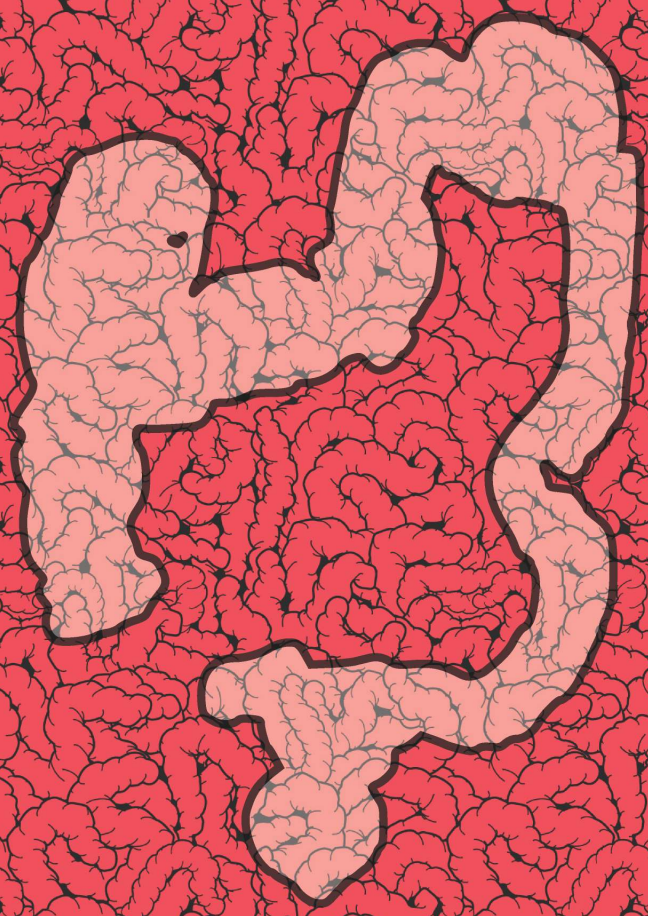
Computer Aided Detection of Polyps in CT Colonography



Vincent van Ravesteijn

Computer Aided Detection of Polyps in CT Colonography

Front cover: The background image is courtesy of Trenton Marshall (tm4rt.deviantart.com).

Computer Aided Detection of Polyps in CT Colonography

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op vrijdag 15 november 2013 om
10.00 uur
door

Vincent Frans van RAVESTEIJN

ingenieur in de technische natuurkunde
geboren te Gorinchem.

Dit proefschrift is goedgekeurd door de promotor:

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

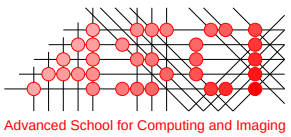
copromotor:

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. J. Stoker	Academisch Medisch Centrum Amsterdam
Prof. dr. ir. F.W. Jansen	Technische Universiteit Delft
Prof. dr. ir. N. Karssemeijer	Radboud Universiteit Nijmegen
Prof. dr. ir. M. Breeuwer	Technische Universiteit Eindhoven Philips Healthcare
Prof. dr. M. Loog	Technische Universiteit Delft University of Copenhagen
Prof. I.T. Young	Technische Universiteit Delft, reservelid

This work is supported by Philips Healthcare, Best, The Netherlands.



This work was carried out in the ASCI graduate school
ASCI dissertation series number 292.

ISBN: 978-94-6191-950-2

Contents

Chapter 1

Introduction

Colorectal cancer is the second leading cause of mortality due to cancer in the Western world [1]. 85% of all diagnosed colorectal cancers arise from adenomatous polyps [1]. Adenomatous polyps are uncontrolled growths in the colon wall consisting of cells with abnormal DNA. At first they are benign, but they can turn into carcinoma over the course of years. Fortunately, screening for colorectal polyps can significantly decrease the incidence of colorectal cancer by facilitating early detection of such polyps [2]. Moreover, any cancers residing in the colon or rectum may also be detected in an early stage, so that they can be treated before metastasizing [3]. Approximately half of the number of polyps encountered in the colon are non-adenomatous polyps and the vast majority of the non-adenomatous polyps are hyperplastic. This means they have normal DNA and it is assumed they have almost no potential to turn malignant [4].

Morphologically, polyps can be divided into flat, sessile and pedunculated polyps which are illustrated in Figure 1.1. When a polyp evolves into a malignant tumor, it grows into the glandular cells of the colon wall and it is called an adenocarcinoma.

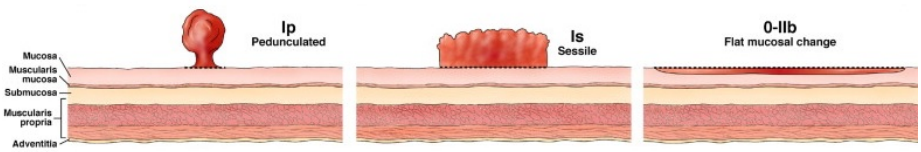


Figure 1.1: A pedunculated polyp (left), a sessile polyp (middle) and a flat polyp without elevation (right). (Images are obtained from Ref. [5].)

1.1 CT Colonography

Various methods are available to screen for colorectal polyps [2]. The most important methods are the fecal occult blood test (FOBT), sigmoidoscopy, optical colonoscopy (endoscopy), double contrast barium enema (DCBE) examination and computed tomography colonography (CTC). However, none of these methods is ideal. Optical colonoscopy and sigmoidoscopy, for example, are very invasive, FOBT and DCBE have a rather low sensitivity, and CTC has the disadvantage of exposing the patient to ionizing radiation. Still, CTC receives a lot of interest as a candidate for screening low-risk patients, because CTC is minimally invasive and promises to be highly sensitive [6, 7]. Furthermore, CTC offers the possibility of virtual colonography [8], automated polyp detection (CAD) [9, 10], and electronic cleansing [11] to ease the task of the radiologist. In fact, CTC is preferred by patients over colonoscopy [3, 12], which is important when screening of low-risk patients is considered. Figure 1.2 shows a polyp as it can be seen on various imaging modalities. An example of virtual colonography and electronic cleansing is shown in Figure 1.3.

In comparison with optical colonoscopy, CTC has the drawback that it only serves as a diagnostic tool. This means that patients have to undergo optical colonoscopy if CTC leads to a positive diagnosis. Common practice nowadays is to refer a patient to colonoscopy when at least one polyp larger than 10 mm is found [6]. Patients who have polyps with a size between 6 and 10 mm are monitored closely. The subsequent treatment of these patients depends largely on the number of polyps found. Polyps smaller than 6 mm are often neglected because they are believed not to evolve in malignant growths and they might even disappear [14, 4].

1.1.1 Automated Polyp Detection

Screening for colorectal polyps by a radiologist without CAD support is time consuming. However, in comparison with a CAD system, a radiologist is more accurate in deciding whether a candidate object is a polyp or not. This is because a CAD system only uses information from the direct vicinity of the candidate and disregards any further 'contextual' information. Conversely, this information is available and important to a radiologist. The impact of using local and global information on the accuracy of screening was investigated by Sluimer et al. [15]. They studied the application of a CAD system for the detection of abnormal tissues in lungs. It was shown that the sensitivity of 1 of the 2 observers decreased by 10% when only local information was presented to the observers. Still, it is known that radiologists do overlook polyps. Therefore, the main motivations for the development of a CAD system is to reduce the time it takes for a radiologist to examine a patient, to improve sensitivity and to limit perceptual errors [9].

From a clinical point of view, the first requirement of a CAD system is to

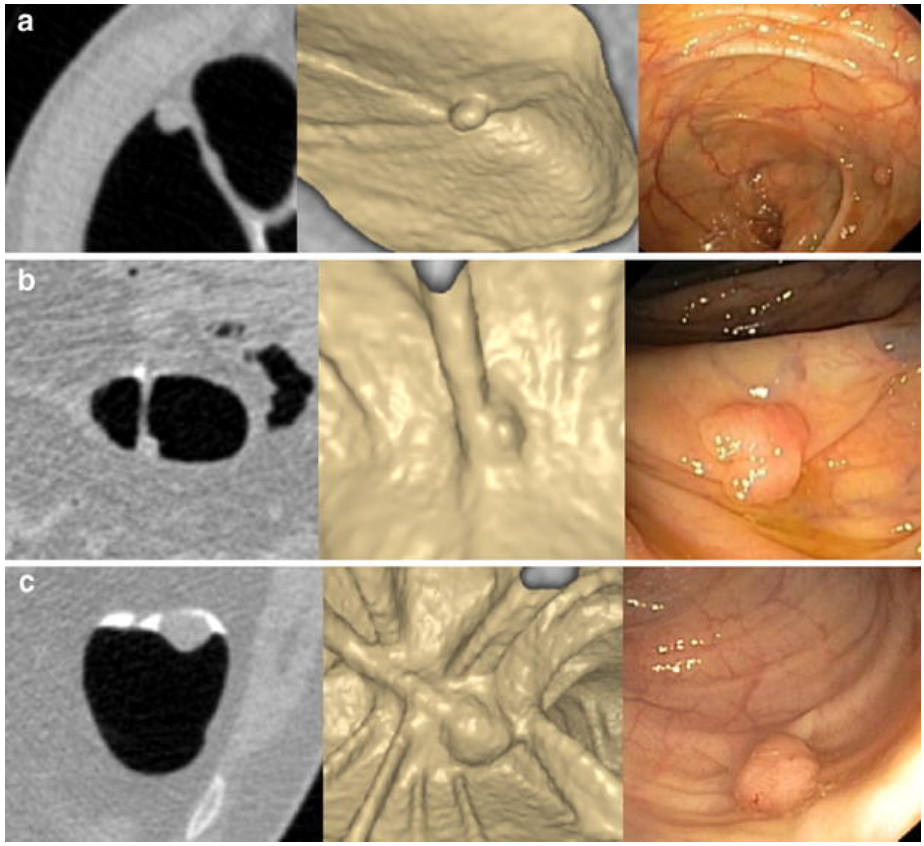


Figure 1.2: Two sessile polyps (a, b) and a pedunculated polyp (c) as can be seen in CT colonography (left), virtual colonography (middle) and optical colonoscopy (right). (Images are obtained from Ref. [13].)

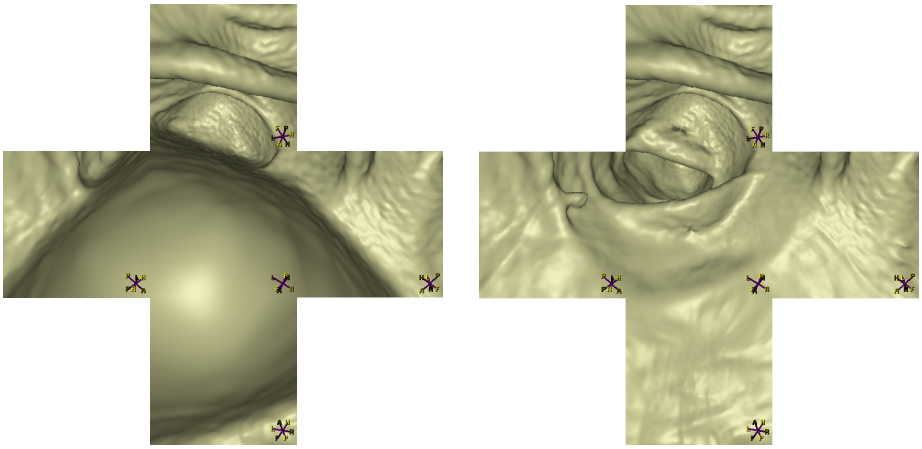


Figure 1.3: Virtual colonography without electronic cleansing (left) and with electronic cleansing (right).

detect polyps with a high sensitivity. Typically, percentages around 80–90% are considered acceptable, since this is about the sensitivity of radiologists as well as colonoscopists [6]. The second demand on a CAD-system is that it should produce as few false positive detections as possible, to limit the number of cases that a radiologist has to review. The acceptance of the number of false positive detections by radiologists depends on several aspects such as the interface used to review the presented candidates and the difficulty of the decisions whether a candidate is a polyp or not [9].

Even though the CAD system may reduce the examination time of the radiologist, it currently operates as a second reader in clinical practice. In this way the CAD system aims only at increasing the sensitivity of the radiologist. The CAD system will not become a first reader until it is proven to be reliable and trusted by radiologists. Besides this, all medical devices and software applications are subject to a thorough validation process in order to get approval for the use of these devices and application in hospitals. In Europe, this is supervised by the EMA, the European Medicines Agency, and in the United States by the FDA, the United States Food and Drug Administration. Therefore, the third demand of a CAD system for use in a medical device is that the system is as transparent and as simple as possible.

1.1.2 Electronic Cleansing

As was the case for colonoscopy, examination by CTC still required the patients to undergo cathartic bowel preparation before the examination. Although such a cathartic bowel preparation ensures optimal image quality, it also leads to

excessive diarrhea and discomfort. It was found that this preparation was one of the most burdensome aspects of CT colonography with a cathartic bowel preparation [12]. Tagging of the bowel content with oral iodine or barium contrast facilitates CTC with non-cathartic bowel preparation.

Evaluating data from CTC in a 3D reading mode with non-cathartic bowel preparation requires that the data is electronically cleansed before evaluation by the radiologist. An electronic cleansing algorithm aims at replacing the tagged materials inside the colon by air, such that 3D visualization of the whole bowel becomes possible [16]. Recently, several studies have shown that the diagnostic accuracy for polyps $\geq 6\,\text{mm}$ remains high while using a 24-hour limited bowel preparation (i.e., the least burdensome type of non-cathartic preparations) [17, 18]. In fact, a limited bowel preparation significantly improves the acceptance and therefore likely the screening adherence [19, 17, 20]. Moreover, Liedenbaum et al. showed that a 24-hour limited iodine-based bowel preparation yields a significantly better subject's acceptance and less burden compared with a 48-hour preparation [21].

Unfortunately, such preparations can adversely affect the 3D image quality. Particularly, untagged stool can cause artefacts like incomplete cleansing or pseudo-soft tissue structures [22, 23]. These artefacts limit a primary 3D reading and hinder 3D problem solving after a primary 2D reading. Still, accurate electronic cleansing can result in shorter reading times in a primary 3D reading strategy and to a higher confidence and less reader effort in a primary 2D reading strategy [24].

1.1.3 Dose Reduction

Apart from the burden associated with the bowel preparation, the acceptance of CT colonography as a screening technique is also influenced by the radiation exposure. Clearly, the radiation burden should be as low as possible to ensure a high benefit-risk ratio. However, a lower-dose scanning protocol will inevitably lead to increased image noise which in turn compromises polyp detection. For ethical reasons it is a general, complex problem to study the performance of both human and computerized observers at increasingly lower CT doses: one cannot simply scan the patient using various doses, nor is it possible to scan a patient at a low-dose for which the diagnostic value is not guaranteed.

1.2 Objectives

The work presented in this thesis aims at improving CT colonography to make large-scale screening feasible. Specifically, computerized techniques from the fields of image processing and pattern recognition will be explored to support this. Section 1.1 described the topics that will be addressed: computer aided polyp detection, electronic cleansing for limited bowel preparations, and dose

reduction. Ultimately, our focus will be on low-dose, 24-hour limited bowel preparation CTC.

To increase the sensitivity and to reduce the working time of the radiologists, automated polyp detection systems are proposed. In Section 1.1.1, three demands for such a system were posed to enable its use in clinical practice. This leads to the next question:

- Can we design a computer aided polyp detection system that has a sensitivity that is at least comparable to the sensitivity of human observers, and has a low complexity, such that it generalizes well, i.e. it has similar performance for comparable data from different medical centers ?

As the subject's preparation is one the most burdensome aspects of CTC, recent clinical research aimed at reducing the subject's preparation. At the moment, the state-of-the-art bowel preparation is 24-hour limited iodine-based [21]. Although this increases the subject's acceptance of CTC, it comes with severely degradated image quality. Especially 3D evaluation of the data is currently impossible for the radiologist. To still be able to employ CTC and virtual colonoscopy, the electronic cleansing algorithms need to be able to cope with this kind of data. Thus:

- Can we design an electronic cleansing algorithm that is able to process data from CTC with a 24-hour limited bowel preparation in such a way that allows effective 3D evaluation of the colon without compromising the observers' sensitivity ?

Lowering the radiation dose reduces the risk of cancer induction, which is particularly relevant in a screening setting. However, the effect of a low radiation dose on the polyp sensitivity of screening, c.q. the CAD system can not be easily assessed. To facilitate exploration of the relation between radiation dose and diagnostic accuracy, we will investigate:

- Can we use computer simulated low-dose CTC to assess the performance of CTC with lower radiation doses ?

1.3 Thesis Outline

The first question of designing a robust, low-complex CAD system will be addressed in Chapters 2–5. Here, an automated polyp detection system is presented which is based on a minimal principal curvature flow algorithm. It is shown that the algorithm combines the detection and segmentation of polyp candidates, thereby making it very robust. The algorithm presented in Chapter 2 still requires a segmentation of the colon surface as a first step, whereas Chapter 3 shows that the algorithm can work directly on the grey value image and thereby also relaxing the need for such an explicit segmentation step. The

robustness of the system is signified by the fact that both types of algorithm perform well in combination with a low-complex pattern recognition step. This is further illustrated in Chapter 4 and 5. The latter considers the algorithm in the context of dissimilarity classification, where the algorithm is presented as a deformation defining the polyp class and shows that the algorithm comprises not only a detection and segmentation step, but a classification step as well.

The second part of this thesis focuses on how the patient's burden can be reduced. Different patient preparation schemes have been proposed and each one poses a different challenge for electronic cleansing. Chapter 6 focuses on how the cleansing algorithm can be adapted to describe thin layers that often occur when the preparation involves a barium solution as a fecal tagging agent. Alternatively, an iodine based type of preparation leads to heterogeneities in the appearance of tagged materials. By using the knowledge obtained in the first part of the thesis, Chapter 7 introduces a principal curvature flow algorithm to resolve such heterogeneities while retaining the colon anatomy.

Finally, Chapter 8 investigates whether the radiation burden can be further reduced in the future. As it is ethically not acceptable to perform experiments with varying doses on subjects, the effect of low-dose CTC is simulated. In the appendix, the performance of a CAD system in low-dose CTC is assessed for a number of radiation dose levels.

Chapter 2

Computer Aided Polyp Detection 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 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. 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.

2.1 Introduction

Colorectal cancer is the second leading cause of mortality due to cancer in the western world [1]. 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 [25, 26, 27]. It has been shown that screening for such polyps can significantly reduce the incidence of colon cancer [2, 28]. 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 [9, 29, 30] and radiologists using CTC [31].

The best indicator of the risk that a polyp is malignant or turns malignant over time is size [32]. The consensus [33] 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 ($\leq$ 5 mm) should not even be reported [34, 35]. There is still debate over the need for polypectomy for 6–9 mm polyps. Surveillance for growth with CT colonography has also been suggested.

2.1.1 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. [36, 37] 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 [36]. Yoshida et al. [38, 39] 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. [40] 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 parameter space of the Hough transformation. Kiss et al. characterized the candidate's shape by comparing the spherical harmonics with those of the polypoid models in a database [41].

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 [38, 39, 42] as well as Jerebko et al. [43]. Wang et al. [44] 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. [45] classified the candidates based on a single heuristically designed score using curvature information of the candidate patches. Göktürk et al. [46] 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. [47] and Zheng et al. [48] used a committee of support vector machines. Neural networks were also used by Jerebko et al. [47] and Näppi et al. [29, 49] for classification, and by Suzuki et al. [50] 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., [46, 38, 51, 40, 9, 44, 43, 41, 10, 52, 53, 42]) reported on the performance of polyp detection algorithms (see Yoshida and Näppi [31] 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 2.2.1).

2.1.2 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 [54]. It will be shown that this two-level classification system is effective over data from various sources.

2.2 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.

2.2.1 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 examination. 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 2.1. 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) 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 data set 'A' has

Table 2.1: Properties of the data sets

Data set	Medical Center	Slice Thick-ness (mm)	Fecal Tagging	Scans per Patient	Number of Patients	Number of Polyps $\geq$ 6 mm	Ref.
'A'	AMC / Amsterdam	3.2	No	2	86	59	[6]
'B'	WRAMC / Wash., DC	1.2	Yes	2	48	28	[3]
'C'	UW / Madison, WI	1.2	Yes	2	141	176	[55][1]
'D'	Charité / Berlin	1.0	Yes	1	32	8	[56][1]

been used during development of the system.

2.2.2 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. 2.1 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 [10, 57] was solved to remove all protruding structures from the mesh that displayed a positive second principal curvature. 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.

[1]Information about the patient preparation can be retrieved from the reference. However, the specific data set we used is not described.

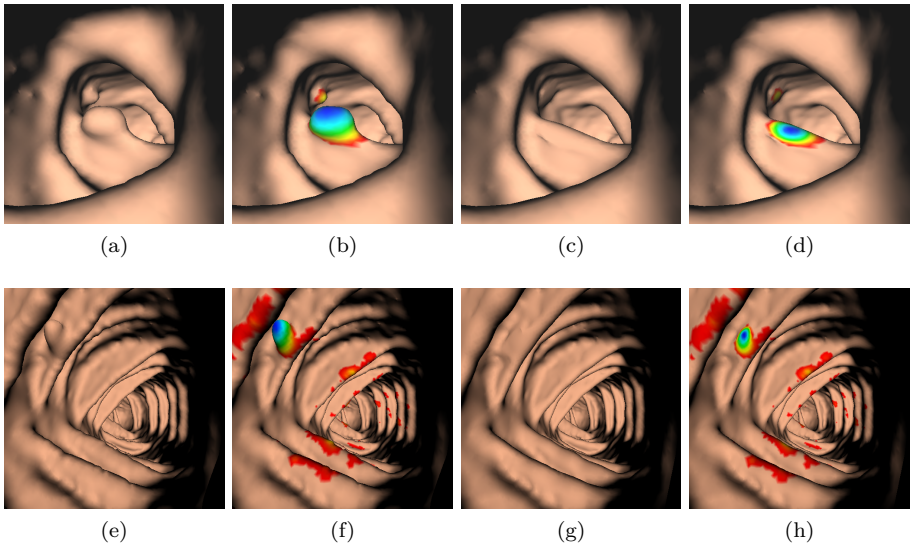


Figure 2.1: The candidate detection method applies a non-linear 'flattening' operation to the colon wall. The protrusion field is defined as the difference in position of the colon wall before (a–b, e–f) and after (c–d, g–h) application of the operation. The coloring (b,d,f,h) indicates the protrusion 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.

2.2.3 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 analyse CTC data. Radiologists using the 3D rendering of the colon (virtual colonography) 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 [39], mean curvature, average principal curvatures and sphericity ratio [36, 37] and spherical harmonics [41]. An alternative method to measure shape, which is based on the protrusion field, will be introduced (see Section 2.2.3, below).

The internal intensity of the candidates has been found before to be a discriminative feature to discard a large number of false detections [42, 51, 44, 43]. 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 object's internal voxel intensities (see Section 2.2.3, 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 [58, 50]). Therefore, a third feature will be proposed to discard such false detections (see Section 2.2.3, below).

Shape Feature from Protrusion Field

Polyps are conventionally characterized by the single largest diameter, excluding the stalk [32, 59]. However, Fig. 2.2(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. 2.2(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 unrelevant 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 with protrusion below T and result in a low response. Thus, this feature favors compact objects with steep edges. Fig. 2.2(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 2.2.2). 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_{t} 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 [60]). Consequently, σ_{t} 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.

Subsequently, the mean ($f_{I,\mathrm{mean}}$), median ($f_{I,\mathrm{median}}$), maximum ($f_{I,\mathrm{max}}$), minimum ($f_{I,\mathrm{min}}$), and standard deviation ($f_{I,\mathrm{std}}$) were determined from the weighted averages of all vertices. The latter four were only used in the classifier selection stage (see Section 2.5.1).

Feature for Suppressing Candidates on the Rectal Enema Tube

The rectal enema tube is a prominent source of false positive classifications [58, 50]. 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. 2.3(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. 2.3(b)):

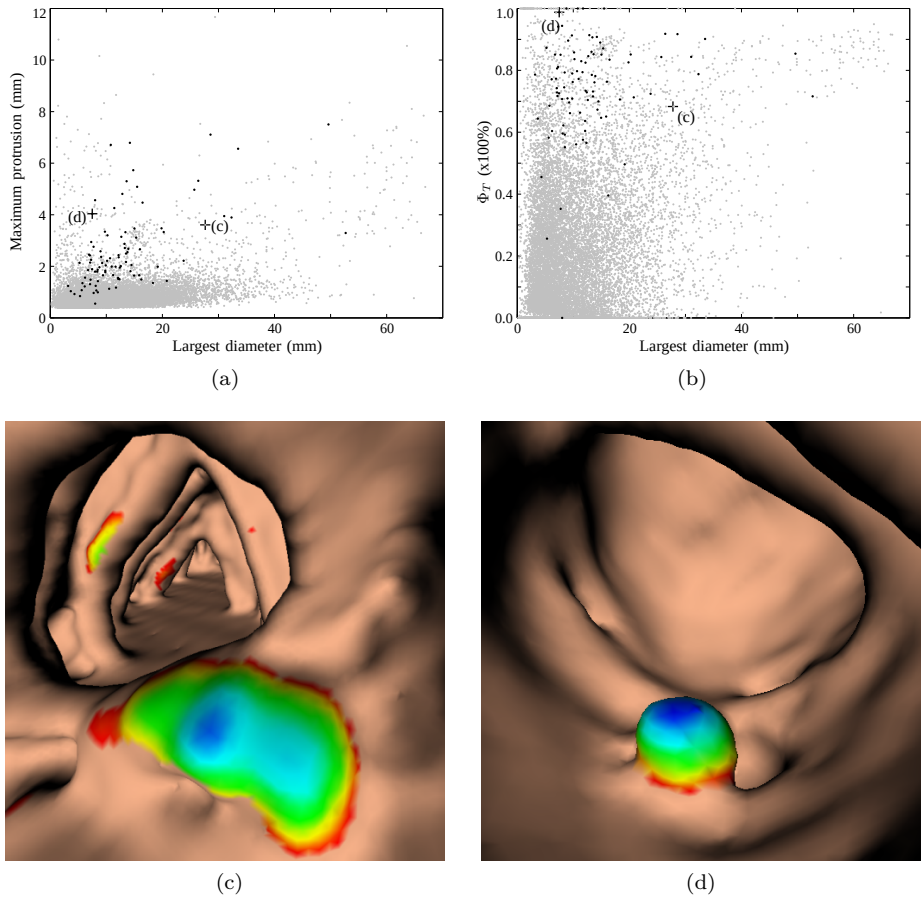


Figure 2.2: (a)–(b) Scatter plots of features calculated for data set 'A'. Grey dots denote false detections and black dots indicate polyps $\geq$ 6 mm. Note that each polyp may appear as two separate dots in the scatter plot, since each patient is scanned twice. (a) The maximum protrusion versus the single largest diameter of a candidate. The threshold of the candidate detection can be seen at a maximum protrusion of 0.4 mm. (b) $\Phi_T(T{=}0.6\,\text{mm})$ versus the largest diameter. (c–d) Two candidates with the same maximum protrusion that are ordered differently according to Φ_T.

$$f_{\mathrm{FOV}} = \frac{1}{4\pi} \sum_{\text{points}\in\text{candidate}} A_{1-\mathrm{ring}} \frac{(\vec{q_i} \cdot \vec{n_i})}{||\vec{q_i}||^3} \tag{2.1}$$

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 $A_{1-\mathrm{ring}}$ is the surface area of the one-ring neighborhood. 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.

Fig. 2.3(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 $A_{1-\mathrm{ring,i}}$. 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.

2.3 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 $\geq$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. 2.4(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

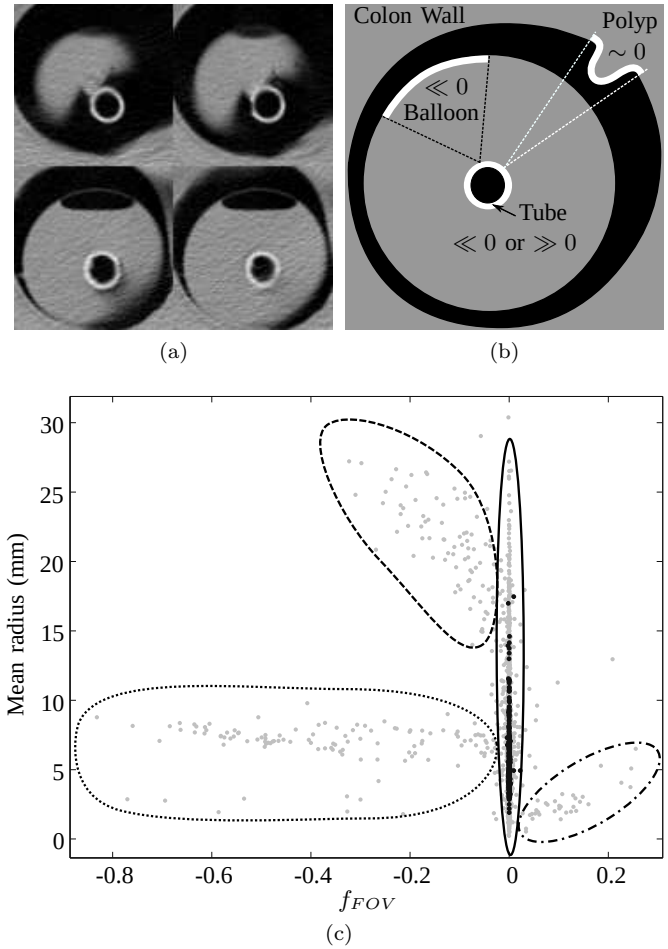


Figure 2.3: (a) Example of a rectal enema tube in data set 'A' as seen in different slices of a CT image. (b) A schematical explanation of the responses of f_{FOV}. (c) A scatter plot of the mean radius versus f_{FOV}. The grey dots are false detections and the black dots are polyps. In the text we identify the four clusters.

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. 2.5(a) and with respect to Φ_T in Fig. 2.5(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 $\Phi_T = 100\%$ in Fig. 2.4(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.

2.4 The Classification System

This section describes a classification system that fulfills the demands derived in the previous section. It is schematically depicted in Fig. 2.6. 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 2.5): the normal-based linear discriminant classifier (LDC) [61], the normal-based quadratic discriminant classifier (QDC) [61] and the logistic discriminant classifier [61].

2.4.1 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

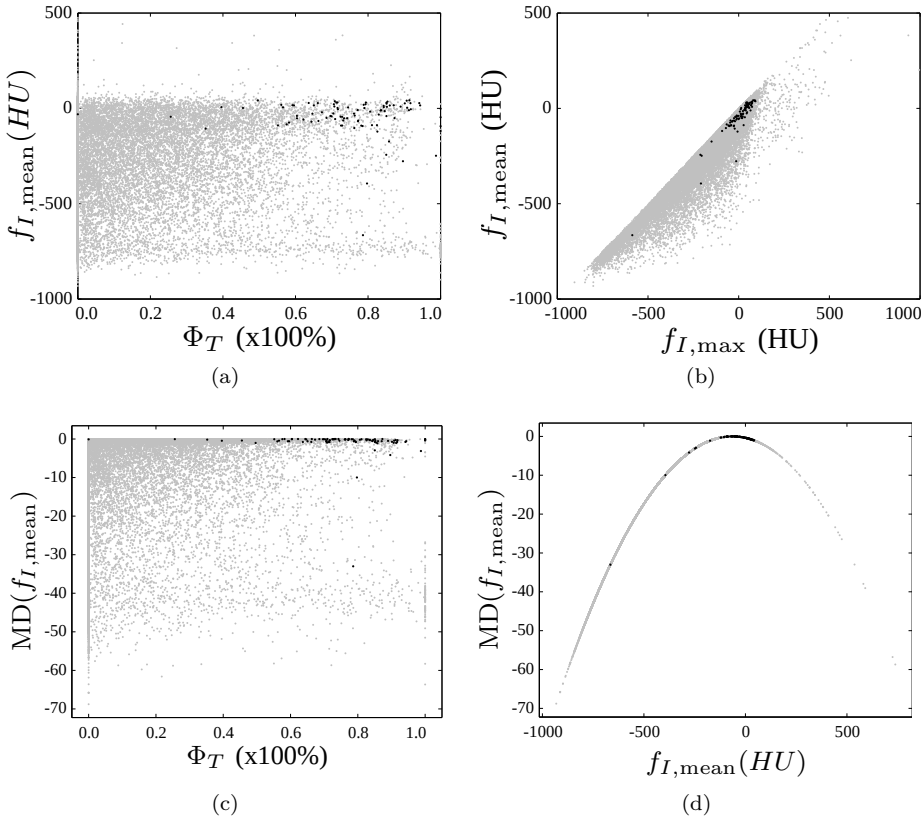
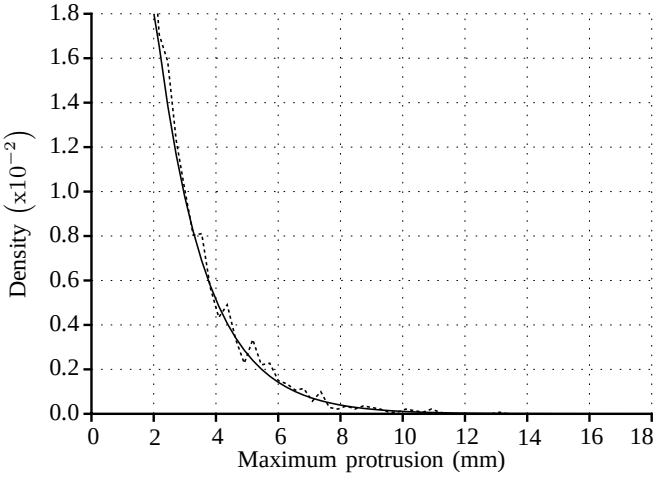
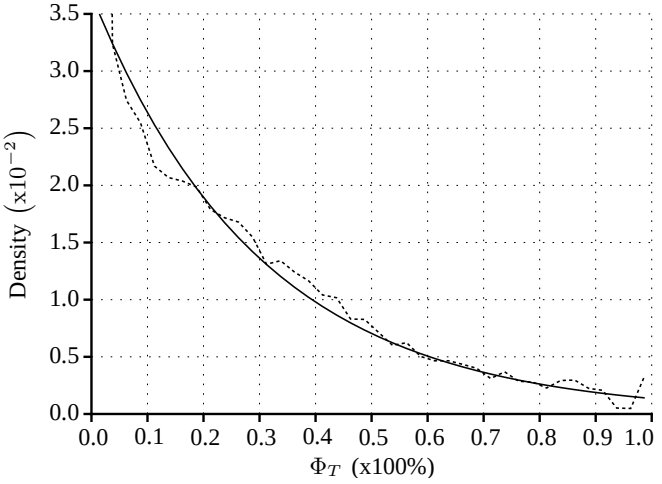


Figure 2.4: Scatter plots demonstrating the distribution of the candidates for data set 'A'. The grey dots are false detections and the black dots are polyps. (a) Mean intensity vs. Φ_T. (b) Mean intensity vs. maximum intensity. (c) The same feature space as (a) with the output of the negated Mahalanobis distance mapping on the vertical axis. This mapping is introduced in Section 2.4.1. (d) The influence of the mapping on $f_{I,\mathrm{mean}}$. Note that candidates with a high and low mean intensity have a lower mapped feature than the polyps.



(a)



(b)

Figure 2.5: 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).

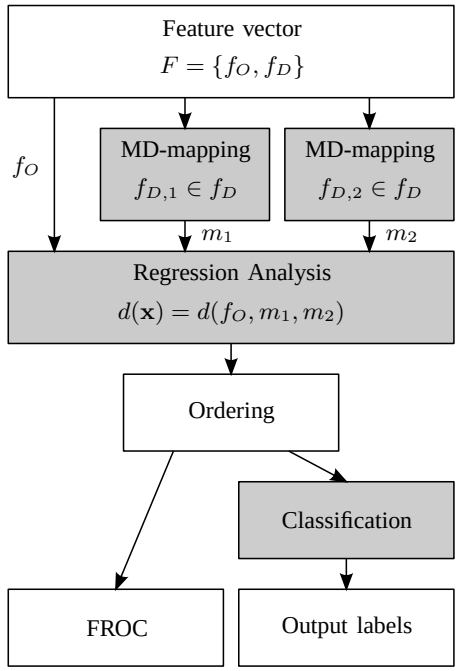


Figure 2.6: Schematic representation of the classification system. The classification starts with a feature vector consisting of features suitable for ordering (f_O) and features suitable for density estimation (f_D). The feature sets $f_{D,1}$ and $f_{D,2}$ are processed through two mappings. An ordering 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.

outliers, i.e. objects not belonging to the target class. This procedure compares to the operation of a Gaussian one-class classifier [54].

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 $f_{I,\mathrm{mean}}$. Effectively, candidates on the rectal tubes as well as candidates with an abnormal intensity are rejected. Fig. 2.4 illustrates the influence of the mapping on $f_{I,\mathrm{mean}}$.

In comparison to Wang et al. [44], 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.

2.4.2 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. 2.5). 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.

2.4.3 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. 2.5). Fig. 2.4 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:

$$\log\left(\frac{p(\mathbf{x}|\omega_p)}{p(\mathbf{x}|\omega_f)}\right) = d(\mathbf{x}) \tag{2.2}$$

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. 2.2 the assumption made by a logistic classifier, which corresponds to sigmoidal posterior probability density functions:

$$p(\omega_f, \mathbf{x}) = \frac{1}{1+\exp\left(d(\mathbf{x})\right)},\ p(\omega_p, \mathbf{x}) = 1 - p(\omega_f, \mathbf{x}). \tag{2.3}$$

The weights of the discriminant function can be determined by a maximum likelihood estimator [61].

2.5 Results

Classifier selection aims at choosing the best method for the regression analysis in our classification system (see Fig. 2.6). Three classifiers will be analyzed: the LDC, the QDC and the logistic classifier (see Section 2.4). 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. 2.6 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 2.5.2.

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 2.5.3 by means of an evaluation using data sets from four different medical centers (see Section 2.2.1).

2.5.1 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

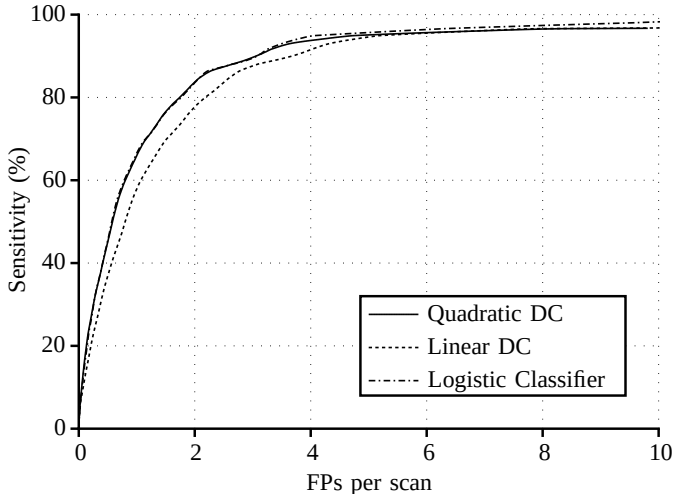


Figure 2.7: FROC curves averaged over all feature sets for the LDC, QDC and logistic classifiers.

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. 2.7. The standard deviation 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.

Table 2.2: Instability of various classifiers

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

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 [62]. Table 2.2 lists the instability measures. The table clearly shows that the logistic classifier and the LDC are the most stable classifiers.

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 matrix for the polyp class. This is reflected by the poor instability of the QDC.

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.

2.5.2 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 2.1).

Fig. 2.8a 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 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. 2.8b. 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' 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.

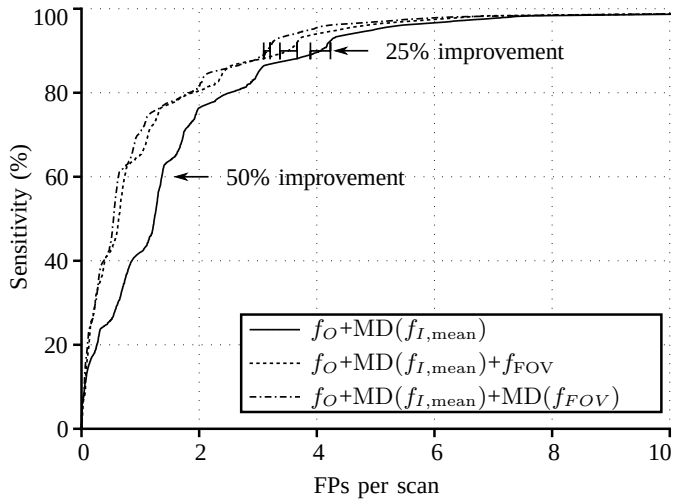
The second step was to compute the same FROC curves with the mapped mean intensity feature. A striking improvement can be seen. This result can be explained by the fact that in this case there are both false detections with lower mean intensity as there are false detections with higher mean intensity.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

2.5.3 Multi Center Evaluation

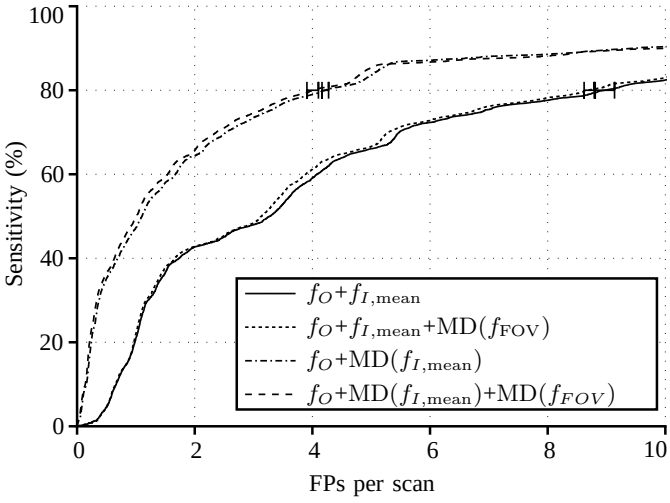
An important aspect of a CAD system for CT colonography is its ability to generalize to data sets differing in a variety of aspects. The generalization power of the presented system will be investigated by FROC analysis and a cross-center evaluation.

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 $\geq$ 6 mm in total (see Table 2.3(a)). The preparation of the patients is at the basis of the differences in the number of missed findings. The patients of data set 'A' had undergone an extensive preparation. This might explain the fact



(a) Data set 'A'



(b) Data set 'C'

Figure 2.8: 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.

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 [24]). 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 [63]; see Table 2.3(b)) .

Fig. 2.9 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. 2.9 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 [64] 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.

Table 2.3: Results of the candidate detection system: (a) polyp findings ≥ 6 mm, (b) polyps ≥ 6 mm, and (c) the number of false detections

Data set	Number of annotations	Number of detections	Detection rate (%)
'A'	94	93	99
'B'	49	38	78
'C'	340	297	87
'D'	8	8	100
Total	491	436	89

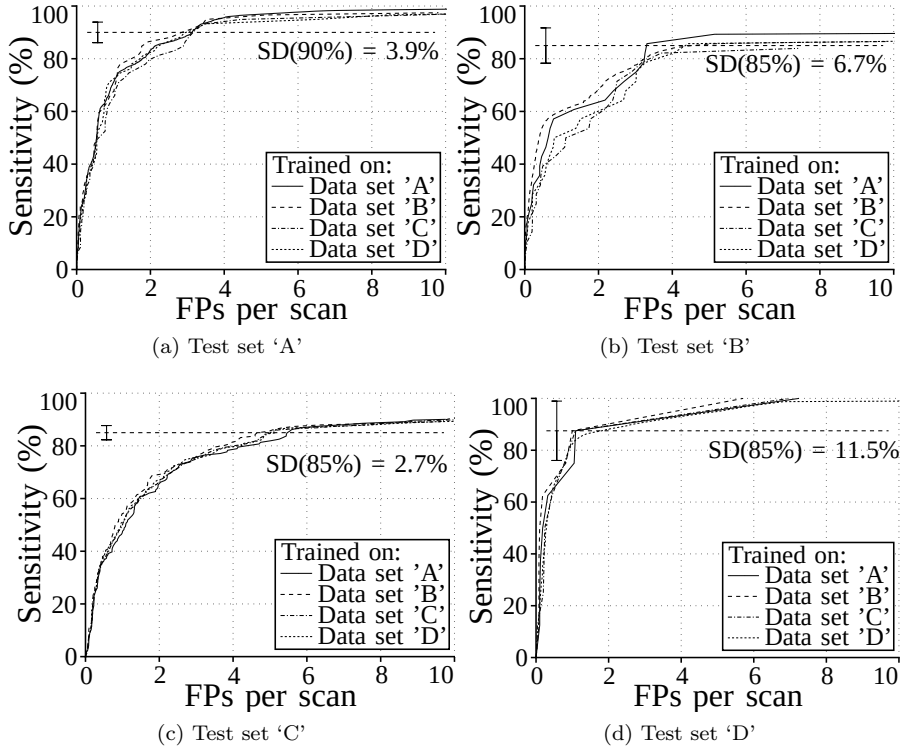
(a) Polyp findings (≥ 6 mm)

Data set	Number of annotations	Number of detections	Detection rate (%)
'A'	59	59	100
'B'	28	28	100
'C'	176	174	99
'D'	8	8	100
Total	271	269	99

(b) Polyps (≥ 6 mm)

Data set	Number of false detections
'A'	28 678
'B'	12 334
'C'	53 698
'D'	8026
Total	102 736

(c) False detections



(a) Test set 'A'

(b) Test set 'B'

(c) Test set 'C'

(d) Test set 'D'

Figure 2.9: 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.

2.6 Discussion / 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 shear 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 [6, 3, 9] and competed with other CAD systems [9, 43, 29, 30]. 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.

Chapter 3

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.

3.1 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. 3.1). There is an increasing interest in computer aided detection (CAD) systems for polyp detection in CTC data to assist the radiologist [53, 40, 65, 66, 9, 29, 67]. 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 [68, 69].

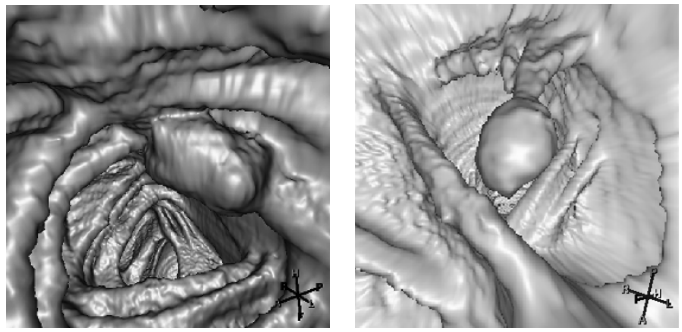
Adenomatous polyps are important precursors to colon cancer and develop due to genetic mutations in the mucosa cells [70]. This process of oncogenesis leads to enhanced cell proliferation causing the polyp to grow and to evolve from a small adenoma into a large adenoma into a carcinoma. This induces a morphological change to the colon surface[1], 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 analyse the local curvature of the colon surface, which is subsequently used to compute shape descriptors such as shape index or curvedness [37, 38]. 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 method's efficiency, we make use of a pattern recognition system introduced by us in Ref. [73]. 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 principal curvature flow.

[1] Not all colonic lesions grow into protruding polyps. It is estimated that approximately 10% of the lesions are so-called flat adenomas [71, 72].



(a) (b)

Figure 3.1: Isosurface renderings (at -650 HU) of the colon surface showing typical polyps in their structured surroundings.

3.1.1 Previous Work

For the detection of candidate regions, Summers et al. [37] 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 neighbourhoods of a triangulated isosurface [36]. Candidates were generated by selecting a range of mean and Gaussian curvature values. Additionaly, a large number of other shape criteria were used ([74]: Table 2), to limit the number of false positive detections. Similarly, Yoshida et al. [38] 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:

$$\begin{aligned} SI &= \frac{1}{2} - \frac{1}{\pi}\arctan\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right), \\ CV &= \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}}, \end{aligned} \tag{3.1}$$

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. [40] 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 (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. [53, 75] 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, 76, 29, 53, 38, 77] apply several conservative thresholds to the mean curvature, principal curvatures, sphericity ratio and/or shape index to generate candidate regions.

3.1.2 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 [59] of lesion size, precise 'delineation' (or segmentation) of the candidate is needed. Although a number of methods are available for segmentation purposes [74, 78, 79], 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 analyse the effect of scale or to determine the optimal scale for polyp detection. It is partly addressed in [45] 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 [80, 81].

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.

A steadily growing number of papers ([9, 82, 83, 38, 84, 85, 30]) 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 as it might add more parameters,
3. it must be able to cope with the whole polyp size range encountered in practice, and
4. it should take into account the increased clinical relevance of larger polyps.

3.1.3 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 original (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 [10] 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. Any (beneficial) information from 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 tuning (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.

3.2 Methodology

3.2.1 Materials

A total of 84 patients with an increased risk for colorectal cancer were consecutively included in a previous study [86]. All data were acquired using a Mx8000

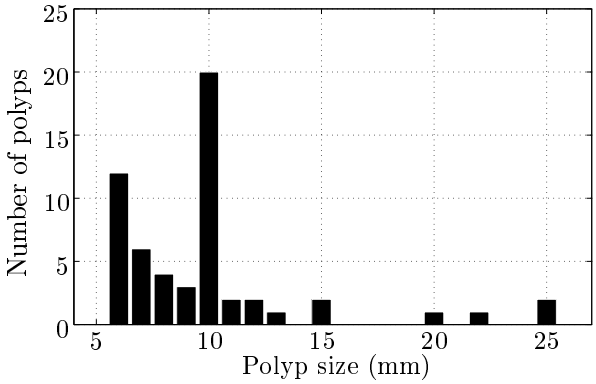
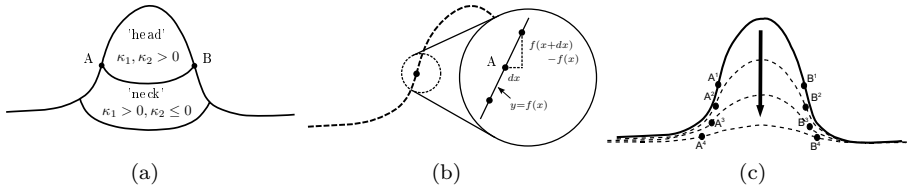


Figure 3.2: Distribution of sizes obtained during colonoscopy of 57 polyps larger than or equal to 6 mm in 84 patients from a previous study [86]. One polyp of 45 mm is not visible in the histogram.

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. 3.2 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.



(a) (b) (c)

Figure 3.3: 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.

3.2.2 Method

A typical polypoid shape is shown in Fig. 3.3(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. 3.3(c)). Note that the convex region expands during the process and will ultimately include the polyp neck as well. After a certain amount of 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. 3.3(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. 3.3(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 given by:

$$\kappa = -\frac{\Delta y}{|\nabla y|}, \tag{3.2}$$

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:

$$\Delta y = \frac{d^2 f}{dx^2} = \lim_{dx \to 0} \frac{\frac{f(x+dx)-f(x)}{dx} - \frac{f(x)-f(x-dx)}{dx}}{dx} = 0. \tag{3.3}$$

A reduction of the protrusion in the head region implies that the value of $f(x + dx)$ in (3.3) 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. 3.3(b) and the head region will expand into the neck region.

The effect of repeatedly reducing the protrusion is illustrated in Fig. 3.3(c). The points with zero second principal curvature shift from A^1 to A^4 and B^1 to B^4. 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.

3.2.3 Second Principal Curvature Flow

A scheme to remove protruding elements from a curve in 2D is the Euclidean shortening flow [87]. A similar approach can be taken in 3D, for which the flow is governed by:

$$\frac{\partial I}{\partial t} = -g(\kappa_1, \kappa_2) \left|\nabla I\right|, \tag{3.4}$$

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:

$$H = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix}, \tag{3.5}$$

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 [88]: 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:

$$\begin{aligned} I_{uu} &= -\kappa_1 \left|\nabla I\right|, \\ I_{vv} &= -\kappa_2 \left|\nabla I\right|. \end{aligned} \tag{3.6}$$

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 [89], e.g. by the mean curvature [90, 91] 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:

$$\frac{\partial I}{\partial t} = \begin{cases} I_{vv} & (\kappa_2 > 0) \\ 0 & (\kappa_2 \leq 0) \end{cases}. \tag{3.7}$$

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.

3.2.4 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 (3.7) 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 (3.7) is discretized by a forward difference scheme:

$$\frac{\partial I}{\partial t} = \frac{I^{t+1} - I^t}{dt} + O(dt). \tag{3.8}$$

The right hand side of (3.7) 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 [92]. Therefore, the first and second order derivatives are computed after a (second order) Taylor expansion in a 3x3x3 neighbourhood [93]. 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:

$$I^{t+1} = \begin{cases} I^t + \frac{\Delta t}{(\Delta x)^2} I_{vv}^{t+1/2} & (\kappa_2 > 0) \\ I^t & (\kappa_2 \leq 0) \end{cases}, \tag{3.9}$$

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 $\frac{1}{6}$ [94]. 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 [87]. For a more formal discussion, see [95] and also [94]. 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 (3.9) 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. 3.3). 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 3.3.2).

Fig. 3.4 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 ($\sigma = 2, 4, 8, 12\,\text{mm}$), 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. 3.11b for the performance of our algorithm on large objects).

3.2.5 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 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:

$$I^{t=t_i}(\vec{x}) = I^{t=t_0}(\vec{x}) + \delta \cdot \nabla I^{t=t_0}(\vec{x}) + \epsilon. \tag{3.10}$$

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:

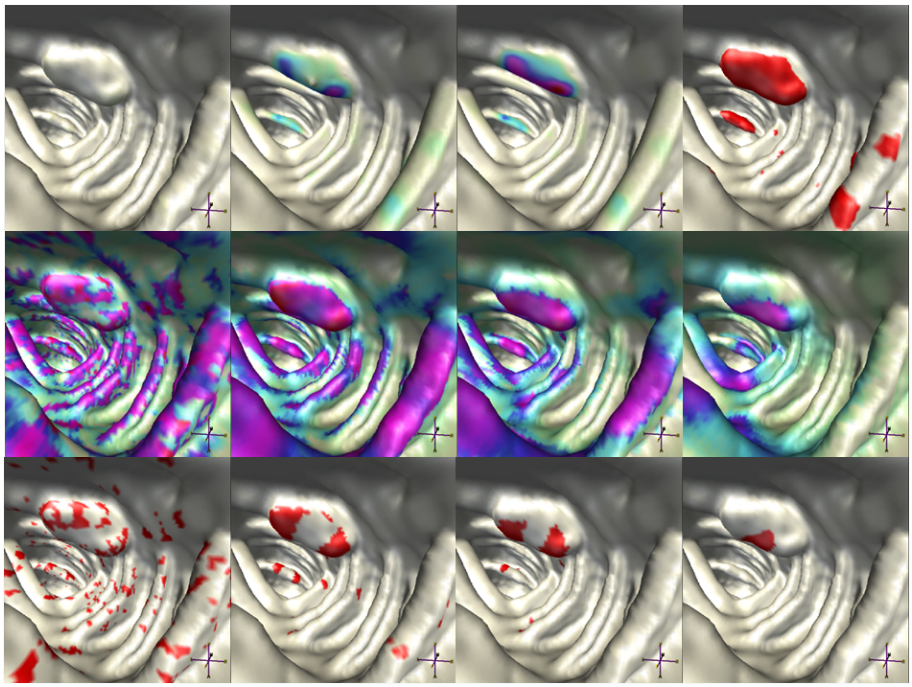


Figure 3.4: Demonstration of polyp detection by the curvature flow (first row). The second and third row show results as obtained by thresholding the Shape Index, computed at different scales. See text for details.

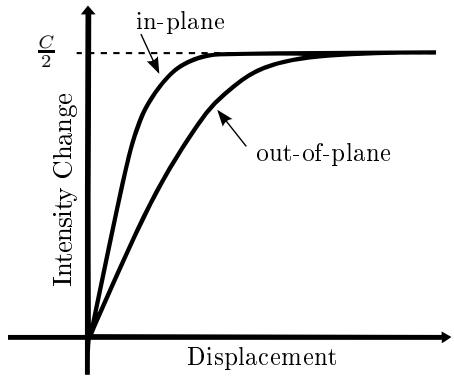


Figure 3.5: Sketch of the relation between colon surface displacement and the observed intensity change for positions halfway the step edge. The relation depends on the apparent local scale of the PSF, i.e. the scale in the direction of the surface normal. Often, the scanner resolution is not isotropic: 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.

$$I^{t=t_\infty}(\vec{x}) = I^{t=t_0}(\vec{x}) - \frac{C}{2}, \tag{3.11}$$

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

The sketch in Fig. 3.5 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 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. 3.5) 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 $\sigma_i = \sqrt{\sigma^2_{\text{target}} - \sigma^2_{\text{apparent},i}}$, 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 [10] for data with an assumed Gaussian PSF [60] with $\sigma = 1.6\,\mathrm{mm}^2$.

3.2.6 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 diameter, excluding the stalk. We compute it automatically using the method described in [79], 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. 3.5). 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.

3.2.7 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. 3.6. 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 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

[2] Halfway the air-tissue transition: $\nabla I^{t=t_0} = \frac{C}{\sigma\sqrt{2\pi}} = \frac{1000}{1.6\sqrt{2\pi}} \approx 250\,\mathrm{HU/mm}$, thus $100\,\mathrm{HU} \triangleq 0.4\,\mathrm{mm}$, i.e. equal to the threshold used in [10].

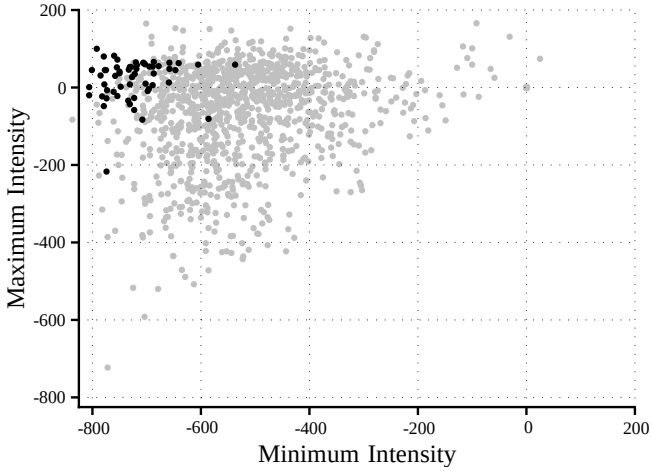


Figure 3.6: Feature space of the maximum and minimum intensities for each candidate region. Annotated polyps are depicted by black dots and have maximum intensities around 0 HU (tissue) and minimum intensities around -650 HU. Only one in every 20 false positives is shown as a grey dot.

independent training data. This strategy mimics the use of a Gaussian one-class classifier [54]. 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 3x3x3 Taylor's 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 3.6 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 [61] is 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.

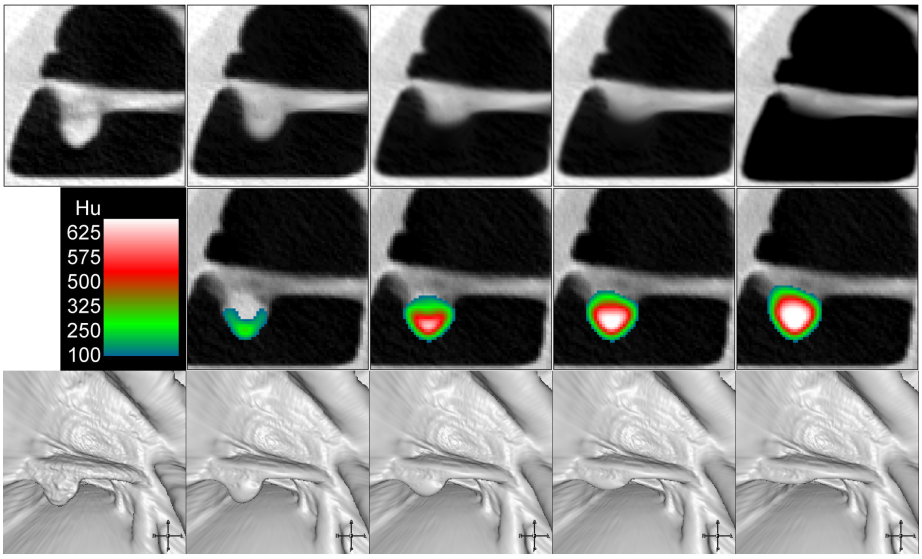


Figure 3.7: Polyp (10 mm) at different stages of the intensity deformation (after 0, 10, 40, 80 and 160 iterations of (3.9)) . First row: original data; second row: overlay showing the intensity changes larger than 100 HU (the color scale was truncated at 650 HU; third row: isosurface renderings (at -650 HU).

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).

3.3 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 [10]

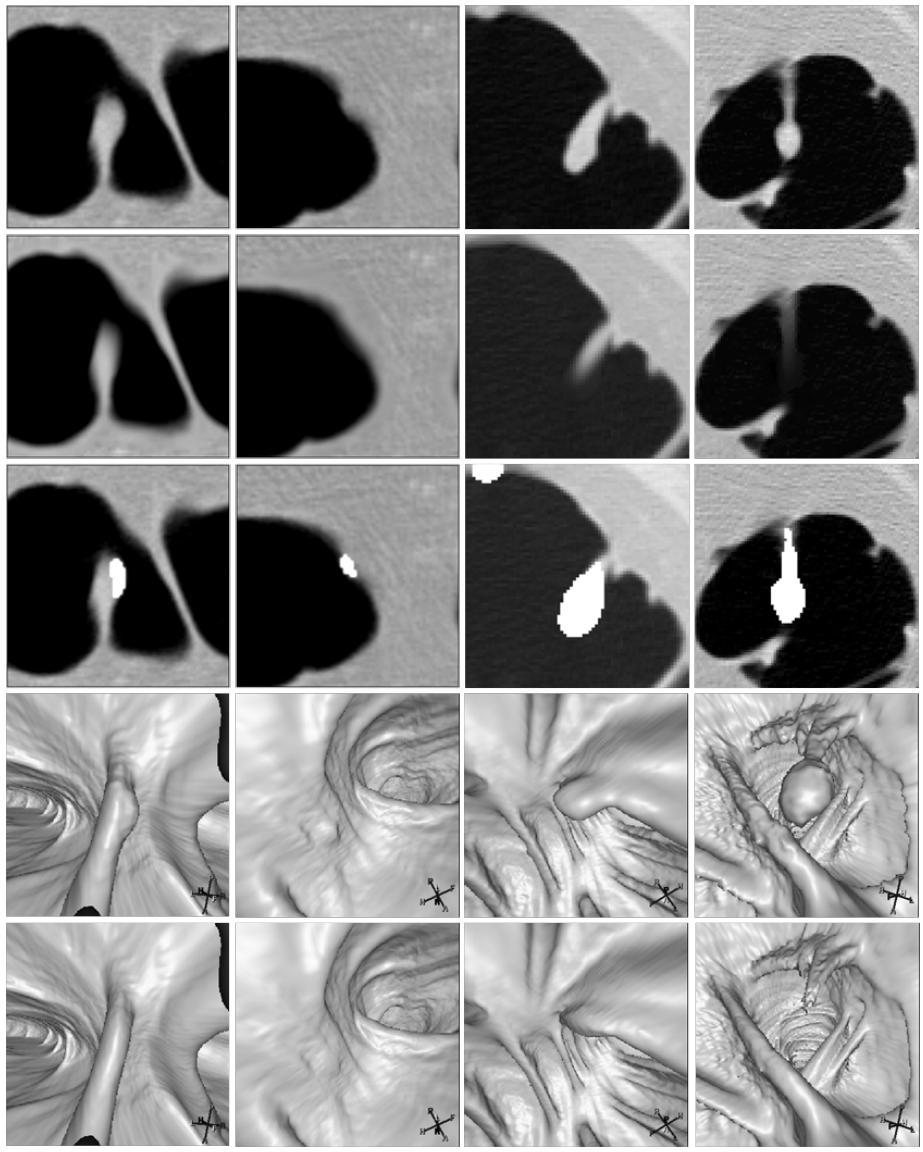


Figure 3.8: Typical results for four polyps. Each column shows the results for a different polyp. The first two rows show grey value cross sections before and after intensity deformation. The third row shows the segmentation masks which are obtained by thresholding the intensity change at a level of 100 HU. The last two rows show isosurface renderings (at -650 HU) of the polyps before and after intensity deformation.

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.

3.3.1 Qualitative Analysis

Fig. 3.7 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 (3.9). 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.

Fig. 3.8 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. 3.9).

3.3.2 Performance of the Candidate Detection

Fig. 3.10 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. 3.10(a) was

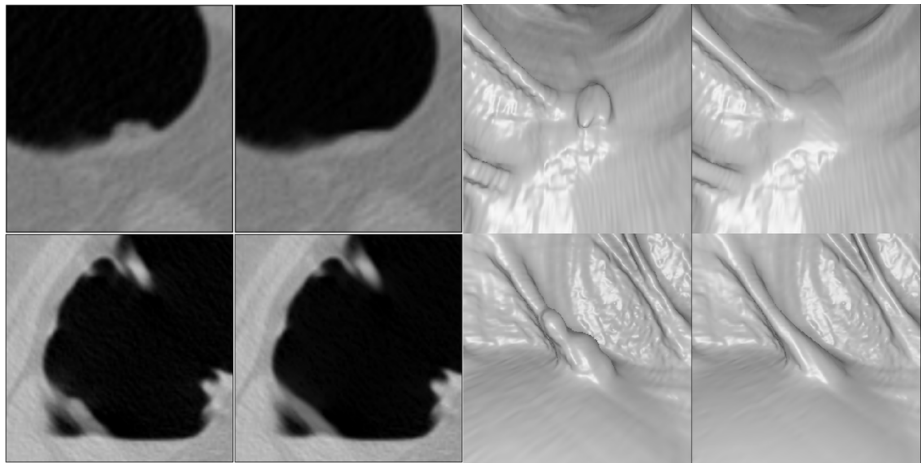


Figure 3.9: Each row shows a false positive. First row: example of stool. Air inside object is clearly visible on first image. Second row: stool on a fold. The original data is shown in the first and third column. The data after deformation by curvature flow is shown in columns two and four.

obtained using the method that involves an explicit (mesh) representation of the colon surface [10] and Fig. 3.10(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 system's performance is really determined after classifying the candidates (see below).

3.3.3 Results after Classification

Fig. 3.11 shows the overall performance of both the proposed and the mesh based method [10]. The figure shows the performance for the detection of polyps for two 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

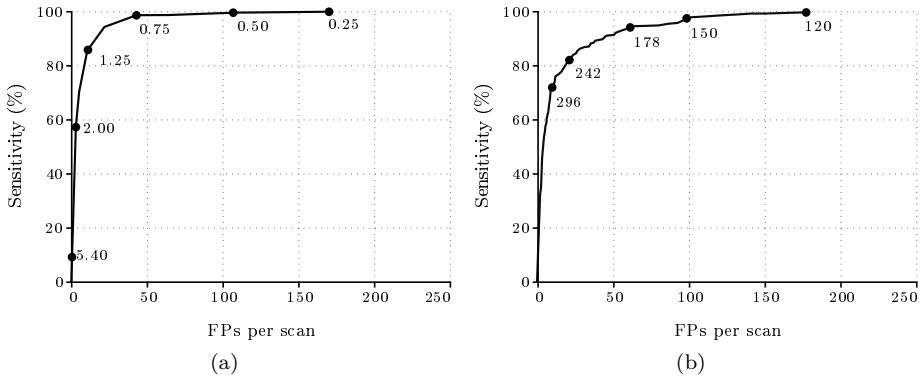


Figure 3.10: FROC curves showing the candidate detection sensitivity versus the number of false positives for (a) the mesh based and (b) the currently proposed technique. The numbers in (a) denote the threshold on the deformation field in mm and in (b) the threshold on the intensity change field in HU.

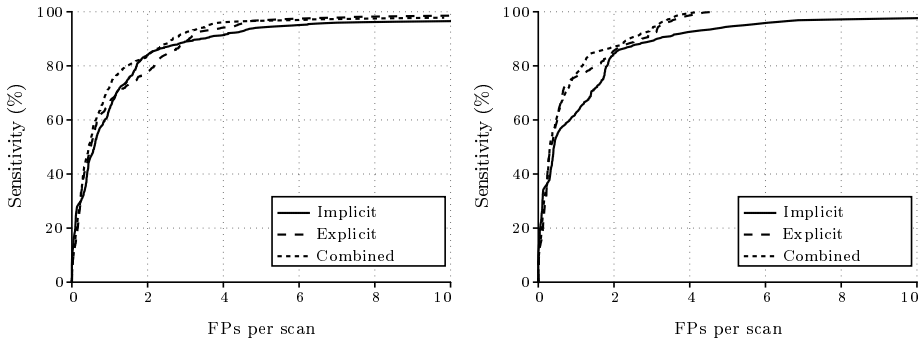
polyps. A sensitivity of 95% for polyps $\geq$ 6 mm is achieved at an average false positive rate of 4-6 per scan. For polyps $\geq$ 10 mm, a sensitivity of 95% is obtained at about 4 false positives per scan.

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. 3.9).

3.3.4 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. 3.12 confirms that 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



(a) Polyps $\geq$ 6 mm (b) Polyps $\geq$ 10 mm

Figure 3.11: 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.

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. 3.11 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 $\geq$ 6 mm.

3.4 Discussion / 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

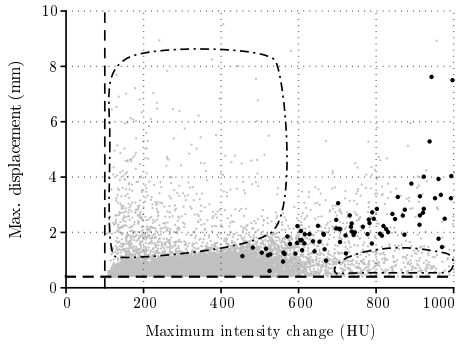


Figure 3.12: Feature space of the maximum displacement (explicit method) vs. the maximum intensity change (implicit method). The black dots correspond to polyps and the grey dots to false detections. Two regions (encircled by dash-dotted lines top-left and bottom-right) with false detections (grey dots) can be observed in which the depicted features are uncorrelated and complementary.

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. 3.8 (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 3x3x3 neighborhood will effectively deal with the noise practically encountered in low-dose (20 mAs) scans.

The method's performance on so-called flat polyps requires further research.

Chapter 4

Combining Mesh, Volume, and Streamlines Representations for Polyp Detection

CT colonography is a screening technique for adenomatous colorectal polyps, which are important precursors to colon cancer. Computer aided detection (CAD) systems are developed to assist radiologists. We present a CAD system that orders the polyps according to clinical relevance (size) and substantially reduces false positives while keeping the sensitivity high. Hereto, we combine protrusion measures derived after a nonlinear partial derivative equation (PDE) is applied to both an explicit mesh and an implicit volumetric representation of the colon wall. Hence, surface as well as intensity characteristics are exploited. The shape of the protruding elements is efficiently described via a technique from data visualization based on curvature streamlines. A low-complex pattern recognition system based on an intuitive feature from the aforementioned representations improves performance to less than 1.6 false positive per scan at 92% sensitivity per polyp.

4.1 Introduction

Colorectal cancer is the second leading cause of death due to cancer in the Western world [1]. It has been shown that screening for adenomatous colorectal polyps, which are important precursors to cancer, and subsequent removal of identified lesions significantly reduces the incidence of colon carcinoma [2, 28]. Computed tomography colonography (CTC) is a rapidly evolving technique that is advocated for screening. To assist the radiologists, effort is put in the development of computer aided detection (CAD) systems [43, 41, 29, 9, 57, 73].

Traditionally, polyps are tentatively detected by curvature derived features. Subsequently, the candidates thus obtained are classified by curvature as well as material structure features. The latter are typically involved to reject falsely detected stool rests (frequently having a granulated grey-value structure due to air bubbles) and false detections emanating from partial volume effects. It was demonstrated that no other features but the aforementioned ones are required for a performance that is comparable to optical colonoscopy [73].

In previous work, it is shown that polyps can be detected equally well as protrusions on an explicit (mesh-surface) representation as on an implicit (grey-level) representation of the colon wall [57]. In this paper, we present a low-complex, unambiguous pattern recognition step, which combines the two approaches. It will be shown why the two techniques are to some extent complementary and it will be demonstrated how techniques from data visualization are efficiently incorporated in our framework.

4.2 Materials

For evaluation, a subset of 28 patients from a larger study [6] is used. All patients adhered to an extensive laxative regime and no fecal tagging agent was administered. The data sets consist of scans in both prone and supine positions; the slice thickness was 3.2 mm. The reference standard is optical colonoscopy. Expert radiologists retrospectively indicated the location of polyps by annotating a point in the 3D data set using the reference standard. 65 polyp annotations were made in the 56 scans, corresponding to 40 polyps larger than or equal to 6 mm. The candidate segmentations were labeled by comparison to these annotations. A polyp was counted as a true positive CAD detection if it was found in at least one of the two scanned positions.

4.3 Protrusion-Based Detection of Polyps

Polyps may be characterized by the condition that the smallest principal curvature is larger than zero. In other words, they are caplike structures, whereas colonic folds are elongated with typically one positive curvature and the other close to zero or (slightly) negative. Because of the cylindrical global nature of

the colon, regions with two positive curvatures are relatively scarce. The principal curvatures of the colon wall can either be computed by differentiation of a mesh representation or directly from the underlying image data [88].

Previously, we introduced two distinct, iterative schemes for the detection of polypoid objects by the amount of protrusion compared to the background. The updating function in both schemes is abstractly defined as:

$$X^{t+1} = X^t + dt \cdot f(-\kappa_2) \tag{4.1}$$

in which $f(-\kappa_2)$ is a function that is designed to operate on an object only if the smallest curvature is positive and dt is a time step [10, 57]. In the explicit method $f(\cdot)$ is related to the force applied to the mesh vertices, whereas in the implicit method this function modifies the intensity of voxels in such a way that 'protruding' intensities are smoothed into the background. Both approaches involve a repeated application of (4.1) until κ_2 is smaller than or equal to zero everywhere or until some convergence criterium is satisfied.

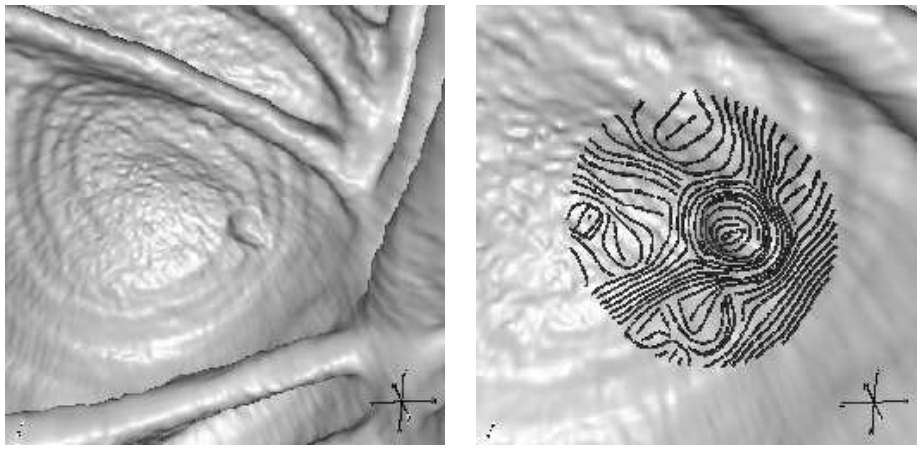
4.3.1 Complementary Protrusion Analysis

It may be concluded from the previous section that polyps are protrusions in both the explicit (mesh) and in the implicit (grey-value) representation of the colon wall. In practice we have found that particularly the false detections of both methods are to some extent uncorrelated. For instance, the explicit method typically had some false detections emanating from partial volume effects (PVE), whereas the implicit method was more robust because it took the internal intensities into account. Reversely, the implicit method, using neighbourhood information, is inherently sensitive to neighbouring structures, whereas such problems are excluded in the explicit method in which feature measurement is confined along the (infinitesimal thin) mesh. Because of the thinness of some folds, the intensities inside folds are influenced by the PVE. As a consequence, the implicit method also detects candidates inside a fold. The mesh method has only limited response at these locations.

The explicit approach directly acts on a representation of the colon wall, whereas the implicit method interacts with the underlying data. These two representations also reflect the features to distinguish true polyps from false detections. Note that the protruding aspect is predominantly represented in the mesh representing the colon wall, whereas the material structure is captured in the underlying data. We conceive a combined approach in which the protrusion extent is better represented in the explicit method, whereas the volumetric and intensity properties are delivered by the implicit method.

4.3.2 Shape Analysis

It may be observed that so far information about the shape of the objects remains limited. Effectively, lesions are detected by the amount of protrusion only,



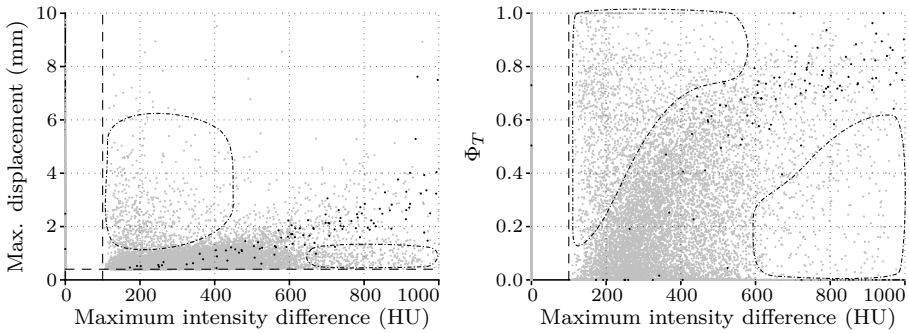
(a) (b)

Figure 4.1: (a) A polyp with its surrounding environment. (b) Example of the curvature streamlines generated in the vicinity of the polyp.

irrespective of shape. An effective representation of shape was recently described in the literature on data visualization, deriving from curvature streamline analysis.

Curvature streamlines (or lines of curvature) are defined as lines that are tangent everywhere to one of the two principal curvature direction vector fields on the surface. Techniques for deriving curvature streamlines on the surface were presented in [96]. In order to capture the essential surface shape information, streamlines were adaptively spaced over the whole surface with spacing dependent on the local principal curvature magnitudes. On less curved surface regions, fewer streamlines were generated than on highly curved surface regions.

Curvature streamlines that are constrained to the colonic wall have the useful characteristic that they tend to encircle polyp necks. The 'winding angle' feature was derived to utilize this characteristic. It is defined as the cumulative signed change of direction along a streamline. At each sample point, the differential change of direction is determined based on the surface normal at that point. Closed streamlines, such as those around polyp necks, have a winding angle of at least 2π. A candidate was assigned the maximum winding angle of a streamline in its vicinity. Initial experimentation showed that this winding angle feature correlated highly with true polyp detections and could thus be useful to reduce the number of false positives (FPs) found by CAD systems [96]. Special care should be taken to ensure that streamlines are sufficiently long in order to fully capture polyp surface geometry [97].



(a) (b)

Figure 4.2: Scatter plots of a larger (equivalent) data set (86 patients) showing (a) the maximum mesh displacement of a candidate, and (b) Φ_T vs. the maximum intensity difference derived from the implicit method. The black dots correspond to polyps and the grey dots to false detections.

4.4 Experiments and Results

The detections on the explicit representation (Section 4.4.1) are at the basis to analyse how protrusions are detected with the implicit method and how the streamline analysis may contribute (Section 4.4.2). In the last section, FROC analysis shows the improvement of each newly added feature.

4.4.1 Combined CAD System Based on Protrusion

The explicit method actually involves two features. First, we use a feature derived from the displacement field of the mesh. This feature measures the percentage of the candidate with a displacement larger than a certain threshold T, further denoted as Φ_T. We use a threshold of 0.6 mm as in [73]. This design favors candidates with steep edges and compact forms. The second feature we use is the mean intensity of the candidate [73]. This feature is calculated by a tonal weighted sum of all voxels included in a segmentation mask. The latter consists of the area included between the original and the displaced mesh.

To analyse the performance of a combined system, the correspondences between candidates found by both methods should be established. The implicit method acts only on these regions in the image where a candidate was found by the explicit method. These regions are obtained by ten times dilation of the binary segmentation mask of the candidate. A corresponding segmentation area for the implicit method is derived from the deformed image by thresholding the intensity difference at a value of 100 Hounsfield unit (HU) (as in [57]). Thus, each candidate from the implicit method is inherently linked to a corresponding

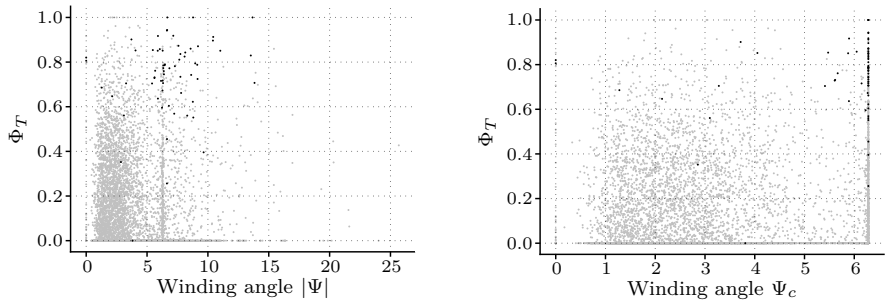


Figure 4.3: Scatter plots showing Φ_T versus (a) $|\Psi|$ and (b) Ψ_c. The black dots correspond to polyps and the grey dots to false detections.

detection on the mesh.

Fig. 4.2 contains two scatter plots of the maximum intensity difference derived from the implicit method versus the maximum displacement of the mesh (Fig. 4.2(a)) and versus Φ_T derived from the displacement field (Fig. 4.2(b)). It can be seen that the maximum mesh displacement and the maximum intensity difference correlate well for polyps (black dots). In both scatter plots two clusters of false detections (grey dots) can also be observed in which the depicted features are uncorrelated (top-left and bottom right in both graphs). One cluster has rather low maximum intensity change, but concurrently quite large maximum displacement of the mesh or Φ_T; another cluster is characterized by a large maximum intensity difference, but a low maximum mesh displacement or Φ_T. The indicated boundaries of the feature space (dashed lines) in Fig. 4.2(a) represent the two thresholds used in the candidate generation.

4.4.2 Streamline Analysis

For all detections on the mesh, a center line through the center of gravity and the center of curvature of the segmentation mask is computed [79]. The intersection of this line and the mesh defines the initial seed point for the streamline analysis. For each detection curvature streamlines are generated within a spherical ROI with 16 mm radius around the seed point. As explained in Section 4.3.2, the winding angle is calculated on these streamlines as the cumulative signed change of direction along the streamline. At each sample point, the differential curvature is derived relative to the surface normal at that point. In other words, when a streamline forms a circle, its absolute winding angle is 2π or more (for polyp characterization the sign of the winding angle is not important). Importantly, an absolute winding angle of more than 2π is not necessarily more 'polyp-like' than a winding angle equal to 2π. Therefore, we clip this feature

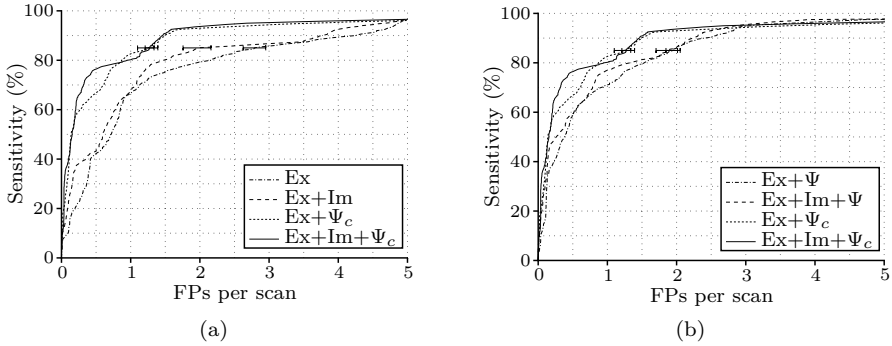


Figure 4.4: FROC curves for polyp detection systems consisting of a combination of the explicit method (Ex), the implicit method (Im), and the streamline analysis (Ψ/Ψ_c).

to a maximum of 2π, i.e. $\Psi_c = \min(|\Psi|, 2\pi)$. This hypothesis is confirmed by FROC analysis as shown in Fig. 4.4(b).

Fig. 4.3 shows scatter plots of (a) the absolute winding angle $|\Psi|$ and (b) the clipped winding angle $|\Psi_c|$ vs. Φ_T (derived from the mesh displacement field). Again, the black dots denote the polyps and the grey dots denote the false detections. Observe that almost all polyps have a winding angle close to or larger than 2π, whereas many false detections have lower winding angles. In other words, the winding angle might indeed help to distinguish between polyps and false detections.

4.4.3 FROC Analysis

Fig. 4.4 shows FROC curves describing the performance of the CAD system. The FROC curves are computed by a ten times repeated ten-fold cross-validation. In all cases, we used a logistic classifier and detections on the rectal tube were discarded.

Initially, the system was based on two features: Φ_T and the mean intensity. The performance of this system (Ex) is shown in Fig. 4.4(a) by the dash-dotted line. Then, the implicit method (Im) was added by means of the maximum intensity difference feature; the resulting performance is given by the large dashed line. Finally, the clipped winding angle feature Ψ_c was included to both previous configurations, represented by the small dashed respectively the solid line. We conclude that 92% of the polyps were detected with less than 1.6 false positives per scan when all four features are included. The error bars denote two times the standard deviation in the number of false positives over all scans at 85% sensitivity. Actually, the standard deviation of the FROC curves is over seven times smaller due to averaging over all scans.

4.5 Conclusions

We present a polyp CAD system that detects polyps based on four intuitive features. The detection consists of solving Equation 4.1 by means of two different approaches, which characterize different aspects of the candidates. In effect, protruding objects are detected by means of deforming an explicit representation of the colon surface or by means of modifying the intensity data containing an implicit representation. We also added a shape-descriptor derived from curvature streamline analysis. It was shown that the feature based on the mesh displacement field, the mean intensity, the maximum intensity difference and the streamline's winding angle are sufficient for optimal performance. We analyzed 56 scans from 28 patients and it was found that over 92% of the polyps were detected with less than 1.6 false positives per scan.

Chapter 5

Recognition of Protruding Objects in Highly Structured Surroundings by Structural Inference

Object recognition in highly structured surroundings is a challenging task, because the appearance of target objects changes due to fluctuations in their surroundings. This makes the problem highly context dependent. Due to the lack of knowledge about the target class, we also encounter a difficulty delimiting the non-target class. Hence, objects can neither be recognized by their similarity to prototypes of the target class, nor by their similarity to the non-target class. We solve this problem by introducing a transformation that will eliminate the objects from the structured surroundings. Now, the dissimilarity between an object and its surrounding (non-target class) is inferred from the difference between the local image before and after transformation. This forms the basis of the detection and classification of polyps in computed tomography colonography. 95% of the polyps are detected at the expense of four false positives per scan.

5.1 Introduction

For classification tasks that can be solved by an expert, there exists a set of features for which the classes are separable. If we encounter class overlap, not enough features are obtained or the features are not chosen well enough. This conveys the viewpoint that a feature vector representation directly reduces the object representation [98]. In the field of imaging, the objects are represented by their grey (or color) values in the image. This sampling is already a reduced representation of the real world object and one has to ascertain that the acquired digital image still holds sufficient information to complete the classification task successfully. If so, all information is still retained and the problem reduces to a search for an object representation that will reveal the class separability.

Using all pixels (or voxels) as features would give a feature set for which there is no class overlap. However, this feature set usually forms a very high dimensional feature space and the problem would be sensitive to the curse of dimensionality. Considering a classification problem in which the objects are regions of interest $\mathcal{V}$ with size N from an image with dimensionality D, the dimensionality of the feature space Ω would then be N^D, i.e. the number of pixels in $\mathcal{V}$. This high dimensionality poses problems for statistical pattern recognition approaches. To avoid these problems, principal component analysis (PCA) could for example be used to reduce the dimensionality of the data without having the user to design a feature vector representation of the object (see for example [99]). Although PCA is designed to reduce the dimensionality while keeping as most information as possible, the mapping unavoidably reduces the object representation.

The use of statistical approaches completely neglects that images often contain structured data. One can think of images that are very similar (images that are close in the feature space spanned by all pixel values), but might contain significantly different structures. Classification of such structured data receives a lot of attention and is motivated by the idea that humans interpret images by perception of structure rather than by perception of all individual pixel values. An approach for the representation of structure of objects is to represent the objects by their dissimilarities to other objects [100]. When a dissimilarity measure is defined (for example the 'cost' of deforming an object into another object), the object can be classified based on the dissimilarities of the object to a set (or sets) of prototypes representing the classes.

Classification based on dissimilarities demands prototypes of both classes, but this demand can not always be fulfilled. For example, the detection of target objects in highly structured surroundings poses two problems. First, there is a fundamental problem describing the class of non-targets. Even if there is detailed knowledge about the target objects, the class of non-targets (or outliers) is merely defined as all other objects. Second, if the surroundings of the target objects is highly structured, the number of non-target prototypes is very large and they all differ each in their own way, i.e. they are scattered all

over the feature space. The selection of a finite set of prototypes that sufficiently represents the non-target class is almost impossible and one might have to rely on one-class classification.

In this paper we link image processing to dissimilarity based pattern recognition. The solution is based on structural inference. Featureless pattern recognition is extended to classification in the absence of prototypes. The role of prototypes is replaced by a single context-dependent prototype that is derived from the image itself by a specific transformation for the application at hand. The approach will be applied in the context of automated polyp detection.

5.2 Automated Polyp Detection

The application that we present in this paper is automated polyp detection in computed tomography (CT) colonography (CTC). Adenomatous polyps are important precursors to cancer and early removal of such polyps can reduce the incidence of colorectal cancer significantly [2, 28]. Polyps manifest themselves as protrusions from the colon wall and are therefore visible in CT. CTC is a minimal-invasive technique for the detection of polyps and, therefore, CTC is considered a promising candidate for large-scale screening for adenomatous polyps. Computer aided detection (CAD) of polyps is being investigated to assist the radiologists. A typical CAD system consists of two consecutive steps: candidate detection to detect suspicious locations on the colon wall, and classification to classify the candidates as either a polyp or a false detection.

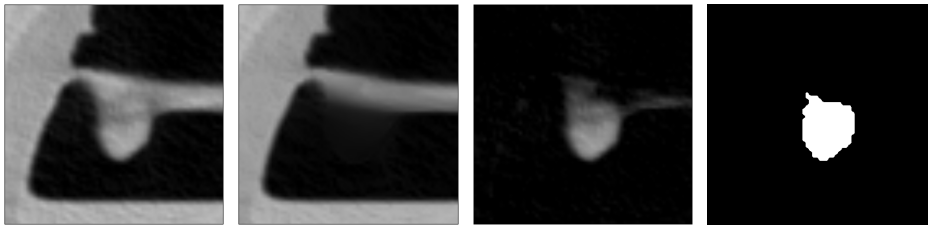
By nature the colon is highly structured; it is curved, bended and folded. This makes that the appearance of a polyp is highly dependent on its surrounding. Moreover, a polyp can even be (partly) occluded by fecal remains in the colon.

5.2.1 Candidate Detection

Candidate detection is based on a curvature-driven surface evolution [10, 57]. Due to the tube-like shape of the colon, the second principal curvature κ_2 of the colon surface is smaller than or close to zero everywhere (the normal vector points into the colon), except on protruding locations. Polyps can thus be characterized by a positive second principal curvature. The surface evolution reduces the protrusion iteratively by solving a non-linear partial differential equation (PDE):

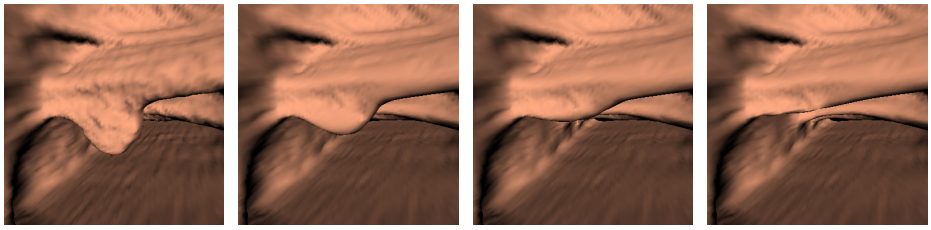
$$\frac{\partial I}{\partial t} = \begin{cases} -\kappa_2 |\nabla I| & (\kappa_2 > 0) \\ 0 & (\kappa_2 \leq 0) \end{cases} \tag{5.1}$$

where I is the three-dimensional image and $|\nabla I|$ the gradient magnitude of the image.



(a) Original (b) Solution (c) Intensity change (d) Segmentation

Figure 5.1: (a) The original CT image (grey is tissue, black is air inside the colon). (b) The result after deformation. The polyp is smoothed away and only the surrounding is retained. (c) The difference image between (a) and (b). (d) The segmentation of the polyp obtained by thresholding the intensity change image.



(a) Original (b) 20 Iterations (c) 50 Iterations (d) Result

Figure 5.2: Isosurface renderings (-750 HU) of a polyp and its surrounding. (a) Before deformation. (b–c) After 20 and 50 iterations. (d) The estimated colon surface without the polyp.

Iterative application of (5.1) will remove all protruding elements (i.e. locations where $\kappa_2 > 0$) from the image and estimates the appearance of the colon surface as if the protrusion (polyp) was never there. This is visualized in Fig. 5.1 and Fig. 5.2. Fig. 5.1(a) shows the original image with a polyp situated on a fold. The grey values are iteratively adjusted by (5.1) . The deformed image (or the solution of the PDE) is shown in Fig. 5.1(b). The surrounding is almost unchanged, whereas the polyp has completely disappeared. The change in intensity between the two images is shown in Fig. 5.1(c). Locations where the intensity change is larger than 100 HU yield the polyp candidates and their segmentation (Fig. 5.1(d)). Fig. 5.2 also shows isosurface renderings at different time-steps.

5.2.2 Related work

Konukoglu et al. [75] have proposed a related, but different approach. Their method is also based on a curvature-based surface evolution, but instead of removing protruding structures, they proposed to enhance polyp-like structures and to deform them into spherical objects. The deformation is guided by

$$\frac{\partial I}{\partial t} = \left(1 - \frac{H}{H_0}\right) |\nabla I| \tag{5.2}$$

with H the mean curvature and H_0 the curvature of the sphere towards the candidate is deformed.

5.3 Structural Inference for Object Recognition

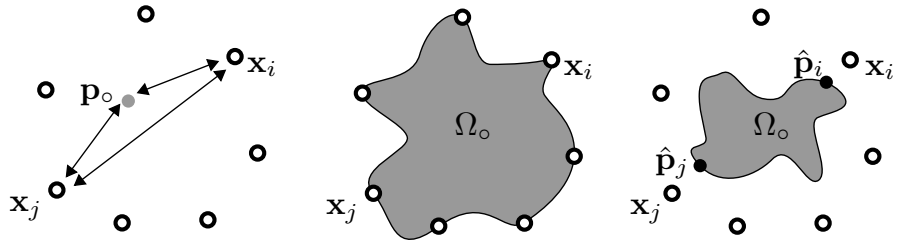
After the candidate detection step, the feature space Ω can be thought to consist of two separate parts. One part consists of all images showing a protruding element and the other part consists of images without any protruding element. It is assumed that the latter part does not contain any images with polyps. On the other hand, not all images with a protruding element do contain polyps as there might be other causes of protrusions, like fecal remains, the ileocecal valve (between the large and small intestine) and small fluctuations of the colon wall. To summarize, the feature space now consists of three different parts:

1. a set $\Omega_\circ \subset \Omega$ containing all images without a polyp (the surrounding class),
2. a set $\Omega_f \subset (\Omega \backslash \Omega_s)$, spanned by all volumes containing a protrusion which is not a polyp (false detection class), and
3. a set $\Omega_t = \Omega \backslash (\Omega_s \cup \Omega_f)$, spanned by all volumes containing a polyp (true detection class).

Section 5.3.1 will describe how the dissimilarities are defined with respect to objects of which the appearance is highly context-dependent. Section 5.3.2 will discuss how the classes are represented.

5.3.1 Dissimilarity Measure

A simple example of dissimilarities between objects is illustrated in Fig. 5.3(a). Two objects $\mathbf{x}_i$ and $\mathbf{x}_j$ both have a certain dissimilarity with a certain prototype $\mathbf{p}_\circ$, respectively $d_{i\circ}$ and $d_{j\circ}$. If the Euclidean distance is used as the measure, the triangle inequality holds in this case. An example of such a situation can be seen in Fig. 5.4(a). For each object on the table the dissimilarity between the image with the object (Fig. 5.4(a)) and without the object (Fig. 5.4(c)) can be defined as the image of the object itself (Fig. 5.4(b)). Considering each object



$d_{i\circ} > 0,\ d_{j\circ} > 0$
$d_{ij} \leq d_{i\circ} + d_{j\circ}$
(a)

$d_{i\circ} = 0,\ d_{j\circ} = 0$
$d_{ij} \not\leq d_{i\circ} + d_{j\circ} = 0$
(b)

$d_{i\circ} > 0,\ d_{j\circ} > 0$
$d_{ij} \not\leq d_{i\circ} + d_{j\circ}$
(c)

Figure 5.3: (a) Feature space of two objects having the same surrounding, which means that the surrounding (the table in Fig. 5.4(a)) reduces to a single point $\mathbf{p}_\circ$. (b) When considering spatial distances between the objects, the surrounding $\mathbf{p}_\circ$ transforms into a blob and all distances between objects within the blob are zero. (c) If the surroundings of each object are different and we are considering the dissimilarity in appearance, the feature space is a combination of (a) and (b).

separately, the surrounding of all objects is exactly the same (i.e. the table is the same). Now the table can thus be defined as the prototype $\mathbf{p}_\circ$ and different objects all have a different dissimilarity to the table.

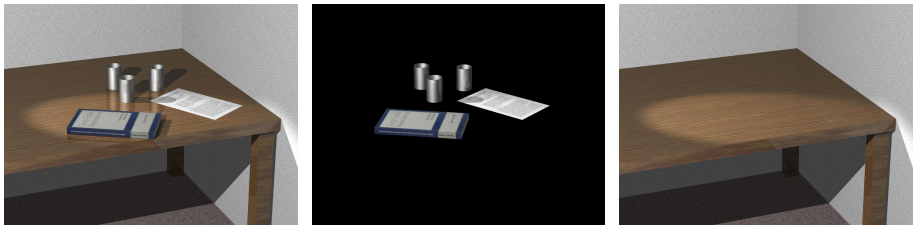
If the dissimilarity between two objects is defined as the spatial distance between the objects, the dissimilarity measure violates the triangle inequality and the measure becomes non-metric [101]. For example, all objects in Fig. 5.4(a) have zero distance to the table, but the distance between two objects might be larger than zero. This is illustrated in Fig. 5.3(b). The prototype $\mathbf{p}_\circ$ is no longer a single point, but is transformed into a blob $\Omega_\circ$ representing all objects with zero distance to the table.

Let us now consider the problem of object detection in structured surroundings. First, as in the first example given above, the dissimilarity of an object to its surrounding is defined by the object itself. Second, although the surroundings may differ significantly from each other, it is known that none of the surroundings contain an object of interest. Thus, as in the second example, the distances between all surroundings can be made zero and we obtain the same blob representation of $\Omega_\circ$. The distance of an object to the surrounding class can now be defined as a minimization over all prototypes from the set of surroundings $\Omega_\circ$

$$d_{i\circ} \triangleq d(\mathbf{x}_i, \Omega_\circ) = \min_k d(\mathbf{x}_i, \mathbf{p}_k) \text{ with } \mathbf{p}_k \in \Omega_\circ.$$

This problem is thus a combination of the two examples and this leads to the

feature space shown in Fig. 5.3(c). Both objects $\mathbf{x}_i$ and $\mathbf{x}_j$ have a related prototype, respectively $\hat{\mathbf{p}}_i$ and $\hat{\mathbf{p}}_j$, to which the dissimilarity is the smallest. Again, the triangle inequality does no longer hold: two images that look very different may both be very close to the surrounding class. On the other hand, two objects that are very similar do have similar dissimilarity to the surrounding class and the compactness hypothesis still holds in the space spanned by the dissimilarities. Moreover, the dissimilarity of an object to its surrounding still contains all information for successful classification of the object, which is easily seen by looking at Fig. 5.4(b).



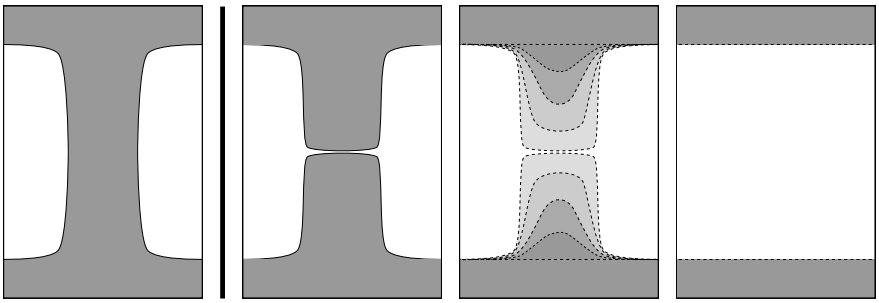
(a) (b) (c)

Figure 5.4: (a) Objects in their surroundings. (b) Objects without their surroundings. All information about the objects is retained, so the objects can still be classified correctly. (c) The estimated surrounding without the objects.

5.3.2 Class Representation

The prototypes $\hat{\mathbf{p}}_i$ and $\hat{\mathbf{p}}_j$ thus represent the surrounding class, but these prototypes are not available a priori. We know that they must be part of the boundary of $\Omega_\circ$ and that the boundary of $\Omega_\circ$ is the set of objects that divide the feature space of images with protrusions and those without protrusions. Consequently, for each object we can derive its related prototype of the surrounding class by iteratively solving the PDE in (5.1). That is, $\Omega_s \triangleq \delta\Omega_\circ \cap (\delta\Omega_t \cup \delta\Omega_f)$ are all solutions of (5.1) and the dissimilarity of an object to its surroundings is the 'cost' of the deformation guided by (5.1). Furthermore, the prototypes of the surroundings class can now be sampled almost infinitely, i.e. a prototype can be derived if it is needed.

A few characteristics of our approach to object detection are illustrated in Fig. 5.5. At the first glance, objects $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively shown in Figs. 5.5(a) and (b), seem to be similar (i.e. close together in the feature space spanned by all pixel values), but the structures present in these images differ significantly. This difference in structure is revealed when the images are being transformed by the PDE (5.1). Object $\mathbf{x}_1$ does not have any protruding elements and can thus be considered as an element of $\Omega_\circ$, whereas object $\mathbf{x}_2$ exhibits two large protrusions: one pointing down from the top, the other pointing up from the bottom. Fig. 5.5(c) shows several intermediate steps in the deformation of this



(a) $\mathbf{x}_1 \in \Omega_\circ$ (b) $\mathbf{x}_2$ (c) Deformation (d) $\hat{\mathbf{p}}_2 \in \Omega_s$

Figure 5.5: (a–b) Two similar images having different structure lead to different responses to deformation by the PDE in (5.1). The object $\mathbf{x}_1$ is a solution itself, whereas $\mathbf{x}_2$ will be deformed into $\hat{\mathbf{p}}_2$. A number of structures that might occur during the deformation process are shown in (c).

object and Fig. 5.5(d) shows the final solution. This illustrates that by defining a suitable deformation, a specific structure can be measured in an image. Using the deformation defined by the PDE in (5.1), all intermediate images are also valid images with protrusions with decreasing protrudedness. Furthermore, all intermediate objects shown in Fig. 5.5(c) have the same solution. Thus, different objects can have the same solution and relate to the same prototype.

Let us propose to use a morphological closing operation as the deformation, then one might conclude that images $\mathbf{x}_1$ and $\mathbf{x}_2$ are very similar. In that case we might conclude that image $\mathbf{x}_2$ does not really have the structure of two large polyps, as we concluded before, but might have the same structure as in $\mathbf{x}_1$, but altered by an imaging artifact. Using different deformations can thus lead to a better understanding of the local structure. In that case, one could represent each class by a deformation instead of a set of prototypes [98]. Especially for problems involving objects in highly structured surroundings, it might be advantageous to define different deformations in order to infer from structure.

An example of an alternative deformation was already given by the PDE in (5.2). This deformation creates a new prototype of the polyp class given an image and the 'cost' of deformation could thus be used in classification. Combining both methods thus gives for each object a dissimilarity to both classes. However, this deformation was proposed as a preprocessing step for current CAD systems. By doing so, the dissimilarity was not explicitly used in the candidate detection or classification step.

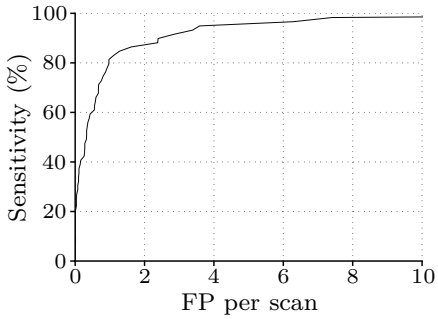


Figure 5.6: FROC curve for the detection of polyps $\geq$ 6 mm.

5.4 Classification

Now we have a very well sampled class of the healthy (normal) images, which do not contain any protrusions. Any deviations from this class indicates unhealthy protrusions. This can be considered as a typical one-class classification problem in which the dissimilarity between the object $\mathbf{x}$ and the prototype $\mathbf{p}$ indicates the probability of belonging to the polyp class. The last step in the design of the polyp detection system is to define a dissimilarity measure that quantifies the introduced deformation, such that it can be used to successfully distinguish the non-polyps from the polyps. As said before, the difference image still contains all information, and thus there is still no class overlap.

Until now, features are computed from this difference image to quantify the 'cost' of deformation. Three features are used for classification: the length of the two principal axes (perpendicular to the polyp axis) [79] of the segmentation of the candidate, and the maximum intensity change. A linear logistic classifier is used for classification. Classification based on the three features obtained from the difference image leads to results comparable to other studies [9, 30, 29]. Fig. 5.6 shows an FROC curve of the CAD system for 59 polyps larger than 6 mm (smaller polyps are clinically irrelevant) annotated in 86 patients (172 scans). Results of the current polyp detection systems are also presented elsewhere [10, 57, 73].

5.5 Conclusion

We have presented an automated polyp detection system based on structural inference. By transforming the image using a structure-driven partial differential equation, knowledge is inferred from the structure in the data. Although no prototypes are available a priori, a prototype of the 'healthy' surrounding class can be obtained for each candidate object. The dissimilarity with the healthy class is obtained by means of a difference image between the image before and

after the transformation. This dissimilarity is used for classification of the object as either a polyp or as healthy tissue. Subsequent classification is based on three features derived from the difference image. The current implementation basically acts like a one-class classification system: the system measures the dissimilarity to a well sampled class of volumes showing only normal (healthy) tissue. The class is well sampled in the sense that for each candidate object, we can derive a healthy counterpart, which acts as a prototype.

Images that are very similar might not always have the same structure. In the case of structured data, it is this structure that is most important. It was shown that the transformation guided by the PDE in (5.1) is capable of retrieving structure from data. Furthermore, if two objects are very similar, but situated in a different surrounding, the images might look very different. However, after iteratively solving the PDE, the resulting difference images of the two objects are also similar. The feature space spanned by the dissimilarities thus complies with the compactness hypothesis.

Until now, only information is used about the dissimilarity to the 'healthy' class. The work of Konukoglu et al. [75] offers the possibility of deriving a prototype for the polyp class given a candidate object just as we derived prototypes for the non-polyp class. A promising solution might be a combination of both techniques; each candidate object is then characterized by its dissimilarity to a non-polyp prototype and by its dissimilarity to a polyp prototype. Both prototypes are created on-the-fly and are situated in the same surrounding as the candidate. In fact, two classes have been defined and each class is characterized by its own deformation.

In the future, the patient preparation is further reduced to improve patient compliance [102]. This will lead to data with increased amount of fecal remains in the colon and this will complicate both the task of automated polyp detection as well as electronic cleansing of the colon [16, 24]. The presented approach to infer from structure can also contribute to the image processing of such data, especially if the structure within the colon becomes increasingly complicated.

Chapter 6

Thin Layer Tissue Classification for Electronic Cleansing

CT colonography (CTC) is a rapidly evolving technique to screen for colorectal polyps. Fecal residue may occlude or, reversely, mimic polyps. Electronic cleansing aims at removing contrast-enhanced fecal residue from the image. However, thin layers of soft tissue (the colon wall or a fold) or residue are easily misclassified by current electronic cleansing methods, thereby causing holes in the colon wall or other artefacts that hamper visualization and automated detection. We present a thin layer model to detect and characterize such layers to support electronic cleansing. It is demonstrated that the model sustains robust estimation of the location and thickness of such a layer. Such thicknesses of thin layers were measured in real data sets. A lower bound on the thickness of such layers exists and was found to be 1.0 mm for our data.

6.1 Introduction

Adenomatous colorectal polyps are important precursors to colorectal carcinoma. Computed tomography colonography (CTC) is a rapidly evolving technique that is advocated to screen for such polyps. The preparation of patients who undergo CTC often consists of cathartic cleansing and the oral administration of a contrast agent for fecal tagging. Subsequently, the patient undergoes CT scanning prior to which the colon is distended by insufflation with room air. Several 'electronic cleansing' algorithms were introduced to automatically segment the colon surface from this data and to facilitate a 3D endoluminal view into the colon [16, 103, 33]. All these methods focused on data from patients who had undergone extensive cathartic cleansing.

Current research aims at increasing the patient's compliance by minimizing the patient preparation. However, omitting the cathartic cleansing complicates the segmentation procedure. Due to the partial volume effect, thin layers of soft tissue surrounded by residue on the one side and air on the other are hard to segment. Likewise, thin layers of contrast-enhanced fecal residue adhering to the colon surface cause a similar problem. The problem of the detection of soft tissue in the presence of fecal residue was studied before in [104], but none of the electronic cleansing algorithms have addressed the problem of thin layers explicitly. Nevertheless, thin layer characterization was studied in other CT applications [105, 106, 107, 108]. All these approaches involved a model-fitting procedure to find the thickness and position of the layer.

This paper describes a novel segmentation method that estimates these features directly from the observed data. Moreover, the potential benefit for CTC will be shown.

6.2 Cleansing of CTC Data

The data may be asserted to comprise three types of materials: air (A), soft tissue (T) and tagged residue (R). The mean intensities of these materials will be denoted by I_A, I_T and I_R respectively. Effectively, electronic cleansing estimates the volume fractions of these materials in each voxel [16]. Subsequently, the colon surface may be segmented by the isosurface of 50% soft tissue.

A previously proposed approach to cleansing of CTC data was based on tissue classification using the measured intensity I and the gradient magnitude I_{w} in the gradient direction [16]. These features form a rotation and scale-invariant feature space as shown in Fig. 6.1a (in which σ_{w} represents the effective scale of measurement and the scale of the point spread function (PSF) in the gradient direction). Each type of edge (material transition) is represented as a different arch-shaped cloud in this feature space [109]. These arches are described by the arch-model [16]. For instance, the A-R transitions appear as a large arch from I_A to I_R and can be modelled by:

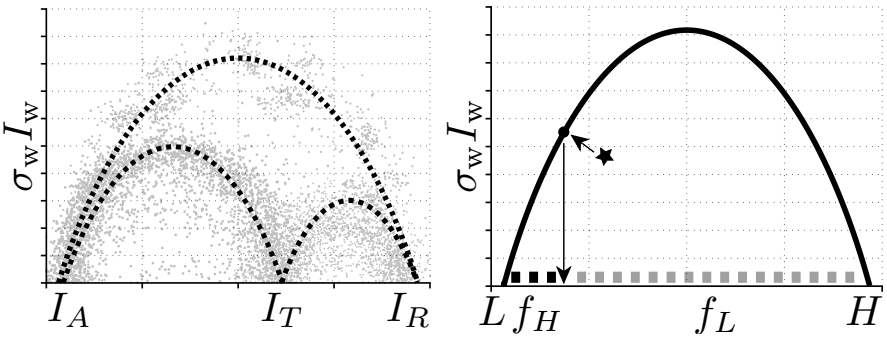


Figure 6.1: (a) Scatter plot of the intensity and scale-invariant gradient magnitude of voxels. The arches represent the three types of edges. (b) The material fractions are determined from the projection of a voxel onto the arch.

$$\frac{I_w}{I_R - I_A} = \frac{\sigma_w^{-1}}{\sqrt{2\pi}} \exp\left(-\operatorname{erf}^{-1}\left(\frac{I - I_M}{I_R - I_M} \right)^2 \right) \tag{6.1}$$

in which $I_M = (I_A + I_R)/2$.

To estimate the material constituency of a voxel, a number of voxels up and down the gradient direction are sampled. The intensity and gradient magnitude of all these voxels form a trace in $\langle I, |\nabla I| \rangle$-space. Fitting the arch model to this measured trace yields the low L and high H material intensities, which are used to classify the transition as one of the three types. Projecting the intensity and the gradient magnitude onto the corresponding arch-model yields the material fractions f_L and f_H as indicated in Fig. 6.1b.

It has been shown that this algorithm produces good results for all 'pure' two-material (2M) mixture transitions [16]. However, specifically a layer of soft tissue between air and residue (an A-T-R transition) is easily misclassified as an A-R transition if the layer of soft tissue is thin. We seek a model that describes such a thin layer geometry for segmentation with improved accuracy.

6.2.1 Thin Layer Model

The novelty of this paper is in the derivation of a mathematical description of a thin layer of soft tissue between air and contrast-enhanced fecal residue. Such a layer in real CTC data is shown in Fig. 6.2a. The gradient magnitude $|\nabla I|$ and the intensity I in the phantom model (Fig. 6.2b) are computed assuming a Gaussian PSF (as done in [16]) with an effective scale σ_w=1 (notice that consequently $\sigma_w I_w$ can simply be written as $|\nabla I|$). The traces in $\langle I, |\nabla I| \rangle$-space for different thicknesses D=$\sigma \ldots 4\sigma$ of the thin layer are shown in Fig. 6.2c. If

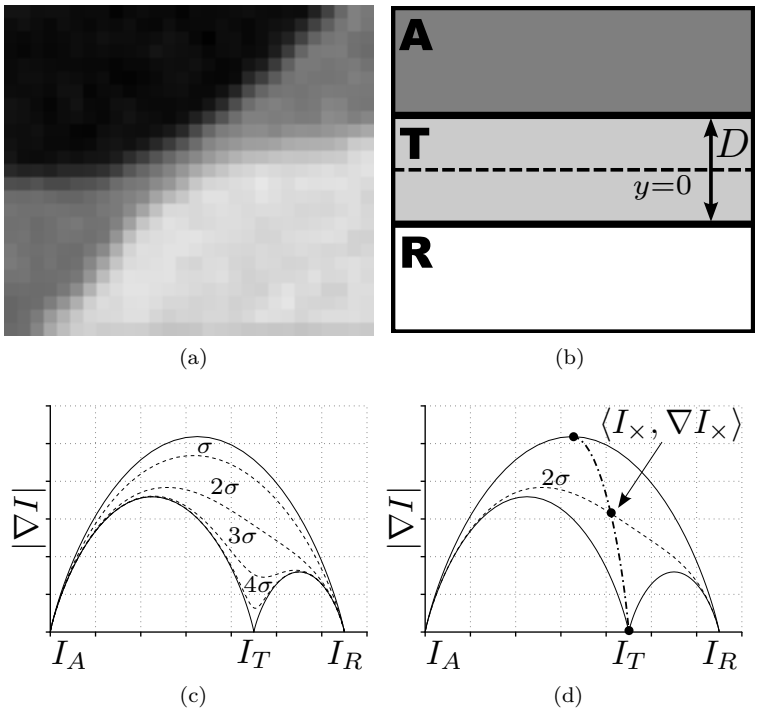


Figure 6.2: (a–b) A thin layer of soft tissue (gray) in CTC data and the phantom model. (c) Profiles for different thicknesses of the layer (dashed) and (d) the model for the middle of a thin layer with varying thickness (dash-dotted).

the thickness $D < \sigma$ the profile resembles the profile of an A-R transition, whereas if the thickness $D > 4\sigma$ the profile approximates that of two separate transitions. Several features for these types of profiles will now be derived to distinguish a thin layer A-T-R transition from an A-R transition.

6.2.2 Description at $y = 0$

Using the assumption of a Gaussian PSF, the intensity at the middle of the layer (y=0) is given by

$$I(y{=}0; D, \sigma) = I_A \int_{-\infty}^{-D/2} g(y'; \sigma) dy' + I_T \int_{-D/2}^{D/2} g(y'; \sigma) dy' + I_R \int_{D/2}^{\infty} g(y'; \sigma) dy'$$

with $g(\cdot;\sigma)$ the Gaussian kernel. Rewriting this leads to an expression for the intensity given a thickness D:

$$\frac{I(y{=}0;D,\sigma)-I_M}{I_T-I_M}=\operatorname{erf}\left(\frac{D/\sigma}{2\sqrt{2}}\right). \tag{6.2}$$

The gradient of the intensity at $y{=}0$ is given by

$$\begin{aligned}\nabla I(y{=}0;D,\sigma){=}&\mathrm{g}(\frac{D}{2};\sigma)\left[(I_R-I_T)+(I_T-I_A)\right]\\ &=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(D/2)^2}{2\sigma^2}\right)(I_R-I_A)\,. \end{aligned}\tag{6.3}$$

The relation between I and ∇I at $y{=}\,0$ is obtained from eqs. 6.2 and 6.3 by eliminating D:

$$\frac{\nabla I(y{=}0;D,\sigma)}{I_R-I_A}= \frac{\sigma^{-1}}{\sqrt{2\pi}}\exp\left(-\operatorname{erf}^{-1}\left(\frac{I(y{=}0;D,\sigma)-I_M}{I_T-I_M}\right)^2\right). \tag{6.4}$$

Eq. 6.4 has the same form as eq. 6.1, which describes the arches of 2M-transitions. The only difference is that the argument of the $\operatorname{erf}^{-1}(\cdot)$ function is scaled by (I_T-I_M) rather than (I_R-I_M). In other words, the model delivered by eq. 6.4 is essentially obtained by scaling the intensity of the arch of a two-material A-R transition by $(I_T-I_M)\,/\,(I_R-I_M)$. Fig. 6.2d visualizes this relation between the gradient and intensity (dash-dotted curve). The dots along the curve indicate the feature values at $y{=}0$ for the thicknesses $D{=}0, 2\sigma$ and ∞.

6.3 Geometric Construction

The model for the middle of the layer of soft tissue can also be obtained by adding parts of the arches of the A-T and T-R transitions. Fig. 6.3 illustrates this construction which consists solely of combining curves that were already presented in Figs. 6.2(c–d).

First, one may realize that the difference between the model of an A-R transition and the model of an A-T transition is the model of a T-R transition. This can be seen in Fig. 6.3(a). Starting from I_A and traversing the large arch ends in the same point as traversing the two smaller arches one after the other. Moreover, traversing from I_A to halfway an A-R transition is the same as traversing to halfway an A-T transition and subsequently traversing half of the T-R transition. This is shown geometrically by shifting half of the T-R arch onto the A-T arch (arrow 1). The same can be done for half of the A-T transition, which can be shifted onto the T-R transition (arrow 2).

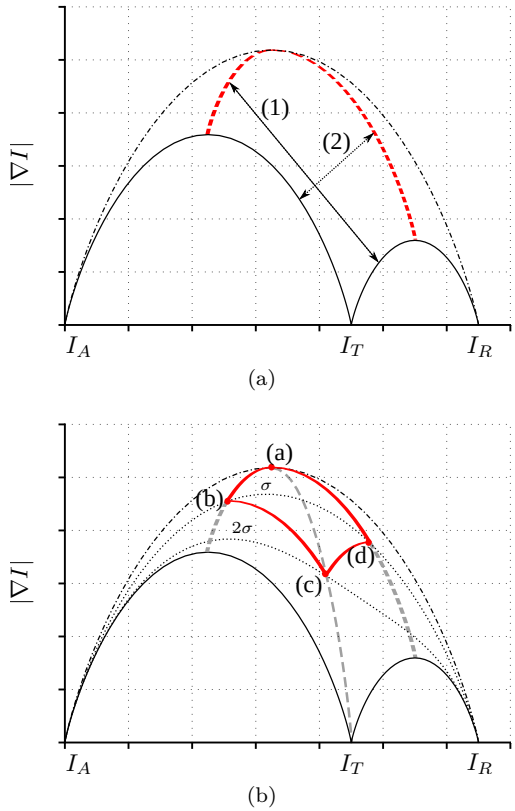


Figure 6.3: (a) Shifting parts of the arches describing two-material transitions. (b) Additive behavior of the arch representation. The model delivered by eq. 6.4 is constructed by adding parts of the T–R transition model (a–b and d–c) and the A–T transition model (b–c and a–d).

The two arches that were shifted previously touch each other halfway an A-R transition. Traversing from this point along the shifted arches is equivalent to replacing either air or residue by soft tissue starting from the middle (i.e. a layer of soft tissue is grown either extending into the air or into the residue). This is shown in Fig. 6.3(b). Let us start halfway the A-R transition. After replacing 1σ of residue with soft tissue (curve a-b) and subsequently replacing 1σ of air with soft tissue (curve b-c), we will end up in point c. This point is now in the middle of a 2σ A-T-R transition. This is confirmed in Fig. 6.3(b) as point c is on the intersection of the model of a 2σ A-T-R transition and the model describing the middle of the layer as derived in the previous section.

To conclude, the profile of the middle of the layer of soft tissue can not only be constructed by scaling half of an A-R transition, but also by adding parts of

the two other two-material transitions.

6.4 Feature Retrieval

The objective now is to derive an expression for the thickness and the location of the thin layer. This is achieved by determining the intersection of a measured profile with the curve described by eq. 6.4. Let us represent the crossing by $\langle I_\times, |\nabla I_\times| \rangle$ (see Fig. 6.2d). Backprojecting this point into the image renders the location where y=0. The thickness of the layer may be obtained by substituting $I_\times$ into eq. 6.2. Solving for D gives

$$D = \sigma 2\sqrt{2} \left\{ \mathrm{erf}^{-1} \left(\frac{I_\times - I_M}{I_T - I_M} \right) \right\}. \tag{6.5}$$

Alternatively, D may be obtained as a function of $|\nabla I_\times|$ by inverting eq. 6.3:

$$D = \sigma 2\sqrt{2} \sqrt{\log \left(\frac{I_R - I_A}{\sigma\sqrt{2\pi}|\nabla I_\times|} \right)}. \tag{6.6}$$

The two approaches will give exactly the same solution since $\langle I_\times, |\nabla I_\times| \rangle$ lies on the curve given by eq. 6.4. It may be noticed that the relation between $|\nabla I|$ and I at y=0 will lead to the best resolution in determining D, because the maximal separation of the thickness profiles is obtained at y=0 (see Fig. 6.2).

6.5 Experiments

The electronic cleansing method described previously [16] was applied to four CT colonography data sets. All A-R transitions detected (as described in Section 6.2) were examined in more detail. While doing so, only those transitions were retained that were at least 5 voxels away from another type of edge. This was done to reject potential three-material 'junctions' [24, 11] not complying to the thin layer model. Furthermore, only objects were included consisting of more than 10 voxels.

From each object-voxel a trace was generated in the positive and negative direction of the gradient vector until the gradient magnitude was smaller than some threshold. The intensities at the lower side (L) and the upper side (H) were at the basis to scale the encountered intensities into the $[0, 1]$ range. Moreover, all traces were made invariant to the effective scale of the measurement by multiplying the sampled gradient magnitude with σ_w. Finally, a third-order polynomial was fitted to the transformed samples to represent the profile around 'y=0'. The crossing point of this polynomial with the function of eq. 6.4 was determined, which, in turn, yielded the location and thickness of the thin layer under consideration (see Section 6.2.1).

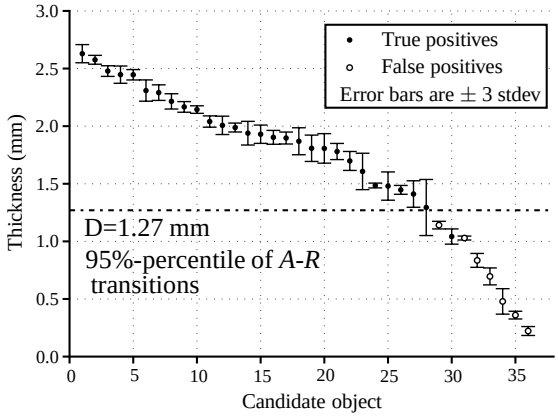


Figure 6.4: Mean and stdev of estimated thicknesses of the candidate objects.

6.6 Results

In total 411 candidate objects were generated of which 80 had a finite measured thickness. From these, 53 candidates resided in the large intestine, whereas the other 27 were associated with the small intestine. The latter fall outside the scope of CTC. 17 candidates from the 53 were found to be heterogeneously composed (by visual inspection), i.e. they comprised either a mixture of different geometries, like a fold separating air and residue and a pure 2M transition (see below for further treatment of these objects), or they consisted of an incompatible geometry like a thin layer of contrast. Thus, 36 candidate objects remained.

The mean measured thickness and corresponding standard deviation for the 36 candidate objects are shown in Fig. 6.4. The standard deviation emanates from averaging over the traces generated from all the object-voxels. From the same data sets, the seven largest objects consisting of traces of 'pure' A-R material transitions were analyzed. These transitions yielded a measurement of only 0.03 mm and a standard deviation ranging from 0.8–1.1 mm. The 95% percentile-level of the thickness distribution was at 1.27 mm. Accordingly, an object is considered an A-T-R transition if the mean estimated thickness of the soft tissue layer is significantly larger than 1.27 mm. Applying this procedure to the data in Fig. 6.4 yields a detection sensitivity of 93% (27/29), and no false positive detections. No previously discarded candidates were considered to contain a thin layer of soft tissue (by visual inspection).

The 19 objects consisting of a mixture of geometries yielded large standard deviations with respect to thickness measurement. These geometries were separated by an EM-clustering algorithm based on the measured thickness for each profile and the spatial relation between the voxels. Fig. 6.5 shows the results

for such an object (a fold surrounded by air and residue, both are not shown). Fig. 6.5a shows a cross-section of the object. For each voxel, represented by a dot, a trace is measured and all measurements constituting the traces are shown in Fig. 6.5b. The result of clustering the traces is indicated by the color. Notice that the grouped voxels indeed form clusters in x-y space and that the traces are well separated: one cluster of traces showing normal two-material A-R transitions (gray) and another cluster showing thin layers of soft tissue with varying thickness (black). The calculated thickness was close to zero for the A-R transition and 1.8 mm for the A-T-R transition. The latter value may be negatively biased, though, due to the profiles near the transition area.

6.7 Conclusions

We introduced a novel method for detection and characterization of thin layers and applied it to CT colonography data. The technique may preclude erroneously removing thin layers of soft tissue (sandwiched between air and contrast-enhanced residue) and disturbing the topology of an object. A mathematical function was derived relating the intensity to the gradient magnitude in the middle of a thin layer as a function of the thickness. Practically, a layer's thickness was obtained by intersecting a measured curve in $\langle I, \sigma_{\mathrm{w}} I_{\mathrm{w}} \rangle$-space with a function describing the middle of the layer. We demonstrated the usefulness of the method by first detection of thin layers of soft tissue and subsequent estimation of the thickness of these layers surrounded by residual matter and air. Observe that the method could equally well be applied to thin layers of contrast material stuck to the colon surface and surrounded by air. In future research, we will include the algorithm in an electronic cleansing algorithm for CT colonography.

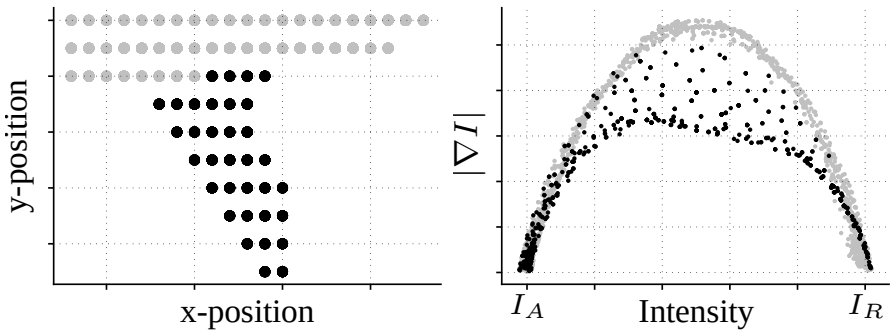


Figure 6.5: Heterogeneous candidate object. (a) Clustering shows two distinct regions. (b) Points on traces for both clusters.

Chapter 7

Electronic Cleansing for 24-H Limited Bowel Preparation CT Colonography Using Principal Curvature Flow

CT colonography (CTC) is one of the recommended methods for colorectal cancer screening. The subject's preparation is one of the most burdensome aspects of CTC with a cathartic bowel preparation. Tagging of the bowel content with an oral contrast medium facilitates CTC with limited bowel preparation. Unfortunately, such preparations adversely affect the 3D image quality. Thus far, data acquired after very limited bowel preparation were evaluated with a 2D reading strategy only. Existing cleansing algorithms do not work sufficiently well to allow a primary 3D reading strategy. We developed an electronic cleansing algorithm, aimed to realize optimal 3D image quality for low-dose CTC with 24-h limited bowel preparation. The method employs a principal curvature flow algorithm to remove heterogeneities within poorly tagged fecal residue. In addition, a pattern recognition based approach is used to prevent polyp-like protrusions on the colon surface from being removed by the method. Two experts independently evaluated 40 CT colonography cases by means of a primary 2D approach without involvement of electronic cleansing as well as by a primary 3D method after electronic cleansing.

The data contained four variations of 24-hr limited bowel preparation and was based on a low radiation dose scanning protocol. The sensitivity for lesions ≥ 6 mm was significantly higher for the primary 3D reading strategy (84 %) than for the primary 2D reading strategy (68 %) ($p = 0.031$). The reading time was increased from 5:39 min (2D) to 7:09 min (3D) ($p = 0.005$); the readers' confidence was reduced from 2.3 (2D) to 2.1 (3D) ($p = 0.013$) on a 3-point Likert scale. Polyp conspicuity for cleansed submerged lesions was similar to not submerged lesions ($p = 0.06$). To our knowledge this study is the first to describe and clinically validate an electronic cleansing algorithm that facilitates low-dose CTC with 24-h limited bowel preparation.

7.1 Introduction

Computed tomography colonography (CTC) is a structural radiological examination of the colorectum and is widely studied for use in colorectal cancer screening [110]. An important issue for large scale application is the adherence, which is closely related to the perceived burden of the employed screening technique [111]. The subject's preparation is one of the most burdensome aspects of CT colonography with a cathartic bowel preparation [12]. Although such a cathartic bowel preparation ensures optimal image quality, it also leads to excessive diarrhea and discomfort. Tagging of the bowel content with oral iodine or barium contrast facilitates CTC with non-cathartic bowel preparation. Recently, several studies have shown that the diagnostic accuracy for polyps ≥ 6 mm remains high while using a 24-h limited bowel preparation (i.e., least burdensome type of non-cathartic preparations) [17, 18]. In fact, a limited bowel preparation significantly improves the acceptance and therefore likely the screening adherence [19, 17, 20]. Liedenbaum et al. showed that a 24-h limited iodine-based bowel preparation yields a significantly better subject's acceptance and less burden compared with a 48-h preparation [21].

Unfortunately, such preparations can adversely affect the 3D image quality. Particularly, untagged stool can cause artifacts like incomplete cleansing or pseudo-soft tissue structures [22, 23]. These artifacts limit a primary 3D reading and hinder 3D problem solving after a primary 2D reading. Still, accurate electronic cleansing can result in shorter reading times in a primary 3D reading strategy and to a higher confidence and less reader effort in a primary 2D reading strategy [24]. Juchems et al. [112] studied reader performance with the use of electronic cleansing and found a significant improvement in polyp sensitivity using 3D reading with electronic cleansing versus 3D reading without electronic cleansing. Importantly, less experienced readers achieve a higher sensitivity with a 3D reading strategy as compared to a 2D reading strategy [113]. Recent guidelines, summarizing the evidence by experts in the field, emphasize the need for both 2D and 3D visualization [114, 115].

Apart from the burden associated with the bowel preparation, the acceptance of CT colonography as a screening technique is also influenced by the radiation exposure. The radiation burden should be as low as possible to ensure a high benefit-risk ratio. At the same time, a low-dose scanning protocol leads to increased image noise which complicates electronic cleansing and significantly affects polyp detection [116].

We developed a new electronic cleansing algorithm, aimed to realize optimal 3D image quality for low-dose CTC with 24-h limited bowel preparation. We hypothesize that electronic cleansing does not lead to a degradation of a polyp's conspicuity in a 3D viewing mode and, even stronger, that it enables a primary 3D reading strategy. To our knowledge, no earlier study described an electronic cleansing algorithm for low-dose CTC with 24-h limited bowel preparation.

7.1.1 Related Work

Much of the previous technical work on electronic cleansing has been validated on data obtained with extensive patient preparation. The following summary was largely adapted from [117].

Initially, Lakare et al. [118] addressed the cleansing problem by exploiting the unique local signature caused by partial volume voxels bordering on the fluid mask. The intensity profile is considered a unique property of each type of material transition. Typical edge profiles that are present between materials, e.g., air and tagged material, are identified by rigorously "exploring" some 3D CT data sets in a separate learning phase prior to the actual cleansing. During cleansing, for each edge voxel, the profile is selected from the learning set that fits best to the encountered intensity profiles. A transfer function is defined for each such profile in order to remove the partial volume problem during rendering. In later work, Lakare et al. [119] created 23-D feature vectors of local data values that are reduced to five dimensions by principal component analysis. Clustering takes place in this low-dimensional space using a vector similarity measure. A threshold on the average intensity of each class is used to classify voxels to be tagged residue.

Zalis et al. [120, 121] constructed a binary subtraction mask to segment the bowel content and addressed the partial volume problem using a colon-surface reconstruction routine. Data values represent a distance measure to the subtraction mask. Wang et al. [103] presented an improved electronic colon cleansing method based on a partial volume image segmentation framework, which is based on the well established statistical expectation-maximization algorithm.

Franaszek et al. [122] developed a segmentation procedure which represents individual air- and fluid-filled regions by a graph that enables identification and prevention of undesired leakage through the colon wall. The proposed hybrid algorithm uses modified region growing, fuzzy connectedness, and level set segmentation. Wang et al. [123] also investigated a maximum a posteriori expectation–maximization image segmentation algorithm which simultaneously

estimates tissue mixture percentages within each image voxel and statistical model parameters for the tissue distribution.

Cai et al. [22, 23] developed an electronic cleansing method, called structure analysis cleansing. A structure enhancement function and a local roughness measure are integrated into the speed function of a level set method for delineating the tagged fecal material.

To our opinion, the variety of presently proposed algorithms reflect that a perfect solution has not been found yet. Accordingly, incomplete processing is still reported to leave artifacts [24]. A specifically noticeable problem is posed by the distracting bumps emanating from locations where air, soft tissue and tagged material meet. Lately, an electronic cleansing algorithm was proposed aiming to improve the accuracy at such three material junctions [16, 24, 117]. The method is automated and adapts to patient specific conditions, such as the local variation of the tagged material density. The algorithm assumes that the measured CT value arises due to a combination of three materials: air, tagged material and soft tissue. It estimates the percentages of these materials in each voxel. Subsequently, the tagged material fraction is simply 'replaced' by air, to arrive at a new, cleansed CT value. Unfortunately, this method still assumes that the three materials constituting the junction have a homogeneous composition which makes the method not adequate to deal with the limited bowel preparation data.

Recent work by Cai et al. describes an electronic cleansing technique based on a mosaic decomposition method for use with a limited bowel preparation [22, 23, 124]. This method solved the artifacts often associated with a limited bowel preparation. It was tested on cases that underwent a 48-h bowel preparation consisting of a low-fiber, low-residue diet and oral administration of Omnipaque with a total ingested amount of 75 ml of iodine (300 mg I/ml concentration). No clinical evaluation of this cleansing method was performed yet.

Visualization techniques can be found in [125, 126, 8].

7.1.2 Objective

The objective of this work is to enable 3D visualization of CT colonography data for use with a 24-h limited bowel preparation and a reduced radiation dose. This is, to the best of our knowledge, currently not feasible. Examples of typical artifacts prevalent in such data are shown in Fig. 7.1.

The properties that an electronic cleansing algorithms must possess to allow a limited patient preparation are:

- ability to cope with heterogeneous bowel content;
- reconstruction of a smooth colon surface that does not distract the radiologist;
- preservation of polyps in badly tagged regions.

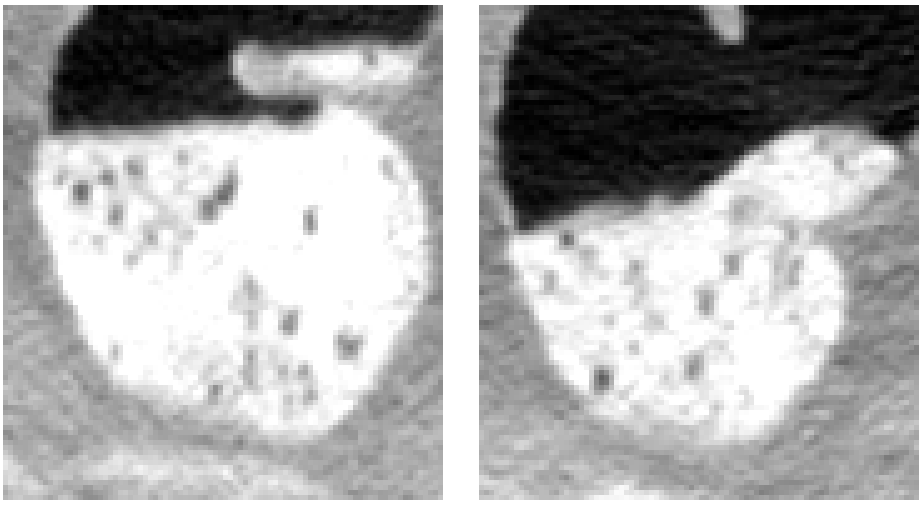


Figure 7.1: Examples of typical cases emanating from a limited bowel preparation and a low radiation dose.

Heterogeneous mixing of materials as well as highly curved and irregular material interfaces as depicted in Fig. 7.1 complicates a simple solution to electronic cleansing. In order to visualize the colon wall such that it accurately represents the shape of the colon wall and that it does not show distracting artifacts to the radiologist, a sub-voxel precision is required. Thereby, the main challenge is to retain all polyps in the data, even when polyps are submerged in a heterogeneously tagged fluid. To secure the preservation of polyps, we will make use of the extensive knowledge about the appearance of colonic polyps that was acquired in developing techniques for computer aided detection of polyps [57, 10].

In this paper we present a plug-in pre-processing method to the electronic cleansing method by Serlie et al. [16, 24, 117]. The pre-processing acts as a preconditioning step which fills the inhomogeneities in poorly tagged fecal material to meet the input requirement of Serlie's electronic cleansing method. Several novelties are introduced in order to do so. First, the removal of heterogeneities in poorly tagged fecal matter by a principal curvature flow algorithm. Second, a robust reconstruction of the colon surface, including all polyp candidates, using a logistic classifier. This corrects the negative side effects of the first step such as erroneously removal of polyp-like protrusions on the bowel wall. Third, we demonstrate the effects of our technique on the sensitivity, reading time, and lesion conspicuity in a clinical evaluation on data acquired with a reduced radiation dose after a 24-h limited bowel preparation.

7.2 Materials and methods

7.2.1 CT Colonography Data

The new electronic cleansing method will be evaluated on four subject groups, each with a different variation of limited bowel preparation scheme and a low radiation dose scanning protocol (Table 7.1, groups 'A'–'D'). Fifteen CT colonography examinations were randomly selected from one former study (group 'A') [17], supplemented by all (three times fifteen) examinations from another recent study (groups 'B'–'D') [21]. Each subject was scanned in prone and supine position. CT reconstruction of all data sets were done using standard filtered back projection. All participants underwent a low-fiber diet starting the day before the CT colonography examination. The fifteen subjects from the first study were scanned at a tube current of 40 reference mAs whereas the subjects from the second study were scanned at a tube current of 25 reference mAs. Notice that both radiation levels are lower than in previous studies on electronic cleansing. The amount of tagging agent (meglumine-ioxithalamate; Telebrix Gastro 300 mg I/ml; Guerbet, Cedex, France) ranged from 4 x 50 ml (group 'A') to 3 x 25 ml (group 'D'). This is relevant as it is well known that the tagging agent has a laxative effect and therefore a low tagging dose is preferred. We are unaware of any electronic cleansing algorithm that has been developed and evaluated for CTC with 24-h limited bowel preparation as in the aforementioned studies [21, 17].

One third of these data, i.e. five subjects per group, was randomly selected for development and training of the algorithm. The other ten subjects of each group were reserved to be included in the evaluation study. Colonoscopy served as the reference standard for all cases. All annotated lesions were reviewed by a research fellow to determine whether the polyps measuring ≥ 6 mm were (partially) covered by fecal material or surrounded exclusively by air. Considering prone and supine positions as separate cases yielded three lesions in the training set that were covered by fecal matter. The test set contained 66 lesions; 58 of them were not covered by fecal matter; eight were covered by fecal matter. The effective radiation dose of the employed protocols for an average person (hermaphrodite of 70 kg) is: 3 mSv for 40 reference mAs and 2 mSv for 25 reference mAs [127, 128].

The training set based on data sets 'A'–'D' did not contain sufficient polyps for reliably training the classifier. For the training of the classifier we also included data from another study [3], *only* to extend the number of samples in the polyp class, see Table 7.1 (group 'O'). We showed in earlier work on computer aided detection of polyps that our classifier is robust against this approach. This study originally concerned 1233 patients in total that all adhered to an extensive laxative regime, including contrast agents for stool tagging (50 ml barium [Scan C, Lafayette Pharmaceuticals] and 120 ml of diatrizoate meglumine and diatrizoate sodium [Gastrografin, Bracco Diagnostics]). CT colonography was

Table 7.1: Overview of the data sets used in this paper.

Data set	Number of Subjects	Radiation Level	Tagging Agent	Number of Lesions in Test Set			Ref.
	Training / Testing / Total	reference mAs	ml	≥ 6 mm	≥ 10 mm	Submerged	
'A'	5 / 10 / 50	40	4 x 50	15	8	7	[17]
'B'	5 / 10 / 15	25	3 x 50	7	4	0	[21]
'C'	5 / 10 / 15	25	4 x 25	2	1	1	[21]
'D'	5 / 10 / 15	25	3 x 25	10	7	0	[21]
'O'	15 / 0 / 1233	100	620	0	0	0	[3]
Total	35 / 40 / 1328			34	20	8	

performed in prone and supine position at a tube current of 100 reference mAs. The reference standard was optical colonoscopy. In total 210 polyps larger than or equal to 6 mm were identified as such. Radiologists had retrospectively indicated the location of polyps based on the reference standard. A research fellow selected all those patients that were considered to harbor polyps larger than 6 mm and that were fully submerged in the contrast medium in either scan position. In the end, this delivered 15 such patients with 15 polyps. More details about these data may be retrieved from [3]. Notice that these data were only used for the development of the classifier and not in evaluating the performance of the new electronic cleansing method.

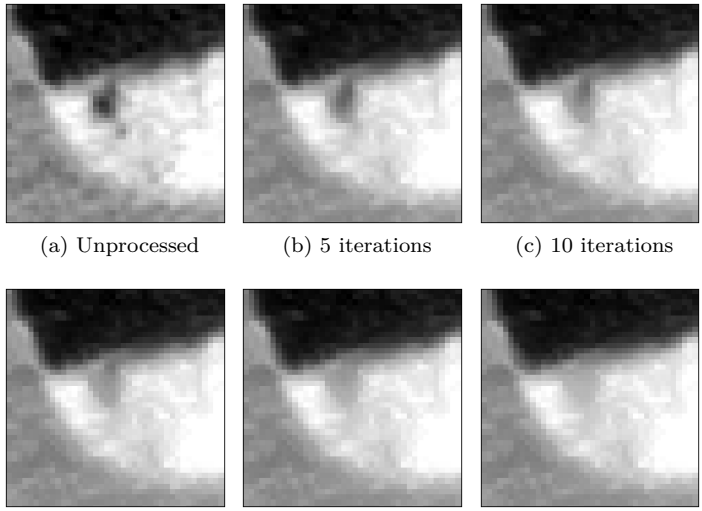
7.2.2 Step 1: Filling inhomogeneous tagging

The normal anatomy of the colon surface can be globally considered a cylindrical structure which is interrupted by indentations, the so-called haustral folds. Colorectal polyps/cancer typically appear as spherical protrusions into the bowel lumen. Analysis of the surface curvature (cylindrical, ridge-like, or cap-like) facilitates the distinction of those structures, which is common practice in CT colonography. We exploit the fact that the local shape of all surfaces can be characterized by the principal curvatures of an isophote surface patch, even the dark objects formed by inhomogeneous tagging. The principal curvature flow method described below incorporates this a prior knowledge of the local shape of the colon surface.

Now, observe that heterogeneities within poorly tagged matter appear as irregularly shaped dark grey objects on a white background (properly tagged fecal remains). Although these heterogeneities have an irregular shape – they consist of locally convex patches (see Fig. 7.2) which are characterized by two negative principal curvatures, $\kappa_2 \leq \kappa_1 \leq 0$. We present a principal curvature flow algorithm inspired by Van Wijk et al. [57] in which these convex regions shrink by evolving the structure in such a way that negative first principal curvatures κ_1 are raised until they approach zero curvature. During the evolution, the shape of the dark objects become strictly convex before disappearing completely. This process is illustrated in Fig. 7.3. Specifically, since the local shape characteristics of the inhomogeneities are determined by the first principal curvature, the locally convex (dark) regions shrink by raising the intensities in areas where the first principal curvature was negative according to

$$\frac{\partial I}{\partial t} = g(\kappa_1, \kappa_2)\,|\nabla I|\,, \tag{7.1}$$

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. $g(\kappa_1, \kappa_2)$ is a continuous function to avoid a discontinuous deformation, especially at locations where the sign of κ_1 changes. Moreover, it



(a) Unprocessed (b) 5 iterations (c) 10 iterations

(d) 15 iterations (e) 25 iterations (f) 70 iterations

Figure 7.2: The image shows poorly tagged fecal matter. The heterogeneity is resolved by iteratively applying the principal curvature flow algorithm. Below the figure is the number of iterations of the algorithm. After 70 iterations, the image does not change significantly.

must be small on folds with a small negative value of κ_1 so that the deformation on such locations is negligible. Reversely, the response to local inhomogeneities with two large negative principal curvatures should be large. In the end $g(\kappa_1, \kappa_2)$ was chosen such that the following partial differential equation (PDE) was solved.

$$\frac{\partial I}{\partial t} = \begin{cases} I_{uu} & (\kappa_1 < 0) \\ 0 & (\kappa_1 \geq 0) \end{cases} . \tag{7.2}$$

in which I_{uu} is the second derivative in the direction of the first principal curvature [88] and I_{uu} is related to $\kappa_1$as follows

$$I_{uu} = -\kappa_1 \left|\nabla I\right| .$$

This shows that the intensity will increase until the largest curvature vanishes and there only is positive curvature left. This also means that the image intensities are not conserved, which is different from, for example, a diffusion process. Fig. 7.2 shows a typical example. Considering the normal shape of the colon, the effect of this evolution scheme is negligible around the colon surface bordering air as the first principal curvature is large at polyps and haustral folds and close to zero in all other parts of the surface. For the submerged colon

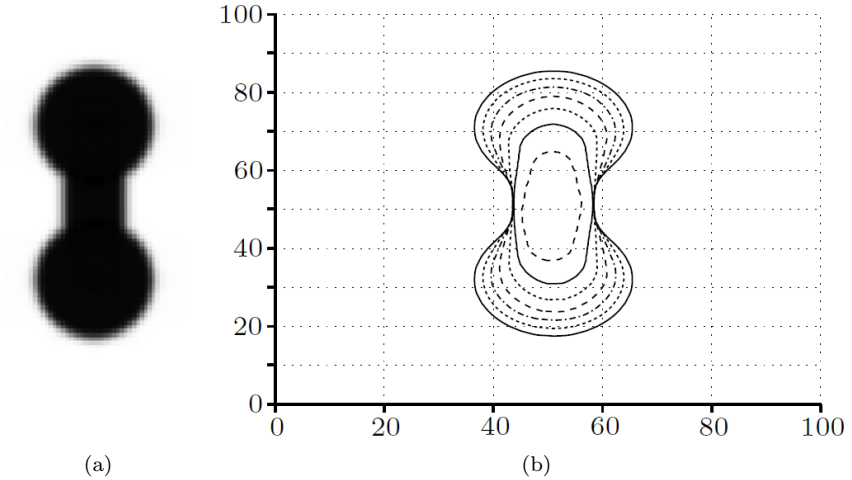


Figure 7.3: A dumbbell object that simulates an heterogeneity in tagged fecal matter. The diagram shows the evolution of a cross-section of the object. The first iterations raise the intensities of the voxels in the local convex parts of the object. This moves the object's contour inwards. After that, the whole object is shrunken until it completely vanishes. Each line represents the shape after a certain number of iterations.

surface in which the sign of the curvatures are inverted with respect to colon-air interface things work out differently. A side effect to this evolution is that it affects the shape of submerged polyps, i.e. these polyps are gradually removed whereas the remainder of the submerged colon surface remains undisturbed.

The number of iterations is largely determined by the size of an object. In general, it takes only a few iterations for small objects, but more for large objects to reach convergence. As such the method adapts to the scale of an object (see also Ref. [57]). Consequently, there is some variation over a whole data set. In practice, about 100 iterations appeared to work well.

7.2.3 Step 2: Reconstruction of Colon Surface

The reconstruction phase addresses the issue that submerged polyps should be retained in the data even though they also display negative κ_1. The approach to do so is to first obtain an accurate estimate of the colon surface without protrusions by applying the original electronic cleansing algorithm (see Section 7.2.4) to the pre-processed data. Subsequently, this estimate can be used to delineate the polyp candidates and to determine whether a candidate is connected to the colon wall.

Previously, a principal curvature flow algorithm served as an initial candidate detection step for automated polyp detection [10, 57]. Characteristic features were derived from the candidates after which polyps were identified by a sophisticated multi-stage logistic classifier. We adapted this classifier to the specific nature of the problem at hand. Automatic polyp detection requires that the sensitivity for polyps is maximal at the expense of as few false positive detections as possible. Here, the aim is to reconstruct the normal bowel wall and leave as little fecal material as possible. The classification system must reconstruct all polyps to permit a 100 % sensitivity for polyps in subsequent reading strategies. To achieve this, (non polypoid) protrusions of the normal bowel may be restored as well, but (preferably) not the fecal residue. Therefore, the classifier does not need to be as strict in separating polyps from other protrusions as in automated polyp detection. We realize that this requires a different way of setting the decision boundary of the classification system. The reconstruction technique was developed using a training set of five subjects from each of the four patient groups 'A'–'D' in Table 7.1 supplemented by fifteen subjects from group 'O' to raise the number of samples of the polyp class (excluded from the experiments below). By adding the polyps from group 'O' to training set, it becomes possible to get a sufficiently good description of the polyp class.

Practically, candidates for restoration are defined as regions with an increase in intensity of more than 200 HU. This value is (approximately) the lower bound for the difference between tagged material and tissue. The limit of 200 HU is comparable to the threshold that was previously used for polyp candidate selection [57]. The objects not retained after this step typically correspond to noise and (irrelevant) small protrusions of the bowel wall, i.e. smaller than a 6 mm polyp.

Among the candidates detected one finds "floating debris". In addition, candidates may (partially or fully) cover areas with a high signal intensity surrounded by material with even higher intensities. Reversely, there may also be areas with low intensities, e.g. due to presence of air. Therefore, we impose the requirement that candidate areas must be connected to the colon wall and contain intensities in the soft tissue range. To do so, the electronic cleansing algorithm (see below) tentatively identifies the colon's inner surface after application of second order curvature flow. Subsequently, we propagated [129] the bowel surface into connected candidate voxels with signal values higher than -200 HU. Notice that this may yield smaller candidate objects while air is discarded. The parameters associated with all these heuristic conditions were set to have 100 % sensitivity for polyps in the training set. One might observe that candidate objects of higher intensity are not considered in this step of the algorithm, because they will be removed by the cleansing step (see below) which is designed to remove all (tagged) high intensities voxels.

Subsequently, a simple classifier will be applied to discard most of the remaining false positives, caused by untagged stool, mixing with air, etc. Conventionally, automated polyp detection relies on the two properties used by

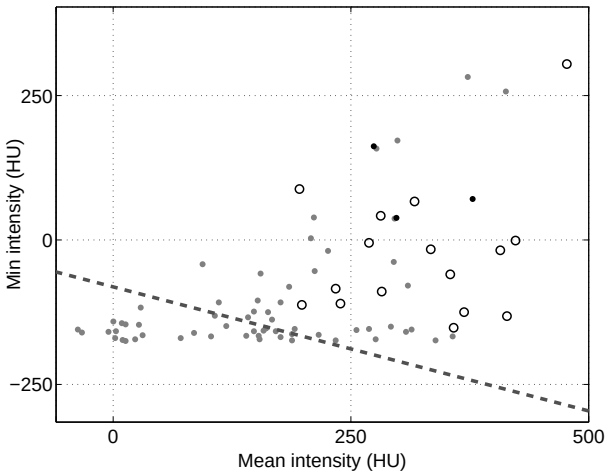


Figure 7.4: Feature space indicating the objects from the training set; black circles are lesions from subject group 'O', black dots are lesions from the training data set from subject groups 'A'–'D' and grey dots are non-polypoid structures. The dashed line indicates the classifier.

radiologists: the shape of a candidate and the intensity distribution inside a candidate. Unfortunately, the shape of a candidate appeared to be a non-distinguishing feature for the current problem. We attribute this to the variety in shapes of poorly tagged material due to which it may closely resemble both polyps as well as non polypoid structures. Therefore, we rely on intensity features of the candidate objects represented by the candidate's mean intensity and its minimum intensity. We use a linear logistic classifier that acts on these features, trained on the objects in our training set. This classifier was used previously for automated polyp detection and is robust against severe class imbalances [73]. Again, we require that the sensitivity for polyps is 100 % in the training set. Also, we do not have to be very strict in the classification as restoring normal (healthy) bowel wall does not affect the goals of this work. To guarantee high lesion sensitivity while the number of training examples is low, the resulting decision boundary was shifted to at least three times the standard deviation away from the mean of the lesion examples composed by subjects from training data sets 'A'–'D'. Fig. 7.4 shows a feature space of the previously mentioned characteristics measured on the training set as well as the resulting decision boundary.

Finally, the restoration of detected objects simply boils down to re-substituting the original intensity values (i.e. prior to principal curvature flow).

Table 7.2: Likert scales used for assessment of reading effort (a), confidence (b), and lesion conspicuity (c).

(a)

Effort	
Very difficult	1
Difficult	2
Easy	3
Very easy	4

(b)

Confidence	
Not confident	1
Medium confident	2
Very confident	3

(c)

Conspicuity	
Inadequate	1
Bad	2
Moderate	3
Sufficient	4
Good	5

7.2.4 Electronic Cleansing

The final step of the method is to perform electronic cleansing, for which we use the technique described by Serlie et al. [16] (see the Appendix for a more extensive description of this method). It was found that this method lead to shorter evaluation time, lower assessment effort, and greater observer confidence than CT colonography without electronic cleansing on patients that underwent extensive bowel preparation [24]. This rigorous preparation simplifies electronic cleansing somewhat because the fecal remains typically have a more homogeneous appearance than in a limited bowel preparation. The principal curvature flow algorithm effectively acts as a pre-processing step that removes heterogeneities from tagging. As such the data become suitable for processing by this electronic cleansing algorithm.

7.2.5 Running Times

The algorithm ran on a Dell Precision T3600 workstation incorporating a quad-core, 2.33 GHz Intel Xeon processor and 3.25 GB Ram system memory. All pre-processing as described in this paper took approximately one minute per data set (thus, per patient position). In addition to that, the (standardly used) electronic cleansing algorithm also took about one minute per data set.

7.3 Experiments and Results

7.3.1 Experimental Setup

Assessing the performance of an electronic cleansing algorithm is a complex problem. Effectively, one would like to determine the performance of reading

Table 7.3: Sensitivity and accuracy for 2D and 3D readings. The sensitivity is determined on a per-lesion basis. Accuracy represents the diagnostic accuracy per patient. The results for both observers are combined, hence doubling the total number of lesions and cases in this table (68 lesions ≥ 6 mm, 40 lesions ≥ 10 mm, 80 cases).

	3D Reading			
	Sensitivity (lesions%)		Accuracy (cases%)	
	≥ 6 mm	≥ 10 mm	≥ 6 mm	≥ 10 mm
Group 'A'	**80** (24/30)	**88** (14/16)	**85** (17/20)	**90** (18/20)
Group 'B'	**71** (10/14)	**88** (7/8)	**100** (20/20)	**100** (20/20)
Group 'C'	**100** (4/4)	**100** (2/2)	**95** (19/20)	**100** (20/20)
Group 'D'	**95** (19/20)	**93** (13/14)	**100** (20/20)	**100** (20/20)
Overall	**84** (57/68)	**90** (36/40)	**95** (76/80)	**98** (78/80)
	2D Reading			
	Sensitivity (lesions%)		Accuracy (cases%)	
	≥ 6 mm	≥ 10 mm	≥ 6 mm	≥ 10 mm
Group 'A'	**70** (21/30)	**75** (12/16)	**85** (17/20)	**90** (18/20)
Group 'B'	**57** (8/14)	**75** (6/8)	**100** (20/20)	**100** (20/20)
Group 'C'	**0** (0/4)	**0** (0/2)	**80** (16/20)	**85** (17/20)
Group 'D'	**85** (17/20)	**100** (14/14)	**100** (20/20)	**100** (20/20)
Overall	**68** (46/68)	**80** (32/40)	**91** (73/80)	**94** (75/80)

electronically cleansed data as well as the extent to which the algorithm modifies polyp shape so that it is not detected anymore. The former is accomplished by a clinical evaluation (part I, below), the latter by a polyp conspicuity study (part II). After all, if the electronic cleansing algorithm would change polyp shape in a destructive manner, this might lead to a different conspicuity. This evaluation of the method was performed much in the same way as by Serlie et al. [24].

Part I

Two observers independently evaluated the 40 cases by means of a primary 2D approach without involvement of electronic cleansing as well as by a primary 3D method after electronic cleansing (employing the unfolded cube fly-through technique [8]). As the focus of this study was to evaluate the electronic cleansing algorithm, the readers were not offered the possibility to observe the uncleansed

3D data. Both observers had a previous experience of evaluating more than 200 colonoscopy verified CT colonography cases. The cases were evaluated twice, in two sessions. In the first session, cases were randomly assigned to be evaluated either by means of the primary 3D or the primary 2D approach. Subsequently, in the second session, the alternative reading method was used. In both sessions, the cases were presented in random order. There was a six-week interval before reading the identical case for the second time to avoid a recall bias. For each lesion, the size, location, and morphology was annotated. Furthermore, the observers rated, per case, their assessment effort on a 4-point Likert scale and their confidence in the reading on a 3-point Likert scale (Table 7.2a and 7.2b). All evaluation times were recorded.

The performance of the observers in reading the data was evaluated by assessing the sensitivity of lesion detection as well as the diagnostic accuracy of case classification. The latter represents the fraction of correctly classified cases. The sensitivity and diagnostic accuracy of the observers was determined by an independent research fellow (prior experience: >200 colonoscopy verified CT colonography cases) in comparison to the colonoscopy data. For primary 2D and primary 3D reading, the per-lesion (polyps and cancers) sensitivity and the per-case diagnostic accuracy were compared with a Generalized Estimating Equation (GEE) analysis [130]. This test corrects for related misses and findings of the two observers. For submerged lesions we performed McNemar's test because there were insufficient lesions to do a GEE analysis. McNemar's test was also used to test the differences in polyp sensitivity of the two observers separately. The reading efforts and reading confidences were assessed by means of a Wilcoxon signed-rank test. For comparison of the reading times a paired t-test was performed.

Part II

The conspicuity of the polyps with and without electronic cleansing was examined by the same two observers more than six weeks after all readings of part I were completed. Prone and supine acquisitions were considered as separate cases in this part of the evaluation. The polyps initially covered by tagged material were shown after electronic cleansing; all polyps surrounded by air were presented as is. The lesions were presented in a 3D reading to the same two observers in random order. The observers indicated the level of conspicuity on a 5-point Likert scale ranging from "inadequate" to "good" (Table 7.2c). For the cases classified as "inadequate" or "moderate", the observers indicated, if possible, the cause. The lesion conspicuity of the polyps residing in air were compared with the polyps partly or fully covered by fecal material using a Wilcoxon signed-rank test.

The results of the evaluation are presented in Sections 7.3.3 and 7.3.4. For all calculations a p-value <0.05 is considered to indicate a statistically significant difference.

7.3.2 Pictorial results

Fig. 7.5 displays a typical instance of the 3D unfolded cube fly-through visualization without and with the proposed pre-processing, i.e. principal curvature flow followed by colon surface reconstruction. It can be observed that the view of the colon surface is severely hampered in the visualization of the original data (see Fig. 7.5a). These artifacts emanate from heterogeneously tagged fecal matter. These artifacts made it impossible to evaluate such limited preparation data in a primary 3D way and hinder 3D problem solving in primary 2D reading approaches. The heterogeneity filter removes the non-polyp-like objects as can be seen in the image resulting after electronic cleansing. Fig. 7.6 shows examples of the pre-processing, reconstruction and cleansing of typical structures that are covered by fecal remains. The first two rows show examples of heterogeneities of which at least one is close to the air-fluid border and at least one is close to the colon wall. The last row shows a polyp on a fold.

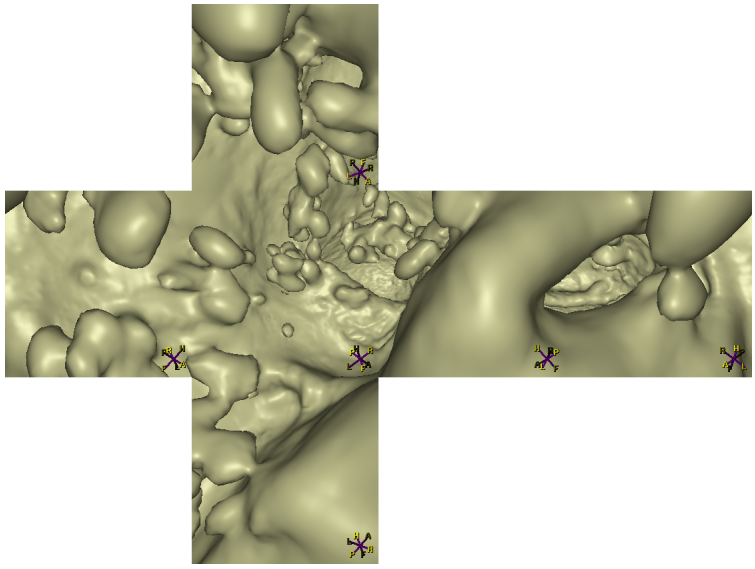
Table 7.4: Overall sensitivities for submerged and not submerged polyps ≥ 6 mm.

Submerged?	3D Reading		2D Reading	
	Sensitivity (lesions%)		Sensitivity (lesions%)	
	Observer I	Observer II	Observer I	Observer II
No	**78** (21/27)	**85** (23/27)	**63** (17/27)	**63** (17/27)
Yes	**100** (7/7)	**86** (6/7)	**86** (6/7)	**86** (6/7)
All	**82** (28/34)	**85** (29/34)	**68** (23/34)	**68** (23/34)

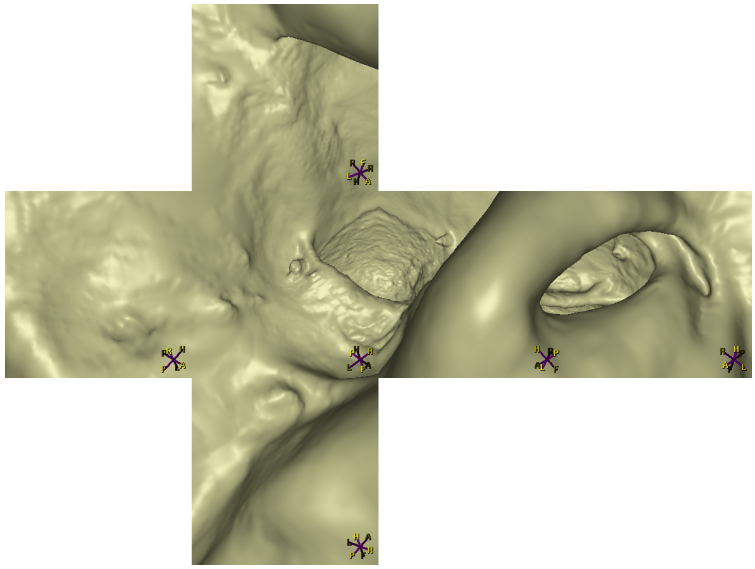
7.3.3 2D and 3D Case Reading Results (part I)

The results are collated in Tables 7.3 and 7.4. The overall sensitivity for lesions ≥ 6 mm was significantly higher for the primary 3D reading strategy after electronic cleansing than for the primary 2D approach ($p = 0.031$). The overall sensitivity for lesions ≥ 10 mm was also higher with primary 3D reading, but the difference was not significant ($p = 0.160$). The per-case accuracies were not significantly different. In total (for the two readers combined) there were eleven false positive findings ≥ 6 mm for primary 2D reading (94% per-case specificity) and twelve for primary 3D reading (96% per-case specificity). We also tested the differences in polyp sensitivity of the two observers separately. For lesions ≥ 6 mm, the p values were 0.18 and 0.07 for observer I and II, respectively (i.e. both not significant).

The sensitivity for primary 3D reading after electronic cleansing was significantly higher and the reader confidence was significantly lower compared with



(a) Visualization without pre-processing



(b) Visualization with pre-processing

Figure 7.5: Result of the 3D unfolded-cube fly-through visualization of the cleansing before (a) and after (b) the electronic cleansing as proposed in this paper.

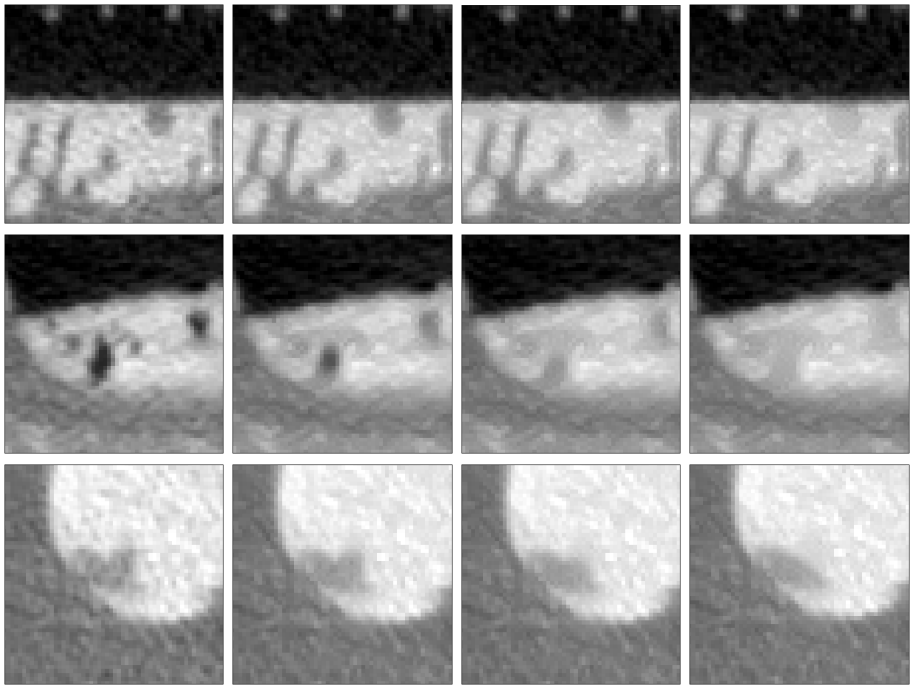


Figure 7.6: Examples of principal curvature flow for objects that are covered by fecal remains: (top row) floating stool; (middle row) heterogeneities, floating stool of which some are close to the colon wall; (bottom row) polyp on a fold. The first column shows the original data, columns 2–4 shows the data after various number of iterations.

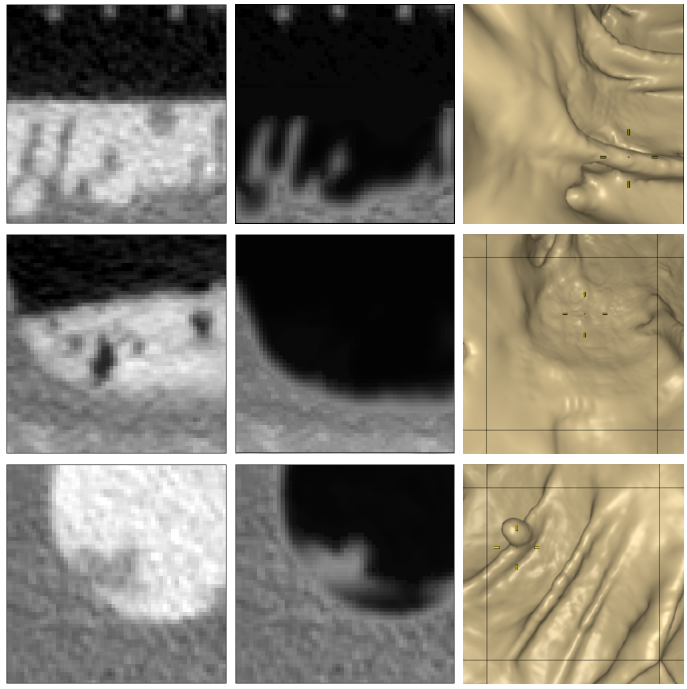


Figure 7.7: Results after cleansing of the examples shown in Figure 7.6. Note that the polyp on the last row re-appears in the reconstruction step as described in Section 7.2.3.

Table 7.5: Average reading times of the two observers.

Observer	2D	3D	Difference
I	7:38	7:32	-0:06
II	3:40	6:47	+3:07
I + II	5:39	7:09	+1:30

Table 7.6: Conspicuity scoring. Lesions in prone and supine acquisitions are considered separately. The scores are based upon the Likert scale for conspicuity in Table 7.2c.

Submerged lesion ?	Median score for observer		
	I	II	I + II
No ($N = 58$)	4	3	3
Yes ($N = 8$)	2.5	2.5	2.5

primary 2D reading. The effort values were slightly, but not significantly decreased for 3D reading compared with primary 2D reading, see Fig. 7.8. The mean effort value (4-point Likert scale) was reduced from 2.8 (2D) to 2.6 (3D) ($p = 0.060$). The average confidence value (3-point Likert scale) was reduced from 2.3 (2D) to 2.1 (3D) ($p = 0.013$).

Also, the overall reading time for both observers combined is significantly lower for 2D reading compared with 3D reading ($p = 0.005$), see Table 7.5. The 2D reading time is particularly lower for observer II ($p = 0.0001$); for observer I this difference is not significant.

7.3.4 Conspicuity Scoring Results (part II)

Table 7.6 shows the median values of the conspicuity scores for the two observers. The combined median conspicuities (for observers I + II) do not differ more than half a scale point and are not significantly different between the submerged and not submerged lesions ($p = 0.092$).

7.4 Discussion

This study shows that our proposed electronic cleansing method enables primary 3D reading of low-dose CT colonography with a 24-h limited bowel preparation. Even for the most limited patient preparation (group 'D') the sensitivity and accuracy remained high. The sensitivity of the primary 3D reading was significantly higher compared with the primary 2D reading for lesions ≥ 6 mm. The

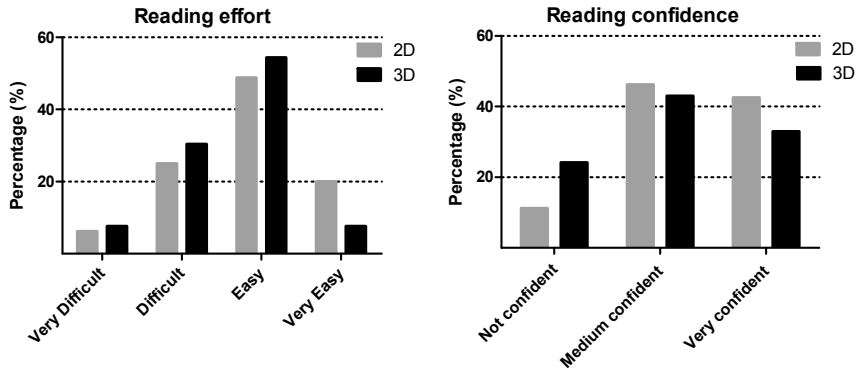


Figure 7.8: Comparison of the reading effort and confidence values of the observers for 2D and 3D reading.

sensitivity of the primary 3D reading was also higher than the sensitivity of the primary 2D reading for lesions ≥ 10 mm, but the difference was not significant. Reader confidence was significantly lower and reading time was significantly higher compared with a 2D primary reading; there was no significant difference in reader effort. There was no significant difference in 3D lesion conspicuity for submerged lesions after electronic cleansing compared with polyps residing in air. The higher sensitivity that we found using the primary 3D reading strategy has been reported previously as well [131, 132]. However, certainly not all studies point towards an improved detection with a primary 3D reading [133, 134]. We noticed that in our data some lesions, which were surrounded by poorly tagged stool, were difficult to see in a 2D reading, but conspicuous enough in 3D to initiate a careful 2D characterization.

The decreased reader confidence for primary 3D reading may seem in conflict with the improved sensitivity results. We believe that the decreased reader confidence, and also the higher reading times in a primary 2D reading strategy, relate to the prior experience of observer II. Observer II had scored image quality in 900 CT colonography cases in a 2D fashion, compared to a smaller number of 200 CT colonography cases that were read in both primary 2D and primary 3D manner [20]. Another explanation for reduced reader confidence may be the result of small holes that were occasionally present in the colon wall as well as on a few spots a large number of small irregularities that prevented a 2D verification for all of them. The latter might have required extra reading time, although primary 3D evaluation was more often found to take longer than primary 2D evaluation [113, 135, 133]. On the other hand, the holes and irregularities did not seem to have a negative influence on the diagnostic performance. Furthermore, it should be noted that although the primary 3D reading time was

longer compared with the 2D primary reading time, these reading times are well below the average reading times reported in the literature [3, 136]. Hence, our cleansing algorithm facilitates time-efficient primary 3D reading as well. In this study we did not evaluate primary 3D reading strategy without electronic cleansing. Furthermore, a separate study on the effect of the low radiation dose on the data quality was outside the scope of this study and can be found elsewhere [137, 116].

This study has a few limitations. As mentioned above, the rendered colon wall sometimes contained small holes. For these data sets the number of iterations of the principal curvature flow may have been too high. Although these holes typically appear in areas with very thin soft tissue structures (not likely to contain polyps or cancers), we would have preferred to avoid this effect. The small irregularities on the bowel wall are likely the result of poorly tagged bowel content. Areas inside the colon where the intensity of the tagged material is below 200 HU will remain problematic since such poorly tagged material simply bears a very close resemblance to soft tissue. Furthermore, data on submerged lesions as well as the number of readers were limited.

7.5 Conclusion

Thus far, data from CT colonography with a reduced radiation dose and with very limited bowel preparation were evaluated using a 2D reading strategy only. Existing cleansing algorithms did not work sufficiently well to allow a primary 3D reading strategy. This study shows that the presented cleansing method allows a primary 3D reading strategy with high lesion sensitivity for low-dose CT colonography with 24-h limited bowel preparation. Moreover, the results indicate that the primary 3D evaluation significantly outperforms the primary 2D evaluation with respect to sensitivity for polyps ≥ 6 mm. To our knowledge this study is the first to describe an electronic cleansing algorithm that has been verified to be able to handle such very limited prepared data.

Chapter 8

Simulation of Scanner- and Patient-specific Low-dose CT Imaging from Existing CT Images

Purpose: *Simulating low-dose Computed Tomography (CT) facilitates in-silico studies into the required dose for a diagnostic task. Conventionally, low-dose CT images are created by adding noise to the recorded projection data. However, this is not always achievable in practice as the raw data are simply not available. This paper aims to present a new method for simulating patient-specific, low-dose CT images without the need of the original projection data.*

Methods: *The low-dose CT simulation included the following steps: (1) computation of a virtual sinogram from a high dose CT image by means of the Radon transform; (2) simulation of an 'reduced'-dose sinogram with appropriate amounts of noise; (3) subtraction of the high-dose virtual sinogram from the reduced-dose sinogram; (4) reconstruction of a noise volume via filtered back-projection; (5) addition of the noise image to the original high-dose image. The required scanner-specific parameters were retrieved from calibration images of a water cylinder. The apodization window was obtained from the noise power spectrum (NPS) in a small region of interest in the center of those images. Furthermore, the bowtie filter attenuation characteristics were derived from the pixel variance. Finally, the X-ray tube output parameter (reflecting the photon flux) and the detector read-out noise were computed from the pixel variance at various exposure levels. The low-dose simulation was evaluated by comparing the noise distribution in simulated images with experimentally acquired data.*

Results: *We found that the models used to recover the scanner specific parameters fitted accurately to the calibration data. What is more, the retrieved apodization window of the reconstruction filter, the bowtie filter, the X-ray tube output parameter and the detector read-out noise were comparable to values reported in literature. Finally, the simulated low-dose images accurately reproduced the noise characteristics in experimentally acquired low-dose-volumes.*

Conclusion: *The developed methods truthfully simulate low-dose CT imaging, without requiring projection data. The scanner-specific parameters can be estimated from experimentally acquired calibration data. The new methodology could aid in further optimizing CT protocols by facilitating in-silico studies on dose dependency of low contrast object detectability.*

8.1 Introduction

Since its invention by Sir Geoffrey Hounsfield about 30 years ago, computed tomography (CT) has established itself as one of the most important medical imaging modalities [138]. In fact, the number of CT examinations is still increasing [139]. An important disadvantage of CT, however, is the exposure to ionizing radiation that is inherent to the technique. Accordingly, it is common practice to keep the radiation dose as low as reasonable achievable (ALARA). Unfortunately, lowering the dose yields a lower signal-to-noise ratio and thus a poorer image quality which may hamper subsequent diagnosis. Optimization of the dose/quality trade-off is a far from trivial problem as one cannot simply expose subjects to a range of radiation doses for ethical reasons. Imaging animals or cadavers as such is not a fully acceptable solution due to the sheer disparity with the clinical case. Therefore, a lower-dose CT image is usually simulated by adding noise to the underlying projection data, i.e. the sinogram [140, 141, 142, 143, 144, 145]. Subsequently, the lower-dose image is reconstructed from these noisy projections using the scanner's software. However, this approach is not always achievable in practice as the projection data are often simply not available. Retrospective, investigation of the influence of low-dose imaging might be permitted if one could generate such data directly from existing images. Obviously, this requires that the simulation process complies with the physics of image formation to produce reliable lower-dose CT volumes.

8.1.1 Related Work

Mayo et al. [144] and Frush et al. [145] were among the first to simulate low-dose CT images. They added Gaussian noise to the projection data, after which the images were generated by means of the scanner's reconstruction software. Any such approach assumes that the number of photons hitting the detector is large. However, when only a low number of photons is detected, the properties of the noise in the sinograms become much more complex. Then, the readout noise

becomes significant and the detection of photons is best described by compound Poisson statistics [146, 147, 148].

Extensive research was done to analytically describe the noise power spectrum (NPS) of CT images [149, 150, 151]. Particularly, the latter work presented techniques to estimate the reconstruction kernel. These methods proved very valuable, since manufacturers are often reluctant to disclose the kernels. Other scanner-specific parameters, such as the bowtie filter and the readout noise, were derived from the projection data [143, 148, 146]. To the best of our knowledge none of the latter parameters were derived from actual image data.

Previous work on low-dose CT simulation also delivered quantitative measures for validation [145, 144, 143, 142, 152]. Boedeker et al. [153] and Faulkner et al. [150] proposed to use the NPS and the noise equivalent quanta (NEQ) to describe the noise properties in CT images, whereas Joemai et al. [142] used the NPS and variance to validate their low-dose CT model.

Recently, Kim and Kim [154] presented a method to simulate low-dose CT scans from the reconstructed images themselves. Although their work gave promising results, it did not incorporate a procedure to calibrate certain scanner-specific parameters which are neither documented in the literature nor supplied by the manufacturer. Additionally, the validation only showed agreement in the average noise level, while the noise properties in CT are known to be highly space-variant. In this paper, we overcome these shortcomings.

8.1.2 Objective

This paper presents a novel method to simulate lower-dose CT images from existing patient data without the need of having the original projection data. The method first creates a virtual sinogram from a high-dose CT image, which is processed to yield one with a lower dose. A lower-dose CT volume is computed from both the original and lower-dose virtual sinograms. The method assumes that a virtual X-ray source produces monochromatic photons with energy equal to the effective energy of the original, polychromatic X-ray tube (as in Refs. [154, 146]). Furthermore, we use CT images of a water barrel for estimating several scanner and scanning parameters (if these are not already available): the X-ray tube's photon flux, the bowtie filter, the reconstruction filter, and the readout noise level. By forward modeling the entire acquisition process, a spatially varying noise distribution is generated. The noise characteristics depend on the actual object that is being imaged as well as the aforementioned acquisition parameters. The entire approach is validated by means of the NPS derived from separate CT images of a water barrel and a pelvic phantom.

The paper is organized as follows. Section 8.2 describes the low-dose simulation method. Subsequently, Section 8.3 goes into how several system parameters can be estimated from CT images. Section 8.4 lists the experiments that are done in order to measure the parameters (8.4.2) and to validate the model (8.4.3). The outcome is discussed in Section 8.5.

8.2 Lower-dose CT Simulator

The lower-dose CT simulator consists of nine steps (see Fig. 8.1):

1. an attenuation image μ_{high} is constructed from a high-dose CT image J_{high};
2. a virtual sinogram R_{high} is generated from μ_{high} by means of the Radon transform [155, 156]. Note that computing the Radon transform requires interpolation, hence this virtual sinogram is slightly more blurred than the true sinogram;
3. a virtual noiseless measurement, defined by the number of detected photons $N_{\text{det,high}}$, is generated from R_{high};
4. a virtual noisy measurement N_{red} is created with appropriate amounts of Poisson and Gaussian noise – reflecting the quantum and readout noise components [138] as well as the noise already present in the original high-dose image;
5. a virtual 'reduced'-dose sinogram R_{red} is computed from N_{red}. This ensures that all smoothing effects that are inherent to the discrete Radon transform (step 2) and the discrete inverse Radon transform (step 7) do not impose additional blurring to the object being imaged;
6. a noise sinogram R_{noise} is obtained by subtracting the virtual sinogram R_{high} from the reduced-dose sinogram R_{red};
7. a noise attenuation image μ_{noise} is reconstructed by means of the inverse Radon transform —via filtered back-projection (FBP)— from the noise sinogram R_{noise};
8. a noise image J_{noise} is constructed from μ_{noise}; and
9. a low-dose attenuation image J_{low} is formed by adding J_{noise} —which contains noise only— to the original high-dose CT image J_{high}.

In the next subsections, we will detail steps 1–9.

8.2.1 The Virtual Sinogram (Steps 1–2)

The attenuation of X-ray radiation on a path from source to detector is described in a discretized form of Lambert-Beer's law by:

$$N_{\text{det}} = N_0 e^{-\sum_{q=0}^{q_{\text{end}}} \mu(q\Delta s)\Delta s}, \tag{8.1}$$

in which N_{det} denotes the number of photons that hit the detector, N_0 the number of photons that would hit the detector in case no object is present, Δs

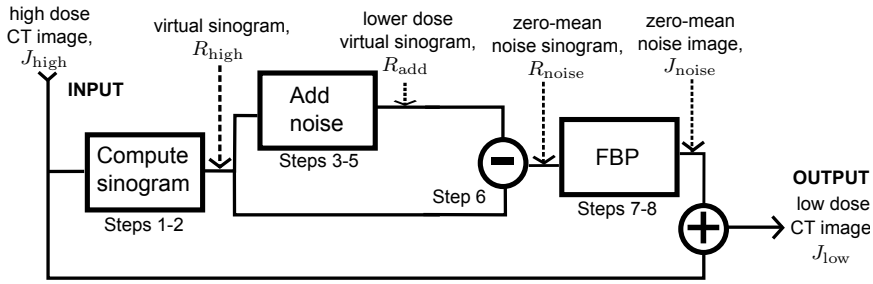


Figure 8.1: Schematic overview of the low-dose CT simulator.

the step size, q the position on the path, $q_{\text{end}}\Delta s$ the end-position of the path and $\mu(q\Delta s)$ represents the (position dependent) attenuation coefficient, also called attenuation image. Henceforth, the summation in the exponent of this equation will be called an attenuation projection, p.

The attenuation image μ_{high} is calculated from the input, high-dose image J_{high} in step 1 via:

$$\mu_{\text{high}} = \frac{\mu_{\text{water}} J_{\text{high}}}{1000} + \mu_{\text{water}}, \tag{8.2}$$

in which we use for μ_{water} the attenuation coefficient of water at the effective energy of an X-ray tube of 80 keV [157, 158].

The aggregate of such attenuation projections in a parallel-beam scanner geometry can be approximated from the high-dose attenuation image using the Radon transform:

$$P_{\text{high}}(k,l) = \sum_{n=1}^{N_{\text{pix}}} \sum_{m=1}^{N_{\text{pix}}} \mu_{\text{high}}(n,m) \cdot \\ \cdot \delta(n\cos(k\Delta\theta) + m\sin(k\Delta\theta) - l\Delta t)\Delta x \Delta y, \tag{8.3}$$

where $\delta(\cdot)$ is the delta function, $k\Delta\theta$ the gantry angle, $l\Delta t$ the position of a point on a straight line intersecting the isocenter, (n,m) the pixel coordinates, N_{pix} the image size, and Δx and Δy the pixel sizes.

For a fan-beam geometry, however, Eq. 8.3 is not valid. Therefore, the projections $P_{\text{high}}(k,l)$ need to be reordered in such a way that they correspond to the projections in fan-beam geometry. In the latter geometry, let $i\Delta\gamma$ represent the angular position on the detector ring (fan angle) and $j\Delta\beta$ the offset angle from the line through the source and isocenter(gantry angle). The relation between a parallel projection $P_{\text{high}}(k,l)$ and a fan-beam projection $R_{\text{high}}(i,j)$ can be derived to be [155]:

$$R_{\text{high}}\left(\frac{1}{\Delta\beta}\left[k\Delta\theta - \arcsin\left(\frac{l\Delta t}{D_{\text{si}}}\right)\right], \frac{1}{\Delta\gamma}\arcsin\left(\frac{l\Delta t}{D_{\text{si}}}\right)\right) = P_{\text{high}}(k,l), \tag{8.4}$$

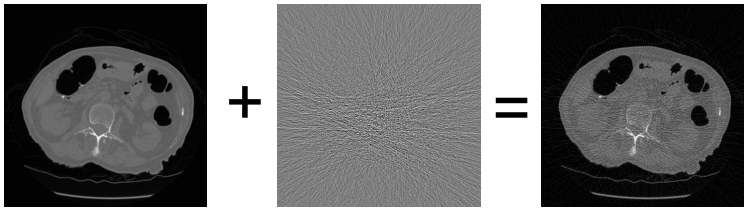


Figure 8.2: A low dose CT image J_{low} (right) is simulated by adding a patient-specific (zero-mean) noise image J_{noise} (middle) to the original high-dose image J_{high} (left).

where D_{si} is the (X-ray) source to isocenter distance. $R_{\text{high}}(i,j)$ should be considered an approximate, *virtual* sinogram, particularly since the noise is strongly reduced due to all the interpolations and averaging involved in its calculation. Eqs. 8.3 and 8.4 together form step 2. The associated, virtual transmission $T_{\text{high}}(i,j)$ is calculated by:

$$T_{\text{high}}(i,j) = e^{-R_{\text{high}}(i,j)}. \tag{8.5}$$

8.2.2 Adding Noise to the Virtual Sinogram (Step 3)

Above, we introduced the idea to create a low-dose attenuation image by adding a space-variant object-dependent noise image to the high-dose attenuation image (see Fig. 8.2). As such, the pixel variance of the low-dose image $\text{var}[\mu_{\text{low}}(n,m)]$ is given by

$$\text{var}[\mu_{\text{low}}(n,m)] = \text{var}[\mu_{\text{high}}(n,m)] + \text{var}[\mu_{\text{noise}}(n,m)], \tag{8.6}$$

where $\text{var}[\mu_{\text{high}}(n,m)]$ and $\text{var}[\mu_{\text{noise}}(n,m)]$ are the pixel variances of the high-dose image and the noise image, respectively.

As a result of filtered back-projection, *any* attenuation image $\mu(n,m)$ is a weighted sum of attenuation projections $R(i,j)$:

$$\mu(n,m) = \left(\frac{\pi\Delta t}{M}\right) \sum_{j=1}^{M} \sum_{i=-N}^{N-1} c_{\text{tot}}(n,m,i,j) R(i,j), \tag{8.7}$$

in which $c_{\text{tot}}(n,m,i,j)$ represents the coefficients of the reconstruction filter (including all the interpolation steps), $2N$ is the number of detector elements, and M is the number of gantry angles composing a full revolution.

Since the noise in the projections is assumed to be independent, the pixel variance equals

$$\text{var}[\mu(n,m)] = \left(\frac{\pi\Delta t}{M}\right)^2 \sum_{j=1}^{M} \sum_{i=-N}^{N-1} c_{\text{tot}}(n,m,i,j)^2 \text{var}[R(i,j)]. \tag{8.8}$$

According to Eq. 8.8, the correct noise characteristics are created when each attenuation projection contains an appropriate amount of noise. Thus, the following condition should be satisfied:

$$\text{var}[R_{\text{low}}(i,j)] = \text{var}[R_{\text{high}}(i,j)] + \text{var}[R_{\text{noise}}(i,j)], \tag{8.9}$$

in which $\text{var}[R_{\text{high}}(i,j)]$ is the variance in the high-dose sinogram, $\text{var}[R_{\text{low}}(i,j)]$ the variance in a sinogram acquired at the specified dose and $\text{var}[R_{\text{noise}}(i,j)]$ the variance in the noise sinogram.

In general, the variance of an attenuation projection $\text{var}\left[R(i,j)\right]$ can be approximated by a first order Taylor series approximation [159, 148, 160] as:

$$\text{var}[R(i,j)] \approx \frac{1}{N_{\text{det}}(i,j)} + \frac{\sigma_e^2}{N_{\text{det}}^2(i,j)}. \tag{8.10}$$

Eq. 8.10 consists of two terms. The first reflects the quantum (photon) noise, which obeys a Poisson distribution. The second term represents the readout noise, which is modeled by zero-mean Gaussian noise $N\left(0,\sigma_e^2\right)$ with variance σ_e^2. The number of detected photons $N_{\text{det}}(i,j)$ is given by

$$N_{\text{det}}(i,j) = N_0(i)T(i,j), \tag{8.11}$$

where $T(i,j)$ is the transmission at fan angle $i\Delta\gamma$ and gantry angle $j\Delta\beta$. $N_0(i)$ is a function of the fan angle $i\Delta\gamma$ due to the bowtie filter. The bowtie filter is incorporated by means of its transmission coefficients as a function of fan angle. Furthermore, N_0 can be defined as a function of protocol- and scanner-dependent variables:

$$N_0(i) = \frac{Kwd_{\text{fan}}I\tau}{M}T_B(i), \tag{8.12}$$

in which w is the collimation (width of the fan beam), d_{fan} the detector size in the fan angle direction at the isocenter, τ the rotation time, I the tube current, $T_B(i)$ the transmission of the bowtie filter at fan angle $i\Delta\gamma$, and K a constant reflecting the X-ray tube output in photons/(mAs.mm^2). The parameters K and $T_B(i)$ are scanner-specific and need to be estimated using calibration scans (if not known a-priori), for which a procedure is detailed in Section 8.4.2. The other parameters in Eq. 8.12 can be typically retrieved from the literature [161, 138] or are included in the DICOM-header. Note that the radiation dose is steered via the exposure $I\tau$. Essentially, step 3 consists of calculating N_{det} using Eqs. 8.5, 8.11 and 8.12.

8.2.3 Model Implementation (Steps 4–6)

This section describes how $R_{\text{red}}(i,j)$ is created in such a way that Eq. 8.9 is fulfilled. Note that the noise sinogram $R_{\text{noise}}(i,j)$ is obtained via $R_{\text{noise}}(i,j)$ = $R_{\text{red}}(i,j)$ - $R_{\text{high}}(i,j)$. Here, $R_{\text{high}}(i,j)$ is the virtual sinogram (obtained

via Eq. 8.4) and not the real sinogram associated with $\mu_{\text{high}}(i,j)$. Due to the interpolations and averaging in the calculation of the discrete Radon transform, the noise of $R_{\text{high}}(i,j)$ is assumed to be negligible, so that

$$\text{var}[R_{\text{noise}}(i,j)] \approx \text{var}[R_{\text{red}}(i,j)].$$

Eq. 8.10 indicates that the quantum noise and the readout noise can be added to yield the total noise. Accordingly, $R_{\text{red}}(i,j)$ is calculated in two steps.

First, only the quantum noise inherent to the detected number of photons is simulated. Essentially, this is implemented by drawing samples from a Poisson distributions with with expectation value $N_{\text{red,q}}(i,j)$. $N_{\text{red,q}}(i,j)$ is calculated from Eq. 8.11.

Substituting Eqs. 8.10, 8.11 and 8.12 into Eq. 8.9 yields an explicit relation between I_{low}, chosen a-priori, I_{high}, given by the high-dose image, and I_{red}, which is the tube current that yields the correct amount of noise to be used for creating the sinogram:

$$I_{\text{red}} = \frac{I_{\text{high}} I_{\text{low}}}{I_{\text{high}} - I_{\text{low}}} \tag{8.13}$$

in which I_{high} and I_{low} are the tube currents of the high-dose and the low-dose image to be simulated. Note that this equation essentially compensates for quantum noise already present in the high-dose image.

Second, readout noise is simulated by repeatedly drawing samples $N_{\text{red,r}}(i,j)$ from a Gaussian distribution with zero mean and variance σ^2_{red}. Given that $N_{\text{red,r}}$ is calculated as described previously, σ^2_{red} is computed by:

$$\frac{\sigma^2_{\text{red}}}{N^2_{\text{red,q}}} = \frac{\sigma^2_e}{N^2_{\text{det,low}}} - \frac{\sigma^2_e}{N^2_{\text{det,high}}}. \tag{8.14}$$

This equation is simplified by substituting Eqs. 8.11, 8.12 and 8.13 into Eq. 8.14 and dropping all redundant terms to yield:

$$\sigma^2_{\text{red}} = \sigma^2_e \frac{I^2_{\text{high}} - I^2_{\text{low}}}{(I_{\text{high}} - I_{low})^2}. \tag{8.15}$$

As before, this equation compensates for Gaussian noise already present in the high-dose data. The virtual noisy measurement (step 4) is found by adding the read-out noise to the Poisson process

$$N_{\text{red}}(i,j) = N_{\text{red,q}}(i,j) + N_{\text{red,r}}(i,j). \tag{8.16}$$

Hence, the virtual 'reduced' dose sinogram (step 5) becomes

$$R_{\text{red}}(i,j) = -\ln\left(N_{\text{red}}(i,j)/N_0(i)\right), \tag{8.17}$$

which yields the noise sinogram (step 6)

$$R_{\text{noise}}(i,j) = R_{\text{red}}(i,j) - R_{\text{high}}(i,j). \tag{8.18}$$

8.2.4 Reconstructing the Image from the Noisy Sinogram (Steps 7–9)

In step 7, we opt to reconstruct $\mu(i,j)$ in a parallel-beam geometry which is conventionally done by manufacturers [155], because it is computationally much less expensive and the image is not blurred space-variantly. Therefore, the fan-beam projections $R_{\text{noise}}(i,j)$ need to be reordered into parallel-beam projections $P_{\text{noise}}(k,l)$ by inverting Eq. 8.4:

$$P_{\text{noise}}\left(\frac{1}{\Delta\theta}\left(j\Delta\beta + i\Delta\gamma\right), \frac{1}{\Delta t} D_{\text{si}} \sin(i\Delta\gamma)\right) = R_{\text{noise}}(i,j). \tag{8.19}$$

Just as before, uniformly sampled parallel projections $P(k,l)$ are obtained by linear interpolation of $R(i,j)$. Subsequently, filtered back-projection is used to construct a noise image via:

$$\mu_{\text{noise}}(n,m) = \left(\frac{\pi\Delta t}{M}\right) \sum_{k=1}^{M_{par}} \sum_{l=-N_{par}}^{N_{par}-1} P_{\text{noise}}(k,l) \cdot \\ \cdot c(n\Delta x \cos(k\Delta\theta) + m\Delta y \sin(k\Delta\theta) - l\Delta t), \tag{8.20}$$

in which c are the reconstruction filter coefficients, and M_{par} and N_{par} the number of gantry angles and detector elements in the parallel beam geometry. Note that c only contains the coefficients of the apodized ramp filter while $c_{\text{tot}}(i,j,n,m)$ also incorporates the necessary interpolation steps. Clearly, the reconstruction filter coefficients are scanner- and protocol-specific. A calibration procedure for obtaining the filter coefficients is described below. Next, $\mu_{\text{noise}}(n,m)$ was scaled to Hounsfield units in step 8 by:

$$J_{\text{noise}}(n,m) = 1000\left(\frac{\mu_{\text{noise}}(n,m) - \mu_{\text{water}}}{\mu_{\text{water}}}\right). \tag{8.21}$$

Finally, the low-dose image can be obtained in step 9 as:

$$J_{\text{low}}(n,m) = J_{\text{high}}(n,m) + J_{\text{noise}}(n,m).$$

8.3 Parameter Estimation

This section describes how the required patient- and scanner-specific parameters may be computed from calibration scans:

1. The reconstruction filter coefficients in $c(l)$ as well as $c_{\text{tot}}(n,m,i,j)$,
2. the bowtie filter transmission $T_B(i)$,
3. the X-ray tube output parameter K, and
4. the readout noise variance σ_e^2.

8.3.1 The reconstruction filter coefficients

Eq. 8.7 can be rewritten in matrix notation in which a system matrix $\mathbf{C}_{\text{tot}}$ contains the filter coefficients $c_{\text{tot}}(n, m, i, j)$:

$$\boldsymbol{\mu} = \mathbf{C}_{\text{tot}}\mathbf{R}, \tag{8.22}$$

where $\boldsymbol{\mu}$ is a N_{pix} vector (the reconstructed image), $\mathbf{C}_{\text{tot}}$ is $N_{\text{pix}} \times M2N$ matrix, and $\mathbf{R}$ is a $M2N$ vector holding the attenuation projections. N_{pix} is the number of pixels in the image to be reconstructed.

In our implementation $\mathbf{C}_{\text{tot}}$ actually consists of a series of matrix multiplications, each representing a different processing step:

$$\mathbf{C}_{\text{tot}} = \mathbf{C}_{\text{back}}\mathbf{C}_{\text{filt}}\mathbf{C}_{\text{fan2par}}. \tag{8.23}$$

Here, $\mathbf{C}_{\text{fan2par}}$ is a $M_{\text{par}}2N_{\text{par}} \times M2N$ matrix that implements the fan-beam to parallel-beam transformation; $\mathbf{C}_{\text{filt}}$ is a $M_{\text{par}}2N_{\text{par}} \times M_{\text{par}}2N_{\text{par}}$ Toeplitz matrix performing the high-pass filtering and $\mathbf{C}_{\text{back}}$ is a $N_{\text{pix}} \times M_{\text{par}}2N_{\text{par}}$ matrix that incorporates the back-projection step. Observe that $\mathbf{C}_{\text{fan2par}}$ and $\mathbf{C}_{\text{back}}$ can be derived from the beam geometry of the simulation. The matrix $\mathbf{C}_{\text{filt}}$ has the filter coefficients $c(l)$ in its rows. Essentially, $c(l)$ is the $M_{\text{par}}2N_{\text{par}}$ 'core' back-projection filter that needs to be estimated in order to compute $c_{\text{tot}}(n, m, i, j)$.

Conventionally, the reconstruction filter $c_{\text{tot}}(n, m, i, j)$ is derived from the NPS in a region of interest (ROI) [154, 153, 150, 149]. If all fan-beam projections $R(i, j)$ used to reconstruct the ROI contain white noise (and aliasing is negligible), the NPS becomes radially symmetric [149]. This is approximately the case in the center of a water cylinder that is placed in the center of the scanner. The pixels inside such a ROI are reconstructed by the projections located at the center of the detector array. The expected number of detected photons is constant, since the water cylinder is locally approximately flat, hence the noise level in each of these projections is the same. Therefore, the noise is approximately white as the amount of cross-talk is negligible [138].

Here, the NPS, which is the Fourier transform of the auto-correlation function, is radially symmetric and given by:

$$\text{NPS}(\omega_r) = H^2_{\text{tot}}(\omega_r)S(\omega_r) \tag{8.24}$$

where $\omega_r = \sqrt{\omega_x^2 + \omega_y^2}$, ω_x and ω_y are the frequencies in Cartesian coordinates, $H_{\text{tot}}(\omega_r)$ the modulation transfer function of the scanner (see below), and $S(\omega_r)$ the NPS of the projections. The assumption that $R(i, j)$ contains white noise makes that the NPS becomes [151, 149]

$$S(\omega_r) \approx \frac{1}{\omega_r}, \tag{8.25}$$

and which is valid when the sampling of the gantry angles of $R(i, j)$ is uniform and sufficiently dense.

H_{tot} is modeled to consist of three elements, namely one apodized ramp filter and two interpolation filters. The apodized ramp filter H_{filter} ensures a mathematically correct reconstruction up to the cut-off frequency of the apodization filter. Furthermore, one interpolation filter H_{fan2par} represents the transformation of the fan-beam rays to uniformly sampled parallel-beam rays, and the other one H_{back} reflects the interpolation along the path during the back-projection. Consequently, when aliasing is ignored, H_{tot} is given by:

$$H_{\text{tot}}(\omega_r) = H_{\text{filter}}(\omega_r) H_{\text{back}}(\omega_r) H_{\text{fan2par}}(\omega_r), \tag{8.26}$$

where H_{filter} is the Fourier transform of $c(l)$, the 'core' back-projection filter, which can be further decomposed into

$$H_{\text{filter}}(\omega_r) = \omega_r H_{\text{apo}}(\omega_r), \tag{8.27}$$

with H_{apo} a cut-off window designed to avoid ringing artifacts near large tissue transitions and to suppress noise in the image. The goal is to determine the shape of H_{apo}, which we approximate by:

$$H_{\text{apo}}(\omega_r) = \frac{\left(a + b\cos\left(\frac{\pi\omega_r}{f_{\text{par}}}\right)\right)}{a+b}, \tag{8.28}$$

where a and b are two filter parameters. For some values of a and b, $H_{\text{apo}}(\omega_r)$ is equal to cut-off windows found in literature [151].

The NPS is estimated in a ROI using the periodogram, which is an estimate of the NPS and is defined by,

$$\text{NPS}(\omega_x, \omega_y) = \sum_{i=1}^{N} \frac{|\mathcal{F}\{J_{\text{noise}}(x,y)\}|^2}{N} \tag{8.29}$$

in which N is the number of images used to estimate the NPS, J_{noise} is a zero-mean noise image (e.g. obtained by subtracting repeated images of the same slice) and $\mathcal{F}\{\}$ symbolizes the Fourier transform. We assume that only linear interpolation is used and therefore Eq. 8.24 becomes (after filling in the previous equations):

$$\text{NPS}(\omega_r) = \omega_r \left(\frac{a + b\cos\left(\frac{\pi\omega_r}{f_{\text{par}}}\right)}{a+b}\right)^2 \text{sinc}^4\left(\frac{\omega_r}{f_{\text{fan}}}\right) \text{sinc}^4\left(\frac{\omega_r}{f_{\text{par}}}\right) \tag{8.30}$$

in which f_{fan} is the Nyquist frequency of the detector array with detectors of size d_{fan} at the isocenter and f_{par} is the Nyquist frequency of the re-binned detectors of size d_{par}.

Essentially, the parameters a and b are estimated by fitting the model described in Eq. 8.30 to the NPS measured in a ROI in the center of a water cylinder placed in the center of the scanner. For that we use the Levenberg-Maquardt optimization algorithm.

Thereafter, $c(l) = \mathcal{F}^{-1}\{H_{\text{apo}}(\omega_r)\}$ and $c_{\text{tot}}(n, m, i, j)$ is obtained via Eq. 8.23.

8.3.2 The bowtie filter

We will now demonstrate how the bowtie filter transmission can be estimated in a least-squares sense from the pixel variance measurements in a phantom. This is an additional novelty of our work, which is required if projection data cannot be obtained.

Assuming that phantom images are acquired at a dose that is high enough to ignore the electronic noise, then the pixel variance $\mathrm{var}[\mu(n,m)]$ can be rewritten by substituting Eqs. 8.10, 8.11 and 8.12 into Eq. 8.8:

$$\mathrm{var}[\mu(n,m)] = \left(\frac{M^2\pi^2\Delta t^2}{Iwd_{\mathrm{fan}}K\tau}\right)\sum_{j=1}^{M}\sum_{i=-N}^{N-1}\frac{c_{\mathrm{tot}}(n,m,i,j)^2}{T_B(i)T(i,j)}. \tag{8.31}$$

Apart from K and $T_B(i)$, the other parameters in this equation are known (c_{tot} may be estimated following the procedure from the previous section and T can be obtained via Eq. 8.5). Writing this equation in matrix notation yields:

$$\mathrm{var}[\boldsymbol{\mu}] = A\mathbf{C}_{\mathrm{tot}}^2\frac{1}{\mathbf{T}_B\cdot\mathbf{T}}, \tag{8.32}$$

where $\mathrm{var}[\boldsymbol{\mu}]$ is an N_{pix} vector (with N_{pix} again the number of pixels); $A = \left(M^2\pi^2\Delta t^2\right)/\left(Iwd_{\mathrm{fan}}K\tau\right)$ is a scalar; $(\cdot)^2$ reflects element-wise square (of the $N_I \times M2N$ matrix $\mathbf{C}_{\mathrm{tot}}$, see above); $\frac{1}{(\cdot)}$ corresponds to an element-wise division; $()\cdot()$ corresponds to an element-wise multiplication;$\mathbf{T}_B$ is an $M2N$ vector representing the bowtie filter transmission, and $\mathbf{T}$ is an $M2N$ vector containing the transmission values.

Eq. 8.32 can be rewritten by putting the elements of $\mathbf{T}$ into a $M2N \times M2N$ diagonal matrix $\mathbf{D}_T$ and by using the property that $\mathbf{T}_B$ does not depend on the gantry angle into:

$$\mathrm{var}[\boldsymbol{\mu}] = \mathbf{C}_{\mathrm{tot}}^2\frac{A}{\mathbf{D}_T\mathbf{D}_M\mathbf{T}_B}, \tag{8.33}$$

in which $\mathbf{D}_M$ is a $M2N \times 2N$ matrix that replicates the bowtie filter for all gantry angles and $\mathbf{T}_B$ is now a $2N$ vector representing the transmission of the bowtie filter. In fact $\mathbf{D}_M$ consists of M 'stacked' $(2N \times 2N)$ unit matrices.

Clearly, Eq. 8.33 is a linear equation that might be solved analytically. Unfortunately, the system matrix and the number of parameters are very large, and therefore such a purely analytic approach is computationally very expensive. Therefore, we opt to model the bowtie filter's transmission by:

$$T_B(i) = o_{\mathrm{low}} + (1 - o_{\mathrm{low}})\left(\sum_{q=0}^{4} a_q \cos\left(\frac{\pi q i}{N}\right)\right)^2, \tag{8.34}$$

with a_q and o_{low} parameters of the model. The summation in the equation represents a truncated fourth order Fourier series. Only cosines are used as the

bowtie filter is a symmetric function and the sum is squared to ensure that the transmission is not negative. Furthermore, we assume that the transmission in the center of the bowtie filter is one ($T_B(0) = 1$), which is imposed by constraining $a_0 = 1 - \sum_{q=1}^{4} a_q$. Thereby, we assume that the transmission also has a lower limit (as the bowtie filter itself has a finite thickness). To that end, o_{low} is a constant representing the minimal transmission. Furthermore, the weighted-sum construction ensures that the transmission can indeed converge to 1.

We assert that T_B is a monotonically decreasing function to both sides of the filter. Therefore, we devised the following simple penalty term:

$$P(\mathbf{a}) = \sum_{i=0}^{N-1} \frac{\partial T_B(i,\mathbf{a})}{\partial i} u\left(\frac{\partial T_B(i,\mathbf{a})}{\partial i}\right), \tag{8.35}$$

with $u(\cdot)$ the Heaviside function and $\mathbf{a}$ the 4 parameter vector of the model (c.f. Eq. 8.34). Essentially, Eq. 8.35 sums all positive derivative values over half the filter (which is symmetric by definition).

Finally, the filter parameters are estimated by solving:

$$\mathbf{a}, A = \arg\min_{\mathbf{a},\mathbf{A}} \left\{ \sum_{k=1}^{N_{\text{pix}}} \chi_k(\mathbf{a}, A, k) + \beta P(\mathbf{a}) \right\}, \tag{8.36}$$

with

$$\chi_k(\mathbf{a}, A, k) = \left(\text{var}[\mu_k]_{\text{experimental}} - \text{var}[\mu_k, \mathbf{a}, A]_{\text{model}}\right)^2, \tag{8.37}$$

where A is the scaling constant from Eq. 8.32 which is essentially a gain factor that needs to be simultaneously estimated and β the weight of the penalty term

The required pixel variances to solve this equation will be obtained from a central region in images of a water cylinder that is repeatedly scanned (see below).

8.3.3 The X-ray tube output parameter and the readout noise level

Finally, we will demonstrate how the X-ray tube output parameter K and the variance of the readout noise σ_e^2 can be estimated from the pixel variance $\text{var}[\mu(n,m)]$ measured at different tube currents. Therefore, Eq. 8.8 will be simplified in such a way that K and σ_e^2 can be derived from a linear fit.

First, substituting Eq. 8.10 into Eq. 8.8 yields:

$$\text{var}[\mu(n,m)] = \left(\frac{\pi \Delta t}{M}\right)^2 \sum_{j=1}^{M} \sum_{i=-N}^{N-1} c_{\text{tot}}(n,m,i,j)^2 \left(\frac{1}{N_{\text{det}}(i,j)} + \frac{\sigma_e^2}{N_{\text{det}}(i,j)^2}\right). \tag{8.38}$$

Subsequently, substituting Eqs. 8.11 and 8.12 into 8.38 and reshuffling the variables yields

$$\mathrm{var}[\mu(n,m)] = \left(\frac{\pi^2\Delta t^2}{IM^2wK\tau}\right)\sum_{j=1}^{M}\sum_{i=-N}^{N-1}\frac{c_{\mathrm{tot}}(n,m,i,j)^2}{T_B(i)T(i,j)} + \\ + \left(\frac{\pi\sigma_e\Delta t}{IKMw\tau}\right)^2\sum_{j=1}^{M}\sum_{i=-N}^{N-1}\frac{c_{\mathrm{tot}}(n,m,i,j)^2}{T_B(i)^2T(i,j)^2}. \quad (8.39)$$

Consequently, the next relation emerges when the exposure $I\tau$, the X-ray tube output K, and the variance of the readout noise σ_e^2 are separated from all other variables of Eq. 8.39:

$$\mathrm{var}[\mu(n,m)] = \frac{C_p(n,m)}{KI\tau} + \frac{C_e(n,m)}{K^2I^2\tau^2}\sigma_e^2, \quad (8.40)$$

in which $C_p(n,m)$ and $C_e(n,m)$ are given by

$$C_p(n,m) = \left(\frac{\pi^2\Delta t^2}{M^2w}\right)\sum_{j=1}^{M}\sum_{i=-N}^{N-1}\frac{c_{\mathrm{tot}}(n,m,i,j)^2}{T_B(i)T(i,j)}, \quad (8.41)$$

and

$$C_e(n,m) = \left(\frac{\pi^2\Delta t^2}{w^2M^2}\right)\sum_{j=1}^{M}\sum_{i=-N}^{N-1}\frac{c_{\mathrm{tot}}(n,m,i,j)^2}{T_B(i)^2T(i,j)^2}. \quad (8.42)$$

$C_p(n,m)$ and $C_e(n,m)$ can be calculated using the reconstruction filter weights c_{tot} (from Section 8.3.1), the bowtie filter (from Section 8.3.2), and the transmission (calculated via Eq. 8.5). Eq. 8.40 is ill-posed whenever the variance is measured at a single exposure. Therefore, acquisitions are obtained at multiple exposures. Next, the model is fitted in a least-squares sense using the following non-linear model,

$$\{K,\sigma_e^2\} = \arg\min_{K,\sigma_e^2}\left\{\sum_{k=1}^{N_{\mathrm{pix}}}\sum_{i=1}^{N_I}\chi_{k,i}(K,\sigma_e^2)\right\}, \quad (8.43)$$

with

$$\chi_{k,i}(K,\sigma_e^2) = \left(\mathrm{var}[\mu_{k,i}]_{\mathrm{measured}} - \mathrm{var}\left[\mu_{k,i}|K,\sigma_e^2\right]_{\mathrm{model}}\right)^2, \quad (8.44)$$

in which the model is expressed by Eq. 8.40, N_{pix} and N_I define the number of pixels and exposures respectively. The initial parameters are obtained by first solving a simpler problem that emerges when Eq. 8.40 is approximated by two linear equations. If the tube current is very high, the contribution of the

read-out noise to the pixel variance is often neglected. In this case, the second term in the right hand side of Eq. 8.40 is ignored, hence K can be approximated from:

$$\mathrm{var}[\mu(n,m)] \approx \frac{C_p(n,m)}{I\tau K}. \tag{8.45}$$

With K known, σ_e^2 can be estimated using new images that are acquired at a lower dose level. Reshuffling Eq. 8.40 gives:

$$\sigma_e^2 = \frac{K^2\tau^2 I^2}{C_e(n,m)} \left(\mathrm{var}[\mu(n,.m)] - \frac{C_p(n,m)}{KI\tau} \right). \tag{8.46}$$

These estimates for K and σ_e^2 are the initial parameters for minimizing Eq. 8.43.

8.4 Results

8.4.1 Measurement Data

CT images of a water cylinder 34 cm in diameter and a pelvic phantom were acquired on a Philips Brilliance 64 CT scanner at the Academical Medical Center in Amsterdam, The Netherlands. A modified CT colon protocol was used, since the intended application is CT colonography. The modifications only concerned the tube current, which was adjusted to control the dose level and the acquisition mode, which was sequential for the water cylinder (i.e. imaging the exact same plane) and the pelvic phantom. Tables 8.1 and 8.2 list the parameter settings. Note that the scan protocol parameters are controlled by the user whereas the geometry parameters are scanner dependent.

Table 8.2: Scanner's geometry parameters.

Scanner's geometry parameters	
Number of detector rings	64
Source to isocenter distance, D_{si} (mm)	570
Source to detector distance, D_{sd} (mm)	1040
Field of measurement:D_{FOM} (mm)	500
Number of detectors, N_{det}	672
Detector size, d_{det} (mm)	1.41
Detector size at iso-center, d_{fan} (mm)	$\sim$0.77
Sampling after rebinning, d_{par} (mm)	$\frac{d_{\mathrm{fan}}}{2}$
Number of gantry angles per revolution, M	1160

Table 8.1: Scan protocol parameters.

Scan protocol parameters			
Phantom type	water cylinder		pelvic phantom
Acquisition mode	sequential		sequential
Kernel	B		B
X-ray tube Voltage (kV)	120		120
Slice Thickness:w (mm)	0.68		0.68
Collimation	40 x 0.625		64 x 0.625
Matrix	512 X 512		512 X 512
Field of measurement: D_{FOM} (mm)	500		500
Set type	Calibration/ Training	Test	Test
Field of view	350	350	350
Pixel sizes: d_{pix}(mm)	0.68	0.68	0.68
Exposure: $I\tau$ (mAs)	250, 120, 60, 30	210, 170, 120, 85, 60, 30, 20, 15	80, 40,15
Number of rotations	25 (except for 250 mAs: 100)	13	128
Copies per rotations	40	40	1
Number of slices	1000 (except for 250 mAs: 4000)	520	128

CT images of the water cylinder were used to estimate the unknown, scanner-specific parameters: H_{apo}, $T_B(i)$, K and σ_e^2. Here, the settings listed under 'Calibration/Training' (Table 8.1) were used. Subsequently, separate images of the water cylinder and images of the pelvic phantom were used to validate

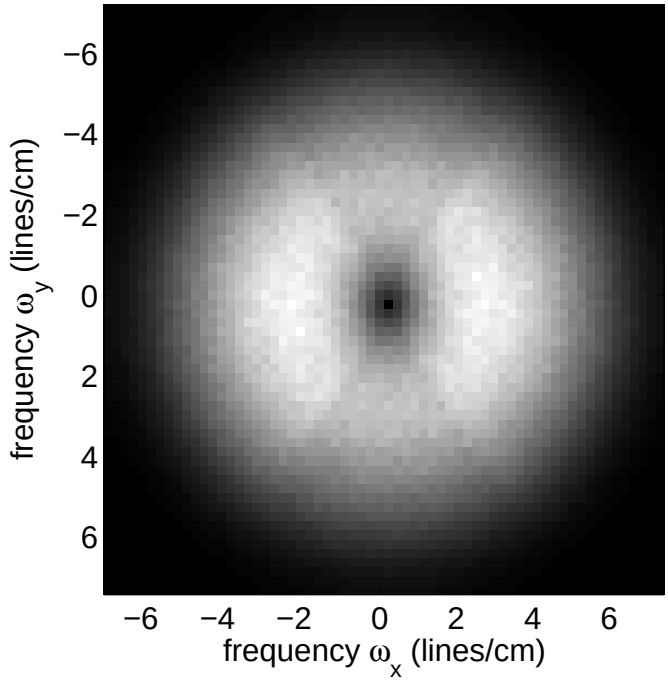


Figure 8.3: The NPS calculated from a 64x64 ROI in the center of a 34cm water cylinder. The NPS was normalized such that it ranged from 0 to 1.

the low-dose simulation model (settings listed under 'Test'). Therefore, simulated and measured noise characteristics were compared by means of the pixel variances and the NPS.

8.4.2 Parameter Estimation

The Reconstruction Filter Coefficients

The volumes emanating from successive X-ray tube rotations at 250 mAs were pairwise subtracted to yield 2000 zero-mean noise images (i.e. corresponding slices from successive rotations). The NPS was computed in a small rectangular ROI consisting of 64x64 pixels in the center of the images. Fig. 8.3 shows the result.

Subsequently, the parameters of the apodization window H_{apo} were estimated by fitting our model (Eq. 8.30) to the NPS. The left plot of Fig. 8.4 demonstrates how well the model fits the data, while the right plot shows in blue the apodization window which is used in the remainder of the paper

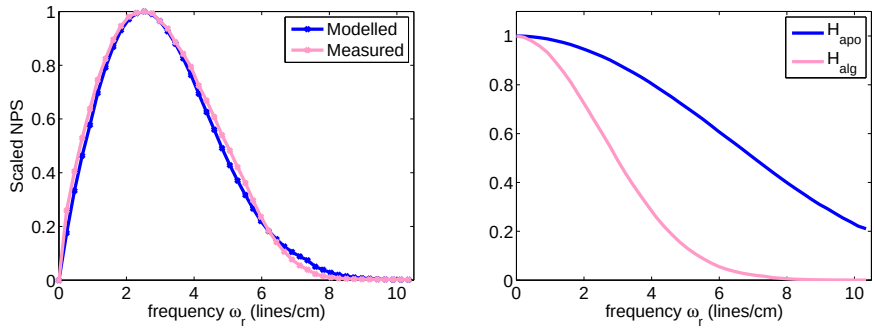


Figure 8.4: Left: the (angular averaged) radial NPS calculated from the images of a water cylinder (pink) and the fitted model, i.e. Eq. 8.30 (blue). Right: the ensuing apodization window H_{apo} (blue, Eq. 8.28) and the algorithmic transfer function, H_{alg}.

(Eq. 8.28). In the same plot, the algorithmic transfer function, H_{alg}, which is defined by $H_{\text{alg}}(\omega) = \left(H_{\text{apo}}(\omega_r) H_{\text{back}}(\omega_r) H_{\text{fan2par}}(\omega_r)\right)^2$ and represents the total apodization, is depicted in pink.

The Bowtie Filter

The 2000 zero-mean noise images from the previous step were used to calculate pixel variances (see Fig. 8.5). Subsequently, only those variances were retained up to 154 mm from the center, i.e. inside the water cylinder. As such, artifacts at the boundary of the cylinder and problems due to signal clipping outside the cylinder are avoided. Furthermore, samples up to 10 pixels (7 mm) from the center were discarded, since the variance of the central pixels cannot be estimated with sufficient precision. Samples were collected from within the remaining region along 80 evenly distributed radial lines drawn outward from the center (see Fig. 8.5). The bowtie filter parameters $\mathbf{a}$ were estimated for each of the 80 radial segments separately (via Eq. 8.36), after which the associated bowtie filter transmissions were averaged. We took this approach for computational reasons since the matrices in Eq. 8.32 are extremely large. The initial parameter setting for every estimation was $\mathbf{a}_{\text{init}} = [0.5, 0, 0, 0]$, which corresponds to a single cosine (see below). o_{low} was set to 0.15, which is comparable to the minimum found in other scanners [146, 143].

Fig. 8.6 shows the estimated bowtie transmission. The left plot gives the profile of the mean bowtie transmission, its 95% confidence interval (all in blue) as well as the initialization (pink). Notice that the bowtie transmission was estimated more precisely for the central detectors than for the ones at the periphery (reflected in the smaller confidence interval). Particularly, the variance

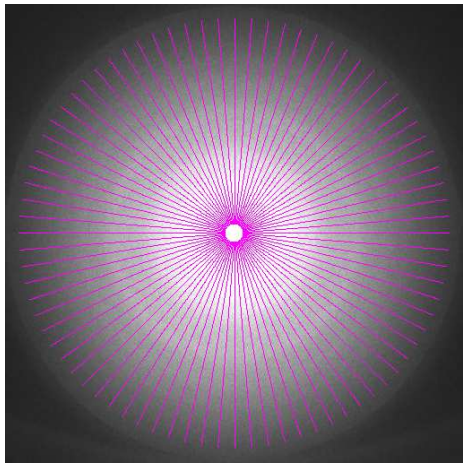


Figure 8.5: Variance image and all of the radial segments. The bowtie filter was calculated for each such segment, i.e. for each separate line on either side of the center.

is large for the range of $200 < |\text{DetectorID}|$, which corresponds to the edges of the water cylinder.

The right plot displays the variance (blue) measured along one of the radial segments depicted in Fig. 8.5. Additionally, the variance as a function of position was computed by means of Eq. 8.33 with (pink) and without the bowtie filter (green). This shows that adding the bowtie filter to the model enabled it to describe the measured variance more accurately.

The shape of the transmission profile was similar to previously estimated bowtie filters of CT-scanners [143, 146].

The X-ray Tube Output Parameter and the Readout Noise Level

The variance per pixel was estimated from the water cylinder images for each exposure level in the training set. As specified in Section 8.4.2, the analysis was restricted to pixels positioned within the 7 mm (10 pixels)–154 mm distance interval from the center.

Fig. 8.7 shows the ratio (ρ) of measured variances at 250 and 120 mAs as a function of the distance to the image center. Neglecting the read-out noise, it follows from Eq. 8.45 that

$$\rho = \frac{\text{var}[\mu(r)]_{I_{\text{low}}}}{\text{var}[\mu(r)]_{I_{\text{high}}}} = \frac{I_{\text{high}}}{I_{\text{low}}} = \frac{250}{120} = 2.08.$$

Fig. 8.7 essentially shows that this assumption was only approximately valid at these dose levels as a paired t-test showed that the pink line varied significantly

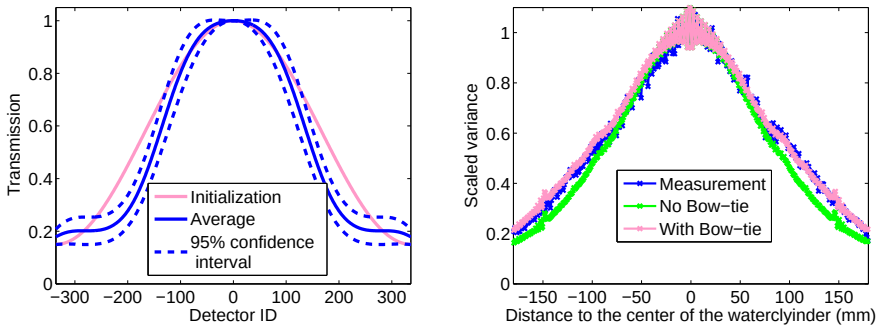


Figure 8.6: left: The mean estimated bowtie filter (blue solid), the 95% confidence intervals(blue dashed) and the bowtie used as initialization (pink). right: the values of the pixels measured (blue) on the diagonal of the variance image depicted in Fig. 8.5, the analytically computed variance with (pink) and without (green) bowtie filter. The variance was scaled by dividing the original by the mean.

from the blue one ($p < 0.05$). Nonetheless, the difference was only 2%.

Next, Fig. 8.8 (left) depicts estimates of K obtained at different exposure levels. K was estimated using Eq. 8.45 by fitting a line through all data points, after which the slope of the line corresponds to $1/(I\tau K)$. Clearly, the estimation of K appears to stabilize for higher exposure levels and increasingly deviates from stability as the exposure level decreases. We attribute this to the increasing importance of the read-out noise due to which Eq. 8.45 is not valid anymore. Henceforth, the estimated value for K at 250 mAs is used as an initialization to compute the initial value of σ_e^2.

The read-out noise σ_e^2 was computed using Eq. 8.46, again based on the 7 mm (10 pixels)–154 mm distance interval from the center. Fig. 8.9 shows histograms of estimated σ_e^2 values at 120 mAs and 30 mAs. Clearly, the distribution is wider at 120 mAs (reflecting less precise estimation), because the total noise is dominated by the Poisson component. Note that the variance can take on negative values due to the subtraction involved in Eq. 8.46. Simultaneously, observe that the mean is larger than zero in each case.

Figure 8.8 displays the read-out noise and uncertainty as a function of exposure in the training data. Table 8.3 lists the estimated values of K and σ_e^2 at each exposure level. The weighted average of estimates of σ_e^2 over all exposure levels is used to initialize the minimization procedure described by Eq. 8.43. Here, the weights were inversely proportional to the variances in the estimations of σ_e^2.

Finally, Fig. 8.10 displays a contour plot depicting the density distribution as a function of the pixel-variances and C_p within the 7 mm (10 pixels) –154 mm

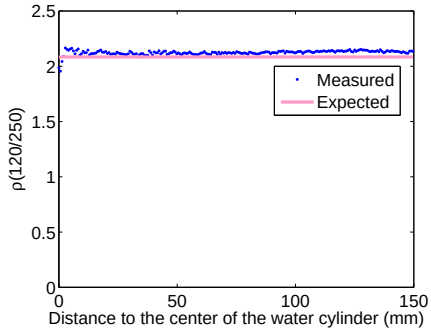


Figure 8.7: The ratio ρ between pixel variances at 120 mAs and 250 mAs as a function of distance to the center of the image. Measurements are depicted in blue, the expected ratio $250/120 = 2.08$ is indicated in pink.

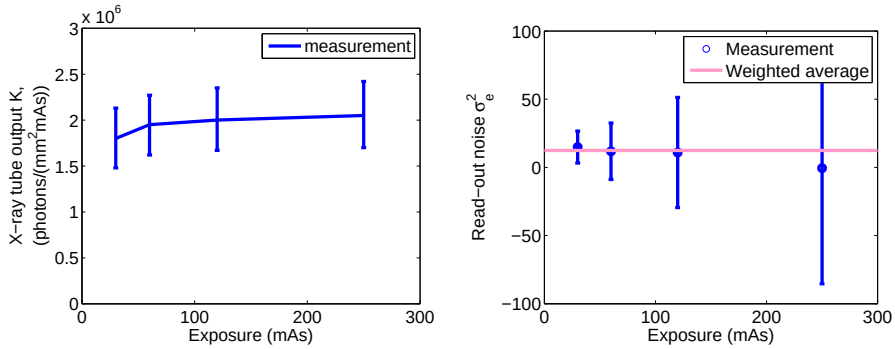


Figure 8.8: left:K as a function of the exposure. right:σ_e^2 as a function of the exposure. The error-bars indicate 97.5 percentile of the estimation K and the ± 1 times the standard deviation of the estimationσ_e^2.

Table 8.3: Estimates for K and σ_e^2 at each exposure level. σ_e^2 is computed using K estimated at 250 mAs. K is expressed in the number of photons/(mm^2.mAs.revolution).

Exposure (mAs)	250	120	60	30
$K\ (\times 10^6)$	2.08	2.03	1.98	1.84
σ_e^2	-0.6	10.9	11.8	14.9

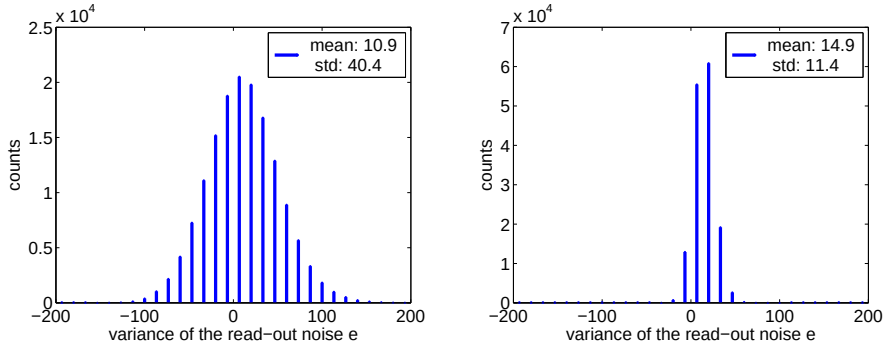


Figure 8.9: Histograms of σ_e^2 estimated in the pixels of the water cylinder at 120 mAs (left) and 30 mAs (right).

ROI of the 250 mAs images. Notice how the pink line (solid), described by Eq. 8.45, is nicely in the center and its 95% confidence interval (dashed) delineates the point cloud. The density of points is much higher in the lower left than the upper right of the figure because more measurement points originate from the periphery of the water cylinder, which have a lower variance. The final values of K and σ_e^2 were estimated to be $2.17 * 10^6$ photons/(mm^2.mAs) and 23.2. These values will be used in the remaining experiments.

8.4.3 Validation of the Low-dose CT Model

The noise characteristics from simulated low-dose CT images were compared with those produced by experimental low-dose CT-scans to validate our method. The scan parameters to do so are collated in Table 8.1 (The 'Test' column under 'water cylinder' and under 'pelvic phantom'). The highest exposure level given in the table served as input to simulate low-dose images. The noise characteristics of all generated low-dose CT images were assessed by means of the pixel variance describing the noise 'strength', and the NPS quantifying the noise structure. Both were computed in a number of ROIs.

34cm Water Cylinder

Fig. 8.11 shows a CT image indicating the positions of the ROIs that were used to compute the NPS. Each ROI had a size of 64x64 pixels.

Fig. 8.12 shows the standard deviation as a function of the distance to the center at different exposure levels: 60 mAs (lower) and 21 mAs (upper), respectively. The pink and blue lines correspond to the angular averaged experimental and the simulated data. Additionally, Table 8.4 gives the relative root-mean-squared difference ϵ_σ between the simulated and experimental noise levels as a

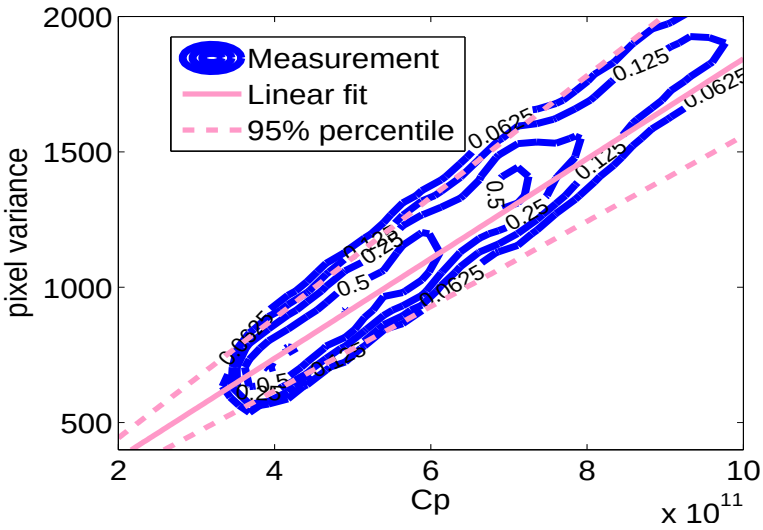


Figure 8.10: A contour plot depicting the density distribution as a function of the pixel variance and C_p, with superimposed a line fit to all data points (continuously drawn) as well as the 95% confidence interval of the fitted line. The slope of the line is equal to $1/(250K)$.

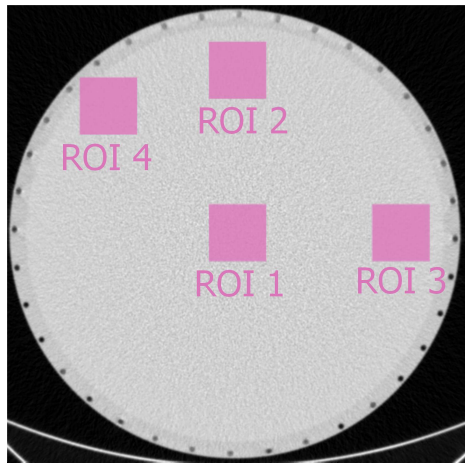


Figure 8.11: Image depicting four ROIs that were used to validate the low-dose CT simulation.

Table 8.4: The relative RMS difference between simulated and experimentally acquired noise levels ϵ_σ as a function of the exposure.

Exposure (mAs)	210	170	120	85	60	42	30	21	15
$\epsilon_\sigma(10^{-2})$	1.9	1.4	1.5	1.5	1.4	2.1	2.3	2.3	2.5

Table 8.5: Relative RMS difference between measured and simulated NPS ϵ_{NPS} for the ROIs depicted in Fig. 8.11.

ROI number	1	2	3	4
ϵ_{NPS}	0.16	0.16	0.12	0.14

function of the exposure:

$$\epsilon_\sigma = \sqrt{\frac{1}{NM}\sum_{n=1}^{M}\sum_{m=1}^{M}\frac{(\sigma_{\text{simu}}(n,m)-\sigma_{\text{exp}}(n,m))^2}{\sigma^2_{\text{simu}}(n,m)}}, \tag{8.47}$$

in which $\sigma_{\text{simu}}(n,m)$ and $\sigma_{\text{exp}}(n,m)$ are the simulated and experimentally acquired standard deviations at pixels (n,m) and N and M indicate the number of pixels on both axes.

Fig. 8.12 and Table 8.4 demonstrate that the simulated noise level closely approximates the experimental noise level in the water cylinder.

Fig. 8.13 shows contour plots of the NPS calculated from ROIs depicted in Fig. 8.11. The blue and pink lines correspond to the experimental and the simulated NPS, respectively. Table 8.5 gives the relative root-mean-squared difference ϵ_{NPS} between the experimental (NPS_{exp}) and simulated NPS (NPS_{simu}) for each ROI, i.e.:

$$\epsilon_{\text{NPS}} = \sqrt{\frac{\sum_{n=1}^{M}\sum_{m=1}^{M}(\text{NPS}_{\text{simu}}(\omega_n,\omega_m)-\text{NPS}_{\text{exp}}(\omega_n,\omega_m))^2}{\sum_{n=1}^{M}\sum_{m=1}^{M}\text{NPS}^2_{\text{simu}}(\omega_n,\omega_m)}}. \tag{8.48}$$

Fig. 8.13 and Table 8.5 signify the close resemblance in shape between the noise of the experiment and the simulation.

Pelvic Phantom

Fifteen ROIs were selected in which the noise properties were analyzed, see Fig. 8.14. Each ROI was composed of 41 x 41 pixels. The standard deviation in each ROI was determined for both the experimental scans and the simulated scans at 80, 40 and 15 mAs. For the simulations at 40 and 15 mAs, the scans at 80 mAs were used as the high-dose image. A validation at 80 mAs was possible by computing the noise properties directly from the zero-mean noise image μ_{noise}

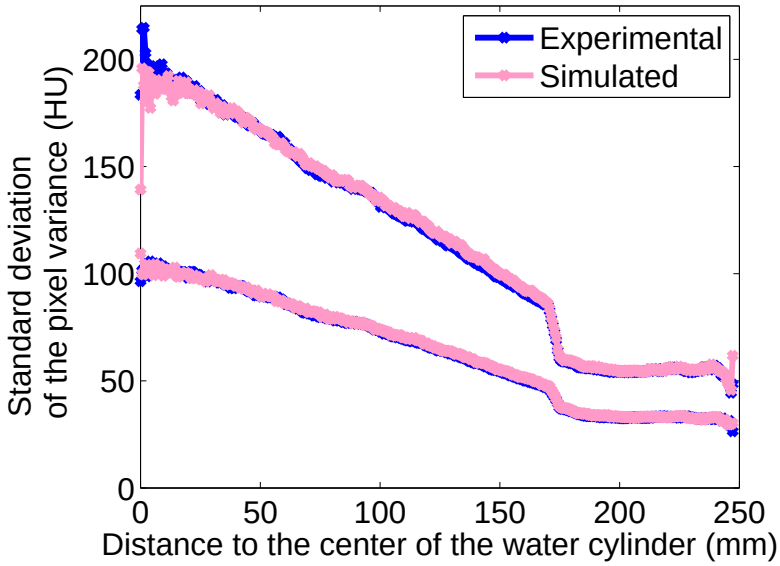


Figure 8.12: The standard deviation of the noise in a 34 cm water cylinder as a function of the distance to the center. A comparison between simulations (pink) and experiments (blue) at 60 mAs (lower lines) and 21 mAs (upper lines). The simulations were based on original images acquired at 250 mAs

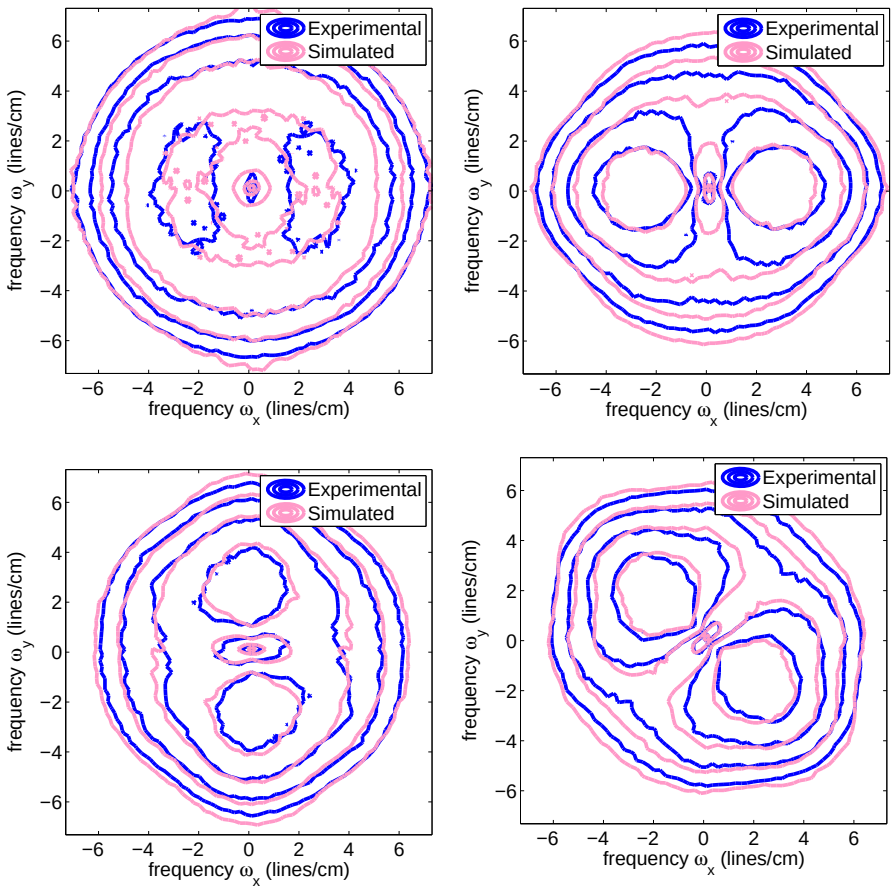


Figure 8.13: Contour plots of the experimental (blue) and the simulated (pink) NPS at 15 mAs in ROI 1 (top left), ROI 2 (top right), ROI 3 (left) and ROI 4 (right) of Fig. 8.11.

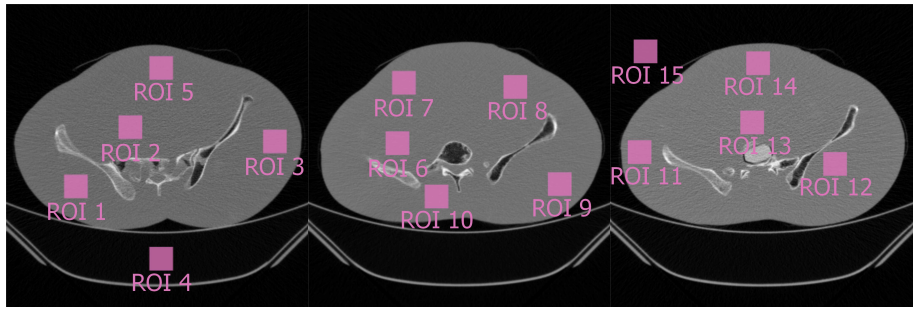


Figure 8.14: Three slices of the pelvic phantom imaged at 80 mAs in which ROIs were selected for the assessment of the low-dose simulation method.

Table 8.6: The parameters and the correlation of the linear fit illustrated in Fig. 8.15.

Exposure (mAs)	slope	offset	correlation
80	0.995	1.71	0.994
40	1.006	0.94	0.998
15	1.018	1.75	0.994

that was simulated assuming the original image was acquired at infinite dose. For each slice and exposure level, 64 simulations were created to take variations in noise realizations into account. On average, the standard deviation of the simulated images deviated 5.3%, 2.4% and 4.3% from the standard deviation of the experimental images for the experimental acquired scans at 80, 40 and 15 mAs, which was within the 95% confidence interval of the estimated standard deviation.

Fig. 8.15 and Table 8.6 show that the simulated scans closely resembled the noise strength in the experimental scans. Nonetheless, the difference between the simulations and experimental scans was larger for the measurements based on the pelvic phantom than the ones based on the water cylinder. We attributed this to the presence of bone-like structures, that cause beam hardening, which was not taken into account in our method.

Furthermore, the NPS was computed in four arbitrarily selected ROIs from experimental and simulated scans at 15 mAs, see Fig. 8.16. Once more, the figures show how well the simulation technique approximates the experimental acquired low-dose images.

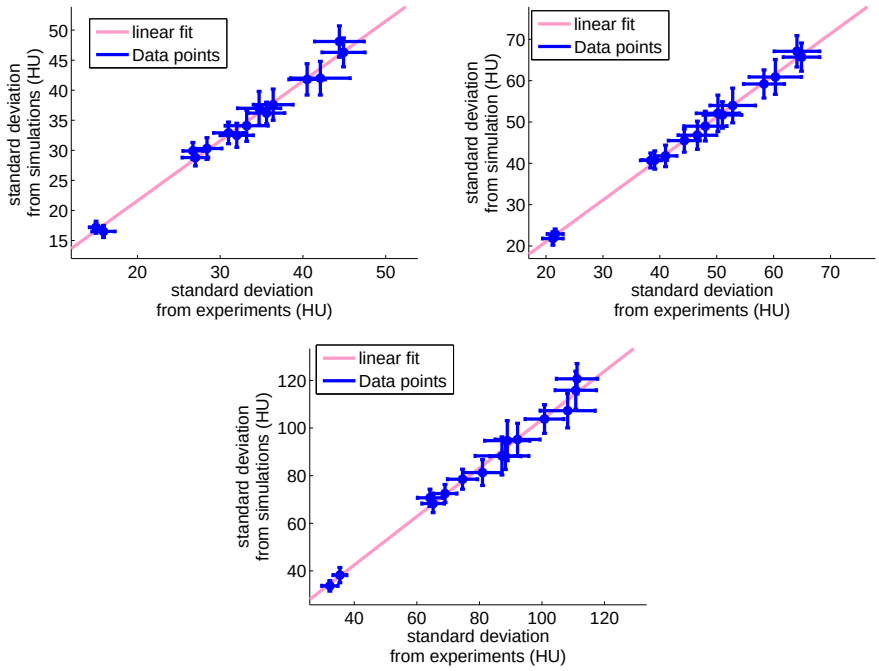


Figure 8.15: Standard deviations obtained from the simulated images as a function of the ones obtained from the experimentally acquired data at 80 mAs (top left), 40 mAs (top right) and 15 mAs (bottom). Each point indicates the average standard deviation in one of the ROIs depicted in Fig. 8.14. The bars depict the 95% confidence interval of the estimation. The pink line represents the linear fit, the parameters of the fit are listed in Table 8.6.

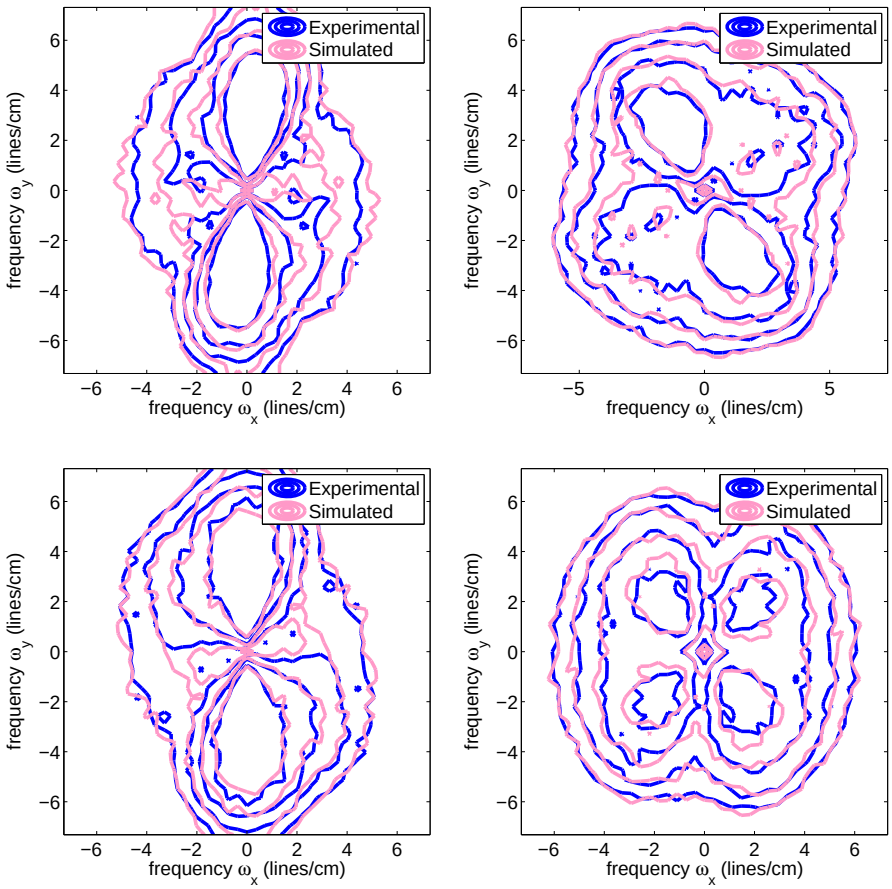


Figure 8.16: Contour plots of the experimental (blue) and the simulated (pink) NPS from ROIs 3, 7, 12, 14 of the pelvic phantom (depicted in Fig. 8.14) at 15 mAs.

8.5 Discussion

We presented a novel method to simulate patient-specific, low-dose CT images from existing high-dose images. Scanner-specific parameters i.e. the apodization window of the reconstruction filter, the bowtie filter, the X-ray tube output parameter and the read-out noise were estimated by means of calibration images of a water cylinder. Therefore, new estimators were developed that used reconstructed images and did not require projection data. The low-dose simulation was evaluated by comparing the noise characteristics in simulated low-dose images with experimentally acquired data.

We demonstrated that the models used to recover the scanner-specific parameters accurately described the calibration data. The estimated reconstruction filter corresponded well to the smooth kernels found earlier [153, 154]. We recognize that a more general model such as an higher order cosine series is necessary to describe the very sharp kernels. The shape of the bowtie filter estimated by us closely resembles the bowtie filter used in a Siemens scanner [146, 143, 162]. Furthermore, we estimated the X-ray tube output parameter at $K = 2.17 * 10^6$ photons/mm^2.mAs and the read-out noise variance at $\sigma_e^2 = 23.2$. These estimated values were of the same order of magnitude as reported previously: Massoumzadeh et al. [143] estimated $2.7 * 10^6$ photons/mm^2.mAs for K and values from 40 through 200 for σ_e^2; Faulkner et al. [150] estimated $4.15 * 10^6$ and $6.46 * 10^6$ photons/mm^2.mAs for K. Ma et al. [148] found 10 for σ_e^2. Finally, we showed that the simulated low-dose images accurately reproduced the noise in experimental low-dose volumes.

A limitation of our approach is in the assumption that monochromatic photons are produced by a virtual X-ray source. As such, our method does not take beam hardening effects into account, which likely cause the encountered deviations in the noise characteristics between simulated and experimental scans of the pelvic phantom. At the same time, the differences between the simulated and experimental data were relatively small even in the presence of bony structures and at relatively low exposures of 15 mAs. Essentially, a polychromatic approach would require a spectral dependency in our framework, particularly concerning the x-ray tube output N_0 and the calculation of attenuation projections. We consider this an important topic for further research.

Another limitation is that our method does not take tube current modulation into account. The tube current modulation essentially adjusts the tube current to the part of the body being imaged and the size of the patient. A variation per slice can be simply incorporated in our method by adjusting N_0 to the actually used tube current I_{high} which may be stored in the DICOM-header. However, it might not be easy to recover complexer variations of N_0, e.g. as a function of the gantry angle. We did not take this into account as the effect on the noise characteristics was proven to be relatively small [160, 156].

Another limitation is that we did not model the correlation of noise due to z-interpolation that is inherent to helical data acquisition. Finally, if iterative

reconstruction algorithms are introduced into commercial CT-scanners in the future, then the low dose CT simulator needs to be updated accordingly.

In summary, the developed methods truthfully simulated low-dose CT imaging without requiring projection data. This new technology might facilitate large-scale studies into the diagnostic accuracy for lower CT dose. In turn, it could aid in further reducing the radiation risks of CT-examinations.

8.A Appendix: CAD for Low-Dose CT Colonography

8.A.1 Goal

This Appendix presents the results of an in silico study to the performance of our computer aided polyp detection system for lower dose CTC data.

8.A.2 Experiments

We used a data set that was used previously by us to validate the automated polyp detection algorithm (data set 'C' in Ref. [73]). The data from 61 out of 141 randomly selected patients were used. Notice that these scans were acquired at relatively high radiation dose levels of around 75 mAs [55]. In this data set, 71 polyps ≥ 6 mm were found; 15 of which were ≥ 10 mm..

Lower dose CT scans of these patients were simulated at three dose levels: 20 mAs, 10 mAs, and 5 mAs. The data was interpreted by the CAD system presented previously [73, 57].

8.A.3 Results

Fig. 8.17 shows FROC curves for the original data set and the three simulated dose data sets. Apparently, the 5 mAs data set has a sensitivity reduced by approximately just 10% compared to the original data. The 15 lesions that remained undetected even in the original data concerned tumors with lobulated shapes, polyps covered by fecal remains, 'non-protruding' polyps annotated as flat lesions by the radiologists and polyps that were located between haustral folds (see also Chapter 2).

8.A.4 Discussion/Conclusion

Comparing the individual FROC curves shows that the sensitivity is reduced by 10% as a consequence of reducing the radiation dose by a factor of around fifteen. Based on this gradual decline in performance, we conclude that the CAD system is robust to adding simulated noise. When comparing the results with previously presented results (see Fig. 2.9(c)), it should be noted that the overall sensitivity is lower even for the FROC curve without adding simulated noise. This can be explained by an unfavorable selection of subjects with relatively more polyps covered by fecal remains, flat polyps and tumors which cannot be detected by the CAD system.

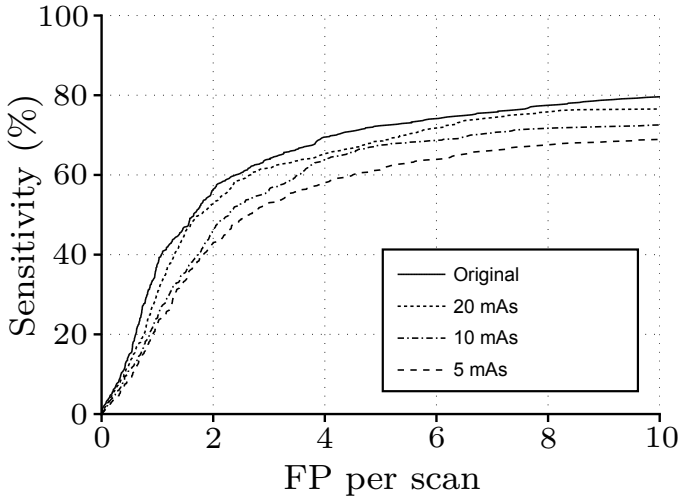


Figure 8.17: FROC curves showing the performance of the polyp detection system for the same data set at varying dose levels.

Chapter 9

Conclusion

We designed, implemented and evaluated new techniques for computer aided detection (CAD), electronic cleansing (EC) and low-dose simulation based on high-dose reconstructed images to be used in CT colonography (CTC). Automated polyp detection was presented as a technique to assist the radiologist by indicating suspicious locations. Electronic cleansing allows for 3D reading of CTC data even if the patients did not undergo an extensive bowel preparation. Furthermore, we investigated how the techniques are affected by a reduced radiation dose.

9.1 Automated Polyp Detection

A method based on a minimal principal curvature flow was presented which sustains polyp detection based on the protruding character irrespective of the polyp's size and shape. The method modifies the appearance of the colon surface at locations of protruding objects only. This was achieved by finding a steady state solution of a nonlinear partial differential equation with the CTC image as input. The method allows for a simple segmentation of polyp candidates by applying a single threshold on an intensity change field and it does not make a specific choice for the scale that is used in the involved first and second 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 and the convergence criteria of the posed partial differential equation adapts to the local data.

We developed a classification system based on logistic regression that makes use of a measure for the detection of polyp candidates that 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, and thereby limits the risk of missing large polyps. As such, the ordering correlates with the clinical relevance of the candidates. 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 shear 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 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 of low-complexity and proved to be stable while only using a small number of features: the candidate's protrudedness from the colon wall, a feature derived from the protrusion field that was sensitive for candidates that had steep edges and large protrusion, and the internal intensity distribution. The features were divided into two types, namely features that directly allowed 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. The Mahalanobis distance mapping in conjunction with logistic regression is generally applicable to obtain a clinically relevant ordering of the candidates.

The 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. The observed sensitivity was comparable to the sensitivity of radiologists using CTC [6, 3, 9] and competed with other CAD systems [9, 43, 29, 30]. 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 an essential prerequisite for application in large-scale screening for colorectal polyps.

To conclude, we introduced a low-complex CAD system in which the candidate detection, segmentation and part of the classification steps were combined into a single step. The characteristics of the remaining classification problem led to a low-complex classification system that proved to generalize to data from different medical centers. This fullfilled the requirement of a robust low-complex automated polyp detection system.

9.2 Electronic Cleansing

We extended an existing electronic cleansing system to be able to cope with CTC data with reduced radiation dose and limited bowel preparation. Based

on knowledge obtained from automated polyp detection, a pre-processing step was added to transform the data such that heterogeneities in tagged material were removed while leaving the colon surface unchanged. This is done by iteratively removing inhomogeneities using a principal curvature algorithm. As shown for polyp detection, this algorithm is able to remove inhomogeneities (or protrusions) while retaining the colon wall. The latter is retained because of its specific characteristic that one of the principal curvatures (depending on the application) is smaller than or equal to zero everywhere.

An evaluation study showed that our proposed method enables primary 3D reading of low-dose CT colonography with a 24-hour limited bowel preparation. For a variety of subject preparations, the sensitivity and accuracy remained high. The sensitivity of the primary 3D reading was significantly higher compared with the primary 2D reading for lesions ≥ 6 mm. The sensitivity of the primary 3D reading was also higher than the sensitivity of the primary 2D reading for lesions ≥ 10 mm, but the difference was not significant. Reader confidence was significantly lower and reading time was significantly higher compared with a 2D primary reading; there was no significant difference in reader effort. Also, there was no significant difference in 3D lesion conspicuity for submerged lesions after electronic cleansing compared with polyps residing in air. Furthermore, it should be noted that although the primary 3D reading time (7:09) was longer compared with the 2D primary reading time (5:39), these reading times are well below the average reading times reported in the literature. Hence, our cleansing algorithm facilitates time-efficient primary 3D reading as well.

The study had several limitations. Artefacts might still occur in the visualization but they did not seem to influence the observers' performance. Areas inside the colon where the intensity of the tagged material is below 200 HU will (probably) always remain problematic since such poorly tagged material simply bears a very close resemblance to soft tissue. Another limitation was that the number of submerged polyps in the evaluation of polyp conspicuity was limited, especially in the most limited bowel preparation groups. This limitation arose because the fraction of polyps submerged by tagged material is in general small.

Thus far, CT colonography data acquired after very limited preparation were evaluated using a 2D reading strategy only. Existing cleansing algorithms did not work sufficiently well to allow a primary 3D reading strategy. Our study showed that the presented cleansing method allowed a primary 3D reading strategy with high lesion sensitivity for low-dose CT colonography with 24-hour limited bowel preparation. Moreover, the results proved that the primary 3D evaluation significantly outperformed the primary 2D evaluation with respect to lesion sensitivity.

9.3 Dose Reduction

To evaluate the performance of CAD systems for polyp detection for low-dose CT colonography, we developed a method to add patient and scanner specific noise to existing CT colonography data sets. Using this method, it was not necessary to have the raw projection data, which often is simply not available. The method was patient specific because the added noise realizations were based on the virtual sinograms obtained from the high-dose image. The method was also scanner specific as all scanner parameters were extracted from calibration images. As such, the apodization window of the reconstruction filter, the bowtie filter, the X-ray tube output, and the detector's read-out noise were estimated from calibration images without the need to use the raw projection data. The low-dose simulation was evaluated by comparing the standard deviation of the noise and the noise power spectrum in simulated low dose images with experimentally acquired data. To conclude, we showed that the developed methods truthfully simulated low-dose CT imaging.

This new technology facilitated to retrospectively analyze the behaviour of the CAD system for low dose CTC. In an experiment we analyzed 61 patients from a data set previously used in the evaluation of the CAD system. It showed that the CAD system gradually degraded with increasing noise levels and that it is feasible to investigate the use of the CAD system for lower dose CTC.

9.4 Future Directions

The results confirm the hypothesis that the principal curvature flow algorithm used for polyp detection and electronic cleansing is quite robust against increasing noise levels. However, studies on polyp detection for lower dose CTC data are still limited and need to be extended to larger data sets. At the same time, even while the CAD system is found to be robust against noise, there still is a debate on how to incorporate the CAD system into clinical practice. Moreover, no studies are done into the combined usage of electronic cleansing and CAD for increasing noise levels.

To further improve the electronic cleansing, dual-energy CT colonography is a promising technique to be able to distinguish between poorly tagged materials and tissue materials. At the moment no clinical evaluations have taken place indicating that more development is needed to evaluate whether this technique can indeed improve the diagnostic value of CTC. As dual-energy CTC records two images of lower radiation dose, further research is also needed to evaluate whether the trade-off between additional multi-spectral information and increasing noise levels will be benificial in practice. Our low-dose simulation might be used to investigate this trade-off before large-scale evaluations take place.

Although research into CT colonography has overcome some serious limiti-

ations, there will always be the remaining burden of the radiation dose limiting the spread of its use. This is the main reason that magnetic resonance (MR) colonography is an important research topic. Unfortunately, MR colonography also has serious drawbacks. Due to varying contrast over a single image, the lower resolution of MRI compared to CT, and motion artefactsdue to the longer acquisition times, the image processing and visualization techniques developed for CT colonography data are not directly applicable to MR colonography data. In that context, segmentation of the colon and automated detection of polyps remain important subjects of current research.

Bibliography

[1] "Colorectal cancer facts & figures," Tech. Rep. No. 8617.00, American Cancer Society, Atlanta, 2005.

[2] J. T. Ferrucci, "Colon cancer screening with virtual colonoscopy: Promise, polyps, politics," *Am. J. Roentgenol.*, vol. 177, pp. 975–88, 2001.

[3] 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*, vol. 349, pp. 2191–200, 2003.

[4] P. J. Pickhardt, J. R. Choi, I. Hwang, and W. R. Schindler, "Nonadenomatous polyps at CT colonography: Prevalence, size distribution, and detection rates," *Radiology*, vol. 232, pp. 784–90, 2004.

[5] K. S. Nanda and M. J. Bourke, "Endoscopic mucosal resection and complications," *Techniques in Gastrointestinal Endoscopy*, vol. 15, no. 2, pp. 88–95, 2013.

[6] R. E. van Gelder, C. Y. Nio, J. Florie, J. F. Bartelsman, P. Snel, S. W. de Jager, S. J. van Deventer, J. S. Laméris, P. M. Bossuyt, and J. Stoker, "Computed tomographic colonography compared with colonoscopy in patients at increased risk for colorectal cancer," *Gastroenterology*, vol. 127, no. 1, pp. 41–8, 2004.

[7] 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*, vol. 341, no. 20, pp. 1496–1503, 1999.

[8] F. M. Vos, R. E. van Gelder, I. W. O. 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*, vol. 228, pp. 878–85, 2003.

[9] 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*, vol. 129, no. 6, pp. 1832–44, 2005.

[10] 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 *Proc. MICCAI'06*, vol. LNCS 4191, pp. 471–8, 2006.

[11] I. W. O. Serlie, R. Truyen, J. Florie, F. H. Post, L. J. van Vliet, and F. M. Vos, "Computed cleansing for virtual colonoscopy using a three-material transition model," in *Proc. MICCAI'03*, vol. LNCS 2879, pp. 175–83, 2003.

[12] R. E. van Gelder, E. Birnie, J. Florie, M. P. Schutter, J. F. Bartelsman, P. Snel, J. S. Laméris, G. J. Bonsel, and J. Stoker, "CT colonography and colonoscopy: Assessment of patient preference in a 5-week follow-up study," *Radiology*, vol. 233, no. 2, pp. 328–37, 2004.

[13] M. H. Liedenbaum, A. H. de Vries, S. Halligan, P. M. M. Bossuyt, A. H. Dachman, E. Dekker, J. Florie, S. S. Gryspeerdt, S. Jensch, C. D. Johnson, A. Laghi, S. A. Taylor, and J. Stoker, "CT colonography polyp matching: differences between experienced readers," *Eur. Radiol.*, vol. 19, pp. 1723–1730, 2009.

[14] N. H. Hyman, P.Anderson, and H. Blasyk, "Hyperplastic polyposis and the risk of colorectal cancer," *Dis Colon Rectum*, vol. 47, pp. 2101–4, 2004.

[15] I. C. Sluimer, P. F. van Waes, M. A. Viergever, and B. van Ginneken, "Computer-aided diagnosis in high resolution CT of the lungs," *Med. Phys.*, vol. 30, no. 12, pp. 3081–90, 2003.

[16] I. W. O. Serlie, F. M. Vos, R. Truyen, F. H. Post, and L. J. van Vliet, "Classifying CT image data into material fractions by a scale and rotation invariant edge model," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2891–904, 2007.

[17] M. H. Liedenbaum, A. H. de Vries, C. I. Gouw, A. F. van Rijn, S. Bipat, E. Dekker, and J. Stoker, "CT colonography with minimal bowel preparation: evaluation of tagging quality, patient acceptance and diagnostic accuracy in two iodine-based preparation schemes," *Eur. Radiol.*, vol. 20, pp. 367–76, Feb. 2010.

[18] E. Neri, F. Turini, F. Cerri, P. Vagle, and C. Bartolozzi, "CT colonography: same-day tagging regimen with iodixanol and reduced cathartic preparation," *Abdom. Imaging*, vol. 34, pp. 642–7, Sep.–Oct. 2009.

[19] K. Nagata, T. Okawa, A. Honma, S. Endo, S. E. Kudo, and H. Yoshida, "Full-laxative versus minimum-laxative fecal-tagging CT colonography using 64-detector row CT: prospective blinded comparison of diagnostic performance, tagging quality, and patient acceptance," *Acad. Radiol.*, vol. 16, pp. 780–9, Jul. 2009.

[20] E. M. Stoop, M. C. de Haan, T. R. de Wijkerslooth, P. M. Bossuyt, M. van Ballegooijen, C. Y. Nio, M. J. van de Vijver, K. Biermann, M. Thomeer, M. E. van Leerdam, P. Fockens, J. Stoker, E. J. Kuipers, and E. Dekker, "Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial," *The Lancet Oncology*, vol. 13, no. 1, pp. 55–64, 2012.

[21] M. H. Liedenbaum, M. J. Denters, F. M. Zijta, V. F. van Ravesteijn, F. M. Vos, S. Bipat, E. Dekker, and J. Stoker, "Reducing the oral contrast dose in CT colonography: evaluation of faecal tagging quality and patient acceptance," *Clin. Radiol.*, vol. 66, pp. 30–7, 2011.

[22] W. Cai, M. E. Zalis, J. Näppi, G. J. Harris, and H. Yoshida, "Structure-analysis method for electronic cleansing in cathartic and noncathartic CT colonography," *Med. Phys.*, vol. 35, pp. 3259–77, Jul. 2008.

[23] W. Cai, H. Yoshida, M. E. Zalis, J. Näppi, and G. J. Harris, "Informatics in radiology: electronic cleansing for noncathartic CT colonography: a structure-analysis scheme," *Radiographics*, vol. 30, pp. 585–602, May.–Jun. 2010.

[24] I. W. O. Serlie, A. H. de Vries, F. M. Vos, Y. Nio, R. Truyen, J. Stoker, and L. J. van Vliet, "Lesion conspicuity and efficiency of CT colonography with electronic cleansing based on a three-material transition model," *Am. J. Roentgenol.*, vol. 191, no. 5, pp. 1493–502, 2008.

[25] B. C. Morson, "Evolution of cancer of the colon and rectum," *Cancer*, vol. 34, pp. 345–9, 1974.

[26] J. H. Bond, "Clinical evidence for the adenoma-carcinoma sequence, and the management of patients with colorectal adenomas," *Semin Gastrointest Dis*, vol. 11, pp. 176–84, 2000.

[27] P. J. Pickhardt, "CT colonography (virtual colonoscopy) for primary colorectal screening: Challenges facing clinical implementation," *Abdom. Imaging*, vol. 30, pp. 1–4, 2005.

[28] 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*, vol. 124, pp. 544–60, 2003.

[29] J. Näppi and H. Yoshida, "Fully automated three-dimensional detection of polyps in fecal-tagging CT colonography," *Acad. Radiol.*, vol. 14, pp. 287–300, 2007.

[30] 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.*, vol. 191, pp. 168–74, 2008.

[31] H. Yoshida and J. Nppi, "CAD in CT colonography without and with oral contrast agents: Progress and challenges," *Computerized Medical Imaging and Graphics*, vol. 31, pp. 267–84, 2007.

[32] 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*, vol. 236, pp. 872–78, 2005.

[33] 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*, vol. 236, pp. 3–9, 2005.

[34] 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*, vol. 109, pp. 2213–21, 2007.

[35] 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.*, vol. 17, pp. 575–79, 2007.

[36] 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*, vol. 208, pp. 331–7, 1998.

[37] 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*, vol. 219, pp. 51–9, 2001.

[38] H. Yoshida and J. Näppi, "Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps," *IEEE Trans. Med. Imag.*, vol. 20, no. 12, pp. 1261–74, 2001.

[39] 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*, vol. 22, no. 4, pp. 963–79, 2002.

[40] 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*, vol. LNCS 3749, pp. 859–67, 2005.

[41] 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.*, vol. 13, no. 9, pp. 1062–71, 2006.

[42] J. Näppi and H. Yoshida, "Feature-guided analysis for reduction of false positives in CAD of polyps for computed tomographic colonography," *Med. Phys.*, vol. 30, no. 7, pp. 1592–601, 2003.

[43] A. Jerebko, S. Lakare, P. Cathier, S. Periaswamy, and L. Bogoni, "Symmetric curvature patterns for colonic polyp detection," in *Proc. MICCAI'06*, vol. LNCS 4191, pp. 169–76, 2006.

[44] 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.*, vol. 32, no. 12, pp. 3602–16, 2005.

[45] P. Sundaram, A. Zomorodian, C. Beaulieu, and S. Napel, "Colon polyp detection using smoothed shape operators: Preliminary results," *Med. Image Anal.*, vol. 12, no. 2, pp. 99–119, 2008.

[46] S. B. Göktürk, C. Tomasi, B. Acar, C. F. Beaulieu, D. S. Paik, R. B. Jeffrey Jr., 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.*, vol. 20, no. 12, pp. 1251–60, 2001.

[47] 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.*, vol. 10, pp. 154–60, 2003.

[48] Y. Zheng, X. Yang, and G. Beddoe, "Reduction of false positives in polyp detection using weighted support vector machines," *Proc. 29th IEEE EMBS*, pp. 4433–36, 2007.

[49] 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.*, vol. 20, no. 9, pp. 886–99, 2001.

[50] 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.*, vol. 33, no. 10, pp. 3814–24, 2006.

[51] L. Bogoni, P. Cathier, M. Dundar, A. Jerebko, S. Lakare, J. Liang, S. Periaswamy, M. E. Baker, and M. Macari, "Computer-aided detection (CAD) for CT colonography: A tool to address a growing need," *The British Journal of Radiology*, vol. 78, pp. 57–62, 2005.

[52] 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.*, vol. 17, no. 3, pp. 662–8, 2007.

[53] 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*, vol. 87, no. 10, pp. 2407–16, 2007.

[54] D. M. J. Tax, *One-class Classification*. PhD thesis, Delft University of Technology, Delft, The Netherlands, June 2001.

[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," *N Engl J Med*, vol. 357, pp. 1403–12, 2007.

[56] Institut für Radiologie, Klinik für Strahlenheilkunde, "Virtuelle Koloskopie." `http://radiologie.charite.de/colon/`. Accessed: 15-04-2013.

[57] C. van Wijk, V. F. van Ravesteijn, F. M. Vos, and L. J. van Vliet, "Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow," *IEEE Trans. Med. Imag.*, vol. 29, no. 3, pp. 688–98, 2010.

[58] G. Iordanescu and R. M. Summers, "Reduction of false positives on the rectal tube in computer-aided detection for CT colonography," *Med. Phys.*, vol. 31, no. 10, pp. 2855–62, 2004.

[59] 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.*, vol. 190, no. 5, pp. 1279–85, 2008.

[60] I. W. O. Serlie, F. M. Vos, H. W. Venema, and L. J. van Vliet, "CT imaging characteristics," tech. rep., 2006.

[61] A. R. Webb, *Statistical Pattern Recognition.* John Wiley & Sons, second ed., 2002.

[62] M. Skurichina, *Stabilizing Weak Classifiers.* PhD thesis, Delft University of Technology, Delft, The Netherlands, October 2001.

[63] M. A. Barish, J. A. Soto, and J. T. Ferrucci, "Consensus on current clinical practice of virtual colonoscopy," *Am. J. Roentgenol.*, vol. 184, pp. 786–92, 2005.

[64] C. Chatfield, *Statistics for Technology.* Chapman & Hall, third ed., 1983.

[65] K. Suzuki, H. Yoshida, J. Näppi, S. Armato, and A. 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.*, vol. 35, no. 2, pp. 694–703, 2008.

[66] 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.*, vol. 55, no. 3, pp. 888–901, 2008.

[67] S. Wang, J. Yao, and R. Summers, "Improved classifier for computer-aided polyp detection in CT colonography by nonlinear dimensionality reduction," *Med. Phys.*, vol. 35, no. 4, pp. 1377–86, 2008.

[68] 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.*, vol. 12, no. 4, pp. 479–86, 2005.

[69] J. 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*, p. 633, 2004.

[70] P. M. Calvert and H. Frucht, "The genetics of colorectal cancer," *Ann. Intern. Med.*, vol. 137, pp. 603–12, 2000.

[71] R. M. Soetikno, T. Kaltenbach, R. V. Rouse, W. Park, A. Maheshwari, T. Sato, S. Matsui, and S. Friedland, "Prevalence of nonpolypoid (flat and depressed) colorectal neoplasms in asymptomatic and symptomatic adults," *JAMA*, vol. 299, no. 9, pp. 1027–35, 2008.

[72] E. Gorgun and J. Church, "Flat adenomas of the large bowel: A single endoscopist study," *Dis Colon Rectum*, vol. 52, no. 5, p. 972, 2009.

[73] V. F. van Ravesteijn, C. van Wijk, F. Vos, R. Truyen, J. F. Peters, and L. J. van Vliet, "Computer aided detection of polyps in CT colonography using logistic regression," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 120–31, 2010.

[74] 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.*, vol. 23, no. 11, pp. 1344–52, 2004.

[75] 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.*, vol. 26, no. 12, pp. 1649–56, 2007.

[76] D. S. Paik, C. F. Beaulieu, G. D. Rubin, B. Acar, R. B. Jeffrey Jr, 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.*, vol. 23, no. 6, pp. 661–75, 2004.

[77] 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*, vol. LNCS 2208, pp. 85–93, 2001.

[78] B. Acar, C. F. Beaulieu, S. B. Göktürk, C. Tomasi, D. S. Paik, R. B. Jeffrey Jr, J. Yee, and S. Napel, "Edge displacement field-based classification for improved detection of polyps in CT colonography," *IEEE Trans. Med. Imag.*, vol. 21, no. 12, pp. 1461–67, 2002.

[79] 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," in *Proc. MICCAI'05*, vol. LNCS 3749, pp. 712–9, 2005.

[80] J. Liu, J. Yao, and R. Summers, "Scale-based scatter correction for computer-aided polyp detection in CT colonography," *Med. Phys.*, vol. 35, no. 12, pp. 5664–71, 2008.

[81] 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*, vol. 7259, p. 725923, 2009.

[82] 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.*, vol. 61, no. 9, pp. 758–63, 2006.

[83] P. Cathier, S. Periaswamy, A. K. Jerebko, M. Dundar, J. Liang, G. Fung, J. Stoeckel, T. Venkata, R. Amara, A. Krishnan, R. B. 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*, vol. 1268 of *International Congress Series*, pp. 978–82, 2004.

[84] S. Taylor, R. Greenhalgh, R. Ilangovan, E. Tam, V. Sahni, D. Burling, J. Zhang, P. Bassett, P. 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*, vol. 247, no. 1, pp. 133–40, 2008.

[85] N. Petrick, M. Haider, R. Summers, S. Yeshwant, L. Brown, E. Iuliano, A. Louie, J. Choi, and P. Pickhardt, "CT colonography with computer-aided detection as a second reader: observer performance study," *Radiology*, vol. 247, no. 1, pp. 148–56, 2008.

[86] 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. M. Bossuyt, and J. Stoker, "Computed tomographic colonography compared with colonoscopy in patients at increased risk for colorectal cancer," *Gastroenterology*, vol. 127, no. 1, pp. 41–8, 2004.

[87] 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*, vol. 21, no. 3, pp. 187–205, 1997.

[88] L. J. van Vliet and P. Verbeek, "Curvature and bending energy in digitized 2D and 3D images," in *Proc. SCIA'93*, pp. 1403–10, 1993.

[89] P. J. Olver, G. Sapiro, and A. Tannenbaum, "Invariant geometric evolutions of surfaces and volumetric smoothing," *SIAM Journal on Applied Mathematics*, vol. 57, no. 1, pp. 176–94, 1997.

[90] G. Huisken, "Flow by mean curvature of convex hypersurfaces into spheres," *J. Differ. Geom.*, vol. 20, pp. 237–68, 1984.

[91] K. A. Brakke, *The Motion of a Surface by Its Mean Curvature.* PhD thesis, Princeton University, 1978.

[92] H. Scharr, *Optimal Operators in Digital Image Processing.* PhD thesis, Ruprecht-Karls-Universität Heidelberg, 2000.

[93] J. P. Thirion and A. Gourdon, "Computing the differential characteristics of isointensity surfaces," *Computer Vision and Image Understanding*, vol. 61, no. 2, pp. 190–202, 1995.

[94] J. van Kan, A. Segal, and F. Vermolen, *Numerical Methods in Scientific Computing.* VSSD, Delft, The Netherlands, ISBN 9071301508, first ed., 2005.

[95] W. Ames, *Nonlinear Partial Differential Equations in Engineering*, vol. 1. New York: Academic Press, 1972.

[96] 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. Graphics*, vol. 12, no. 5, pp. 885–92, 2006.

[97] L. Zhao, C. P. Botha, R. Truyen, F. M. Vos, and F. H. Post, "Efficient seeding and defragmentation of curvature streamlines for colonic polyp detection," in *Proc. SPIE Medical Imaging 2008* (X. Hu and A. Clough, eds.), vol. 6916, 2008.

[98] R. P. W. Duin and E. Pekalska, "Structural inference of sensor-based measurements," in *Proc. S+SSPR'06*, vol. LNCS 4109, pp. 41–55, 2006.

[99] M. W. A. Caan, K. A. Vermeer, L. J. van Vliet, C. B. L. M. Majoie, B. D. Peters, G. J. den Heeten, and F. M. Vos, "Shaving diffusion tensor images in discriminant analysis: A study into schizophrenia," *Med. Image Anal.*, vol. 10, pp. 841–49, 2006.

[100] E. Pekalska and R. P. W. Duin, *The Dissimilarity Representation for Pattern Recognition, Foundations and Applications*. Singapore: World Scientific, 2005.

[101] E. Pekalska and R. P. W. Duin, "Learning with general proximity measures," in *Proc. PRIS'06*, pp. IS15–IS24, 2006.

[102] P. Lefere, S. Gryspeerdt, M. Baekelandt, and B. van Holsbeeck, "Laxative-free CT colonography," *Am. J. Roentgenol.*, vol. 183, pp. 945–8, 2004.

[103] Z. Wang, Z. Liang, X. Li, L. Li, B. Li, D. Eremina, and H. Lu, "An improved electronic colon cleansing method for detection of colonic polyps by virtual colonoscopy," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1635–46, 2006.

[104] J. Näppi and H. Yoshida, "Adaptive correction of the pseudo-enhancement of CT attenuation for fecal-tagging CT colonography," *Med. Image Anal.*, vol. 12, pp. 413–26, 2008.

[105] Y. Cheng, Y. Sato, H. Tanaka, T. Nishii, N. Sugano, H. Nakamura, H. Yoshikawa, S. Wang, and S. Tamura, "Accurate thickness measurement of two adjacent sheet structures in CT images," *IEICE Trans. Inf. & Syst.*, vol. E90-D, no. 1, pp. 271–81, 2007.

[106] S. Prevrhal, J. C. Fox, J. A. Shepherd, and H. K. Genant, "Accuracy of CT-based thickness measurement of thin structures: Modeling of limited spatial resolution in all three dimensions," *Med. Phys.*, vol. 30, no. 1, pp. 1–8, 2003.

[107] J. M. Reinhardt, N. D. D'Souza, and E. A. Hoffman, "Accurate measurement of intrathoracic airways," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 820–7, 1997.

[108] G. J. Streekstra, P. Brascamp, C. van der leij, R. ter Wee, S. D. Strackee, M. Maas, and H. W. Venema, "Cartilage thickness measurement in the sub-millimeter range," in *Proc. MICCAI'04*, vol. LNCS 3217, pp. 950–8, 2004.

[109] G. Kindlmann and J. W. Durkin, "Semi-automatic generation of transfer functions for direct volume rendering," in *Proc. IEEE Symp. Volume Visualization*, pp. 79–86, October 1998.

[110] E. G. McFarland, J. G. Fletcher, P. Pickhardt, A. Dachman, J. Yee, C. H. McCollough, M. Macari, P. Knechtges, M. Zalis, M. Barish, D. H. Kim, K. J. Keysor, C. D. Johnson, and the American College of Radiology, "ACR colon cancer committee white paper: status of CT colonography 2009," *J Am Coll Radiol.*, vol. 6, pp. 756–72, Nov. 2009.

[111] T. R. de Wijkerslooth, M. C. de Haan, E. M. Stoop, P. M. Bossuyt, M. Thomeer, M.-L. Essink-Bot, M. E. van Leerdam, P. Fockens, E. J. Kuipers, J. Stoker, and E. Dekker, "Burden of colonoscopy compared to non-cathartic CT-colonography in a colorectal cancer screening programme: randomised controlled trial," *Gut*, vol. 61, no. 11, pp. 1552–9, 2012.

[112] M. S. Juchems, A. Ernst, P. Johnson, S. Virmani, H.-J. Brambs, and A. J. Aschoff, "Electronic colon-cleansing for CT colonography: Diagnostic performance," *Abdom. Imaging*, vol. 34, pp. 359–64, 2009.

[113] A. H. de Vries, M. H. Liedenbaum, S. Bipat, R. Truyen, I. W. O. Serlie, R. H. Cohen, S. G. van Elderen, O. Kesselring, W. de Monyé, L. te Strake, T. Wiersma, and J. Stoker, "Primary uncleansed 2D versus primary electronically cleansed 3D in limited bowel preparation CT-colonography. is there a difference for novices and experienced readers?," *Eur. Radiol.*, vol. 19, pp. 1939–50, Aug. 2009.

[114] E. Neri, T. Mang, M. Hellstrom, A. Mantarro, L. Faggioni, and C. Bartolozzi, "How to read and report CTC," *Eur. Radiol.*, vol. Epub ahead of print, 2012.

[115] D. Burling, "CT colonography standards," *Clin. Radiol.*, vol. 65, no. 6, pp. 474–80, 2010.

[116] A. H. de Vries, H. W. Venema, J. Florie, C. Y. Nio, and J. Stoker, "Influence of tagged fecal material on detectability of colorectal polyps at CT: phantom study," *Am. J. Roentgenol.*, vol. 191, no. 4, pp. W181–W189, 2008.

[117] I. W. O. Serlie, F. M. Vos, R. Truyen, F. H. Post, J. Stoker, and L. J. van Vliet, "Electronic cleansing for computed tomography (CT) colonography

using a scale-invariant three-material model," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 6, pp. 1306–17, 2010.

[118] S. Lakare, D. Chen, L. Li, A. Kaufman, M. Wax, and Z. Liang, "Electronic colon cleansing using segmentation rays for virtual colonoscopy," in *Proc. SPIE Med. Imag., Physiol. Funct. Multidimens. Images*, pp. 412–8, 2002.

[119] S. Lakare, D. Chen, L. Li, A. Kaufman, M. Wax, and Z. Liang, "Robust colon residue detection using vector quantization based classification for virtual colonoscopy," in *Proc. SPIE Med. Imag., Physiol. Funct. Multidimens. Images*, pp. 515–20, 2003.

[120] M. Zalis, J. Perumpillichira, C. D. Frate, and P. Hahn, "CT colonography: Digital subtraction bowel cleansing with mucosal reconstruction initial observations," *Radiology*, vol. 226, no. 3, pp. 911–7, 2003.

[121] M. Zalis, J. Perumpillichira, and P. Hahn, "Digital subtraction bowel cleansing for CT colonography using morphological and linear filtration methods," *IEEE Trans. Med. Imag.*, vol. 23, no. 11, pp. 1335–43, 2004.

[122] M. Franaszek, R. Summers, P. Pickhardt, and J. Choi, "Hybrid segmentation of colon filled with air and opacified fluid for CT colonography," *IEEE Trans. Med. Imag.*, vol. 25, no. 3, pp. 358–68, 2006.

[123] S. Wang, L. Li, H. Cohen, S. Mankes, J. Chen, and Z. Liang, "An EM approach to MAP solution of segmenting tissue mixture precentages with application to CT-based virtual colonography," *Med. Phys.*, vol. 35, no. 12, pp. 5787–98, 2008.

[124] W. Cai, J.-G. Lee, M. E. Zalis, and H. Yoshida, "Mosaic decomposition: An electronic cleansing method for inhomogeneously tagged regions in noncathartic CT colonography," *IEEE Trans. Med. Imag.*, vol. 30, pp. 559–74, Mar. 2011.

[125] L. Hong, S. Muraki, A. E. Kaufman, D. Bartz, and T. He, "Virtual voyage: interactive navigation in the human colon," in *SIGGRAPH*, pp. 27–34, 1997.

[126] C. F. Beaulieu, R. B. Jeffrey, C. Karadi, D. S. Paik, and S. Napel, "Display modes for CT colonography. part II," *Radiology*, vol. 212, pp. 203–212, 1999.

[127] G. Stamm and H. Nagel, "CT-Expo – a novel program for dose evaluation in CT (in German," *Fortschr Roentgenstr*, vol. 174, pp. 1570–6, 2002.

[128] ICRP, "The 2007 recommendations of the International Commission on Radiological Protection," *Ann. ICRP*, vol. 37, no. 103, pp. 2–4, 2007.

[129] L. J. van Vliet and B. J. H. Verwer, "A contour processing method for fast binary neighbourhood operations," *Pattern Recognition Letters*, vol. 7, no. 1, pp. 27–36, 1988.

[130] J. Hardin and J. Hilbe, *Generalized Estimating Equations*. London: Chapman and Hall/CRC, 2003.

[131] P. J. Pickhardt, A. D. Lee, A. J. Taylor, S. J. Michel, T. Winter, A. Shadid, R. J. Meiners, P. J. Chase, J. L. Hinshaw, J. G. Williams, T. M. Prout, S. H. Husain, and D. H. Kim, "Primary 2D versus primary 3D polyp detection at screening CT colonography," *Am. J. Roentgenol.*, vol. 189, pp. 1451–6, Dec. 2007.

[132] M. Chaparro, J. Gisbert, L. Del Campo, J. Cantero, and J. Maté, "Accuracy of computed tomographic colonography for the detection of polyps and colorectal tumors: a systematic review and meta-analysis," *Digestion*, vol. 80, no. 1, pp. 1–17, 2009.

[133] R. E. van Gelder, J. Florie, C. Y. Nio, S. Jensch, S. W. de Jager, F. M. Vos, H. W. Venema, J. F. Bartelsman, J. B. Reitsma, P. M. Bossuyt, J. S. Laméris, and J. Stoker, "A comparison of primary two- and three-dimensional methods to review CT colonography," *Eur. Radiol.*, vol. 17, no. 5, pp. 1181–92, 2007.

[134] C. D. Johnson, J. G. Fletcher, R. L. MacCarty, J. N. Mandrekar, W. S. Harmsen, P. J. Limburg, and L. A. Wilson, "Effect of slice thickness and primary 2D versus 3D virtual dissection on colorectal lesion detection at CT colonography in 452 asymptomatic adults," *Am. J. Roentgenol.*, vol. 189, no. 3, pp. 672–80, 2007.

[135] D. K. Lenhart, J. Babb, J. Bonavita, D. Kim, E. J. Bini, A. J. Megibow, and M. Macari, "Comparison of a unidirectional panoramic 3D endoluminal interpretation technique to traditional 2D and bidirectional 3D interpretation techniques at CT colonography: preliminary observations," *Clin. Radiol.*, vol. 65, pp. 118–25, Feb. 2010.

[136] C. D. Johnson, M.-H. Chen, A. Y. Toledano, J. P. Heiken, A. Dachman, M. D. Kuo, C. O. Menias, B. Siewert, J. I. Cheema, R. G. Obregon, J. L. Fidler, P. Zimmerman, K. M. Horton, K. Coakley, R. B. Iyer, A. K. Hara, R. A. Halvorsen, G. Casola, J. Yee, B. A. Herman, L. J. Burgart, and P. J. Limburg, "Accuracy of CT colonography for detection of large adenomas and cancers," *N Engl J Med*, vol. 359, no. 12, pp. 1207–17, 2008.

[137] J. Florie, R. van Gelder, M. Schutter, A. van Randen, H. Venema, S. de Jager, V. van de Hulst, A. Prent, S. Bipat, P. Bossuyt, L. Baak, and

J. Stoker, "Feasibility study of computed tomography colonography using limited bowel preparation at normal and low-dose levels study," *Eur. Radiol.*, vol. 17, no. 12, pp. 3112–22, 2007.

[138] W. A. Kalender, *Computed Tomography*. Publicis MCD Verlag, 2000.

[139] OECD, "OECD health data (www.oecd.org)," tech. rep., The Organisation for Economic Co-operation and Development, November 2012.

[140] J. Florie, R. E. Gelder, M. P. Schutter, A. van Randen, H. W. Venema, S. de Jager, V. P. M. van der Hulst, A. Prent, S. Bipat, P. M. M. Bossuyt, L. C. Baak, and J. Stoker, "Feasibiliy study of computed tomography colonography using limited bowel preparation at normal and low-dose levels study," *Eur. Radiol.*, vol. 17, no. 12, pp. 3112–22, 2007.

[141] R. E. Gelder, H. W. Venema, J. Florie, C. Y. Nio, I. W. O. Serlie, M. P. Schutter, J. C. van Rijn, F. M. Vos, A. S. Glas, P. M. M. Bossuyt, J. F. W. Bartelsman, J. S. Laméris, and J. Stoker, "CT colonography: Feasibility of substantial dose reduction – comparison of medium to very low doses in identical patients," *Radiology*, vol. 232, no. 2, pp. 611–20, 2004.

[142] R. M. S. Joemai, J. Geleijns, and W. J. H. Veldkamp, "Development and validation of a low dose simulator for computed tomography," *Eur. Radiol.*, vol. 20, no. 4, pp. 958–66, 2010.

[143] P. Massoumzadeh, S. Don, C. F. Hildebolt, K. T. Bae, and B. R. Whiting, "Validation of CT dose-reduction simulation," *Med. Phys.*, vol. 36, no. 1, pp. 174–89, 2009.

[144] J. R. Mayo, K. P. Whittall, A. N. Leung, T. E. Hartman, C. S. Park, S. L. Primack, G. K. Chambers, M. K. Limkeman, T. L. Toth, and S. H. Fox, "Simulated dose reduction in conventional chest CT: validation study," *Radiology*, vol. 202, no. 2, pp. 453–7, 1997.

[145] D. P. Frush, C. C. Slack, C. L. Hollingsworth, G. S. Bisset, L. F. Donnelly, J. Hsieh, T. Lavin-Wensell, and J. R. Mayo, "Computer-simulated radiation dose reduction for abdominal multidetector CT of pediatric patients," *Am. J. Roentgenol.*, vol. 179, no. 5, pp. 1107–13, 2002.

[146] B. R. Whiting, P. Massoumzadeh, O. A. Earl, J. A. O'Sullivan, D. L. Snyder, and J. F. Williamson, "Properties of preprocessed sinogram data in X-ray computed tomography," *Med. Phys.*, vol. 33, no. 9, pp. 3290–303, 2006.

[147] K. Yang, A. L. C. Kwan, S. Y. Huang, N. J. Packard, and J. M. Boone, "Noise power properties of a cone-beam CT system for breast cancer detection," *Med. Phys.*, vol. 35, no. 12, pp. 5317–27.

[148] J. Ma, Z. Liang, Y. Fan, Y. Liu, J. Huang, W. Chen, and H. Lu, "Variance analysis of X-ray CT sinograms in the presence of electronic noise background," *Med. Phys.*, vol. 39, no. 7, pp. 4051–65, 2012.

[149] M. F. Kijewski and P. F. Judy, "The noise power spectrum of CT images," *Phys. Med. Biol.*, vol. 32, no. 5, pp. 565–75, 1987.

[150] K. Faulkner and B. M. Moores, "Analysis of X-ray computed tomography images using the noise power spectrum and autocorrelation function," *Phys. Med. Biol.*, vol. 29, no. 11, pp. 1343–52.

[151] S. J. Riederer, N. J. Pelc, and D. A. Chesler, "The noise power spectrum in computed X-ray tomography," *Phys. Med. Biol.*, vol. 23, no. 3, pp. 446–54.

[152] A. J. Britten, M. Crotty, A. Kiremidjian, H. Grundy, and E. J. Adam, "The addition of computer simulated noise to investigate radiation dose and image quality in images with spatial correlation of statistical noise: an example application to X-ray CT of the brain," *Br. J. Radiol.*, vol. 77, no. 916), pages = 323–8, year = 2004.

[153] K. L. Boedeker, V. N. Cooper, and M. F. McNitt-Gray, "Application of the noise power spectrum in modern diagnostic MDCT: part I. measurement of noise power spectra and noise equivalent quanta," *Phys. Med. Biol.*, vol. 52, no. 14, pp. 4027–46, 2007.

[154] C. W. Kim and J. H. Kim, "Application of CT simulation technique for virtual ultra-low-dose trial in CT colonography," in *Abdominal Imaging, Computational and Clinical Applications*, vol. 7601 of *Lecture Notes in Computer Science*, pp. 49–57, 2012.

[155] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging.* IEEE Press, 1988.

[156] J. L. Prince, *Medical Imaging, Signals and Systems.* Pearson Prentice Hall, 2006.

[157] S. S. M. Hubbel, J. H., "Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients (version 1.4)." http://physics.nist.gov/xaamdi, 2004. Accessed: 2013, March 3.

[158] A. Wunderlich and F. Noo, "Image covariance and lesion detectability in direct fan-beam X-ray computed tomography," *Phys. Med. Biol.*, vol. 53, no. 10, pp. 2471–93, 2008.

[159] A. Papoulis, *Probability, random variables and stochastic processes.* McGraw-Hill, 1984.

[160] A. Macovski, *Medical Imaging Systems*. Prentice-Hall, 1982.

[161] ImPACTGroup, "Comparitive specifications: 64 slice CT scanners," Tech. Rep. CEP08027, Centre for Evidence-based Purchasing, 2009.

[162] J. Wang, H. Lu, Z. Liang, D. Eremina, G. Zhang, S. Wang, J. Chen, and J. Manzione, "An experimental study on the noise properties of X-ray CT sinogram data in Radon space," *Phys. Med. Biol.*, vol. 53, no. 12, pp. 3327–41, 2008.

Summary

CT colonography (CTC) is a minimally invasive method for detection of colorectal polyps and colon cancer. Limitations of CTC related to the efficiency as well as the sensitivity of radiologists. Additionally, the patient's preparation was considered burdensome and the X-ray radiation that is inherent to the technique increases the risk of cancer induction. In this thesis, computerized techniques from the fields of image processing and pattern recognition are proposed in order to increase the efficiency and the acceptance of CT colonography.

An automated polyp detection method is introduced to assist the radiologist. This computer aided detection (CAD) system can be used as a second reader, since it is highly sensitive and may therefore enhance the observer's sensitivity. Additionally, the reading time may be reduced as only a few false positives are detected. The CAD system finds and classifies polyp candidates based on a measure indicating the amount of protrusion into the colon. Such protruding candidates are found by using a second principal curvature flow algorithm, which makes use of the knowledge about the normal colon shape. For classification of the candidates, a low-complex pattern recognition system is designed which is shown to be highly sensitive as well as robust to data from different medical centers.

Furthermore, an extended electronic cleansing algorithm is proposed that facilitates 3D reading of data from patients adhering to a limited bowel preparation. The electronic cleansing algorithm relies on a preprocessing step using the same principal curvature flow technique that was previously introduced for automated polyp detection. As such, data from patients with a limited bowel preparation can be assessed with an unfolded cube fly-through visualization method, while it does not degradate the radiologist's detection performance.

Lastly, the effect of reduced radiation dose is investigated. Therefore, a technique is developed for simulating low-dose Computed Tomography (CT) scans from reconstructed high-dose images. Essentially, this enables in-silico studies into the minimal dose for a particular diagnostic task. It is used to investigate the effectiveness of the automated polyp detection system when the radiation dose is minimized.

In conclusion, this thesis presents novel techniques and results that open

the way to large-scale screening of colorectal polyps and colon cancer using CT colonography.

Samenvatting

CT colonografie (CTC) is een minimaal invasieve techniek om preventief poliepen en kanker in de darm op te sporen. Het beoordelen van de beelddata dat uit dit onderzoek komt, was een arbeidsintensieve taak en er werden soms poliepen gemist. Daarnaast was de voorbereiding van patienten aanvankelijk belastend en had het onderzoek een hoge stralingsbelasting. In dit proefschrift worden nieuwe technieken onderzocht in de gebieden van beeldverwerking en patroonherkenning met als doel om de efficiëntie en acceptatie van CT colonografie te verbeteren.

Als eerste is een algoritme ontwikkeld om automatisch poliepen te detecteren. Dit algoritme detecteert poliepen en de bevindingen kunnen door de radioloog vervolgens nader geinspecteerd worden. Aangezien het algoritme zeer nauwkeurig is, kan dit de nauwkeurigheid van de radioloog helpen te verbeteren. Ook zou het de benodigde tijd van de radioloog kunnen reduceren. Dit systeem wordt een CAD systeem genoemd. CAD is het acronym van de engelse benaming: computer aided detection. Het CAD systeem vindt en classificeert poliepen op basis van hoeveel ze van de darmwand naar buiten steken. Hierbij wordt gebruik gemaakt van een algoritme dat de tweede principale kromming van het darmoppervlak iteratief reduceert. Dit wordt gedaan omdat het bekend is dat de normale vorm van de darm geen delen bevat waarbij deze kromming positief is. Met andere woorden, de normale darm bevat geen bolvormige uitsteeksels. Voor de uiteindelijke classificatie wordt een patroonherkenningssysteem met lage complexiteit gebruikt, voor welke is aangetoond dat het zowel zeer nauwkeurig is, als dat het robuust is tegen het feit dat de data uit verschillende ziekenhuizen kan komen.

Ook wordt een uitbreiding voorgesteld voor het algoritme dat fecale resten in de beelddata identificeerd en uit het beeld verwijderd. Dit algoritme maakt het mogelijk om de darm uit de beelddata te extraheren en door middel van 3D visualisatie aan de radioloog te presenteren zonder dat het zicht op de darm door fecale resten wordt belemmerd. De uitbreiding op dit algoritme maakt dit nu ook mogelijk voor data afkomstig van patienten die een zeer beperkte patientvoorbereiding hebben ondergaan zonder dat de nauwkeurigheid van de radiologen achteruit gaat. Deze uitbreiding is gebaseerd op een soortgelijk al-

goritme als dat was ontwikkeld voor automatische poliepdetectie.

Tenslotte is er onderzocht wat de invloed van een lagere stralingsdosis is op de nauwkeurigheid van CT colonografie. Hiervoor is een techniek ontwikkeld die lage dosis CT simuleert op basis van beschikbare hoge dosis data, waardoor het mogelijk wordt om in silico studies te doen naar de prestatie van de algoritmes op lage dosis data, zonder patienten op lage doses te hoeven scannen. Deze techniek is gebruikt om een schatting te maken van de prestatie van het CAD systeem wanneer de stralingsdosis wordt verlaagd.

Concluderend, dit proefschrift presenteert nieuwe technieken en resultaten die ertoe bijdragen dat CT colonografie gebruikt kan gaan worden om op grote schaal preventief onderzoek te doen om vroegtijdig poliepen in de darm en darmkanker te ontdekken.

List of Publications

Journal publications

[1] **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.*, vol. 29, no. 1, pp. 120–131, 2010.

[2] C. van Wijk, **V.F. van Ravesteijn,** F.M. Vos, and L.J. van Vliet, "Detection and Segmentation of Colonic Polyps on Implicit Isosurfaces by Second Principal Curvature Flow," *IEEE Trans. Med. Imag.,* vol. 29, no. 3, pp. 688–698, 2010.

[3] M.H. Liedenbaum, M.J. Denters, A.H. de Vries, **V.F. van Ravesteijn**, S. Bipat, F.M. Vos, E. Dekker, and J. Stoker, "Low-fibre Diet in CT Colonography Limited Bowel Preparation: Influence on Image Quality and Patient Acceptance," *AJR,* vol. 195, no. 1, pp. W31–W37, 2010.

[4] M.H. Liedenbaum, M.J. Denters, F.M. Zijta, **V.F. van Ravesteijn**, S. Bipat, F.M. Vos, E. Dekker, and J. Stoker, "Reducing the Oral Contrast Dose in CT Colonography: Evaluation of Faecal Tagging Quality and Patient Acceptance," *Clinical Radiology,* vol. 66*,* pp. 30–37, 2011.

[5] **V.F. van Ravesteijn**, T.N. Boellaard, M.P. van der Paardt, I.W.O. Serlie, M.C. de Haan, J. Stoker, L.J. van Vliet, and F.M. Vos, "Electronic Cleansing for 24-H Limited Bowel Preparation CT Colonography Using Principal Curvature Flow," *IEEE Trans. Biomed. Eng.,* vol. 60, no. 11, pp. 3036–3045, 2013.

[6] R.E. Naziroglu**, V.F. van Ravesteijn**, L.J. van Vliet, G.J. Streekstra, F.M. Vos, "Simulation of Scanner- and Patient-specific Low-dose CT Imaging from Existing CT Images," *submitted.*

Conference proceedings

[7] 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, pp. 471–478, 2006.

[**8**] L. Zhao, **V.F. van Ravesteijn**, C.P. Botha, R. Truyen, F.M. Vos and F.H. Post, "Surface Curvature Line Clustering for Computer-Aided Diagnosis in CT Colonography," in: *Eurographics Workshop on Visual Computing for Biomedicine, VCBM 2008* (C.P. Botha, G. Kindlmann, W.J. Niessen and B. Preim, Proc., Delft, The Netherlands, Oct. 6–7), pp. 53–60, 2008.

[**9**] **V.F. van Ravesteijn**, F.M. Vos, I.W.O. Serlie, R. Truyen, and L.J. van Vliet, "Thin Layer Tissue Classification for Electronic Cleansing of CT Colonography Data," in: *International Conference on Pattern Recognition, ICPR 2008* (Proc. 19th, Tampa, Florida, USA, Dec. 8–11), IEEE Catalog Number: CFP08182, ISBN: 978-1-4244-2175-6, ISSN: 1051-4651, Paper MoBT6.3, 2008.

[**10**] **V.F. van Ravesteijn**, F.M. Vos, and L.J. van Vliet, "Recognition of Protruding Objects in Highly Structured Surroundings by Structural Inference," in: *Scandinavian Conference on Image Analysis, SCIA 2009* (A.B. Salberg, J.Y. Hardeberg and R. Jenssen (eds.), Proc. 16th, Oslo, Norway, Jun. 15–18), LNCS 5575, pp. 41–50, 2009.

[**11**] **V.F. van Ravesteijn**, L. Zhao, C. Botha, F. Post, F.M. Vos, and L.J. van Vliet, "Combining Mesh, Volume, and Streamline Representations for Polyp Detection In CT Colonography," in: *International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009* (Proc. IEEE, Boston, Massachussets, USA, Jun. 28–Jul. 1), IEEE Catalog Number: CFP09BIS-CDR, ISBN: 978-1-4244-3932-4, ISSN: 1945-7936, pp. 907–910, 2009.

Other

[**12**] F.M. Vos, C. van Wijk, **V.F. van Ravesteijn**, I.W.O. Serlie, S.E. Grigorescu, F.H. Post, R. Truyen, J. Stoker, and L.J. van Vliet, "Recent Advances in Techniques for CT Colonography: Electronic Cleansing and CAD," in: *CARS 2008 - CAD: Special Session on Abdominal CAD, Int J CARS*, vol. 3 (Suppl 1), pp. S188–S189, 2008.

[**13**] **V.F. van Ravesteijn**, L.J. van Vliet, and F.M. Vos, "Thin Layer Tissue Classification for Electronic Cleansing of CT Colonography Data," in: *MICCAI 2008 Workshop on Computational and Visualization Challenges*

in the New Era of Virtual Colonoscopy (H. Yoshida (ed.), Proc., New York, USA, Sept. 6), pp. 78–84, 2008.

[14] **V.F. van Ravesteijn**, F.M. Vos, I.W.O. Serlie, and L.J. van Vliet, "Thin Layer Tissue Classification for Electronic Cleansing of CT Colonography Data," in: *Advanced School for Computing and Imaging, ASCI 2008* (G.J.M. Smit, D.H.J. Epema, M.S. Lew (eds.), Proc. 14th Annual Conference, Heijen, The Netherlands, June 11–13), ASCI, Delft, pp. 355–361, 2008.

[15] **V.F. van Ravesteijn**, F.M. Vos, I.W.O. Serlie, and L.J. van Vliet, "Thin Layer Tissue Classification for Electronic Cleansing of CT Colonography Data", *2nd Dutch Biomedical Engineering (BME) Conference 2009*, (Egmond aan Zee, The Netherlands, Jan. 22–23), 2009.

[16] **V.F. van Ravesteijn**, F.M. Vos, and L.J. van Vliet, "Detection of Protruding Objects in Highly Structured Surroundings," in: *Advanced School for Computing and Imaging, ASCI 2009* (Th. Gevers, H.J. Bos and L. Wolters (eds.), Proc. 15th Annual Conference, Zeewolde, The Netherlands, June 3–5), ASCI, Delft, 2009.

[17] **V.F. van Ravesteijn**, C. van Wijk, J. Stoker, L.J. van Vliet, and F.M. Vos, "Recent Advances in Automated Lesion Detection for CT Colonography by Second Order Curvature Flow (Featured Lecture)," in: *Proc. of MICCAI 2010 Workshop: Virtual Colonoscopy & Abdomonal Imaging,* Beijing, China, pp. 13–18, 2010.

[18] **V.F. van Ravesteijn**, "Combining Mesh, Volume and Streamline Representations for Polyp Detection in CT Colonography," in: *Advanced School for Computing and Imaging, ASCI 2010* (T. Kielemann, M. van Kreveld, W. Niessen (eds.), Proc 16th Annual Conference, Veldhoven, The Netherlands, Nov. 1–3), ASCI, Delft, 2010.

[19] **V.F. van Ravesteijn**, "Automated Polyp Detection for CT Colonography," in: *Proc. of 3rd Dutch Biomedical Engineering Conference, BME 2011*, (Egmond aan Zee, The Netherlands, Jan. 20–21), 2011.

[20] T.N. Boellaard, **V.F. van Ravesteijn**, M. van der Paardt, F.M. Vos, L.J. van Vliet, J. Stoker, "Electronic Cleansing for Limited Bowel Preparation CT Colonography Using Minimal Principal Curvature Flow," Radiological Society of North America 2012 Scientific Assembly and Annual Meeting; November 27- December 2, Chicago IL, 2012.