

Automated Detection of Polyps for CT Colonography

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op maandag 1 maart 2010 om 12.30 uur
door

Cornelis van WIJK

ingenieur in de luchtvaart en ruimtevaart
geboren te Zwammerdam.

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. L.J. van Vliet

copromotor:

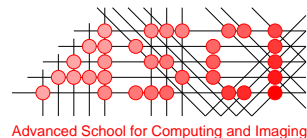
dr. F. M. Vos

Samenstelling van de promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. ir. L.J. van Vliet	Technische Universiteit Delft, promotor
dr. F.M. Vos	Technische Universiteit Delft, copromotor
Prof. dr. ir. M.J.T. Reinders	Technische Universiteit Delft
Prof. dr. ir. J.H.C. Reiber	Leids Universitair Medisch Centrum
Prof. dr. J. Stoker	Academisch Medisch Centrum Amsterdam
dr. B. van Ginneken	Universitair Medisch Centrum Utrecht
Prof. dr. P.F. Whelan	School of Electronic Engineering, Dublin
Prof. dr. I.T. Young	Technische Universiteit Delft, reservelid



This work was supported by the Dutch Ministry of Economic Affairs, through their Innovation-Driven Research Programme (IOP Beeldverwerking), project number IBV02020.



This work was carried out in graduate school ASCI.

Copyright © 2009 by Cornelis van Wijk
All rights reserved.

ISBN 978-90-9025155-4

Contents

1	Introduction	1
1.1	CT colonography	2
1.2	Automated detection	4
1.3	Thesis organization	7
2	On normalized convolution for the measurement of image derivatives in highly structured surroundings	9
2.1	Introduction	10
2.2	Methods	10
2.2.1	A least squares approach to normalized convolution	10
2.2.2	Normalized convolution and Gaussian derivatives	12
2.2.3	Local confidence values	13
2.3	Results	13
2.4	Conclusions	17
3	Segmentation and size measurement of polyps in CT colonography	19
3.1	Introduction	20
3.2	Polyp segmentation	21
3.3	Experiments and results	23
3.4	Conclusions & future work	27
4	Detection of protrusions on curved folded surfaces applied to auto- mated detection in CT colonography	29
4.1	Introduction	30
4.2	Methods	32
4.2.1	Surface evolution	33
4.3	Results	35
4.3.1	Experimental data	35
4.3.2	Candidate generation	35
4.3.3	CAD performance	36
4.4	Conclusions	37

5	Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow	41
5.1	Introduction	42
5.1.1	Previous work	43
5.1.2	Problem definition	44
5.1.3	Objective	45
5.2	Methodology	46
5.2.1	Materials	46
5.2.2	Method	46
5.2.3	Second principal curvature flow	48
5.2.4	Implementation	49
5.2.5	Candidate segmentation	51
5.2.6	Features for classification	53
5.2.7	Classifier training	54
5.3	Experiments and results	57
5.3.1	Qualitative analysis	57
5.3.2	Performance of the candidate detection	59
5.3.3	Results after classification	59
5.3.4	A combined approach	60
5.4	Discussion and conclusion	61
6	Automatic polyp size measurement for CT colonography based on a protrusion estimation method	63
6.1	Introduction	64
6.2	Materials & methods	65
6.3	Results	69
6.3.1	Accuracy and measurement variability	69
6.3.2	Intra-observer variability	70
6.3.3	Inter-observer variability	71
6.3.4	Orientation of the phantom in the scanner	71
6.3.5	Slice thickness	71
6.4	Discussion	73
7	Computer aided detection of polyps in CT colonography using logistic regression	77
7.1	Introduction	78
7.1.1	Related work	78
7.1.2	Objective	80
7.2	Data description and feature design	81
7.2.1	CT colonography data	81
7.2.2	Candidate detection	83

7.2.3	Features	83
7.3	Characteristics of the feature space	88
7.4	The classification system	90
7.4.1	Mahalanobis distance mapping	90
7.4.2	Normal-based discriminant classifiers	92
7.4.3	Logistic discriminant classifier	93
7.5	Results	94
7.5.1	Classifier selection: performance and stability	94
7.5.2	Outlier rejection by Mahalanobis distance mapping	96
7.5.3	Multi center evaluation	98
7.6	Discussion and conclusion	101
8	Conclusions	103
8.1	Improvement compared to detection based on shape index	104
8.2	Protrusion detection	104
8.3	Classification	105
8.4	Evaluation	105
8.5	Challenges for future research	106
	Summary	107
	Samenvatting	109
	Bibliography	111
	List of publications	125
	Curriculum Vitae	127
	Dankwoord	129

1 Introduction

Polyps are small protruding mounds that may develop throughout the intestinal system. The ones that are located in the large bowel, or colon, are referred to as colorectal polyps. Colorectal cancer that may develop from them is one of the most commonly diagnosed type of cancer, responsible for about 12% of all cancer related deaths. For males it is the third ranked cause of cancer related mortality, after lung and prostate cancer. For female it is also ranked third, after breast and lung cancer. In The Netherlands, each year about 9000 people are diagnosed with the disease and more than 4000 die as a result from it. (source: KWF kanker bestrijding).

About 95% of all cases of colon cancers arise from adenomatous polyps, that are initially benign [11]. Due to genetic mutations, such polyps develop from the top layer of epithelium cells that make up the colon surface. By the process of oncogenesis a polyp may evolve stepwise from small tubular adenomas to large adenomas and eventually to carcinomas [17]. That is, due to cell proliferation, the colon wall thickens and bulges out and thereby undergoes a morphological change without which their detection by means of imaging techniques such as computerized tomography (CT) or optical colonoscopy would not be possible.

Polyps are initially not cancerous. Unfortunately, there is a chance of gradual development into malignancy, and, this chance is related to its size. For polyps of about five millimeter in diameter the transition from healthy tissue to malignancy may take up more than ten years. However, for polyps of about ten millimeter in diameter this time span is reduced to five years [71] and for even larger polyps the transition may take place in an even shorter time. The good news is, that, the long premalignant polypoid stadium, offers a time window for screening and removal, and thus, prevention.

It is proposed that polyps with a diameter smaller than 6 mm require no further action, whereas polyps equal to and larger than 10 mm should be removed by colonoscopy [22, 146]. There is debate over the need for polypectomy for 6-9mm polyps. Surveillance for growth with CT colonography has been suggested as a safe alternative [12, 146].

Another concensus is reached on the effectiveness of screening of high risk and symptomatic patient groups, including those with a known hereditary risk for colon cancer. There is, unfortunately, much less agreement on the need for screening of asymptomatic populations [54].

A standard method to examine a patient is by means of optical colonoscopy. This is accomplished by inserting a colonoscope into the anus and then advancing it slowly into the rectum and through the colon. Images of the colon wall are projected on a

video screen or can be observed directly through the scope (figure 1.1). Unfortunately, the technique has a couple of drawbacks. First, there is a small risk for complications. Moreover, direct visual inspection requires the colon to be cleansed. To that end, patients undergo an extensive laxative preparation, which is the main cause for low patient acceptance [123]. Also, polyps may be missed due to the limited opening angle of the camera in combination with the highly structured nature of the colon. In an ideal situation colonoscopy is only applied to patients known to harbor (large) polyps, so that they may be instantly removed.

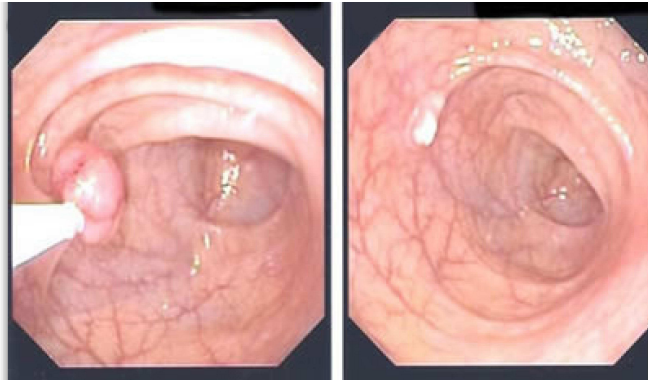


Figure 1.1: Images from inside the colon during a colonoscopy. In the left image a polyp is being removed using a polypectomy snare. The image on the right shows the colon after removal of the polyp.

An innovation which partly avoids these drawbacks is CT colonography, which is a radiological technique which employs computerized tomography. The methods discussed in this thesis apply to the data as obtained by the latter technique. It is briefly described below.

1.1 CT colonography

CT colonography (CTC) was presented in 1994 [131] as a technique with which a 3D image of a patient's abdomen is recorded by a CT scanner. Unlike optical colonoscopy, the technique is non-invasive¹ and does not require sedation. CTC has been studied extensively over the last years [5, 29, 34, 39, 41, 48, 53, 93, 145].

Traditionally, the images are visually inspected by a combination of slice by slice inspection and volume renderings [96]. Two factors hamper the inspection. Segments of the colon may be collapsed and stool may be present in the colon. Stool typically has an attenuation similar to tissue and thus may yield false interpretation of the colon surface

¹Apart from a small radiation dose

1.1. CT COLONOGRAPHY

location. In order to reduce the sensitivity to these artifacts, it is clinical practice to scan patients twice (in prone and supine position). Typical inspection times per patient is about 20-30 minutes [96, 132].

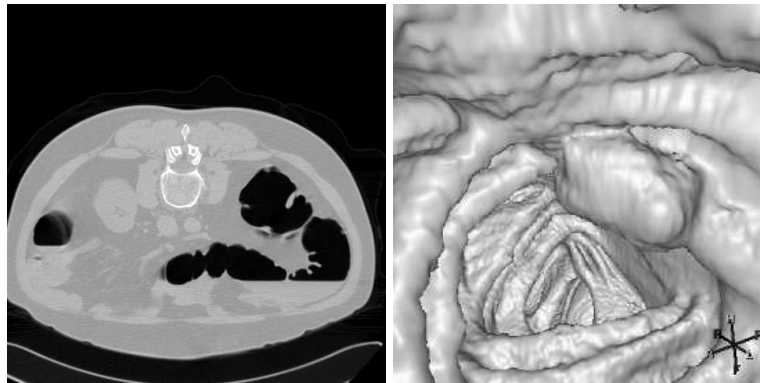


Figure 1.2: Grey image slice through a 3D volume obtained with CT Colonography (left). Isosurface rendering at -750 HU of the colon (right).

An important aspect that is used in deciding on a patient's treatment is the polyp size. It is measured from the largest object diameter in cross sectional views or in volume renderings. For this digital calipers are used (See figure 1.3). (In optical colonoscopy size is also an important aspect, but the technique provides some additional textural information, such as the blood vessel structure that may aid in the decision taking).

Patients harboring polyps larger than or equal to ten millimeter are to be scheduled for optical colonoscopy, such that the polyp can be removed immediately. Such polyps inhibit a large risk for cancer. Smaller polyps, with a size between five² and ten millimeter, inhibit a smaller risk. They are removed nevertheless, in case the patient is scheduled for optical colonoscopy due to the presence of large polyps. Otherwise the patient is rescheduled for a scan at a later time. Polyps with a size smaller than five (or six) millimeter are considered to inhibit a small risk and they cannot be found with high sensitivity and specificity. They are usually ignored.

The main drive behind the development of CT colonography compared to optical colonoscopy is the expected higher patient acceptance [122]. It was concluded that patients with an increased risk for colorectal cancer preferred CT colonography above colonoscopy even if there was a 20% chance for subsequent colonoscopic investigation. A role of CT colonography in screening is to pre-select patients with polyps such that only patients with polyps are sent to colonoscopy. Another advantage of CT colonography is that it aids colonoscopy by localizing the lesion and hence increasing the overall sensitivity.

²Sometimes a size of six millimeter is used.

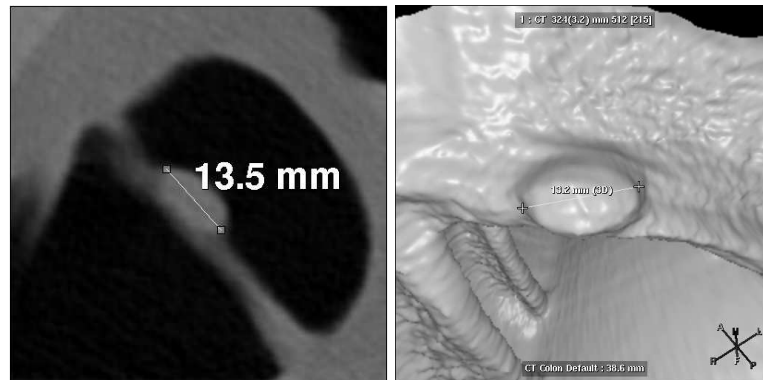


Figure 1.3: Manual size measurement in CT Colonography using digital calipers. The operator selects two points on each side of the polyp. The distance between the points is used as a measure for polyp size. The size is either measured in 2D cross sectional views (a) or in 3D isosurface renderings (b).

The sensitivity and specificity of colonoscopy and CT colonography was assessed in a number of studies [5, 29, 33, 34, 39, 41, 48, 53, 93, 145]. The performance of CT colonography is generally compared to a golden standard obtained with colonoscopy. It is estimated that CT colonography has a sensitivity of around 85-95% [5, 122]. A large study on 3000+ patients from an increased risk population concludes that CT colonography and colonoscopy have a similar sensitivity [55].

A drawback of manual inspection of CT data is the large amount of data which has to be analyzed. A typical CT dataset consists of 250-700 slices (depending on scanner type and settings). As discussed above, it is clinical practice to scan patients twice, in prone and supine position, to increase sensitivity, but simultaneously doubling the amount of data. The development of CT colonography for large screening programs preferably involves automation to facilitate the large amount of data. That is, a Computer Aided Detection System (CAD) is needed to detect suspicious sites on the colon surface. Ideally, instead of inspection of the full colon only a few suspicious sites are presented to the radiologist.

This thesis discusses several aspects of such a system.

1.2 Automated detection

A number of key issues are associated with the automation of polyp detection. Similarly to manual inspection, patient preparation is important. The use of extensive laxative preparation may remove most fecal remains, however, the ones that remain are often difficult to distinguish from polyps or tissue. This may be resolved by the use of tagging. Tagging has the advantage that fecal remains and fluid are easily discriminated

from tissue by means of their intensity (which is increased). However, for automated detection, the use of a contrast agents introduces the risk that polyps are fully or partially covered by (bright) fluid, which hampers its detection. In such situations, digital techniques [99] are required to remove the tagging from the images before applying detection algorithms. The techniques presented in this thesis are intended to be applied to data that have been digitally cleansed first. The multicenter validation (chapter 7), includes results obtained from data without applying digital cleansing.

Another issue, specific to automated detection, is false detections on either the ileocecal valve, or the rectal tube (with insufflated balloon-tipped catheter), which both have characteristics similar to polyps. This thesis has ignored the issue of detections on ileocecal valve. It is addressed in a number of other studies [68, 110]. In chapter 7 the issue of false detections on the rectal tube is discussed.

Currently, there is an ongoing debate [32, 38, 89, 92, 102, 119] on the prevalence and clinical significance of so called 'flat' polyps. The term 'flat' usually defines elevations less than 1 cm in diameter with a polyp height that is less than half of its width and which have a plaguelike morphology. Because these lesions are generally less conspicuous than polypoid lesions, they can be more difficult to detect both in optical colonoscopy as well as CTC. In [89] it is argued that, although flat lesions remain a diagnostic challenge they do not represent a major drawback to widespread CTC screening. In [102] it is argued that "completely flat lesions are exceedingly rare". This thesis does not specifically address the detection of flat polyps, but, the techniques developed in chapters 4 and 5 are designed to detect any elevations from surrounding surfaces before ordering them based on size and intensity measures. As such flat polyps may be detected by techniques proposed in this work.

The aim of automated detection of colorectal polyps is to reduce inspection time, without sacrificing detection sensitivity and specificity. This is achieved by presenting to the expert, only the most suspicious sites, allowing the expert to skip inspection of the obvious polyp free parts of the colon. In this context, two distinct roles for Computer Aided Detection (CAD) are acknowledged [39, 83, 86, 116]. In situations where CAD is 2nd reader, after a human expert, the sites are to be presented in an orderly fashion, such that the most prominent ones missed by the expert are shown first. In the other case, where CAD is to be first reader, an absolute measure of 'polypness' is required, instead of a relative ranking. Additionally, as a first reader, information is required, which allows performing a diagnosis. That is, information which relates to the chance to develop cancer, such as a measure for the polyp size.

The polyp detection pipeline consists of three steps (figure 1.4): **segmentation** of the colon wall from the CT images; **detection** (and segmentation) of suspicious sites (polyp candidates); and ranking followed by **classification** [43, 103, 107, 141, 143].

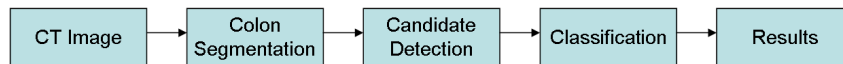


Figure 1.4: Three steps of a typical detection scheme.

The first step is relatively simple due to the high contrast between tissue outside and air inside the colon which allows for a segmentation of the colon from the 3D CT volume by a fixed threshold. The threshold is roughly set to halfway the value for tissue and air. User interaction is sometimes used to discard air in the small intestines, lungs or other air containing organs.

Proper patient preparation is important to avoid two main causes for failing segmentation. First of all, remaining stool has very little contrast with tissue and, if present in the colon, may lead to false positive detections or polyps submerged in stool may be missed. The use of contrast agent to tag stool present in the colon is common practice nowadays. Advanced segmentation techniques were proposed to sustain 3D viewing [98]. Secondly, proper distention is vital to avoid collapsed segments. This thesis does not further address the issue of colon segmentation. The reader is referred to [69].

The second step, the detection of suspicious sites on the colon wall, is performed to discard large parts of colon wall, which are 'obviously' non polypoid. The aim of this step is to retain a high sensitivity. The specificity may be rather low, as it is to be improved in a subsequent supervised pattern recognition step. A large number of methods have been proposed for candidate detection. The most common approach is to focus on the characteristic protruding shape of polyps. It involves measures that describes the local shape of the colon wall and they are often compared to values expected for polypoid shapes[20, 49, 56, 75, 76, 82, 103, 111, 143, 147]. Chapters 2 and 3 of this thesis contribute to such an approach.

Others have looked at probabilistic models [70] or statistical methods to analyze template similarity measures [35]. A number of papers focus on analyzing the deformation properties of the colon wall [1, 62, 63, 130] and again others have proposed to incorporate wall thickness measures [79].

In the chapters 4 to 6 a novel approach is proposed which, instead focuses on the amount of colon surface displacement due to polyp growth.

The third step of the detection pipeline involves the discrimination of the candidate regions into polyps and non polyps. A number of papers proposed to include a processing step to reduce the number of false positive findings in the previous step [67, 72] or have looked into false positive reduction by focussing explicitly on detection of the ileocecal valve and or rectal tube [7, 67, 110]. The most successful strategy seems to be one, involving supervised pattern recognition using gold standard expert labeling of CT data sets, for which both an optical colonoscopy and CT colonography ground truth is available [50, 52, 113].

1.3 Thesis organization

This thesis focuses on the second and third step of the detection pipeline: finding candidate sites followed by ranking and classification.

The analysis of the colon surface shape requires the computation of image derivatives. Their measurement is particularly complicated due to the highly folded structure. This prevents the use of large filter kernels. In **chapter 2** a technique is proposed that uses normalized convolution with a specially devised weighting term in order to optimize the trade off between noise suppression and structure induced bias.

A new algorithm for segmentation of polyps is presented in **chapter 3**. It operates on a triangulated isosurface and takes into consideration the local mesh orientation and vertex position. The algorithm starts with and expands an initial seed, located somewhere on the protruding surface. Based on the resulting segmentation the algorithm estimates the size of the object.

Using a similar explicit triangulation of the colon surface, in **chapter 4** a new method is proposed for the detection of candidate sites. It is based on the notion that polyps have introduced a deformation to the colon surface. The method estimates the colon surface dislocation and candidate sites are obtained by selecting the regions with a dislocation larger than an optimized threshold. The method is tested in the context of a supervised classification scheme, based on features obtained from the deformation field and the grey level image.

Chapter 5 proposes a method that is based on the same principle. Complementary to the previous approach, this method operates directly on the grey level voxels, rather than a triangulated isosurface. It is proved, that the second principle curvature is sufficient to estimate the amount of deformation. A classification scheme based on linear logistic regression is proposed that explicitly keeps large polyps away from the decision boundary. Again, the method is assessed in the context of a supervised classification scheme.

In **chapter 6** a polyp segmentation method is evaluated. Unlike the method from chapter 3, the segmentation is obtained directly from the deformation field. The performance of the method is assessed by comparison to expert size measurements on phantom data and true polyps.

Candidate detection typically renders a lot of candidates to sustain maximum sensitivity. Hence, the number of objects from the target class (polyps) is relatively low. This large imbalance of the prevailing classes typically hampers classifier design and training. Furthermore, the classifier should take into account the increased clinical relevance of larger polyps. The last **chapter 7** in this thesis discusses the consequences of these characteristics for the design of the classification system. A novel, low-complex, classification system is proposed that orders the polyps according to clinical relevance. This chapter also serves to demonstrate the overall performance of a CAD system based on the techniques presented in this thesis.

2 On normalized convolution for the measurement of image derivatives in highly structured surroundings

This chapter discusses the trade-off between noise reduction while retaining the image structures when computing image differentials. A scheme is presented which allows to incorporate confidence values into the measurement. This scheme is evaluated with an application for finding protruding regions in 3D CT images of the human colon from differentials up to second order.

Based on:

C. van Wijk, R. Truyen, R.E. van Gelder, L.J. van Vliet, F.M. Vos, *On normalized convolution to measure curvature features for automatic polyp detection*, MICCAI 2004 [128]

2.1 Introduction

Noisy data asks for a certain amount of regularization, whereas thin or small image structures require a very small filter kernel. Violating the first requirement yields a noisy result (stochastic error) whereas violating the second causes a substantial bias in the derivatives (systematic error). For example, both errors hamper the curvature measurement in 3D CT images of the human colon in which case finding a trade-off between the conflicting requirements is very difficult due to the presence of small folded structures on the colon wall of only a few voxels wide (see Figure 2.1).

In this chapter we present a novel method to adapt the size and shape of the filter kernels to the local image data. The method avoids the systematic error due to mixing of nearby image structures and is optimized for noise reduction. However, using irregular shaped filter kernels requires a space-variant normalization of the derivative filters. Therefore we present an intuitive framework for deriving normalized differential convolution of arbitrary order (Section 2.2.1). In section 2.2.3 we present a scheme to compute space-variant kernels from the local image structure.

Derivatives in 3D images can be computed by convolution with derivatives of Gaussian kernels. In order to adapt the Gaussian (derivative) kernels to the local geometry they are multiplied with a confidence function which is extracted from the local image structure. This additional weighting requires re-normalization as well as a (re-)orthogonalization. The technique which takes care of both is normalized convolution ([28, 61])

The performance of the new method is assessed on both simulated as well as CT data. For the detection of polyps the resulting image derivatives can be combined into principal curvatures, κ_1 and κ_2 (Thirion and Gourdon [120]). Based on the principal curvatures a number of polyp detectors can be constructed. Yoshida [141] uses primarily the shape index and curvedness. The shape index is given by $SI = \frac{1}{2} - \frac{1}{\pi} \text{atan}\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right)$ and the curvedness is given by $CV = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}}$.

2.2 Methods

2.2.1 A least squares approach to normalized convolution

The following assumes a 2D image (extension to 3D image space is straightforward). Consider a local neighbourhood of $N \times N$ pixels f_i that is modeled by a Taylor expansion around the center of the local neighbourhood (indicated by 0):

$$f_i = I(0) + x_i I_x(0) + y_i I_y(0) + \frac{x_i^2 I_{xx}(0)}{2!} + \frac{y_i^2 I_{yy}(0)}{2!} + \frac{2x_i y_i I_{xy}(0)}{2!} + R(i) \quad (2.1)$$

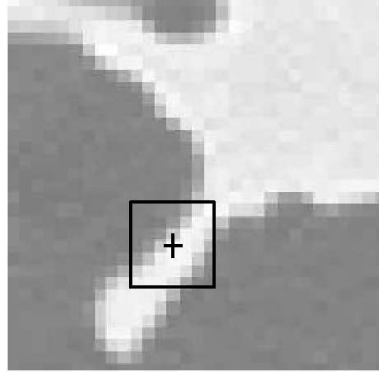


Figure 2.1: Small kernel overlaying multiple image structures

in which I indicates the 'true', underlying image function and i is a linear index. Using terms up to the second order and substituting $\eta_1 = I(0)$, $\eta_2 = I_x(0)$, ..., Equation 2.1 is rewritten as:

$$\begin{bmatrix} f_1 \\ \dots \\ f_{N^2} \end{bmatrix} \approx \begin{bmatrix} 1 & x_1 & y_1 & 0.5x_1^2 & 0.5y_1^2 & x_1y_1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{N^2} & y_{N^2} & 0.5x_{N^2}^2 & 0.5y_{N^2}^2 & x_{N^2}y_{N^2} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \dots \\ \eta_6 \end{bmatrix} \quad (2.2)$$

The local neighbourhood can be depicted as a point in an N^2 -dimensional space spanned by the orthonormal basis $\{\mathbf{e}_i\}$. A new set of basis vectors $\mathbf{b}_j = \{\mathbf{1}, \mathbf{x}, \mathbf{y}, \frac{\mathbf{xx}}{2}, \frac{\mathbf{yy}}{2}, \mathbf{xy}\}$ are the basis functions of the Taylor expansion (i.e. the columns of the matrix in Equation 2.2). Thus, $\{\eta_1, \eta_2, \eta_3, \eta_4, \eta_5, \eta_6\}$ are the coordinates of the signal on the new basis and directly yield the first and second order derivatives. It can be stated that:

$$f_e^i = \mathbf{B}\eta_b^j + \mathbf{r} \quad (2.3)$$

Equation 2.3 merely rewrites Equation 2.2, implying that the signal f on basis \mathbf{e}_i is approximated by the so-called basis tensor \mathbf{B} times the coordinates of f on basis \mathbf{b}_j , (η_b^j) , with a residual \mathbf{r} . It must be emphasized that, in general, the basis functions can be freely selected and need not be orthonormal. Our basis was merely chosen to comply with the Taylor expansion. The objective now is to find the new coordinates η_b^j by minimizing the error $\varepsilon = \|f - \mathbf{B}\eta\| = (f - \mathbf{B}\eta)^2$. The result is the general least squares solution to 2.3:

$$(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T f_e^i = \eta_b^j \quad (2.4)$$

with $(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T$ the pseudo-inverse of \mathbf{B} .

To reduce the influence of points further away from the neighbourhood center we multiply the set of equations in eq. 2.2 by a (rotation invariant) matrix $\hat{\mathbf{A}}$, with $\mathbf{A} = \hat{\mathbf{A}}^T \hat{\mathbf{A}}$. The $N^2 \times N^2$ diagonal matrix \mathbf{A} contains spatial weights and is called the applicability

function.

$$\hat{\mathbf{A}}f_e^i = \hat{\mathbf{A}}\mathbf{B}\eta_b^j. \quad (2.5)$$

Multiplication with $\hat{\mathbf{A}}$ is allowed as long as it does not yield a singular system of equations. Similarly, each equation in (2.5) can be multiplied again by other weights. It is now clear how confidence levels assigned to each neighbour can be incorporated. The result is a double weighted least squares solution:

$$(\mathbf{B}^T \mathbf{A} \mathbf{C} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{A} \mathbf{C} f_e^i = \eta_b^j \quad (2.6)$$

with the diagonal matrix $\mathbf{C} = \hat{\mathbf{C}}^T \hat{\mathbf{C}}$ holding the confidence value of each neighbor.

2.2.2 Normalized convolution and Gaussian derivatives

In the previous section the signal was expanded using a Taylor polynomial. However, the choice for the basis is not limited to it. The advantages of the Taylor expansion is that image derivatives are obtained directly from the coefficients (η). Other basis functions are also possible and one in particular leads to the well known Gaussian n-jet.

The Gaussian kernel is given by the functional

$$G(x, \sigma) = e^{-\frac{x^2}{2\sigma^2}} = e^{-\hat{x}^2} \quad (2.7)$$

with $\hat{x} = \frac{x}{\sqrt{2}\sigma}$

The n-th Gaussian derivatives can be written as the Gaussian function times a Hermite polynomial of order n. One aspect of Hermite polynomials is that they belong to a family of orthogonal functions on the infinite interval $(-\infty, \infty)$ with the weighting function e^{-x^2} . Therefore, using Hermite polynomials, together with an applicability function given by $\hat{A} = e^{-\frac{x^2}{2}}$, $A = \hat{A}^T \hat{A}$ leads to kernels given by:

$$(\mathbf{B}^T \mathbf{A} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{A} \quad (2.8)$$

The matrix $\mathbf{B}^T \mathbf{A} \mathbf{B}$ is diagonal as each element is given by:

$$\sum_x H_i(x) A H_j(x) = 2^{i+j} i! \sqrt{\pi} \delta_{ij} \quad (2.9)$$

One consequence of the orthogonality of Gaussian basis functions is, the the projections of the signal onto each basis function can be computed independently. Additionally the term $\mathbf{B}^T \mathbf{A}$ give the well know Gaussian derivative kernels: $H(\hat{x})_n e^{-\hat{x}^2}$.

Unfortunately, when the the confidence values assigned to the local neighborhood are not constant, the orthogonality is lost and the inversion in equation 2.8 is not trivial. It needs to be recomputed for each local neighborhood separately.

2.2.3 Local confidence values

The framework presented in the previous section accommodates normalized space variant-kernels. The confidence values which are inserted into the regularization are computed locally and will adapt the kernel to the local image structure. The goal is to assign high confidence to voxels on the image structure under consideration and a low confidence to other structures. Such structures might be neighboring folds, changes in tissue structures, the opposite side of a fold, etc.

We propose the following scheme to compute the confidence values.

1. Segment the air to find the air-tissue interface. Usually this is achieved by simple thresholding. We use a dynamic threshold [9] to allow for a correct segmentation of geometries affected by partial volume effects.
2. Compute for all voxels the distance to the air-tissue interface. We perform two distance transforms. One to compute the distance to air. From this we subtract a second distance transform, the distance to tissue. This operation results in positive values for air and negative values for tissue. On the colon wall the values of the distance transform are zero.
3. Compute the gradient of the distance transform which will act as a normal vector field. We will use these normals to distinct between different structures.

Steps 1 to 3 can be computed for the entire image at once. In contrast the following step is a local one to be incorporated in the convolution process. To distinguish between different geometries one can remark that the surface normal of the structure under consideration will differ from that of the direct neighboring structures.

1. Assign neighboring voxels to belong to the current structure by taking the inner product of the normal at the neighborhood center with the normal of a neighbor. A threshold on this value (e.g. >0) classifies the neighbor and sets its confidence value to zero or one.

An example of a region selected by the above scheme is given in Figure 2.2 . Note that the confidence values are weighted with the applicability function in the regularization process.

2.3 Results

The performance of the space-variant filtering is assessed on both simulated objects as well as CT data. Two test images were created to test the computation of radii of curvature with both the isotropic method as well as the new method. The first image,

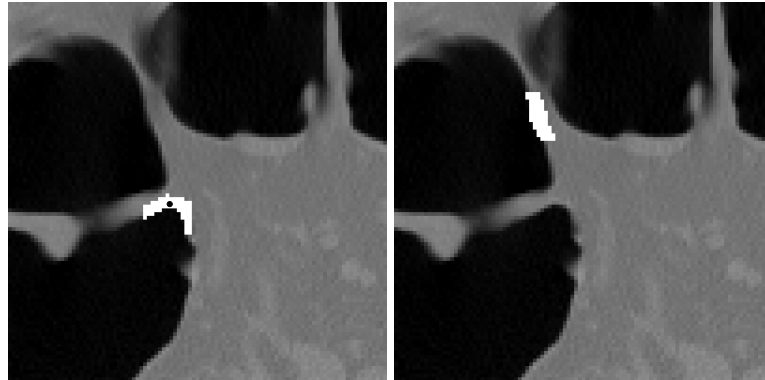


Figure 2.2: *Constructed confidence values for a neighborhood of 13^2 pixels. The neighborhood center is indicated by the black dot.*

displayed in Figure 2.3a, is a 3D cylinder (only cross-section shown) which was constructed using the error function with a σ of 2. The cylinder has a radius of 18 pixels. Gaussian noise was added to the images. The standard deviation of the noise was 4% of the contrast (intensity difference between air and tissue). The second test image contains two 3D cylinders, their centers separated by 40 pixels. The image was constructed by multiplying two separate cylinder images after which noise was added.

Figure 2.3 shows that noise affects the derivative computation at small scales (a and b). Increasing the (isotropic) scale of the operator improves the results (c), but adjacent structures inside the footprint of the filter spoil the final result (d).

The isotropic Gaussian derivative filtering fails to return the correct curvatures. In this paper we propose to improve the curvature measurement by introducing space variant kernels. The performance is compared to the isotropic method in Figure 2.4a. The result of the isotropic method are repeated on the left cylinder. The results obtained with the proposed method are plotted on the right cylinder. It is clear that in the region where the two cylinders are close together the method using isotropic kernels fails to give correct curvature values, while the new method returns correct results

The new method does not suffer from the systematic error introduced when using isotropic filters. The cost is a small increase in a stochastic error due to the fact that the incorporation of confidence levels into the filtering in effect reduces the number of voxels used to suppress noise. However, the specific choice of confidence levels based on the local structure allows to discard just those voxels which would have introduced a systematic error. In other words our method optimizes the trade-off between noise reduction and preservation of image structure.

The shape index is computed from the principal curvatures and is often used to select polyp candidates by means of thresholding. Applying such classification to the image in Figure 2.4a yield Figure 2.4b. The isotropic method will result in a classification of a

2.3. RESULTS

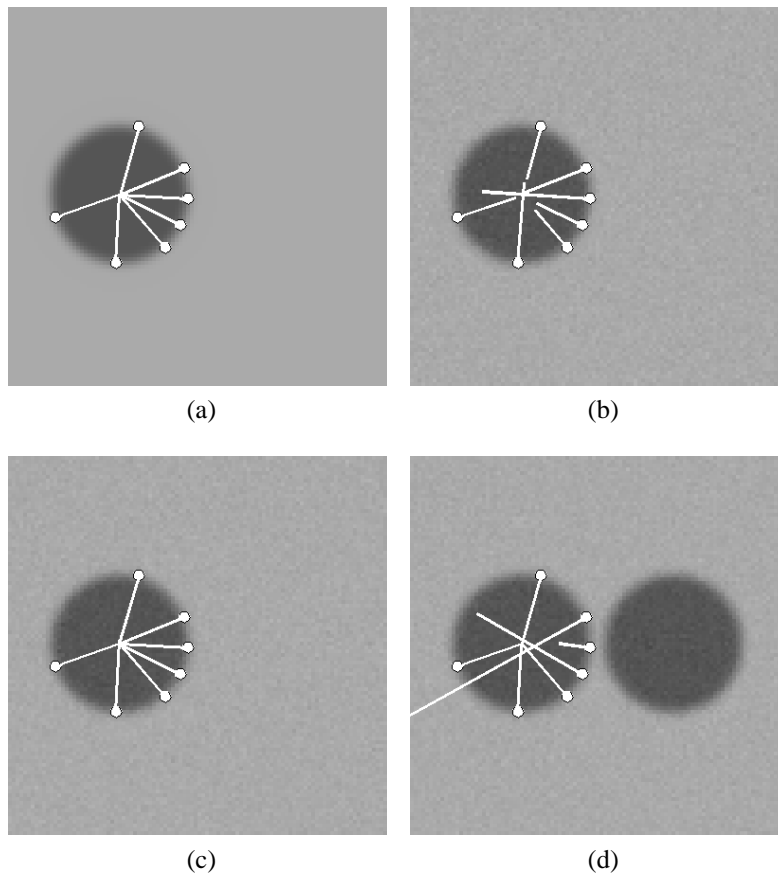


Figure 2.3: *Trade-off between noise suppression and resolution. On several positions on the edge the normal direction (line direction) and radius of curvature (line length) are plotted. From left to right: (a) noise free image, small scale $\sigma = 1$. (b) Gaussian noise added, $\sigma = 1$. (c) computation at larger scale suppresses the noise, $\sigma = 3$. At larger scale ($\sigma = 3$) incorrect curvature and gradient direction are obtained close to neighboring structures (d).*

large part of the cylinder to a ridge-like structure. The new method correctly classifies all the voxels to a rut-like structure (8c).

To demonstrate the performance of the method on CT data, a scheme similar to [76] is applied. Yoshida et al. use the the shape index and curvedness to select the set of polyp candidates. In [76] thresholds were presented for the shape index (between 0.9 and 1.0) and for curvedness (0.05 mm^{-1} and 0.25 mm^{-1}). We applied the same scheme using hysteresis thresholding [76] to investigate the performance with respect to the candidate selection step. Initial test results on a few patients show promising results. The power of the method is clearly demonstrated in Figure 2.5 and Table 2.1 which are obtained by

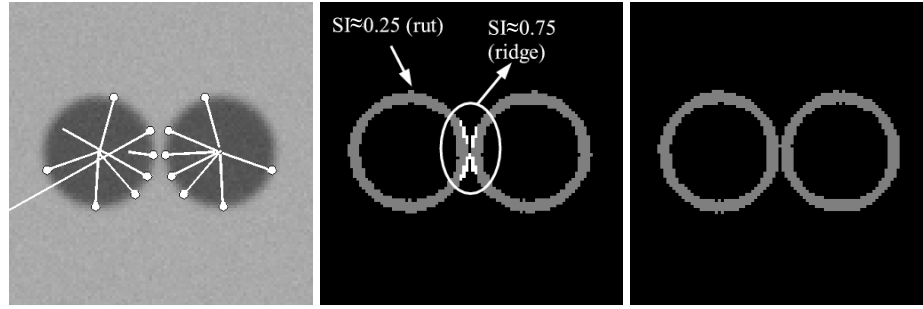


Figure 2.4: *Radius of curvature (left) and shape index (middle, right) computed on two cylinders (Only a cross section of the cylinders is shown). On several positions on the edge the gradient (line direction) and radius of curvature (line length) are plotted. Left cylinder: isotropic method. The gradient direction is obtained using isotropic Gaussian kernels ($\sigma = 3$). Right cylinder: both gradient direction and radius of curvature are obtained with space variant kernels ($\sigma = 3$). The middle image shows the classification by shape index computed by the isotropic method. The isotropic method classifies large part of the cylinders to a ridge like structure. The new method (right) using space variant kernels classifies all voxels correctly.*

applying the method to a small dataset ($200 \times 200 \times 100$ voxels) containing one polyp (approx. 4 mm). The new method detects the polyp and finds one false positive. The isotropic method detects three false positives and misses the true positive. From the demonstration of our method both on simulated data as well as CT data we feel confident that space-variant kernels will yield fewer false positives. Especially for small polyps the new method is likely to increase the sensitivity. However, we are aware that the performance of the operator can only be assessed by statistical validation on a large number of datasets.

id	method	cluster size	label
1	isotropic	9	false positive
2	isotropic	28	false positive
3	isotropic	86	false positive
4	space-variant	31	true positive
5	space-variant	3	false positive

Table 2.1: *Detection results. The isotropic method detects 3 false positives. The space variant method detection the true positive and one (small) false positive. The results were obtained by filtering with $\sigma = 3$.*

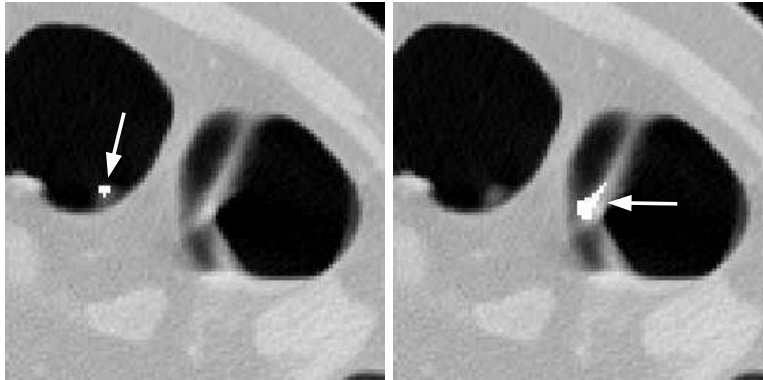


Figure 2.5: *One slice from the 3D Ct dataset. Voxels labelled as belonging to polyps (white). The new method finds the polyp (left). The isotropic method fails to find the polyp and selects a false positive. The results were obtained by filtering with $\sigma = 3$.*

2.4 Conclusions

The measurement of curvature in CT data for the detection of polyps is difficult due to the highly folded colon. Therefore noise suppression with larger isotropic filters is not possible. We have shown that with a specific formulation of normalized convolution using a local Taylor expansion space-variant kernels can be used. In addition we have shown that space-variant kernels can be constructed which discards just those voxels belonging to neighboring image geometries. Thereby the derivative filtering optimizes the trade-off between noise suppression and preservation of local image structure.

The assessment of the method by simulated images shows that the space-variant filtering outperforms isotropic filtering. Also, on CT data the new method seems to indicate a higher sensitivity and higher specificity. However, the authors do realize an investigation on more data is needed to be conclusive on the overall improvement of polyp detection.

3 Segmentation and size measurement of polyps in CT colonography

In this chapter a new method is proposed for the automatic measurement of polyp size. It operates on a triangulated isosurface and takes into consideration the local orientation and position of the mesh. The algorithm starts with and expands an initial seed, located somewhere on the protruding surface. Based on the resulting segmentation the algorithm estimates the size of the object. We assess the performance by comparison to expert size measurements on phantom data.

Based on:

J.J. Dijkers, C. van Wijk, F. M. Vos, J. Florie, Y.C. Nio, H.W. Venema, R. Truyen, L.J. van Vliet, *Segmentation and size measurement of polyps in CT colonography*, MICCAI 2005 [26]

3.1 Introduction

Colorectal cancer is one of the most commonly diagnosed types of cancer. Specifically, the American Cancer Society predicts 145,000 new cases and 56,000 deaths from colorectal cancer for 2005 [3]. Polyps are a well-known precursor to such carcinoma. Not surprisingly, it has been shown that early removal of polyps ensures a decrease in incidence [121].

In recent years, CT colonography has been proposed as a noninvasive alternative to traditional polyp detection by colonoscopy [42, 96]. In CT colonography, the colon structure is often visualized from an endoluminal perspective by means of surface or volume rendering. Recently, methods have been proposed to support the inspection by a computer aided detection (CAD) system indicating suspect locations [106, 141]. The size of a detected polyp is an important aspect for diagnosis and decision making. It is generally accepted that polyps with diameter $< 5\text{mm}$ require no direct further action, whereas larger polyps should be removed via colonoscopy. Typically, the size of polyps is measured in colonoscopy by comparison with an open biopsy forceps. In CT colonography, it is usually measured in reformatted images, in which the largest polyp diameter is selected for size measurement. However, polyp sizes thus measured by human experts can show significant inter- and intra-observer variability.

Clearly, an automated method is needed to enable more accurate measurement of polyp size. As a side effect, such a procedure is also useful in CAD algorithms. Automated polyp detection is usually based on sophisticated pattern recognition techniques that take into account many features measured on tentatively selected candidates (e.g. size, area, average shape index etcetera). Proper segmentation is crucial to perform reliable feature measurement.

The existing methods for colonic polyp segmentation (such as Summers et al. [52, 103] and Yoshida et al. [77]) are especially designed to work directly on the 3D CT data. Such an approach is hindered by not operating on a specifically defined region of interest c.q. the colon surface. Hence, segmentation of polyps which are by definition protrusions of the colon surface is not a trivial task.

In this paper we present a new method for semi-automatic segmentation of polyp-like structures. Additionally, a technique is described to automatically measure polyp sizes using this algorithm. Our method assumes that the colon surface has been identified as a region of interest. Moreover, it is asserted that a candidate location has been identified; in our system by a vertex detection step based on the measured shape index [128]. We will compare the size measurement by our algorithm with that of physicians in a set of phantom objects (in which the size is known a priori).

3.2. POLYP SEGMENTATION

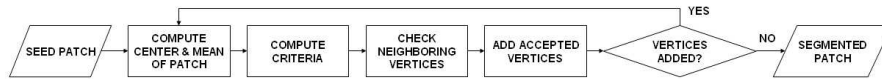


Figure 3.1: Schematic overview of segmentation procedure.

3.2 Polyp segmentation

The description below assumes that the colon wall is described by a triangular mesh. However, the basic ideas of the method are not restricted to a mesh based surface representation of the colon wall; they can as well be implemented to work on a voxel based model. An additional advantage of our method is that the segmentation requires no user input, that is all parameters are drawn from the underlying data.

Outline of segmentation procedure. Ideally, a polyp could be described as a rather spherical, symmetric mound on a background shape (see e.g. Figure 3.2a). One could intuitively delineate a polyp by the inflection points on both sides. However, these points may not be easily identifiable due to the curvature of the background shape (e.g. a fold).

Hence, we model a polyp to have a symmetry axis that goes through the center point (P_c) in which the apical surface normals converge, and the mean position (P_m) calculated from the polyps surface points. The edge of the polyp is defined by the points at which the surface normals tend to deflect from the center point (we will formalize this below).

Initially, a single position or a small patch indicates a point on the polyp candidate [128]. Since the center and mean points may not be robustly determined from such a seed patch, the polyp segmentation procedure is set up as an iterative process. During each cycle of this process neighboring vertices are added if certain criteria are met. The process terminates when no more points are added. An overview of the procedure is shown in figure 3.1.

Computing the center and mean points. As depicted in Figure 3.2a, the surface normals on the polyp apex tend to converge in a center point. This point (P_c) is found by minimizing the sum of the distances (d_i) to all normals (\vec{n}_i). The surface normals are calculated by Gaussian derivatives the underlying 3D CT data at a scale of 2mm. This scale was determined experimentally such that no polyps are missed. The distances can be computed according to:

$$d_i = \|\vec{n}_i \times (P_i - P_c)\| / \|\vec{n}_i\| \quad (3.1)$$

where P_i is a point on the patch and \times denotes the vector outer product. Additionally, a mean point (P_m) is associated with a patch. The position of the mean is simply computed by averaging the positions of all vertices: $P_m = \frac{1}{N} \sum P_i$. The mean and the center points define a centerline (dashed in Figure 3.2b). Henceforth it is called the polyp axis.

Adding points to a seed patch Points are to be added to a seed patch until the local

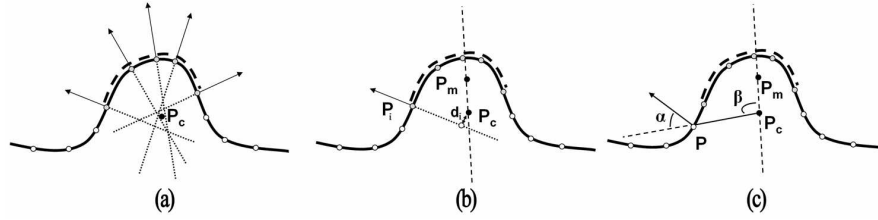


Figure 3.2: Schematic representation of a patch (dashed curve) on the colon wall. Figure (a) shows how convergent normals define a center point; figure (b) shows how the minimized distance d_i is defined for surface point P_i ; figure (c) shows how the angles α and β are defined.

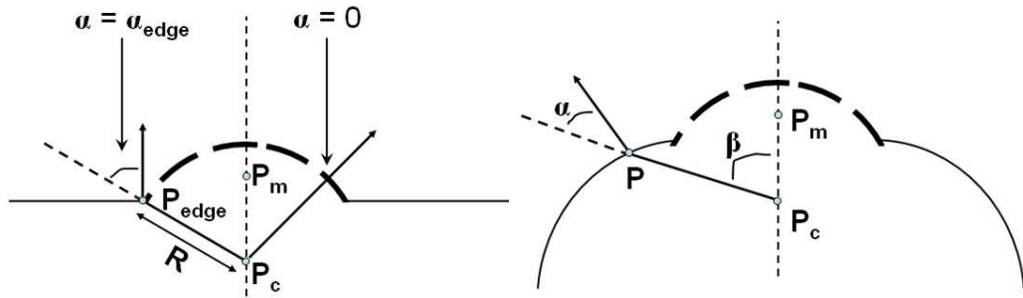


Figure 3.3: Schematic representation of a polyp (dashed curve) on a flat background (left) and on a fold (right).

surface normal tends to deviate from P_c . To formalize the stopping criterion, consider first a sphere on a flat background. Let us define α as the angle between the line from the center point (P_c) to the vertex (P_i) and the normal at the position of the vertex (see Figure 3.2c). Clearly, on top of the polyp α is small (exactly zero on a spherical cap, see Figure 3.3). The angle α increases while moving to periphery of the polyp. Right outside the polyp the angle is given by (compare with Figure 3.3):

$$\alpha_{edge} = \arccos \left[\frac{(P_{edge} - P_c) \cdot \vec{n}}{\|P_{edge} - P_c\| \cdot \|\vec{n}\|} \right] \quad (3.2)$$

in which P_{edge} is defined as in Figure 3.3 and \vec{n} is the normal at point P_{edge} . We assume that the ideal threshold-value lies somewhere between these extreme values (respectively 0 and α_{edge}). The required midway point is closely approximated by the angle calculated via (compare with Figure 3.3):

$$\alpha_{mid} = \arccos [(P_m - P_c) / R] \quad (3.3)$$

Thus, $\alpha < \alpha_{mid}$ yields a safe stopping criterion for adding neighboring vertices to a polyp on a flat background. On a fold, however, the angle α remains small (see Fig.

3.3. EXPERIMENTS AND RESULTS

3.3). Let us define β as the angle between the polyp axis and the line between the vertex and the center point (as in Figure 3.2c):

$$\beta = \arccos \left[\frac{(P - P_c) \cdot (P_m - P_c)}{\|P - P_c\| \cdot \|P_m - P_c\|} \right] \quad (3.4)$$

At the edge of the polyp β is given by $\beta_{edge} = \alpha_{edge}$. Typically, β continues to increase while moving onto the fold. Consequently, $\beta < \beta_{edge}$ yields a logical stopping criterion for a polyp on a fold. It should be noticed that the two posed criteria are mutually exclusive: the sidepoints of a polyp on a fold do not meet the criterion of $\alpha < \alpha_{mid}$. On the other hand, points besides a polyp on flat background do not fulfill $\beta < \beta_{edge}$. Also, the angles α_{mid} and β_{edge} are both dependent on the shape of a polyp. Flatter polyps tend to have lower values for α_{mid} and β_{edge} than more protruding polyps. In other words, the threshold values automatically depend on the polyp shape.

All vertices neighboring a seed patch that match the conditions are accepted and added at once to yield a new seed. Consequently, the outcome does not depend on the order in which points are processed. Clearly, if none of the vertices match the criteria, no points are added and the current patch is considered the final, segmented polyp. Otherwise, all steps are iterated.

Automated size measurement. The size measurements for polyps are based upon the segmented patches. The edges of these patches are projected along the polyp axis onto a plane. An ellipse is fitted to these points in 2D space by computation of the first and second order moments. This is in accordance with the current medical practice in the Academic Medical Center where the polyp size is characterized by its largest diameter.

3.3 Experiments and results

The performance of the method was assessed by comparing the automated size measurement with those of radiologists using scans of a colon phantom. We have looked into several aspects to test our approach:

- Inter-observer variability of radiologists
- Intra-observer variability of radiologists and our method
- Accuracy and precision of the radiologists and our method

Experimental data. All data was acquired using a Mx8000 multislice CT-scanner (Philips Medical Systems, Best, the Netherlands) using the same scanning protocol for all scans (scan parameters: 120 kV, 100 mAs, 4 x 2.5 mm collimation, pitch 1.25, standard reconstruction filter, and a 180o interpolation algorithm).

Object	Lucite Objects (lengthxheight) [mm]	Plasticine Objects (lengthxheight) [mm]
1	10.0 x 5.0	19 x 9
2	10.0 x 2.5	17 x 8
3	8.0 x 4.0	14 x 10
4	8.0 x 2.0	14 x 8
5	6.0 x 3.0	12 x 8
6	5.0 x 2.5	11 x 11
7	4.0 x 2.0	11 x 5
8	-	6 x 5

Table 3.1: Dimensions of the phantom objects; of all lucite objects two specimens were used.

The phantom consisted of a lucite cylinder into which fabricated polyps were inserted. At first, the phantom contained 10 hemispherical lucite objects of various sizes, and 4 objects with reduced height (2 mm) in order to mimic flat lesions. Subsequently, 8 asymmetric objects from plasticine were inserted in the phantom (maximum width 6-19 mm). The size of all objects (see Table 3.1) was measured by sliding calipers. The phantom was placed in a cylinder, 34 cm in diameter that was filled with water to arrive at a signal to noise ratio comparable to that in patient data. The two phantoms with lucite and plasticine polyps respectively were scanned twice: in the axial plane of the cylinder, and an orthogonal plane (see Figure 3.5).

The size of all objects was measured in the CT scans by two radiologists and by our automated method. The radiologists measured the objects in multiplanar reformatted CT images. Each object was measured twice, along the main axes as perceived by the physician. The largest value was taken as the polyp size. The automatic measurements were done as described previously. For that purpose, an arbitrary seed point was manually indicated somewhere on the polyp surface.

Inter-observer variability of the radiologists. Figure 3.4a,b contains the graphs displaying the measurements of one radiologist against those of the other. Clearly, radiologist A tends to measure larger diameters compared to radiologist B. The average difference of their measurements was 1.2mm for the lucite objects and 3.1mm for the plasticine objects. The standard deviation of the absolute difference was 0.7mm for the lucite objects and 2.7mm for the plasticine objects.

Intra-observer variability of the radiologists and the automatic method. The intra-observer variability is assessed by the difference in size measured in the axial scan versus the measurement on the same object in the orthogonal scan. It must be conceded

3.3. EXPERIMENTS AND RESULTS

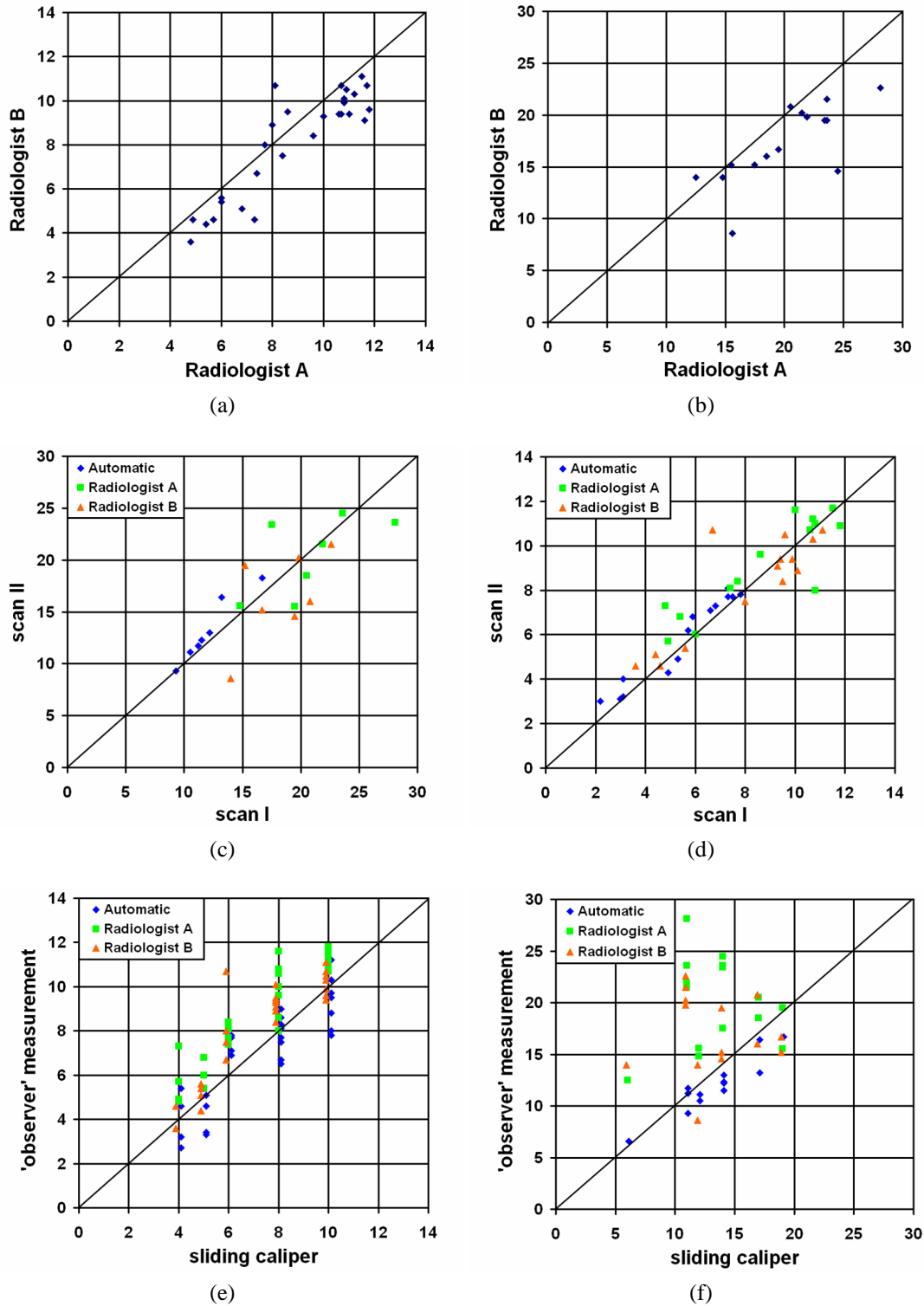


Figure 3.4: Inter-observer variability in polyp size for lucite (a) and plasticine (b) objects. Intra-observer variability in polyp size for lucite (c) and plasticine (d) objects. Accuracy and precision of size measurement for lucite (e) and plasticine (f) objects. 25

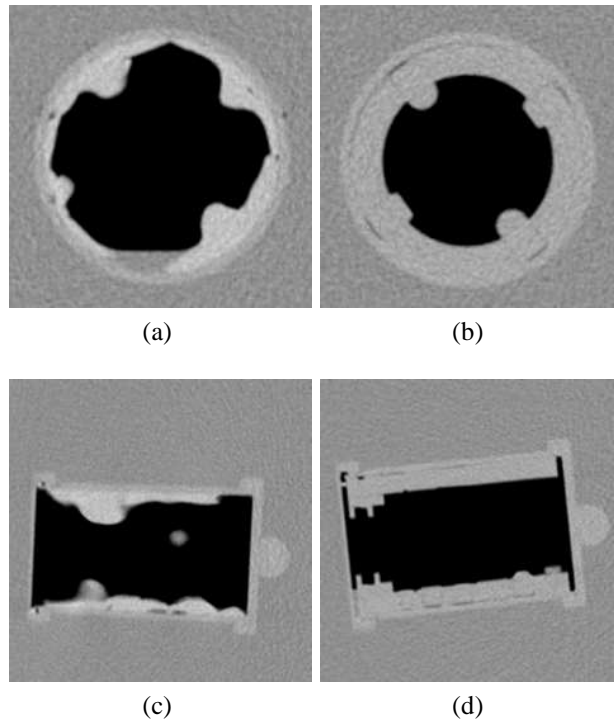


Figure 3.5: Single slice from each scan. From a-d: lucite polyps axially, plasticine polyps axially, plasticine polyps orthogonally and lucite polyps orthogonally.

that this involves two different scans of the same object. We assume, however, that it allows for a good approximation of the intra-observer variability.

Figure 3.4c,d shows the measurements of the observers in one scan versus the measurement in the other. Apparently, neither the radiologists nor the automatic method shows a bias. The average absolute difference between the two measurements on the lucite objects was 1.0mm and 0.8mm for the two radiologists and 0.5mm for the automatic technique. The standard deviation of the absolute difference was 0.9, 1.0 and 0.5mm respectively. For the plasticine objects the average absolute differences were 2.6, 3.2 and 1.1mm and the standard deviations 2.2, 2.1 and 1.1mm respectively.

Accuracy and precision. The accuracy and precision of the observers is defined by comparison to sliding calipers (see Figure 3.4e,f). The measurements of both radiologists appear to be slightly biased. This can be explained by the procedure of always selecting the larger of two measurements as the size of the polyp. The bias in the automatic method is less pronounced, but not completely absent. It can be explained by

3.4. CONCLUSIONS & FUTURE WORK

	Lucite		Plasticine	
	mean diff. [mm]	std. dev. [mm]	mean diff. [mm]	std. dev. [mm]
Rad. A	1.6	0.9	7	5
Rad. B	0.8	1.0	4	5
Automatic	-0.2	1.2	-1	1.2

Table 3.2: Mean difference and standard deviation of difference between observers and sliding calipers measurements for the lucite objects and the plasticine objects.

notifying that all points (except due to noise) on a segmented hemispherical polyp surface are projected inside a circle with the diameter of the polyp. An ellipse fitted through the contour points yields a small underestimation of the true size. Clearly, one might correct for all these biases in a calibration step. Specifically noticeable, is the higher precision of the automatic method on the plasticine objects, indicated by the smaller spread of values around the line of identity.

As shown in table 3.2, the automatic method shows a smaller systematic error than the radiologists. There is no significant difference in the precision (std. dev.) for the (symmetric) lucite objects between the automatic system and the radiologists. However, for the irregular plasticine objects the precision of the automatic system remains the same, whereas the precision of the radiologists decreases significantly.

3.4 Conclusions & future work

The size of a colonographically detected polyp is important for diagnosis and decision making. The size measurement by human observers is generally considered to be imprecise and inaccurate. In this paper we presented a method for the automatic segmentation of polyp-like structures. The polyp size was automatically derived from the segmentation result. It was shown that our algorithm yields a smaller bias than the measurements from radiologists: on average 1mm or less for the automatic method and between 1 and 7mm for the radiologists, depending on the irregularity of the object. Even more important, the algorithm is consistent irrespective of the polyp shape. As opposed to that, the radiologists show a four times larger variation for the irregularly shaped objects. It is this irregularity which occurs in practice.

A good polyp segmentation algorithm is also useful for automatic polyp detection algorithms. It allows for extraction of features such as volume, surface area, average grey-value etcetera. Such features may improve the specificity of CAD algorithms.

4 Detection of protrusions on curved folded surfaces applied to automated detection in CT colonography

In this chapter a new method is proposed for the detection of polyp candidate sites on the colon surface. It is based on the notion that polyp tissue growth introduces a local deformation of the colon surface. The method estimates the original 'undeformed' surface position by solving a nonlinear partial differential equation. Candidate sites are obtained by comparing the two surfaces. The method is assessed by a supervised classification, based on features obtained from the deformation field and the grey level CT image.

Published as:

C. van Wijk, V.F. van Ravesteijn, F.M. Vos, R. Truyen, A.H. de Vries, J. Stoker, L.J. van Vliet, *Detection of protrusions in curved folded surfaces applied to automated polyp detection in CT colonography, MICCAI 2006 [130]*

4.1 Introduction

CT colonography is a modern, noninvasive method to inspect the large bowel. It enables to screen for colorectal polyps by way of images rendered from an endoluminal perspective. Polyps are well-known precursors to colon cancer. The size of a detected polyp is an important indication for diagnosis and decision making. It is generally accepted that polyps with a diameter smaller than 5mm require no direct further action, whereas polyps larger than 10mm should be removed. The policy with patients harboring polyps with a size between 5mm and 10mm is to have a follow-up CT-scan several years later.

Unfortunately, current colonographic visualization techniques are still rather time consuming. More important, large polyps are sometimes missed. Therefore, methods have been proposed to support the inspection by way of computer aided diagnosis (CAD). A large number of such schemes are proposed in the literature [2, 36, 47, 60, 79, 103, 107, 140]. Like most CAD systems, automated polyp detection usually consist of three basic steps: (1) segmentation of the colon wall; (2) candidate generation and (3) supervised pattern recognition.

A good approximation to the true colon wall (defining the region of interest) is obtained rather easily due to the large contrast between tissue and air/ CO_2 inside the colon. However, partial volume effects may affect the image intensity at thin colonic folds. Still, most techniques use a thresholding at a fixed value of around -650 Hu.

Finding candidate objects on the colon surface is a much more challenging task. Summers et al. [103, 104, 107] propose to use methods from differential geometry. In [104] a triangle mesh is extracted from 3D CT data after which principal curvatures were computed by fitting a 4th order b-spline to local neighborhoods with a 5 mm radius. Candidates were generated by selecting regions with a positive¹ mean curvature.

Yoshida et al. [140, 141] use the shape index and curvedness to find candidate objects on the colon wall. The shape index and curvedness are functions of the principal curvatures κ_1 and κ_2 :

$$SI = \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right) \text{ and } CV = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}} \quad (4.1)$$

and are computed using a Gaussian-shaped window (aperture).

Alternatively, Kiss et al. [58] generate candidates by searching for convex regions on the colon wall. Their method fits a sphere to the surface normal field. The side on which the center of the fitted sphere is found (in tissue or in air) determines the classification of the surface as convex or concave. Roughly 90% of the colon wall is classified as concave, that is as 'normal'. To the remaining part of the colon surface a generalized Hough transformation is applied using a spherical model. Candidate objects

¹Throughout this paper we assume that the surface normal is pointing into the colon.

are generated by finding local maxima in the parameter space created by the Hough transformation.

Simply selecting regions on the colon that protrude inwards yield too many candidates. Therefore, thresholds on mean curvature, principal curvatures, sphericity ratio and/or shape index are used as restrictive criteria. Unfortunately, their values are sensitive to e.g. the CT image noise level and the size of the local neighborhood used to compute them. Generally, the thresholds are set conservative in order not to harm the sensitivity.

All of the above CAD schemes are based on the modelling of an approximately spherical polypoid shape, although, many polyps are often far from symmetric, let alone spherical. Therefore, the candidate generation step of these schemes is characterized by low specificity and much effort is needed to improve specificity while preserving a high sensitivity.

The problems associated with modelling polyps as spherical protrusions are presented in figure 4.1. It shows the (κ_1, κ_2) -space. The horizontal axis shows the first principal curvature, κ_1 . The vertical axis shows the second principal curvature, κ_2 . Since the first principal curvature is by definition larger than or equal to the second ($\kappa_1 \geq \kappa_2$) all convex points on the colon lie inside the region given by $\kappa_2 > 0$ and $\kappa_1 \geq \kappa_2$. Points on perfectly spherical protrusions lie on the line $\kappa_1 = \kappa_2$ ($SI = 1$). On the other hand, perfect cylindrical folds lie on the line $\kappa_2 = 0$ ($SI = 0.75$).

Objects with a larger radius yield smaller principal curvature values and therefore show up closer to the origin. Additionally larger polyps tend to be more asymmetric and therefore more fold-like when described by the shape index. Both large and small polyps are found close to the borders defined by the thresholds on SI and CV as indicated in figure 4.1. As a consequence the criteria used to limit the number of candidate detections are at least as stringent for the larger polyps as for the smaller ones. Notice that this behavior is in conflict with clinical decision making, which dictates that large polyps are more important than smaller ones, since the former have a larger probability to develop into colon cancer.

To comply with clinical practice, one needs a candidate generation step steered by parameters that directly reflect polyp size, such that variations in thresholds only affect the CAD system's sensitivity for small polyps.

In this paper, we introduce a method to estimate the physiological, background shape of the colon wall. Polyp candidates are detected as a local deviation from the background. Our method locally flattens/deforms the colon wall in order to 'remove' the protrusions. The amount of displacement needed for this deformation is used as a measure of 'protrudness' of the underlying lesion. Regions where this measure has a high value are considered as candidate polyps, after which the protrusion measure together with a few additional features is used in supervised pattern recognition.

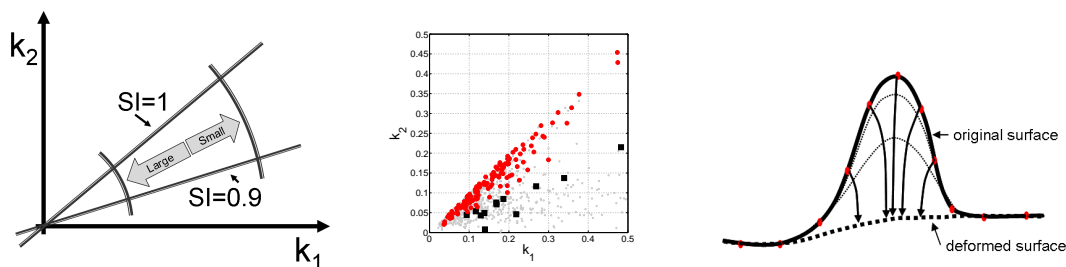


Figure 4.1: Left: the four thick lines enclose a region in the (κ_1, κ_2) -space. The curved lines represent thresholds on the curvedness, while the straight lines enclose the region given by $0.9 < SI \leq 1$. The center image shows the κ_1 and κ_2 values as measured on candidate objects. (See results section) The light grey objects are false positives. The circles are TP that have a largest shape index of at least 0.9. The black squares represent TP as well but have a largest shape index smaller than 0.9. The right image shows a typical polyp shape.

4.2 Methods

A typical polypoid shape is shown in Figure 4.1(right). Suppose that the points on convex parts of the polyp (the polyp head) are iteratively moved inwards. In effect this will 'flatten' the object. At a certain amount of deformation the surface flattening is such that the complete protrusion is removed. That is, the surface looks like as if the object was never there. This is the key concept on which the method is based.

A more formal presentation follows from the description of the surface shape using the principal curvatures. Protrusions are defined as those regions on the surface where the second principal curvature is larger than zero (This implies of course that the first and largest principal curvature is larger than zero as well).

The method then deforms the surface until the second principal curvature is smaller or equal to zero. Clearly, this will only affect structures that are curved in two directions like polyps and will not deform curved structures like folds. Folds typically bend in one direction only and have a first principal curvature larger than zero and a second principal curvature around zero.

Figure 4.1 illustrates how polyps modelled as spherical mounds are found near the line $\kappa_1 = \kappa_2$ with $\kappa_2 > 0$. From here on we drop the spherical model and note that polyp 'heads' are characterized by a $\kappa_2 > 0$. Consequently all regions on the colon wall where $\kappa_2 > 0$ are considered as candidate objects.

The requirement $\kappa_2 > 0$ is less strict than used in most other state of the art systems, which put restrictions on the curvature values. For example by limiting the allowed shape index values to: $SI > 0.8$. One might argue that this will lead to many more candidates. This is unarguably true. We will show, however, that the proposed method

does not require any thresholds other than $\kappa_2 > 0$. Moreover, the deformation method described below leads to a quantitative measure of the polyp protrudness and therefore permits ordering of the generated candidate objects in a way intuitive to the radiologist. In effect this will reduce the number of candidates that are passed to the classification.

4.2.1 Surface evolution

The method employs surface evolution on triangle meshes [24]. The triangle mesh is generated by the marching cubes algorithm applied to the 3D CT data using a threshold of -750 Hu. A typical mesh size consist of around 10^6 vertices.

In [24] a method was presented to rapidly remove rough features (noise) from irregularly triangulated data. It was based on the diffusion equation:

$$\frac{\partial X_i}{\partial t} = \lambda L(X_i), \text{ with } L(X_i) = \left(\frac{1}{N_1} \sum_{j \in 1ring} X_j \right) - X_i \quad (4.2)$$

where $L(X_i)$ is a discrete (1-ring) estimate of the Laplacian at vertex i . X are the positions of the mesh points, N_1 is the number of vertices in the 1-ring neighborhood of vertex X_i and λ is the diffusion coefficient. The solution at time t was found using a backward Euler method which translated the problem into a matrix-vector equation

$$(\mathbf{I} - \lambda dt \mathbf{L}) \bar{X}^{t+1} = \bar{X}^t \quad (4.3)$$

The matrix $\mathbf{M} = \mathbf{I} - \lambda dt \mathbf{L}$ is sparse and its structure is given by the mesh one-ring relations, \bar{X} is a vector containing all mesh points and \mathbf{I} is the identity matrix. This system can be solved efficiently using the bi-conjugate gradient method [91].

In [24] the diffusion was applied to all mesh points. A well known effect of prolonged diffusion on the complete mesh is global mesh shrinking and in [24] a solution was proposed by compensating for the reduction of the mesh volume.

We, however, apply the diffusion only to a limited number of mesh points, namely the points where $\kappa_2 > 0$. The majority of points have negative or zero second principal curvature and remain at their original position. They provide the boundary conditions for the other points. Therefore, in contrast to the method suggested in [24] global shrinking is not an issue and we can search for the steady state solution of the diffusion equation:

$$\frac{\partial X_i}{\partial t} = L(X_i) = 0 \quad (4.4)$$

The discrete Laplacian estimates the new position of vertex X_i by a linear combination of its 1-ring neighbors, X_j . Rewriting equation 4.4 then yields a matrix-vector equation:

$$\left(\frac{1}{N_1} \sum_{j \in 1ring_i} X_j \right) - X_i = \mathbf{M} \bar{X} = 0 \quad (4.5)$$

Fortunately, \mathbf{M} is sparse and its structure is given by the 1-ring mesh relations. The number of non-zero elements on each row equals the number of 1-ring member vertices. Like the backward Euler formulation this equation can also be solved efficiently using the bi-conjugate gradient method [91].

It is well known that the solution to the Laplace equation minimizes the membrane energy subject to the imposed boundary conditions. However, our objective is not to minimize the mean curvature, but to minimize the second principal curvature. Therefore, we extend the above equation by introducing a 'force' term. The resulting equation is a Poisson equation:

$$L(\bar{X}) = \bar{F}(\kappa_2) \quad (4.6)$$

This equation reads as follows: the new positions of the mesh points are found by initially moving each mesh vertex to a position as prescribed by the Laplacian operator. Subsequently, the term on the right hand side 'pushes back' the point such that the resulting second principal curvature is zero.

The force term \bar{F} is designed to depend on κ_2 and is updated after solving equation 4.6. In other words we solve (4.6) iteratively. The force term is initialized with $\bar{L}(X)$ such that we start with:

$$\bar{F}^{t=0} = L(\bar{X}) \quad (4.7)$$

Thus, the 'force field' \bar{F} initially balances the displacement prescribed by the Laplacian and leaves the mesh unaltered. After each iteration \bar{F} is updated with:

$$\bar{F}^{t+1} = \bar{F}^t - \kappa_2^t \frac{A_{1ring}}{2\pi} \bar{n} \quad (4.8)$$

where A_{1ring} is the surface area of the 1-ring neighborhood and \bar{n} is the vertex normal. The last term can be interpreted as a correction term. Note that if κ_2 is positive $\|\bar{F}\|$ should be relaxed. On the other hand, the magnitude of the reduction term additionally depends on the sampling density of the mesh. If the sampling is dense and A_{1-ring} small the magnitude of the correction term should be small. Since κ_2 equals the reciprocal of the radius of the surface tangent circle ($R = \frac{1}{\kappa_2}$) in κ_2 -direction, the term $\frac{2\pi}{\kappa_2^2}$ is half of the area of the fitting sphere. Therefore, the displacement R needed to remove the curvature in second principal direction is normalized by the ratio of these two areas. The estimated displacement is given by:

$$d_{est} = R \frac{A_{1ring}}{2\pi/\kappa_2^2} = \kappa_2 \frac{A_{1ring}}{2\pi} \quad (4.9)$$

The resulting displacement of the mesh points yields a deformed mesh which is an estimate of how the colon wall looks like in the absence of protrusions. The amount of displacement of each mesh point (e.g. in millimeters) is a quantitative measure of the 'protrudness'. Candidate objects are generated by applying a threshold on the displacement field.

4.3 Results

The performance of the method was tested using clinical data from a large, previous study [123]. Automatic polyp detection was executed in three steps: (1) segmentation of the colon wall via the marching cubes algorithm; (2) candidate generation using protrudness; (3) supervised pattern recognition involving a linear classifier and only a few features.

4.3.1 Experimental data

A total of 249 consecutive patients at increased risk for colorectal cancer were included in a previous study [123]. These patients underwent CT colonography before colonoscopy, which served as the gold standard. All patients were scanned in both prone and supine position. The size of a polyp identified during CT colonography was measured in reformatted images, in which the largest polyp diameter was selected for size measurement. Polyp size was also measured during colonoscopy by comparison with an open biopsy forceps of known size. The colonoscopy findings were matched with the CT data. 13 patients were selected that contain 1/3 of the polyps larger than 5 mm from the complete study. This yielded 64 polyps larger than 5 mm. 34 of 64 polyps could be identified in both the prone and the supine CT scan and 30 were identified on either scan but not on the other. Consequently, there were 98 example objects in total.

Figure 4.4 (left) shows a histogram of the CT size-measurements. The majority of objects have a size between 5 and 7 mm. 42 are smaller than 5 mm and are considered as clinically unimportant. 28 are in the range [5,6]mm, 63 are in the range (6,10]mm and 32 are larger than 10mm.

4.3.2 Candidate generation

A typical result is given in figure 4.2. It shows three renderings of the colon wall surface. In the left picture an isosurface volume rendering of a 7 mm large polyp is shown. The polyp is situated on a folded colon structure. The middle picture shows the deformed mesh (visualized by a mesh rendering). The protrusion is 'removed', demonstrating how the colon may have looked like in the absence of the polyp. The right image shows the original mesh with the segmentation obtained by thresholding the protrusion measure at a value of 0.1mm

An other example is presented in figure 4.3. The left picture shows an isosurface volume rendering. In the center of the picture a large (14mm) non-spherical polyp is situated between two folds. The middle picture shows the segmentation as obtained by hysteresis thresholding the shape index using values 0.8 and 0.9 and curvedness values in the range: $0.05 < CV < 0.25$. The same thresholds were used in [141] to generate

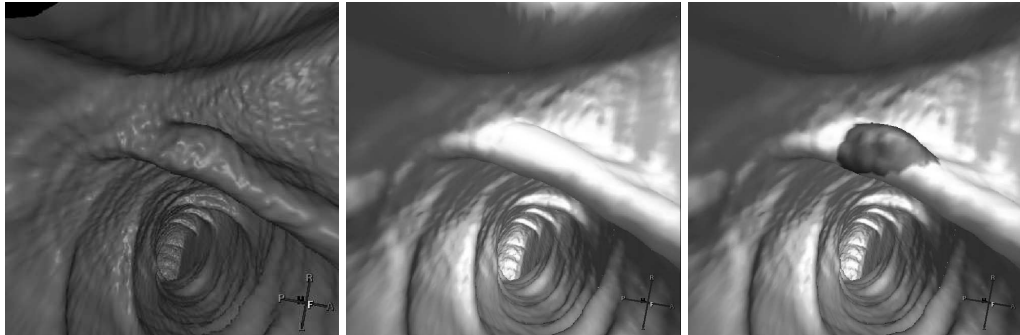


Figure 4.2: Example of the deformation applied to a polyp on a fold. Left: the original colon surface with a polyp on a fold (isosurface volume rendering). Middle: the deformed surface mesh with the polyp 'removed' (mesh rendering). Right: the obtained segmentation by thresholding the displacement field (mesh rendering).

candidate objects. The shape index and curvedness were computed by fitting a 2nd order polynomial to the mesh using a local neighborhood with 5mm radius.

Due to the irregularity of the protrusion several segments of the polyp have been found. The right picture shows the segmentation as a result of thresholding the protrusion measure (value 0.1mm). A more coherent segmentation has been obtained.

Our method applied to the 13 patients yielded 1578 candidate objects (including the true positives), which is ± 60 per dataset. A total of 3 polyps (between 5mm and 6mm) were missed in the candidate generation step. In contrast, if a threshold of 0.9 on the shape index was used to segmented candidates a total of 16 polyps would have been missed. (See figure 4.1 (middle)). From the segmented candidates the vertex with largest shape index value was found. For this vertex the κ_1 and κ_2 values are plotted in figure 4.1 (middle). This plot does not include the 3 missed polyps.

4.3.3 CAD performance

The candidate generation step was used as an input to supervised pattern recognition. In figure 4.4 we show ROC curves based on a linear classifier applied to four features: the maximum protrusion found on a candidate object; the size of the object measured as proposed in [26]; volume, obtained from the enclosed volume between the original mesh and the deformed mesh and the percentage of SI values for the vertices on the segmented surface patch, that is within the range of $0.65 < SI < 0.85$. The latter value is expected to attain high values on folded structures.

Three lines in figure 4.4 (right) show the performance of the system for different size classes. The data was generated in a leave-one-patient out manner.

The large polyps (>10mm) are found with 100% sensitivity at a FP rate of 13 per

4.4. CONCLUSIONS

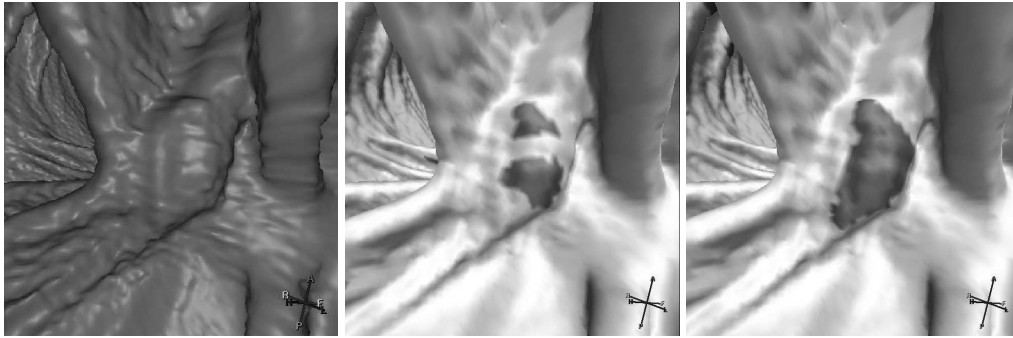


Figure 4.3: Segmentation of irregular objects. Left: isosurface volume rendering; Middle: segmentation obtained by hysteresis thresholding on the shape index; Right: segmentation obtained by thresholding the protrusion measure.

dataset and 90% sensitivity at 2 FP per dataset. The ROC for polyps larger than 6 mm (including those larger than 10mm) show that 80% sensitivity is obtained at the cost of 4 FP per dataset. A 95% sensitivity is obtained at the cost of 10 FP per dataset. The results for polyps larger than 5mm are similar to the results for polyps larger than 6mm. However, 100% sensitivity is not reached. This is due to the fact that in the candidate generation step 3 polyps between 5mm and 6mm have been missed.

The plots in figure 4.4 demonstrate that the specificity of the CAD system is even higher if the sole purpose of the system is for diagnostic purposes only, namely to decide whether a patient has or has not any polyps. The left plot shows how many false positives have been assigned a higher posterior probability than the first large polyp ($\geq 10\text{mm}$). It follows that for 7 out of 13 patients the first object is a polyp. For one patient three false positives have a higher posterior probability than the first polyp. Note that the total number of patients is 11 (and not 13). This is due to the fact that 2 patients did not have polyps larger than 10mm. The right plot shows the number of false positives with a higher posterior probability than any of the polyps larger than 6mm (including those larger than 10mm). Note that including the smaller polyps improves the results. Apparently some of the small polyps are assigned a higher posterior probability than the large ones.

4.4 Conclusions

We have presented a method to detect protruding objects on curved surfaces. It was used to generate candidate objects for automated polyp detection. The method works by locally flattening the colon wall in order to 'remove' protrusions. Actually, the colon surface is deformed until the second principal curvature is smaller than or equal to zero. Therefore, only those structures are affected that are curved in two directions,

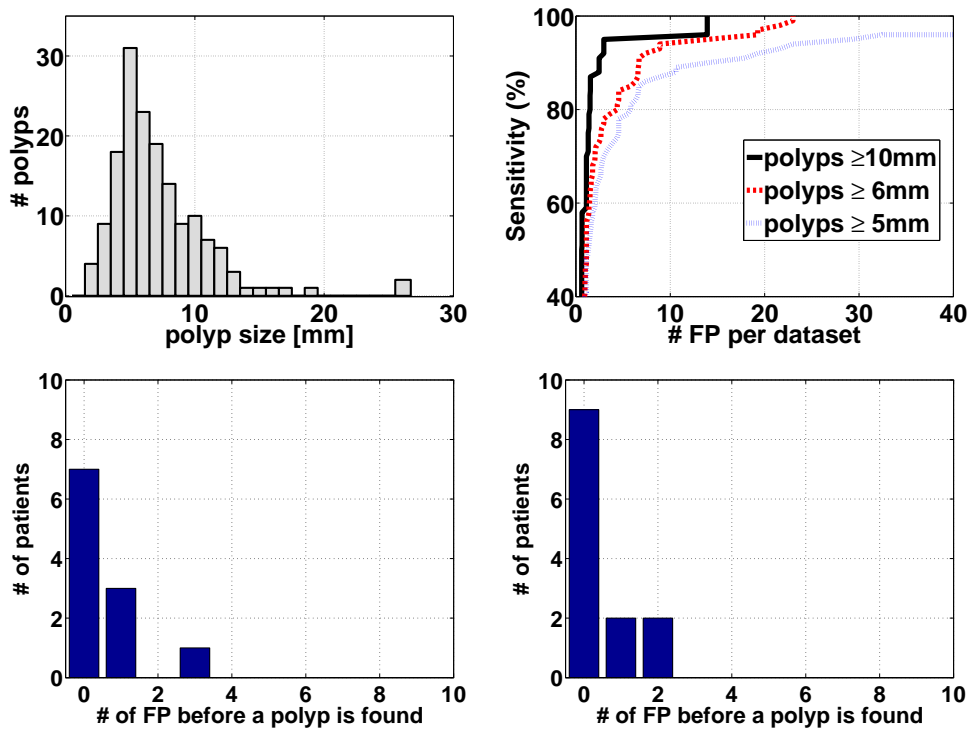


Figure 4.4: Left two: Polyp size distribution (left) and ROC curve for the supervised pattern recognition based on protrusion plus three additional features (right). Right two: Histogram of the number of false positives per patient presented to an expert before a true positive polyp is found (ordering by the posterior probabilities). The left picture shows the number of FP before a polyp larger than 10mm is considered (if there is one in the dataset). For 7 patients the first object presented is a large polyp. For one patient 3 FP are presented before a large TP is presented. The right plot shows the the number of FP before a polyp larger than 5 mm is encountered.

4.4. CONCLUSIONS

like polyps. Folds remain unaltered. The amount of displacement needed for flattening/deformation is used as a measure of 'protrudness' of the object.

A threshold on the deformation field is the only parameter needed for candidate generation. This is a clear advantage over methods that involve many restriction criteria. Another advantage is that the deformation field immediately allows for the computation of additional features such as the object's volume.

We have shown that a simple linear classifier involving only four features already yields 95% sensitivity at the cost of about 10 FP per dataset.

Clearly, the algorithm must be extensively tested. We do realize an investigation on more data involving more complex classifiers is needed to be conclusive on the overall improvement of polyp detection. However, the current results give an indication that the protrusion measure may enhance polyp detection schemes.

5 Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow

Today's computer aided detection (CAD) systems for CT colonography (CTC) enable automated detection and segmentation of colorectal polyps. We present a paradigm shift by proposing a method that measures the amount of protrudedness of a candidate object in a scale adaptive fashion. One of the main results is that the performance of the candidate detection depends only on one parameter, the amount of protrusion. Additionally the method yields correct polyp segmentation without the need of an additional segmentation step. The supervised pattern recognition involves a clear distinction between size related features and features related to shape or intensity. A Mahalanobis transformation of the latter facilitates ranking of the objects using a logistic classifier. We evaluate two implementations of the method on 84 patients with a total of 57 polyps larger than or equal to 6 mm. We obtained a performance of 95% sensitivity at 4 false positives per scan for polyps larger than or equal to 6 mm.

Accepted for publication:

C. van Wijk, V. F. van Ravesteijn, F. M. Vos, L. J. van Vliet, *Detection and Segmentation of Colonic Polyps on Implicit Isosurfaces by Second Principal Curvature Flow*, IEEE-TMI

5.1 Introduction

Protrusions of a surface embedded in a 3D image are difficult to detect. The challenge increases even further if the surface itself is highly structured and interacts with the protruding elements. Such a problem is the detection of polyps in CT colonography (CTC), a minimal invasive technique for examining the colon surface (cf. Fig. 5.1). There is an increasing interest in computer aided detection (CAD) systems for polyp detection in CTC data to assist the radiologist [21,58,62,78,105,112,133]. Such a CAD system typically consists of three consecutive steps: colon segmentation; detection of polyp candidates; and supervised classification of candidates as polyps or non-polyps [51,138].

Adenomatous polyps are important precursors to colon cancer and develop due to genetic mutations in the mucosa cells [17]. This process of oncogenesis leads to enhanced cell proliferation causing the polyp to grow and to evolve from a small adenoma into a large adenoma into a carcinoma. This induces a morphological change to the colon surface¹, that manifests itself as tissue protruding into the lumen. The deformation is an important property which is used in the detection by radiologists as well as gastroenterologists.

Practically all CAD systems for polyp detection analyze the local curvature of the colon surface, which is subsequently used to compute shape descriptors such as shape index or curvedness [103,141]. Computation of the curvature values is typically done in 'one shot' on a single predetermined scale, which is defined as the effective size of the area over which the image features are calculated. We will maintain this definition throughout the paper.

We propose a new paradigm for the detection and segmentation of polyps that effectively copes with the highly structured environment. The novelty of the approach is in computing an intensity change field, which removes protruding elements from the underlying data, while leaving the highly structured folds intact. The deformation algorithm is described by a partial differential equation (PDE) that is steered by the second principal curvature.

In order to demonstrate the efficiency of the method, we make use of a pattern recognition system introduced by us in [125]. The paper involved polyp detection based on the explicit representation of the colon surface. The method proved to generalize well and lead to satisfying results. It encouraged us to further investigate the candidate detection system. Presently, we propose a technique based on an implicit representation of the colon surface, which enables a number of improvements over the explicit model. A concise description of the classifier is contained, since it is only indirectly related to the paper's main objective. This allows us to fully go into all facets associated with second

¹Not all colonic lesions grow into protruding polyps. It is estimated that approximately 10% of the lesions are so-called flat adenomas [?,38].

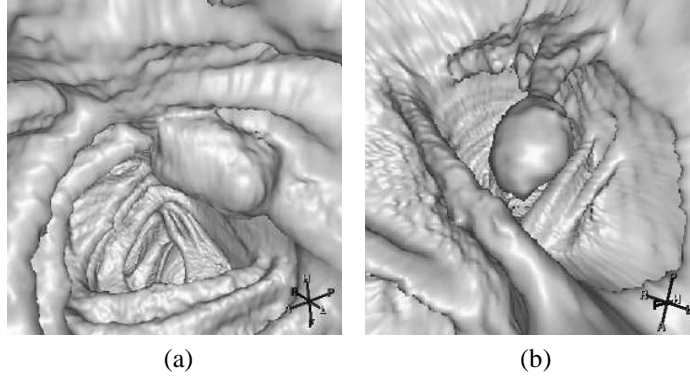


Figure 5.1: Isosurface renderings (at -650 HU) of the colon surface showing typical polyps in their structured surroundings.

principal curvature flow.

5.1.1 Previous work

For the detection of candidate regions, Summers et al. [103] proposed to use the mean and Gaussian curvature. They were computed by methods from differential geometry, by fitting a 4th order b-spline to local 5 mm radius neighborhoods of a triangulated isosurface [104]. Candidates were generated by selecting a range of mean and Gaussian curvature values. Additionally, a large number of other shape criteria were used ([137]: Table 2), to limit the number of false positive detections. Similarly, Yoshida et al. [141] used the shape index and curvedness to find candidate objects on the colon surface. The shape index SI and curvedness CV are functions of the principal curvatures of the surface:

$$\begin{aligned} SI &= \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right), \\ CV &= \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}}, \end{aligned} \quad (5.1)$$

with κ_1 and κ_2 the maximum and minimum principal curvature respectively. A Gaussian-shaped window (aperture) of fixed size was used to compute the curvatures from the 3D CT data.

Alternatively, Kiss et al. [58] proposed to use a sphere fitting method to generate candidates. The colon surface was classified as convex depending on the side on which the center of the fitted sphere was found, either in tissue or in air. This method classifies roughly 90% of the colon surface as concave, that is as 'normal'. To the remaining part

of the colon surface a generalized Hough transformation was applied using a spherical model. Candidate objects were generated by finding local maxima in the parameter space created by the Hough transformation.

Konukoglu et al. [62, 63] proposed a method that is in some sense the inverse of the approach that is proposed in the current paper. Effectively, a wall evolution algorithm is described based on a level-set formulation that regularizes and enhances polyps as a preprocessing step to CTC CAD algorithms. The underlying idea is to evolve the polyps towards spherical protrusions on the colon wall while keeping other structures, such as haustral folds, relatively unchanged. Thereby, the performance of CTC CAD algorithms is potentially improved, especially for smaller polyps.

Conventionally, the shape-based candidate detection methods [37, 62, 78, 82, 103, 141] apply several conservative thresholds to the mean curvature, principal curvatures, sphericity ratio and/or shape index to generate candidate regions.

5.1.2 Problem definition

We identify a number of challenges that are associated with the detection of polyp candidates. First, optimization of the parameters is always complicated by the limited availability of training examples. This may lead to overtraining for a specific patient population, patient preparation, scanning hardware or scanning protocol. Thus, it is preferred to keep the number of restrictive criteria to a minimum.

Second, to achieve good discrimination power and accurate measurement [127] of lesion size, precise 'delineation' (or segmentation) of the candidate is needed. Although a number of methods are available for segmentation purposes [1, 26, 137], adding such a separate step would introduce more parameters to the CAD pipeline and should be avoided. Fuzzy segmentation methods using sophisticated pattern recognition techniques might preclude this problem.

A third challenge is associated with the computation of the first and second order derivatives, which are needed to compute the principal curvatures and to characterize local shape. The derivative operators must act on a range of sizes and should not have optimal performance for a specific size only. Ideally, the scale should adapt to the underlying image structure. To our knowledge no research has been performed to either analyze the effect of scale or to determine the optimal scale for polyp detection. It is partly addressed in [111] by performing the curvature computation on a high resolution triangulated isosurface mesh thereby limiting the low pass filtering across the isosurface. Furthermore, some research on scale selection for CTC in general has been performed in [27, 66].

Last, detecting large polyps is (clinically) more important than detecting smaller ones. One would like to have this built into the CAD system. In other words, the detection method must perform optimal for large polyps.

5.1. INTRODUCTION

A steadily growing number of papers ([18, 40, 84, 105, 108, 115, 141]) report on the performance of specific polyp detection algorithms. Unfortunately, a proper comparison of algorithms is complex due to differences in prevalence, patient preparation, scanning protocol, and determination of the ground truth.

We aim to convey some general requirements for polyp detection systems:

1. It should not involve many parameters which need to be tuned in the presence of a limited number of polyps;
2. A separate segmentation step should be avoided when it adds more parameters;
3. It must be able to cope with the whole range of polyp sizes encountered in practice,;
4. It should take into account the increased clinical relevance of larger polyps;
5. It should be robust to variations in the imaging process (e.g. radiation dose, but also scan resolution, orientation and patient preparation)

5.1.3 Objective

We aim to introduce a new paradigm for the detection of protruding regions on highly structured surfaces embedded in a 3D image. Polyps are assumed to have introduced a deformation to the originally healthy colon surface. We will describe a novel method to reconstruct the data without these protrusions. Effectively, the technique aims to ‘undo’ the deformation by modifying the underlying intensities so that the protruding shape is no longer there.

The proposed method does not require any assumptions on the lesion shape such as axial-symmetry, sphericity or lesion size, other than that it protrudes. It works well for highly irregular protruding objects. Lesion candidates are generated using only a single threshold. Small variations of the threshold affect the detection sensitivity of the smaller polyps first. Additionally, the resulting segmentations include the complete object (both head and neck).

In earlier work [130] we proposed a scheme that operated on an explicit representation of the colon surface, which was obtained by a triangulation of the isosurface at -650 HU. Only information of this particular isophote was used to estimate the structured surface without the protrusions. Possible beneficial information from other isophotes, with higher or lower intensities, was ignored. The scheme proposed in this paper differs fundamentally by acting on an implicit representation of the colon surface. That is, it uses information from other isophotes as well. Consequently, there is no need for optimizing the intensity level of the isosurface. Another advantage of this method is that topological complexities and complex mesh processing tasks, such as mesh generation

and mesh smoothing, are avoided. We will compare both methods and demonstrate that the two techniques are to some extent complementary. Moreover, exploiting the complementary aspects will be shown to lead to improved sensitivity.

5.2 Methodology

5.2.1 Materials

A total of 84 patients with an increased risk for colorectal cancer were consecutively included in a previous study [122]. All data were acquired using a Mx8000 multislice CT scanner (Philips Healthcare, Best, The Netherlands) using the same scanning protocol for all scans (120 kV, 100 mAs, 4x2.5 mm collimation, pitch 1.25, standard reconstruction filter). Slice thickness was 3.2 mm. All patients adhered to an extensive laxative regime without taking a tagging agent with their diet. All patients underwent CT colonography before colonoscopy. The patients were scanned in both prone and supine position; thus, a total of 168 scans were used in our study. The findings of colonoscopy served as the golden standard. Polyp size was also measured during colonoscopy by comparison with an open biopsy forceps of known size. A research fellow annotated the location of polyps in all CT scans. For the 84 patients, 108 polyps were annotated. The number of polyps with a size larger than or equal to 6 mm was 57 and the number of polyps larger than or equal to 10 mm was 32. Fig. 5.2 shows a histogram of the optical colonoscopy size-measurements. Conventionally, polyps which are smaller than 6 mm are considered clinically unimportant. Therefore, they were not used in the performance analysis. The peak at 10 mm polyp size is caused by the fact that in clinical practice only a few bins are used: smaller than 6 mm, between 6 and 10 mm and larger than or equal to 10 mm.

Experts labelled the polyps in CT data based on the optical colonoscopy findings without using CAD. A candidate generated by the CAD system was labelled as a true positive if an annotation was within 5 mm from any of the voxels in the candidate and was not closer to any other candidate. A margin of 5 mm was used to accommodate inaccurate localization by the expert. Especially for the explicit method, such a margin is needed to accommodate annotation inside the polyp. To be able to make a proper comparison between the two methods, the same margin is used for both techniques.

5.2.2 Method

A typical polypoid shape is shown in Fig. 5.3(a). Suppose that the points on the convex region of the polyp (the polyp head) are iteratively moved inwards. In effect this process will ‘flatten’ the object (Fig. 5.3(c)). Note that the convex region expands during the process and will ultimately include the polyp neck as well. After a certain amount of

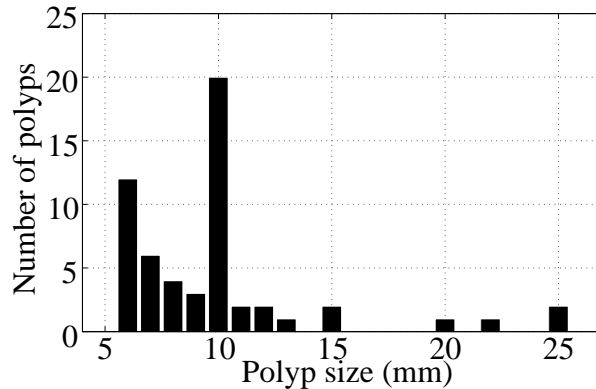


Figure 5.2: Distribution of sizes obtained during colonoscopy of 57 polyps larger than or equal to 6 mm in 84 patients from a previous study [122]. One polyp of 45 mm is not visible in the histogram.

deformation, the surface flattening is such that the protrusion is completely removed. That is, the surface looks like as if the object was never there. This is the key concept on which the method is based.

Before formalizing on the operator we first have a closer look at the second order differential properties of the implicit surface embedded in a three-dimensional voxel space. The colon can be considered as a long elongated structured tube. For a perfect cylinder shape the principal curvatures are smaller than or equal to zero everywhere. However, the colon contains many folds, i.e. structures which are bended only in one direction: the first principal curvature is larger than zero, whereas, the second principal curvature is close to zero. Protruding objects, such as polyps, have positive values for the first and second principal curvature. Therefore, an operator is designed to affect only on points with a positive second principal curvature and in such a way that the second principal curvature decreases. Repeated application of the operator will eventually yield an image where the second principal curvature is smaller than or equal to zero everywhere.

Consider once more the schematic representation of a polyp in Fig. 5.3(a). The distinction between the head ($\kappa_1 > 0, \kappa_2 > 0$) and neck ($\kappa_1 > 0, \kappa_2 \leq 0$) regions of the object is made by the sign of the second principal curvature. On the line connecting the inflection points A and B in the figure (separating the regions 'head' and 'neck') the Gaussian curvature is zero. The proposed method initially adapts the head region only. It will now be demonstrated that such adaptation leads to an expansion of this region.

To that end, Fig. 5.3(b) shows a planar cross section through A, spanned by the local gradient vector and the direction of the second principal curvature. Let us merely consider the curve emanating from this cross section. The steepness of this curve corresponds to its first derivative; the curvature corresponds to its second derivative and is

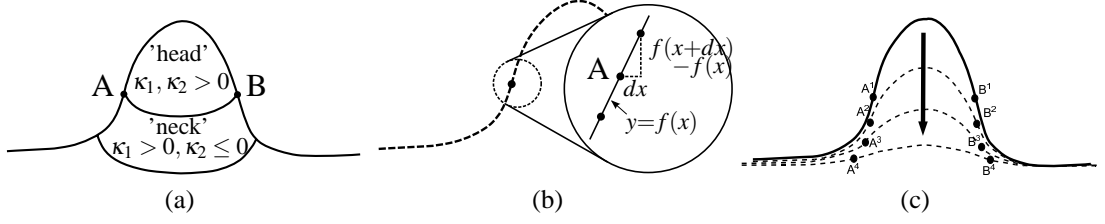


Figure 5.3: Schematic illustration of the deformation process. (a) Three regions (head, neck and periphery) are distinguished. (b) The second principal curvature κ_2 is zero at the border between the head and neck region. (c) The head region expands during the deformation process.

given by:

$$\kappa = -\frac{\Delta y}{|\nabla y|}, \quad (5.2)$$

in which Δy represents the second derivative of the curve. By convention κ has a sign opposite to that of the second derivative. Observe that this curvature is positive on the 'head' side from A and negative on the 'neck' side from A; the curvature equals zero in A. At the position of A the second derivative is:

$$\Delta y = \frac{d^2 f}{dx^2} = \lim_{dx \rightarrow 0} \frac{\frac{f(x+dx) - f(x)}{dx} - \frac{f(x) - f(x-dx)}{dx}}{dx} = 0. \quad (5.3)$$

A reduction of the protrusion in the head region implies that the value of $f(x+dx)$ in (5.3) is lowered. Consequently, the second derivative in A (Δy) becomes negative, and the curvature (κ) positive. Thus, the zero crossing of the second derivative will shift outwards in Fig. 5.3(b) and the head region will expand into the neck region.

The effect of repeatedly reducing the protrusion is illustrated in Fig. 5.3(c). The points with zero second principal curvature shift from A¹ to A⁴ and B¹ to B⁴. Eventually, the protrusion is flattened over the complete shape, i.e. both the head and neck regions. Although the initial delineation of the head region of the structure (in which the deformation is started) may be affected by noise, the area of operation eventually spreads to the entire polyp area. It is this property that makes the procedure robust. The results section contains some examples to illustrate the method's efficacy.

5.2.3 Second principal curvature flow

A scheme to remove protruding elements from a curve in 2D is the Euclidean shortening flow [80]. A similar approach can be taken in 3D, for which the flow is governed by:

$$\frac{\partial I}{\partial t} = -g(\kappa_1, \kappa_2) |\nabla I|, \quad (5.4)$$

with κ_1 and κ_2 the first and second principal curvatures, $|\nabla I|$ the gradient magnitude of the input image I , and $g(\cdot)$ a curvature dependent function characterizing the flow. The principal curvatures can be derived from the trace of the Hessian matrix H :

$$H = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix}, \quad (5.5)$$

with x , y and z the image coordinates and I_{ij} the second derivative $I_{ij} = \partial^2 I / \partial i \partial j$. In gauge coordinates the Hessian is a diagonal matrix with terms [126]: I_{gg} , I_{uu} and I_{vv} . The first term is the second derivative in gradient direction; the second and third terms are the second derivatives in the directions of the principal curvatures of the isosurface perpendicular to the gradient vector. The latter two relate to the principal curvatures of the isosurface:

$$\begin{aligned} I_{uu} &= -\kappa_1 |\nabla I|, \\ I_{vv} &= -\kappa_2 |\nabla I|. \end{aligned} \quad (5.6)$$

With the definition of inward normals, the second principal curvature in the colon is everywhere smaller than or equal to zero, except on protruding regions. Here, both the first and second principal curvatures are positive and the corresponding second derivatives are negative.

$g(\kappa_1, \kappa_2)$ may be defined in various ways [81], e.g. by the mean curvature [13, 44] or the Gaussian curvature. We require that $g(\kappa_1, \kappa_2)$ is continuous, especially at locations where the sign of κ_2 changes, to avoid a discontinuous deformation. Moreover, it must be small on folds with a small positive value of κ_2 so that the deformation on such locations is small. Reversely, the response to polyps with two large principal curvatures should be large. Accordingly, we solve the following nonlinear PDE:

$$\frac{\partial I}{\partial t} = \begin{cases} I_{vv} & (\kappa_2 > 0) \\ 0 & (\kappa_2 \leq 0) \end{cases}. \quad (5.7)$$

Thus, only at protruding regions the image intensity is reduced by an amount proportional to the local second derivative in the direction of κ_2 .

5.2.4 Implementation

The proposed method is applied to voxels on and around the colon surface. This region of interest (ROI) is defined by a mask. First, a binary image is obtained by thresholding the CT image at -650 HU. Subsequently the mask is generated by applying the exclusive

or (XOR) operation to an eroded and a dilated version of the binary image. The number of iterations for the dilation and erosion should be such that the full air-colon transition is included in the resulting mask image. We used a conservative value of 10 mm for the radius of the erosion and dilation kernels.

The partial differential equation (5.7) is solved for the voxels in the ROI defined previously. The intensities of voxels outside the ROI are not altered and serve as Dirichlet boundary conditions. The left hand side of (5.7) is discretized by a forward difference scheme:

$$\frac{\partial I}{\partial t} = \frac{I^{t+1} - I^t}{dt} + O(dt). \quad (5.8)$$

The right hand side of (5.7) requires computation of first and second order derivatives. The first order derivative is determined by the local orientation of the normal field. An accurate estimate is required to prevent diffusion of information across isophotes, leading to blurry effects. Unfortunately, simple central difference derivative operators are known to have rather poor rotation invariance [94]. Therefore, the first and second order derivatives are computed after a (second order) Taylor expansion in a 3x3x3 neighborhood [120]. They are used to compute I_{vv} .

The image values are modified in a semi-implicit manner comparable to a Gauss-Seidel scheme, meaning that some of the underlying intensity values are at time $t + 1$, while others are at time t :

$$I^{t+1} = \begin{cases} I^t + \frac{\Delta t}{(\Delta x)^2} I_{vv}^{t+1/2} & (\kappa_2 > 0) \\ I^t & (\kappa_2 \leq 0) \end{cases}, \quad (5.9)$$

in which $I_{vv}^{t+1/2}$ indicates that it is computed with information from time steps t and $t + 1$. For Laplace's equation, numerical stability is guaranteed if the term $\Delta t / (\Delta x)^2$ is smaller than $\frac{1}{6}$ [124]. Therefore, the maximum time step for which stability is attained depends on the direction in which the voxel size is smallest (typically in-plane): $(\Delta t)_{\max} = \frac{1}{6} \cdot (\Delta x)^2$. Note that this is a conservative value since we only use the principal second derivative, I_{vv} , instead of the full Laplacian: $I_{gg} + I_{uu} + I_{vv}$. The aspects of stability, convergence and correctness for similar problems have been elaborately discussed in [80]. For a more formal discussion, see [4] and also [124]. In practice, we have never encountered a problem concerning the stability and convergence of the solution.

Summarizing, the algorithm acts only on the head regions in which $\kappa_2 > 0$. A new intensity is assigned by (5.9) to each voxel within such a region. Subsequently, the principal curvatures are recomputed. Some of the voxels which initially had zero or negative second principal curvature will now be in the head region and will be added to the area of operation. In this way, during iteration, the area of operation will expand from the head into the neck region.

An obvious stopping criterion would be to track the amount of intensity change during iterations and stop when the amount of intensity change at a particular iteration is

lower than some predefined value. Unfortunately, this leads to an underestimate of the protrusion of large objects, with a low value for the second derivative even when the protrusion may be quite large. In our implementation, we have taken a heuristic approach. After each iteration, the number of voxels that are added to the convex region is counted. The algorithm stops when this number is zero.

A crucial property of the method is that the effective kernel scale increases with each iteration. Such adaptation occurs since the curvature calculation continuously uses the result from the previous step. In effect, the scale 'adapts' to the underlying image structure, because a small protrusion will require less iterations to be flattened into the background than a large one. In other words, the effective scale varies locally as the number of iterations needed to reach a 'steady state' differs from location to location. Simultaneously, the area of operation, which is delimited by zero second principal curvature, also changes during iterations. By definition, the head region of a structure is adapted first, but subsequently the area of operation extends to the neck region (see Fig. 5.3). Existing methods typically estimate curvature values in 'one shot' by selecting one scale of derivative operators a priori. A limitation of the current method may be associated with protruding objects with small κ_2 . Such structures deform slowly due to small curvature. It will be demonstrated that the detection of large polyps is not hampered by this limitation (see Section 5.3.2).

Fig. 5.4 demonstrates that the method works well also for highly irregular shapes. The first row shows the isosurface (rendered at -650 HU) at different stadia of the deformation process. During the first iterations only the two protruding regions on the left and right side of the polyp are affected. In later stages these two regions merge and also the middle part is deformed. The steady state solution and the resulting segmentation by thresholding is shown in the last two pictures of the first row. The second row shows the shape index (SI) computed from Gaussian derivatives obtained using different scales ($\sigma = 2, 4, 8, 12$ mm), red corresponds to $SI = 1$, magenta to $SI = 0.75$ (e.g. on folds). The third row shows the regions with SI larger than 0.8. The example demonstrates that scale has a profound effect on the resulting SI values. All polyps in our dataset that are larger than 10 mm have multiple separated head regions when 'observed' at a small scale (see Fig. 5.11b for the performance of our algorithm on large objects).

5.2.5 Candidate segmentation

The steady state yields new intensities for voxels, particularly in protruding regions. We will now demonstrate that the intensity change is a measure for the amount of displacement of the isosurface.

Let \vec{x} represent a position in which the intensity $I^{t=t_0}(\vec{x})$ is halfway the intensities of the colon lumen and the tissue. Furthermore, the algorithm is asserted to displace the isosurface through \vec{x} by a small amount δ (smaller than the width of the point spread

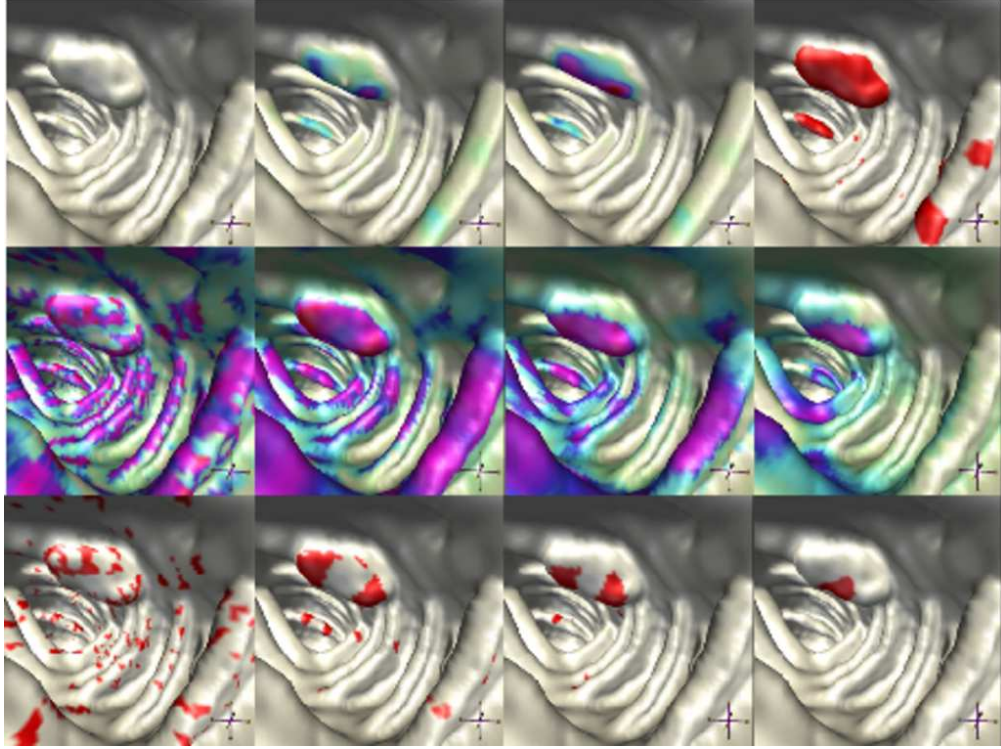


Figure 5.4: Demonstration of polyp detection by the curvature flow (first row). The second and third row show results as obtained by thresholding the Shape Index, computed at different scales. See text for details.

function (PSF)) after some iterations at $t = t_i$. Then, the intensity $I^{t=t_i}(\vec{x})$ can be computed via a first order Taylor series expansion:

$$I^{t=t_i}(\vec{x}) = I^{t=t_0}(\vec{x}) + \delta \cdot \nabla I^{t=t_0}(\vec{x}) + \varepsilon. \quad (5.10)$$

Notice that δ refers to a hypothetical step size corresponding to a small displacement of the isosurface. Reversely, a small change in intensity relates linearly to the amount of displacement. However, large displacements of the isosurface cannot be described as such. The intensity change levels off for displacements larger than the PSF width:

$$I^{t=t_\infty}(\vec{x}) = I^{t=t_0}(\vec{x}) - \frac{C}{2}, \quad (5.11)$$

in which C denotes the total contrast over the transition from lumen to tissue (typically around 1000 HU).

The sketch in Fig. 5.5 illustrates the relation between the intensity change (before and after deformation) and the colon surface displacement, halfway the air-tissue transition. Clearly, the intensity change is monotonically increasing with increasing displacement

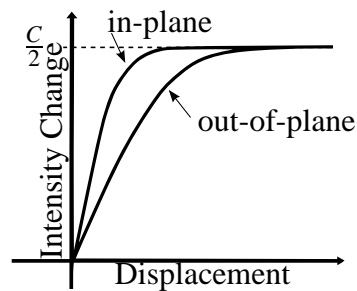


Figure 5.5: Sketch of the relation between colon surface displacement and the observed intensity change for positions halfway the step edge. The relation depends on the apparent local scale of the PSF, i.e. the scale in the direction of the surface normal. Often, the scanner resolution is not isotropic: the in-plane resolution is larger than the out-of-plane resolution. As a consequence, the relation depends also on the direction of surface displacement.

of the isosurface. This would permit a segmentation by a simple threshold on the intensity change if the data were isotropic, but unfortunately CT data often are not. The in-plane resolution is frequently higher than the resolution in scanning direction (z). In other words, the apparent scale of the PSF σ_{apparent} depends on the direction of the colon surface normal. Consequently, the relation between intensity change and colon surface displacement (cf. Fig. 5.5) depends on the orientation of the protruding structure. To solve this problem, the derivative kernels are made anisotropic such that the apparent scale will be isotropic and equal to a certain target scale σ_{target} . The kernel scale σ_i , in the direction $i \in \{x, y, z\}$, is computed by $\sigma_i = \sqrt{\sigma_{\text{target}}^2 - \sigma_{\text{apparent},i}^2}$, in which $\sigma_{\text{apparent},i}$ is the apparent (anisotropic) scale of the PSF. Polyp candidate regions are segmented by thresholding the intensity change field, followed by a labelling operation. The threshold value is 100 HU corresponding to the threshold of 0.4 mm surface displacement as used in [130] for data with an assumed Gaussian PSF [97] with $\sigma = 1.6 \text{ mm}^2$.

5.2.6 Features for classification

For each candidate object, five features are computed. These features relate to the two properties that are primarily used by a radiologist: shape of a candidate and intensity distribution inside a candidate. We explicitly make this distinction since only size descriptors permit a ranking of the candidate objects in a way that relates to clinical relevancy. Accordingly, size related features will be treated differently than the other features in the pattern recognition step. Conventionally, polyp size is defined as the single largest

²Halfway the air-tissue transition: $\nabla I^{t=t_0} = \frac{C}{\sigma\sqrt{2\pi}} = \frac{1000}{1.6\sqrt{2\pi}} \approx 250 \text{ HU/mm}$, thus $100 \text{ HU} \triangleq 0.4 \text{ mm}$, i.e. equal to the threshold used in [130].

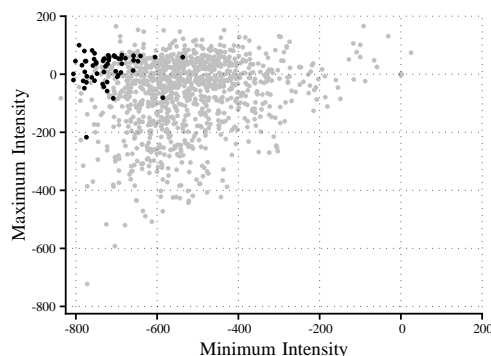


Figure 5.6: Feature space of the maximum and minimum intensities for each candidate region. Annotated polyps are depicted by black dots and have maximum intensities around 0 HU (tissue) and minimum intensities around -650 HU. Only one in every 20 false positives is shown as a grey dot.

diameter, excluding the stalk. We compute it automatically using the method described in [26], which not only returns the largest diameter (LongAxis), but also the shortest diameter (ShortAxis). These are the first two size related features that are used in the classification. Notice that their ratio incorporates shape information. The third feature is the maximum intensity change (MaxIntChange) within each segmented region (candidate). It directly relates to the isosurface displacement (cf. Fig. 5.5). For larger polyps the values of this feature will be large and vice versa. The fourth and fifth features used for classification are the 5 and 95 percentile intensities inside the candidate. We employ these percentile values and not the minimum and maximum intensities to increase the robustness against noise. For simplicity, we will refer to these two features as the minimum (MinHU) and maximum (MaxHU) intensity values inside the objects. Notice that all features depend on the intensity change field since all are computed over the segmented volume of a candidate. Only the MaxIntChange feature is directly derived from the intensity change field in the segmented volume, the others are computed from the original CT data.

5.2.7 Classifier training

It was mentioned previously that the intensity features do not directly allow for an ordering of the candidates. As an example, consider the feature space of MinHU and MaxHU shown in Fig. 5.6. The black dots denote true positive candidates and the grey dots denote false positive candidates.

The distribution of polyps is somewhat Gaussian, and there is a large overlap with the non-polyps. The latter do not show a simple distribution in this space. For these reasons, these two features are not used *directly* for classifier training. Instead, we compute

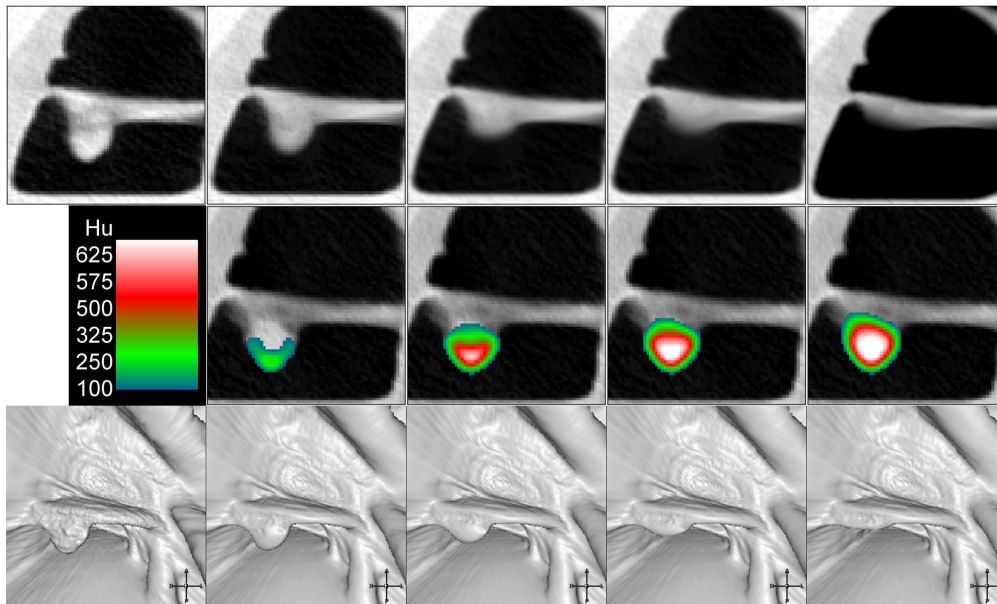


Figure 5.7: Polyp (10 mm) at different stages of the intensity deformation (after 0, 10, 40, 80 and 160 iterations of (5.9)) . First row: original data; second row: overlay showing the intensity changes larger than 100 HU (the color scale was truncated at 650 HU); third row: isosurface renderings (at -650 HU).

the Mahalanobis distance to the polyp class center. Such a mapping orders the candidates by the distance to the center of the Gaussian, i.e. the center of the polyp class yield zero Mahalanobis distance. Notice that the center and width of the Gaussian are to be determined on independent training data. This strategy mimics the use of a Gaussian one-class classifier [114]. Complementary, the remaining features (MaxIntChange, LongAxis, ShortAxis) relate to size and are directly used to order the candidates. The ranking of the candidates imposes that changes in the decision boundary affects the classification in an ordered fashion.

It may be expected that far more small candidates are detected than large ones due to noise and the small ‘effective’ scale on small objects. Consider a connected number of pixels affected by positively signed noise. Such coherent regions may mimic small objects with positive principal curvature. The derivatives computed from the $3 \times 3 \times 3$ Taylor expansion experience a small amount of regularization. Consequently, the little blurring may leave small noise protrusions on an otherwise smooth surface. This is confirmed by the distribution of the false positive candidates with respect to the MaxIntChange feature, which resembles an exponential distribution. Concurrently, we have observed that the polyps denoted by black dots in 5.6 are approximately uniformly distributed. Therefore, the ratio of the posterior probabilities must follow an exponential decay as a function of MaxIntChange. This is a situation in which a logistic classifier [135] is

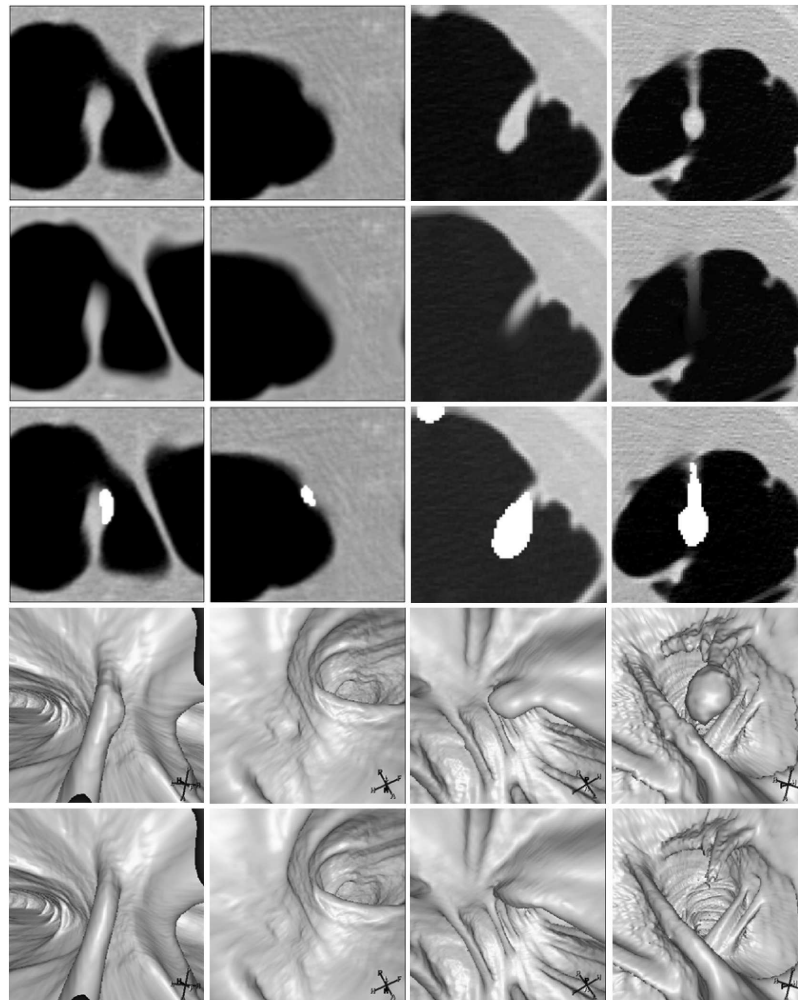


Figure 5.8: Typical results for four polyps. Each column shows the results for a different polyp. The first two rows show grey value cross sections before and after intensity deformation. The third row shows the segmentation masks which are obtained by thresholding the intensity change at a level of 100 HU. The last two rows show isosurface renderings (at -650 HU) of the polyps before and after intensity deformation.

optimal.

The linear logistic classifier involves estimating the posterior probabilities $p(\omega_i|x)$ instead of the class distributions $p(x|\omega_i)$. These posterior distributions are assumed to be the sigmoidal functions. This is a valid assumption when the classes are Gaussian distributed, or, as in our case, one of the class distributions is exponentially decreasing, while the other is more or less uniformly distributed. A maximum likelihood estimation is performed to find the linear direction in the data that best fits these assumed sigmoidal distribution functions. Using the posterior probabilities instead of the class-dependent distribution functions makes this classifier less sensitive to the large class imbalance.

As such, the problem is treated as a regression problem rather than a traditional two-class pattern recognition task. In other words, one searches for a linear direction in which the sigmoidal pdfs best describe the data. The performance of the classifier will be assessed by a 5-fold, 10 times repeated cross validation (see below).

5.3 Experiments and results

The proposed method is applied to the detection of colonic polyps in CT colonography data of 84 patients (see above). We will first show qualitative results. The sensitivity and specificity of the candidate detection step of the CAD system will be given for varying thresholds on the MaxIntChange feature. The results of the complete CAD system after classifier training will be given at the end of this section. We will include the results obtained by the method that involves an explicit (mesh) representation of the colon surface [130] for comparison. The FROC curves were calculated from a leave-one-patient out cross-validation. A polyp was counted as a true positive CAD detection if it was found in at least one of the two scanned positions (prone or supine).

The mean computation time per patient on a PC with a Pentium 4 processor (3.0 GHz) and 2 GB memory was 4 minutes.

5.3.1 Qualitative analysis

Fig. 5.7 illustrates how the intensities are modified during the deformation process and how this affects the position of the isosurface. The first row of grey valued images show cross sections through the polyp after 0, 10, 40, 80 and 160 iterations of (5.9). The second row shows images with an overlay of a color map of the intensity change for voxels with a change of more than 100 HU. The color bar gives an indication of the amount of change in the polyp compared to its surroundings (< 100 HU; the scale of the color bar was truncated at 650 HU). To appreciate the three dimensional structure, the last row shows isosurface renderings (at -650 HU.) that clearly show the gradual deformation of the polyp, while its surroundings stay almost unaltered.

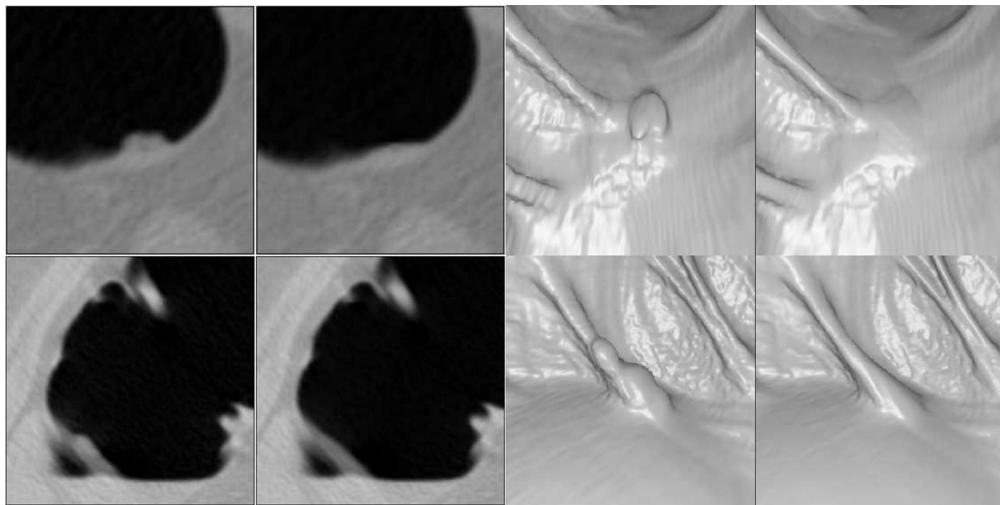


Figure 5.9: Each row shows a false positive. First row: example of stool. Air inside object is clearly visible on first image. Second row: stool on a fold. The original data is shown in the first and third column. The data after deformation by curvature flow is shown in columns two and four.

Fig. 5.8 shows the final outcome for a number of other polyps. The first two rows show grey-valued cross sections, respectively before and after the intensity deformation. The third row shows an overlay of the segmentation as obtained by thresholding the intensity change between the images in the first two rows at a level of 100 HU. The bottom two rows show isosurface renderings (at -650 HU) of the polyps before and after the deformation. The images demonstrate that the intensity deformation method yields probable estimates of the colon surface. This even applies to objects situated in highly structured surroundings, such as the polyp in the first column. The second column shows the result for a 6 mm polyp. It is situated on an almost flat background. The isosurface rendering containing the colon surface after deformation shows hardly any residual protrusion. The third column displays an elongated polyp on a strongly folded part of the colon. After deformation some residual protrusion can still be observed, albeit small compared to the original protrusion. The same holds for the polyp in the fourth column. This is a classical pedunculated polyp on a narrow stem. The head region is removed, while the stem remains.

Approximately 60% of the false positives are stool and 30% of the false positives are on folds. Among the remaining false positives are detections on the illeocecal valve. All these objects had a shape and structure that closely resemble a polyp (two examples are contained in Fig. 5.9).

5.3. EXPERIMENTS AND RESULTS

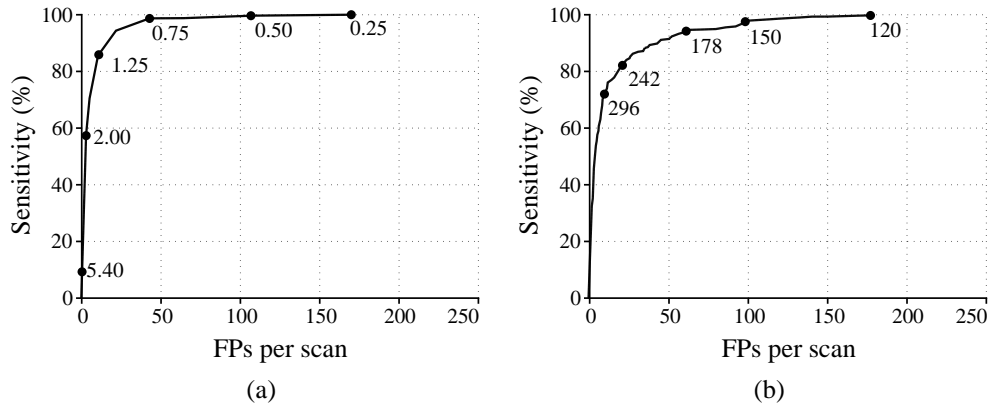


Figure 5.10: FROC curves showing the candidate detection sensitivity versus the number of false positives for (a) the mesh based and (b) the currently proposed technique. The numbers in (a) denote the threshold on the deformation field in mm and in (b) the threshold on the intensity change field in HU.

5.3.2 Performance of the candidate detection

Fig. 5.10 serves to show that our choice of thresholds is not affecting the detection sensitivity. Both figures (a and b) contain a free-response receiver operating characteristic (FROC) curve for the candidate detection step. Fig. 5.10(a) was obtained using the method that involves an explicit (mesh) representation of the colon surface [130] and Fig. 5.10(b) was based on the method presented in the current paper. The independent variable along the curves is the threshold on the displacement of the mesh, respectively the intensity change. In either case a lower threshold returns more candidate objects. Reversely, as the threshold is increased, fewer candidates are found, but also some polyps may be missed. For the full CAD system (see below) we have chosen a threshold for which at least 100% sensitivity is achieved on an independent training set. For the mesh based method this resulted in a threshold of 0.4 mm displacement. For the intensity deformation method we use a threshold of 100 HU on the intensity change. The smaller number of false positives of the mesh representation is due its description by fewer points (about 500000) than the implicit representation (about 10 million points). Notice that the large number of false positives at this stage is irrelevant: the overall performance of the system is determined after classifying the candidates (see below).

5.3.3 Results after classification

Fig. 5.11 shows the overall performance of both the proposed and the mesh based method [130]. The figure shows the performance for the detection of polyps for two

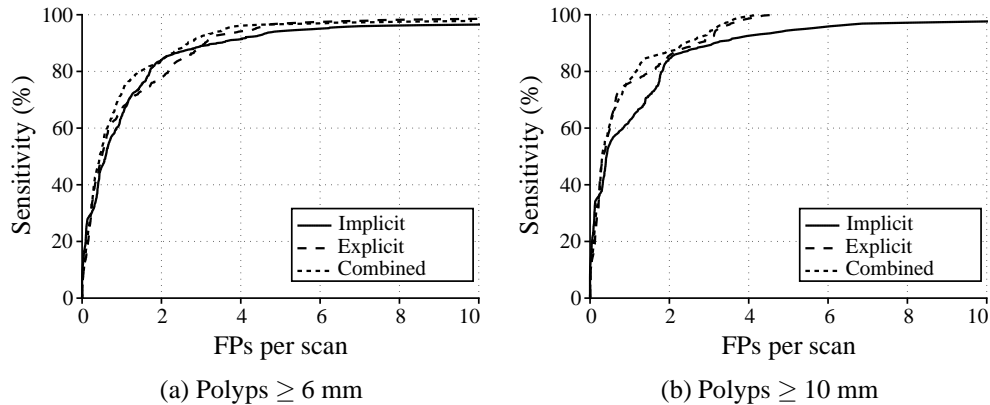


Figure 5.11: FROC curves depicting the performance of classification for the mesh based (explicit) and the currently proposed (implicit) technique. The FROC curves were computed by a five times repeated ten-fold cross-validation.

size ranges: larger than or equal to 6 mm (including those larger than 10 mm), and larger than or equal to 10 mm. Apparently, the performance of the two methods is comparable. Both techniques perform better on the larger polyps. A sensitivity of 95% for polyps ≥ 6 mm is achieved at an average false positive rate of 4-6 per scan. For polyps ≥ 10 mm, a sensitivity of 95% is obtained at about 4 false positives per scan.

For our data, approximately 50% of the false positives are stool and 40% are on folds. Among the remaining false positives are detections on the illeocecal valve. All these objects have a shape and internal structure that closely resemble a polyp (two examples are contained in Fig. 5.9).

5.3.4 A combined approach

In practice we found that particularly the false detections of both methods were to some extent uncorrelated. For instance, the mesh based method typically had false detections emanating from the partial volume effect (PVE) as it operates on a single isophote, whereas the current method was more robust because it took the full transition (air-tissue) into account. Reversely, the current method is inherently sensitive to intensity variations within tissue, especially in thin folds, whereas such problems are excluded in the mesh based method in which feature measurement is confined to the isosurface.

The two methods were combined as follows. The location of the candidates of both methods were compared. A consensus voting was used to accept candidates only if an overlapping candidate was found by the other method, in which case they were linked. Candidates with a vote from only one method were discarded. Fig. 5.12 confirms that

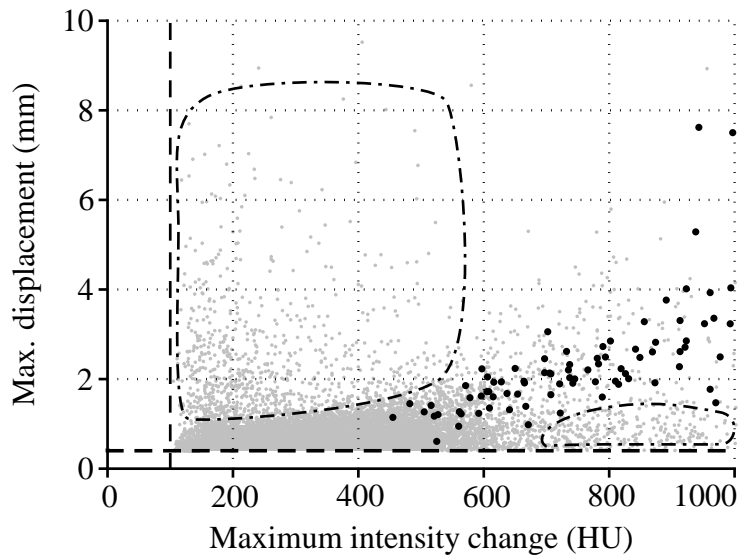


Figure 5.12: Feature space of the maximum displacement (explicit method) vs. the maximum intensity change (implicit method). The black dots correspond to polyps and the grey dots to false detections. Two regions (encircled by dash-dotted lines top-left and bottom-right) with false detections (grey dots) can be observed in which the depicted features are uncorrelated and complementary.

there is complementary information in the two methods. It contains a scatter plot of the MaxIntChange feature versus the maximum displacement of the mesh as obtained by the mesh based method. It can be seen that these correlate well for polyps (black dots). Two regions with false detections (grey dots) can also be observed in which the depicted features are uncorrelated (top-left and bottom-right in both graphs). One region has rather low MaxIntChange, but concurrently quite large maximum displacement of the mesh; another region is characterized by a large MaxIntChange, but a low maximum mesh displacement.

Fig. 5.11 also contains an FROC curve of the combined approach. It demonstrates improved performance by exploiting the complementary aspects of the two approaches particularly on polyps ≥ 6 mm.

5.4 Discussion and conclusion

A novel method was presented which detects polyps based on their protruding character irrespective of the actual shape. The method modifies image intensities at locations of

protruding objects. This is achieved by finding a steady state solution of a nonlinear PDE with the recorded image as input. We showed that the intensity change relates to the displacement of iso-contours. We also demonstrated how this relation is made invariant to the anisotropic resolution and sampling of the scanner. This allows for a simple segmentation of polyp candidates by applying a single threshold on the intensity change field. We proposed a measure for the detection of polyp candidates, which directly relates to polyp size, and not to polyp shape. This measure orders detected structures according to size which, in effect, keeps increasingly larger objects further away from the decision boundary. In other words, this limits the risk of missing large polyps. Also, our method does not make a specific choice for the scale for the computation of the 1st and 2nd order derivative operators. The iterative character of the method changes the intrinsic scale of the image (local and anisotropic): the aperture of observation (window size of the operation times the number of iterations) inherently increases.

We have chosen to adapt the convergence criteria of the posed PDE to the local data. Effectively, the deformation of a region stops when it does not expand anymore. This yields a stopping criterion which is data dependent and does not need user interaction. However, the criterion is rather strict as can be seen from Fig. 5.8 (third column), in which case the protrusion was not completely removed. A high noise level might prevent the algorithm from segmenting the entire polyp area. The (second order) Taylor expansion in a $3 \times 3 \times 3$ neighborhood will effectively deal with the noise practically encountered in low-dose (20 mAs) scans.

The performance of the method on so-called flat polyps requires further research.

6 Automatic polyp size measurement for CT colonography based on a protrusion estimation method

In this chapter a polyp segmentation algorithm is evaluated. The method is based on the technique proposed in chapter 4. The performance of the method is assessed by comparison to expert size measurements on phantom data and true polyps.

Published as:

C. van Wijk, J. Florie, C. Y. Nio, E. Dekker, A. H. de Vries, H. W. Venema, L. J. van Vliet, J. Stoker, F. M. Vos, *Protrusion method for automated estimation of polyp size on CT colonography*, AJR 2008 [127]

6.1 Introduction

Computed tomography (CT) colonography is a minimally invasive procedure that is advocated for the detection of colorectal cancer and polyps [131]. The size of a detected polyp is of primary importance for diagnosis and decision making since it relates to the risk of malignancy [74]. Accordingly, it is proposed that polyps with a diameter smaller than 6 mm require no further action, whereas polyps equal to and larger than 10 mm should be removed by colonoscopy [22, 146]. There is debate over the need for polypectomy for 6-9mm polyps. Surveillance for growth with CT colonography has been suggested as a safe alternative [146]. A reliable measurement technique is required in this scenario.

The focus in this paper is on accuracy and measurement variability. Accuracy is defined as the mean difference between a measurement method and the reference standard. A systematic error is a *significant* mean difference, which may be due over- or under-estimation by the method under investigation. The measurement variability is defined as the standard deviation of the mean difference. Notice that a method may be highly accurate, but at the same time have a large measurement variability or vice versa.

Lesion size is best defined as the single largest diameter of the polyp head, excluding the stalk. It is usually measured in 2D reformatted images or in endoluminal 3D display [15,90,144]. In either case, significant measurement variability was reported contingent on the experience of the observer and the viewing display used [15].



Figure 6.1: Phantom colon with plasticine objects. Photograph shows two halves of a cylinder inserted into a tight-fit second cylinder before scanning.

Automated techniques were introduced in order to enhance the measurement reliability [14, 26]. It was reported that automatic and manual 3D measurements were more accurate than manual 2D measurement on polyps in a human colectomy specimen [118]. The measurements were done on the resected specimen that was insufflated and submerged in a container with 0.9% saline solution. However, later it was found that 3D measurement had the largest systematic error in a study that included polyps

from a CT colonography study in which colonoscopy was the reference standard [16]. (Semi-)automatic measurement was found to have superior inter - and intra¹ -observer variability compared to manual 2D measurement with spherical, polyp-like phantom objects [14]. In other work, though, it was observed that automated and manual approaches have comparable inter-observer agreement [118]. The latter observation was confirmed in [16]. Several factors might explain the conflicting results: the types of objects used (phantom objects versus patient data), noise characteristics and scanner resolution, reader variations (inter and intra), and reliability of the reference standard in patient studies (colonoscopy).

This paper studies the accuracy and measurement variability of an automatic measurement technique [130] under varying scanning conditions using both phantom data and patient data. The performance of the algorithm was compared to both 2D and 3D manual measurement by human observers. For the phantom data this was done for two different slice thicknesses and two orientations of the phantom data in the scanner. All data was acquired using a 64 slice CT scanner. We hypothesize that the measurement variability of automatic measurement will be higher than the intra and inter-observer variability of a human reader.

6.2 Materials & methods

Phantom Data

A phantom consisted of an air-containing, lucite cylinder (length 10 cm; internal diameter 5 cm, see Fig. 6.1) into which artificial polyps were inserted. First, scans of the phantom containing seven hemispherical lucite objects (diameter 4-10mm) were made. Subsequently, 15 asymmetric objects from plasticine (largest diameter 4-19mm) were inserted in the phantom. The orientation of the main axes of the latter objects was arbitrarily chosen to be either parallel or orthogonal to the main axis of the phantom. The size of all objects (see Table 6.1) was measured by sliding calipers defining the reference standard. All scans were made with the phantom placed in a cylinder, 34 cm in diameter that was filled with water.

Patient Data

Polyps were included from scans of 10 patients (6 male, 4 female; mean age 59 years, range 30 - 74) selected from an ongoing CT colonography study. This study concerned patients at increased risk for colorectal cancer. All these patients underwent CT colonography succeeded by colonoscopy which was video taped. The selection of

¹The method presented in [14] is semi-automatic and requires the user to indicate a start point for polyp segmentation & size measurement. Presumably, the user interaction causes some measurement variability

CHAPTER 6. AUTOMATIC POLYP SIZE MEASUREMENT FOR CT COLONOGRAPHY BASED ON A PROTRUSION ESTIMATION METHOD

Acrylic Resin	Plasticine	Polyps
4	4	4.5
5	6	5
6	6	5.5
8	6	6.5
8	8	7
10	8	7.5
10	9	7.5
	10	8
	12	8
	12	8.5
	13	9.5
	13	10
	14	10.5
	14	11.5
	19	11.5
	19	13.5

Table 6.1: Reference Sizes (mm) of Phantom Objects and Polyps

patients for the present study was based on the polyp size measured during colonoscopy, which was required to be larger than 5 mm in diameter, irrespective of shape or location. All such polyps in patients that were examined in the period from 31 March 2006 to 30 August 2006 were included. For the present study the size of the polyps was re-measured on the colonoscopy video by two experienced gastroenterologists who were aware of the aim of the present study. The gastroenterologists were blinded for the CT colonography size measurements (see below) as well as the initial colonoscopy size measurements. The retrospective measurements were performed by comparison to an open biopsy forceps (size 8 mm) and to a caliper tool (size 10 mm) if available (4/16 cases). The mean retrospective size measurement served as the reference standard for polyp size. In the 10 patients 16 polyps were present, 9 polyps between 6 and 10 mm in diameter in 8 patients and 7 polyps with a diameter of 10 mm or larger in 5 patients (Table 6.1).

The colonoscopy findings were matched with the colonography data by a research fellow who was not involved in the present study. The CT colonography study was approved by the medical ethical committee of the hospital. The patients were informed a priori by letter as well as verbally of the study purpose and gave written consent.

CT Imaging

CT scanning of the phantom as well as the patients was performed on a 64-slice CT

6.2. MATERIALS & METHODS

scanner (Brilliance, Philips Medical Systems, Best, The Netherlands). The scan parameters: 100 mAs, 120 kV, 64 x 0.625 mm collimation, pitch 0.98, standard reconstruction filter ('C').

The phantom (Fig. 6.1) was scanned in two positions: parallel and under an angle of 45 degrees with respect to the axis of the scanner. This was done to obtain an oblique orientation of the main polyp axes with respect to the scanning direction. The field-of-view was fixed at 300 mm. The phantom in parallel orientation was scanned once with a slice thickness of 3.0 mm; the slice thickness was 0.9 mm for all other scans.

All patients drank 4 liter polyethylene glycol solution (KleanPrep, Helsinn Birex Pharmaceuticals Ltd, Dublin, Ireland), which is a hyperosmolar cathartic agent, combined with 4*50 ml tagging material (meglumine joxitalamate, 300 mg I/ml, Telebrix, Guerbet, Roissy, France) for bowel preparation, starting on the day before the examination. The colon was distended by automatic insufflation of CO₂ to a maximum of 20 mm Hg or maximum patient tolerance. The patients were scanned in both prone and supine position. The field of view varied between 286 and 440 mm². The slice thickness was 0.9 mm.

Automatic polyp measurement

The automated size measurement method is part of a Computer Aided Detection (CAD) scheme for automatic polyp detection [130]. The scheme uses a method which estimates the deformation the colon wall in order to digitally remove a presumed lesion. Schematically the method is explained in Fig. 6.2. Suppose that the points on the convex parts of the polyp (i.e. the 'protruding' part) are iteratively moved inwards. Effectively, this will 'flatten' the object. After a certain amount of deformation the surface flattening is such that the protrusion is removed. Thus, the surface looks like as if the object was never there. The amount of displacement is a measure of 'protrudedness'. A polyp is delimited by thresholding the deformation field. The size measurement is obtained by 'fitting' an ellipse [118]. The size of a polyp is measured by the largest ellipse diameter. The automatic polyp measurement was implemented on a proprietary, experimental version of the ViewForum workstation (ViewForum 6.1; Philips Medical Systems; Best; The Netherlands).

Manual polyp measurement

An abdominal radiologist (observer 1) and a research fellow (observer 2) independently measured the size of all objects. Observer 1 had a previous experience of more than 500 colonoscopy verified CT colonography examinations at the start of the study. Observer 2 had a previous experience of more than 350 of such examinations. Both had no knowledge of the reference standard, were blinded to any measurements by themselves for the same object (in the other scans) as well as each other's measurements. The measurements were performed on a commercially available image processing workstation (ViewForum, Philips Medical Systems, Best, The Netherlands).

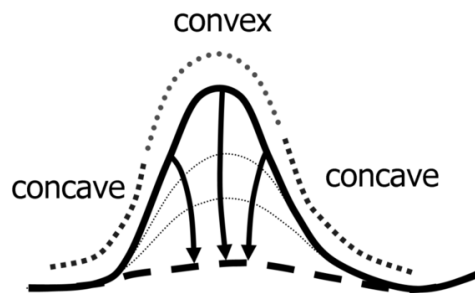


Figure 6.2: Diagram shows cross section of an idealized polyp with deformed surface for estimation of protrusion.

The 3D display was obtained by isosurface volume rendering into an enhanced 3D viewing method (unfolded cube images) [132]. The transfer function comprised a threshold at -650 HU (making the voxels below this threshold completely transparent). The 3D measurements were obtained using electronic calipers. The viewing software used for this study allowed manual navigation for optimizing the endoluminal vantage point and the placement of caliper points in the 3D space. The observers were instructed to maneuver orthogonally over the object and measure the maximum diameter.

The 2D measurement also required navigating orthogonally over an object. Subsequently, a reformatted cross section through the object was shown, that could be rotated to identify the longest linear dimension. A window/level setting of 1300/0 for the phantom and 1250/-50 for the patient data was applied. The difference accounts for the slight increased attenuation of the plasticine, which was measured to be 100 HU. The observers could freely zoom in/out. The size of the object was determined using electronic calipers. The interval between 2D and 3D measurement on the same object was a few hours during which approximately 100 other measurements took place. This setup was chosen to avoid observer bias. The cases were presented in random order. Thus, the order in which 2D measurements were done differed from the order in which the 3D measurements were done. The objects in the phantom and the true polyps were measured twice by both observers using both methods. The interval between two measurements (2D and 3D) on the same object in the same scan was at least four weeks. Recall bias was further reduced as both observers evaluated more than 50 other CT colonography examinations during the interval periods.

Outcome parameters and statistical analysis

The measurements were used for these assessments: The accuracy and measurement variability of the observers and the algorithm was determined by comparing the first measurements on the phantom objects in parallel orientation with the reference standard. Additionally, the accuracy of the measurements on true polyps was determined

6.3. RESULTS

by comparison to the retrospective colonoscopic size measurements. Moreover, we counted how often the critical category of a polyp was changed by the measurement (for instance a polyp measured to be 6-9 mm by an observer and 10 mm or larger by the reference standard). The intra-observer variability was determined by comparing the initial measurements on the scan with the phantom in parallel orientation with those made on the same scan four weeks later. In the same way the intra-observer variability of the measurements on the true polyps was determined. Inter-observer variability was explored by comparing the first measurements of the two observers on the phantom in parallel orientation. Likewise, the inter-observer variability of the first measurements on the true polyps was determined. The variability due to differences in the orientation of the phantom in the scanner was assessed by comparing the first measurements from the observers and the algorithm on the scans with the phantom in parallel and oblique orientation. The influence of slice thickness was studied by comparing the measurements from the observers and the algorithm on the phantom scans with slice thickness of 0.9 mm and 3.0 mm. Student's t-test was applied to assess any systematic mean difference between paired measurements. The standard deviation of the mean difference was calculated to express the measurement variability. A Bartlett test [101] was first applied to test the assumption that standard deviations across measurement series were equal (e.g. the standard deviations as in Fig. 6.3a). If the 0-hypothesis of equal variance was rejected, then the (squared) standard deviations were compared by means of the F-test.

Bland-Altman plots² were used to visualize potential trends in the difference between measurements and/or trends in the standard deviation. A trend in the Bland-Altman plot was detected by a linear regression analysis. Such a trend was considered significant if the regression coefficient differed significantly from zero by a t-test. The outcomes were stratified by observer and by the type of measurement method (2D or 3D). In any case, a p-value less than 0.05 was considered to indicate a significant difference.

6.3 Results

For clarity only the most prominent outcomes are reported; any comparison that is not explicitly reported, did not yield a significant difference. Results on phantom data is consistently reported first, followed by results on patient data. Notice, that the Bland-Altman plots (figures 6.3 to 6.5) explicitly include mean differences and corresponding standard deviations in the plot legends.

6.3.1 Accuracy and measurement variability

Phantom data

²The Bland-Altman plot depicts horizontally the average of two corresponding measurements and vertically the difference.

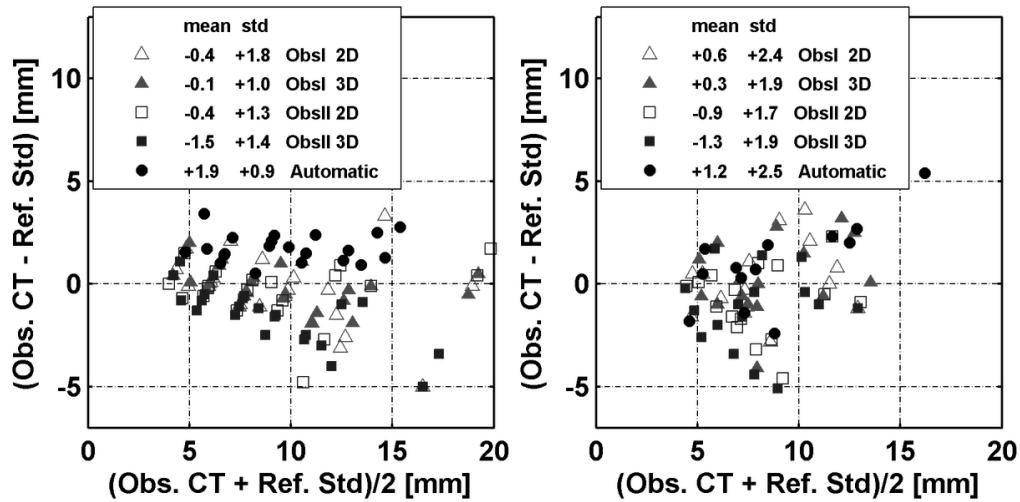


Figure 6.3: Observer performance on CT with respect to reference standard. a and b, Bland-Altman plot shows data for phantom (a) and for true polyps (b).

Fig. 6.3(a) shows the Bland-Altman plot in which the measurements on the phantom data are compared to the reference standard. A systematic error was found for the automatic measurement and for observer 2, 3D ($p < 0.01$, for both). The measurement variability of the automatic method was significantly smaller than all those of the observers, except for observer 1, 3D (automatic: 0.86mm; O1: 1.76mm (2D), 0.96 (3D); O2: 1.34mm (2D), 1.45 mm (3D)).

Patient data

The Bland-Altman plot on the patient data is shown in Fig. 6.3(b). On patient data, observer 2 made systematic errors with both 2D and 3D measurement. There were no statistically significant differences between the automatic method and both 2D and 3D manual measurement regarding measurement variability. Linear regression revealed a significant trend in the automatic measurements, i.e. a significantly larger measurement error with a larger polyp size. All the approaches changed the critical category of 4/16 polyps, except for observer 2 3D, by whom 1/16 polyps changed category.

6.3.2 Intra-observer variability

Phantom data

Fig. 6.4(a) shows Bland-Altman plots of intra-observer variability on the phantom data. The phantom objects resulted in intra-observer variability that was not significantly different between the observers for both 2D and 3D. Patient data. Likewise, Fig. 6.4(b)

6.3. RESULTS

shows the outcome for the polyps. Again, the intra-observer variability was not significantly different between the observers for 2D, but the variability was significantly different (p-value: 0.035) for 3D measurement.

6.3.3 Inter-observer variability

Phantom data

Fig. 6.4(c) shows the Bland-Altman plot illustrating the inter-observer variability for 2D and 3D measurement for the phantom data. A statistically significant mean size difference of 1.43 mm on the phantom objects was found between the 3D measurements of the two observers ($p < 0.01$). The inter-observer variability of 2D measurement was significantly larger than the variability of 3D measurement (p-value $\ll 0.01$).

Patient data

Fig. 6.4(d) shows the Bland-Altman plot of the inter-observer variability for 2D and 3D measurement on true polyps. Observer II underestimated the polyp size compared to observer I (by 1.48 mm for 2D and 1.54 mm for 3D). The inter-observer variability of 2D and 3D measurement were not significantly different for the true polyps.

6.3.4 Orientation of the phantom in the scanner

Phantom data³.

There were no significant mean size differences between the measurements on the phantom objects in different orientation, neither for the automatic method nor for the observers. Fig. 6.5a shows the Bland-Altman plot illustrating the variability due to varying object orientation. The measurement variability of the automatic measurements was significantly smaller than the variability of the observers for 2D measurement (for both: $p < 0.001$). The variability of the automatic measurement was also smaller than the variability of the 3D measurements by the observers, but the difference was not significant.

6.3.5 Slice thickness

Phantom data⁴.

Fig. 6.5b shows the Bland-Altman plot illustrating the variability due to different slice thicknesses. Statistically significant mean size differences were only found between the 3D measurements both observers ($p < 0.01$). The variability of the automatic measurement was not significantly different from that of the observers.

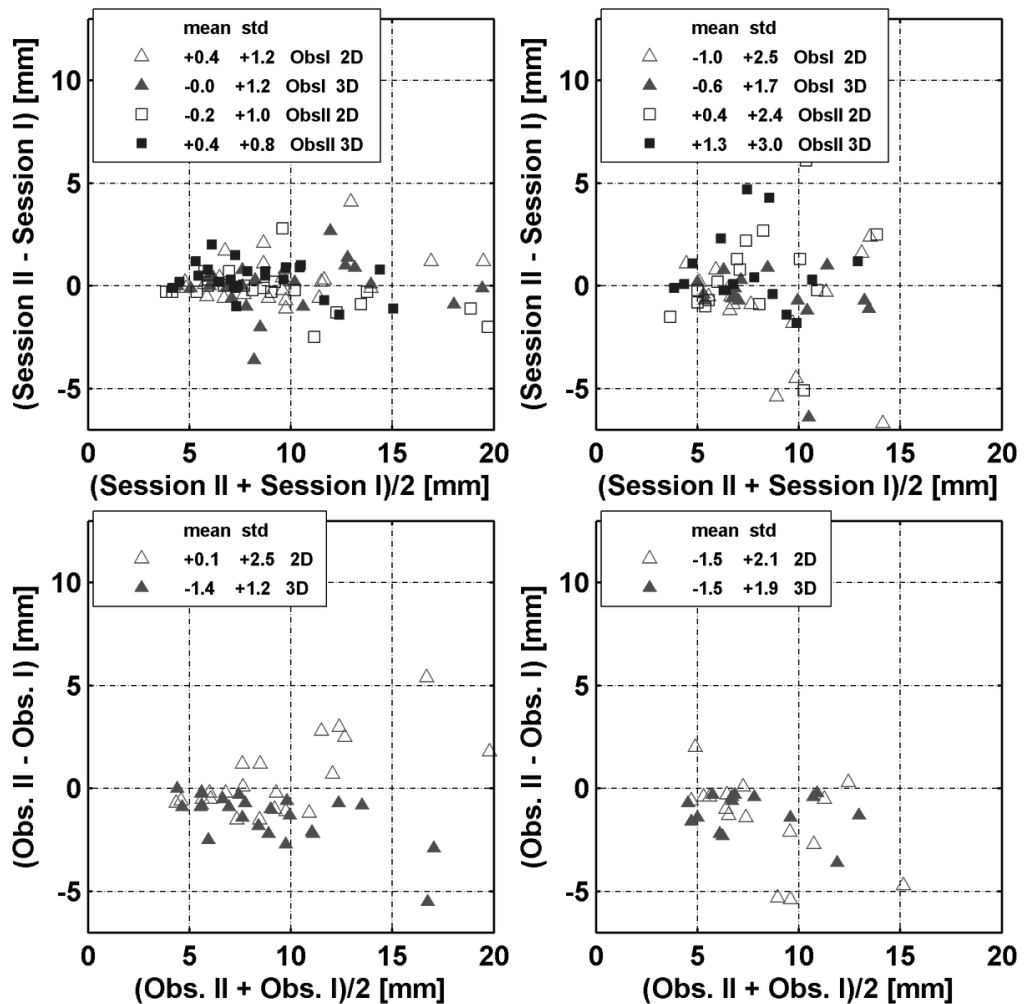


Figure 6.4: Intra observer and interobserver variability of 2D and 3D measurements. a and b, Bland-Altman plots show intra observer variability of two repeated measurements of same phantom object (a) and of same true polyp (b). c and c, Bland-Altman plots show interobserver variability of corresponding measurements of phantom object (c) and of true polyp (d).

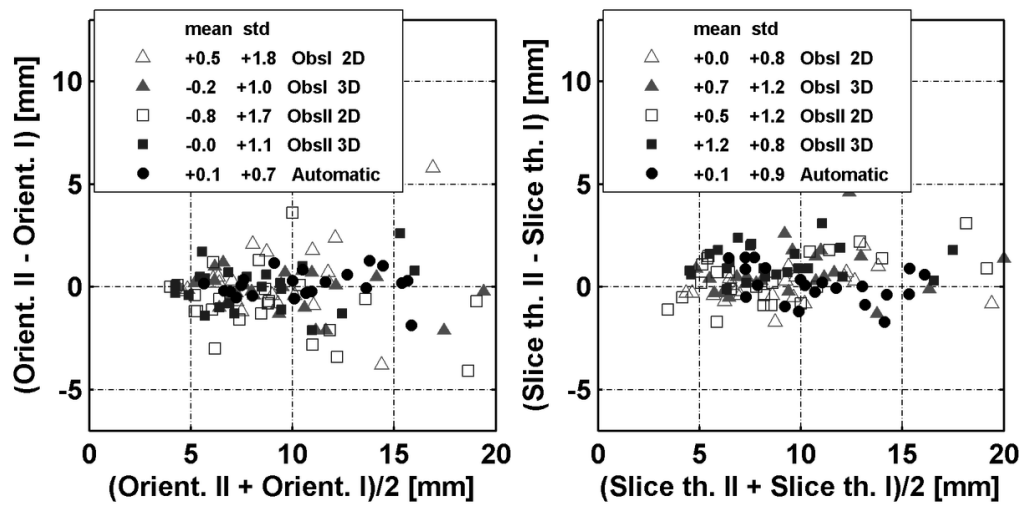


Figure 6.5: a and b, Bland-Altman plot shows measurement variability due to change in orientation of phantom in scanner for automated method and manual measurement by both observers (a). Bland-Altman plot shows measurement variability in phantom data due to differences in scanner anisotropy for automated method and manual measurements by both observers (b).

6.4 Discussion

Our study shows that for phantom data the measurement variability of the automatic method was smaller than the measurement variability of the observers when either the orientation⁵ of a phantom or the slice thickness⁶ was varied. This may show that the reader variation of the automatic method is less sensitive to such data variations. Moreover, the automatic approach had a smaller variation than the observers in comparison to the reference standard⁷. This difference could be explained by the intra observer variation, which is nonexistent for the automatic method by definition. We found that one observer made a systematic error (i.e. consistent undersizing, for both 2D and 3D measurement) on the patient data. We attribute this to a different perception of the polyp boundaries by this observer. The automatic method had the largest measurement variability, although it did not differ significantly from any manual approach. The large variability may be explained due to a systematic error which increases with polyp size (Fig. 6.3b). The measurement variability (Fig. 6.3b) on polyps is considerably larger than the corresponding variability found on the phantom objects (Fig. 6.3a). The increase could be explained by imprecision and inaccuracy in the reference standard (colonoscopy).

⁵Significant for 2D, not significant for 3D measurement

⁶Significant for 3D, not significant for 2D measurement.

⁷Significant for one observer, not for the other.

Conventionally, polyp size is measured by the single largest polyp diameter either in 2D or 3D, although other ways to quantify size (e.g. by means of the volume) were explored as well (see e.g. [14, 139]). Conflicting results were reported previously [16, 90, 118] regarding whether 2D or 3D measurement is preferred and how such manual measurement relates to automatic methods. Pickhardt et al found that manual 3D measurement is significantly more accurate than manual 2D measurement [90]. The observed underestimation of 2D measurement was attributed to the suboptimal alignment of the standard orthogonal MPR's to the polyp axis. The study was based on phantom data as well as patient data (colonoscopy served as the reference for the polyps). We reduced the pitfall of selecting the proper cross-section in 2D measurement by letting the observers navigate orthogonally over a polyp in the 3D display. A reformatted cross section through the object was shown, that could be manipulated to find the longest dimension of the object. Simultaneously, the orientation of the reformatted cross-section was visualized in the 3D display, but the measurement had to be done in the reformatted image.

Burling et al found that the greatest measurement error was made by the manual 3D approach [16]. Burling indicated that 3D measurement is prone to subjective cursor placement, e.g. due to varying angle and direction from which a lesion is viewed [15]. Both studies by Burling included true polyps and colonoscopy served as the golden standard. In our experiments, the observers were aware of the difficulty of positioning electronic calipers in 3D views. Accordingly, they checked their placement carefully. Then again, Taylor et al found that automated and manual 3D polyp measurements were more accurate than manual 2D measurement in a human colectomy specimen (irrespective of observer experience) [118].

We opted to include both phantom data and patient data since observable differences regarding 2D versus 3D measurement might relate to the types of objects that were used (phantom vs true polyps) and the accuracy of the reference standard (sliding calipers vs. colonoscopic size measurement). Our 2D measurements on the phantom data show a larger variability than our 3D measurements as orientation of the objects in the scanner is varied. Moreover, we found larger variability for 2D measurement than for 3D measurement on the phantom data by comparison to the reference standard. The latter finding confirms previous reported results of Pickardt et al [90]. We hypothesize that more complex manipulations are needed for 2D (manual) measurement. Still, results for the true polyps did not reveal differences, which could be due to the inaccuracy of colonoscopic measurement (see below). The encountered indifference concerning patient data agrees with Taylor et al [118].

One of the observers made a systematic error (under estimation) for the polyps. Such a difference in observer measurement confirms the findings of Burling et al [15].

The automatic measurements on the phantom show a smaller variability than both manual methods, in the comparison with the reference standard. However, the automatic

6.4. DISCUSSION

method has a systematic error on the phantom data. The reported bias of the automatic measurement may be partially explained by the low threshold (-750 HU) applied to segment the phantom (compared to a value of -650 HU for manual measurement as in [132].) Certainly, the objects will appear larger as the threshold is lowered, but its value of -750HU yields optimal sensitivity and specificity for automated polyp detection in a CAD system [130].

Any systematic error might be corrected for by a proper calibration. This holds for the measurements by observers as well as an automatic approach. The systematic difference between the observers regarding 2D as well as 3D measurement signifies that separate correction values may be needed. In other words, it might indicate that observers need to be calibrated individually to avoid systematic errors. Calibration would require a procedure in which an established collection of objects (the size of which is known exactly) is measured to determine the accuracy attained by an observer. Subsequently, an eventual systematic error should be subtracted from all subsequent measurements.

Previously, Burling et al [14] described a “fully automatic” technique that is initiated by two software seeds and proceeds in a region growing scheme. Another (technical) paper introduced an automatic technique that starts by placing a seed point on the polyp, from which a patch is grown over the polyp surface [25]. The current method merely requires user interaction to indicate a specific protrusion. By definition, repeated measurement using our approach, either initiated by the same or another observer, yields exactly the same result irrespective of seed placement. Since other programs use different methods to measure size, the current results for automated measurement are limited to the software used by us.

A limitation of our work is in the restricted number of polypoid objects that we used. Clearly, an unlimited number of shapes may be encountered in clinical practice. For practical reasons we selected a limited number of phantom shapes that we considered relevant for the hypothesis tested. Notice that no criteria regarding lesion shape were applied to select the true polyps. Also, a limitation is in the slightly denser material of phantom objects compared to true polyps (by approximately 100 HU). We used $w/l = 1300/0$ in the phantom and $1250/-50$ on patient data to have a similar appearance upon 2D measurement and minimize the effect. We hypothesize that the influence on the automated method may be neglected since the algorithm does not use the underlying CT values.

Another limitation is the precision of the reference standard for the colonoscopy measurements, which increases the total variation for all readers (automatic and manual) on patient data. It is well known that the colonoscopy measurements come with errors [30, 95]. Consequently, the reported standard deviations of measured true polyp size compared to the reference standard may be too pessimistic.

We conclude that our work indicates that there is reduced variability in measured polyp size by the automatic method in phantom data. The automatic measurement has

a variability that is in the same range as manual methods on patient data. A clear advantage of the automated method is that it does not suffer from intra-observer variation. Moreover, the automatic method may be calibrated once, whereas each observer may require individual calibration. Therefore, automatic size measurement may well contribute to a practical evaluation strategy.

Acknowledgments

We would like to thank Philips Medical Systems Nederland B.V. for providing the View-Forum prototyping software and the feedback on our work.

7 Computer aided detection of polyps in CT colonography using logistic regression

We present a computer aided detection (CAD) system for computed tomography colonography that orders the polyps according to clinical relevance. The CAD system consists of two steps: candidate detection and supervised classification. The characteristics of the detection step lead to specific choices for the classification system. The candidates are ordered by a linear logistic classifier (logistic regression) based on only three features: the protrusion of the colon wall, the mean internal intensity and a feature to discard detections on the rectal enema tube. This classifier can cope with a small number of polyps available for training, a large imbalance between polyps and non-polyp candidates, a truncated feature space, unbalanced and unknown misclassification costs, and an exponential distribution with respect to candidate size in feature space. Our CAD system was evaluated with data sets from four different medical centers. For polyps larger than or equal to 6 mm we achieved sensitivities of respectively 95%, 85%, 85%, and 100% with 5, 4, 5, and 6 false positives per scan over 86, 48, 141, and 32 patients. A cross-center evaluation in which the system is trained and tested with data from different sources showed that the trained CAD system generalizes to data from different medical centers and with different patient preparations. This is essential to application in large-scale screening for colorectal polyps.

Based on:

V. F. van Ravesteijn, C. van Wijk, F. M. Vos, R. Truyen, J.F. Peters, L. J. van Vliet, Computer-Aided Detection of Polyps in CT Colonography Using Logistic Regression, IEEE-TMI, 29(1), 2010

7.1 Introduction

Cancer is the second leading cause of mortality due to cancer in the western world [3]. Paradoxically, perhaps, is that it is preventable for a large part or at least curable if detected early. Adenomatous colorectal polyps are considered important precursors to colon cancer [11, 73, 85]. It has been shown that screening for such polyps can significantly reduce the incidence of colon cancer [31, 136]. Computed tomography (CT) colonography (CTC) is a rapidly evolving technique for screening, but the interpretation of the data sets is still time-consuming. Computer aided detection (CAD) of polyps may enhance the efficiency and also increase the sensitivity. This is specifically important for large-scale screening. Recent studies show that the sensitivity of CAD systems is already comparable to the sensitivity of optical colonoscopy [78, 105, 109] and radiologists using CTC [142].

The best indicator of the risk that a polyp is malignant or turns malignant over time is size [90]. The consensus [146] is that patients with a polyp of at least 10 mm must be referred to optical colonoscopy for polypectomy and it is advised that diminutive polyps (≤ 5 mm) should not even be reported [88, 117]. There is still debate over the need for polypectomy for 6–9 mm polyps. Surveillance for growth with CT colonography has also been suggested.

7.1.1 Related work

CAD algorithms for polyp detection in CT colonography usually consist of candidate detection followed by supervised classification. Candidate detection aims at 100% sensitivity for polyps larger than 6 mm which goes at the expense of hundreds of false positives (FPs) per scan. The task of supervised classification is to reduce the number of detections to about a handful without sacrificing the sensitivity too much.

For the detection of polyp candidates, Summers et al. [103, 104] proposed to use methods from differential geometry in which the principal curvatures were computed by fitting a fourth order B-spline to local neighborhoods with a 5 mm radius. Candidates were generated by selecting regions of elliptic curvature with a positive mean curvature [104]. Yoshida et al. [141, 143] used the shape index and curvedness to find candidate objects on the colon wall. The shape index and curvedness are functions of the principal curvatures of the surface, which were computed in a Gaussian-shaped window (aperture). Alternatively, Kiss et al. [59] generated candidates by searching for convex regions on the colon wall. Their method fitted a sphere to the surface normal field. The type of material in which the center of the fitted sphere was found (in tissue or in air) determined the classification of the surface as either convex or concave. As a result, roughly 90% of the colon wall was labeled as concave, that is 'normal'. Subsequently, a generalized Hough transformation using a spherical model was applied to the convex surface regions. Candidate objects were generated by searching for local maxima in the

Table 7.1: Properties of the data sets

Data set	Medical Center	Slice Thickness (mm)	Fecal Tagging	Scans per Patient	Number of Patients	Number of Polyps ≥ 6 mm	Ref.
'A'	AMC / Amsterdam	3.2	No	2	86	59	[123]
'B'	WRAMC / Washington, DC	1.2	Yes	2	48	28	[87]
'C'	UW / Madison, WI	1.2	Yes	2	141	176	[55]
'D'	Charité / Berlin	1.0	Yes	1	32	8	[45]

parameter space of the Hough transformation. Kiss et al. characterized the shape of the candidate by comparing the spherical harmonics with those of the polypoid models in a database [57].

Apart from the different candidate detection algorithms, there is a wide variety in the design of the pattern recognition system, ranging from low-complex systems like linear discriminant classifiers to classification systems using multiple neural networks. Yoshida and Näppi used linear and quadratic discriminant classifiers [77, 141, 143] as well as Jerebko et al. [49]. Wang et al. [134] uses a two-level classifier with a further unspecified linear discriminant classifier in the second level. The first level of this classifier consisted of a normalization procedure, which was specially designed and had four parameters. Sundaram et al. [111] classified the candidates based on a single heuristically designed score using curvature information of the candidate patches. Göktürk et al. [35] employed a support vector machine for classification, in which it was assumed that after a transformation by the kernel function, the data were linearly separable. This implicitly required minimal mixing between polyps and false detections. Jerebko et al. [50] and Zheng et al. [148] used a committee of support vector machines. Neural networks were also used by Jerebko et al. [50] and Näppi et al. [65, 78] for classification, and by Suzuki et al. [113] for the reduction of false detections on the rectal enema tube.

To conclude, many different proposals for a classification system for computer aided detection of polyps have been presented. However, the motivation for a specific design of the classification system is often unclear. Moreover, proper comparison between classification systems is difficult due to the different candidate detection systems and feature extraction methods. One may reason that the optimization of complex classification systems (with large number of parameters or features) may be complicated by the limited availability of training examples. This could lead to overtraining to a specific patient population or patient preparation.

A steadily growing number of papers (e.g. [10, 23, 35, 49, 57, 59, 62, 77, 105, 130, 134, 141]) reported on the performance of polyp detection algorithms (see Yoshida and Näppi [142] for a review on CAD systems for CTC). However, the results can not easily be compared due to large differences in the data sets used for evaluation (see also Section 7.2.1).

7.1.2 Objective

Candidate detection typically renders a lot of candidates to sustain maximum sensitivity. Hence, the number of objects from the target class (polyps) is relatively low. This large imbalance of the prevailing classes typically hampers classifier design and training. A further complication is that the misclassification costs for objects from the two classes are unknown and certainly very different. This paper discusses the consequences of these characteristics for the design of the classification system.

We aim to design a novel, low-complex, classification system that orders the polyps

according to clinical relevance. It implicitly takes into account that the misclassification costs of polyps increase with lesion size. In other words, larger polyps are more important than smaller ones and the problem is not considered as a mere two-class classification task, but rather as a regression problem. With this in mind, we distinguish two types of features in the design of the classification system. First, there are features that facilitate an ordering of the candidates. These are the features that directly relate to the lesion size. Second, there are features which will be shown to render a Gaussian distribution. In order to keep the classifier simple and to prevent the use of complex combination strategies, these features are mapped into features of the first type by a Mahalanobis distance (MD) mapping. This strategy is used to discard outliers and mimics the use of a Gaussian one-class classifier [114]. It will be shown that this two-level classification system is effective over data from various sources.

The technical novelty of our paper is to approach the classification task as a regression problem. Such a strategy requires that features are ordered according to relevance. A mechanism is introduced to map features that are not ordered as such into features that do have the ordering property. It will be demonstrated that the Mahalanobis distance to the target class mean is appropriate for the current problem. Imposing the ordering may be achieved for any other problem provided that the distance to the most typical representation of the target class can be defined.

7.2 Data description and feature design

A CAD system for CTC starts with the acquisition of CT colonography data. In these data, candidate objects are detected and segmented. The segmented candidates are typically characterized by features describing, for instance, the candidate's shape and its internal intensity distribution. Such data serve as input for the classification system. All preprocessing steps will be addressed in this section.

7.2.1 CT colonography data

Data sets from four different medical centers were used to evaluate the performance of our system. Data sets from different sources differ in polyp prevalence, the patient preparation, the scanning protocol, the protocol for determining the ground truth, and the type of rectal tube used for colon distension during CT examination¹. An arbitrary number of patients were randomly selected from each source, irrespective of the number of polyps and their shape. The most important characteristics of the data sets are shown in Table 7.1. More details may be retrieved from the references included in the table. All patients adhered to an extensive laxative regime. The reference standard (ground truth)

¹Information about the patient preparation can be retrieved from the reference included in Table 7.1. However, the specific data set we used is not described.

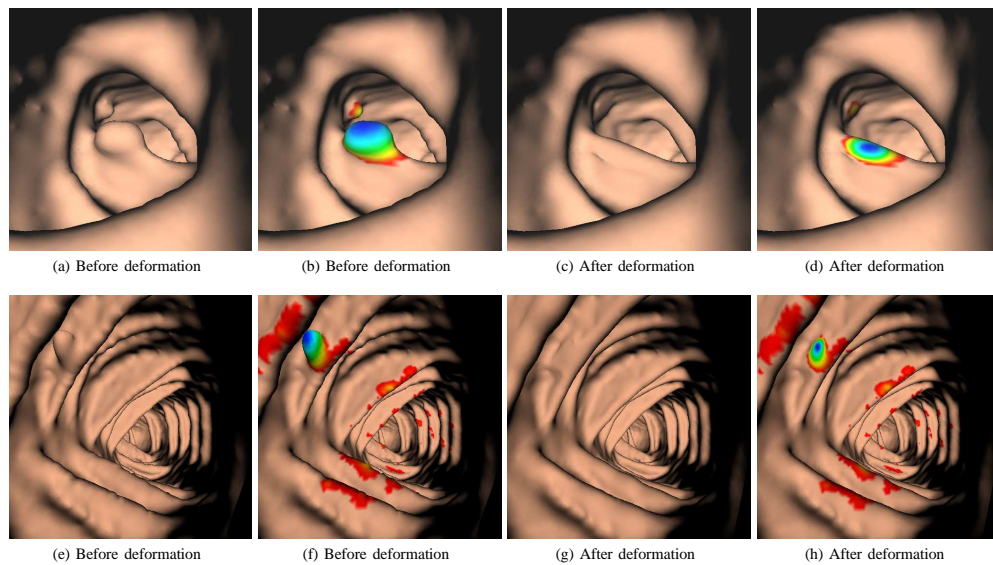


Figure 7.1: The candidate detection method applies a non-linear 'flattening' operation to the colon wall. The protrusion field is defined as the difference in position of the colon wall before (a–b,e–f) and after (c–d, g–h) application of the operation. The coloring (b,d,f,h) indicates the protrusion of the mesh vertices of detected candidates (blue denotes a large protrusion and red denotes a protrusion of 0.2 mm, i.e. the low hysteresis threshold). Notice that the folds are hardly affected by the operation.

for data sets 'A', 'B' and 'C' was optical colonoscopy. An expert radiologist served as the reference for data set 'D'. Radiologists retrospectively indicated the location of polyps by annotating a point in the 3D data set based on the reference standard. The candidate segmentations (see below) were labeled by comparison to these annotations. Data sets 'A', 'B' and 'C' consisted of scans in both prone and supine positions. A polyp was counted as a true positive CAD detection if it was found in at least one of the two scanned positions. Only dataset 'A' has been used during development of the system.

7.2.2 Candidate detection

Polyps are often described as objects that protrude from the colon wall. For that reason, the candidate detection method is designed to detect all objects that protrude from the colon wall, irrespective of their shape. Suppose that the points on the convex parts of a protruding object are iteratively moved inwards. Effectively, this will 'remove' the object. After a certain amount of deformation, the protrusion is completely removed and the colon wall appears 'normal'. The amount of deformation as a result of the operation is a measure of 'protrudedness'. Fig. 7.1 illustrates this process by showing images before and after application of the non-linear 'flattening' operation.

Practically, the colon wall was represented by a triangle mesh, which was obtained by thresholding the CT colonography data at -750 Hounsfield units (HU). A non-linear PDE [130] was solved to remove all protruding structures from the mesh that displayed a positive second principal curvature. A similar approach that acts directly on the grey valued image is presented in [129]. In this procedure, the global shape of the colon including the folds was retained, since these structures display a second principal curvature that is smaller than or equal to zero. The protrusion field was computed by the position difference of the mesh vertices before and after processing. Subsequently, hysteresis thresholding was applied to this field to detect and segment the candidates. The high threshold on the protrusion was 0.4 mm and determines the sensitivity. The value of 0.4 mm was selected since it yields 100% sensitivity per polyp annotation in our training set. All retained regions of the colon surface were augmented by adding the adjacent mesh points with a protrusion of at least 0.2 mm (the low threshold). The regions thus obtained form the segmented candidates.

7.2.3 Features

Radiologists that evaluate CTC data primarily use two properties of a candidate for classification: the shape and the voxel intensities inside the candidate. There is still debate about the optimal way to analyze CTC data. Radiologists using the 3D rendering of the colon (virtual colonoscopy) detect polyps based on shape, but they will often fall back to the 2D representation (grey values) before a final decision is made. Using

the 2D representation, both the internal intensities and the shape are assessed, although shape is often hard to extract from the grey-value images. The features used in the presented CAD system are based on the same two properties that are primarily used by radiologists.

Shape was previously described by the shape index and curvedness [143], mean curvature, average principal curvatures and sphericity ratio [103, 104] and spherical harmonics [57]. An alternative method to measure shape, which is based on the protrusion field, will be introduced (see Section 7.2.3, below).

The internal intensity of the candidates has been found before to be a discriminative feature to discard a large number of false detections [10, 49, 77, 134]. It may be expected that due to the partial volume effect false detections arise that have low internal intensity. False detections that are stool often have air inside, which also lowers the intensity. Such information about the candidates will be included through statistics on the voxel intensities inside the object (see Section 7.2.3, below).

At last, it was experimentally found that many false positives turned out to be detections on the rectal enema tube (RET) (previously also reported in [46, 113]). Therefore, a third feature will be proposed to discard such false detections (see Section 7.2.3, below).

Shape feature from protrusion field

Polyps are conventionally characterized by the single largest diameter, excluding the stalk [90, 127]. However, Fig. 7.2(a) shows that this measure does not distinguish polyps from false detections well. It appears that especially among the less protruding candidates (≤ 2 mm), the candidates with the larger diameters are predominantly false detections. Alternatively, it might be natural to select the maximum protrusion of a candidate as a feature, but it appears that a lot of polyps have only modest protrusion. As an illustration, Figs. 7.2(c) and (d) show two candidates that have approximately the same maximum protrusion but a completely different appearance. The first candidate (candidate 'c') has a large diameter, but does not resemble a polyp at all, whereas the second candidate (candidate 'd') with a small diameter does so. To conclude, a large diameter relative to the maximum protrusion indicates a non-polypoidal shape (candidate 'c') and a small diameter or a relative low protrusion points to a small clinically irrelevant candidate. A feature that is derived from the thresholded protrusion field should therefore include the size of a candidate as well as the ratio between the largest diameter and the maximum protrusion. Moreover, the feature should characterize the whole segmented area instead of the extrema (like the largest diameter or the maximum protrusion).

We designed a feature that takes into account both the protrusion as well as the lateral size of the object. Effectively, it measures the percentage of the area of the candidate that has a protrusion larger than a certain threshold T . This feature is further denoted as Φ_T . A large circumference as well as shallow edges lead to relatively large areas

7.2. DATA DESCRIPTION AND FEATURE DESIGN

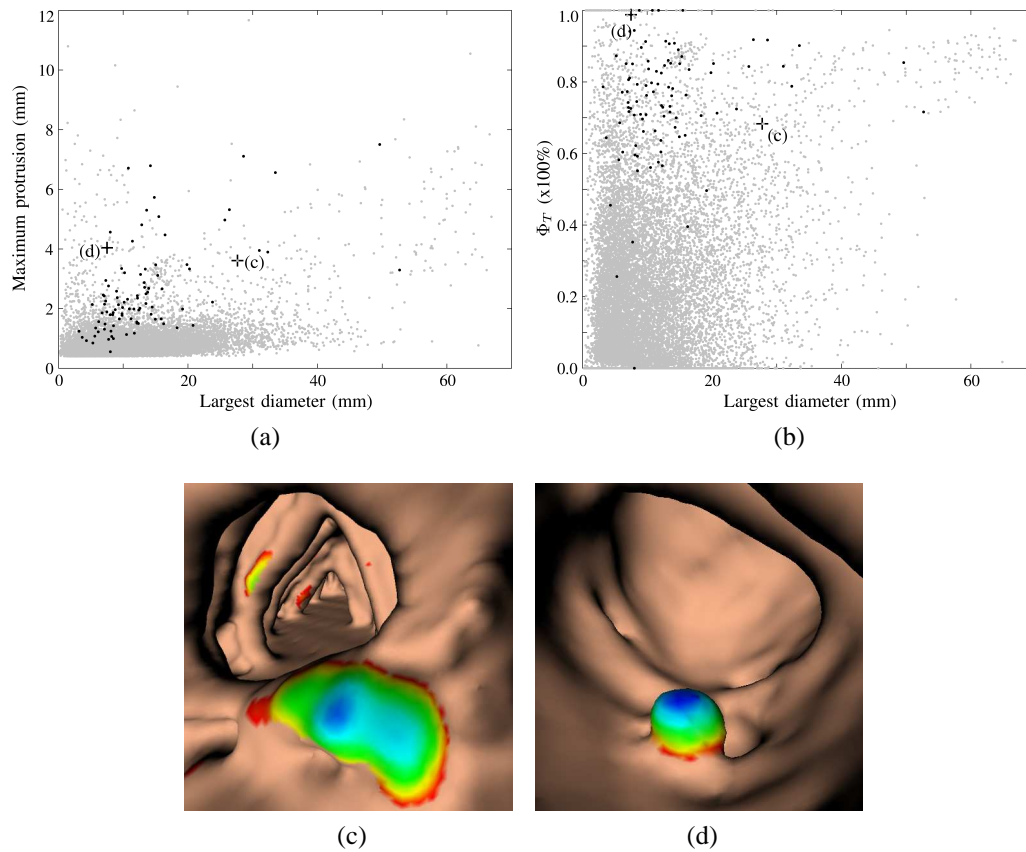


Figure 7.2: (a)–(b) Scatter plots of features calculated for data set 'A'. Grey dots denote false detections and black dots indicate polyps ≥ 6 mm. Note that each polyp may appear as two separate dots in the scatter plot, since each patient is scanned twice. (a) The maximum protrusion versus the single largest diameter of a candidate. The threshold of the candidate detection can be seen at a maximum protrusion of 0.4 mm. (b) Φ_T ($T=0.6$ mm) versus the largest diameter. (c–d) Two candidates with the same maximum protrusion that are ordered differently according to Φ_T .

with protrusion below T and result in a low response. Thus, this feature favors compact objects with steep edges. Fig. 7.2(b) shows that according to Φ_T ($T=0.6$ mm) candidate 'd' is indeed favored over candidate 'c'. Ordering the candidates based on Φ_T is thus expected to improve the performance of the CAD system over simply using the maximum diameter alone.

Intensity features

Consider all mesh vertices that are part of the segmentation mask of a candidate object (see Section 7.2.2). For each vertex, a weighted average of colon wall intensities was calculated along the line segment from the vertex under consideration to the center of mass of the candidate's vertices. The weight of the intensity of each voxel depends on the Gaussian scaled squared-distance between the intensity and the maximum intensity along the line segment. The tonal scale σ_I used for weighting was set to 140 HU. This value is substantially larger than two times the image noise (previously measured to be 43.4 HU for data acquired with 50 mAs [97]). Consequently, σ_I facilitated that the edges of the candidate contributed less to the weighted average than the internal voxels of the candidate. In other words, the candidate's true internal intensity was emphasized. The center of mass falling inside the polyp is supported by the smooth apex of polyps.

Subsequently, the mean ($f_{I,\text{mean}}$), median ($f_{I,\text{median}}$), maximum ($f_{I,\text{max}}$), minimum ($f_{I,\text{min}}$), and standard deviation ($f_{I,\text{std}}$) were determined from the weighted averages of all vertices. The latter four were only used in the classifier selection stage (see Section 7.5.1).

Feature for suppressing candidates on the rectal enema tube

The rectal enema tube is a prominent source of false positive classifications [46, 113]. This is because the tube's attenuation in CT is similar to that of tissue. Moreover, the size and shape (25 mm in diameter) resembles a large polyp. Cross-sectional examples of a rectal enema tube are shown in Fig. 7.3(a). To suppress the false detections on the rectal tubes, a feature has been developed to distinguish these false detections from the other candidates. For each candidate it was measured how much 'field-of-view' (FOV) the candidate 'blocks' as seen from the rectal enema tube (Fig. 7.3(b)):

$$f_{\text{FOV}} = \frac{1}{4\pi} \sum_{\text{points} \in \text{candidate}} A_{1\text{-ring}} \frac{(\vec{q}_i \cdot \vec{n}_i)}{\|\vec{q}_i\|^3} \quad (7.1)$$

in which \vec{q}_i is the vector from a mesh point i of the candidate to an arbitrary point on the rectal tube, \vec{n}_i is the vertex normal, and $A_{1\text{-ring}}$ is the surface area of the one-ring neighborhood defined as the average area of the cells adjacent to the point of interest. A positive value means that the candidate is bended away from the tube and a negative value indicates that the candidate is bended towards the tube.

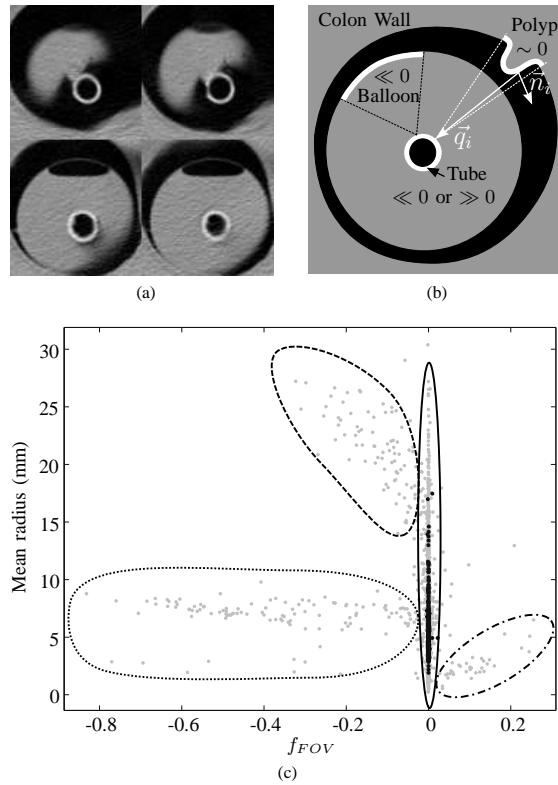


Figure 7.3: (a) Example of a rectal enema tube in data set 'A' as seen in different slices of a CT image. (b) A schematic explanation of the responses of f_{FOV} . (c) A scatter plot of the mean radius versus f_{FOV} . The grey dots are false detections and the black dots are polyps. In the text we identify the four clusters.

Fig. 7.3(c) shows a scatter plot of false detections (grey) and true polyps (black) with f_{FOV} on the horizontal axis and with the mean radius of the candidates on the vertical axis. The mean radius is calculated as a weighted sum of the distances of all mesh points i to the center of gravity of the candidate, $\|\vec{r}_i\|$, weighted by the area of the one-ring neighborhood $A_{1-ring,i}$. Apparently, four clusters are identifiable in this feature space: candidates at the end of the tube have negative values for f_{FOV} and a rather small mean radius (dotted line); candidates on the balloon also yield negative f_{FOV} , but come with a large mean radius (dashed line); candidates inside the tube have positive response for f_{FOV} (dash-dotted); and candidates that are not related to the tube have negligible blocking and form an elongated cluster centered at $f_{FOV}=0$ (solid line). To conclude, non-zero values of this feature tend to indicate detections on the rectal enema tube.

7.3 Characteristics of the feature space

A first prerequisite for clinical application is that the system has high sensitivity for the detection of polyps. To limit the risk of missing a polyp in the candidate detection step, this step unavoidably yields a large number of detections. Consequently, the number of objects from the two classes is severely unbalanced. For instance, only 0.3% of the candidates detected in data set 'A' were polyps ≥ 6 mm. Any classifier relies heavily on the few polyp examples. Complex classifiers may not be expected to generalize well to other data sets, because they are typically sensitive to small changes in training data. Furthermore, the misclassification costs for objects from the two classes are unbalanced and unknown: a missed polyp is far more troublesome than a false positive classification. Finally, it has to be realized that the size of a polyp indicates the risk of it becoming malignant.

A part of the feature space is presented in Figs. 7.4(a–b) by two scatter plots. It can be seen that the distribution of the polyps is rather uniform with respect to Φ_T , though it appears truncated at a certain level ($\Phi_T \approx 55\%$). This occurs because polyps < 6 mm are not clinically relevant and were therefore excluded a priori (i.e. not annotated in the data). The false detections display a different behavior. As our focus is on irregularities on the colon surface (protruding objects), it may be expected that far more candidates with small protrusion are detected than candidates with large protrusion, e.g. due to natural fluctuations of the colon wall and noise. This can also be seen in the distribution of the candidates with respect to the maximum protrusion in Fig. 7.5(a) and with respect to Φ_T in Fig. 7.5(b) (dotted curves). An exponential decaying function fitted to the distribution is also shown (solid curves). Thus, one must not only reckon with many false detections, the false detections are also unevenly distributed in the feature space. Finally, it can be observed that the classes largely overlap and that the way the candidates were generated imposes abrupt cluster boundaries, which may hamper density based classifiers. The abrupt cluster boundaries can be seen at $\Phi_T = 0\%$ and $\Phi_T = 100\%$

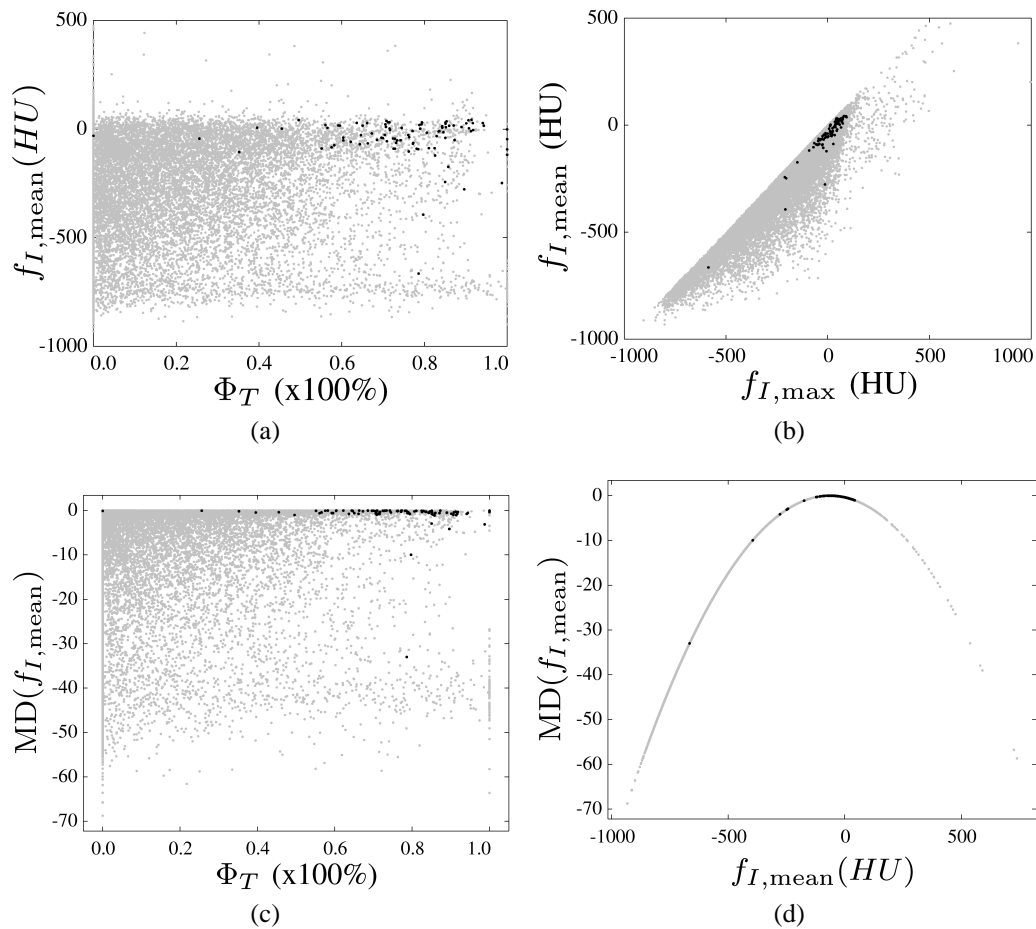


Figure 7.4: Scatter plots demonstrating the distribution of the candidates for data set 'A'. The grey dots are false detections and the black dots are polyps. (a) Mean intensity vs. Φ_T . (b) Mean intensity vs. maximum intensity. (c) The same feature space as (a) with the output of the negated Mahalanobis distance mapping on the vertical axis. This mapping is introduced in Section 7.4.1. (d) The influence of the mapping on $f_{I,\text{mean}}$. Note that candidates with a high and low mean intensity have a lower mapped feature than the polyps.

in Fig. 7.4(a).

We approach the classification problem not just as a two-class classification task, but rather as a regression problem. In other words, the classification system should be designed to facilitate a clinically relevant ordering of the candidates. Ideally, this means that the polyps should be ranked above the false detections and that the larger polyps are ranked above the smaller polyps. The classifier that is used in the regression analysis should be robust to the large class imbalance, the uneven distribution of candidates in the feature space, and the abrupt boundaries in the feature space. Moreover, the classification system as a whole must be low-complex in order to be robust to variations in the data sets from different sources.

7.4 The classification system

This section describes a classification system that fulfills the demands derived in the previous section. It is schematically depicted in Fig. 7.6. The input feature vector consists of two types of features, namely those suitable for ordering the candidates (f_O) and those allowing for density estimation and outlier rejection (f_D). The features of the first type are directly used in the regression analysis, whereas the other features are mapped first by a Mahalanobis distance mapping. Subsequently, regression analysis leads to an ordering. The ordering can then be used to compute FROC curves to estimate the performance. Three discriminant classifiers will be applied in the regression problem (see Section 7.5): the normal-based linear discriminant classifier (LDC) [135], the normal-based quadratic discriminant classifier (QDC) [135] and the logistic discriminant classifier [135].

We did not opt for support vector machine (SVM) classifiers due to the large class overlap. Due to this large overlap, it is not expected that a unique classification boundary can be found confidently. Moreover, we did not opt for neural networks too because, obviously, multi-layer neural networks based solutions may increase complexity. On the other hand, one can think of low-complex neural networks, like single layer networks with sigmoidal transfer functions (as used in [50, 65, 78]). However, these are known to be closely related to the logistic classifier.

7.4.1 Mahalanobis distance mapping

Let us assume that, for a certain subset of features, a Gaussian properly describes the distribution of the objects from the target class, i.e. the polyps. One might say that the mean of this distribution corresponds to a typical representation of a polyp (“the most polyp-like polyp”). Moreover, the Mahalanobis distance to the mean of the polyp class may act as an efficient feature to reject outliers, i.e. objects not belonging to the target class. This procedure compares to the operation of a Gaussian one-class classifier [114].

7.4. THE CLASSIFICATION SYSTEM

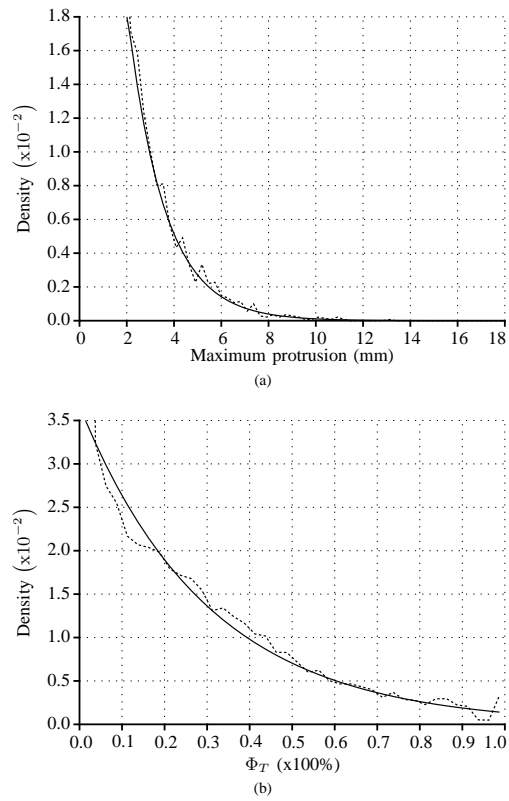


Figure 7.5: Distribution of (a) the maximum protrusion and (b) Φ_T of the false detections in data set 'A' (dotted curves). Exponential decaying functions were fitted to the distributions (solid curves).

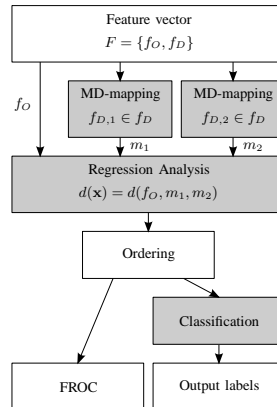


Figure 7.6: Schematic representation of the classification system. The classification starts with a feature vector consisting of features suitable for ordering (f_O) and features suitable for density estimation (f_D). The feature sets $f_{D,1}$ and $f_{D,2}$ are processed through two mappings. An ordering of the candidates is determined by regression that incorporates both the features f_O and the outputs of the mappings, m_1 and m_2 . The ordering may be thresholded for classification in order to construct FROC curves.

Instead of comparing this distance to a preset threshold, the (negated) Mahalanobis distance is used as a feature. The mean of the polyp class was derived from the train data set. Consequently, this acts as a mapping transforming one or more features into a single feature. The output feature is suitable for ordering the candidates, since zero Mahalanobis distance (the mean of the Gaussian) is considered most polyp-like. The feature can thus be used in the regression analysis. In practice, the mapping was applied to f_{FOV} and $f_{I,mean}$. Effectively, candidates on the rectal tubes as well as candidates with an abnormal intensity are rejected. Fig. 7.4 illustrates the influence of the mapping on $f_{I,mean}$.

In comparison to Wang et al. [134], our mapping replaces the normalization procedure of their two-level classifier. This allows us to use a standard technique from statistical pattern recognition to determine the parameters of the mapping.

7.4.2 Normal-based discriminant classifiers

Let us consider the linear normal-based discriminant classifier (LDC) to represent a common, low-complexity type of classifier. Such an LDC includes a weighted sum of the covariance matrices of both classes, in which the weights are the prior probabilities. In the case of a large class imbalance, however, as in the polyp detection problem, the prior of the minority class is extremely small. As a consequence, the weighted sum is almost identical to the covariance matrix of the majority class and the covariance matrix

of the minority class is neglected. In other words, contrary to common preference, the detection of objects from the minority (target) class is largely based on information of the objects from the majority (outlier) class. One might conceive this as the opposite of a one-class classifier, which typically uses information about the target class only.

One might consider a quadratic normal-based discriminant classifier (QDC) instead, since it does not weight the covariance matrices by the prior probabilities. One underlying problem here is that the classes have non-Gaussian distributions. In order to capture a polyp inside the tip of the quadratic decision boundary, simultaneously an exponentially increasing number of false positives are included (see Fig. 7.5). The more conservative linear decision boundary will make a different error to detect such a polyp, but this error is less pronounced. What is more, the quadratic classifier depends strongly on the covariance matrix of the polyp class. This covariance matrix might be somewhat unstable, however, due to the limited number of polyps.

7.4.3 Logistic discriminant classifier

It was previously demonstrated that the false detections are distributed in an exponential fashion with respect to size and Φ_T (see Fig. 7.5). Fig. 7.4 illustrated that the polyps are somewhat uniformly distributed. This implies that the ratio of the posterior probabilities must also follow an exponential function, which is represented in the next relation:

$$\log \left(\frac{p(\mathbf{x}|\omega_p)}{p(\mathbf{x}|\omega_f)} \right) = d(\mathbf{x}) \quad (7.2)$$

in which $d(\mathbf{x})$ is the linear discriminant function of the feature vector and ω_p and ω_f denote the polyp class and the false detection class, respectively. One can recognize in Eq. 7.2 the assumption made by a logistic classifier, which corresponds to sigmoidal posterior probability density functions:

$$p(\omega_f, \mathbf{x}) = \frac{1}{1 + \exp(d(\mathbf{x}))}, p(\omega_p, \mathbf{x}) = 1 - p(\omega_f, \mathbf{x}). \quad (7.3)$$

The linear logistic classifier estimates the posterior probabilities $p(\omega_i|x)$ instead of the class-dependent distributions $p(x|\omega_i)$ [135]. These posterior distributions are assumed to be the sigmoidal functions. This is a valid assumption when e.g. the classes are distributed Gaussian, or, as in this case, one of the distributions is exponentially decreasing while the other is more or less uniform. Then, a maximum likelihood (ML) estimation is made to find the linear direction in the data that best fits these assumed sigmoidal posterior functions. This ML estimator will give the weights of the discriminant function $d(\mathbf{x})$. Using the posterior probabilities instead of the class-dependent distribution functions makes this classifier less sensitive to the large class imbalance.

7.5 Results

Classifier selection aims at choosing the best method for the regression analysis in our classification system (see Fig. 7.6). Three classifiers will be analyzed: the LDC, the QDC and the logistic classifier (see Section 7.4). The specific choice will be based on two types of analysis: FROC analysis using a variety of sets of features in order to select the best classifier for the problem (instead of the best classifier for a specific feature set), and stability analysis by bootstrapping the training set.

The feature vector F in Fig. 7.6 consists of three features: Φ_T , $f_{I,\text{mean}}$ and f_{FOV} . Φ_T is related to the size of the candidates and is therefore directly used in the regression analysis, thus $f_O = \{\Phi_T\}$. The Mahalanobis distance mapping is applied to the other two features prior to the regression analysis. It is applied to $f_{D,1} = \{f_{I,\text{mean}}\}$ to sort all candidates based on the mean intensity in order of increasing distance to the normal tissue values of polyps; and to $f_{D,2} = \{f_{FOV}\}$ to aid discarding the candidates on the rectal tube. The added value of these features and the influence of the mappings will be analyzed in Section 7.5.2.

In practice, the usefulness of a CAD system depends on whether it will generalize to data sets from different sources. The robustness of the complete system will be tested in Section 7.5.3 by means of an evaluation using data sets from four different medical centers (see Section 7.2.1).

7.5.1 Classifier selection: performance and stability

The performance of the classifiers was analyzed by means of FROC analysis. The FROC curves were calculated for a large pool of different feature sets to secure that the classifier selection step is not dependent on a certain choice of features. The FROC curves were calculated from a repeated ten-fold cross-validation. Only data set 'A' was used in this learning phase to remain completely independent of the other data sets.

The aggregate of the different sets of features employed in the experiment will be called the feature pool. This pool was not created in order to select the best features, but merely to study the performance of the classifiers without choosing a specific feature set first. If some feature set were chosen first (before the classifier selection step), one might select the best classifier for the specific set of features and not necessarily the classifier which is best for the problem at hand. The feature pool consisted of 29 sets of features chosen from a total of nine different features: three protrusion-based features Φ_T with various thresholds T : 0.5, 0.6 and 0.7 mm; the features related to the intensity (i.e. the mean, maximum, minimum and median intensity and the standard deviation of the intensity) and f_{FOV} to discard candidates on the rectal tubes. Each set contained at most five features of which one was chosen from the set of protrusion-based features.

An FROC curve was computed for each classifier and for each set of features from the pool. The average FROC curve for a classifier is shown in Fig. 7.7. The standard devia-

7.5. RESULTS

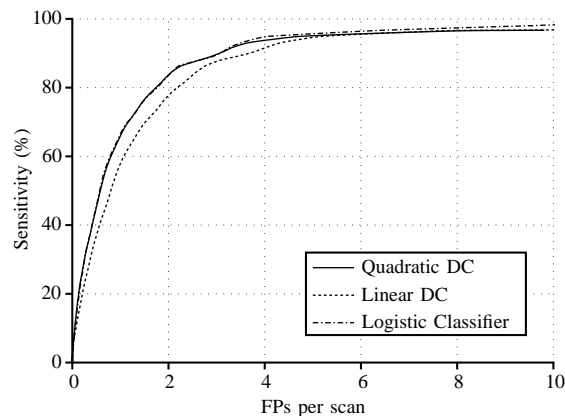


Figure 7.7: FROC curves averaged over all feature sets for the LDC, QDC and logistic classifiers.

Table 7.2: Instability of various classifiers

Classifier	Instability	Percentage (%)
Logistic	33.7	0.11
QDC	220.0	0.76
LDC	15.6	0.05

tion that was derived from the variation between the FROC curves for different feature sets was less than 0.03 FPs per scan for sensitivities below 95%. The FROC curves reveal that the logistic classifier and the QDC do not differ in their performance as their FROC curves almost completely overlap. The performance of LDC was significantly worse by approximately 15 times the standard deviation.

The second criterion used for classifier selection was the stability of the classifiers. This stability was assessed by means of bootstrapping the training set. This results in a perturbed orientation of the classifiers, which consequently leads to a number of differently classified candidates. The average number of different decisions is then used as a measure of instability [100]. Table 7.2 lists the instability measures. The table clearly shows that the logistic classifier and the LDC are the most stable classifiers. The instability has been measured for a sensitivity of 85%, but the results generalize well to other sensitivity levels, i.e. different locations of the decision boundary.

More specifically, it is noticeable that the LDC is much more stable than the QDC. This is explained by the covariance matrix estimated by the LDC being nearly identical to the covariance matrix of the majority class, which barely changes due to bootstrapping. On the other hand, the QDC also estimates a covariance matrix for the polyp class. Because of the low number of polyps, bootstrapping leads to a different covariance ma-

trix for the polyp class. This is reflected by the poor instability of the QDC. The logistic classifier is expected to be more stable since it poses an assumption onto the relative posterior probabilities of the two classes rather than estimating both (class-dependent) probability distribution functions.

To conclude, it is shown that the logistic classifier combines a good performance in terms of FROC analysis with a good stability value. Therefore, the logistic classifier will be used as the regressor in the classification system.

7.5.2 Outlier rejection by Mahalanobis distance mapping

Let us now look into the performance of outlier rejection by the Mahalanobis distance mapping. The starting point of our analysis is the FROC curve generated by the logistic classifier using Φ_T with a threshold T of 0.6 mm, and $f_{I,\text{mean}}$ (prior to mapping). FROC curves are computed for data sets 'A' and 'C'. Among other differences, these data sets differ in the type of rectal tubes used and the administration of a fecal tagging agent (see also Table 7.1).

Fig. 7.8(a) shows the FROC curves for data set 'A'. In this data set, no fecal tagging agent was administered to the patients. As a consequence, only false detections with low mean intensities were present. This means that this feature is already suitable for ordering the candidates. Mapping $f_{I,\text{mean}}$ did not result in a significantly different FROC curve; for this reason and for the purpose of clarity the curves with the 'unmapped' $f_{I,\text{mean}}$ are not shown. The solid curve is the FROC curve of a system with only the $\text{MD}(f_{I,\text{mean}})$ and Φ_T . The dotted line is obtained when the feature f_{FOV} is added directly, without prior Mahalanobis distance mapping; the dash-dotted FROC curve is the outcome when a mapped version of this feature is used instead. The improvement by adding this feature may be a reduction up to 25–50% of the number of false positives depending on the required sensitivity (see arrows). The error bars denote two times the standard deviation of the number of false positives over all scans.

The results for data set 'C' are shown in Fig. 7.8(b). In contrast to data set 'A', patients from this data set were administered a fecal tagging agent. As a consequence, it may be expected that the Mahalanobis distance mapping of $f_{I,\text{mean}}$ has a larger influence due to the presence of both candidates with a low mean intensity as candidates with a high mean intensity. Here again, the solid curve corresponds to classification using Φ_T and $f_{I,\text{mean}}$. Similar to the analysis of data set 'A', the feature f_{FOV} is added and the MD-mapping is applied to this feature and to $f_{I,\text{mean}}$. In contrast to the rectal tubes in data set 'A', the tubes in this data set did not have a balloon attached, but included a marker of high attenuation material. Because of this, less candidates on the rectal tubes were found and those which were found could often be easily discarded by means of their intensity. As a consequence, adding the feature f_{FOV} may be expected not to improve the performance. This is confirmed by the dotted line, indicating no significant improvement. Again, for the purpose of clarity, the FROC curves with the 'unmapped'

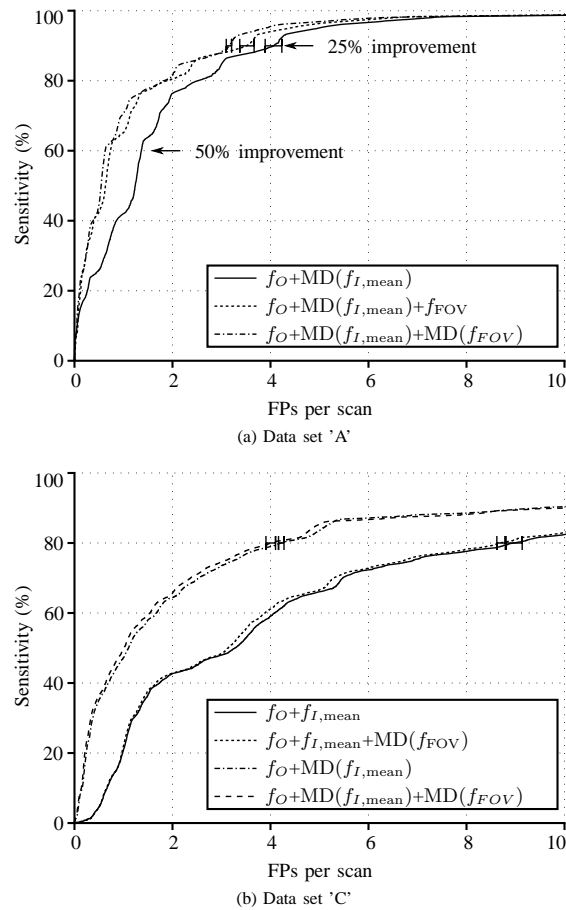


Figure 7.8: FROC curves that indicate the added value of the feature f_{FOV} and the use of the Mahalanobis distance mapping. (a) Data set 'A' with and without f_{FOV} . Using the Mahalanobis distance mapping leads to a small increase in performance. (b) Data set 'C' with and without f_{FOV} and with the unmapped and mapped mean intensity feature. The graph reveals that it is an absolute necessity to apply the mapping in the case of fecal-tagged data.

Table 7.3: Results of the candidate detection system

Data set	Polyp findings (≥ 6 mm)			Polyps (≥ 6 mm)			Number of false Detectitons
	Number of annotations	Number of detections	Detection rate (%)	Number of annotations	Number of detections	Detection rate (%)	
'A'	94	93	99	59	59	100	28 678
'B'	49	38	78	28	28	100	12 334
'C'	340	297	87	176	174	99	53 698
'D'	8	8	100	8	8	100	8026
Total	491	436	89	271	269	99	102 736

f_{FOV} are not shown in this figure, as they do not differ significantly. Observe that adding f_{FOV} does not lead to worse results.

The second step was to compute the same FROC curves with the mapped mean intensity feature. A striking improvement can be seen. This result can be explained by the fact that in this case there are both false detections with lower mean intensity as there are false detections with higher mean intensity. According to these results, only the mapped features will be used in further FROC analyzes.

7.5.3 Multi center evaluation

An important aspect of a CAD system for CT colonography is its ability to generalize to data sets differing in a variety of aspects. The generalization power of the presented system will be investigated by FROC analysis and a cross-center evaluation.

The patients from data sets 'A', 'B' and 'C' were scanned in both prone and supine positions. At the basis of this conventional approach is that a polyp is not always visible in both CT scans, e.g. due to suboptimal distension or remaining fluid rests. Consequently, a polyp may not be annotated in both scans. Let us initially focus on the annotated polyp 'findings' to assess the performance of the candidate detection step.

The candidate detection returned 88.8% (436/491) of the annotated findings ≥ 6 mm in total (see Table 3). The preparation of the patients is at the basis of the differences

7.5. RESULTS

in the number of missed findings. The patients of data set 'A' had undergone an extensive preparation. This might explain the fact that the system detected almost all annotations in this data set (93/94). On the other hand, data set 'B' appeared to contain a large amount of residual fluid (confirmed by [99]). Consequently, many polyps were obscured by fecal remains, reducing the detection rate to 77.6% (38/49). Data set 'C' had less contrast-enhanced fluid in the colon, which resulted in a higher detection rate of 87.4% (297/340). The percentage of polyps detected in either scan was 99.0% (269/271) (sensitivity is conventionally measured in this way [8]).

Fig. 7.9 shows the results of the cross-center evaluation. It is generally known that a large amount of features decreases the generalization power of a classifier, especially when the data sets differ as much as the four data sets of our study. Therefore, we consciously limited the number of features in this evaluation to the three features described before: Φ_T with a threshold 0.6 mm, $MD(f_{I,\text{mean}})$, and $MD(f_{FOV})$. Each graph in Fig. 7.9 corresponds to one test set; the line styles in the figures indicate the specific data set on which the classifier was trained. In the case of testing and training on the data from the same medical center, a ten-fold, repeated cross-validation was performed. The standard deviation indicated in the graphs is estimated as the standard deviation of a binomial distribution [19] and depends on the number of polyps and the sensitivity. This standard deviation characterizes the variation in the FROC curves when a new subset is drawn from the same distribution.

It can be seen that in all graphs, the FROC curves for classifiers trained on the different data sets are generally within one standard deviation from each other. In other words, the same performance is attained no matter on which data set the classifier is trained. Concurrently, there are small differences in the performance of the CAD system for the four data sets. Despite this, all yield a sensitivity larger than 85% at the cost of five false positive detections per scan. Four polyps in data set 'B' remained undetected at 86% (25/29) sensitivity. The missed polyps were all reviewed by a fellow researcher with a background in CAD of polyps in CTC. All missed polyps were covered by contrast-enhanced material in at least one of the two scans and were annotated in only one position. Consequently (no electronic cleansing was used), the CAD system did not get a second chance of finding these polyps. In data set 'C', fourteen polyps remained undetected by the CAD system at 90% sensitivity. The false negatives consisted of tumors with lobulated shapes, polyps covered by fecal remains, 'non-protruding' polyps annotated as a flat polyp by the radiologists and polyps that were located between haustral folds. Even though data set 'D' contained only one scan per patient, the FROC curves for this data set compete with the FROC curves for the other data sets.

In conclusion, the FROC curves for the different data sets show that the CAD system is independent on the specific data set used for training. The differences between the curves are a result of the administration of a fecal tagging agent, the preparation of the patients and natural fluctuations in the appearance of the polyps in the data sets.

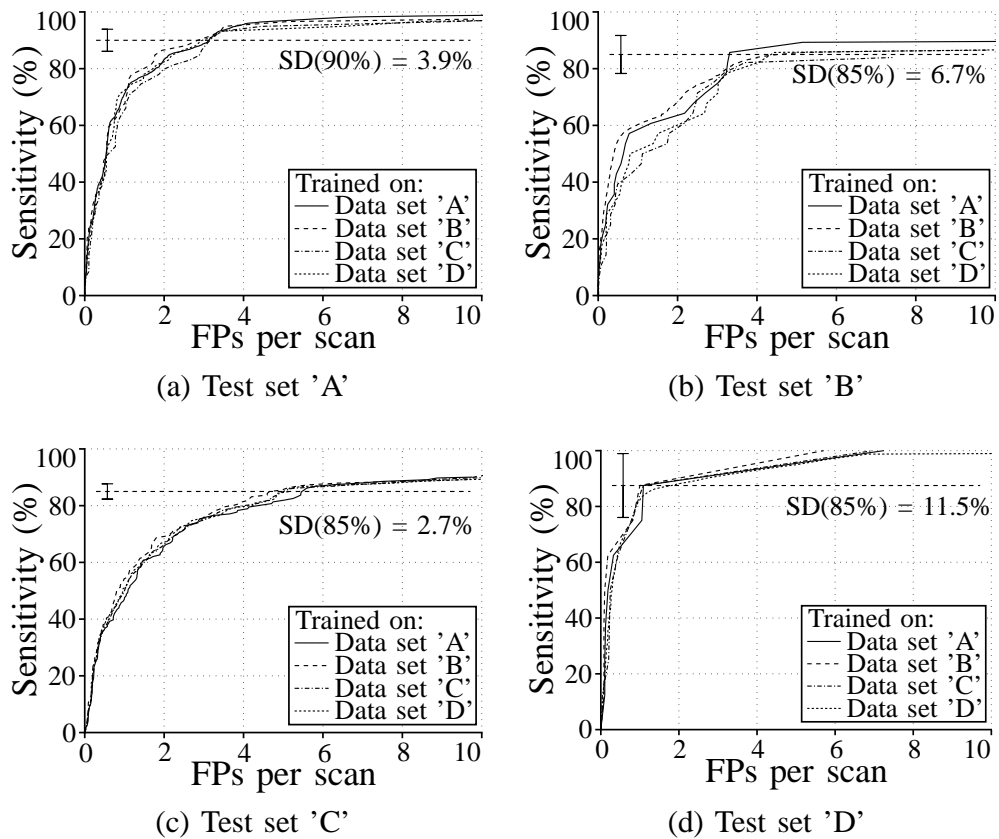


Figure 7.9: Each graph shows the results of classifying a certain data set, using four different classifiers that are each trained on one of the four data sets. The line style indicates the data set on which is trained. When the same data set is used for training and classifying, a ten-fold, repeated cross-validation was used.

7.6 Discussion and conclusion

We developed a classification system based on logistic regression for computer aided detection of polyps in CT colonography data. Typically, there are unbalanced and unknown misclassification costs and a huge class imbalance. The latter occurs because there are only a few examples of the abnormality class in a sheer endless sea of normal 'healthy' samples. Our classification system can cope with the aforementioned characteristics by carrying out a regression analysis instead of classifying the candidates into one of the two classes. The ordering correlates with the clinical relevance of the candidates. The exponential distribution of the candidates and the small number of polyps available for training led to the use of the logistic classifier for regression. The logistic classifier is low-complex and proved to be stable.

Candidates were detected based on their protrudedness from the colon wall. A feature derived from the protrusion field was sensitive for candidates that had steep edges and large protrusion. Other features used were the internal intensity distribution, and a feature to discard detections on the rectal tubes.

The features were divided into two types of features, namely features that allowed directly an ordering of the candidates and features that were well described by a Gaussian density distribution. The features of the second type were mapped by a Mahalanobis distance mapping to impose an ordering. This mapping was chosen because it emulates a Gaussian one-class classifier. In this way, outlier rejection was incorporated into the classification system.

After discarding the candidates on the rectal tubes, polyps and non-polyps could be distinguished using only information about the protrusion and the internal intensity of the candidates. The observed sensitivity was comparable to the sensitivity of radiologists using CTC [87, 105, 123] and competed with other CAD systems [49, 78, 105, 109]. It was also shown that the CAD system generalizes well to data sets from different medical centers.

To conclude, we introduced a low-complex CAD system that took into account all the characteristics of the classification problem. These characteristics will frequently occur in medical image processing problems. The Mahalanobis distance mapping in conjunction with logistic regression is generally applicable to obtain a clinically relevant ordering of the candidates. For automatic polyp detection, the generalization to data sets from different medical centers and with different patient preparations is essential to application in large-scale screening.

Acknowledgment

The authors would like to thank Dr. R. Choi, Virtual Colonoscopy Center, Walter Reed Army Medical Center, Washington, DC; Dr. P. Rogalla, Charité Hospital, Humboldt

CHAPTER 7. COMPUTER AIDED DETECTION OF POLYPS IN CT
COLONOGRAPHY USING LOGISTIC REGRESSION

University Berlin, Germany; and Dr. P.J. Pickhardt, University of Wisconsin Medical School, Madison, WI for providing the data sets.

8 Conclusions

Adenomatous polyps are small protruding mounds that may develop throughout the entire intestinal system. The ones that are located in the large bowel, or colon, are referred to as colorectal polyps. Such polyps may develop into colon cancer, one of the most commonly diagnosed types of cancer. Fortunately, there is a time window for early detection and removal of colorectal polyps, and thus prevention of cancer.

The main goal of this thesis was to develop methods for automatic detection of colorectal polyps from CT data. These methods cover three important aspects in the detection pipeline: **detection** of suspicious sites on the colon wall, **segmentation** of the site, permitting size assessment and feature extraction, and, the **classification** of these sites into polyps and non polyps, or the ranking based on a measure of polypness. Various techniques have been developed:

- Space variant filtering improves the measurement of image derivatives in highly structured environments. This aids in the detection of initial polyp candidates in strongly folded segments of the colon.
- Polyps can be accurately segmented by means of constrained propagation over the colon surface. The developed method starts from an initial seed that is obtained either by user interaction or by an automated method.
- The amount of colon wall deformation due to polyp tissue growth can be accurately measured. Two novel methods have been developed that detect and segment polyps with high sensitivity using a single threshold.
- Correct segmentation of polypoid objects is important in clinical decision making. The methods for protrusion estimation yield accurate segmentations that compare to manual size measures.
- A low-complex classification system was developed. It is based on logistic regression that effectively orders the polyps according to clinical relevance.

8.1 Improvement compared to detection based on shape index

The measurement of curvature in CT data for the detection of polyps is a difficult task for two reasons. The protrusions are not only embedded in a highly folded colon surface, but the protruding objects also have an irregular shape which gives rise to an enormous spread in curvature values depending on the amount of the regularization. Whereas the irregular shape and size would require a rather large filter that matches the size of the unknown underlying object, the highly folded nature of the colon restricts the size in order to avoid mixing with surrounding structures. Therefore, noise suppression with large isotropic filters is not possible. We have shown that with a specific formulation of normalized convolution using a local Taylor expansion, space-variant kernels can be constructed. In addition we have shown that these local kernels should be constructed by discarding voxels belonging to neighboring image geometries. Thereby the derivative filtering optimizes the trade-off between noise suppression and preservation of local image structure. The assessment of the method by simulated images shows that the space-variant filtering outperforms isotropic filtering.

In chapter 3 we presented an algorithm for the automatic segmentation of polyp-like structures on triangulated isosurfaces. It was shown that our algorithm yields a smaller bias than the measurements from radiologists: on average 1mm or less for the automatic method and between 1 and 7mm for the radiologists, depending on the irregularity of the object.

8.2 Protrusion detection

In chapter 4 and 5 two novel methods have been developed to detect protruding objects. They aim to estimate the deformation of the colon surface that is introduced by polyp growth. This is achieved by finding a steady state solution of a nonlinear PDE with the recorded image as input.

It was shown that the displacement of iso-contours relates to a change in image intensities in protruding regions in the image. This relation was made invariant to the anisotropic resolution and sampling. A segmentation of polyp candidates was obtained by applying a single threshold to the deformation field.

From the segmentations several features were extracted that directly relate to polyp size, and not to polyp shape. As a consequence ordering with respect to size is possible which, in effect, keeps increasingly larger objects further away from decision boundaries. In other words, this limits the risk of missing large polyps. Also, the method does not make a specific choice for the scale of the computation of the 1st and 2nd order derivative operators. The iterative character of the method automatically changes the

intrinsic scale of the image (local and anisotropic): the aperture of observation (window size of the operation times the number of iterations) increases until convergence.

Another advantage of the new approach is the fact that as a by-product good segmentation and size measures of the polypoid objects are obtained. In chapter 6 the segmentation method was evaluated and compared to a manual measurement. The study shows that for phantom data the measurement variability of the automatic method was smaller than the measurement variability of the observers when either the orientation of a phantom or the slice thickness was varied. This indicates that the automatic method is less sensitive to such data variations. Moreover, the automatic approach had a smaller variation than the observers in comparison to the reference standard.

8.3 Classification

A novel classification system based on logistic regression was proposed. It orders the candidates by a linear logistic classifier (logistic regression) based on only three features. This classifier can cope with a small number of polyps available for training, a large imbalance between polyps and non-polyp candidates, a truncated feature space, unbalanced and unknown misclassification costs, and an exponential distribution with respect to candidate size in feature space.

A clear distinction was made between those features that are suitable for ordering the candidates according to size, and those allowing for density estimation and outlier rejection. The features of the second type are transformed into a single feature by a Mahalanobis distance mapping. Together with features of the first type they are used in a regression analysis. The outcome leads to an ordered set from which FROC curves were extracted to estimate the classification performance.

8.4 Evaluation

We tested the robustness of the CAD system based on methods from chapters 4 to 7 by a cross-center evaluation in which the system is trained and tested with data from four different medical centers (307 patients). For polyps larger than or equal to 6 mm we achieve sensitivities of respectively 95%, 85%, 85%, and 100% with 5, 4, 5, and 6 false positives per scan over 86, 48, 141, and 32 patients.

Note that the data differs not only in patient preparation, but also in scanning protocol. Permutation of the training set among the different centers showed that the CAD system generalizes well under these varying conditions. The observed sensitivity was comparable (>85%) to the sensitivity of radiologists using CTC and competed with other CAD systems with only a limited number of false positives.

8.5 Challenges for future research

A number of challenges lie ahead in computerized screening for colorectal polyps.

Towards screening. CAD systems have not yet established themselves in the clinic. Several recent studies indicate a complementary role in CT colonography, when CAD is used as second reader [6, 10]. However, more research is needed to evaluate the use of CAD systems under varying conditions such as patient preparation and scanner types.

Colon Cleansing Advances in CT Colonography includes the use of contrast agents to tag fecal remains inside to colon. This allows for accurate localization of the colon wall with limited patient preparation. The evaluation of such data requires computerized techniques for both enhanced 3D visualization and automated detection. Digital cleansing aims to segment the colon surface in the presence of tagged intraluminal remains. Currently, the effect of digital cleansing on the performance of automated detection is unclear. Further research is needed to achieve good performance as digital cleansing is combined with CAD.

Flat Polyps Current techniques to detect colorectal cancer, such as optical colonoscopy and CT Colonography, are based on the assumption that the neoplasms are polypoid. However, recent studies have demonstrated that colorectal cancer can also arise from nonpolypoid colorectal neoplasms [102]. The latter types are more difficult to detect because there are only subtle changes to the normal mucosa. The same holds for automated detection of those so called flat polyps. The detection of flat polyps will require new techniques including new features for classification. There is no clear indication that CT imaging is sufficiently discriminating as these lesions do not lead to large deformations nor image contrast. Several papers indicate that there is a high risk of flat polyps to be cancerous at the time of diagnosis [64, 92]. This may inspire new developments in MRI colonography, which, despite its reduced imaging resolution, may provide alternative or additional contrast.

Summary

Automated Detection of Polyps for CT Colonography

In this thesis algorithms are proposed for the automated detection of colorectal polyps in 3D CT images. Polyps develop due to excessive tissue growth resulting in the colon surface to bulge out into the lumen. As a result the colon shape changes. This change may be detected by automated techniques. It usually involves three steps: colon segmentation, candidate detection and candidate classification. This thesis focuses on the second and third step: finding candidate sites followed by ranking and classification. The most common approach in computer aided detection of polyps is to focus on the characteristic protruding shape of polyps. Although we contributed to this approach as well, the main contribution of this thesis is the invention of a novel approach which measures the amount of colon surface displacement due to polyp growth.

Image derivatives play a crucial role in measuring local shape. Their measurement at or near the colon surface is particularly difficult due to its highly folded structure. This prevents the use of large filter kernels. We extend the technique of normalized convolution with a specially devised weighting term in order to optimize the trade off between noise suppression and structure induced bias.

To assess the clinical importance of detected objects, a reliable measure of polyp size is needed. We present a new algorithm for the segmentation of polyps in 3D CT images. It operates on a triangulated isosurface and takes into consideration the local orientation and position of the mesh. The algorithm starts with and expands an initial seed, located somewhere on the protruding surface. Based on the resulting segmentation the algorithm estimates the size of the object. We assess the performance by comparison to expert size measurements on phantom data and true polyps.

Through the invention of a method called “second principal curvature flow” applied to an explicit triangulation of the colon surface, we propose a new method for the detection of candidate sites. It is based on the notion that polyps have introduced a local deformation of the colon surface. The method estimates the original ‘undeformed’ surface position by solving a nonlinear partial differential equation. Candidate sites are obtained

by selecting the regions where the mesh displacement is larger than an optimized threshold. The method is assessed by a supervised classification, based on features obtained from the deformation field and the grey level image.

The same principle can also be applied directly to the 3-D volumetric data, in which case the method operates directly on the grey level voxels, rather than a triangulated isosurface. It is shown, that the use of the second principal curvature is sufficient to estimate the amount of deformation due to the growth of a polyp. A classification scheme based on linear logistic regression is proposed that explicitly keeps large polyps away from the decision boundary. Again, the method's performance is assessed by means of supervised classification.

The aforementioned detection method has another advantage over the traditional approach based on polyp shape: as a byproduct a good polyp segmentation is obtained. The accuracy of the segmentation is assessed by a comparison to expert size measurements of phantom data and true polyps. The conclusion of this work is that for phantom data, the automated method shows a reduced variability with respect to manual size measures. For patient data the automatic method shows a variability that is in the same range as manual measurement.

We conclude, by proposing a method that orders the candidates by a linear logistic classifier (logistic regression). It uses only three features: the protrusion of the colon wall, the mean internal intensity and a feature to discard detections on the rectal enema tube. This classifier can cope with a small number of polyps available for training, a large imbalance between polyps and non-polyp candidates, a truncated feature space, unbalanced and unknown misclassification costs, and an exponential distribution with respect to candidate size in feature space. Our complete CAD system (detection + classification) was evaluated with data sets from four different medical centers (307 patients). For polyps larger than or equal to 6 mm we achieve sensitivities >85%, with 6 false positives per scan. A cross-center evaluation in which the system is trained and tested with data from different sources showed that the trained CAD system generalizes to data from different medical centers and with different patient preparations.

Samenvatting

Automatische Detectie van Poliepen in CT Colonografie

In dit proefschrift worden algoritmen voorgesteld ten behoeven van de automatische detectie van colorectale poliepen in 3D CT beelden. Poliepen ontwikkelen zich door excessieve weefsel groei, met als resultaat een uitstulping van de darmwand. Als resultaat hiervan verandert de vorm van de darm. Deze verandering kan worden gedetecteerd met behulp van automatische technieken. Meestal gebeurt dit in drie stappen: darm segmentatie, kandidaat detectie en classificatie. Dit proefschrift spitst zich toe op de tweede en derde stap: het vinden van kandidaat locaties in de darm, gevolgd door een ranking en classificatie. De meest gebruikelijke methoden in gecomputeriseerde detectie richten zich op de karakteristieke vorm van poliepen. Alhoewel dit proefschrift ook bijdraagt aan deze methoden, is de grootste bijdrage de uitvinding van een nieuwe aanpak waarbij de verplaatsing van de darmwand, als gevolg van poliep groei, gemeten wordt.

Beeld afgeleiden spelen een cruciale rol in het meten van vorm. Het meten ervan op of bij de darmwand is moeilijk in het bijzonder, door zijn sterk gevouwen structuur. Dit belet het gebruik van grote filters. In dit proefschrift wordt de techniek van genormaliseerde convolutie gebruikt, met een speciaal toegepaste wegingsterm, om de afweging tussen ruis onderdrukking en door structuur geïnduceerde afwijkingen te optimaliseren.

Om de klinische belangrijkheid van gedetecteerde objecten te beoordelen is een betrouwbare maat voor poliep grootte nodig. Dit proefschrift presenteert een nieuw algoritme voor het segmenteren van poliepen in 3D CT beelden. Het werkt op een getrianguleerd iso intensiteitsoppervlak en houdt rekening met de lokale oriëntatie en positie van het mesh. Het algoritme groeit een gebied, startende vanuit een initieel punt, gepositioneerd ergens op het uitstulpende poliep achtige object. Het algoritme gebruikt de resulterende segmentatie om de grootte van het object te schatten. De prestatie van het algoritme wordt beoordeeld door expert metingen te vergelijken met automatisch gemeten grootte voor zowel fantoom objecten als ook poliepen.

Door middel van de uitvinding van de methode genaamd “second principal curvature

flow”, toegepast op een getrianguleerd iso intensiteitsoppervlak, stelt dit proefschrift een nieuwe methode voor voor de detectie van poliep kandidaat locaties. Het is gebaseerd op de observatie dat poliepen de darmwand lokaal deformeren. De methode schat de locatie van de originele “onvervormde” darmwand door het oplossen van een niet lineaire partiele differentiaalvergelijking. Kandidaat locaties worden verkregen door het selecteren van die gebieden waar de mesh verplaatsing groter is dan een geoptimaliseerde drempelwaarde. De methode is beoordeeld door middel van gesuperviseerde classificatie op basis van eigenschappen verkregen uit het deformatie veld en de grijswaarden van het beeld.

Hetzelfde principe kan worden toegepast op de 3D grijswaarde data, in welk geval de methode direct gebruikt maakt van de voxel data in plaats van een getrianguleerd iso intensiteitsoppervlak. In dit proefschrift wordt aangetoond dat het gebruik van tweede hoofdkromming voldoende is om de hoeveelheid vervorming door poliep groei te schatten. Een classificatie gebaseerd op lineaire logistische regressie is voorgesteld, waarbij grote poliepen expliciet ver weg van beslissingsgrenzen wordt gehouden. Ook deze methode is beoordeeld aan de hand van gesuperviseerde classificatie.

De bovengenoemde detectie methode heeft nog een ander voordeel boven traditionele methoden gebaseerd op poliep vorm: als bijproduct wordt een goede poliep segmentatie verkregen. De nauwkeurigheid van deze segmentatie is beoordeeld door een vergelijk met expert grootte metingen van fantomen en poliepen. De conclusie van dit werk is dat voor fantoom data de automatische methode een kleinere variatie vertoont. Voor patiënt data laat de automatische methode een variabiliteit zien die vergelijkbaar is als handmatig meten.

Het proefschrift concludeert met een voorstel van een methode waarbij de kandidaten worden gerangschikt door een lineaire logistische classificator (logistische regressie). Er worden drie eigenschappen gebruikt: een maat voor de uitpuiling van de darmwand; de gemiddelde intensiteit in het object en eigenschappen die detecties op de rectale buis kan onderscheiden. Deze classificator kan overweg met kleine aantallen poliepvoorbeelden bij training, grote klasse onbalans tussen poliep en niet poliep kandidaten, een afgekapte eigenschappenruimte, ongebalanseerde en onbekende classificatie kosten en een exponentiele verdeling met betrekking tot de kandidaat grootte. Het complete CAD systeem (detectie + classificatie) is geëvalueerd met data uit vier verschillende centra (307 patiënten). Voor poliepen grote dan 6mm wordt een gevoeligheid bereikt van >85% met 6 fout positieven per scan. Een centra-overschrijdende evaluatie waarbij het systeem is getraind en getest met data van verschillende origine laat zien dat het getrainde CAD systeem goed gerealiseerd voor data van verschillende centra en patiënt voorbereiding.

Bibliography

- [1] B. Acar, C.F. Beaulieu, S.B. Göktürk, C. Tomasi, D.S. Paik, R.B. Jeffrey, J. Yee, and S. Napel. Edge displacement field-based classification for improved detection of polyps in CT colonography. *IEEE Trans. Med. Imag.*, 21:1461–1467, Dec 2002.
- [2] B. Acer, S. Napel, D. Paik, B. Gokturk, C. Tomasi, and C.F. Beaulieu. Using Optical Flow Fields for Polyp Detection in Virtual Colonoscopy. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention Conference (MICCAI)*, oktober 2001.
- [3] Colorectal cancer facts & figures. Technical Report No. 8617.00, American Cancer Society, Atlanta, 2005.
- [4] W.F. Ames. *Nonlinear Partial Differential Equations in Engineering*, volume 1. New York: Academic Press, 1972.
- [5] R.B. Arnesen, E. von Benzon, S. Adamsen, L.B. Svendsen, H.O. Raaschou, and O.H. Hansen. Diagnostic performance of computed tomography colonography and colonoscopy: a prospective and validated analysis of 231 paired examinations. *Acta Radiologica*, 48:831–837, Oct 2007.
- [6] M.E. Baker, L. Bogoni, N. A. Obuchowski, C. Dass, R. M. Kendzierski, E. M. Remer, D. M. Einstein, P. Cathier, A. Jerebko, S. Lakare, A. Blum, D. F. Caroline, and M. Macari. Computer-aided Detection of Colorectal Polyps: Can It Improve Sensitivity of Less-Experienced Readers? Preliminary Findings. *Radiology*, 245(1):140–149, 2007.
- [7] A. Barbu, L. Bogoni, and D. Comaniciu. Hierarchical part-based detection of 3D flexible tubes: application to CT colonoscopy. In *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, volume 9, pages 462–470, 2006.
- [8] M.A. Barish, J.A. Soto, and J.T. Ferrucci. Consensus on Current Clinical Practice of Virtual Colonoscopy. *Am. J. Roentgenol.*, 184(3):786–792, 2005.
- [9] J. Bernsen. Dynamic thresholding of grey level images. In *ICPRA'86: Proc. Int. Conf. on Pattern Recognition*, pages 1251–1255, 1986.

-
- [10] L. Bogoni, P. Cathier, M. Dundar, A. Jerebko, S. Lakare, J. Liang, S. Periaswamy, ME Baker, and M. Macari. Computer-aided detection (CAD) for CT colonography: a tool to address a growing need. *British Journal of Radiology*, 78(Special Issue 1):S57, 2005.
- [11] J.H. Bond. Clinical evidence for the adenoma-carcinoma sequence, and the management of patients with colorectal adenomas. *Seminars in Gastrointestinal Disease*, 11(4):176–184, 2000.
- [12] J.H. Bond. Colon polyps and cancer. *Endoscopy*, 33:46–54, 2003.
- [13] K.A. Brakke. *The Motion of a Surface by Its Mean Curvature*. PhD thesis, Princeton University, 1978.
- [14] D. Burling, S. Halligan, M.E. Roddie, J. McQuillan, L. Honeyfield, H. Amin, J. Dehmeshki, S.A. Taylor, and E.G. McFarland. Computed Tomography Colonography: Automated Diameter and Volume Measurement of Colonic Polyps Compared With a Manual Technique-In Vitro Study. *Journal of Computer Assisted Tomography*, 29(3):387, 2005.
- [15] D. Burling, S. Halligan, S. Taylor, D.J. Brennan, D.G. Altman, P. Bassett, W. Atkin, and C.I. Bartram. Polyp Measurement Using CT Colonography: Agreement with Colonoscopy and Effect of Viewing Conditions on Interobserver and Intraobserver Agreement. *Am. J. Roentgenol.*, 186(6):1597–1604, 2006.
- [16] D. Burling, S. Halligan, SA Taylor, L. Honeyfield, and ME Roddie. CT colonography: automatic measurement of polyp diameter compared with manual assessment; an in-vivo study. *Clin. Radiol.*, 62(2):145–151, 2007.
- [17] P.M. Calvert and H. Frucht. The genetics of colorectal cancer. *Ann Intern Med*, 137:603–612, 2000.
- [18] P. Cathier, S. Periaswamy, A.K. Jerebko, M. Dundar, L. Liang, G. Fung, J. Stoeckel, T. Venkata, R. Amara, A. Krishnan, R. Bharat Rao, A. Gupta, E. Vega, S. Laks, A. Megibow, M. Macari, and L. Bogoni. CAD for polyp detection: An invaluable tool to meet the increasing need for colon-cancer screening. In *Proc. CARS'04*, volume 1268 of *International Congress Series*, pages 978–82, 2004.
- [19] C. Chatfield. *Statistics for Technology*. Chapman & Hall, third edition, 1983.
- [20] T.A. Chowdhury, P.F. Whelan, and O. Ghita. The use of 3D surface fitting for robust polyp detection and classification in CT colonography. *Comput Med Imaging Graph*, 30:427–436, Dec 2006.

- [21] T.A. Chowdhury, P.F. Whelan, and O. Ghita. A fully automatic cad-ctc system based on curvature analysis for standard and low-dose ct data. *IEEE Trans. Biomed. Eng.*, 55(3):888–901, 2008.
- [22] AH Dachman and ME Zalis. Quality and consistency in CT colonography and research reporting. *Radiology*, 230(2):319–23, 2004.
- [23] J. Dehmeshki, S. Halligan, S.A. Taylor, M.E. Roddie, J. McQuillan, L Honeyfield, and H. Amin. Computer assisted detection software for CT colonography: Effect of sphericity filter on performance characteristics for patients with and without fecal tagging. *Eur. Radiol.*, 17(3):662–668, 2007.
- [24] M. Desbrun, M. Meyer, P. Schröder, and A. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. In *SIGGRAPH 99*, 1999.
- [25] J.J. Dijkers. Segmentation, size measurement and detection of polyps in CT colonography. Master’s thesis, Delft University of Technology, 2005.
- [26] J.J. Dijkers, C. van Wijk, F.M. Vos, J. Florie, Y.C. Nio, H.W. Venema, R. Truyen, and L.J. van Vliet. Segmentation and size measurement of polyps in CT colonography. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, 8:712–719, 2005.
- [27] A. Douiri, M. Siddique, X. Ye, G. Beddoe, and G. Slabaugh. Enhanced detection in CT colonography using adaptive diffusion filtering. *Proc. SPIE Med. Imag. ’09*, 7259:725923, 2009.
- [28] Gunnar Farneäck. Spatial domain methods for orientation and velocity estimation. Lic. Thesis LiU-Tek-Lic-1999:13, Dept. EE, Linköping University, SE-581 83 Linköping, Sweden, March 1999. Thesis No. 755, ISBN 91-7219-441-3.
- [29] H.M. Fenlon, D.P. Nunes, P.C. Schroy, M.A. Barish, P.D. Clarke, and J.T. Ferrucci. A comparison of virtual and conventional colonoscopy for the detection of colorectal polyps. *N. Engl. J. Med.*, 341:1496–1503, Nov 1999.
- [30] M.B. Fennerty, J. Davidson, S.S. Emerson, R.E. Sampliner, L.J. Hixson, and H.S. Garewal. Are endoscopic measurements of colonic polyps reliable? *Am. J. Gastroenterol.*, 88:496–500, 1993.
- [31] Joseph T. Ferrucci. Colon cancer screening with virtual colonoscopy: Promise, polyps, politics. *Am. J. Roentgenol.*, 177:975–988, 2001.
- [32] J. Fidler and C. Johnson. Flat polyps of the colon: accuracy of detection by ct colonography and histologic significance. *Abdom Imaging*, 34(2):157–71, 2009.

-
- [33] J.G. Fletcher, C.D. Johnson, T.J. Welch, R.L. MacCarty, D.A. Ahlquist, J.E. Reed, W.S. Harmsen, and L.A. Wilson. Optimization of ct colonography technique: prospective trial in 180 patients. *Radiology*, 216:704–711, 2000.
- [34] T. Gluecker, G. Dorta, W. Keller, P. Jornod, R. Meuli, and P. Schnyder. Performance of multidetector computed tomography colonography compared with conventional colonoscopy. *British Medical Journal*, 51:207–211, Aug 2002.
- [35] Salih Burak Göktürk, Carlo Tomasi, Burak Acar, Christopher F. Beaulieu, David S. Paik, R. Brooke Jeffrey Jr., Judy Yee, and Sandy Napel. A statistical 3-D pattern processing method for computer-aided detection of polyps in CT colonography. *IEEE Trans. Med. Imag.*, 20(12):1251–1260, 2001.
- [36] S.B. Gokturk, C. Tomasi, B. Acar, C.F. Beaulieu, D.S. Paik, R.B. Jeffrey, J. Yee, and S. Napel. A statistical 3-d pattern processing method for computer-aided detection of polyps in ct colonography. *IEEE Trans. Med. Imag.*, 20:1251–1260, Dec 2001.
- [37] S.B. Göktürk, C. Tomasi, B. Acar, D.S. Paik, C.F. Beaulieu, and S. Napel. A learning method for automated polyp detection. In *Proc. MICCAI'01*, volume LNCS 2208, pages 85–93, 2001.
- [38] E. Gorgun and J. Church. Flat adenomas of the large bowel: A single endoscopist study. *Diseases of the Colon & Rectum*, 52(5):972, 2009.
- [39] S. Halligan, D.G. Altman, S. Mallett, S.A. Taylor, D. Burling, M. Roddie, L. Honeyfield, J. McQuillan, H. Amin, and J. Dehmeshki. Computed tomographic colonography: assessment of radiologist performance with and without computer-aided detection. *Gastroenterology*, 131:1690–1699, Dec 2006.
- [40] S. Halligan, S. A. Taylor, J. Dehmeshki, H. Amin, X. Ye, J. Tsang, and M. E. Roddie. Computer-assisted detection for ct colonography: external validation. *Clin. Radiol.*, 61(9):758–63, 2006.
- [41] A.K. Hara, C.D. Johnson, J.E. Reed, D.A. Ahlquist, H. Nelson, R.L. MacCarty, W.S. Harmsen, and D.M. Ilstrup. Detection of colorectal polyps with CT colography: initial assessment of sensitivity and specificity. *Radiology*, 205:59–65, Oct 1997.
- [42] L. Hong, S. Muraki, A. Kaufman, D. Bartz, and T. He. Virtual voyage: Interactive navigation in the human colon. In *Proc. of ACM Siggraph*, pages 27–34. Citeseer, 1997.

- [43] W. Hong, F. Qiu, and A. Kaufman. A pipeline for computer aided polyp detection. *IEEE Trans Vis Comput Graph*, 12:861–868, 2006.
- [44] Gerard Huisken. Flow by mean curvature of convex hypersurfaces into spheres. *J. Differ. Geom.*, 20:237–68, 1984.
- [45] Klinik für Strahlenheilkunde Institut für Radiologie. <http://radiologie.charite.de>: Virtuelle Koloskopie.
- [46] G Iordanescu and R.M. Summers. Reduction of false positives on the rectal tube in computer-aided detection for CT colonography. *Med. Phys.*, 31(10):2855–2862, 2004.
- [47] Näppi J. and H. Yoshida. Automated detection of polyps with ct colonography: Evaluation of volumetric features for reduction of false-positive findings. *Acad. Radiol.*, 9:386–397, 2002.
- [48] S. Jensch, R.E. van Gelder, J. Florie, M.A. Thomassen-de Graaf, J.V. Lobé, P.M. Bossuyt, S. Bipat, C.Y. Nio, and J. Stoker. Performance of radiographers in the evaluation of CT colonographic images. *AJR Am J Roentgenol*, 188:W249–255, Mar 2007.
- [49] A. Jerebko, S. Lakare, P. Cathier, S Periaswamy, and L. Bogoni. Symmetric curvature patterns for colonic polyp detection. In *Proc. MICCAI’06*, volume LNCS 4191, pages 169–176, 2006.
- [50] A. K. Jerebko, J. D. Malley, M. Franaszek, and R. M. Summers. Multiple neural network classification scheme for detection of colonic polyps in CT colonography data sets. *Acad. Radiol.*, 10:154–160, 2003.
- [51] A. K. Jerebko, J. D. Malley, M. Franaszek, and R. M. Summers. Support vector machines committee classification method for computer-aided polyp detection in CT colonography. *Acad. Radiol.*, 12(4):479–86, 2005.
- [52] A.K. Jerebko, R.M. Summers, J.D. Malley, M. Franaszek, and C.D. Johnson. Computer-assisted detection of colonic polyps with CT colonography using neural networks and binary classification trees. *Med. Phys.*, 30:52–60, Jan 2003.
- [53] C.D. Johnson, R.L. MacCarty, T.J. Welch, L.A. Wilson, W.S. Harmsen, D.M. Ilstrup, and D.A. Ahlquist. Comparison of the relative sensitivity of CT colonography and double-contrast barium enema for screen detection of colorectal polyps. *Clin. Gastroenterol. Hepatol.*, 2:314–321, Apr 2004.
- [54] Coebergh J.W. Colorectal cancer screening in europe: first things first. *Eur. J. Cancer*, 40:638–642, 2004.

-
- [55] D.H. Kim, P.J. Pickhardt, A.J. Taylor, W.K. Leung, T.C. Winter, J.L. Hinshaw, D.V. Gopal, M. Reichelderfer, R.H. Hsu, and P.R. Pfau. CT colonography versus colonoscopy for the detection of advanced neoplasia. *The New England journal of medicine*, 357(14):1403, 2007.
- [56] S.H. Kim, J.M. Lee, J.G. Lee, J.H. Kim, P.A. Lefere, J.K. Han, and B.I. Choi. Computer-aided detection of colonic polyps at CT colonography using a Hessian matrix-based algorithm: preliminary study. *AJR Am J Roentgenol*, 189:41–51, Jul 2007.
- [57] G. Kiss, S. Drisis, D. Bielen, F. Maes, J. van Cleynenbreugel, G. Marchal, and P. Suetens. Computer-aided detection of colonic polyps using low-dose CT acquisitions. *Acad. Radiol.*, 13(9):1062–1071, 2006.
- [58] G. Kiss, J. Van Cleynenbreugel, S. Drisis, D. Bielen, G. Marchal, and P. Suetens. Computer aided detection for low-dose CT colonography. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, 8:859–867, 2005.
- [59] G Kiss, J. van Cleynenbreugel, S. Drisis, D. Bielen, G Marchal, and P. Suetens. Computer-aided detection for low-dose CT colonography. In *Proc. MICCAI'05*, volume LNCS 3749, pages 859–867, 2005.
- [60] G. Kiss, J. van Cleynenbreugel, G. Marchal, and P. Seutens. Computer aided detection in ct colonography via spin images. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention Conference (MICCAI)*, pages II–805, 2004.
- [61] H. Knutsson and C.-F. Westin. Normalized convolution - a technique for filtering incomplete and uncertain data. In K. A. Hø gda, B. Braathen, and K. Heia, editors, *SCIA'93, Proceedings of the 8th Scandinavian Conference on Image Analysis*, volume 2, pages 997–1006, TromsøNorway, 1993. Norwegian Society for Image Processing and Pattern Recognition.
- [62] E. Konukoglu, B. Acar, D.S. Paik, C.F. Beaulieu, and S. Napel. HDF: Heat diffusion fields for polyp detection in CT colonography. *Signal Processing*, 87(10):2407–2416, 2007.
- [63] E. Konukoglu, B. Acar, D.S. Paik, C.F. Beaulieu, J. Rosenberg, and S. Napel. Polyp enhancing level set evolution of colon wall: method and pilot study. *IEEE Trans. Med. Imag.*, 26:1649–1656, Dec 2007.

- [64] S. Kudo, H. Kashida, T. Tamura, E. Kogure, Y. Imai, H. Yamano, and A.R. Hart. Colonoscopic diagnosis and management of nonpolypoid early colorectal cancer. *World Journal of Surgery*, 24(9):1081–1090, 2000.
- [65] M. A. Kupinski, D. C. Edwards, M. L. Giger, and C. E. Metz. Ideal observer approximation using bayesian classification neural networks. *IEEE Trans. Med. Imag.*, 20(9):886–899, 2001.
- [66] J. Liu, J. Yao, and R.M. Summers. Scale-based scatter correction for computer-aided polyp detection in CT colonography. *Med. Phys.*, 35(12):5664–71, 2008.
- [67] G. Lordanescu and R.M. Summers. Reduction of false positives on the rectal tube in computer-aided detection for CT colonography. *Med. Phys.*, 31:2855–2862, Oct 2004.
- [68] L. Lu, A. Barbu, M. Wolf, J. Liang, L. Bogoni, M. Salganicoff, and D. Comaniciu. Simultaneous detection and registration for ileo-cecal valve detection in 3d ct colonography. In *ECCV '08: Proceedings of the 10th European Conference on Computer Vision*, pages 465–478, Berlin, Heidelberg, 2008. Springer-Verlag.
- [69] Y. Masutani, H. Yoshida, P. MacEneaney, and A.H. Dachman. Automated segmentation of colonic walls for computerized detection of polyps in CT colonography. *JCAT*, 25:629–638, 2001.
- [70] J. Melonakos, P. Mendoga, R. Bhotka, and S. Sirohey. A probabilistic model for haustral curvatures with applications to colon CAD. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, 10:420–427, 2007.
- [71] R Midgley and D Kerr. Colorectal cancer. *Lancet*, 353:391–399, 1999.
- [72] A.A. Miranda, T.A. Chowdhury, O. Ghita, and P.F. Whelan. Shape filtering for false positive reduction at computed tomography colonography. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, 9:84–92, 2006.
- [73] Basil C. Morson. Evolution of cancer of the colon and rectum. *Cancer*, 34:345–349, 1974.
- [74] T. Muto, H. J. Bussey, and B. C. Morson. The evolution of cancer of the colon and rectum. *Cancer*, 36(6):2251–2270, 1975.
- [75] J. Näppi, H. Frimmel, and H. Yoshida. Virtual endoscopic visualization of the colon by shape-scale signatures. *IEEE Trans Inf Technol Biomed*, 9:120–131, Mar 2005.

- [76] J. Näppi and H. Yoshida. Automated detection of polyps with ct colonography: Evaluation of volumetric features for reduction of false-positive findings. *Acad. Radiol.*, 9:386–397, 2002.
- [77] J. Näppi and H. Yoshida. Feature-guided analysis for reduction of false positives in CAD of polyps for computed tomographic colonography. *Med. Phys.*, 30(7):1592–1601, 2003.
- [78] J. Näppi and H. Yoshida. Fully automated three-dimensional detection of polyps in fecal-tagging CT colonography. *Acad. Radiol.*, 14:287–300, 2007.
- [79] J.J. Näppi, H. Frimmel, A.H. Dachman, and H. Yoshida. Computerized detection of colorectal masses in CT colonography based on fuzzy merging and wall-thickening analysis. *Med. Phys.*, 31:860–872, Apr 2004.
- [80] W.J. Niessen, B.M. ter Haar Romeny, L.M.J. Florack, and M.A. Viergever. A general framework for geometry-driven evolution equations. *Int. J. Comput. Vision*, 21(3):187–205, 1997.
- [81] P.J. Olver, G. Sapiro, and A. Tannenbaum. Invariant geometric evolutions of surfaces and volumetric smoothing. *SIAM Journal on Applied Mathematics*, 57(1):176–94, 1997.
- [82] D.S. Paik, C.F. Beaulieu, G.D. Rubin, B. Acar, R.B. 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. Imag.*, 23:661–675, Jun 2004.
- [83] N. Petrick, M. Haider, R.M. Summers, S.C. Yeshwant, L. Brown, E.M. Iuliano, A. Louie, J.R. Choi, and P.J. Pickhardt. CT colonography with computer-aided detection as a second reader: observer performance study. *Radiology*, 246:148–156, Jan 2008.
- [84] N. Petrick, M. Haider, R.M. Summers, S.C. Yeshwant, L. Brown, E.M. Iuliano, A. Louie, J.R. Choi, and P.J. Pickhardt. Ct colonography with computer-aided detection as a second reader: observer performance study. *Radiology*, 247(1):148–56, 2008.
- [85] P.J. Pickhardt. CT colonography (virtual colonoscopy) for primary colorectal screening: Challenges facing clinical implementation. *Abdominal Imaging*, 30:1–4, 2005.
- [86] P.J. Pickhardt. Virtual colonoscopy: issues related to primary screening. *Eur. Radiol.*, 15 Suppl 4:D133–137, Nov 2005.

- [87] P.J. Pickhardt, J.R. Choi, I. Hwang, J.A. Butler, M.L. Puckett, H.A. Hildebrandt, R.K. Wong, P.A. Nugent, P.A. Mysliwiec, and W.R. Schindler. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *N Engl J Med*, 349:2191–2200, 2003.
- [88] P.J. Pickhardt, C. Hasssan, A. Laghi, A Zullo, D.H. Kim, and S. Morini. Cost-effectiveness of colorectal cancer screening with computed tomography colonography. *Cancer*, 109:2213–2221, 2007.
- [89] P.J. Pickhardt and D.H. Kim. Colorectal cancer screening with CT colonography: key concepts regarding polyp prevalence, size, histology, morphology, and natural history. *American Journal of Roentgenology*, 193(1):40, 2009.
- [90] P.J. Pickhardt, A.D. Lee, E.G. McFarland, and A.J. Taylor. Linear polyp measurement at CT colonography: In vitro and in vivo comparison of two-dimensional and three-dimensional displays. *Radiology*, 236:872–878, 2005.
- [91] W.H. Press, S.A. Teukolski, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992.
- [92] BJ Rembacken, T. Fujii, A. Cairns, MF Dixon, S. Yoshida, DM Chalmers, and ATR Axon. Flat and depressed colonic neoplasms: a prospective study of 1000 colonoscopies in the uk. *The Lancet*, 355(9211):1211–1214, 2000.
- [93] B.M. Sallam, A. Pilch-Kowalczyk, K. Gruszczyska, J. Baron, and F. Pugliese. Diagnostic performance of CT colonography in a population with high prevalence of large bowel disease. *Med. Sci. Monit.*, 13 Suppl 1:105–110, May 2007.
- [94] H. Scharr. *Optimal Operators in Digital Image Processing*. PhD thesis, Ruprecht-Karls-Universität Heidelberg, 2000.
- [95] R.E. Schoen, L.D. Gerber, and C. Margulies. The pathologic measurement of polyp size is preferable to the endoscopic estimate. *Gastrointestinal Endoscopy*, 46(6):492–496, 1997.
- [96] I. Serlie, F.M. Vos, R.E. van Gelder, J. Stoker, R. Truyen, F.A. Gerritsen, Y. Nio, and F.H. Post. Improved visualization in virtual colonoscopy using image-based rendering. In *Proc. Joint Eurographics-IEEE TCVG Symposium on Visualization*, pages 137–146, 2001.
- [97] I. W. O. Serlie, F. M. Vos, H. W. Venema, and L. J. van Vliet. CT imaging characteristics. Technical Report QI-2006-01, 2006. <http://www.ist.tudelft.nl/>.

-
- [98] IWO Serlie, R. Truyen, J. Florie, F. Post, L. van Vliet, and F. Vos. Computed cleansing for virtual colonoscopy using a three-material transition model. In *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, pages 175–183. Springer-Verlag; 1999, 2003.
- [99] Iwo W. O. Serlie, Ayso H. de Vries, Frans M. Vos, Yung Nio, Roel Truyen, Jaap Stoker, and Lucas J. van Vliet. Lesion conspicuity and efficiency of CT colonography with electronic cleansing based on a three-material transition model. *AJR*, 2008.
- [100] Marina Skurichina. *Stabilizing Weak Classifiers*. PhD thesis, Delft University of Technology, Delft, The Netherlands, October 2001.
- [101] G.W. Snedecor and W.G. Cochran. *Statistical Methods*. Iowa State University Press, 1989.
- [102] Roy M. Soetikno, Tonya Kaltenbach, Robert V. Rouse, Walter Park, Anamika Maheshwari, Tohru Sato, Suzanne Matsui, and Shai Friedland. Prevalence of nonpolypoid (flat and depressed) colorectal neoplasms in asymptomatic and symptomatic adults. *JAMA*, 299(9):1027–1035, 2008.
- [103] R. M. Summers, C. D. Johnson, L. M. Pusanik, J. D. Malley, A. M. Youssef, and J. E. Reed. Automated polyp detection at CT colonography: Feasibility assessment in a human population. *Radiology*, 219(1):51–9, 2001.
- [104] R. M. Summers, W. S. Selbie, J. D. Malley, L. M. Pusanik, A. J. Dwyer, N. A. Courcoutsakis, D. J. Shaw, D. E. Kleiner, M. C. Sneller, C. A. Langford, S. M. Holland, and J. H. Shelhamer. Polypoid lesions of airways: Early experience with computer-assisted detection by using virtual bronchoscopy and surface curvature. *Radiology*, 208(2):331–7, 1998.
- [105] R. M. Summers, J. Yao, P. J. Pickhardt, M. Franaszek, I. Bitter, D. Brickman, V. Krishna, and J. R. Choi. Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. *Gastroenterology*, 129(6):1832–44, 2005.
- [106] R.M. Summers. Challenges for computer-aided diagnosis for ct colonography. *Abdominal Imaging*, 27:268–274, 2002.
- [107] R.M. Summers, C.F. Beaulieu, L.M. Pusanik, J.D. Malley, R.B. Jeffrey, D.I. Glazer, and S. Napel. Automated polyp detector for CT colonography: feasibility study. *Radiology*, 216:284–290, Jul 2000.

Bibliography

- [108] R.M. Summers, L.R. Handwerker, P.J. Pickhardt, R.L. Van Uitert, K.K. Deshpande, S. Yeshwant, J. Yao, and M. Franaszek. Performance of a previously validated ct colonography computer-aided detection system in a new patient population. *Am. J. Roentgenol.*, 191(1):168–74, 2008.
- [109] R.M. Summers, L.R. Handwerker, P.J. Pickhardt, R.L. van Uitert, K.K. Deshpande, S. Yeshwant, J. Yao, and M. Franaszek. Performance of a previously validated ct colonography computer-aided detection system in a new patient population. *AJR*, 191:169–174, 2008.
- [110] R.M. Summers, J. Yao, and C.D. Johnson. CT colonography with computer-aided detection: automated recognition of ileocecal valve to reduce number of false-positive detections. *Radiology*, 233:266–272, Oct 2004.
- [111] P. Sundaram, A. Zomorodian, C. Beaulieu, and S. Napel. Colon polyp detection using smoothed shape operators: preliminary results. *Med. Imag. Anal.*, 12(2):99–119, 2008.
- [112] 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):694–703, 2008.
- [113] K. Suzuki, H. Yoshida, J. Näppi, and A.H. Dachman. Massive-training artificial neural network (MTANN) for reduction of false positives in computer-aided detection of polyps: Suppression of rectal tubes. *Med. Phys.*, 33(10):3814–3824, 2006.
- [114] David M. J. Tax. *One-class Classification; Concept-learning in the absence of counter-examples*. PhD thesis, Delft University of Technology, Delft, The Netherlands, June 2001.
- [115] R. Taylor, S.A. amd Greenhalgh, R. Ilangovan, E. Tam, V.A. Sahni, D. Burling, J. Zhang, P. Bassett, P.J. Pickhardt, and S. Halligan. Ct colonography and computer-aided detection: effect of false-positive results on reader specificity and reading efficiency in a low-prevalence screening population. *Radiology*, 247(1):133–40, 2008.
- [116] S.A. Taylor, S.C. Charman, P. Lefere, E.G. McFarland, E.K. Paulson, J. Yee, R. Aslam, J.M. Barlow, A. Gupta, D.H. Kim, C.M. Miller, and S. Halligan. CT colonography: investigation of the optimum reader paradigm by using computer-aided detection software. *Radiology*, 246:463–471, Feb 2008.

-
- [117] S.A. Taylor, A. Laghi, P. Lefere, S. Halligan, and J. Stoker. European Society of Gastrointestinal and Abdominal Radiology (ESGAR): consensus statement on CT colonography. *Eur. Radiol.*, 17(2):575–579, 2007.
- [118] S.A. Taylor, A. Slater, S. Halligan, L. Honeyfield, M.E. Roddie, J. Demeshski, H. Amin, and D. Burling. CT colonography: automated measurement of colonic polyps compared with manual techniques—human in vitro study. *Radiology*, 242(1):120, 2006.
- [119] S.A. Taylor, N. Suzuki, G. Beddoe, and S. Halligan. Flat neoplasia of the colon: Ct colonography with cad. *Abdom Imaging*, 34(2):173, 2009.
- [120] J.P. Thirion and A. Gourdon. Computing the differential characteristics of iso-intensity surfaces. *Computer Vision and Image Understanding*, 61(2):190–202, 1995.
- [121] R.F. Thoeni and I Laufer. *Polyps and cancer, Textbook of Gastrointestinal Radiology*. Philadelphia: W.B. Saunders, 1994.
- [122] E.R. van Gelder. *CT Colonography for Screening of Patients at Increased Risk for Colorectal Cancer*. PhD thesis, 2004.
- [123] R.E. Van Gelder, C.Y. Nio, J. Florie, J.F Bartelsman, P. Snel, S.W. De Jager, S.J. Van Deventer, J.S. Lameris, P.M. Bossuyt, and J. Stoker. Computed tomographic colonography compared with colonoscopy in patients at increased risk for colorectal cancer. *Gastroenterology*, 127(1):41–8, 2004.
- [124] J. van Kan, A. Segal, and F. Vermolen. *Numerical Methods in Scientific Computing*. VSSD, Delft, The Netherlands, ISBN 9071301508, first edition, 2005.
- [125] V. F. van Ravesteijn, C. van Wijk, F. M. Vos, R. Truyen, J. F. Peters, J. Stoker, and L. J. van Vliet. Computer aided detection of polyps in CT colonography using logistic regression. *IEEE Trans. Med. Imag.*, 29:120–131, 2010.
- [126] L. J. van Vliet and P. W. Verbeek. Curvature and bending energy in digitized 2D and 3D images. *Proc. 8th Scandinavian Conf. on Image Analysis.*, pages 1403–10, 1993.
- [127] C. Van Wijk, J. Florie, C. Y. Nio, E. Dekker, A. H. De Vries, H. W. Venema, L. J. van Vliet, J. Stoker, and F. M. Vos. Protrusion method for automated estimation of polyp size on CT colonography. *Am. J. Roentgenol.*, 190(5):1279–85, 2008.
- [128] C. Van Wijk, R. Truyen, R.E. van Gelder, L.J. van Vliet, and F.M. Vos. On normalized convolution to measure curvature features for automatic polyp detection.

- Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv, MICCAI 2004, pages pt. I, 200–208, 2004.
- [129] 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. *accepted for publication in IEEE Trans. Med. Imag.*
- [130] C. Van Wijk, V.F. Van Ravesteijn, F.M. Vos, R. Truyen, A.H. de Vries, J. Stoker, and L.J. van Vliet. Detection of protrusions in curved folded surfaces applied to automated polyp detection in CT colonography. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, 9:471–478, 2006.
- [131] D.J. Vining, D.W. Gelfand, R.E. Bechtold, E.S. Scharling, E.K. Grishaw, and R.Y. Shifrin. Technical feasibility of colon imaging with helical ct and virtual reality. *AJR*, 162:104, 1994.
- [132] F.M. Vos, R.E. van Gelder, I.W. Serlie, J. Florie, C.Y. Nio, A.S. Glas, F.H. Post, R. Truyen, F.A. Gerritsen, and J. Stoker. Three-dimensional display modes for ct colonography: Conventional 3d virtual colonoscopy versus unfolded cube projection. *Radiology*, 228:878–858, 2003.
- [133] S. Wang, J. Yao, and R.M. Summers. Improved classifier for computer-aided polyp detection in ct colonography by nonlinear dimensionality reduction. *Med. Phys.*, 35(4):1377–86, 2008.
- [134] 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 colonography. *Med. Phys.*, 32(12):3602–3616, 2005.
- [135] Andrew R. Webb. *Statistical Pattern Recognition*. John Wiley & Sons, second edition, 2002.
- [136] S. Winawer, R. Fletcher, D. Rex, J. Bond, R. Burt, J. Ferrucci, T. Ganiats, T Levin, S. Woolf, D. Johnson, L. Kirk, S. Litin, and C. Simmang. Colorectal cancer screening and surveillance: Clinical guidelines and rationale – update based on new evidence. *Gastroenterology*, 124:544–560, 2003.
- [137] J. Yao, M. Miller, M. Franaszek, and R. M. Summers. Colonic polyp segmentation in CT colonography-based on fuzzy clustering and deformable models. *IEEE Trans. Med. Imag.*, 23(11):1344–52, 2004.
- [138] J.H. Yao, S. Campbell, A. K. Hara, and R. M. Summers. Progressive feature vector selection scheme for computer-aided colonic polyp detection. *RSNA Scientific Assembly and Annual Meeting Program*, 633:633, 2004.

-
- [139] S.C. Yeshwant, R.M. Summers, J. Yao, D.S. Brickman, R.J. Choi, and P.J. Pickhardt. Polyps: linear and volumetric measurement at ct colonography. *Radiology*, 241:802–811, 2006.
- [140] H. Yoshida, Y. Masutani, P. MacEneaney, D.T. Rubin, and A.H. Dachman. Computerized detection of colonic polyps at CT colonography on the basis of volumetric features: pilot study. *Radiology*, 222:327–336, Feb 2002.
- [141] H. Yoshida and J. Näppi. Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. *IEEE Trans. Med. Imag.*, 20(12):1267–1274, 2001.
- [142] H. Yoshida and J. Näppi. CAD in CT colonography without and with oral contrast agents: Progress and challenges. *Computerized Medical Imaging and Graphics*, 31:267–284, 2007.
- [143] H. Yoshida, J. Näppi, P. MacEneaney, D.T. Rubin, and A.H. Dachman. Computer-aided diagnosis scheme for detection of polyps at CT colonography. *Radiographics*, 22(4):963–979, 2002.
- [144] B.M. Young, J.G. Fletcher, S.R. Paulsen, F. Booya, C.D. Johnson, K.R. Johnson, Z. Melton, D. Rodysill, and J. Mandrekar. Polyp measurement with ct colonography: multiple-reader, multiple-workstation comparison. *AJR*, 188(1):122–129, 2001.
- [145] J.Y. Yun, H.J. Ro, J.B. Park, J.B. Choi, J.E. Chung, Y.J. Kim, W.H. Suh, and J.K. Lee. Diagnostic performance of CT colonography for the detection of colorectal polyps. *Korean J Radiol*, 8:484–491, 2007.
- [146] M.E. Zalis, M.A. Barish, J.R. Choi, A.H. Dachman, H.M. Fenlon, J.T. Ferrucci, S.N. Glick, A. Laghi, M. Macari, E.G. McFarland, M.M. Morrin, P.J. Pickhardt, J. Soto, and J. Yee. CT colonography reporting and data system: A consensus proposal. *Radiology*, 236:3–9, 2005.
- [147] L. Zhao, C.P. Botha, J.O. Bescos, R. Truyen, F.M. Vos, and F.H. Post. Lines of curvature for polyp detection in virtual colonoscopy. *IEEE Trans Vis Comput Graph*, 12:885–892, 2006.
- [148] Y. Zheng, X. Yang, and G. Beddoe. Reduction of false positives in polyp detection using weighted support vector machines. *Proc. 29th IEEE EMBS*, pages 4433–4436, 2007.

List Of Publications

1. F. Scarano, C. van Wijk, and L.L.M. Veldhuis: Traversing field of view and AR-PIV for mid-field vortex wake investigation in a towing tank. *Experiments in Fluids* 33 (2002) 950-961.
2. F.M. Vos, C. van Wijk, V.F. van Ravesteijn, I.W.O. Serlie, S.E. Grigorescu, F.H. Post, R. Truyen, J. Stoker, L.J. van Vliet, Recent advances in techniques for CT colonography: electronic cleansing and CAD, *International journal of computer assisted radiology and surgery*, 3(Suppl 1), S188-S198, 2008
3. C. van Wijk, J. Florie, C.Y. Nio, E. Dekker, A.H. de Vries, H.W. Venema, L.J. van Vliet, J. Stoker, F.M. Vos, Protrusion method for automated estimation of Polyp Size on CT Colonography, *Am. J. Roentgenol.*, vol. 190, no. 5, 2008, 1279 - 1285
4. V.F. van Ravesteijn, C. van Wijk, F.M. Vos, R. Truyen, J.F. Peters, J. Stoker, L.J. van Vliet, Computer Aided Detection of Polyps in CT Colonography using Logistic Regression, *IEEE Transaction on Medical Imaging*, accepted, 2009
5. C. van Wijk, V.F. van Ravesteijn, F.M. Vos, L.J. van Vliet, Detection and Segmentation of Colonic Polyps on Implicit Isosurfaces by Second Principal Curvature Flow, *IEEE Transaction on Medical Imaging*, accepted, 2009
6. C. van Wijk. A Vortex wake Investigation of an A340 Model Using PIV in a Towing Tank. MSc Thesis, Delft University of Technology, 2001
7. L.L.M. Veldhuis, F. Scarano, and C. van Wijk: Vortex wake investigation of an Airbus A340 model using PIV in a Towing Tank, C-WAKE Report PR 1.1.3-TUD, 29-10-2001
8. F. Scarano, C. van Wijk, L.L.M. Veldhuis; Traversing field-of-view and AR-PIV for mid field wake vortex investigation in a towing tank. In: J. Kompenhans (eds.); . 4th International Symposium on Particle Image Velocimetry(Gottingen, Germany, 9/17/2001), ISBN: CD-Rom.

9. L.L.M. Veldhuis, F. Scarano, C. van Wijk; PIV-survey of the Vortex wake structure behind an Airbus A340 in a Towing tank. In: conference on capacity and vortex wakes. (Imperial college London, 12/9/2001).
10. C. van Wijk, R. Truyen, R.E. van Gelder, L.J. van Vliet, and F.M. Vos, On Normalized Convolution to Measure Curvature Features for Automatic Polyp Detection, in: Christian Barillot, David R. Haynor, Pierre Hellier (eds.), Medical Image Computing and Computer-Assisted Intervention MICCAI 2004 (Proc. 7th Int. Conf. Saint-Malo, France, Sep.26-29) Part 1, Lecture Notes in Computer Science, vol. 3216, Springer Verlag, Heidelberg, 2004, 200-208.
11. J.J. Dijkers, C. van Wijk, F.M. Vos, J. Florie, Y.C. Nio, H.W. Venema, R. Truyen, and L.J. van Vliet, Segmentation and size measurement of polyps in CT colonography, Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005 (Proc. 8th Int. Conf., Palm Springs, CA, USA, Oct. 26-29), Lecture Notes in Computer Science, vol. 3749, Springer Verlag, Heidelberg, 2005, 712-719.
12. C. van Wijk, V.F. van Ravesteijn, F.M. Vos, R. Truyen, A.H. de Vries, J. Stoker, and L.J. van Vliet, Detection of Protrusions in Curved Folded Surfaces Applied to Automated Polyp Detection in CT Colonography, in: Rasmus Larsen, Mads Nielsen, Jon Sporring (eds.), Medical Image Computing and Computer-Assisted Intervention MICCAI 2006 (Proc. 9th Int. Conf., Copenhagen, Denmark, Oct.1-6) Part II , Lecture Notes in Computer Science, vol. 4191, Springer Verlag, Heidelberg, 2006, 471-478.
13. C. van Wijk, J. Florie, C.Y. Nio, E. Dekker, A.H. de Vries, H.W. Venema, L.J. van Vliet, J. Stoker, F.M. Vos, Protrusion method for automated estimation of Polyp Size on CT Colonography, Am. J. Roentgenol., vol. 190, no. 5, 2008, 1279 - 1285
14. V.F. van Ravesteijn, C. van Wijk, F.M. Vos, R. Truyen, J.F. Peters, J. Stoker, L.J. van Vliet, Computer Aided Detection of Polyps in CT Colonography using Logistic Regression, IEEE Transaction on Medical Imaging, accepted, 2009
15. C. van Wijk, V.F. van Ravesteijn, F.M. Vos, L.J. van Vliet, Detection and Segmentation of Colonic Polyps on Implicit Isosurfaces by Second Principal Curvature Flow, IEEE Transaction on Medical Imaging, accepted, 2009
16. K. van Wijk, Method for protrusion detection in 3D medical data, 16-5-2007, NL1032602, patent

Curriculum Vitae

Kees van Wijk was born in Zwammerdam, The Netherlands on Februari 9th 1974. In 1993 he obtained his grammar school diploma at the scholengemeenschap Albanianea. The same year he started his education in Aerospace Engineering at the Delft University of Technology. In 2001 he obtained the title of 'ingenieur' (MSc) with his thesis on 'A Vortex wake Investigation of an A340 Model Using PIV in a Towing Tank', supervised by Leo veldhuis en Fulvio Scarano.

In 2003 he started a PhD project at the Quantitative Imaging group (f.k.a Pattern Recognition group) at Delft University of Technology on the topic of developing automatic detection techniques for CT colonography. He was supervised by Prof. dr. ir. Lucas J. van Vliet and dr. Frans M. Vos. This work was financially supported by the IOP.

In november 2007 he started his current position at Medis medical imaging systems. His work involves the development of new image processing techniques in the field of neuro and cardiovascular imaging.

Dankwoord

Aan het begin van het laatste deel van mijn proefschrift, wil ik Lucas en Frans bedanken voor hun begeleiding en vooral hun eindeloze geduld de laatste jaren, bij het voltooien van mijn proefschrift. Lucas; dank voor je inhoudelijke scherpste, de discussie; je had altijd snel door waar het om ging. Je gaf mij stevast de juiste richting om weer verder te gaan. Dank ook voor alle kinderspullen die we kregen bij de geboorte van Lennart. Hoeveel kilometers zou hij op jullie driewielers hebben afgelegd? Frans; dank voor je geduld bij het schrijven en corrigeren van de publicaties. Je rol als intermediaator/brug tussen de medische en technische wereld was erg fijn. Je was mijn weg in het AMC. Je enthousiasme voor het onderwerp heeft mij telkens weer geprikkeld. Zo ook je enthousiasme voor het schaatsen. Het waren gezellige avonden!

Michael, Bernd, Tuan, jullie hebben als kamergenoten mij binnen QI (PH+1) geïntroduceerd en vlug op weg geholpen binnen de vakgroep en natuurlijk met diplib. Gea en Martijn dank voor jullie gezelligheid. Simona, dank voor het Italiaanse chocolade brood, Heerlijk!

Een promotie doe je niet alleen, Vincent, je kwam als student op mijn onderwerp afstuderen. Het is leuk te zien dat jij ook de stap naar promoveren hebt gemaakt. Je hebt mij enorm geholpen bij dit proefschrift.

Frank ik wil je bedanken voor het kritisch meedenken en het luisteren in de tijd dat we samen op de vakgroep werkzaam waren. Maar vooral ook de tijd erna, in moeilijke momenten van "proefschrift-dip", heb je mij weer moed ingesproken, zie hier! het heeft geholpen.

Hans Reiber, graag wil ik je bedanken voor de tijd en ruimte die ik kreeg bij het afronden van mijn proefschrift.

Leo en Fulvio; jullie waren mij al bekend voor ik aan dit avontuur begon. Jullie inspireerde en enthousiasmeerde mij in het verdiepen in en uitpluizen van technische vraagstukken. De stap van PIV naar CTC is niet zo groot!

Als laatste wil ik eerst mijn twee wereldwonderen bedanken, Lennart en Jacco. Waar zou ik zijn, zonder mijn boefjes? Waarschijnlijk was ik dan allang klaar geweest met dit proefschrift. Maar jullie zijn het meer dan waard!

En nu ... op naar nieuwe avonturen: in de cardiovasculaire beeldverwerking; in Bodegraven, met de jongens; en overal en altijd met jou: mijn Eliaan!