
B-3001 Heverlee, Belgium

Telephone: +32-(0)16-32.17.13 Fax: +32-(0)16-32.17.23

KU LEUVEN

Computer Aided Diagnosis for Virtual Colonography by Geometric Model Fitting

Gabriel Kiss, Johan Van Cleynenbreugel, Maarten Thomeer, Guy Marchal, Paul Suetens

Abstract— We present a system for Computer Aided Diagnosis in Virtual Colonography based on simple geometric model fitting. Our approach extends surface normal analysis and sphere fitting methods. We label locations in the volume data, which have a high probability of being colonic polyps, and present them in a user-friendly way. The method was tested on a study group of 30 patients. Using normal colonoscopy as standard of reference, true positive and false positive findings were determined. The detection rate for polyps larger than 9mm was 100% comparable to that of human readers. We introduce sphere fitting and normal distribution analysis as additional steps to reduce the number of false positive findings generally encountered with surface normal methods. Initial results show that Computer Aided Diagnosis is feasible and that our method holds potential for screening purposes.

Index terms—CT Colonography, Computer Aided Diagnosis, Screening, Colonoscopy, Model Fitting

I. INTRODUCTION

Colorectal carcinoma is the fourth most common cancer and the second leading cause of cancer-related death in the industrialized world [1], with a 4-6% lifetime risk in the general population [2]. Therefore early detection and treatment of adenomatous colonic polyps is widely considered to be crucial to prevent colon cancer [3]. Fecal occult blood testing, barium enema, sigmoidoscopy and colonoscopy are the classical ways of detecting colonic lesions. However, when one considers widespread population screening, sub-optimal patient acceptance is a major problem. E.g. conventional colonoscopy requires colonic cleansing, while colonic wall perforation may occur, hence increasing colateral risks. Furthermore results obtained are heavily dependent on operator experience. For such reasons the need for less invasive, faster and cheaper procedures for colonic screening has arisen.

Virtual colonography (VC), hinging on volumetric image data, can be less invasive. VC may be based on ultrasound, on magnetic resonance imaging (MRI) or on Xray computed tomography (CT). Although the diagnostic quality of MRI data and CT data is equal in this context,. CT still is the modality of choice, despite ionizing radiation. Indeed with MRI, major patient discomfort results from the need to completely fill the colon by a fluid contrast medium. In case of CT the colon is insufflated with air. VC applied to CT images is referred to as CTC (CT colonography). The visualization component of CTC covers a range of techniques employed to generate view^{2}s of the colonic wall. Among them are common representations such as 2D axial slices, 3D multi-planar and curvilinear reformatted images, but also more experimental techniques as virtual double contrast imaging [4], Axial 3D [5]. or 3D endoluminal images (generally known as virtual colonoscopy, e.g. [16]). In one way or another, CTC visualization is about providing an expert human reader with optimal visual access to the CT image data.

However given the expected increase in image volume sizes (soon 1024 voxels in each dimension), access to image contents might benefit from Computer Aided Diagnosis (CAD). The ultimate goal of CTC CAD is the automatic detection and classification of polyps. A subgoal in the short run is the determination of locations, which have a high probability of corresponding to colonic polyps. Such locations are then presented to the human reader, still requiring discrimination between true and false positive findings. Research projects such as [6] already demonstrated the potential of improving current diagnostic performance and speed in CTC by CAD.

This work is part of the GOA/99/05 project: "Variability in Human Shape and Speech", financed by the Research Fund, K.U. Leuven, BELGIUM.

G. Kiss, J. Van Cleynenbreugen, P. Suetens are members of the Laboratory for Medical Imaging Research (directors: Paul Suetens & Guy Marchal), a cooperation between the Department of Electrical Engineering, ESAT (Kasteelpark Arenberg 10, B-3001, Heverlee), and the Department of Radiology, University Hospital Gasthuisberg (Herestraaat 49, B-3000, Leuven), of the Katholieke Universiteit Leuven, Belgium.

M. Thomeer, G. Marchal are members of the Department of Radiology, University Hospital Gasthuisberg (Herestraaat 49, B-3000, Leuven), of the Katholieke Universiteit Leuven, Belgium.

E-mail: Gabriel.Kiss@uz.kuleuven.ac.be

In contrast to CAD for mammography (e.g. [7]) or for the detection of lung nodules (e.g. [8]), CAD for CTC is a relatively new research topic. As illustrated in Figure 1, the colon has a very complex spatial structure containing patches of smooth colonic wall next to haustral folds, as well as occasional residual fluid or stool and polyps. Unravelling the complexity of this structure does require time-consuming algorithms. Recent research (e.g. [6][9][10][11][12] [13]) mainly draws from one of three approaches: surface curvature calculation, surface normal analysis and sphere fitting methods. Figure 2 summarizes these methods. From a conceptual point of view, such approaches belong to the object recognition strategy "Fitting Models to Photometry", defined in [15]. Examples of this strategy instantiate relatively simple geometric models to fairly complex image data in order to retrieve particular objects. Here, the latter are polyps while the former describe their expected shape.

II. METHOD

We have developed a geometric model fitting method for CTC CAD by combining and adapting the approaches of surface normal analysis and sphere fitting. Our method contains four major steps (see Figure 3): A) segmentation, B) polyp candidate generation, C) polyp center generation, and finally D) polyp extraction, normal distribution analysis and polyp presentation. Throughout these steps the processing is dependent on a number of (threshold) parameters, which are summarized in Table 1. Finally, a software phantom was used to determine initial values for the threshold values. It consists of a perfect cylinder simulating a haustral fold and a perfect semi-sphere simulating a polyp. The diameter of the semi-sphere was 10 mm similar to the targeted polyps. The phantom was also used to validate the hypothesis beyond each of the steps of our method. Figure 7 clarifies these ideas.

A) Segmentation

In this step the colonic wall is determined. Since CTC images show a large contrast between (insufflated) colonic air and the colonic wall, classic region-growing algorithms (e.g. see [14]) can be used successfully. Interactively, a seed voxel is chosen and a connected region A of air voxels (determined by $T_{segment}$) is grown outwards in a breadth first manner. The colonic wall is then defined as the set C of voxels adjacent to A, and having an intensity value higher than $T_{segment}$. For each voxel of A all 26 neighbors are taken into consideration, this can lead to a wall thickness of two voxels.

CTC patient preparation may influence this segmentation technique. Indeed wet preparation can lead to a substantial amount of residual fluid left in the colon. Multiple seed points are needed then in order to deal with fluid filled colonic segments and with collapsed regions (due to the inadequate distension). Therefore the final result C of the segmentation may be a set of disjunctive regions. We have to mention that in such cases C does not only contain voxels belonging to the colonic wall but also additional voxels belonging to colonic fluid. Furthermore the submerged parts of the colonic wall are missing. A possible solution here is the use of digital cleansing prior to CAD [5].

B) Polyp candidate generation

Only a small number of voxels of C will turn into actual polyp candidates. The spherical nature of the polyps suggests to look for convex surfaces patches, as opposed to general colonic wall, which is mostly concave. Estimating the curvature properties could be done by using tools from differential geometry such as calculating principal, Mean and Gaussian curvatures. However we employ a simple analytic geometry approach to analyze surface normals.

As illustrated in Figure 4 we compare normals in a current voxel of C (say $\mathbf{p_1}$) and in its neighboring voxels in C (say $\mathbf{p_2}$). The normal to the colonic wall is calculated by a 3D-image gradient Zucker-Hummel operator [14]. It is oriented from air towards tissue. Let $\mathbf{g_1}$ and $\mathbf{g_2}$ be the image gradients in $\mathbf{p_1}$ and $\mathbf{p_2}$ respectively. The equation of the plane tangent to C at $\mathbf{p_1}$ is then given by $\mathbf{p} \bullet \mathbf{g_1} = \mathbf{p_1} \bullet \mathbf{g_1}$ where $\bullet$ is the dot product: The line through $\mathbf{p_2}$ along the direction $\mathbf{g_2}$ has equation $\mathbf{p} = \mathbf{p_2} + r\mathbf{g_2}$ **;** its intersection point $\mathbf{p_3}$ with the tangent plane corresponds to $r = (\mathbf{p_1}-\mathbf{p_2}) \bullet \mathbf{g_1} / \mathbf{g_2} \bullet \mathbf{g_1}$. Figure 4 explains how the relative location of voxels $\mathbf{p_2}$ and $\mathbf{p_3}$ is exploited to assess convexity or concavity of the colon's surface in $\mathbf{p_1}$. Furthermore in case of presumed convexity, a T_{convex} threshold is applied to $\|\mathbf{p_2}-\mathbf{p_3}\|$. For $\mathbf{p_1} \in C$, let B the bounding box around $\mathbf{p_1}$ defined by T_{box}. For each $\mathbf{p_2} \in B \cap C$ we repeat the plane intersection procedure and compute the values V_c and V_t. V_c is the number of voxels situated in $B \cap C$ that satisfy T_{convex}, while V_t is the total number of voxels in $B \cap C$. Finally those $\mathbf{p_1}$ for which V_c/V_t is higher than T_{hits} are declared polyp candidates.

This step will generate all the voxels belonging to convex surfaces. As a result it will return all the voxels situated on polyps but also on haustral folds (and even on the colonic wall due to local wall irregularities) since they have similar convexity properties. Therefore the next step is discrimination between voxels belonging to real polyps and those belonging to haustral folds or colonic wall.

C) *Polyp center generation*

This step exploits the spherical nature of polyps versus the cylindrical nature of haustral folds. A 3D Hough transform is applied to all previously detected polyp candidates to look for sphere centers in the CT image volume.

Given a fixed radius (T_{radius}) of a fitting sphere, each polyp candidate is assumed to be on the surface of such sphere, while its image gradient is pointing towards the center of the sphere, see Figure 5.
In this way a center voting procedure generates a 'center map': for each center we store the number of its polyp candidate 'voters' as well as their (normalized) image gradients, called 'normals' of the center, hereafter. In this way each center in the map can be visualized as shown in Figure 6. Because folds have a cylindrical shape the centers given by such structures will be dispersed along a line. For polyps however these centers will converge towards a small area.(see Figure 7 also).
The value for T_{radius} needs to be selected in close correspondence with the size of the polyps to be detected. Of course multiple passes for different T_{radius} values can be applied to the same set of polyp candidates.
D). *Polyp extraction, normal distribution analysis and presentation*
Given the center map from the previous step, we present polyp candidates on 2D slices to the reader. Therefore local maxima are extracted from the center map, after non-maximum suppression. If none of these exceed $T_{extract}$ the patient is declared negative.
However deciding solely on the number of polyp candidate voters may lead to too much false positive findings per case. Therefore we also take into account the spatial distribution of the normals of each center. Normals generated by haustral folds are almost coplanar, while for polyps a large variance in the normal distribution may be expected. This idea is presented in Figure 6. We measure this spatial distribution by taking a reference plane defined by the center point and two of its normals, randomly chosen. For every other normal we calculate the distance to this reference plane. The thus obtained maximum distance D_{max} gives a good measure of the distribution. An even more accurate measure is the minimum of all D_{max} distances, obtained by varying the pair of normals to span up reference planes. A further thresholding (TD_{max}) is applied and the remaining centers are declared as polyps and are presented to the reading radiologist for evaluation.

III PATIENT SCANNING AND RESULTS

Thirty patients, 15 normal and 15 with 35 polyps of various sizes (Table 2) underwent CT colonography prior to conventional colonoscopy. Informed consent was obtained from all patients. The patient preparation consisted in the oral administration of 3 to 5 liter of precolon, an in-house developed tagging agent. In some cases the use of polyethylene glycol electrolyte solution was preferred. Immediately before CT colonography a bowel relaxant was injected intravenously. CO2 was insufflated using a bag system.
CT colonography was performed on a multi-detector CT (Volume Zoom Multi Slice Helical CT from Siemens) using 4x1 mm detector configuration, 7 mm table feet per 0.5 s tube rotation, 0.8 mm reconstruction increment as well as 60 effective mAs and 120 keV. Patients were scanned in both supine and prone positions, in breadth holds of 20 to 30 seconds.
On average the size of the acquired data sets was 252.9 MB. The image processing was done on an Intel Pentium III system running at 533 MHz and having 512 MB of RAM.
We used the described CAD method on all the patients with the same parameter settings. Using conventional colonoscopy as standard of reference true positive (TP) and false positive (FP) findings were determined for each patient. These values were then compared with the results obtained by a radiologist (MT, the third author) using the CTC method of Axial 3D [5].
The total number of polyps was 35. The detection rate differentiated on polyp size is presented in Table 2. The average computation time on our system was 24.26 minutes and it was mostly made up of the polyp candidate generation. The total number of false positives was 122, which gives us a mean value of 4.06 false positive findings per case. For the first 18 patients a comparison between the number of false positives with and without the use of normal distribution analysis was also estimated (Table 3). Results show a reduction from 147 to 64 false positives in this way. Overall the causes of false positives were colonic wall (38.52%), haustral folds (41.8%), colonic stool or fluid (12.29%), the insuflation tube (4.9%) and the ileocecal valve (2.49%).

IV DISCUSSION AND CONCLUSIONS

Although the size of significant polyp is a debate subject between different radiologists and also between radiologists and gastroenterologists, our primary goal was the detection of polyps bigger than 10 mm. A secondary goal was a reasonable number of false positives per case. Using the polyp presentation application the radiologist can quickly go through the polyp candidate list and discriminate between real and false positive findings.

From literature, we can deduce that there is a strong correlation between sensitivity and the number of false positives generated. For example Tomasi et al. [10] are using a graph method optimized for the detection of polyps larger than 10mm. They achieve a sensitivity of 100% with a number of false positives (FP) as high as 50 per data set. Yoshida et al. [11] reported a sensitivity of 90% for polyps between 7-12mm and 1 FP per case. To achieve a

sensitivity of 100% they reported an increase towards 1.5 FP per case. Beaulieu et al. [9] developed three different methods for polyp detection. Their contour normal method has a sensitivity of 96.4% for polyps larger than 5 mm at the cost of 25 FP per data set. The sphere fit method returns a number of 47 FP, while the surface curvature method has a low number of FP of only 4, but a decreased sensitivity as well (3 out of 7 polyps detected).

We have to stress that we are looking at CAD as a prospective tool and not as a retrospective one. We used the same parameter settings for all of our patients, positives or negatives. However if we would count only the number of false positives until all the polyps bigger than 10mm have been identified, then we would only have 20 FP findings in 15 cases, so an average of 1.33 FP per case comparable with results given by other authors. Of course cut off values can be determined only for retrospective studies and thus, they are not useful in the case of large scale-screenings.

By combining the surface normal and sphere fitting methods we tried to extract and use the advantages of both methods. In fact using the sphere fitting as an additional step we are able to make a better differentiation between polyps and haustral folds. The novelty of our method consists in using a Hough transform based method for sphere fitting and the introduction of the normal distribution analysis, which causes an important reduction of more than 50% of the number of detected FP. Previously methods based on sphere fitting and graph based clustering were applied by Tomasi [10]. While the sphere fitting is a linear problem, the graph searching is a complex method proportional in time to the number of edges in the graph. By incorporating gradient information Hough based methods are reducing the sphere patch detection to a linear problem.

Further improvements to our method are possible. First, the segmentation step can be improved by applying digital cleansing prior to CAD, in order to remove the colonic stool and fluid, and thus reducing the number of false positive findings given by these structures. Also a better colonic distension could eliminate many collapsed colonic regions and the need of multiple seed points. Alternatively a fully automatic segmentation of the colonic lumen as proposed by Chen [16] could be a viable solution to solve the segmentation problem.

The preliminary results of our experiments are at least encouraging. What they show is that our method is feasible, and it may become useful for clinical studies if the number of false positive findings will be further reduced and the detection rate for polyps between 5 and 9 mm will increase. We can conclude that CAD will probably become the most common way of doing CTC, improving on current accuracy, efficiency and costs.

REFERENCES

[1] J.D. Potter et al., "Colon cancer: a review of the epidemiology", Epidemol Rev 1993; 15:449-545

[2] G. Stevenson, "Radiology in the detection and prevention of colorectal cancer", Eur J Cancer 1995;31A:1121-1126

[3] S.J. Winawer, A.G. Zauber, M.N. Ho, et al, " Prevention of colorectal cancer by colonoscopic polyectomy: the National Polyp Study Workgroup", N Engl J Med 1993; 329:1977-1981

[4] W. Luboldt, O. Luz, R. Vonthein, M. Heuschmid, M. Seemann, J. Schaefer, D. Stueker and C.D. Claussen, "Three-Dimensional Double Contrast MR Colonography: A Display Method Simulating Double-Contrast Barium Enema", AJR 2001; 176:930-932

[5] G. Kiss, J. Van Cleynebreugel, M. Thomeer, P. Suetens and G. Marchal, "An Image Processing and Visualization System for Virtual Colonography", KUL/ESAT/PSI/0103, June 2001, KU Leuven, Belgium

[6] R.M. Summers, C.D. Johnson, L.M. Pusanik, J.D. Malley, A.M. Youssef and J.E. Reed, "Automated Polyps Detection at CT Colonography: Feasibility Assessment in a Human Population, Radiology. 2001 Apr; 219(1): 51-9.

[7] J. Nappi and P.B. Dean, "A multiscale algorithm for segmenting calcifications from high-resolution mammographic specimen radiographs", J Digit Imaging, May 2000, 13(2 Suppl1):130-2

[8] Y. Ukai, N. Niki, H. Satoh, K. Eguchi, K. Mori, H. Ohmatsu, R. Kakinuma, M. Kaneko, and N. Moriyama, Computer Aided Diagnosis System for Lung Cancer Based on Retrospective Helical CT Images, CARS proceedings, June 2000, pp 767-772

[9] C. F. Beaulieu, "Computer Aided Detection of Colonic Polyps", Second International Symposium on Virtual Colonoscopy, BOSTON, October 2000, p.73-77

[10] C. Tomasi, et al., "A Graph Method for the Conservative Detection of Polyps in the Colon", Second International Symposium on Virtual Colonoscopy, BOSTON, October 2000, p.105

[11] H.Yoshida et al., "Computer Aided Detection of Colonic Polyps in CT Colonography", Second International Symposium on Virtual Colonoscopy, BOSTON, October 2000, p.104

[12] D. S. Paik et al., "Computer Aided Detection of Polyps in CT Colonography: Free Response ROC Evaluation of Performance", RSNA, Chicago, 2000

[13] R.M. Summers, C.F. Beaulieu, L.M. Pusanik, J.D. Malley, R.B. Jeffrey, D.I. Glazer and S. Napel, "Automated polyp detector in CT colonography: feasibility study", Radiology 2000 Jul;216(1):284-90

[14] D.M. Ballard, C.M. Brown, "Computer Vision", Prentice-Hall, pp 123-166

[15] P.Suetens, P. Fua, and A. J. Hanson, "Computational Strategies for Object Recognition", ACM Computing Surveys, 24(1), pp 5-61 March 1992

[16] D.Chen, Z. Liang, M.R. Wax, L. Li, B. Li and A.E. Kaufman, "A novel approach to Extract Colon Lumen from CT Images for Virtual Colonoscopy, IEEE Trans Med Imaging. 2000 Dec;19(12):1220-6

Threshold parameters	**Threshold symbol**	**Step**	**Short description**
Colonic wall	$T_{segment}$	Segmentation	Differentiate between air, colonic wall and surrounding tissue
Bounding box size	T_{box}	Polyp candidate generation	Defines the 3D volume of interest we take into consideration for a colonic wall voxel
Convexity	T_{convex}	Polyp candidate generation	Distance between a neighbor point and the intersection point of the local gradient with the tangent plane
Hits/total ratio	T_{hits}	Polyp candidate generation	Minimum number of neighbors relative to the total number of neighbors that have to satisfy the convexity threshold
Sphere radius	T_{radius}	Polyp center generation	Distance between the current point and the center of an imaginary sphere
Extraction	$T_{extract}$	Polyp extraction	Minimal value in the center map to be considered as polyp candidate
Distribution	TD_{max}	Polyp extraction	Threshold which measures the distribution of the normals on a sphere

Table 1. Threshold parameters, their name, symbol, appearance and meaning.

Polyps	**Total**	**TP**	**$TP_{cut\ off}$**	**FP**	**$FP_{cut\ off}$**
< 5mm	11	0	0	122	20
Flat lesions	4	0	0		
5-9 mm	5	2	0		
>= 10mm	15	15	15		
Overall	35	17	15		

Table 2. Results of the algorithm used for computer aided diagnosis in colorectal polyp detection in relation to polyp size and cut off value. Cut off chosen retrospectively in such a manner that all polyps larger than 10 mm are identified.

False positives	Wall	Folds	Stool	Valve	Tube	Total
NO normal distrib.analysis	67	47	18	4	11	147
Normal distribution analysis	25	21	10	3	5	64

Table 3. Comparison between CAD methods, when not using and then using normal distribution evaluation. False positives are reduced to 43.54% of the original value.

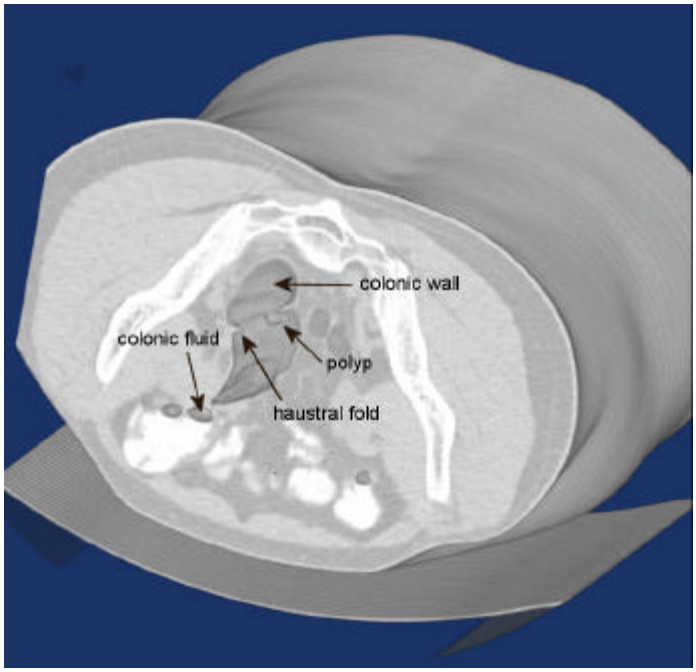

Figure 1. Illustration of the colonic region obtained from a CTC acquisition in prone position. The complex and irregular shape of the colon contains patches of smooth colonic wall next to haustral folds, as well as occasional residual fluid and polyps.

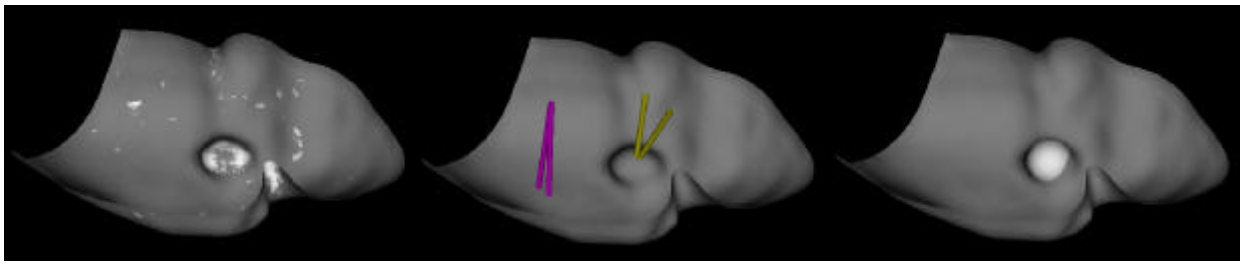

Figure 2. Figure showing the basic principle of different CAD methods. First values for mean curvature are presented on a smoothed surface extracted from a real colon. Polyps and haustral folds have different curvature values than normal colonic wall, as a consequence curvature based methods are looking for regions with high differences in curvature and label them as polyp candidates. Secondly the basic idea of surface normal methods is shown. Normals are computed for voxels belonging to colonic wall and the intersection point between neighbors is inspected. Finally the principle of sphere fitting methods is presented: a sphere is partially fitted to the polyp. Polyps are more complex structures than simple spheres, that is why sphere fitting methods are looking for sphere patches in close vicinity.

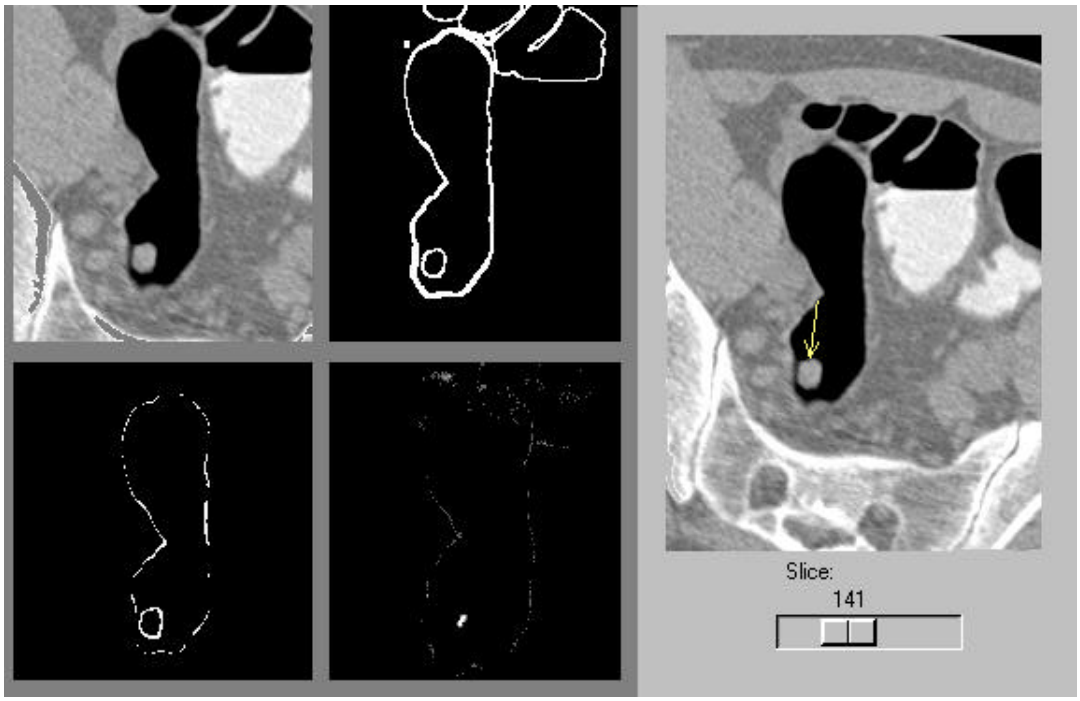

Figure 3. The figure illustrates the four steps of our algorithm. Although our approach is truly 3D, we only show results on 2D slices. On the left a 2D region of interest is presented first as an original slice, then the results of segmentation, polyp candidate generation and polyp center generation are shown. On the right the result of the polyp presentation module is depicted. Each polyp candidate is highlighted by an arrow when scrolling through the polyp list.

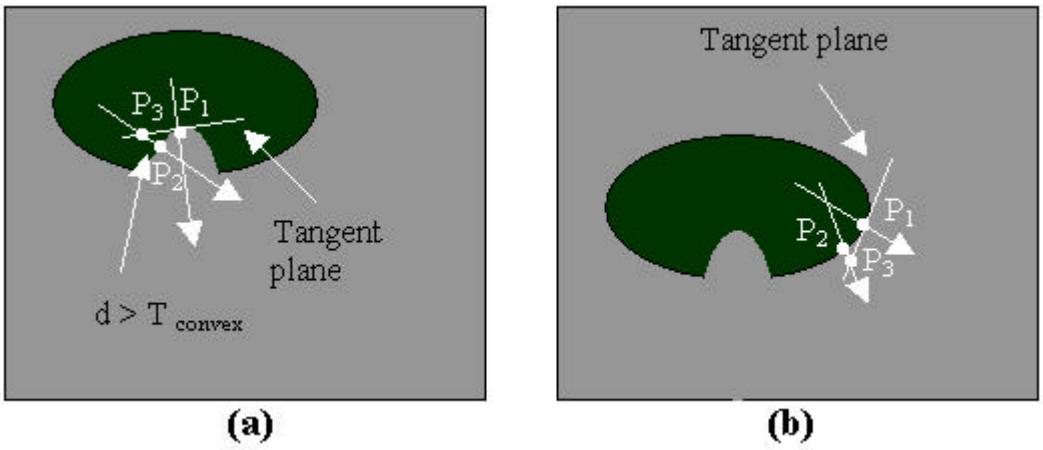

Figure 4. Difference between convex and concave surfaces. For convex surfaces (a) point 3 is situated in negative direction relative to point 2, while for concave surfaces (b) point 3 is situated in positive direction relative to point 2 on the line given by the local gradient. For convex surfaces T_{convex} is applied.

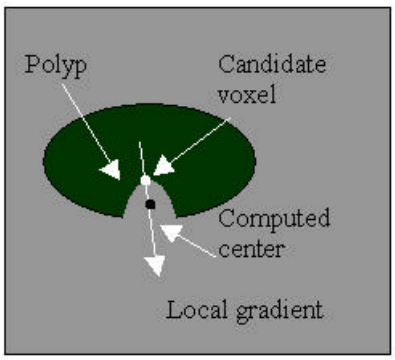

Figure 5. Sphere center computation.

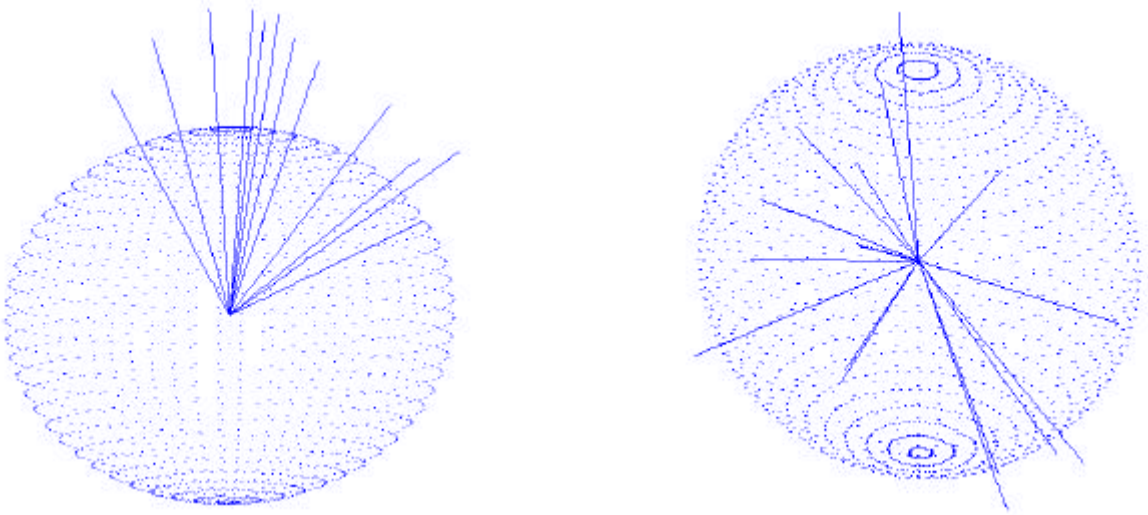

Figure 6. The difference in normal distribution between a haustral fold (left) and a polyp (right). To give spatial hints a sphere of 6 mm diameter is plotted aroun the respective centers.

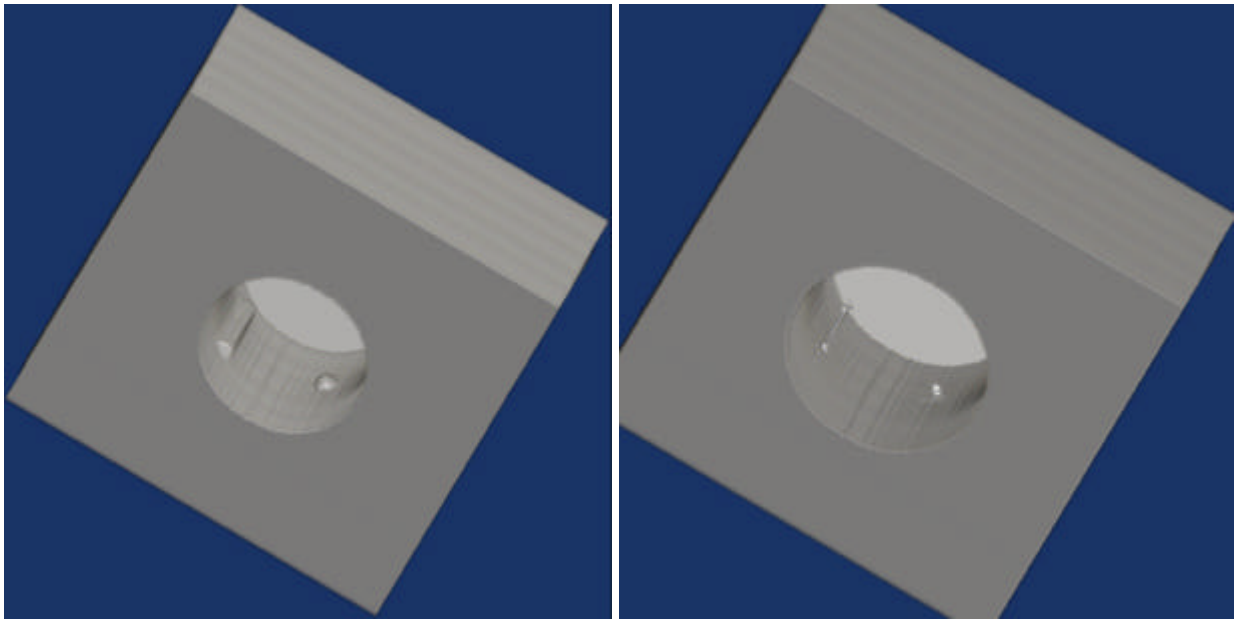

Figure 7. Original software phantom (left) and results from applying steps A, B and C on the phantom (right). The dispersion of sphere centers for the simulated haustral folds is observable as opposed to the convergence on the polyp.