Automatic Colonic Polyp Detection using Curvature Analysis for Standard and Low-Dose CT Data
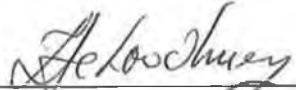
Tarik A. Chowdhury M.Eng.
tarik@eeng.dcu.ie

This thesis is submitted in partial fulfilment of the requirements for the Doctor of Philosophy (Ph.D.) research program at Dublin City University

Supervised by **Prof. Paul F. Whelan**

School of Electronic Engineering
Dublin City University

September 2006

I hereby certify that this material, which I now submit for assessment on the programme of study to the award of Doctor of Philosophy (Ph.D.) is entirely my own work and has not been taken from the work of others save to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _Tarik A. Chowdhury_   ID No.: _501 625 43_

Tarik A. Chowdhury (M.Eng.)

Date: _25/09/2006_

Acknowledgements

Contents

Contents

List of Figures

List of Figures

List of Figures

List of Figures

List of Tables

List of Tables

Abstract

Colon cancer is the second leading cause of cancer related deaths in the developed nations. Early detection and removal of colorectal polyps via screening is the most effective way to reduce *colorectal cancer* (CRC) mortality. *Computed Tomography Colonography* (CTC) or Virtual Colonoscopy (VC) is a rapidly evolving non-invasive technique and the medical community view this medical procedure as an alternative to the standard colonoscopy for the detection of colonic polyps. In *CTC* the first step for automatic polyp detection for 3D visualization of the colon structure and automatic polyp detection addresses the segmentation of the colon lumen. The segmentation of colon lumen is far from a trivial task as in practice many datasets are collapsed due to incorrect patient preparation or blockages caused by residual water/materials left in the colon. In this thesis a robust multi-stage technique for automatic segmentation of the colon is proposed that maximally uses the anatomical model of a generic colon. In this regard, the colon is reconstructed using volume by length analysis, orientation, length, end points, geometrical position in the volumetric data, and gradient of the centreline of each candidate air region detected in the CT data. The proposed method was validated using a total of 151 standard dose (100mAs) and 13 low-dose (13mAs-40mAs) datasets and the collapsed colon surface detection was always higher than 95% with an average of 1.58% extra colonic surface inclusion.

The second major step of automated *CTC* attempts the identification of colorectal polyps. In this thesis a robust method for polyp detection based on surface curvature analysis has been developed and evaluated. The convexity of the segmented colon surface is sampled using the surface normal intersection, Hough transform, 3D histogram, Gaussian distribution, convexity constraint and 3D region growing. For each polyp candidate surface the morphological and statistical features are extracted and the candidate surface is classified as a polyp/fold structure using a Feature Normalized Nearest Neighbourhood classifier. The devised polyp detection scheme entails a low computational overhead (typically takes 3.60 minute per dataset) and shows 100% sensitivity for polyps larger than 10mm, 92% sensitivity for polyps in the range 5 to 10mm and 64.28% sensitivity for polyp smaller than 5mm. The developed technique returns in average 4.01 false positives per dataset.

The patient exposure to ionising radiation is the major concern in using *CTC* as a mass screening technique for colonic polyp detection. A reduction of the radiation dose will increase the level of noise during the acquisition process and as a result the quality of the CT data is degraded. To fully investigate the effect of the low-dose radiation on the performance of automated polyp detection, a phantom has been developed and scanned using different radiation doses. The phantom polyps have realistic shapes (sessile, pedunculated, and flat) and sizes (3 to 20mm) and were designed to closely approximate the real polyps encountered in clinical CT data. Automatic polyp detection shows 100% sensitivity for polyps larger than 10mm and shows 95% sensitivity for polyps in the range 5 to 10mm. The developed method

was applied to CT data acquired at radiation doses between 13 to 40mAs and the experimental results indicate that robust polyp detection can be obtained even at radiation doses as low as 13mAs.

Chapter 1

Introduction

Colon cancer is the second leading cause of cancer-related deaths in the developed nations [1, 2, 3, 4, 5]. Bowel cancer statics in the UK show that 35,600 new cases were diagnosed in 1999 and it is the contributing factor for 16,170 deaths in 2001 [2]. In the United Stated, 130,200 new cases were diagnosed and 56,300 deaths were reported from colon cancer [4] in 2000. In Ireland, 2720 (1232 women and 1488 men) [3] deaths were reported during the period of 1998 to 2000. Statistics in Ireland show [3] that colon cancer was the second leading cause of cancer related deaths for women and the third leading cause of cancer related deaths for men. Greenlee et. al. [4] demonstrate that the probability of developing invasive colon cancer in a life time is 1 in 18(5.55%). In general, colon cancer develops as an intestinal polyp (adenoma) which is an abnormal growth of the colonic tissue. Over time, some of these polyps may become cancerous. Early detection and removal of polyps via screening is the most effective way to reduce the *colorectal cancer* (CRC) mortality [6, 7, 8, 9]. Screening techniques that are available for detection of colonic polyps are *Fecal Occult Blood Test* (FOBT), Flexible Sigmoidoscopy, Barium Enema, Colonoscopy and DNA Stool Test. Colonoscopy is widely considered as the most sensitive method for detection of colonic neoplasia among all above mentioned screening technique [10, 11]. Colonoscopy is performed by experienced gastroenterologists using a colonoscope. The colonoscope is a flexible tube consisting of a camera and a light unit. The gastroenterologist inserts the colonoscope through the anus of the patient and advances it to the end of the colon (cecum). The light and camera in the endoscope allows the gastroenterologist to look into the scope or at a TV monitor. Magnified images allows the gastroenterologist to see any changes in the tissue of the colon. If abnormalities or polyps are found in the colon, the gastroenterologist uses the channels of the endoscope to obtain biopsies (small pieces

of tissue) or remove polyps. The gastroenterologist also uses the enclosure channels to introduce or withdraw fluid or air from the colon. The whole procedure takes about 20 minute to one hour.

Before performing colonoscopy each patient under goes a bowel preparation which includes dietary control and bowel cleansing. In general, the patient should not eat any solid food at least 24 hours before examination. Two methods are commonly used for bowel cleansing. The first involves drinking about one gallon of an undigestible solution that allows bowel cleansing. The second involves taking a solution called "Fleet Phosphosoda" along with several cups of liquid. Prior to the examination, a combination of sedative and narcotic is used to make the patient relax and insensitive to unpleasant sensations. Thus, colonoscopy is an invasive, time consuming, expensive and complex procedure. In conventional colonoscopy, the associated risk of perforation is 0.005% to 0.09% and the mortality risk is 0.001% [12, 13]. In addition to that, conventional colonoscopy fails to reach the cecum which results in an incomplete examination for 5% of the patients [12, 14].

Computed Tomography Colonography (CTC) also known as Virtual Colonoscopy [15, 16, 17, 18, 19] is a rapidly evolving medical imaging technique for the detection of colorectal polyps. In this method, the CT scan of the abdomen of a patient is performed by an experienced radiologist. CT scan images permit interactive viewing of the colon with two-dimensional (2D) and three-dimensional (3D) image display techniques. This imaging method is being widely investigated as a non-invasive examination procedure for the detection of colorectal polyps and many researchers have advocated *CTC* as the optimal mass screening technique to investigate the colorectal cancer [20]. Since the introduction of *CTC*, a significant number of studies have been conducted to evaluate the performance of this non-invasive medical investigation [17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28]. The radiologists perform a visual examination of either two-dimensional CT images or three-dimensional *CTC*, or both. The *CTC* research shows that 3D *CTC* is as good as optical colonoscopy as a screening tool for the early detection of colorectal cancer [29, 30, 31]. The *CTC* exam is less invasive and takes only a few minutes (typically 12-60 minutes). Recent studies [21, 26, 28, 32, 33] demonstrate that *VC* shows a sensitivity in polyp detection comparable to conventional colonoscopy. In 2004, the complete V3D-Colon Workstation obtained the FDA 510(k) market clearance and started to be used in current clinical investigations.

High resolution CT data offers a shear volume of information to a radiologist to visualize and interpret. Typically, a dataset (supine or prone view) has 200-500 slices of images depending on the patient size and reconstruction interval. A reconstruction interval of 1.5mm generates 200-350 images per dataset. Thus, visual evaluation of high resolution CT data is a time consuming procedure and limited by human factors, such as perceptual errors and eye fatigue [34, 35]. The development of *Computer-Aided Detection* (CAD) methods can improve both the sensitivity and efficiency of *CTC*. In recent years, research has focused on developing automated *CAD* systems to detect colonic polyps. Recent studies [35, 36, 37, 38, 39, 40, 41, 42] indicate that *CAD-CTC* offers similar performance as manually analysed *CTC*. The *CAD-CTC* systems need to offer all facilities provided by the VC systems in terms of data visualization and interpretation but primarily they should be designed to detect and report automatically the colorectal polyps larger than 5mm which are clinically significant. It should be noted that the *CAD-CTC* systems can be used off-line by processing the datasets in batch mode. Thus the realtime operation is not a critical requirement for the developed *CAD-CTC* systems. However recent *CAD-CTC* systems are able to process the CT data significantly faster than the radiologist (typical time required by the radiologist to process one dataset is in the range of 12-20 minutes). Currently the *CAD-CTC* systems are used as a second reader that assists the radiologists to improve the quality of the clinical investigation (it is useful to note that there is no agreement between radiologists in regard to the use of the *CAD-CTC* systems in clinical investigations). The high sensitivity of these CAD based polyp detection methods demonstrates that the *CAD-CTC* can be successfully used in clinical studies. Usually the *CAD-CTC* scheme are multistage and the key components are illustrated in Figure 1.1.

Most documented CAD based polyp detection techniques consist of mainly four main phases: Colon segmentation, polyp candidate generation, feature extraction, and classification (see Figure 1.1). Colon segmentation and the surface generation from the volumetric CT data are the primary steps for automatic polyp detection. In CT data, the air insufflated colon provides a high voxel intensity difference between gas (-1000 *Hounsfield Unit*(HU)) and tissue (40HU) surface. Hence, theoretically, the segmentation of the colon can be done using a manually inserted seed point in conjunction with a simple 3D region growing algorithm. But colon segmentation faces two major problems. Firstly, the colon is not the only air filled area inside the
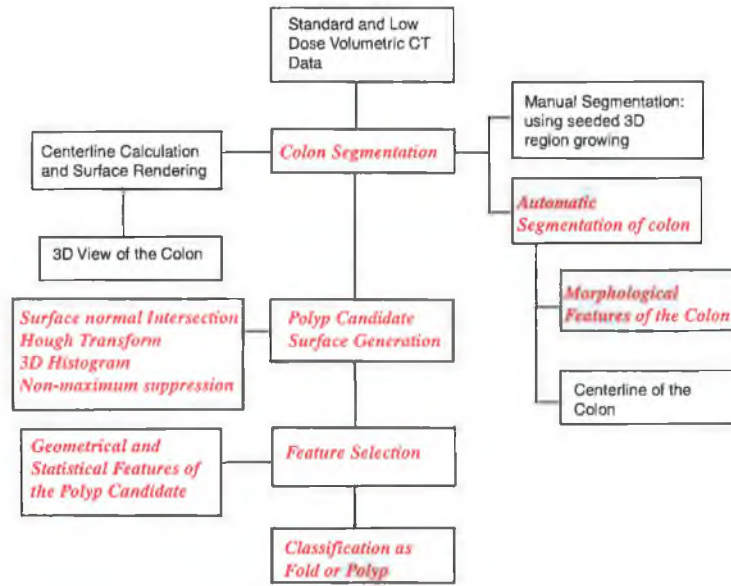
Figure 1.1: Key component of CAD-CTC.

CT data. Secondly, the remaining residual material and water can create multiple collapses in the colon. Thus, the colon segmentation is far from a trivial task and a number of additional processing steps have to be applied in order to perform automatic colon segmentation.

In this research, an automatic segmentation technique of collapsed colons is developed based on the geometrical properties of the colon. Initially, the lung and surrounding air voxels is removed from the volumetric CT data. Then the remaining air voxel regions in the CT data are labelled using a 3D region growing algorithm. The geometrical features like *volume by length* (V/L), orientation, length, end points, position in the volumetric data, and gradient of the centreline of each labelled object are used for colon detection. Figure 1.2 shows the 3D surface of two automatically segmented collapsed colons.

The second step of *CAD-CTC* is the polyp *candidate surface generation* (CSG). The neighbouring voxels of the segmented data having *HU* values higher than a threshold are assigned as the surface of the colon. Figure 1.3a shows a segmented colon and Figure 1.3b illustrate the surface of the segmented colon. In general, polyps are spherical or elliptical in shape and are defined by the abnormal growth of tissue in the inner wall of the colon. Hence, the candidate voxels for polyps have a convex shape compared to the surrounding surface of the colon. But polyps are not the only convex structure, folds have also convex shapes when compared to the

(a)   (b)
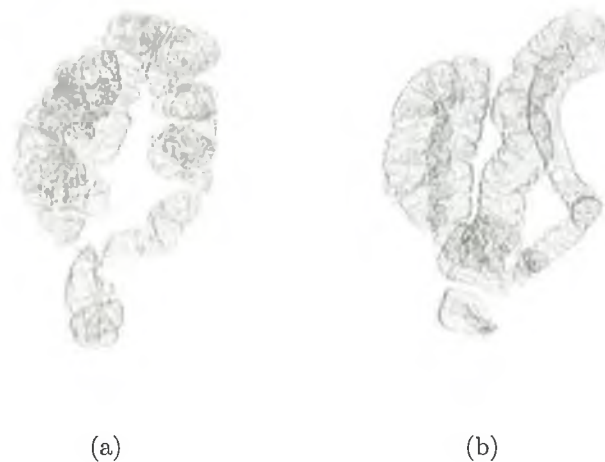
Figure 1.2: 3D surface of two automatically segmented colon.

smooth colon surface. Figure 1.4a shows a CT slice containing a pedunculated polyp. Figures 1.4b, 1.4c show the rendered 3D surface of a polyp and a fold extracted by the candidate surface generation. Therefore, the polyp candidate generation from the colonic wall provides the primary difference between convex and non-convex candidate structures. It is useful to note that a better segmentation of convex surfaces will result in an increased sensitivity and at the same time decreasing the false positive rate incidence in CAD-CTC.

(a)   (b)
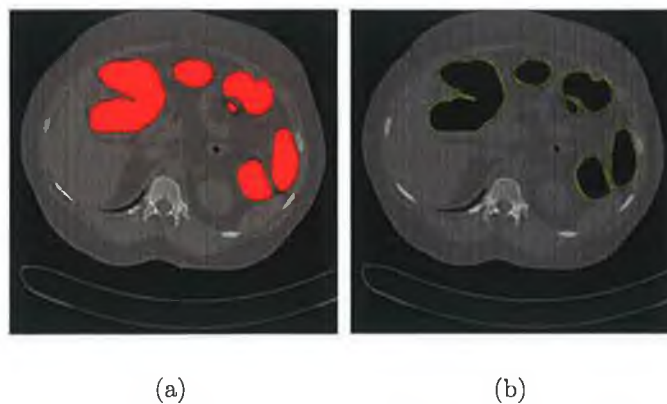
Figure 1.3: (a) Segmented colon. (b) The surface of the segmented colon.

The third step of CAD-CTC is represented by the feature extraction. As indicated above polyps can be typically approximated by spherical or elliptical surfaces whereas the folds can be approximated by cylindrical surfaces. But this geometrical approximation for polyps faces two major problems. Firstly, the size of the polyps

5
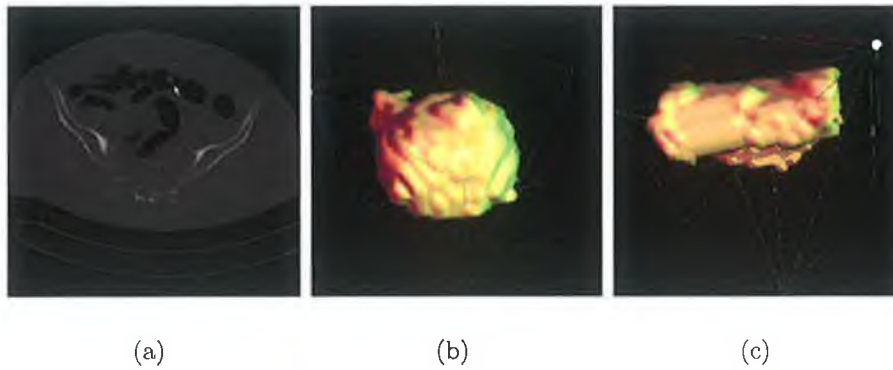
Figure 1.4: Example of polyp and fold in CT data. (a) A pedunculated polyp shown with arrow sign (b) 3D surface of the pedunculated polyp (c) 3D surface of a fold.

varies from 2mm to 20mm or above. Secondly, polyps found in clinical studies have shapes that are significantly different than a spherical/elliptical shape. In general, polyps are classified as sessile, pedunculated (Figure 1.4a), flat and depressed flat. Taken into consideration that the polyps have a large range of shapes and sizes, the calculation and selection of features from candidate surface plays a crucial role in *CAD-CTC*.

The fourth and the last step of *CAD-CTC* is represented by the classification of polyp candidate surfaces into polyps or folds. The selection of an appropriate classifier in conjunction with the appropriate feature normalization scheme are the two main components of the classification process. Several classifiers that are commonly used by different authors *CAD-CTC* are the *Support Vector Machine* (SVM), Neural Network Classifiers and Bayes classifiers.

In the last decade, the introduction of the multi-slice CT scanners boosted the use of CT in the field of medical imaging for the detection of different diseases in human population. The main concern associated with CT investigations is the level of radiation dose associated with modern CT scanners [43, 44]. CT accounts for 4% of the medical radiographic examinations and contributes with 35-40% of the total collective radiation dose received by the patients [45]. Brenner *et al.* [46] stated that the patient exposure to ionising radiation is the major concern in using *CTC* as a mass screening technique for colonic polyp detection. The medical literature indicates that the effective dose used for *CTC* varies from 5 to 20 mSv [32, 47, 48, 49, 50] and this radiation level may result in a 0.02% to 0.05% risk for inducing cancer in patients older than 50 years. It can be concluded that the ionising radiation needs

to be reduced as much as possible in CT colonography. In this regard, by keeping the parameters KVp, and collimation constant, the radiation exposure received by the patient can be reduced by decreasing the tube current (mA). But a low tube current increases the noise in image acquisition and this will require more complex *CAD-CTC* methods to robustly detect the polyps in CT data. To analyze the effect of low dose on colonic polyp detection in CT colonography, a synthetic phantom has been designed and constructed. The phantom (synthetic) polyps were made by latex having a HU value of -90. Polyps have various shapes (Pedunculated, sessile, flat, and flat depressed) and sizes (3mm to 18mm). The phantom was scanned using a 16-slice Somatom Sensation Siemens CT scanner at different radiation doses (mAs) from 100mAs to 13mAs. A statistical polyp detection method was employed to analyze the effect of low-dose on *CAD-CTC* (see Chapter 3). The developed polyp detection algorithm uses the surface normals intersection to calculate the initial center points for candidate polyps. Then the candidate surfaces were created by evaluating the normal distance to the center of the colonic surface. In addition, to remove the non-convex and disjoint surface points, the candidate surface was further processed by 3D region growing, surface normal direction and convexity test. Figure 1.5 presents the 3D surface of two polyps and two folds after candidate surface generation.

(a)   (b)   (c)   (d)
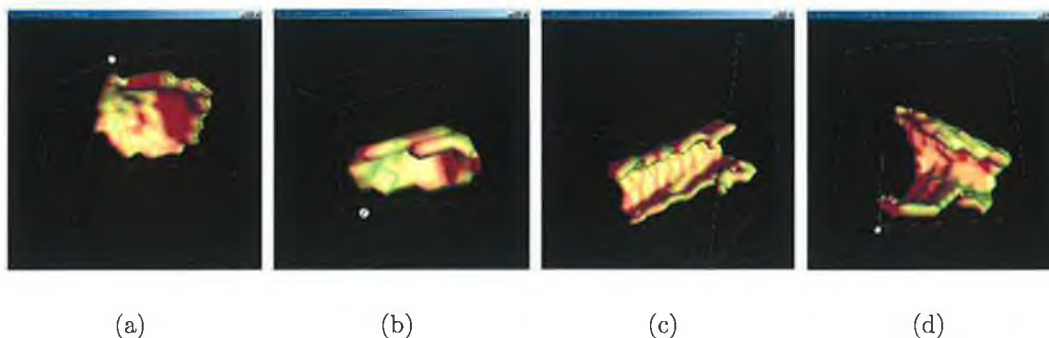
Figure 1.5: 3D surface of polyps and fold after the application of the candidate surface generation (a) and (b) illustrate the 3D surface of two polyps, (c) and (d) depict the 3D surface of two folds.

Since the polyp surfaces are defined by the abnormal growth of tissue in the colon wall and can be approximated by spherical or elliptical shape, the surface normals associated with polyp surfaces show highest concentration in the center of the polyp. On the other hand, folds resemble cylindrical surfaces. Thus, the surface

normals for fold surfaces are evenly distributed along the long axis of the cylinder. In this research work, the geometrical and statistical features evaluate the normal concentration for polyps and folds. Colonic surfaces associated with folds show a higher surface variation when compared with colonic surfaces generated by polyps. To classify the *candidate surface* (CS) as polyp or fold, the surface variation of *CS* is measured using the statistical features (e.g. *Standard Deviation* (SD) of surface number, *SD* of the principle axis of ellipsoid fitting, *SD* of the radius of sphere fitting) that are inputs for a multi-class classifier.

Most authors suggested to divide the polyps into three different classes according to their sizes. First class represents polyps higher or equal than 10mm ($\geq 10mm$). Second class includes polyps between 5 to 10mm ($\geq 5mm \ to \ < 10mm$)). Third class represents polyps smaller than 5mm. This approach was followed in this thesis and the classifiers were trained with polyp and fold surfaces that were segregated by size. The convex structures are classed into polyps and folds using a multi-class *feature normalized nearest neighbourhood*(FNNN) classifier. The developed *CAD-CTC* system has been evaluated on phantom data and patient data that was scanned at normal and low-dose radiation levels.

1.1   Contributions

The main aim of this research is the development of a robust *CAD-CTC* system and to evaluate its performance when applied to standard and low dose CT data. The main contributions resulting from this investigation include:

- The development of an automated technique able to perform robust colon segmentation in *CAD-CTC*. The developed method successfully segmented 146 (96.95%) out of 151 colons with an average of 1.58% extra colonic surface inclusion [51, 52].

- A detailed analysis of the geometrical and statistical features employed for robust classification of polyps and folds. The devised polyp detection scheme shows 100% sensitivity for polyps larger than 10mm, 92% sensitivity for polyps in the range 5 to 10mm and 64.28% sensitivity for polyps smaller than 5mm with an average of 4.01 false positives per dataset [53, 54, 55, 56, 57].

- Design and construction of a phantom to generate synthetic data and evaluate

the effect of low dose on the overall performance of the developed *CAD-CTC* system. It was concluded that a radiation dose as low as 13mAs is feasible to be used in standard *CTC* clinical examinations [58, 59, 60].

- Development of efficient segmentation and classification algorithms suitable for fast polyp detection. The average computation time for polyp detection is 3.60 minute per dataset when the algorithm is run on a standard PC [55].

1.2 Document Organization

Chapter 1 introduces Virtual Colonoscopy and outlines the key components of an automated *CAD-CTC* system. Chapter 2 details the development of an automated colon segmentation method. Chapter 3 describes the algorithm developed for automatic CAD based polyp detection. Chapter 4 discusses the effect of low-dose and CT scanning parameters on the performance of the developed automatic polyp detection scheme. Chapter 5 analyses the performance of the developed *CAD-CTC* system. Finally, Chapter 6 concludes this thesis and discussed the further developments that can be made on the discussed *CAD-CTC* system.

Chapter 2

Segmentation

In *CTC*, the detection of polyps and cancerous lesions depends on the accurate identification of the colon wall and consequently relies heavily on colonic distension and bowel preparation. Currently, two types of bowel preparation are widely used in *CTC*. The first involves a colonic lavage and insufflation with air prior to CT imaging (non-oral contrast enhanced). The second involves colonic lavage and the introduction of an iodinated contrast agent to homogeneously liquefy and opacify the faecal matter prior to air insufflations (oral-contrast enhanced). Most existing automatic colonic surface detection techniques are proposed for the oral-contrast enhanced patient preparation [16, 61, 62] while limited research has addressed the automatic colonic surface detection for the non-oral contrast enhanced patient preparation [63, 64, 65, 66, 67, 68]. The method proposed by Wyatt et al. [63] for automatic segmentation of the colon includes, the removal of surrounding air voxels, distance transformation, seed point selection, labelling of the air voxel area inside of the body, and applying the elongation and location criteria. After the removal of the surrounding air voxels from the CT data, Wyatt et al. employed a distance transform in the binary data to extract the seed points. They used -800HU as threshold for binary image creation. In general the colon is the largest air filled area inside of the abdomen. Therefore, a maximum distance is a good approximation for seed point selection. The detected seed points were used to segment the colon using a 3D region growing algorithm. They also used the elongation criteria for differentiating the large bowel from the stomach. Since the seed point selection was done based on the distance transform, the inclusion of the extra colonic surfaces (small intestine and stomach) were large (28 extra colonic surfaces in 20 datasets) and showed only 40-80% recovery of the colon surface. The method was tested for both oral contrast and without oral contrast patient data and required approximately 60-65 minute to

examine a dataset. Masutani et al.[64] proposed a method to remove the lung tissues, surrounding air voxels, bones from the dataset and then identify the largest air volume as the colon. If collapses appear in the colon, the largest air volume in the CT data was assigned as the colon and the other regions having volume 25% (volume threshold (R_{fc})) of the largest volume were considered as parts of the colon too. As this technique evaluates the air regions only with respect to the R_{fc} threshold, parts of the small bowel can be misinterpreted as part of the colon by the automatic segmentation process. On the other hand, small parts of the collapsed colon may be incorrectly removed. Nappi et al. [65] proposed a different segmentation method that detects the colon as the intersection of the Anatomy Based Extracted (ABE) surface with the Colon Based Extracted (CBE) surface. ABE uses the same volumetric features proposed by Masutani et al.[64]. In CBE method, a 3D region growing was initiated from the rectum and this process continues until a stopping rule that checks for certain experimentally validated conditions is upheld. If the conditions were not met, the region growing process was re-started from an automatically selected new seed point and the stopping rule is re-evaluated. Finally, the intersection surface between ABE and CBE was declared as the colon surface. This method reduced the extra-colonic surface inclusion from 25.6% to 12.6%. Iordanescu et al. [66] proposed an automatic seed placement method using one seed point near the rectum for well-distended colon and two seed points at rectum and cecum for collapsed colon segmentation. Their method has shown that for 83.2% of the datasets the colon segmentation was complete and 9.6% of datsets shows partial colon segmentation. The remaining 7.2% datasets require a manual seeded segmentation. Since, their method used two seed points, the segmentation of the colons with multiple segments (higher than two) requires manual intervention. Li et al. [67] showed a method for automatic seed selection for colon segmentation. The selected seeds are used in conjunction with 3D region growing algorithm for colon segmentation. Their method, initially segment the CT data by applying a 2D region growing algorithm slice by slice. The center point of each 2D segmented area were selected as seed points. Then all the 2D seed points were analysed and refined using shape and size based filters. The filtered 2D points are used to segment the colon using a 3D region algorithm. Their method results in 87.5% colon surface coverage with 6% extra colonic surface inclusion. The proposed shape and size filters can create similar results for collapsed colons and small intestine, hence automatic

segmentation can excludes part of the colon or can include extra colonic surface
or both. Frimmel et al. [68] method uses the centerline and the colon geometry
for automatic segmentation. After the removal of the surrounding air voxels from
the CT data, their method calculate the centerline of air filled regions inside the
abdomen. They calculated the bounding box parameters for each centerline and
used some predefined thresholds to accept or reject the centerline section derived
from the small intestine. Their method shows 96% sensitivity for automatic colon
segmentation.

All of the above segmentation techniques discussed above show different levels
of accuracy and indicate that further investigations are needed in order to obtain
a robust technique for automatic segmentation of collapsed colons especially for
non-oral contrast-enhanced patient preparation. In this chapter a novel method for
automatic segmentation of collapsed colon lumen based on a prior knowledge of the
colon geometrical features and anatomical structure is proposed.

2.1   Materials and method

Prior to their scheduled examination all patients were instructed to take a low-
residue diet for 48 hours followed by clear fluids for 24 hours. Prior to the day
of examination, patients were prescribed one sachet of Pixcolax at 8.00, a second
sachet of Pixcolax at 12.00, a sachet of clean prep in a litre of cold water at 18.00
and a Senokot tablet at 23.00. Before the CT scan, a rectal tube is inserted and
the colon is gently insufflated with room air at the maximum level tolerated by the
patient. All scans were obtained on a commercially available Siemens Somatom 4
slice multidetector Spiral CT scanner. The scanning parameters were 120kVp, stan-
dard dose (100mAs) and low dose (13mAs-40mAs) effective tube current, 2.5mm
collimation, 3mm slice thickness, 1.5mm reconstruction interval, and 0.5s gantry ro-
tation. The data acquisition procedure takes from 10 to 30s, hence, CT acquisitions
were performed in a single breath-hold. The procedure was first performed with
the patient head first supine position and then repeated with the patient head first
prone position. The number of slices varies from 200-350 depending on the height
of the patient. Typical total size of the volumetric data is approximately 150MB.

2.2   Automated Segmentation of Collapsed Colon

In *CTC*, the presence of high contrast gas/tissue interface in the air insufflated colon makes the segmentation of the colon lumen a relatively simple task. However, the automatic segmentation of the entire colon has to address two major problems. Firstly, in CT data the colon is not the only gas filled organ, it also includes the gas filled lung, stomach and small bowel. In particular the small bowel may confuse the automatic colon segmentation process. Secondly, obstructions can occur in the colon itself due to peristalsis, residual faeces, water and insufficient air insufflation. Such obstructions can create multiple collapses in the colon and the complexity of the automatic colon segmentation is significantly increased. Figure 2.1 shows the overview of the proposed algorithm. The proposed method initially removes the surrounding air voxels and lung tissues from the volumetric CT data while the next step identifies and labels all remaining air regions in the volumetric data. Volume by length (V/L) analysis, orientation, length, end points, geometrical position in the volumetric data, and gradient of centreline of each labelled object were used as geometrical features for automatic colon segmentation. Consequently, the proposed automatic segmentation technique includes the outer air segmentation, lung segmentation, labelling, *V/L* analysis, and gradient of centreline calculations.
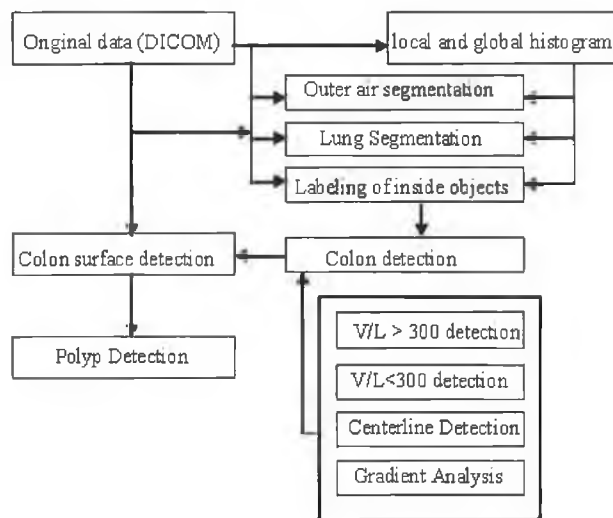
Figure 2.1: Overview of the proposed colon segmentation algorithm.

2.2.1   Surrounding Air Voxel Removal

Colon detection begins with the removal of surrounding air voxels that was performed using a standard seeded 3D region growing algorithm [69]. The seed points for 3D region growing were selected as the left and right-most column voxels from the first slice of the volumetric data. The threshold (T_a) employed to evaluate the similarity measure for region growing was automatically selected from the global histogram. The second peak (T_a) of the global histogram as illustrated in Figure 2.2 was algorithmically detected and used as the threshold for region growing.
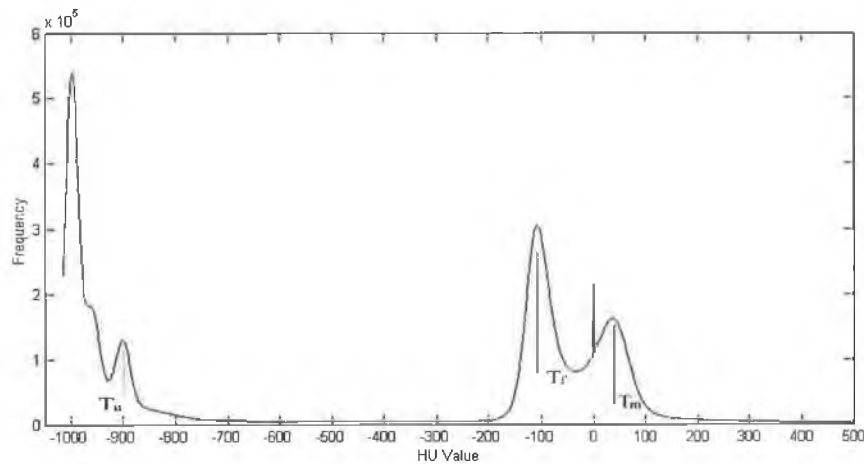
Figure 2.2: Global histogram of the CT volumetric data. T_a is the threshold used for 3D region growing. T_f and T_m represents the histogram peaks for fat and lean tissue respectively.

2.2.2   Lung Detection

In all head first supine or prone volumetric CT datasets the lungs are always visible in the first slice. Consequently, after the removal of the surrounding air voxel, the 3D region growing process starting in the first slice of the volumetric data will segment the lung tissues. To detect the lungs, the algorithm described in this chapter checks for the presence of isolated blood vessels inside the segmented area (see Figure 2.3). If multiple isolated blood vessels are detected, the segmented area is considered to be lung tissue; otherwise it is defined as a candidate region of the colon structure. Based on the analysis of the local histogram of 25 datasets (calculated from the first five slices of the volumetric data), it was found that a threshold greater than -800HU returns the best segmentation for blood vessels from the surrounding lung air. Hence the segmentation threshold was set to -800HU.

<div align="center">(a)   (b)</div>
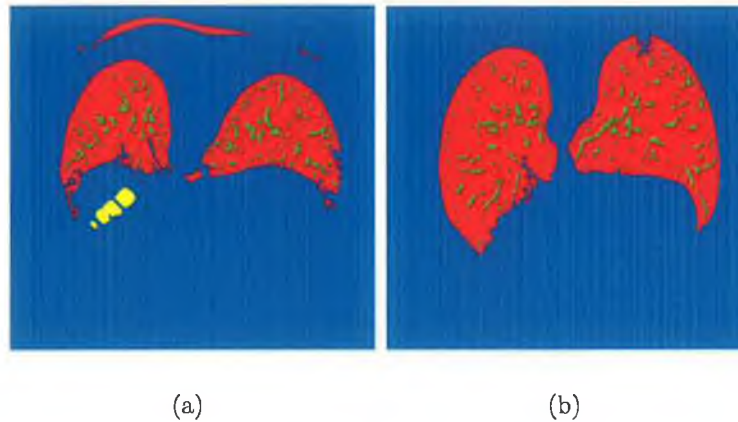
Figure 2.3: Detected lung from the first slice of the volumetric data. (a) Parts of colon (in yellow) and lung. (b) Detected lung (in red). Vessels are marked in green in both (a) and (b) (Results best viewed in colour).

2.2.3   Labelling the Inside Area

Once the lungs have been segmented, the remaining air regions are the colon, small intestine and stomach. In this step, labelling was performed using a 42/46-neighbourhood structured element 3D region growing algorithm (see Figure 2.4). The 42 neighbourhood region growing was used if the voxel width or height was higher than 0.611mm, otherwise a 46 neighbourhood was used (in general, voxel dimensions are: depth 1.5mm, width and height 0.50-0.90mm). The 42/46 neighbourhoods are used to make the region growing approximately isometric. The threshold for region growing was automatically selected from the global histogram and is usually in the range -800HU to -900HU. The labelling of the air regions was performed in two phases. In the first phase, any air voxels (less than the threshold) in the volumetric data initiates the region growing and continue to label all the connected air voxels. The region growing process will stop when no neighbouring voxel with HU values less than the threshold are found. The last voxel where the region growing stopped was considered as the *first end point* (FEP) (see Figure 2.5a) of the labelled region. In the second phase, the region growing process starts from the first end point and labels all the voxels in the region that are already checked in the first phase of labelling. At the end of the second phase, the last voxel where the region growing stopped was considered as the *second end point* (SEP) (see Figure 2.5b). Similarly, all the air regions in the volumetric data will be labelled two times to calculate the two end points. During the labelling process the following information

is also stored for each labelled region: total voxel count, flag value, average *HU* value, region bounding box coordinates and orientation.
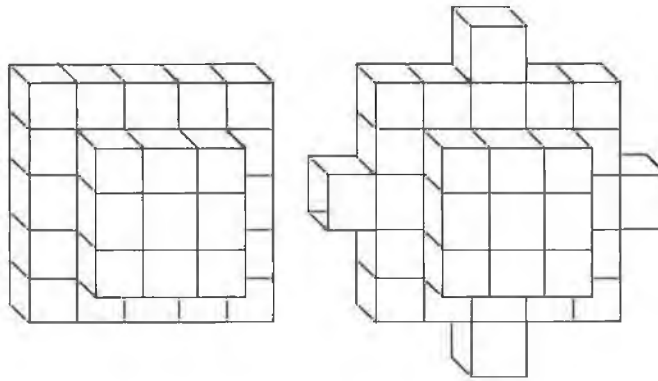
Figure 2.4: Seeds used for region growing. (a) 42 voxels seed. (b) 46 voxels seed.
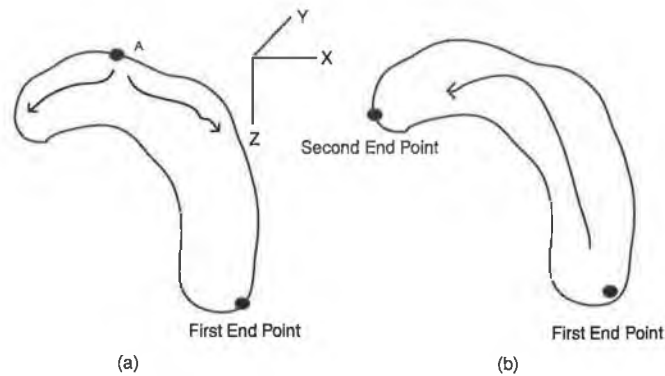
Figure 2.5: Labelling using 3D region growing. (a) Detection of first end point. (b) Detection of second end point.

The *high end point* (HEP) illustrated in Figure 2.6a is detected as the *FEP* at the end of first phase of labelling. In the second phase, region growing starts from the *HEP* and the algorithm is iterated until the *second end point* (SEP) is detected. In this situation, the labelled region will be assigned as $ORIENT\_2$. Similarly, the orientation index for each labelled region is recorded as $ORIENT\_1$, $ORIENT\_3$ or $ORIENT\_4$ as illustrated in Figures 2.6a, 2.6c, 2.6d. It is worth noting that for a well-distended colon supine view the orientation $ORIENT\_4$ will never occur.

2.2.4   Colon Detection

The colon and the small intestine are approximately 1.5m and 7-10m long, respectively [70]. Anatomy of the colon shows that it is shorter and thicker than the
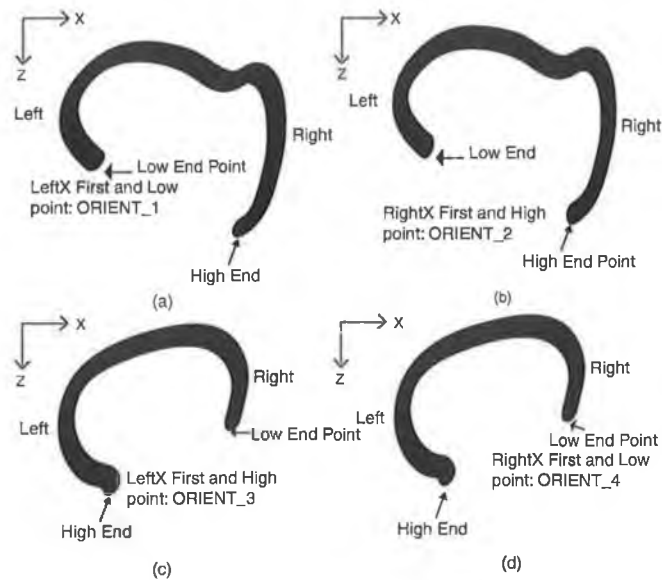
Figure 2.6: The four possible orientations used to differentiate a well-distended colon and a collapsed colon.

small intestine. The volume of each labelled region is calculated using the following equation:

$$Volume = vx * vy * vz * n \tag{2.2.1}$$

where vx, vy and vz are the voxel width, height and depth respectively and n is total number of voxels in the region.

The length of each labelled region was calculated between the two end points using the Dijkstra shortest path algorithm [71]. In Figure 2.7 the V/L analysis for 35 datasets is illustrated. The upper and lower curves represent the colon and small intestine V/L values respectively. As the small bowel is long and thin when compared to the large bowel, the V/L analysis provides a distinctive feature for automatic colon detection.

In general, the V/L value for a well-distended colon is higher than $600mm^2$. To provide a high degree of tolerance in V/L threshold, it was determined experimentally that a well-distended colon must have a V/L value higher than $300mm^2$. The results of developed method indicate that this V/L threshold was robust when the segmentation algorithm was applied to a large number of datasets.

2.2.5   Well Distended (Intact) Colon Detection

The devised algorithm firstly checks whether the colon has a collapsed segment or not. The algorithm is initiated with the detection of the rectum. In general, the

17
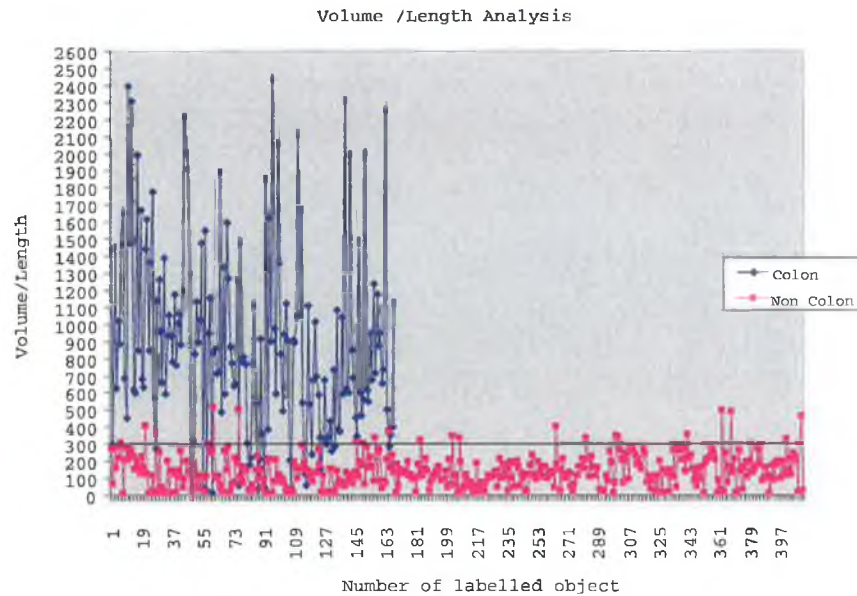
Figure 2.7: Volume/Length analysis provides a distinctive feature to differentiate the colon from small bowel.

rectum is the only air filled area that is located at the end of the dataset. To make sure that the selected object is inside the body, the voxels located around its neighbourhood are tested within a circular region of interest. The colon will be declared as well distended if the selected rectum object fulfills certain conditions.

a) The detected rectum must have a V/L value higher than $300mm^2$, a length higher than 700mm. The detected rectum segments with a length higher than 700mm indicate that are connected with parts of the sigmoid colon and descending colon. In this condition, if collapses appear between transverse and descending colon or transverse and ascending colon, the V/L value of the rectum will be less than the V/L value of the ascending colon and consequently the colon will be assumed to be collapsed.

b) Detected rectum object orientation number (Figure 2.6) must not be $ORIENT\_4$ for supine data and $ORIENT\_1$ for prone data.

c) Validation of the colon geometry. Projection of the well-distended model for colon in the XZ plane is depicted in Figure 2.8. The geometrical approximation of

the colon was calculated dynamically from the labelled regions ($V/L > 100mm^2$) coordinates (left most, right most, top, bottom, front and back) as indicated in Figure 2.8. Sometimes collapses in the sigmoid colon can create colon objects with a V/L value less than $300mm^2$ and to include all these objects the threshold was fixed to $100mm^2$. For a well-distended colon, the air region detected as rectum will have one end point near the rectum and other end point (cecum point) closer to the cecum. To declare the detected labelled air region as well-distended colon, parts of it must fulfill the ascending and descending colon geometry (see Figure 2.8). Otherwise it will be declared as collapsed colon.
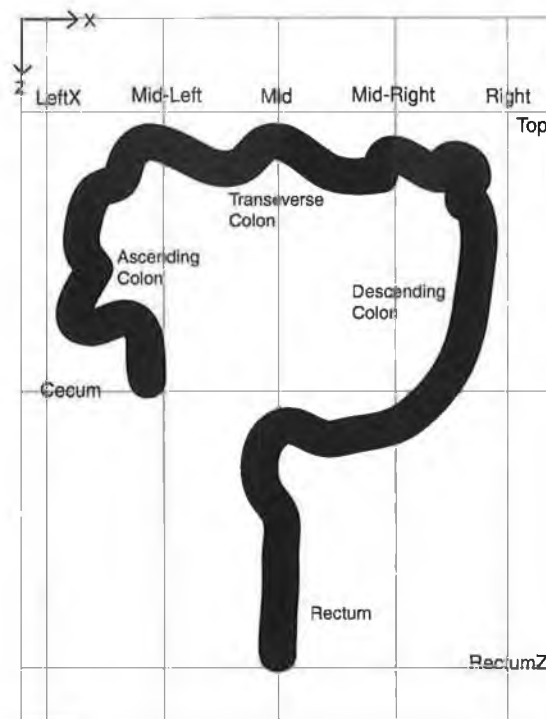
Figure 2.8: An ideal model for a well-distended colon.

2.2.6 Collapsed Colon Detection

Collapsed colon detection is performed in two phases. In the first phase, the large segments (with $V/L > 300$) are detected and in the second phase the small objects (with $V/L < 300$) are detected. The detection of the large segments starts from the rectum. It detects the closest placed large segments using the Euclidean distance between the end points (Figure 2.9) and checks for the condition (a) and (c) which are detailed in Section 2.2.5. This process continues until the conditions (a) and (c) are met. It is worth noting that in some cases detection of the ascending colon appears

after the rectum (clockwise detection) as depicted in Figure 2.9b. This condition occurs if large parts of the descending and sigmoid colon are filled with residual material. In this situation the V/L threshold is automatically changed to $200mm^2$ to meet the geometrical condition (c) detailed in Section 2.2.5. Small-labelled areas (with $V/L < 300$) are either part of the small intestine or the colon. As their anatomical and geometrical properties are quite similar, perfect colon identification is far from a trivial task. The segmentation scheme detailed in this chapter analyses the small segments (with $V/L < 300$) using their position, gradient of the centreline, length, and distance.
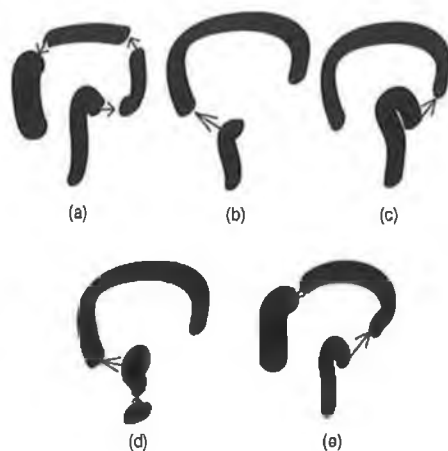
Figure 2.9: Examples for object detection ($V/L > 300$) in collapsed colon for supine data. a) Four objects with $V/L > 300$ in expected direction. b) Two objects with $V/L > 300$ in anti clockwise direction. c) Two objects with $V/L > 300$ in expected direction. d) Anti clockwise direction occurred in the third object. e) Object detected in expected direction.

Centreline and gradient detection: Initially the centrelines of each labelled air region were detected using the method described by Sadleir and Whelan [72]. To reduce the noise in the centreline detection, a three-step procedure was employed [73]. Firstly, a second order low pass filter was applied to remove the high frequency components of the centreline. Secondly, the filtered centreline was down-sampled (typically by a factor of seven). Thirdly, three cubic B-spline interpolations were constructed for the resulting down-sampled set of points (one for each of the three orthogonal directions). The gradient of the centreline was calculated using the first derivative of the interpolated centreline. The *Grad/Num* feature was calculated using the following equation:

$$Grad/Num = 1/n \sum grad \qquad (2.2.2)$$

where *grad* is the gradient in each voxel and n is the number of voxel in the centreline.

Figure 2.10: Example of a colon with three large labelled objects ($V/L > 300$) and few small segments either part of colon or small bowel

Figure 2.10 shows three large segments (with $V/L > 300$) and few small segments (with $V/L < 300$) which are either part of the colon or small bowel. The first step detects the small objects placed between the large segments 1 and 2 depicted in Figure 2.10, and in the second step the small objects between the large segments 2 and 3 will be detected. Detection of these small segments has been done using four different parameters (distance threshold, orientation, length threshold, gradient of centerline) as follows:

(a)   (b)   (c)   (d)
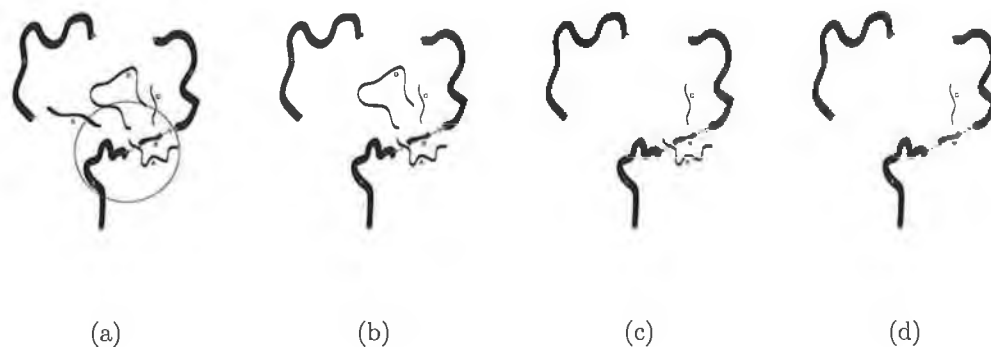
Figure 2.11: Example of removing small intestine. (a) Rejected two small intestine (F and G) due to distance threshold (b) One small intestine (E) rejected due to improper orientation (c) Small intestine (D) removed using the length threshold and (d) Small object with high curvature A is rejected by the gradient threshold.

a) Distance Threshold: The small object ($V/L < 300$) with one or both end points within the circular region of interest illustrated in Figure 2.11a will be detected as candidate colon object. Also, all the small objects ($V/L < 300$) with the exception of F and G in Figure 2.10 will be accepted as colon object (Figure 2.11a).

b) Orientation: The small object E shown in Figure 2.11a have one end point near to the large segment 1 and other end point near to the large segment 3. So, its location violate the geometrical constraint and will be rejected (Figure 2.11b).

c) Length Threshold: The small object D in Figure 2.11b will be rejected because do not pass the length threshold test (see Figure 2.11c). The length threshold was set as twice the distance of the large segments. For instance if the large segments 1 and 2 shown in Figure 2.11a have a distance between the endpoints equal to 50mm, the length threshold will be set to 100mm.

d) Gradient Threshold: As the geometry of the small bowel shows high degree of curvature when compared to the curvature of the colon, the $Grad/Num$ of the small bowel have a higher value than the $Grad/Num$ value of the small colon parts. If the detected small segment has a $Grad/Num$ value higher than a threshold is rejected and declared as part of the small bowel (see Figure 2.11d).

2.3 Results

The segmentation was performed on 151 standard dose (100mAs) and 13 low dose (13mAs to 40mAs) supine and prone patient datasets (87 patient datasets, see Table 2.1). The proposed automatic segmentation method reliably detected 63 standard dose (100mAs) and 5 low-dose well-distended colons without inclusion of any Extra Colonic Surface (ECS) areas. Consequently the colon surface detection was 100% and the ECS error was 0%. The detection of the collapsed colons was performed in several phases. The detection of large segments (with $V/L > 300$) was performed in the first phase and 219 air regions for standard dose were detected in 83 datasets and for low-dose data 20 large segments were detected in 8 datasets. Out of these 239 regions, 238 were colon parts and one was a section of the small bowel. The detection of small regions ($V/L < 300$) was done in the second phase. In total 349 ($V/L < 300$) small air regions were detected in 83 standard dose datasets of which 161 were colon surfaces and 188 were $ECSs$ and 12 colon objects were missed. In 8 low-dose datasets 31 ($V/L < 300$) small objects were detected of which

10 were colon surface and 21 were *ECSs* and 4 colon surfaces were missed. In 83 standard dose collapsed colons, the surface detection was always higher than 95% (see Figure 2.13). Only in five cases it was less than 98% and in three cases it was (98% to 99%). For the remaining datasets the recovery of colon surface was higher than 99% out of 83 collapsed colons. The largest *ECS* inclusion was 14.26% with a mean of 1.58%. In 8 low-dose collapsed colons, five shows 0% *ECSs* surface inclusion with 100% colonic surface detection, one shows 6.5% *ECSs* inclusion, one shows 4.7% *ECSs* inclusion with 96.3% of the colon surface detected and the other one was detected as intact colon and missed 13% of the colonic surface (see Figure 2.14).

Table 2.1: Patient data information

Dose in (mAs)	Number of supine data	Number of prone data
100	81	70
50	0	1
40	1	0
30	1	3
20	0	2
13	0	5
Data Error	4	6
Total	87	87
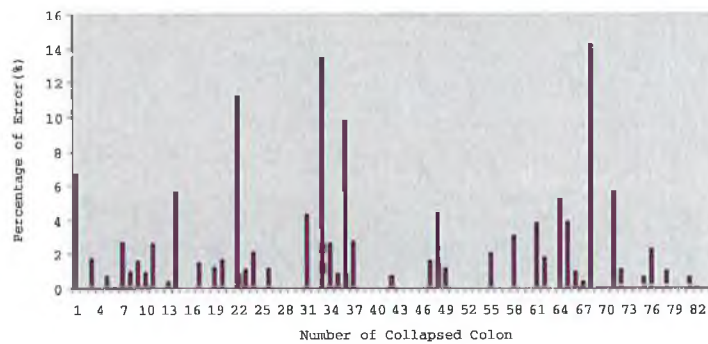
Figure 2.12: Percentage error for *ECS* in 83 collapsed colons. 40 collapsed colons without inclusion of *ECS*.

To examine the performance of automatic colon segmented algorithm, an experienced radiologist from Mater Hospital, Dublin performed a manually seeded segmentation and it was used as the ground truth data. Since the manual mark-

Error Percentage of Colonic Surface Missing in Collapsed Colon
Detection
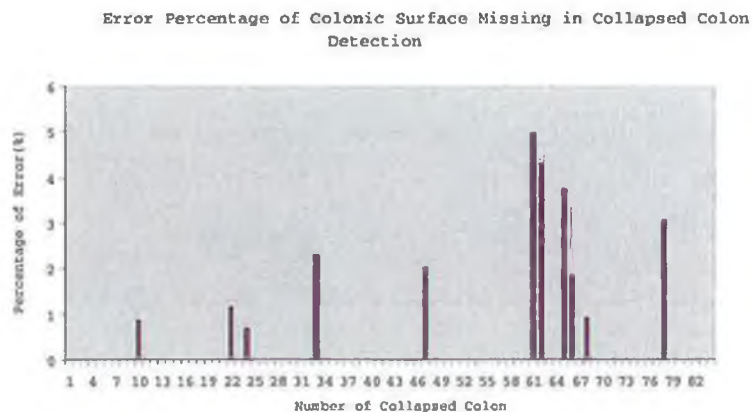
Figure 2.13: Percentage error for undetected colonic surface in 83 collapsed colons.

(a)   (b)

Figure 2.14: (a) 3D surface of the colon segmented by the Radiologist. (b) 3D surface of the colon after automatic segmentation.

ing of the colon area was not feasible due to the enormous amount of data to be analyzed, it has been decided to segment each colon segment individually using a standard seeded region growing algorithm. In this way the radiologist segmented the colon manually using multiple seed points and 3D region growing. To further improve the quality of the ground truth we plan to eliminate the inter and intra observer variability by involving more radiologists in the manual segmentation process. In this research the results returned by the automatic segmentation were compared with the ground truth data side by side using 2D axial views. Any area which was seen in the automatic segmented colon but not found in manually segmented colon was declared as *ECS* and any area excluded by the automated segmentation method was considered as missing. Thus 188 objects ($V/L < 300$) were declared as *ECS's* out of 349 objects ($V/L < 300$) for standard dose datasets and 21 objects were

24

declared as as *ECS's* out of 31 objects ($V/L < 300$) for low-dose datasets.

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)
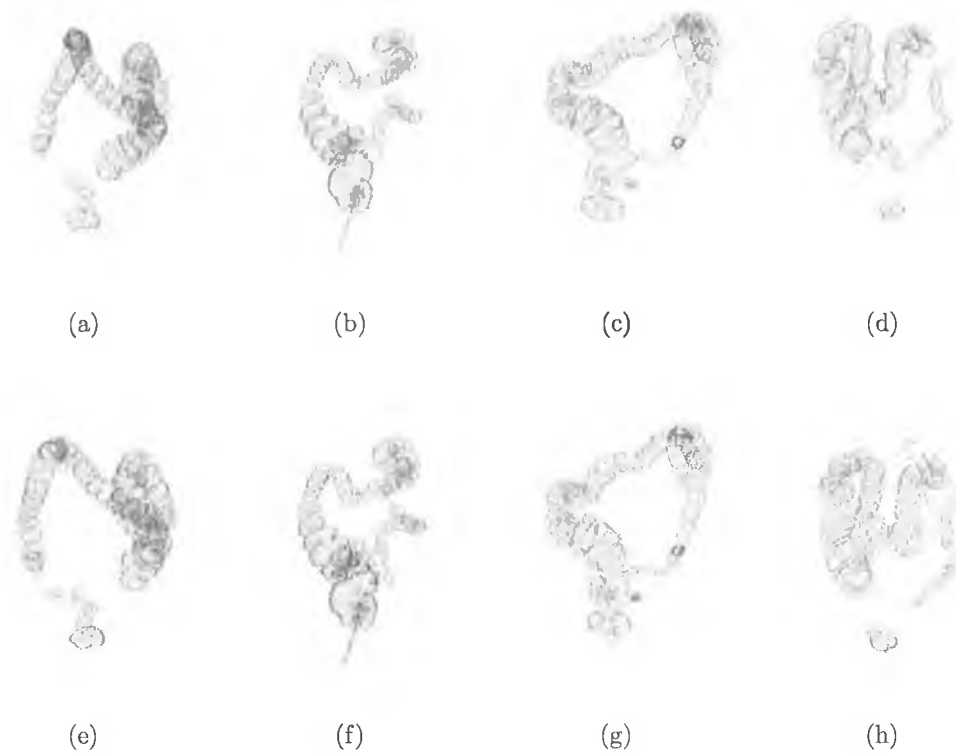
Figure 2.15: Example of manual and automatic segmentation. Figures (a-d) show the 3D surface of the colon segmented by the radiologist. Figures (e-h) show the automatically segmented colon surfaces.

The method proposed by Nappi et al. [65] shows an average of 12.5% *ECS* inclusion with a mean of 0.9% undetected colonic surface which are higher than the proposed algorithm average *ECS* (1.58%) inclusion and mean undetected colonic surface (0.32%). Iordanescu et al. [66] method shows 83.2% success rate for complete automatic segmentation of colons. The developed method provides 94.79% success rate for automatic segmentation of collapsed colon when applied to 96 datasets. Overall sensitivity of automatic colon segmentation is 96.95% in 164 datasets.

The proposed algorithm fails to produce meaningful results when applied to 5 out of 164 datasets due to inappropriate bowel distension (more than 50% of the colon area was filled with fluid and/or residual materials see Figure 2.17). Another advantage of the proposed technique is its low computational cost where the typical processing time for overall segmentation was approximately 3.4min (see Table 2.2) on a Pentium IV 1.6GHz PC with 1024MB RAM.

Table 2.2: Average computation time (in seconds) for well colon segmentation (collapsed and well distended).

Surrounding air voxel removal time	28.65
Lung detection time	13.03
Labelling time	128.32
Length detection (shortest path) time	19.63
Centerline calculation time	15.74
overall time (seconds)	205.37

(a)   (b)   (c)   (d)   (e)

(f)   (g)   (h)   (i)   (j)
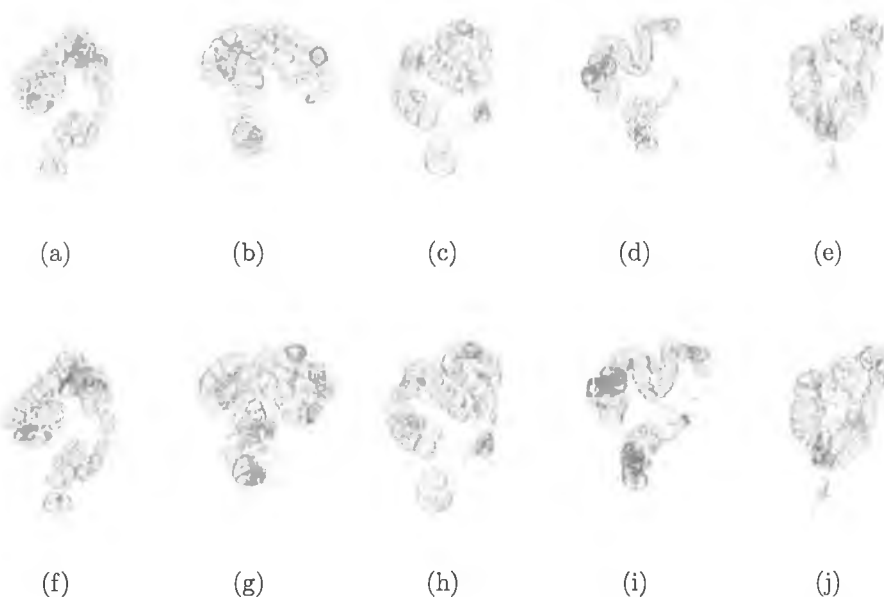
Figure 2.16: Example of manual and automatic segmentation. Figures (a-e) show the 3D surfaces of the colons segmented by the radiologist. Figures (f-j) show the 3D surfaces of colons segmented by the automatic segmentation.

2.4   Discussion

The experimental data indicates that the segmentation algorithm detailed in this chapter returns reliable colon segmentation under all routinely encountered imaging conditions. Well-distended colons have been detected without any inclusion of the small bowel (see Figures 2.18, 2.19 and Table 2.3). When dealing with collapsed colons, the detection of surfaces with a $V/L > 300$ have generated only one false positive in all datasets used in this study (Figures 2.20, 2.21, and 2.22). Small section areas (with $V/L < 300$) include the colon and the small intestines and the final results indicate an average of 1.58% and 1.41% ECS surface inclusion for standard and low dose data respectively and average of 99.68% and 96.52% colon surface

Figure 2.17: Example of a poorly distended collapsed colon.

detection for standard and low dose patient data respectively. Results show that the developed method reliably detects well-distended colons and the large segments in collapsed datasets with a low *ECS* inclusion in small segments detection. Tables 2.3 and 2.4 illustrate the performance of the automatic segmentation technique when applied to standard and low dose datasets.

Table 2.3: Results for standard dose patient data

Rating	Criteria	Results	
		Collapsed colon	*Intact Colon*
Excellent	Includes no small intestine and entire colon segmented	41(49.39%)	63
Good	Includes a small part of the small intestine and segmented entire colon	25(30.12%)	0
Fair	Includes large part of small intestine > 10%	6(7.22%)	0
Poor	Includes small intestine < 5% and missing colon	10(12.04%)	0
Very Poor	Missing >= 5% and includes small intestine	1(1.20%)	0
Total		83	63

In this chapter, a novel scheme for automatic segmentation of collapsed colon is

Table 2.4: Results for low dose patient data

Rating	Criteria	Results	
		Collapsed colon	*Intact Colon*
Excellent	Includes no small intestine and entire colon segmented	5(62.5%)	5
Good	Includes a small part of the small intestine and segmented entire colon	0(0%)	0
Fair	Includes large part of small intestine > 10%	1(12.5%)	0
Poor	Includes small intestine < 5% and missing colon	1(12.5%)	0
Very Poor	Missing >= 5% and includes small intestine	1(12.5%)	0
Total		8	5

(a)   (b)
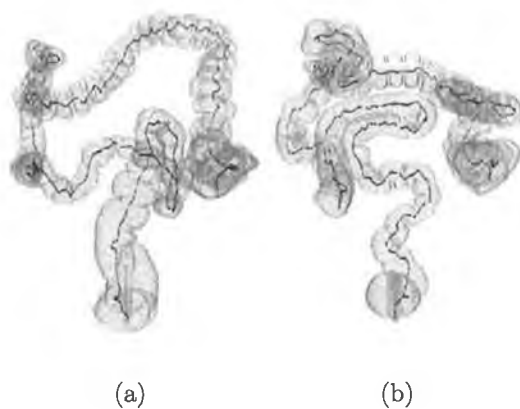
Figure 2.18: Iso-surface of well-distended colons with the centreline superimposed.

(a)   (b)
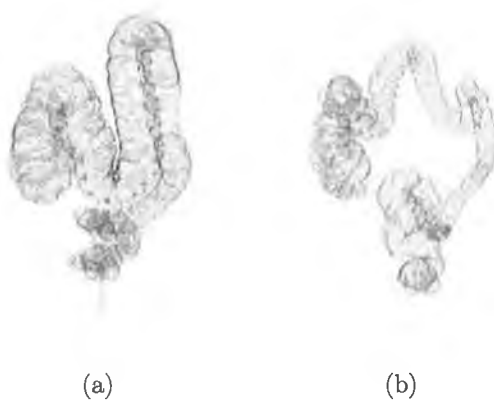
Figure 2.19: Iso-surface of well-distended colons.

Figure 2.20: Collapsed colon (standard dose). (a) Three colon surfaces with $V/L >$ 300 and seven colon surfaces with $(V/L < 300)$ and six *ESCs* with $(V/L < 300)$. (b) Three colon surfaces with $V/L > 300$, twelve colon surfaces with $(V/L < 300)$ and eleven *ECSs* with $(V/L < 300)$. (c) Five colon surfaces of $V/L > 300$, nine colon parts with $(V/L < 300)$, and seven *ESCs* with $(V/L < 300)$.
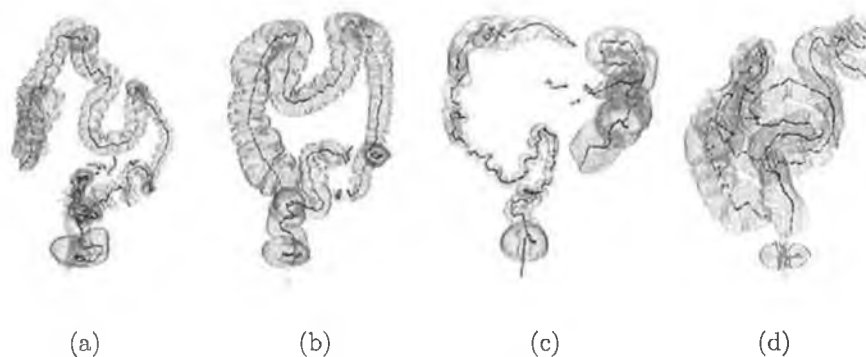
Figure 2.21: Collapsed colon (standard dose). (a) Two colon surfaces with $V/L >$ 300, one colon surface with $(V/L < 300)$ and four *ESCs* with $(V/L < 300)$. (b) Two colon surfaces with $V/L > 300$, one colon part with $(V/L < 300)$, and one extra-colonic surface with $(V/L < 300)$. (c) Two colon surfaces with $V/L > 300$, two colon parts of $(V/L < 300)$, and two *ESCs* with $(V/L < 300)$. (d) Six colon surfaces with $V/L > 300$.

detailed based on the inclusion of geometrical feature such as V/L analysis, orientation, end points, gradient of centreline, and directions (clockwise or anticlockwise). The experimental data indicates that the V/L analysis provides a better approach to discriminate the colon parts from the small bowel. In the calculation of V/L, morphological labelling was used for finding the end points and the volume and the shortest path algorithm was used for finding the length. For well-distended colon detection, the features included in the segmentation process are V/L, length, ori-

(a)   (b)   (c)

Figure 2.22: Collapsed colon (low dose). (a) Three colon surfaces with $V/L > 300$ (b) Two colon surfaces with $V/L > 300$, one colon part with ($V/L < 300$), (c) Three colon surfaces with $V/L > 300$, four colon parts of ($V/L < 300$).

entation, and geometrical position in the volumetric data. In the detection of the large segments of collapsed colon (with $V/L > 300$), the developed method employed the geometrical position, V/L, length and direction as features. The features that are used for collapsed small colon segments ($V/L < 300$) detection are end points, length, distance and gradient of centreline. All threshold parameters used in the automatic segmentation scheme were selected with a high degree of tolerance and they proved to be robust in the segmentation process. Any dataset without a labelled region of length less than 400mm was declared to be a poorly distended dataset and the algorithm rejects the dataset as unsuitable for automated analysis. This condition arises when the datasets have nearly 50% of the regions filled with residual material and fluid.

The developed method for automatic segmentation successfully identified the colonic lumen from volumetric CT data. In 96 supine and prone (88 standard and 8 low-dose) datasets containing collapsed colon data, the segmentation method detects 99.68% of the colonic wall and shows 94.79% sensitivity for collapsed colon detection. The overall sensitivity in colon detection was 96.95%. In 63 datasets the well-distended colons were detected without any inclusion of extra-colonic surface. The performance of the developed algorithm makes it suitable for 3D visualization of the colon surface and advanced polyp detection.

After segmentation the next step in CAD-CTC is automatic polyp detection. The following chapter deals with the development of CAD algorithms for automatic polyp detection in CAD-CTC. The chapter details the development of three different

feature detection schemes for colonic polyp detection in CAD-CTC and the methods are as follows:

- Geometrical features based method.

- Statistical features based method

- 3Db features based method.

Chapter 3

CAD-CTC Polyp Detection

CT Colonography is a rapidly evolving technology for the detection of colorectal polyps and many studies have demonstrated that its sensitivity in polyp detection is comparable to the sensitivity offered by conventional colonoscopy [21, 23, 26, 29, 32, 33, 34]. In this regard, Fenlon et al. [21] indicate that CTC returns 100% sensitivity for the detection of CTC polyps greater than 10mm and 83% sensitivity for detection of polyps in the range 6-9mm polyps. This conclusion is supported by other studies [32, 74, 75, 76] where it is demonstrated that CTC is as good as standard colonoscopy for the detection of colonic polyps. More recently, Pickhardt et al. [29] performed a detailed comparison between CTC and standard Colonoscopy and they concluded that CTC can increase the sensitivity of polyp detection when applied as a second reader with Colonoscopy. In their study, the reported sensitivities for CTC and optical colonoscopy for polyps $\geq 10mm$ were 92.2% and 88.2%, for polyps $\geq 9mm$ were 91.8% and 90.2%, for polyps $\geq 8mm$ were 92.6% and 89.5%, for polyps $\geq 7mm$ were 89.5% and 90.2%, for polyps $\geq 6mm$ were 85.7% and 90.0% respectively. From these results they concluded that the sensitivity in polyp detection offered by CTC matches closely the sensitivity achieved by optical colonoscopy and CTC is feasible to be used in clinical examinations.

Since the introduction of CTC in 1994 [15], a large number of techniques in the fields of 3D visualization, such as the rendering of the colon surface, centerline calculation, and colon wall unfolding were developed to provide the radiologists with all types of 2D and 3D information required to identify the colorectal polyps. [77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87]. However the development of new CT imaging modalities, the high resolution CT data offers a large volume of information that is required to be visualized and interpreted by the radiologists (the typical time required to process a dataset based on a visual examination is in the range 12-60

minutes). As pointed out in the study by Pickhardt et al. [29] the performance of the radiologists can be effected by factors such as perceptual errors [88, 89] and eye fatigue [34, 35]. Johnson et al. [88] study shows that 34% (20 of 59) of the large polyps were missed in CTC due to perceptual errors. Hence, Ven Gelder et al. [89] suggested that the introduction of CAD based automatic polyp detection in CTC is a viable solution to reduce the perceptual errors associated with the visual interpretation of the CTC datasets. Thus, the development of CAD methods can improve both the sensitivity and efficiency of CTC. In the last decade a significant amount of research has been focused on developing automated CAD of colonic polyps and a large number of CAD-based polyp detection techniques have been proposed.

One of the first CAD-CTC systems was proposed by Vining et al. [90] where the detection of colonic polyps was based on surface curvature analysis. In the experimental section of their paper they indicated that the CAD-CTC system achieved 73% sensitivity with 9 to 90 false positives (FP)/dataset.

The polyp detection system developed by Summers et al. [91] attempts to identify the polyps in the CT data using a multi-stage geometrically-driven approach. Initially, they detect the convex surfaces that protrude inward from the colon by applying a kernel filer that is constructed using partial derivatives. After the detection of the candidate surface, they used shape-based criteria derived from the principle curvature (k_{min} and k_{max}), mean curvature (H), sphericity ratio $s = (k_{max} - k_{min})/H$ and minimum polyp size. They used very restrictive sphericity criteria in order to reduce the false positives but their technique shows zero sensitivity for polyps in the range 5-10mm (0 out of 4) and 100% sensitivity for polyps $\geq 10mm$ (6 out of 6). Later, Summers et al. [36] proposed a new method method where they applied a different shape based filter (calculated from k_{min}, k_{max}, and H) to reduce the level of FP but keeping the sensitivity at 100%. One problem with this approach is the fact that the sensitivity and specificity of the system depend on the filter chosen to evaluate the local colon curvature and the reported sensitivities in polyp detection are in the range 29% to 100% with 6 to 20 FPs/dataset.

Yoshida et al. [37, 92] proposed the use of shape index (cup, rut, saddle, ridge, cap), curvedness values (calculated on small volumes of interest) and fuzzy clustering in order to perform candidate polyp surface generation. The principal curvature (k_{min} and k_{max}) derived from the Gaussian and the mean curvature was used to calculate the shape index and curvedness for each colonic wall voxel. They showed

that all types of colon shapes can be mapped in the interval $SI \epsilon [0, 1]$ as follows: cup (0.0), saddle (0.5), ridge (0.75), and cap (1.0). On the other hand, they showed that the curvedness is also an indicator of the variation of the local curvature and they used a predefined threshold with a value between 0.9 to 1.0 for SI and $0.08m^{-1}$ to $0.20mm^{-1}$ for curvedness to generate the initial seed points. The C-Means clustering was used to generate the candidate surfaces and to reduce the incidence of non polyp surface generated by noise. The *CAD-CTC* system [92, 37] employed features such as the shape index, curvedness, magnitude of CT values, CT values, gradient concentration (GC) and direction of the gradient concentration (DGC) calculated from candidate surfaces to classify them into polyps or folds. They reported 95% sensitivity in polyp detection with 1.2 FP per dataset, but the FPs increased with a factor of 1.5 when the sensitivity was increased to 100%.

Paik et al [35, 93] developed a new algorithm called surface normal overlap that was applied to colorectal polyp detection. Their algorithm is based on the assumption that the colorectal polyps are convex structures and the local normal intersection density samples the local convexity for each voxel of the colon wall. The normal overlap technique was used to identify suspicious convex structures while the polyp detection is performed by assessing the deviation of these convex structures from a stochastic model employed to define the shape of a nominal polyp. This algorithm shows 100% sensitivity in detecting polyps larger than 10mm with 7 FP datasets. No experimental data is provided in regard to the sensitivity of their *CAD-CTC* system when applied to the identification of small ($< 5mm$) and mid-sized polyps (between 5-10mm).

Kiss et al. [94, 95, 96] method also employed the surface normal intersection for the detection of convex surface from the colonic wall. To generate the polyp candidate surface, they applied the Hough Transform to calculate the center points and used 3D region growing to find the candidate surface from the convex voxels of the colon wall. Gaussian distribution of the Hough points was used to calculate the normal concentration of the candidate surface. Two different region growing techniques (weighted region growing and greedy region growing) were employed to generate candidate surfaces from the center points and least square ellipsoid fitting was used to calculate the three axes of the candidate surface resulting from these two region growing algorithms. The number of normal intersections for each Hough point, Gaussian distribution, three axes of the surface resulting from the greedy

region growing and three axes of the surface generated by the weighted region grow-ing were used as input features for a probabilistic neural network (PNN) classifier. Their *CAD-CTC* system achieved 90% sensitivity for polyps larger than 6mm with 2.82 FPs/dataset. Recently, a different *CAD-CTC* system has been proposed by Kiss et al. [38] that analyses the slope density function as a discriminative feature to classify the convex candidate surfaces into polyps and folds. The initial stage of their system identifies the candidate surfaces by intersecting the colon wall with a reformatted plane perpendicular on the local normal surface. If the intersection patch between the planar and the colon surface is filled with voxel data the colon surface is concave and is declared part of the healthy colon tissue. Otherwise is a convex surface that is generated either by polyps or folds. The resulting can-didate surfaces are evaluated statistically using the slope density function, which shows peaks for elongated surfaces and smooth values for ellipsoidal surfaces. This property of the SDF is very useful as it provides robust discrimination between the polyps and folds as the folds resemble elongated cylindrical surfaces whereas polyps ellipsoidal surfaces. Their method obtained the following performance in polyp de-tection: 33.33% sensitivity for polyps smaller than 5mm, 85.70% for polyps in the range 6-9mm, 90% for polyps larger than 9mm and 100% sensitivity for cancerous lesions. Kiraly et al. [97] proposed a fast detection method using a gradient-based filter and shows 96% sensitivity for polyp greater than 5mm with 5.76 false positive per dataset.

Acar et al [39, 98] employed a different approach based on the edge flow displace-ment that is applied to obtain robust polyp detection. They developed a method to extract the candidate surfaces based on the Hough Transform that evaluates the normal intersections using the assumption that the normal intersection will be high for convex (cap-like) structures. After the extraction of the candidate surfaces, they scrolled these surfaces with a planar perpendicular on the main axis of the surface and they computed the edge flow from the extremity of the surface towards its center. The divergence of the edge flow is used to determine whether the candidate surface is generated by a polyp or a fold. They applied this technique on 48 datasets and their experiments indicate that their method achieved 35% specificity at a sen-sitivity rate of 100%. This method was further advanced by Gokturk et al. [99] when they applied the randomly oriented triple orthogonal planes at the location of each candidate surface. They applied this approach to sample the sphericity of

the candidate surface based on the fact that any random planar slicing through a spherical surface will generate a circle. The reported experimental results indicated that they achieved 69% specificity at a sensitivity rate of 100%. No detailed analysis with respect to the size of the polyps is provided. Wang et al. [40] stated that the inclusion of morphological and texture features can reduce 10 times the false positives when compared to the standard shape-based approach. Wang et al. [100] combined the texture features and global curvature for automatic polyp detection and shows 100% sensitivity for $\geq 10mm$ with 2.0 false positive per dataset. Jerebko et al. [41] employed a multiple neural network classification scheme to achieve a 36% reduction in FPs and a 20% reduction in false negative (FN)detection. Later, Jerebko et al. [101] employed a support vector machines committee classification scheme to achieve 81% sensitivity with 2.6 false positive per dataset. Iordanescu et al. [102] developed a rectal tube detection method that was applied to reduce the FPs generated by the rectal tube. Li et al. [67] proposed method employed different geometric features such as maximum polyp radius calculated from the minimum curvature, minimum polyp radius derived from the maximum curvature, candidate surface area, roundness of the candidate surface and elongation factor for classification of the candidate surfaces into polyps or folds. Their method achieved 90% sensitivity with 2 FPs per dataset.

All the above mentioned *CAD-CTC* techniques show 100% or close to 100% sensitivities in the detection of polyps $\geq 10mm$, while the sensitivities in the detection of polyps in the range $[5 - 10)mm$ vary from 70% to 95%. The reported false positive rates vary from 2.0 to 90 per dataset. Among all the developed *CAD-CTC* techniques, Yoshida et al. [37, 92] and Kiss et al. [38] methods show best results for sensitivity and false positives incidence per dataset. Yoshida et al. technique achieved a sensitivity of 100% per patient with 2.0 false positive. But it is worth noting that the sensitivity dropped to 90% when it was presented as per polyp. Also the polyps smaller than 5mm were completely ignored in their evaluation. Kiss et al. [38] method shows 90.90% sensitivity for polyps $\geq 9mm$ and 100% sensitivity for colorectal tumors with a false positive rate of 2.48. Their method shows 33.33% sensitivity for polyps $< 6mm$ and 85.70% sensitivity for polyp between $6 - 9mm$ where the data used in their experiments has been acquired with 0.8mm reconstruction interval. It is also useful to note that both Kiss et al. and Yoshida et al. *CAD-CTC* techniques evaluated the difference in the geometrical shapes between polyps and

folds.

In this thesis these geometrically-driven approaches will be further advanced by developing a number of CAD polyp detection techniques where the discrimination between polyps and folds is performed using the features that sample the morphology of the local 3D data. All the proposed polyp detection methods employed different features derived from the colon wall in order to classify optimally the candidate surfaces into polyps and folds. The first polyp detection scheme called geometrical fitting approach evaluates the discriminative power of the features calculated from the colon surface using least square approximation (ellipsoid, sphere, plane) in order to perform polyp identification. The second method uses the statistical features derived from the colonic surface. The third method analyses the 3Decibel (3dB) attenuation on the surface variation curve and surface normal concentration for polyp detection.

3.1 Geometrical Fitting Approach

Figure 3.1, gives an overview of the proposed algorithm. In this section, the segmentation, polyp surface generation and feature extraction phases of the algorithm are discussed in detail.
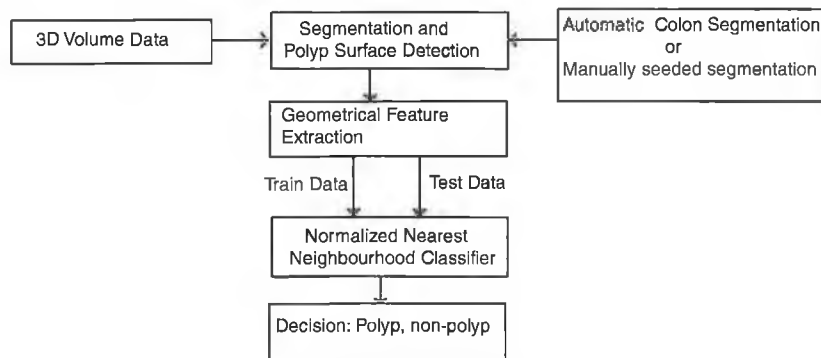
Figure 3.1: Overview of the Geometrical Fitting *CAD-CTC* system.

3.1.1 Segmentation

CTC images provide high contrast between the gas and colon surface. Using a region growing [69] algorithm the gaseous region can be segmented successfully. Sometimes remaining residual material and water can create collapses in the colon and the region growing algorithm may require multiple seed points to segment the

entire colon. The developed *CAD-CTC* system provides both the manual assisted segmentation and the complete automatic segmentation detailed in Chapter 2. The manually placed seed segmentation used -800HU as threshold, as suggested in [72, 90] whereas the automatic colon segmentation detects the threshold (from -900HU to -800HU) automatically from the global histogram (see Figure 2.2). The *colonic wall* (CW) is defined as the adjacent voxels having HU values higher than -800HU or the automatically detected threshold.

3.1.2 Polyp Surface Detection

3.1.2.1 3D Hough Transform

The normal vector for each voxel in the *CW*-set was calculated using the Zuker and Hummel operator [103]. Each voxel in the *CW* creates 7 *Hough points* (HP) (see Figure 3.2) in the normal direction from 2.5mm to 10mm (2.5, 3.75...8.75, 10.0) by varying the parameter t in Eq. 3.1.1,

$$p = p_1 + t \times n \tag{3.1.1}$$

where p_1 is the colon wall voxel and n is the normal vector to that voxel. In Eq. 3.1.1 the value of t starts from 0.1 and increases with the step size of 0.1mm until all the *HP* points situated at distances 2.5, 3.75, 5.0, 6.25, 7.5, 8.75, and 10.00mm are generated. The term Hough Point has been introduced in order to highlight the similarity with the Hough Transform that is applied to identify the 3D spherical objects in CT data.
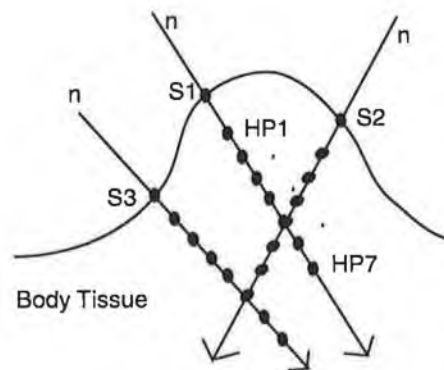
Figure 3.2: Surface normal and the distribution of the 7 *Hough Points* (HP) in the normal direction.

3.1.2.2 3D Histogram

The *HPs* are uniformly distributed from 2.5mm to 10mm along the normal vector direction for each voxel of the colon wall (CW) and the intersections between the *HPs* are recorded (see Figure 3.2) in a 3D histogram. Thus, the 3D histogram records the intersections between the *HPs* that are in fact intersection of the normal vectors. As the normal vectors are determined using 3D local operators their orientation is sensitive to abrupt changes in the 3D structure of the *CW*, and to reduce the level of noise in the histogram a weighted smoothing procedure is applied using the expression illustrated in Eq. 3.1.2,

$$V_{smooth} = \delta \times Voxel + \sum_{0}^{26} \frac{(1 - \delta) \times Voxel_{neighbour}}{26} \tag{3.1.2}$$

where δ is equal to $1/\sqrt{2}$.

3.1.2.3 Non Maximum Suppression

After smoothing, all *HP's* having histogram values higher than 4.0 intersections are considered as *initial candidate center points* (ICCP) of the candidate polyp surfaces. Non maximum suppression was applied in the *ICCP* set to create potential center points. The cluster of surface points was created by including the *HPs* and their corresponding surface voxels within a certain distance from *ICCP* (10mm to 25mm). It is useful to remember that folds are generally shaped like cylinders and show a uniform distribution of the number of intersections generated by the *HPs* along the axis of the cylinder. Conversely, polyps resemble either spherical or ellipsoidal shapes and show a narrow peak in the 3D histogram. A minimum distance of 10mm was experimentally selected in initial clustering to include the highest possible number of surface points in the clustered surface. The distance threshold varies from 10mm to 25mm depending on the histogram value for each center point in *ICCP*. The candidate surface cluster may include surrounding non-convex surface points or disconnected surfaces (Figure 3.3) that may create problems when the candidate surface is analysed to decide if it is a polyp or a fold. To eliminate these undesired surface points from the initial cluster, a *Candidate Surface Processing* procedure is applied. This procedure is described in detail in the next section.

(a)   (b)   (c)
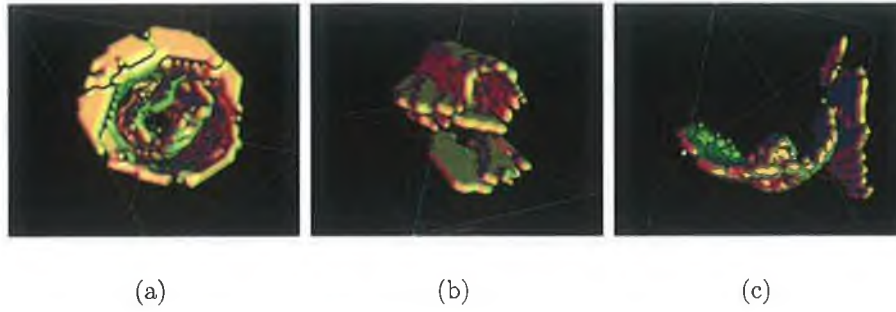
Figure 3.3: 3D Surface after initial clustering (a) 3D surface of an inserted tube, (b) 3D surface of a fold and (c) 3D surface of a polyp

3.1.2.4   Candidate Surface Processing

To remove the non-convex surface points and the disjoint points from the initial cluster, a Candidate Surface Processing procedure that calculates the Gaussian mapping for each cluster and performs a non-convex surface voxel removal test was developed.

1. *Gaussian Center and Radius Detection*: To calculate the center and radius of each cluster, a Gaussian distribution depicted in Eq. 3.1.3 was calculated for each *HP* of the cluster,

$$GM_i = \sum_{j=1}^{N} e^{(-x^2/2.0 \times \sigma)} \tag{3.1.3}$$

where the variable x is the distance between the *HPs*, σ is the standard deviation and is set to 1. The quantity N is the number of *HPs* in the cluster and j takes values between *1...N*.

The *HP* with the highest Gaussian distribution was set as the center of the clustered surface and the Euclidian distance between the center and its corresponding surface point is the radius of the cluster. The Gaussian distribution is an efficient feature that can be used to discriminate between polyps and folds. In this regard, the Gaussian distribution has high values for polyps and low values for folds.

2. *Surface Convexity Test*: Let S be a surface voxel, n be the normal vector at the surface voxel S and Q be the intersection point of the surface normal and the perpendicular line from the center of the cluster to the surface normal (see Figure 3.4). To remove the non-convex points from the initial cluster a simple surface convexity test was employed. In this regard, the non-convex surface point S will be removed from the cluster if the dot product $< \overline{SQ}, \overline{n} >$ is less than zero. In Figure 3.4, the points *s1* and *s4* and their associated *HPs* will be removed from

the cluster as they do not pass the convexity test. The normal distance from the center of the cluster (CP) to the surface normal at position SP and the distance between the surface point (SP) and the intersection point (IP)(see Figure 3.5) were also checked. If the distance between the surface point SP and the intersection point IP is larger than 10mm (the maximum HP distance), the surface point SP is eliminated from the candidate surface.
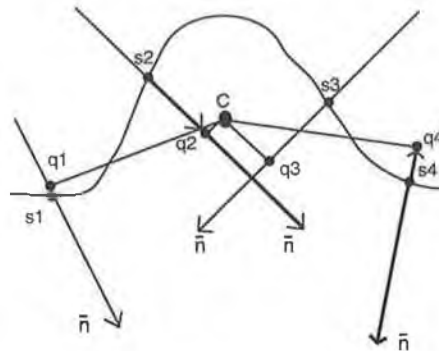
Figure 3.4: Convexity test. The point C is the center of the cluster. The surface points *s2* and *s3* pass the convexity test whereas the surface points *s1* and *s4* and their associated Hough points will be removed from the candidate surface as they do not obey the condition $< \overline{SQ}, \overline{n} >$ less than zero.
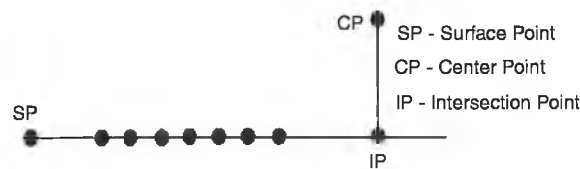
Figure 3.5: SP, CP and IP are the surface point, center point and intersection point respectively. The circles between the SP and IP represent the 7 Hough points for each surface point.

After the removal of the non-convex surface voxels, each cluster was further processed to evaluate discontinuities in the surface under examination. If discontinuities exist in the surface area, the cluster is divided into multiple clusters and their Gaussian map, center and radius are recalculated (see Figure 3.6).

3.1.3 Feature Extraction of Geometrical Fitting

The aim of the method detailed in this section is to calculate the features associated with each cluster surface, which will be considered as input for the classifier. The features must be selected in order to maximize the discriminative power between

(a)   (b)   (c)
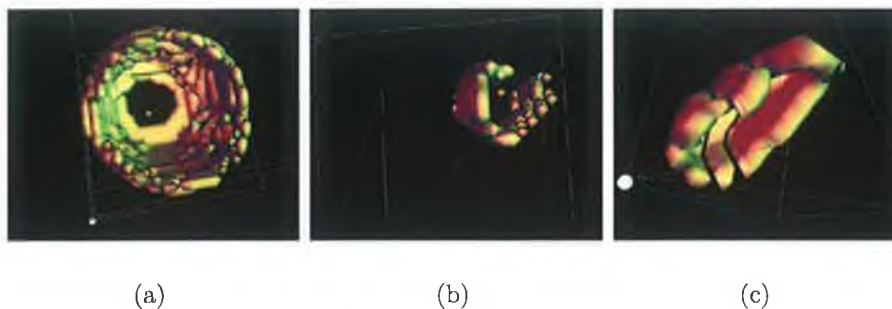
Figure 3.6: 3D surface resulting after the re-clustering phase. (a) 3D surface of the inserted tube illustrated in Figure 3.3a, (b) 3D surface of the fold illustrated in Figure 3.3b, (c) 3D surface of the polyp illustrated in Figure 3.3c.

polyps and folds. Recall that the nominal model for polyp is either spherical or ellipsoidal, while the nominal model for fold is cylindrical [35, 95]. The features computed are: the Gaussian distribution, least square approximation of the sphere fitting radius and error, least square approximation of the three axis of the ellipsoid and ellipsoid fit error. In our experiments we have evaluated a large number of features and in the final implementations we have retained only those that exhibit maximal discrimination between polyps and folds. An automatic feature selection technique would be difficult to be devised since the geometrical features do not respond linearly to polyps having different sizes. Thus, the automatic feature selection method was not examined as part of this study.

The Gaussian distribution which estimates the center and radius of each cluster was calculated in the candidate surface processing (see Section 3.1.2.4). For folds the the value of the Gaussian distribution is considerably smaller than the value of the Gaussian distribution calculated for surfaces generated by polyps (see Figure 3.7). Sphere fitting for each cluster was performed in two phases. Firstly, the error in the least square sphere fitting [104, 105] was calculated using the existing Gaussian center and the Gaussian radius of the cluster. Secondly, the cluster radius and the center point were re-calculated using a least square sphere fitting algorithm [104, 105]. Experimental results indicate that for spherical polyps, the Gaussian radius and the cluster center were very close to those obtained using the least square estimated sphere and the error in fitting is small. For folds the least square estimated radius is higher than the Gaussian radius and the sphere fitting error is significantly higher than the fitting error for polyps. This is illustrated in Figure 3.8 (note that

Gaussian distribution for polyp

(a)

Gaussian distribution for fold
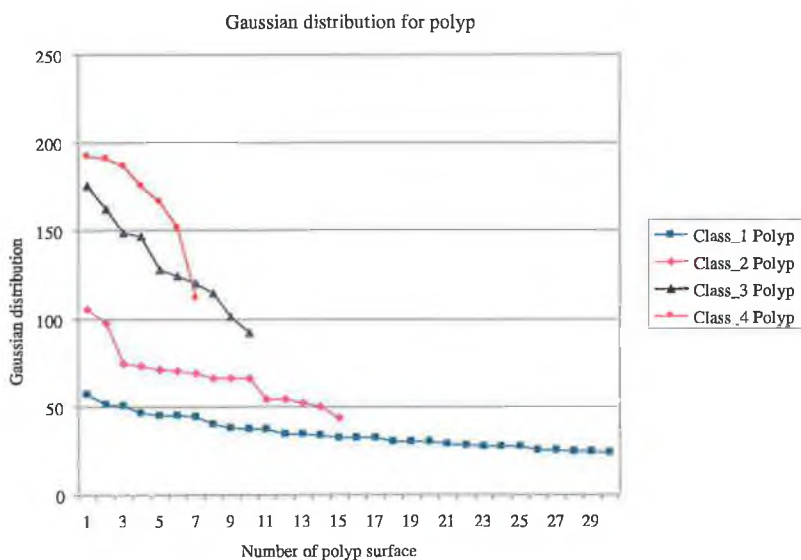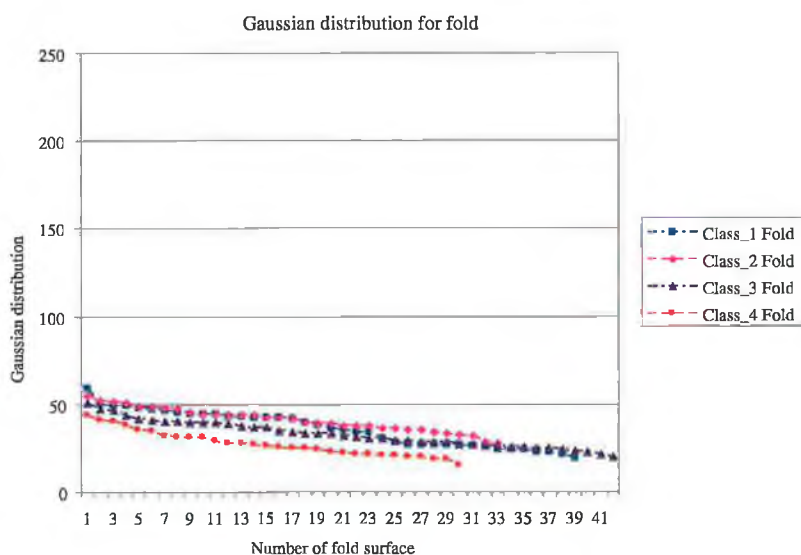
(b)

Figure 3.7: Gaussian distribution. (a) and (b) show the Gaussian distribution for different classes of polyps (a) and folds (b) respectively(polyps and folds classes are sorted by size in ascending order).

polyp and fold classes are ordered by size in the diagram) where the sphere fitting error for a large variety of polyps and folds is plotted. Experimental results also show that the change in the fitted sphere radius for the candidate surface and the half radius surface was significantly higher for folds when compared to polyps (see Figure 3.9).
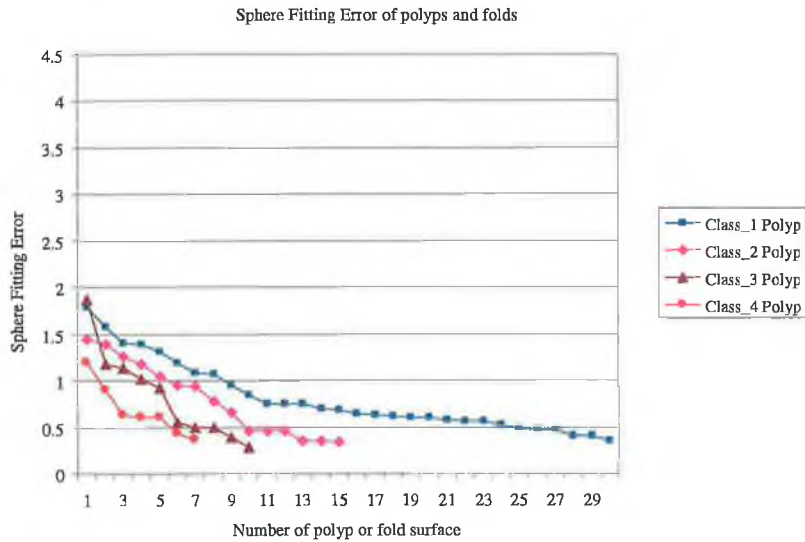
(a)
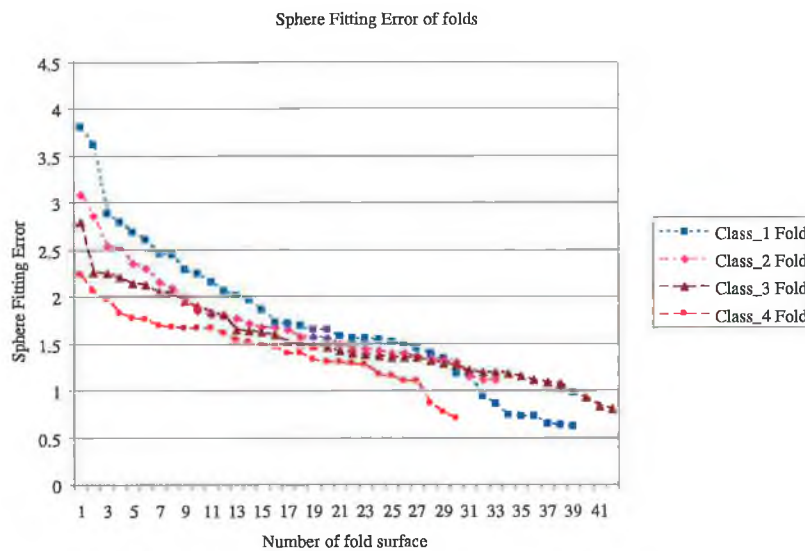
(b)

Figure 3.8: Sphere fitting error analysis. (a) and (b) represent sphere fitting error analysis for different classes of polyps and folds respectively(polyps and folds classes are sorted by size).

The principal axes of the fitted ellipsoid and its associated estimation error [104] were calculated for each polyp candidate surface and its derived half radius surface. The half radius surface voxels are determined from the existing cluster and includes those surface voxels, which have a distance from the center of cluster to the surface normal less than a *half radius threshold* (HRT). The *HRT* is selected in conjunction with the Gaussian distribution value and varies from 2mm for small candidate sur-
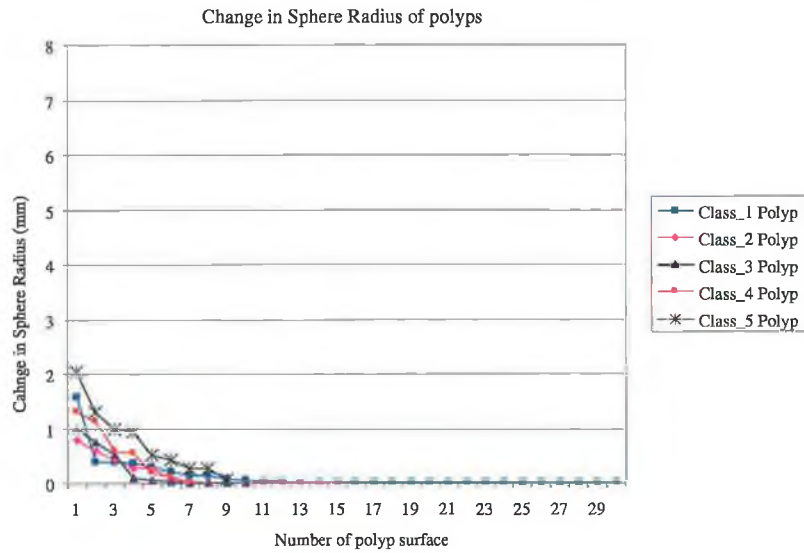
(a)
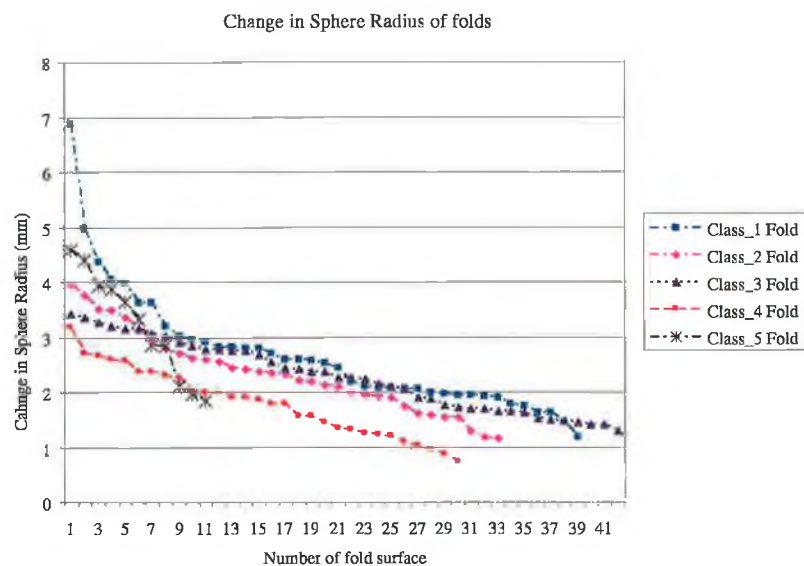
(b)

Figure 3.9: Change in sphere radius. (a) and (b) depict the change in sphere radius for different classes of polyps and folds.

faces to 5mm for large candidate surfaces. The minimum value of HRT (2mm) was experimentally selected. The *Surface Change Rate* (SCR) value computed using the equation 3.1.4 is minimal for polyps (see Figure 3.10) but it is large for folds (see Figure 3.11, 3.12),

$$SCR = (N_T - N_H)/N_H \qquad (3.1.4)$$

where N_T is the number of surface voxels in the cluster and N_H is the number of

surface voxels in the half radius surface.

It was also found that the change in the major axis direction of the fitted ellipsoid for the candidate surface and the half radius surface was significantly higher for folds when compared to polyps (see Figure 3.13). Similarly, change in the ellipsoid fitting error for the candidate surface and the half radius surface was higher for folds when compared to polyps (see Figure 3.14)
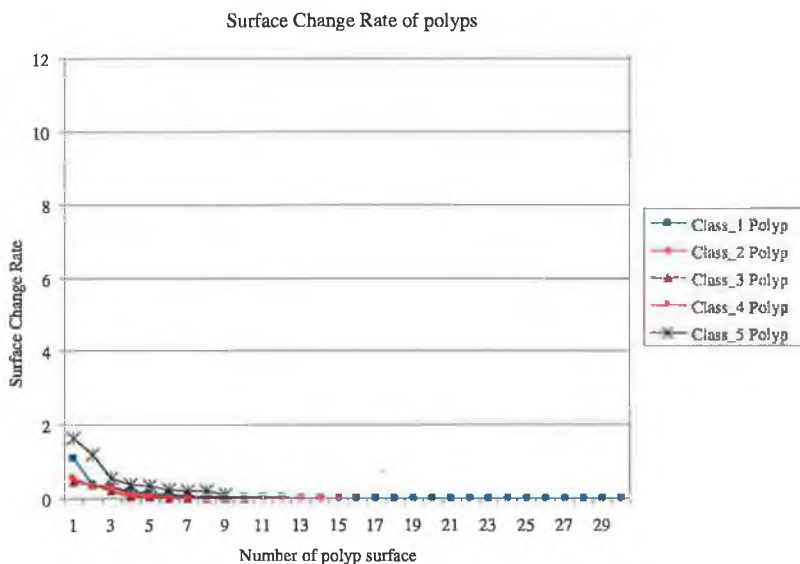
(a)
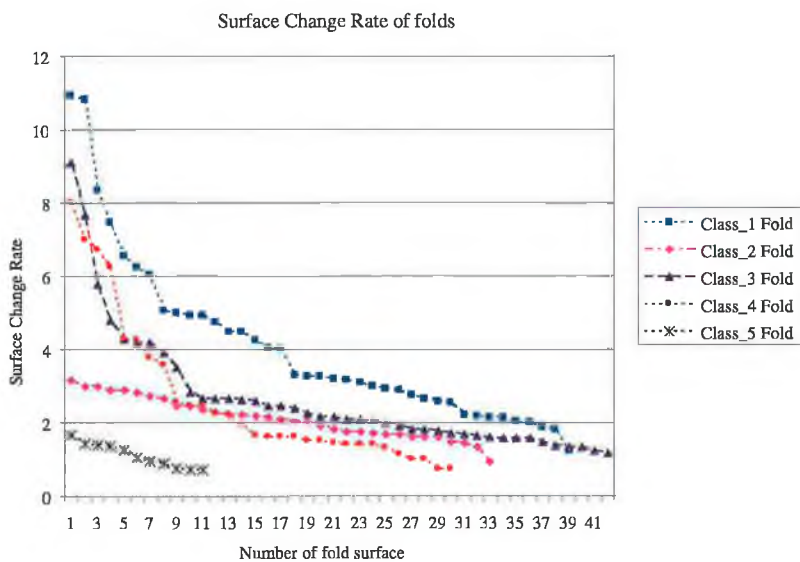
(b)

Figure 3.10: Surface change rate. (a) and (b) show surface change rate for different classes of polyps and folds.
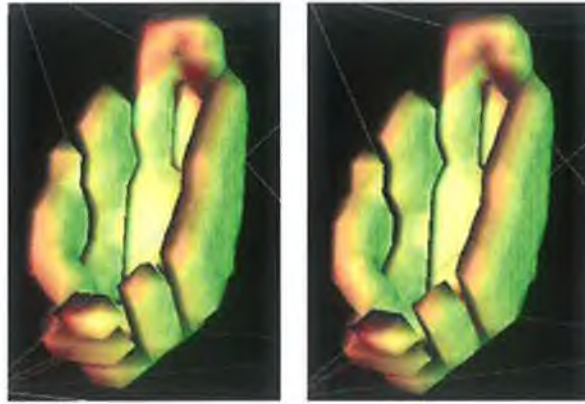
Figure 3.11: 3D surface generation of a polyp (a) and its half radius surface (b). No significant differences between them are noticed.
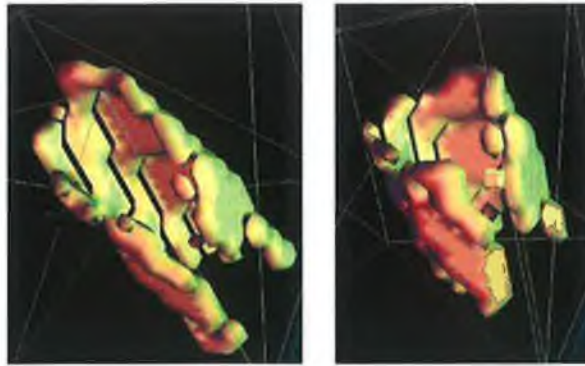
Figure 3.12: 3D surface generation of a fold (a) and its half radius surface (b). It can be noticed a significant difference between them.

The other features that are used for classifying the candidate surface as polyps and folds are sphere radius, change in sphere radius, principle axes of ellipsoid fitting, change in Gaussian distribution. All the above mentioned features exhibit high discrimination between polyps and folds as illustrated in Figures 3.8, 3.10, 3.13 and these features are the inputs for two different classifiers that are used to classify the candidate surfaces into polyps and folds.

3.1.4 Classification

For polyp/fold classification a multiple-class-segregated *feature normalized nearest neighborhood* (FNNN) classifier detailed in [106], *Probabilistic Neural Network* (PNN) [107] were employed. To evaluate the performance of the *FNNN* classifier its performance was compared against the performance of one of the commonly used classification schemes, namely the *PNN* classifier. The *FNNN* classification scheme

Change in Principle Axis of polyps

(a)

Change in Principle Axis of folds
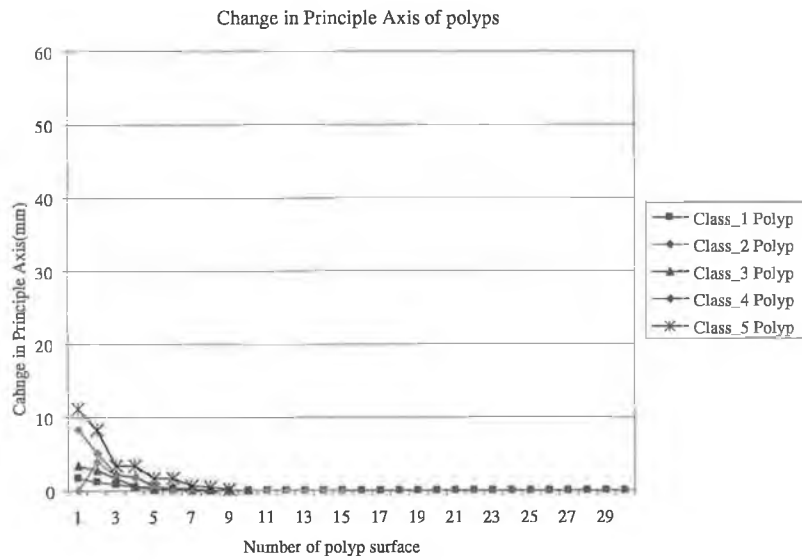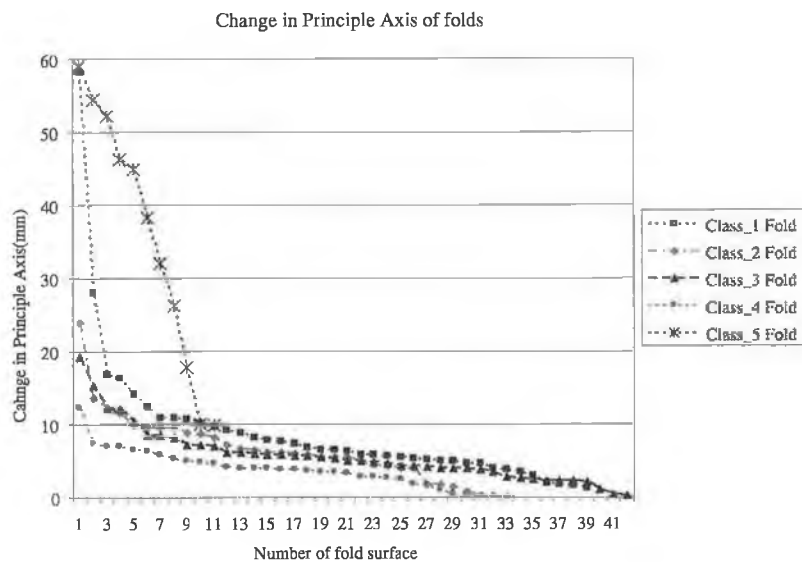
(b)

Figure 3.13: The change in major axis orientation. (a) and (b) display the change in major axis orientation for different classes of polyps and folds.

consists of two stages. Firstly, the training database is created by using the features detailed in the previous section for each class of polyps and folds. Features of each class were normalized in order to avoid the situations where the features with the largest values subdue the remaining ones. The feature normalization scheme was performed in order to normalize each feature to zero mean and unit variance (see Eq. (3.1.5) and (3.1.6))

(a)

(b)
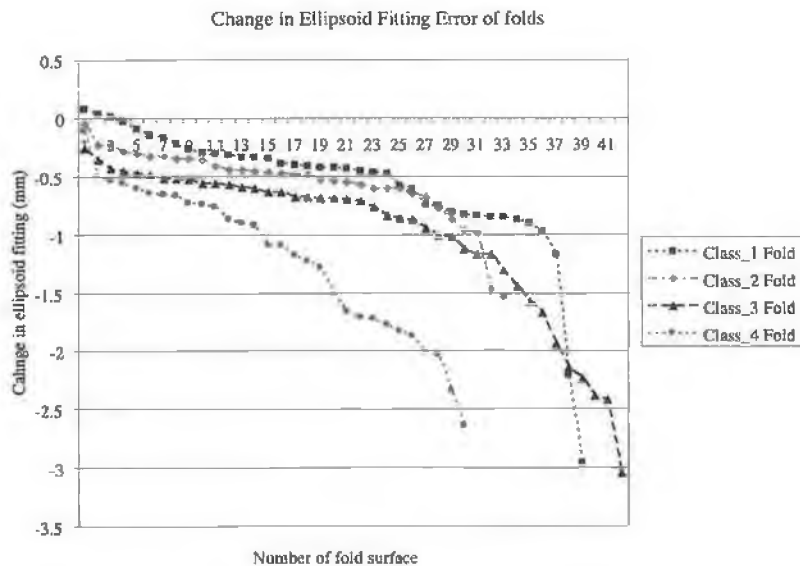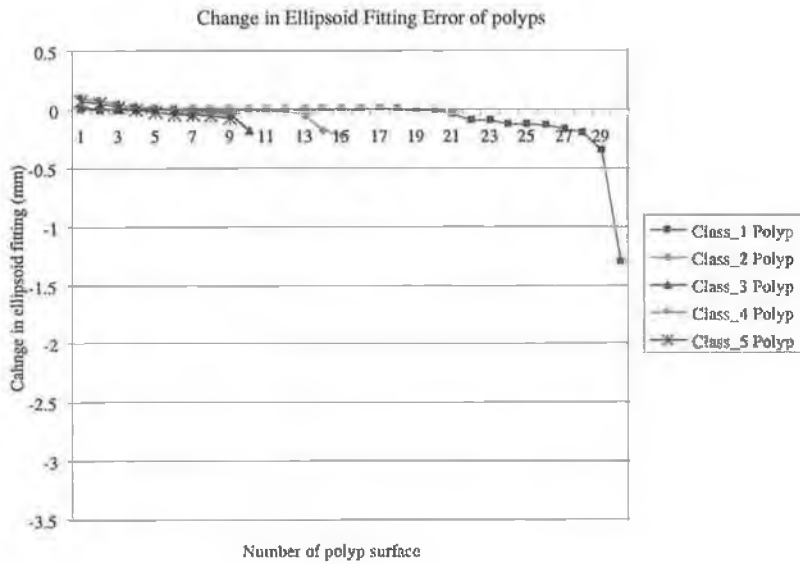
Figure 3.14: Change in ellipsoid fitting error. (a) and (b) display the change in the ellipsoid fitting error for different classes of polyps and folds.

$$m_i = \frac{\sum_{j=1}^{k} x_j[i]}{k} \qquad s_i = \sqrt{\frac{\sum_{j=1}^{k}(x_j[i] - m_i)^2}{k}} \qquad (3.1.5)$$

$$X_j[i] = \frac{x_j[i] - m_i}{s_i} \qquad for \quad j = 1, ..., k, \quad i = 1, ..., n \qquad (3.1.6)$$

where n defines the number of features per pattern, m_i and s_i are the mean and the variance of the ith features, x_j is the unprocessed jth pattern, k defines the num-

ber of patterns contained in the model database and X_j represents the normalized jth pattern. The classification stage computes the Euclidian distance between the input patterns calculated from candidate surfaces and the patterns contained in the database,

$$dist_j = \sqrt{\sum_{i=1}^{n}(X_j[i] - Y[i])^2} \quad for \quad i = 1, ..., n \qquad (3.1.7)$$

where X_j is the jth pattern from the model database and Y defines the pattern from the input surface to be classified. The input is declared as polyp if the $min(dist_j)$ belongs to polyp class, otherwise declared as fold.

Probabilistic neural networks are radial basis networks suitable for a large range of classification problems. *PNN* is constructed on a feed-forward architecture and supervised training algorithm that is based on back propagation. *PNN* allows incremental learning where new training data can be added at any time without requiring retraining of the entire network.

The *FNNN* training databases consist of five polyps and five folds databases. The polyps were classified into small spherical, medium spherical, big spherical, elliptical, and non-spherical polyp. The fold database was also classified as small folds, small convex surface, medium folds, large folds, tube. In Figures 3.8, 3.10 and 3.13 *class_1 polyp, class_2 polyp, class_3 polyp* and *class_4 polyp* represent small, medium, large and elliptical polyps respectively and *class_1 fold, class_2 fold, class_3 fold* and *class_4 fold* represent large folds, medium size folds, small folds and small convex surfaces respectively. *Class_5* polyp and *class_5* fold in Figure 3.10 and 3.13 represent non spherical polyps and inserted tubes respectively. In total 64 polyps and 155 folds were used to train the *FNNN* and *PNN* (this technique has been developed first and it has been trained on a smaller number of false positives than the methods discussed in sections 3.2.2 and 3.3.1). By experimentation it has been demonstrated that the approach of segregation in polyp training by size offered the optimal solution to increase the identification rate especially for small polyps ($< 5mm$) but not at the expense of increasing the level of false positives.

The geometrical feature-based approach is only suitable for the CT data acquired at 3.0mm slice thickness and 1.5mm reconstruction interval which is the standard protocol used in the Mater Hospital Dublin (clinical partner). Modifications in the reconstruction interval generate changes in the surface area for polyp candidate surface and may alter the feature values calculated from the candidate surface.

Hence, the geometrical feature-based technique is suitable only to be applied to CT datasets acquired with the protocol used in the training stage. To overcome these problems statistical feature-based methods were developed where 3D interpolations are applied in order to generate isometric datasets.

3.2   Statistical Feature based method

The statistical feature-based method consists of five steps as illustrated in Figure 3.15. Initially the non-isometric patient data was converted to isometric data by using cubic interpolation. Segmentation of the colon is performed using manually placed seed points in conjunction with 3D region growing (6-neighbouhood) algorithm [69]. Threshold for the region growing was set to -800HU as suggested in 3.1.1. The voxels adjacent to the colon voxel having HU values higher than -800 define the colon wall. Polyp candidate generation includes the application of Hough Transform, 3D Histogram, smoothing of the HP space, initial center point calculation, clustering of the colonic voxels and candidate surface processing. The Hough Transform used to generate the candidate surfaces is similar to the method discussed in the section 3.1.2.1. In this step each surface voxel creates 8 HP in normal direction from 2mm to 10mm. The 3D histogram creation and smoothing is similar with the procedure described in section 3.1.2.2. Initial center points were derived using the non-maximum suppression discussed in 3.1.2.3. Last step of polyp candidate generation is the candidate surface processing which includes the Gaussian center and radius detection and a surface convexity test.
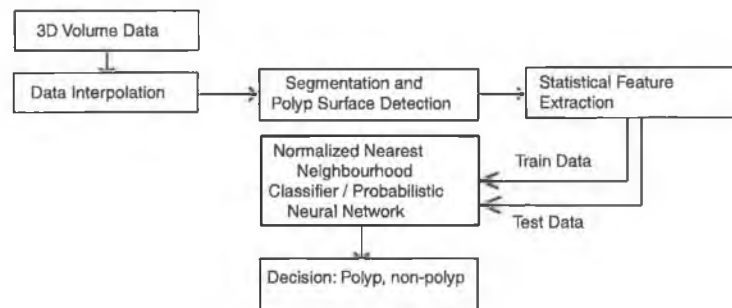
Figure 3.15: Statistical feature-based algorithm for polyp detection.

3.2.1   Candidate Surface Processing

To remove the non-convex surface points from the initial cluster, the Candidate Surface Processing calculates the Gaussian mapping on each cluster to calculate the Gaussian center and radius as discussed in Section 3.1.2.4.

To remove the non convex surface points from the initial cluster two different convexity tests were performed on the initial cluster.

1. The first step involves the convexity test described in Section 3.1.2.4. In this method, a non-convex surface point S will be removed from the cluster if the dot product ($< \overline{SQ}, \overline{n} >$) is less than zero, where $\overline{n}$ is the normal vector.

2. The second step aims to further refine the candidate surface and is based on the surface convexity test proposed by Kiss et al. [94]. For each voxel p_1 belonging to a cluster(W), a bounding box B is defined. As suggested by Kiss et al. [94] the dimension of the bounding box was set to 4. For each voxel $p_2 \epsilon B \bigcap W$, the normal incidence analysis is evaluated (see Figure 3.16) and the values V_c and V_t are computed. V_c defines the number of voxels situated in $B \bigcap W$ that satisfy T_{convex}, while V_t represents the total number of voxels in $B \bigcap W$. Finally, those points ($p1$) where V_c/V_t is higher than T_{hits} remain in the cluster while the others are removed. Threshold values for T_{convex}, T_{hits} were experimentally set to 0.4 and 0.2 respectively (Kiss et al. [94]).
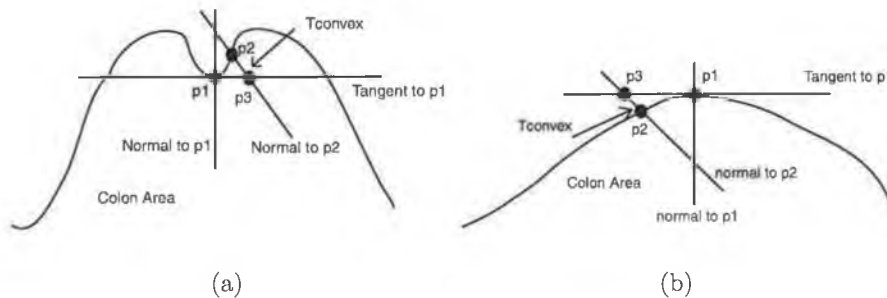
Figure 3.16: Convexity test. (a) voxel p_1 is convex because the intersection between the tangent to p_1 and the normal vector of the neighboring voxel p_2 is inside the colon area. (b) voxel p_1 is non-convex cbecause the intersection between the tangent to p_1 and the normal vector of the neighboring voxel p_2 is ot side the colon area.

After the removal of the non-convex surface voxels from the candidate surface, each cluster was further processed to evaluate discontinuities in the candidate surface. If discontinuities exist in the candidate surface, the cluster was divided into

multiple clusters. Figure 3.17 shows three polyp surfaces and in Figure 3.18 three fold surfaces obtained after candidate surface processing are illustrated.

(a)   (b)   (c)
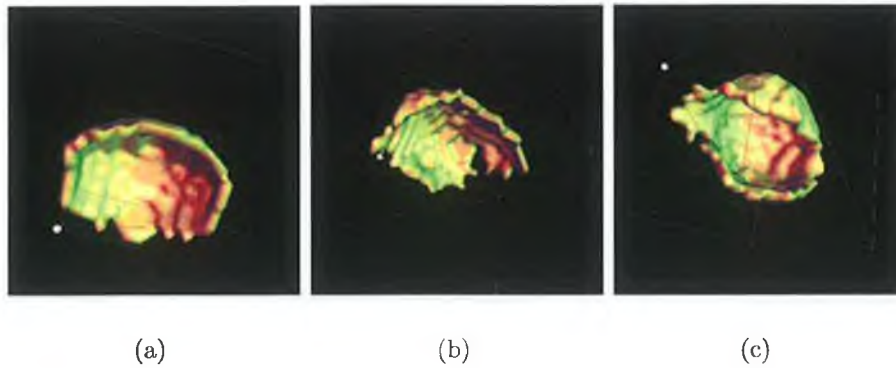
Figure 3.17: 3D surface of three polyps obtained after candidate surface processing.

(a)   (b)   (c)
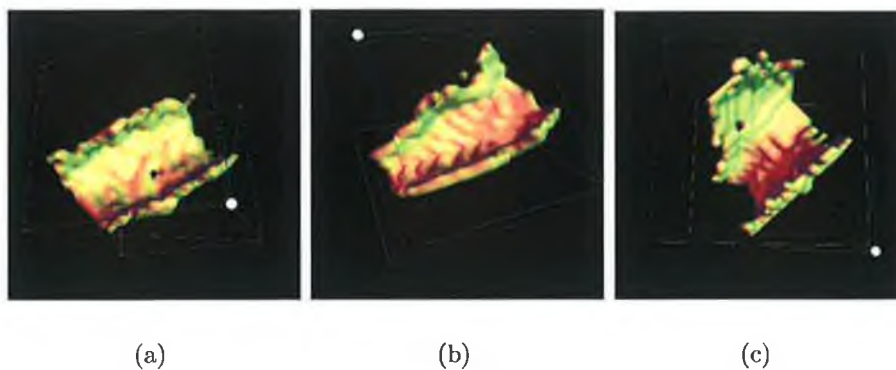
Figure 3.18: 3D surface of three folds obtained after candidate surface processing.

3.2.2   Statistical features extraction

The main objective of this technique is to extract features from the candidate surfaces that offer the best discrimination between polyps and folds. Recall that the polyps can be modelled as spherical or elliptical in shape whereas folds can be modelled as cylindrical. Thus, most of the surface normals of the voxels associated with a polyp surface intersect close to the center of the surface (see Figure 3.19). For fold surfaces, the normals for surface voxels intersect along the principal axis of the cylinder as illustrated in Figure 3.20. To differentiate polyp and fold surfaces, a number of features are calculated from the candidate surface based on the variation of the concentration of the surface normals with respect to the center of the surface. In this regard, a set of statistical features were extracted from the candidate surface.

The statistical features include standard deviation (SD) of surface variation, SD of the three axes of the ellipsoid, SD of the sphere radius, SD of the ellipsoid fitting error, SD of the sphere fitting error, Gaussian distribution, principal axes of the ellipsoid and sphere radius.

        (a)   (b)   (c)
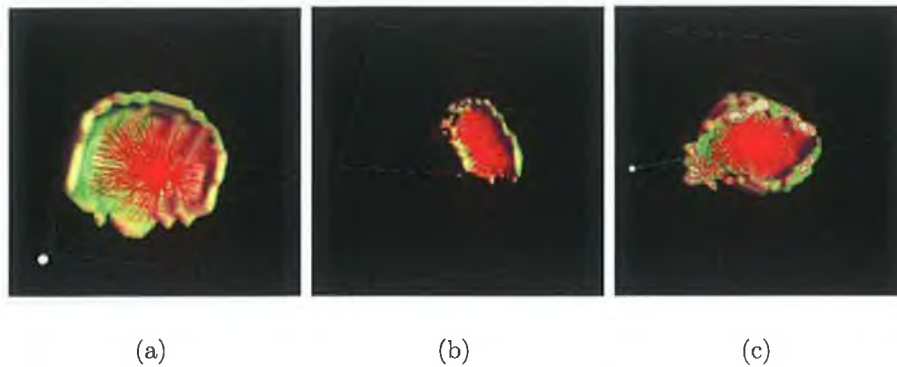
Figure 3.19: Normal concentration for three polyp candidate surfaces.

        (a)   (b)   (c)
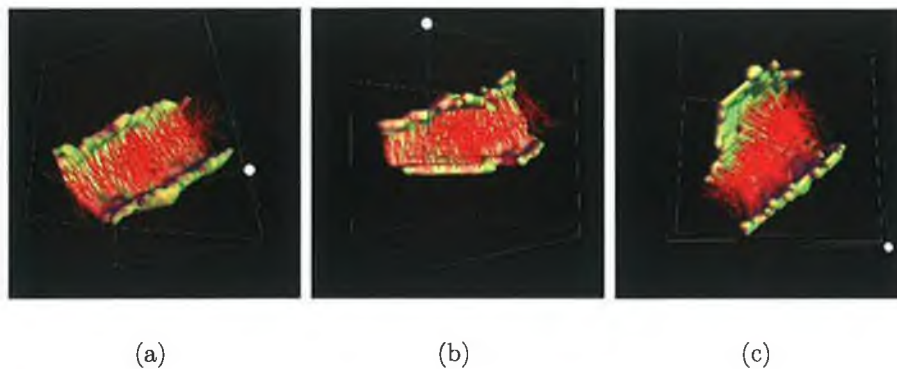
Figure 3.20: Normal concentration for three fold candidate surfaces.

Standard deviation (SD) of the surface variation: The aim of this feature is to evaluate the rate of surface change. In order to evaluate the standard deviation (SD) of the rate of change for a candidate surface the number of surface voxels were calculated at each radius starting from d_{max} towards the minimum radius that was set to 1mm. The goal of this procedure is to determine how many voxels from the candidate surfaces are situated at a particular distance with respect to the center, this will generate the surface number SN. The equations required to calculate the surface number SN_j for each radius are illustrated in equations 3.2.1 to 3.2.3, where N is the number of steps required to sample the surface curvature.

$$Step = (d_{max} - 1.0)/N \qquad (3.2.1)$$

$$R_j = d_{max} - Step \times j \qquad for \quad j = 1, ..., N, \qquad (3.2.2)$$

$$SN_j = \sum_{R_j} Voxel \qquad (3.2.3)$$

Figures 3.21 and 3.22 illustrate the voxel distribution with respect to each radius R_j for different classes of polyps and folds.
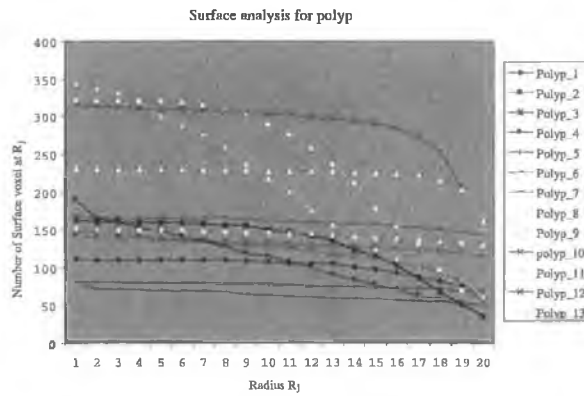
Figure 3.21: Number of surface voxels for each radius (R_j) for polyp classes.
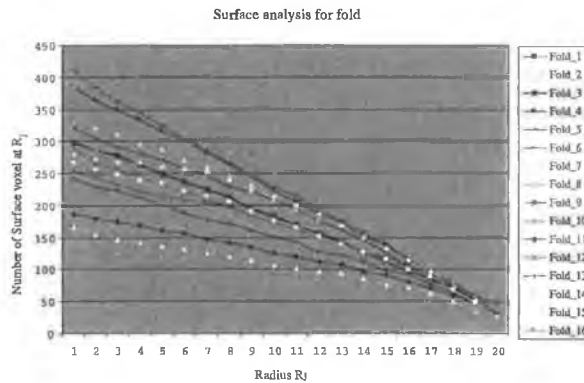
Figure 3.22: Number of surface voxels for each radius (R_j) for fold classes.

From these images (see Figures 3.21 and 3.22) it can be observed that the number of voxels for folds decrease rapidly while for polyps it is almost constant. Thus the surface number can be used to determine the change in curvature and this is best sampled by the standard deviation (SD) that is calculated as illustrated in equations 3.2.4 to 3.2.6.

$$SN_{jmean} = \frac{1}{N} \sum_{j=1}^{N} SN_j \qquad (3.2.4)$$

$$SN_{jnorm} = \frac{SN_j}{SN_{jmean}} \qquad for \quad j = 1, ..., N, \qquad (3.2.5)$$

$$SN_{SD} = \sqrt{\frac{1}{N} \sum_{j=1}^{N} (SN_{jnorm} - SN_{jmean})^2} \qquad (3.2.6)$$

The discrimination offered by the standard deviation (SD) of the surface variation for different classes of polyps and folds is depicted in Fig. 3.23. It can be observed that this feature is quite effective in discriminating polyps from all types of folds.
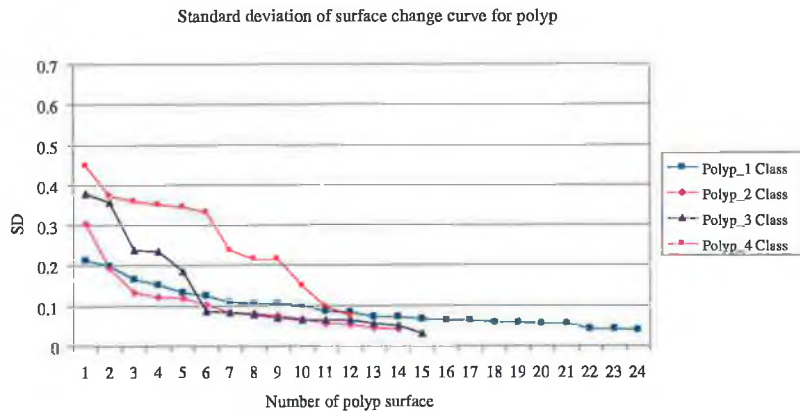
(a)
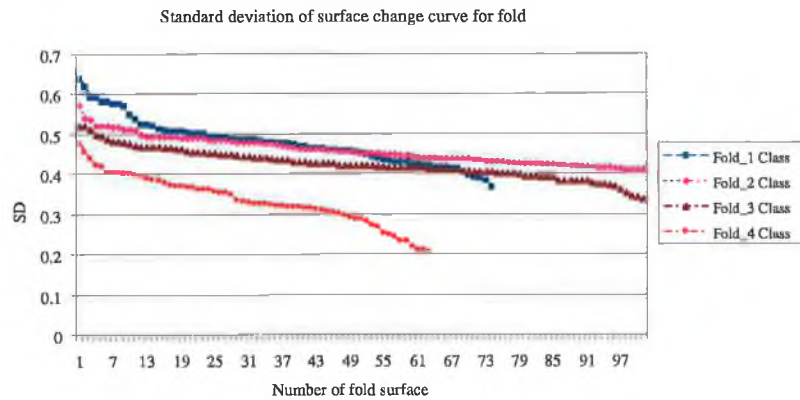
(b)

Figure 3.23: Standard deviation of the surface variation. (a) and (b) show the SD of the surface change for different classes of polyps and folds respectively (classes are sorted in ascending order with respect to the size of the polyps/folds).

SD of the three axes of the ellipsoid: changes in the radius R_j for each candidate surface from 1 to N in equations 3.2.1 to 3.2.3 create N surfaces for each radius R_j.

Let PCS_j^i be the N number of surfaces for a polyp candidate surface PCS^i (where i varies from 1 to the number of candidate surface in the dataset). Least square ellipsoid fitting [104, 105] was employed on each PCS_j^i (where j varies from 1 to N) surface to calculate the three axes of the ellipsoid. Then the SD of the three axes of the ellipsoid for each PCS^i surface are calculated using the equations 3.2.4 to 3.2.6. The discrimination offered by the standard deviation (SD) of the three axes of the ellipsoid for different classes of polyps and folds is depicted in Figures 3.24, 3.25 and 3.26. It can be observed that this feature is effective in discriminating polyps from all types of folds.
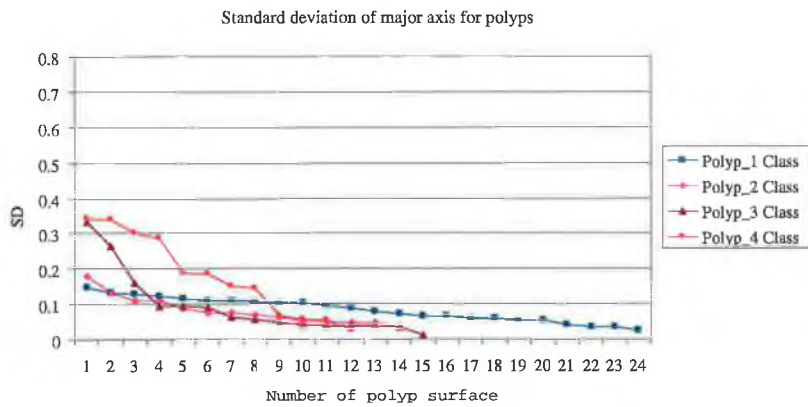
(a)
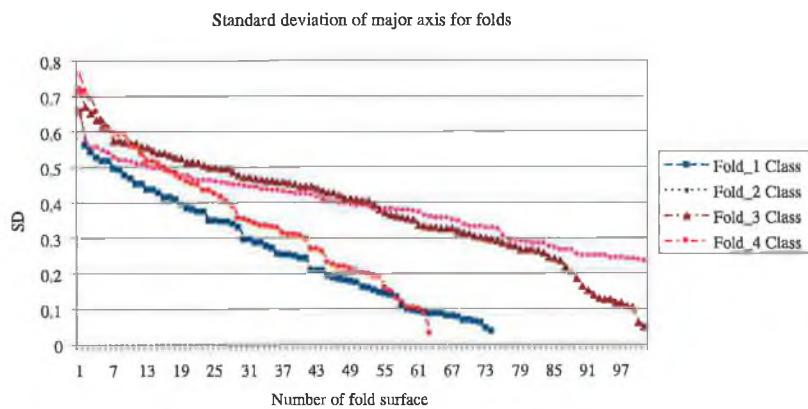
(b)

Figure 3.24: Standard deviation of the major axis of ellipsoid fitting. (a) and (b) show the SD of the major axis of different classes of polyps and folds (classes are sorted in ascending order with respect to the size of the polyps/folds).

Similarly, the SD of ellipsoid fitting error, SD of sphere radius and SD of sphere fitting error were calculated for each candidate surface. Figure 3.27 illustrates the

Standard deviation of second axis for polyps

(a)

Standard deviation of second axis for folds
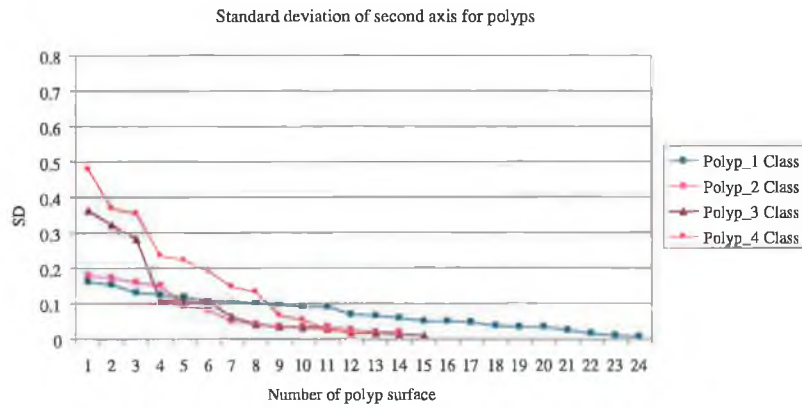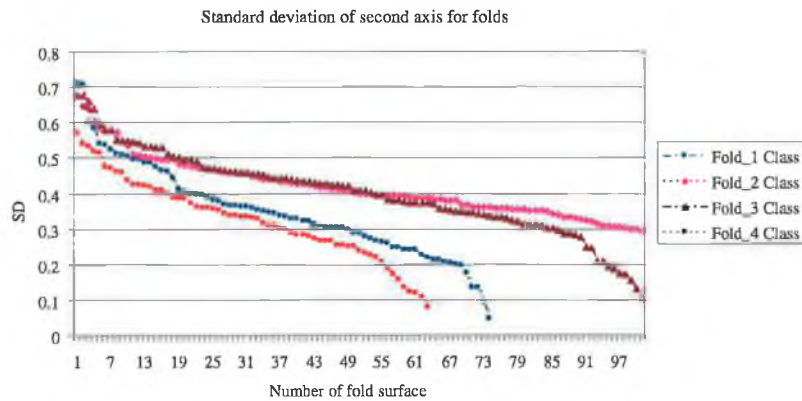
(b)

Figure 3.25: Standard deviation of the second axis of ellipsoid fitting. (a) and (b) show the SD of the second axis of different classes of polyps and folds respectively.

SD of ellipsoid fitting error for different classes of polyps and folds. This feature offers good discrimination between fold and polyp candidate surfaces. Figures 3.28 and 3.29 depict the plot of the SD of the sphere radius and sphere fitting error for polyp and fold surfaces. It can be observed that both SD of the sphere radius and sphere fitting error show effective discrimination between polyp and fold surfaces.

The other features that are used in this method are the Gaussian distribution, length of the ellipsoid major axis and sphere radius. All the above mentioned features were input for the *FNNN* and *PNN* classifiers discussed in Section 3.1.4. The *FNNN* training databases consist of four polyps and four folds databases. The polyps used for training were segregated into small spherical, medium spherical, big spherical, non-spherical polyp. The fold database was also divided into small folds, medium folds, large folds and convex surfaces. In Figures 3.24, 3.25, 3.26,

Standard deviation of third axis for polyps

(a)

Standard deviation of third axis for folds
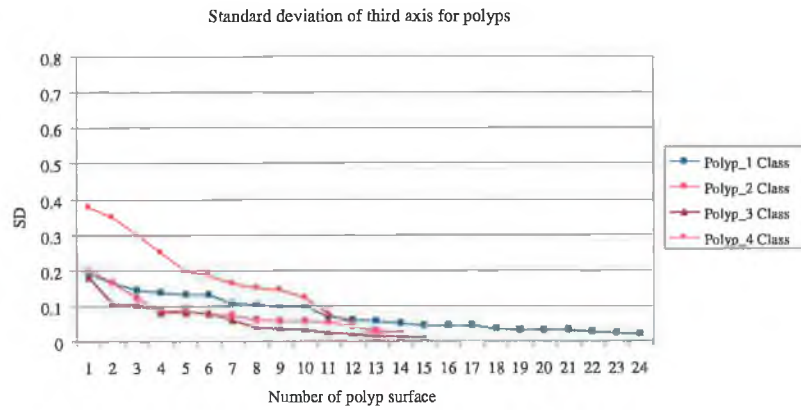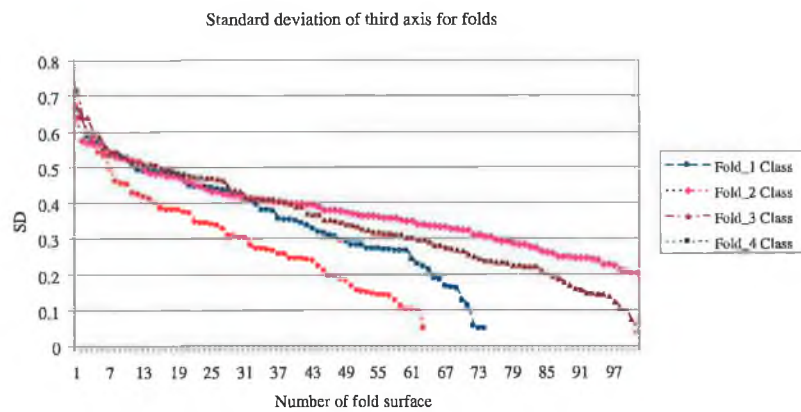
(b)

Figure 3.26: Standard deviation (SD) of the third axis of ellipsoid fitting. (a) and (b) show the SD of the third axis of different classes of polyps and folds respectively.

3.27, 3.28, 3.29 *class_1 polyp, class_2 polyp, class_3 polyp* and *class_4 polyp* represent small, medium, large and non-spherical polyps respectively and *class_1 fold, class_2 fold, class_3 fold* and *class_4 fold* represent large folds, medium size folds, small folds and small convex surface respectively. In total 67 polyps and 348 folds were used to train the *FNNN* and *PNN* classifiers.

3.3   3dB Feature-based approach

The geometrical feature-based approach discussed in the previous section employed least square approximation (ellipsoid, spherical) for analysis of the geometrical shape of the candidate surface in order to extract the features used for polyp and fold classification. The statistical feature-based approach also used least square approximation

Standard deviation of ellipsoid error for polyps

(a)

Standard deviation of ellipsoid error for folds
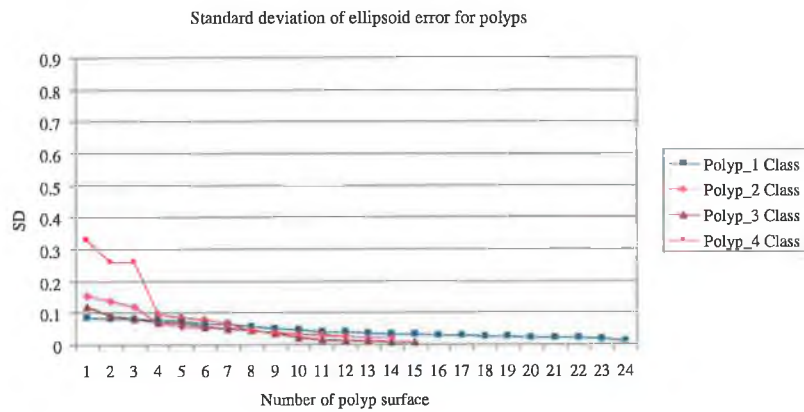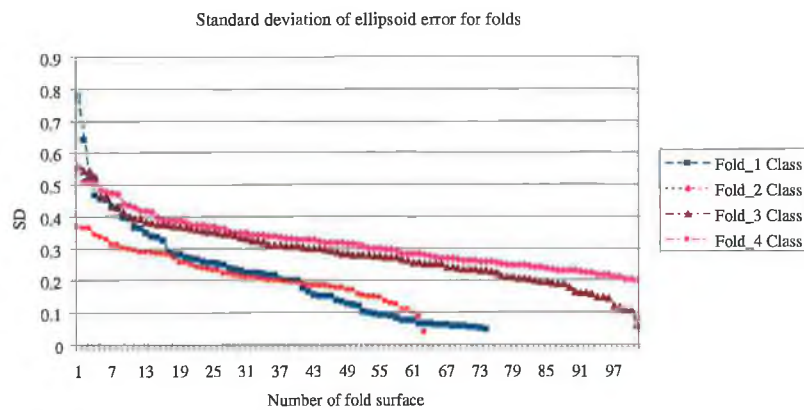
(b)

Figure 3.27: Standard deviation (SD) of the ellipsoid fitting error. (a) and (b) show the SD of the ellipsoid fitting error for different classes of polyps and folds respectively (classes are sorted in ascending order with respect to the size of the polyps/folds).

(ellipsoid, sphere) for calculating the statistical feature for candidate polyp surface. The least square approximation is a maximum likelihood estimator and tries to find the fitted parameter from a particular data set. The least square approximation performs poorly when the data is sparse or noisy. To avoid the problems associated with the least square approximation, the 3dB feature-based method evaluates the geometry of the local colon surfaces by analysing the variation in the candidate surface. The proposed method consists of five steps similar to the statistical feature based method outlined in Figure 3.15. The steps of the algorithm including data interpolation, colon segmentation, polyp candidate surface generation used in the development of the 3dB feature-based are discussed in detail in Section 3.2.1. After
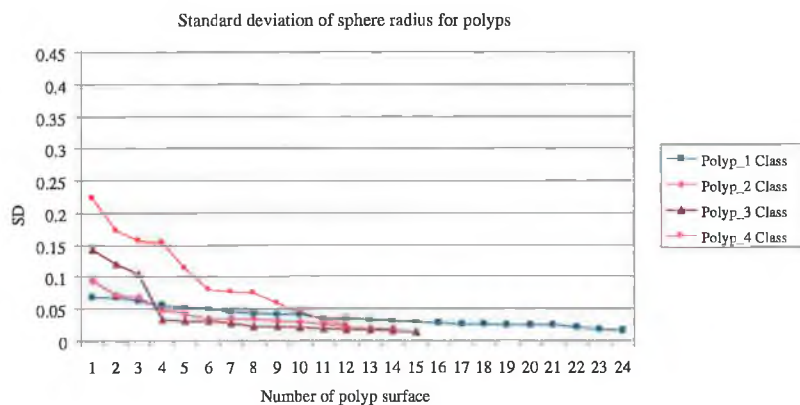
(a)
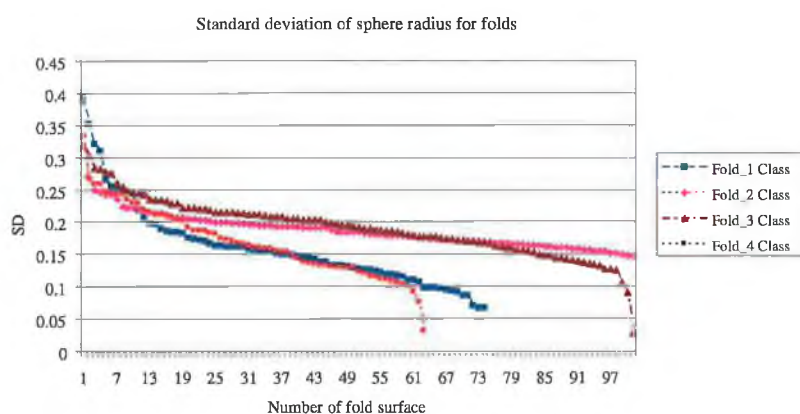
(b)

Figure 3.28: Standard deviation (SD) of the sphere radius of polyp and fold surfaces. (a) and (b) show the SD of the sphere radius for different classes of polyps and folds respectively.

candidate surface generation the features that are calculated for classification are the maximum distance from the cluster center to the surface normal, the standard deviation (SD) of the surface variation, the 3-decibel (dB) attenuation point on the surface change curve and the surface number concentration. In the remainder of this section these features will be presented in detail.

Standard deviation (SD) of the surface variation: The aim of this feature is to evaluate the rate of surface change and its calculation is discussed in Section 3.2.2.

Maximum distance calculation: The maximum distance between the center of the candidate surface and the normal vectors of the candidate surface shows a good discriminative power in separating spherical surfaces from cylindrical surfaces. In this regard, the maximal distance should be significantly higher if the candidate sur-

61

Standard deviation of sphere fitting error for polyps

(a)

Standard deviation of sphere fitting error for folds
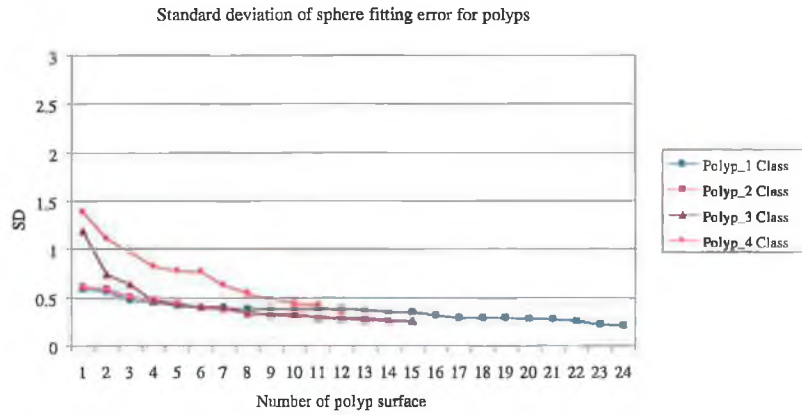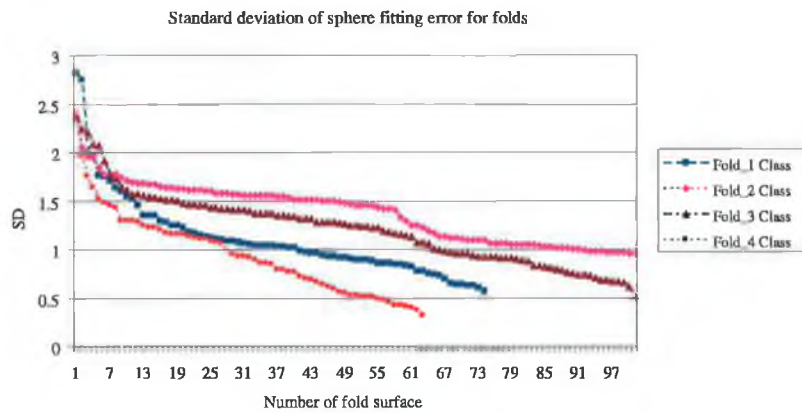
(b)

Figure 3.29: Standard deviation (SD) of the sphere fitting error for polyp and fold surfaces. (a) and (b) show the SD of the sphere fitting error for different classes of polyps and folds respectively.

face belongs to a fold class than in cases when they belong to a polyp class. This can be observed in Figure 3.30. where the maximum distance d_{max} is plotted for different classes of polyps and folds. From Figure 3.30 it can be noticed that this feature is effective in discriminating small/medium polyps ($< 10mm$) when compared to folds. The maximum distance d_{max} does not provide optimal discrimination when the size of the polyp is higher than 10mm (see the plot for class polyp 3 in Figure 3.30).

The 3dB attenuation point on surface change curve: The 3dB point refers to the number of steps required by the SN_j to reach the 3 dB ($3dB = \sum_{j=1}^{N} SN_j/\sqrt{2}$) fall in the total voxel count of the candidate surface. The number of steps required to reach the 3dB point is generally higher for polyps than for folds and this is illustrated

Maximum Distance for polyps and folds
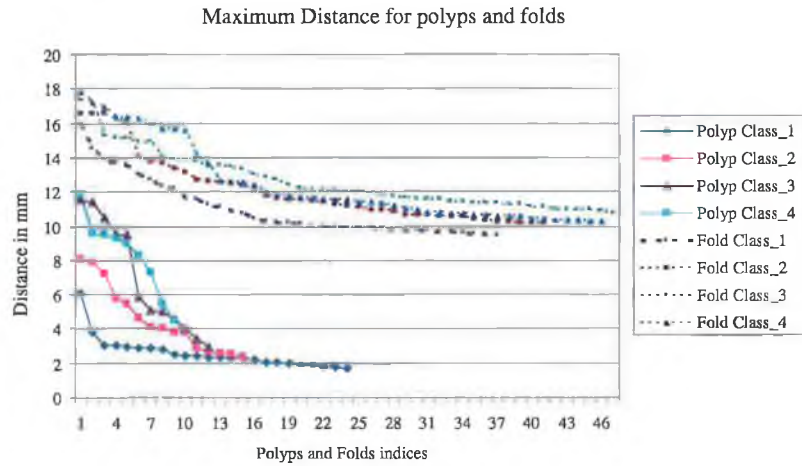
Figure 3.30: Maximum distance d_{max} for different classes of polyps and folds (classes are sorted in ascending order with respect to the size of the polyps/ folds)

3dBstep for polyps and folds
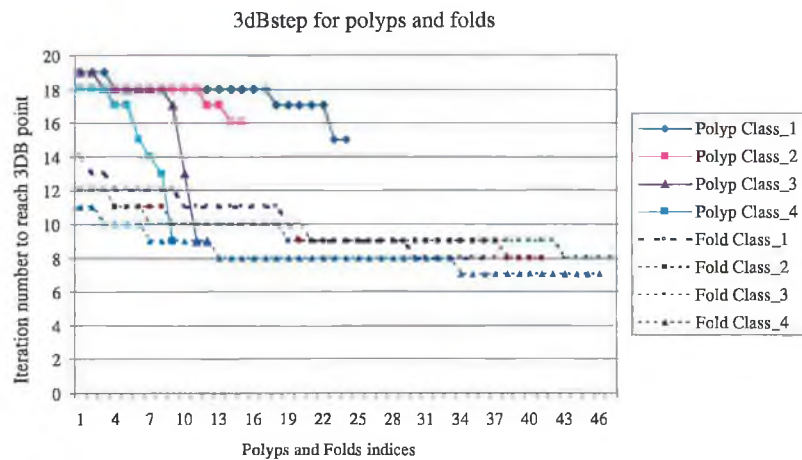
Figure 3.31: The number of steps required in reaching the 3dB point on surface change for different classes of polyps and folds.

in Figure 3.31. This feature is useful in discriminating small and medium polyps from all types of folds.

Surface normal concentration: Recall that for each colon wall voxel 8 HP's were created along the direction of the normal vector from 2.0mm to 10.0mm and the Gaussian distribution has been used to determine the surface center. The normal concentration is given by the number of surface points that generate intersections within 1.25mm from the calculated surface center. As the shape of polyps resembles a spherical surface it is expected that the surface normal concentration to be higher than that calculated for folds (see Figure 3.32). In Figure 3.32 it can be observed that the surface normal concentration offers a good discrimination between large polyps ($\geq 10mm$) and all types of folds. This is very useful as the features
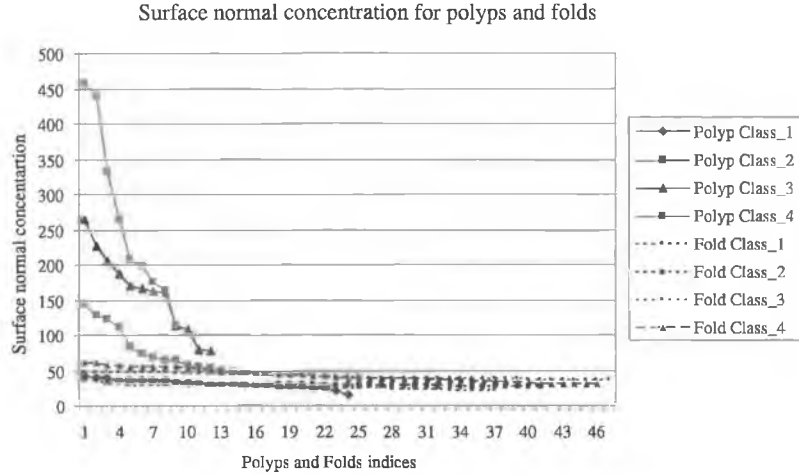
Figure 3.32: Surface normal concentration for different classes of polyps and folds.

discussed before were able to discriminate robustly only small/medium polyps while the discrimination for large polyps was less pronounced.

3.3.1   Polyp/fold classification

To classify the candidate surface into polyps or folds, the calculated features are the inputs for three different classifiers named *FNNN*, *PNN* and Support Vector Machines (SVM). The *FNNN* and *PNN* classifiers are discussed in Section 3.1.4. The Support Vector Machines (SVMs)[108, 109, 110] are powerful tools for data classification. For classification purposes, SVMs find a hypersurface in the space of possible inputs. The hypersurface is generated by the border between positive and negative samples contained in the training set and the classification results are less accurate if the patterns associated with the test data are positioned close to the boundary of the hypersurface. There are several kernels developed for creating different types of hypersurfaces while the linear, radial basis, polynomial and sigmoid are most common *SVM* kernels used for classification. For instance, if a training set of instance-label pairs are (x_i, y_i), $i = 1, ..., l$ where $x_i \epsilon R^n$ and $y \epsilon (1, -1)^l$, the *SVM* kernel is: $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$. The four basic kernels mentioned above can be constructed as follows:

- linear: $K(x_i, xj) = x_i^T x_j$

- polynomial: $K(x_i, x_j) = (x_i^T x_j + r)^d, \gamma > 0$.

- radial basis function (RBF):$K(x_i, x_j) = exp(-\gamma \parallel x_i - x_j \parallel^2), \gamma > 0$.

64

- sigmoid: $K(x_i, x_j) = tanh(\gamma x_i^T x_j + r)$.

The implementation software of the SVM was developed by Gunn [109] and the code was written in MATLAB. The SVM classifier was used to show the robustness of the features calculated from the polyp candidate surface in a non-linear feature space and to provide indicative results for CAD-CTC. For the developed CAD-CTC system the 5-th order polynomial kernel was employed for SVM classification. The training set consists of 81 polyps and 348 folds. The same training data was used to train the PNN and $FNNN$ classifiers.

In the next chapter, the construction of a synthetic phantom and the generation of CT test data (standard and low dose) is provided. The development of a synthetic phantom was an important objective of this research because allowed us to investigate the following problems:

- Exposure to the ionizing radiation does not allow the patients to be scanned successively. Thus, the construction of a synthetic phantom was necessary in order to generate test data where the phantom was scanned in different positions. This CT data can be used to evaluate whether the results returned by the CAD-CTC systems are repeatable with respect to the polyp detection. in CAD-CTC

- Evaluate the influence of noise in low-dose on the overall performance of the CAD-CTC

- Detection of optimum scanning parameters that can be used to generate the low-dose CT data feasible to be used by the CAD-CTC systems

- Development of standard testing datsets that can be used in the development phase of the CAD-CTC.

Chapter 4

Phantom

The major concern associated with CTC is the fact that the patients are subjected to high level of ionising radiation. The medical literature indicates that the effective dose used for CTC varies from 5 to 20 mSv [32, 47, 48, 49, 50, 111] and this radiation level may result in a 0.05% risk for inducing cancer in patients older than 50 years [112]. Brenner *et al.* [112] study indicates that the incidence of induced cancer is in direct relation to the effective radiation dose (see Figure-4.1) and the cancer risk increases with decreasing age [113]. In this regard, Cohen [114] shows that the risk of inducing cancer in patients is significantly lowered when they are subjected to low-level radiation exposure and an important number of studies are dedicated to identify the minimal level of radiation dose that can be feasibly used in CTC [111, 115, 116, 117, 118]. The identification of the optimal scanning parameters (collimation, slice thickness, table speed, reconstruction interval) is a difficult problem and this procedure is applied on synthetic phantoms that are designed to accurately model the human body [119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130]. In this sense, Beaulieu et al. [119] used spherical plastic beads to model polyps while Dachman et al. [120] created false polyps in a pig colon by puckering the mucosa of the colon. Their studies focused on finding the imaging effect of collimation, tube current (pitch) and orientation when they analysed different sizes and types of polyps. Similar studies were performed by Taylor et al. [121] and Springer et al. [122]. Whithing et al. [123] used a different approach and in order to evaluate the artefacts generated by the collimation and the tube current they constructed an air filled acrylic cylinder were synthetic polyps of different sizes were attached on the inner side of the acrylic tube. Laghi et al. [126] and Embleton et al. [127] used synthetic and pig colons and their tests indicate that a collimation of 4×2.5 *to* $1.25mm$ reconstruction interval, tube current of 40 mAs are satisfactory parameters to be

used for clinical *CTC* examinations. Ozgun et al. [128] used latex material to build phantom polyps having dimensions ranging from 1mm to 10mm. Their tests were focused on finding the minimal tube current that allows the detection of polyps larger than 5mm. They reported that the detection of all types of polyps larger than 5mm is feasible only at current tubes in the range 60mAs to 100mAs.
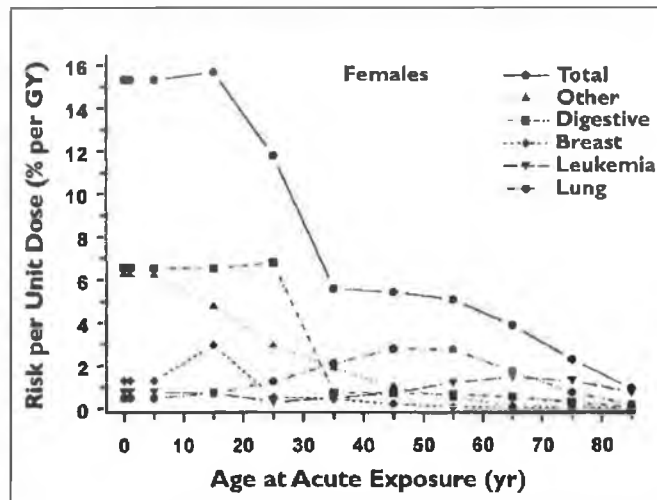
(a)
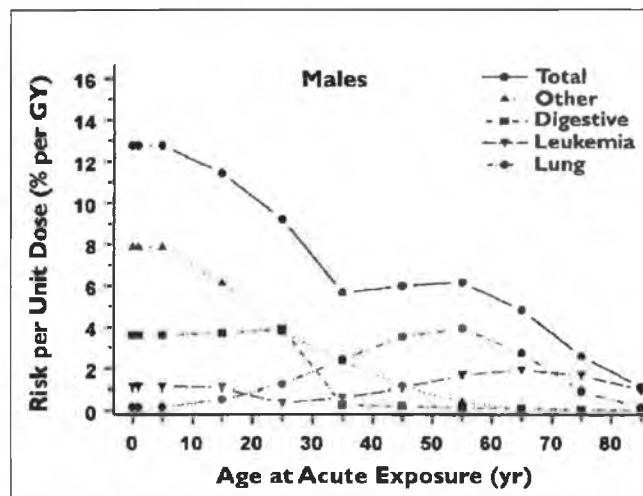
(b)

Figure 4.1: Breakdown by cancer type. (a) and (b) show the lifetime attributable cancer mortality risks as a function of age at a single acute radiation exposure for females (a) and males (b) as estimated by the National Academy of Sciences BEIR V (Biological Effects of Ionizing Radiations) committee [113].

The aim of this chapter is to study the effect of all scanning parameters (mAs, slice thickness, reconstruction interval, field of view, table speed) using a novel

synthetic phantom. The phantom has been specifically designed for *CAD-CTC* to simulate colon polyps with different shapes (pedunculated, sessile and flat) and sizes (3 to 18mm). In this studies the data is evaluated using a developed automated *CAD-CTC* system in order to determine the influence of the scanning parameters on polyp detection. A special emphasis of this study is placed on determining the minimal radiation dose that allows robust identification of colonic polyps but not at the expense of reduced sensitivity in polyp detection.

4.1 Materials and Methods

4.1.1 Phantom design

A synthetic phantom was constructed using a PVC tube, two acrylic tubes, two plastic plates and latex material to emulate the colon wall, polyps and folds. The external PVC tube is 230mm long with a diameter of 300mm. Acrylic tubes are 235mm long and the dimensions of the inner and outer diameters are 40mm and 50mm respectively. *Hounsfield Unit* (HU) values of the PVC tube, acrylic tubes and plastic plates are 1500, 100, 90 respectively. The construction of the synthetic phantom is illustrated in Figure 4.2.

(a) (b)
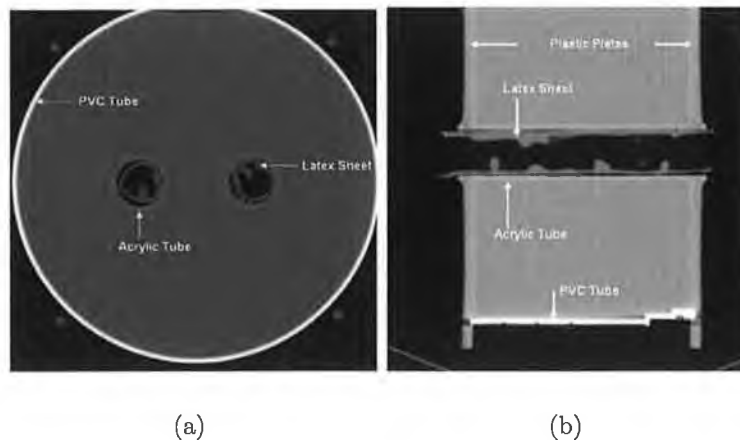
Figure 4.2: Synthetic colon phantom. (a) Longitudinal view. (b) Transversal view.

The polyp inserts for phantom were made by latex material having a *HU* value of -95. We have chosen to use latex as this material allows us to generate very realistic shapes (pedunculated, sessile, flat, flat-depressed) for polyps and folds as illustrated

in Figure 4.3. In addition the *HU* values associated with the latex material approximate well the *HU* values of the colon wall ($\sim$10HU). In *CTC* the large difference between the *HU* values associated with the air voxels (-1000HU) and the *HU* values of the colon tissue is evaluated to identify the surface of the colon wall. The model for polyps was made from clay and liquid latex was poured onto the model to create the latex polyp inserts (see Figure 4.3). To make the surface of the latex sheet more realistic the thickness of the sheet was made uneven. We have created two sheets of latex containing 48 polyps having different sizes (7 flat polyps, 2 depressed flat polyps, 15 non-spherical polyps, 2 pedunculated polyps, 22 spherical/elliptical polyps) and 6 haustral folds. In Figure 4.4 several 3D views of some representative synthetic polyps are depicted.

4.1.2   Image acquisition

The developed phantom described in Section 2.1 was scanned using a 16-slice Siemens Somatom Sensation CT scanner in the Mater Hospital, Dublin, Ireland. The phantom has been scanned in longitudinal (phantom was placed parallel to the CT scanner bed) and transversal directions, where the scanning parameters (slice thickness, field of view, table speed, reconstruction interval and mAs) were varied. All scans were performed at 120kVp and $1.5mm \times 16$ collimation. It is useful to note that the effective radiation dose is influenced by the value of the tube voltage but its relationship with image quality, tissue contrast and image noise is complex and the effect of this parameter would be difficult to be evaluated. Therefore, in this experiments the value of this parameter is maintained constant (120 kVp) and another reason is the fact that this is the standard value of the tube voltage used in clinical examinations. The scanner used in to generate the CT data allows the possibility to adjust the value of collimation to 0.75mm but the value of collimation was fixed at 1.5mm in order to reduce the radiation dose. In this regard, a CT scan performed with 1.5mm collimation and 3mm slice thickness will result in an energy imparted of 7.0mSv while the energy imparted for a CT scan with 0.75mm collimation and 3mm slice thickness is 7.8mSv which is to high to be used safely in clinical studies. The smoothing reconstruction filter used was the B30 filter [131] and this filter has been employed based on its optimal performance in data smoothing and noise removal (this is the filter used in most clinical studies for abdominal CT scans).

In conjunction with our clinical partners from Mater Hospital we have chosen
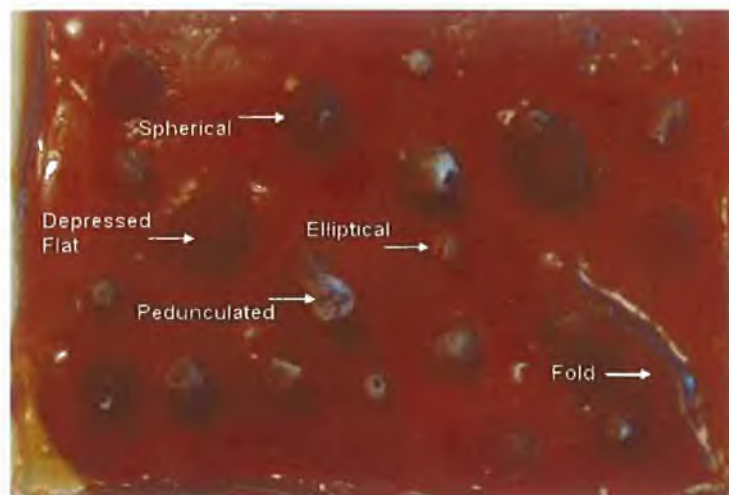
(a)
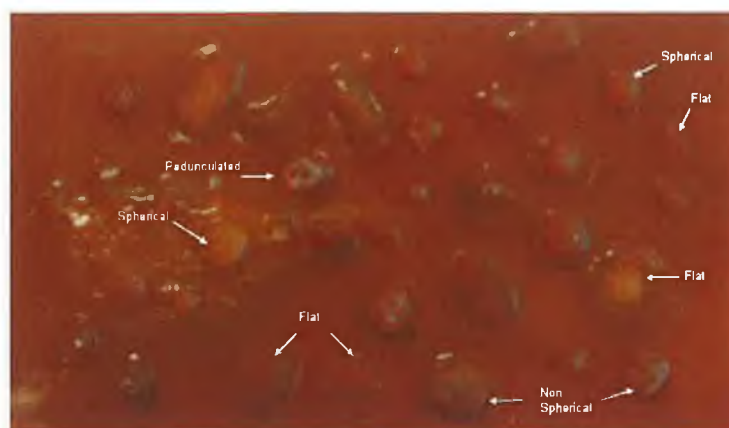
(b)

Figure 4.3: Latex sheet with various types of polyps and folds.

the following spread of parameters: field of view: 325 and 360mm, table speed: 20 to 30 mm/rotation, slice thickness of 2 and 3mm and mAs: 100, 80, 70, 60, 50, 40, 30, 20 and 13 (13 mAs is the minimum value that can be set for Siemens Somatom Sensation CT scanner used in these experiments). These scanning parameters have been divided into six protocols as follows:

- Protocol 1: Collimation 1.5 × 16mm, slice thickness 3mm, reconstruction interval 1.5mm, field of view 325mm, table speed 30mm/rotation, mAs: 100, 80, 70, 60, 50, 40, 30, 20 and 13. This protocol was used to identify the effect of radiation dose and scan orientation (longitudinal and transversal scans) on

(a)   (b)

(c)   (d)
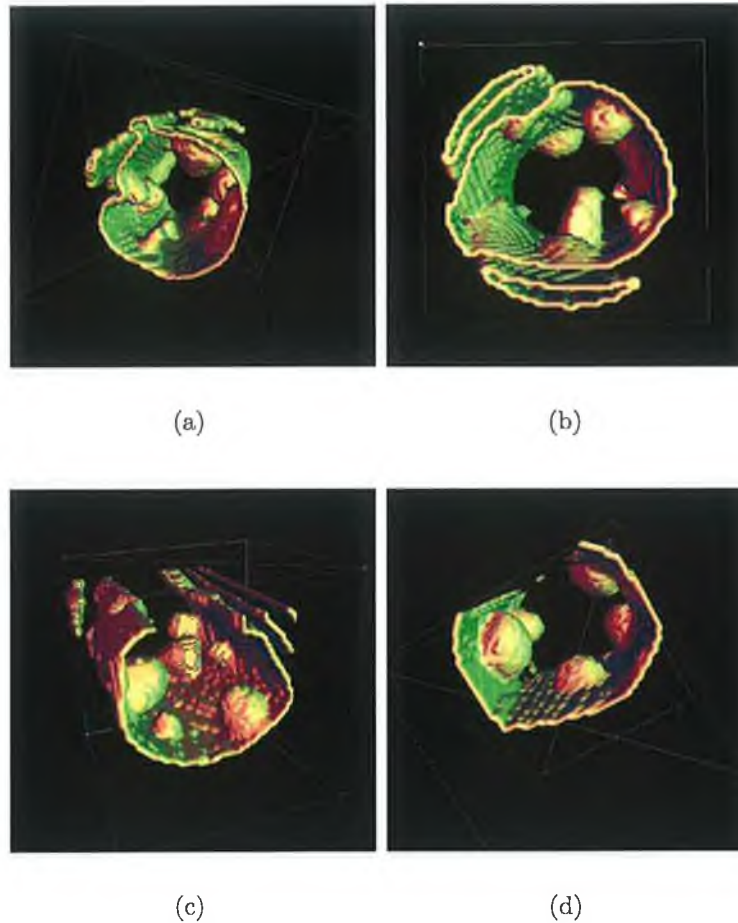
Figure 4.4: 3D longitudinal views of the synthetic polyps made from latex.

the performance of the developed automatic *CAD-CTC* system.

- Protocol 2: Collimation 1.5 × 16mm, slice thickness 3mm, reconstruction interval 1.5mm, field of view 360mm, table speed 30mm/rotation, mAs: 50, 30, 20 and 13. This protocol was employed to evaluate the influence of the field of view and the variation of the radiation dose.

- Protocol 3: Collimation 1.5 × 16mm, slice thickness 3mm, reconstruction interval 1mm, field of view 325mm, table speed 30mm/rotation, mAs: 100, 80, 70, 60, 50, 40, 30, 20 and 13. This protocol was used to analyse the effect of the reconstruction interval and the radiation dose.

- Protocol 4: Collimation 1.5 × 16mm, slice thickness 2mm, reconstruction interval 1mm, field of view 325mm, table speed 30mm/rotation, mAs: 100, 50, 40, 30, 20 and 13. This protocol was used to generate CT data where the effect of the slice thickness and the radiation dose is analysed.

- Protocol 5: Collimation 1.5 × 16mm, slice thickness 2mm, reconstruction interval 0.8mm, field of view 325mm, table speed 30mm/rotation, mAs: 100, 50, 40, 30, 20 and 13. This protocol was employed to analyse the joint effect of the slice thickness, reconstruction interval and radiation dose.

- Protocol 6: Collimation 1.5 × 16mm, slice thickness 3mm, reconstruction interval 1.5mm, field of view 325mm, table speed 20mm/rotation, mAs: 100, 50, 40, 30 and 20. This protocol was used to find the effect of table speed at different radiation doses.

With a multi-slice CT scanner, the selection of the (reconstructed) slice width is independent of patient dose, being solely reliant on collimation selected. Therefore if the collimation remains the same, the selection of a 5mm slice width will generate the same radiation dose as a 3mm slice thickness. It has been found that the Siemens Somatom Sensation 16 slice CT scanner shows similar imparted radiation dose for both 3mm and 5mm (7mSv for both 3mm and 5mm at 100mAs). Thus, it has been decided to ignore the 4mm, 5mm slice thickness in this study since these settings will have virtually no effect on the imparted radiation dose received by the patients.

As already explained, for Siemens Somatom multi-slice CT scanner the variation in the table speed is possible, but by increasing the table speed (reduce the duration of the CT scan) does not vary the patient dose, as this scanner utilises the "effective tube current" model where the mAs is kept constant throughout (a variation in scanning time results in a concomitant variation in mAs). In this study, the table speed has been varied to evaluate the influence of the motion artefacts on the performance of the *CAD-CTC*. Thus, only the 30mm/rotation and 20mm/rotation table speeds were chosen to use in this study.

As mentioned earlier, a collimation of 1.5mm was used for two reasons. Firstly because this setting is recognised as adequate to detect clinically significant colonic polyps (5mm and greater). Secondly while a 0.75 collimation was possible with the Siemens scanner, this setting generates a markedly increased patient dose. A 1.5mm collimation, with 3mm slice width results in a scan time of 10.2sec and an imparted energy of 7.0mSv. The 0.75mm collimation setting with a 3mm slice results in a scan time of 20.14sec and an associated imparted energy of 7.8mSv. This was deemed

to be unacceptable as this radiation dose is too high to be used safely in clinical examinations.

4.1.3   Characterisation of phantom CT data

The method applied for feature detection for phantom polyp is the method based on the statistical features that is discussed in Chapter 3.2. The statistical features include the standard deviation (SD) of surface variation, SD of the three axes of the ellipsoid, SD of the sphere radius, SD of the ellipsoid fitting error, SD of the sphere fitting error, Gaussian distribution, principal axes of the ellipsoid and sphere radius. In this section we evaluate the statistical features for phantom polyps, real polys and folds in order to illustrate the fact that the phantom polyps emulate closely the polyps encountered in clinical studies. Figure 4.5 shows the standard deviation of the surface variation for 45 phantom polyps (14 polyps $\geq 10mm$, 20 polyps between $[5-10)mm$, 5 polyps $< 5mm$ and 6 flat polyps) and 41 real patient polyps and 274 folds with different sizes. The SDs of surface variation for phantom polyps are placed close to those of the real polyps. Similarly, the features: SD of major axis, SD of ellipsoid error, SD of sphere radius and SD of sphere error for phantom exhibit similar characteristics with the real polyps as illustrated in Figures 4.6 to Figure 4.9.
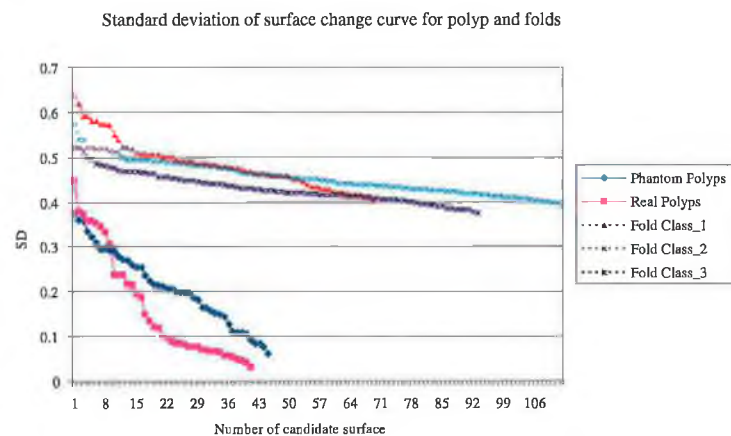
Figure 4.5: Standard deviation of the surface variation for phantom polyps, real polyps and folds.
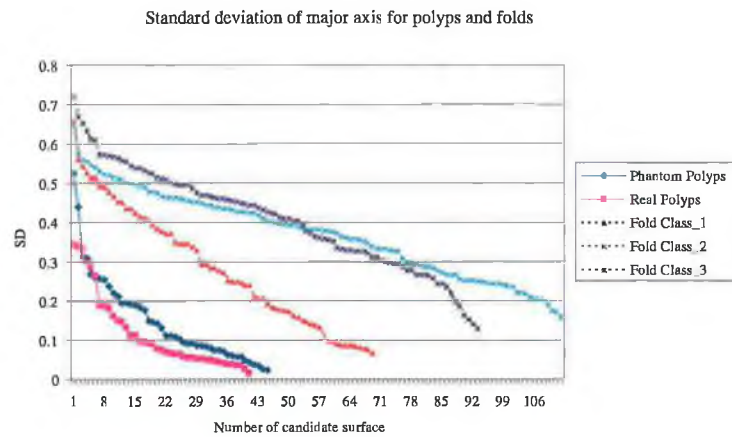
Figure 4.6: Standard deviation of the major axis of ellipsoid fitting for phantom polyps, real polyps and folds.
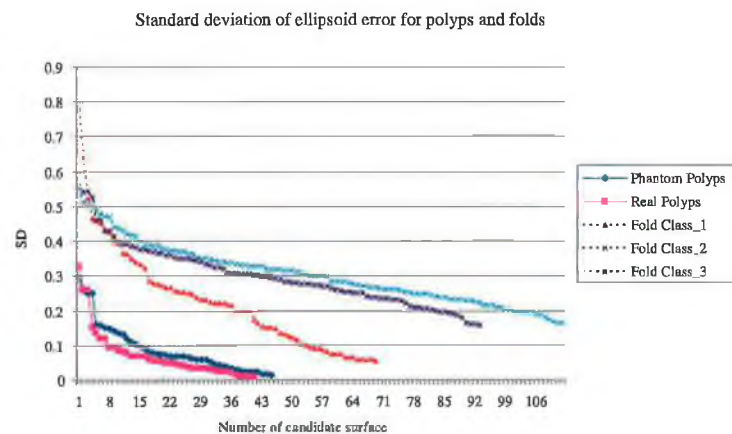
Figure 4.7: Standard deviation (SD) of the ellipsoid fitting error for phantom polyps, real polyps and folds.

4.2   CAD-CTC polyp detection algorithm

The method applied for automatic polyp detection in phantom data was the statistical feature based method discussed in Chapter 3.2.
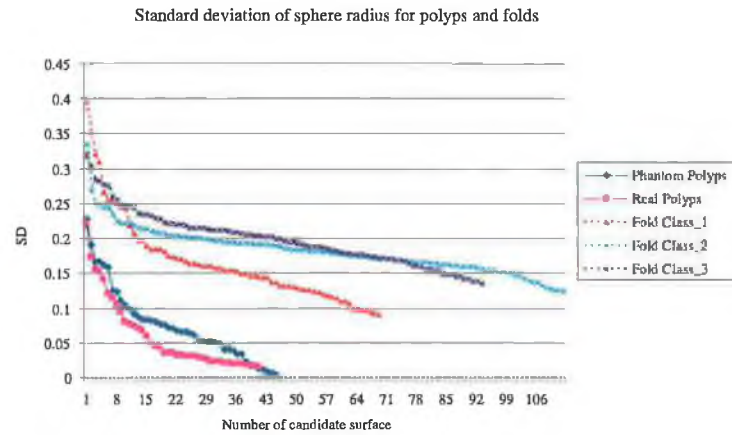
Figure 4.8: Standard deviation (SD) of the sphere radius for phantom polyps, real polyps and folds.
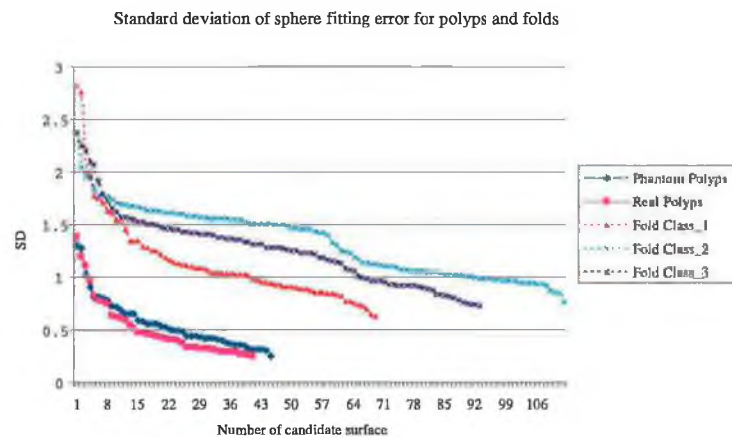
Figure 4.9: Standard deviation (SD) of the sphere fitting error for phantom polyps, real polyps and folds.

4.3   Experiments and results

The aim of this section is to evaluate the influence of the scanning parameters on the overall polyp detection results in *CAD-CTC* systems. In order to evaluate this, the synthetic phantom detailed in Section 2 has been scanned and a total of 46 CT datasets have been acquired using the six protocols mentioned in Section 4.1.2.

When the *CAD-CTC* system has been applied to CT data acquired using the Protocol 1, the results indicate that 100% sensitivity has been achieved for polyps larger than 10mm in both longitudinal and transversal positions for all radiation levels (100 to 13 mAs). For medium size polyps (5mm to 10mm) the sensitivity was

100% in all cases but 20 and 30 mAs, where the sensitivity rate was 95%. The reduction in sensitivity was caused by the undetected polyp illustrated in Figure 4.10a which was situated close to the end plates. The sensitivity in polyp detection when the *CAD-CTC* algorithm was applied to *CT* data acquired using the Protocol 1 is illustrated in Figure 4.11.

(a)   (b)
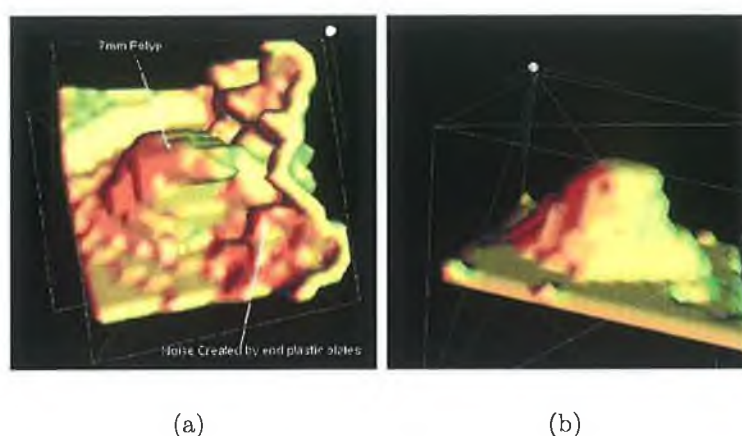
Figure 4.10: (a) Polyp undetected by the *CAD-CTC* algorithm when the CT data was acquired using the Protocols 1, 3 and 6. (b) Polyp undetected by the CAD-CTC algorithm when the data was acquired using the Protocols 2, 4 and 5.

The sensitivity of the *CAD-CTC* technique when applied to CT phantom data acquired using the Protocol 2 is 100% for polyps larger than 10mm. The sensitivity for medium size polyps (5 to 10 mm) dropped to 95% when the phantom was scanned with 30, 20 and 13 mAs. There was only one polyp undetected for data acquired with this protocol and it is illustrated in Figure 4.10b.

For CT data acquired using the Protocol 3, the polyp detection for all scans show 100% sensitivity except the case when the phantom has been scanned with 30mAs. The polyp undetected is illustrated in Figure 4b. The polyp detection sensitivity when the scans were performed using the Protocol 4 is 100% for polyps larger than 10mm for all radiation doses except 100mAs. The sensitivity in polyp detection for medium size polyps is also 100% except in the case where the phantom has been scanned with 30mAs when the sensitivity dropped to 95%. The polyp missed by the *CAD-CTC* system is illustrated in Figure 4a. The sensitivity in polyp detection obtained when the *CAD-CTC* system was applied to CT data scanned using the Protocol 5 is lower than the sensitivity obtained when the Protocols 1 to 4 were employed. The reason for this is that no interpolation was applied to obtain an

Table 4.1: Results of the automated polyp detection for Protocol-1 (Collimation: 1.5×16(mm), Slice Thickness: 3mm, Reconstruction Interval: 1.5mm, Field of view: 325mm, Table Speed: 30mm/rotation.

Protocol	Direction	mAs	Sensitivity(%)					FP
			Large: 14	Medium: 20	Small: 5	Flat: 9	Total: 48	
1	Long	100	14(100)	20(100)	5(100)	4(44.4)	43(89.9)	1
1	Long	80	14(100)	20(100)	5(100)	5(55)	44(91.7)	2
1	Long	70	14(100)	20(100)	5(100)	5(55)	44(91.7)	2
1	Long	60	14(100)	20(100)	5(100)	5(55)	44(91.7)	2
1	Long	50	14(100)	20(100)	5(100)	5(55)	44(91.7)	1
1	Long	40	14(100)	20(100)	5(100)	4(44.4)	43(89.9)	1
1	Long	30	14(100)	19(95)	5(100)	5(55)	43(89.9)	2
1	Long	20	14(100)	19(95)	5(100)	4(44.4)	42(87.5)	2
1	Long	13	14(100)	20(100)	5(100)	4(44.4)	43(89.9)	3
1	Tran	100	14(100)	20(100)	5(100)	2(22.2)	41(85.4)	3
1	Tran	60	14(100)	20(100)	5(100)	3(33.3)	42(87.5)	5
1	Tran	50	14(100)	19(95)	5(100)	2(22.2)	40(83.3)	4
1	Tran	40	14(100)	20(100)	5(100)	2(22.2)	41(85.4)	4
1	Tran	30	14(100)	19(95)	5(100)	2(22.2)	40(83.3)	4
1	Tran	20	14(100)	19(95)	5(100)	3(33.3)	41(85.5)	4
1	Tran	13	14(100)	20(100)	5(100)	3(33.3)	42(87.5)	4

isometric dataset as the reconstruction interval is 0.8mm and the voxel resolution is almost the same in all directions (the lower performance of the *CAD-CTC* system when applied to datasets acquired using the Protocol 5 is justified since the classifier is trained only with interpolated data). Sensitivity achieved for polyp detection when the *CAD-CTC* algorithm has been applied to CT data obtained using the Protocol 6 is 100% for all radiation doses except the case when the data is scanned with 20 mAs. The polyp missed by the polyp detection algorithm is illustrated in Figure 4.10a. Results of the automated polyp detection for all 46 scans used in these experiments are depicted in Figures 4.11 to 4.16. It is useful to note that the overall sensitivity achieved by the developed *CAD-CTC* system is lowered by the inclusion of flat polyps. The sensitivity rate for flat polyps is between 22% to 55% and the developed method has not been designed to detect this class of colorectal polyps. The flat polyps have distinct shapes and their identification should be approached by a *CAD-CTC* system that is specifically designed to deal with this type of polyps [132].
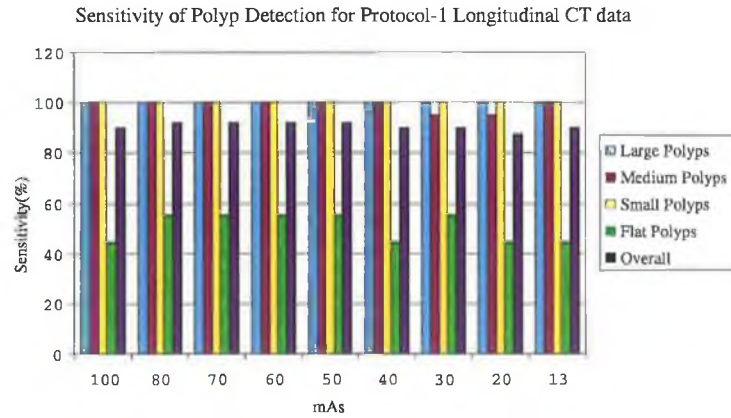
(a)
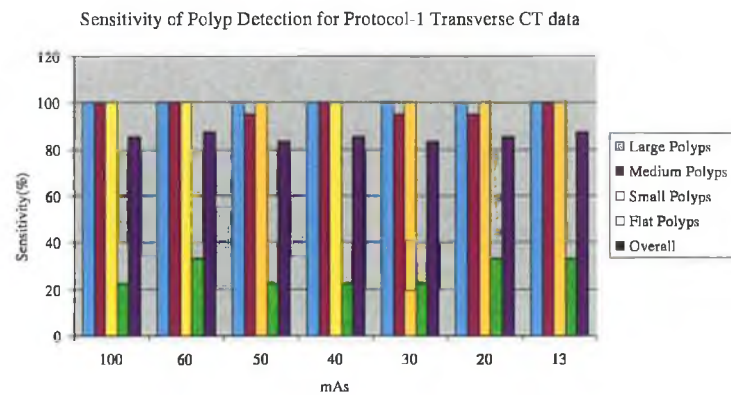
(b)

Figure 4.11: Sensitivity of the polyp detection algorithm when applied to CT data (Protocol-1: Collimation 1.5×16mm, slice thickness 3mm, reconstruction interval 1.5mm, field of view 325mm, table speed 30mm/rotation) acquired at different radiation doses. (a) and (b) show the sensitivities for Protocol-1 longitudinal and transversal CT data respectively.

4.3.1   Effect of slice thickness, reconstruction interval and table speed

To analyse the effect of slice thickness and reconstruction interval, the synthetic phantom has been scanned using protocols where these parameters are varied (Protocols 1,3,4 and 5). An important step preceding the application of the *CAD-CTC* algorithm is data interpolation. All CT datasets were interpolated in order to make them isometric except cases when the phantom was scanned using the Protocol 5. The CT data obtained using the Protocol 5 was not interpolated as the voxel resolu-

78
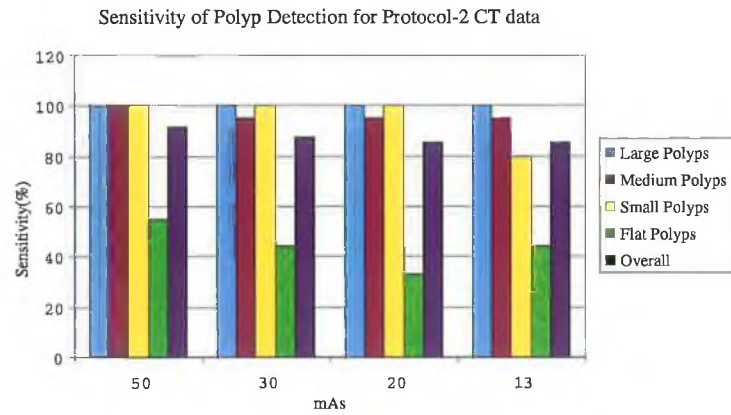
Figure 4.12: Sensitivity of the polyp detection algorithm when applied to Protocol-2 CT data.
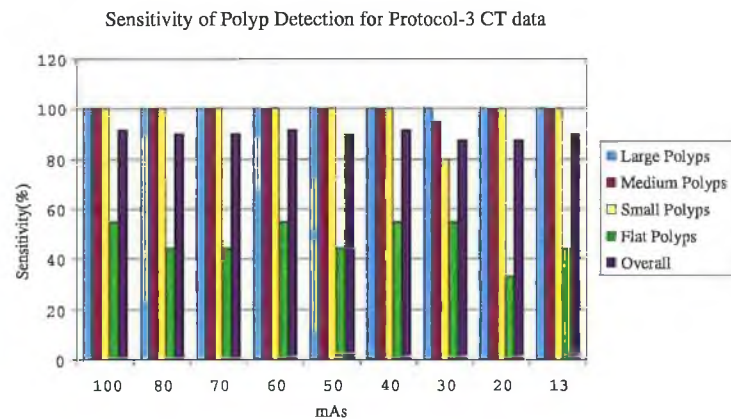
Figure 4.13: Sensitivity of the polyp detection algorithm when applied to Protocol-3 CT data.

tion is almost similar in all directions (voxel width and height: 0.7mm, voxel depth: 0.8mm). The experimental results indicate that the performance of the *CAD-CTC* algorithm is virtually unchanged when it is applied to CT data acquired using the Protocols 1,3 and 4. The results obtained when the algorithm has been applied to data acquired using the Protocol 5 were worse than those obtained when the algorithm was applied to CT data obtained using other protocols. This has been generated by the fact that data interpolation has a smoothing effect on the 3D morphology of the colon wall and another important factor is that the classifier was trained only with interpolated data.

The field of view was set to 360mm for Protocol 2 and to 325mm for other

Sensitivity of Polyp Detection for Protocol-4 CT data
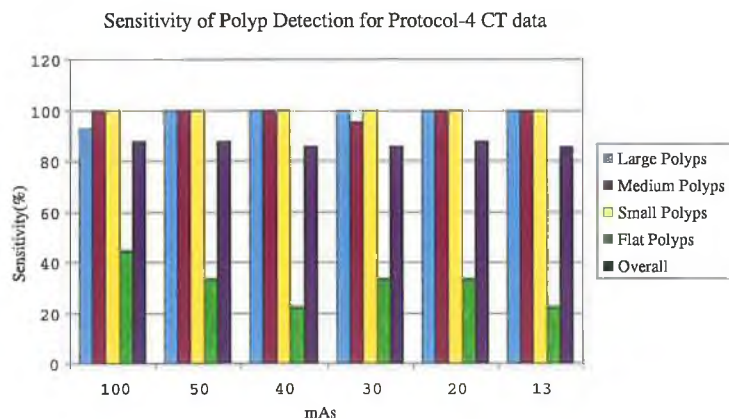
Figure 4.14: Sensitivity of the polyp detection algorithm when applied to Protocol-4 CT data.

Sensitivity of Polyp Detection for Protocol-5 CT data
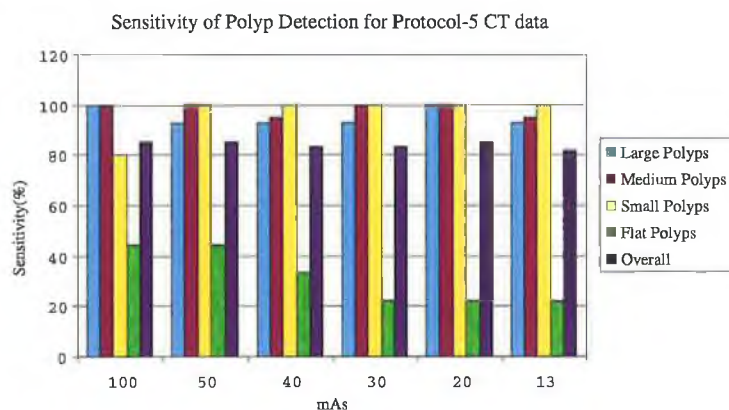
Figure 4.15: Sensitivity of the polyp detection algorithm when applied to Protocol-5 CT data.

protocols. The experimental data indicates that the field of view does not have a significant impact on the performance of the automated polyp detection algorithm.

Another parameter of interest is the table speed. To evaluate the influence of this parameter on the overall polyp detection results, we set this parameter to 20mm/rotation for Protocol 6 and 30 mm/rotation for Protocols 1 to 5. At 30mm/rotation and 20mm/rotation table speeds the energy imparted is 7.0mSv at 100mAs. This parameter has a negligible effect on the radiation dose since the Siemens scanner used in these experiments utilises the "effective tube current" model where a variation in the scan time (the lower the scan time the higher the table speed) implies a concomitant variation in the tube current. For Siemens Somatom 16 slice

80

Sensitivity of Polyp Detection for Protocol-6 CT data
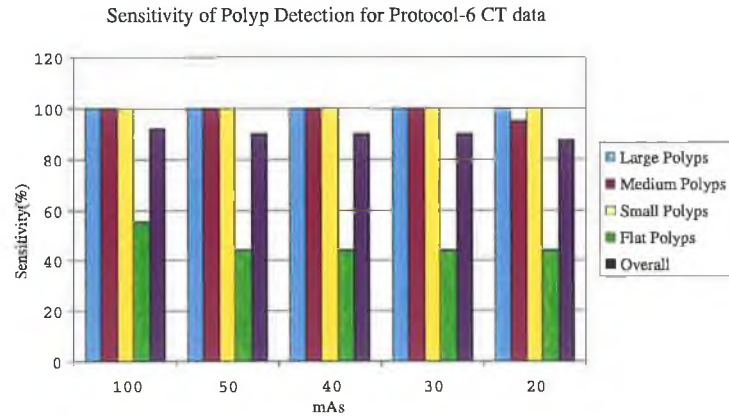
Figure 4.16: Sensitivity of the polyp detection algorithm when applied to Protocol-6 CT data.

CT scanner the lowest mAs that can be set at 20mm/rotation table speed is 20mAs whereas for 30mm/rotation table speed the lowest mAs is 13. This parameter was varied to evaluate only the effect of the motion artefacts and the experimental results indicate that the table speed has a marginal effect on the overall performance of the developed *CAD-CTC* system. Small benefits have been observed when the algorithm has been applied to the detection of small (not clinically significant) and flat polyps.

4.3.1.1   Level of noise and the radiation dose

In this element of the study another aim was to evaluate the correlation between the image noise and the radiation dose. In this regard five circular regions of interest (ROIs) were selected with a radius of 20 voxels that are evaluated for 3 consecutive slices (see Figure 4.17). Since the data is homogenous (the phantom is filled with water) the level of noise can be accurately sampled by calculating the standard deviation (SD) of the voxel distribution within the circular region of interest.

For CT data scanned using the Protocols 1 and 3 the SD increased with a factor of 2.67 (SD = 26.59 for 100mAs and SD = 70.95 for 13mAs) when the scan was performed at 13mAs when compared to the case when the phantom was scanned with 100mAs radiation dose. The relation between the noise level and the radiation dose is illustrated in Figure 4.18.

Figure 4.17: Five regions of interests located on the phantom to evaluate the noise level.
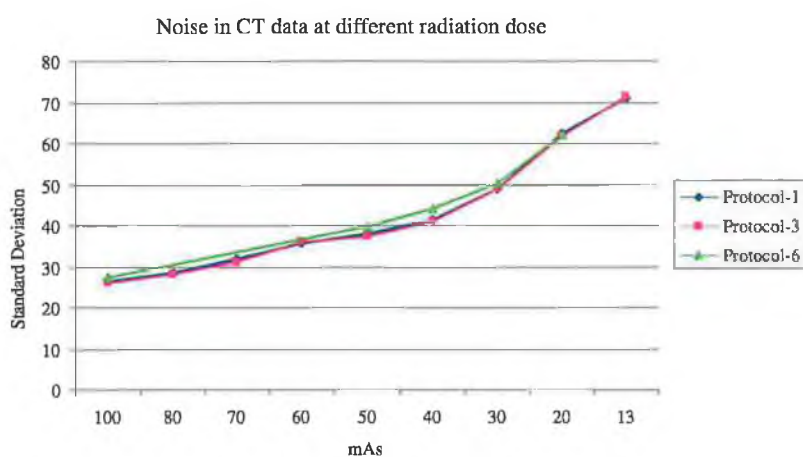
Figure 4.18: The relationship between noise level and the radiation dose.

4.4 Discussion and Conclusion

The experimental data presented in this chapter is obtained by scanning the synthetic phantom described in Section 4.1. Although the phantom was designed to emulate as closely as possible the real clinical conditions it is worth noting that the synthetic data is not affected by factors such as motion artefacts (caused by breathing) or the presence of residual material such as fluid and stool that are currently experienced when analysing real patient data. One of the main aims of this investigation was the development of a study environment that allows us to determine the influence of the scanning parameters on the performance of the polyp detection algorithm. Currently, the performance of the existing CAD-CTC systems is evaluated on real patient data that is supplied by different research organizations that are not

available for computer vision community. Therefore the absence of standard test data makes the performance evaluation of these systems restricted to the scenario they were tested. Thus, another important merit of this investigation is the generation of ground truth synthetic data that can be used to test all developed systems in the same conditions. For comparison purposes the phantom data are made available on request from the following web page: http://www.eeng.dcu.ie/~whelanp/cadctc. Typical size of a CT dataset is in the range (70-125MB). It is useful to note that recently the Walter Reed Army Medical Center (WRAMC) database has been made available to the research community which will help the evaluation of the developed *CAD-CTC* systems but the main advantage of using synthetic data is the generation of unambiguous ground truth data (requires no validation by radiologists) that can be used especially in the development phase of the *CAD-CTC* systems.

The developed *CAD-CTC* system indicates that automated polyp detection is feasible even at radiation doses as low as 13mAs. The sensitivity rate in polyp detection achieved by the developed *CAD-CTC* system is always higher than 90% for polyps larger than 5mm and the overall sensitivity for all types of polyps is higher than 80%. The sensitivity rate would be even higher as the developed method has not been trained for the detection of flat polyps. For the flat polyps the achieved sensitivity is in the range 22% to 55%. In these experiments one polyp (see Figure 4.10a) has been placed closed to the outer plastic plates of the phantom and at low radiation doses the image noise joined the surface of the polyp with the surface of the plastic plate and the classifier assigned this surface to be part of a fold. It is worth mentioning that this situation will not appear in clinical studies.

The main merit of this research work is the development of a realistic phantom that closely simulates the situations encountered in real clinical studies. Thus, the main emphasis was placed on evaluating the influence of the scanning parameters on the performance of the automated polyp detection. From these parameters attention was focused on the radiation dose, as the main concern regarding CT examinations is the exposure of the patients to ionizing radiation. Recent studies demonstrated that CT which accounts for 4% of the medical radiographic examinations contributes 35-40% of the cumulated radiation dose received by the patients [45]. The current study reveals that the reduction of mAs from 100 to 13 reduced the energy imparted from 7.0mSv to 0.92 mSv as it is illustrated in Figure 4.19.

Also another important issue addressed in this chapter is the relationship between

Imparted energy at different radiation dose
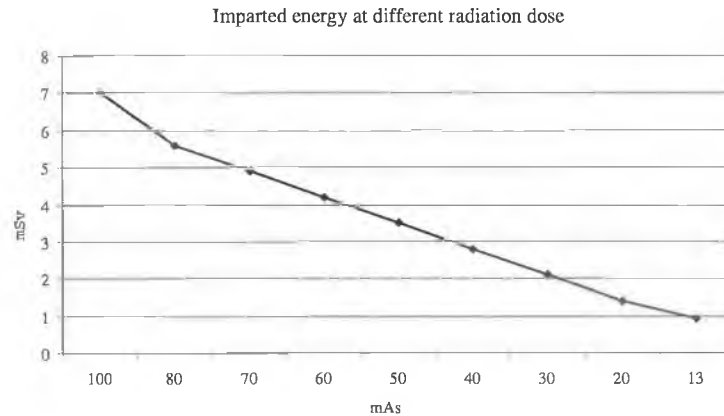
Figure 4.19: Radiation dose received by the patient at different mAs.

the radiation dose and the impact on the performance of the *CAD-CTC* polyp detection algorithm. In this regard, experimental data indicated that the level of image noise when the phantom was scanned with 13mAs was higher with a factor of 2.67 than in the case when the phantom was scanned with 100mAs radiation dose. Although the level of noise significantly increased at low radiation dose the effect on the performance in polyp detection is minimal. The experimental data presented in Figures 4.11 to 4.16 indicates that the sensitivity in polyp detection for polyps larger than 5mm is always above 95%. Results also show a small increase in false positives at 13mAs but the effect on true positive detection rate is not noticeable. The impact of the field of view and the reconstruction interval was negligible and it was virtually eliminated by the smoothing effect of the data interpolation that is applied to make the dataset isometric.

The main conclusion of this chapter in relation to the radiation dose is in line with the reported results provided in a number of publications [129, 88, 121, 125, 133] where is investigated the optimal scanning parameters. Based on our experiments the following low-dose protocol: collimation $1.5 \times 16mm$, slice thickness 3mm, reconstruction interval 1.5mm, table speed 30mm/rotation, radiation dose 13mAs can be potentially used for detection of colorectal polyps larger than 5mm in clinical studies.

In the next chapter, experimental results of automatic polyp detection for the developed *CAD-CTC* system is discussed for standard and low dose real patients datasets, patient datasets with synthetic polyps and phantom datasets.

Chapter 5

Experimental Evaluation

The developed *CAD-CTC* system comprises of an algorithm for automatic colon segmentation (see Chapter 2), three different feature extraction schemes for automatic polyp detection (see Chapter 3) and an analysis of the effects of different CT scanning parameters on the *CAD-CTC* using a synthetic phantom (see Chapter 4). Experimental results of the automatic colon segmentation (Chapter 2) shows that the method is suitable for colon surface generation for automatic polyp detection in *CAD-CTC*. Based on the geometrical shape of the polyp and fold surfaces, three different feature extraction schemes were discussed in Chapter 3. The first method, named *geometrical fitting*, used the least squares approximation in conjunction with surface normal concentration for automatic polyp detection in *CAD-CTC*. The second method calculates the statistical features using the least squares approximation. The third method employed the 3dB point on surface variation curve in conjunction with the calculation of the standard deviation of surface variation and surface normal concentration. In this chapter, the performance of the above mentioned polyp detection techniques will be analysed in detail. For clarity purposes the results were presented in tabular and graphical forms instead of using the ROC curves.

5.1   Results of Geometrical fitting

48 patients' (80 supine and prone) data with 120 polyps and five patients' data with 33 synthetic generated polyps [134] of various sizes were tested using the geometrical-driven method detailed in Section 3.1. For manual assisted colon segmentation, the overall sensitivity of the developed technique when applied to real patient *CTC* data was 74.53% and the rate of false positives per dataset was 3.90 (see Table 5.1). The sensitivity of the *CAD-CTC* system for polyps $\geq 10mm$ was 100%, for polyps

$[5 - 10mm)$ was 91.67% and for polyps $< 5mm$ was 67.95%. For mass and flat polyps the sensitivities were 90.91% and 33.33% respectively. In total 317 folds were detected in 80 patients data (Table 5.2). When the CAD-CTC system applied the automatic colon segmentation algorithm, the overall sensitivity was 70.75% with 3.50 false positives per dataset. The sensitivity for polyps $\geq 10mm$ was 100%, for polyps $[5 - 10mm)$ was 91.67% and for polyps $< 5mm$ was 62.82%. For manual assisted segmentation, the overall sensitivity for synthetic polyps was 90.91% and the rate of false positives was 3.6 per dataset (Table 5.3). The overall sensitivity for polyps greater than 5mm was 100.00% and the sensitivity for polyps less than 5mm was 66.66%. when the automated segmentation was applied the overall sensitivity was 87.88% with 2.2 false positive per dataset. The sensitivity for polyps $\geq 10mm$, $[5 - 10mm)$ and $< 5mm$ were 100%, 94.12% and 66.67% respectively.

Table 5.1: Performance analysis for real patient CT data

Type	Number	FNNN			PNN	
		TP	*Manual Sens.*	*Auto. Sens.*	*TP*	*Sens.*
$< 5mm$	78	53	67.95%	62.82%	49	62.82%
$[5 - 10)mm$	24	22	91.67%	91.67%	16	66.67%
$\geq 10mm$	4	4	100%	100%	4	100%
Flat	3	1	33.33%	33.33%	1	33.33%
Mass	11	10	90.91%	72.73%	8	72.73%
Total	120	90	74.53%	70.75%	78	65%
FP			3.90	3.50		5.83

Table 5.2: Statistics for false positives

Type	Number	
Fold	145	45.74%
Convex Surface	112	35.33%
Residual Material	37	11.67%
Tube	23	7.25%
Total	317	-

The proposed method was also applied to the detection of polyps in phantom data acquired using the Protocol-1 (collimation: 1.5x16mm, slice thickness: 3mm, reconstruction interval: 1.5mm, table speed: 30mm/rotation and mAs: 100, 40, 30, 20, 13) data. It has been already mentioned in Chapter 4 that one polyp in the phantom data was placed close to the plastic plate and at low doses the image noise joined the surface of the polyp with the surface of the plastic plate. Since these

Table 5.3: Performance analysis for synthetic polyp data

Type	Number	FNNN			PNN	
		TP	*Manual Sens.*	*Auto. Sens.*	*TP*	*Sens.*
$< 5mm$	6	4	66.67%	66.67%	2	33.33%
$[5 - 10)mm$	17	17	100%	94.12%	15	88.23%
$\geq 10mm$	9	9	100%	100%	7	77.78%
Flat	1	0	0%	0%	0	33.33%
Total	33	30	90.91%	87.88%	24	72.73%
FP			3.6	2.2		6.4

situation will not be encountered in real clinical studies, this polyp was ignored in the evaluation of the polyp detection techniques and the total number of polyps is 47 instead of 48. The overall sensitivities for 100mAs, 40mAs, 30mAs, 20mAs, and 13mAs phantom data were 87.23%, 87.23%, 82.98%, 87.23% and 82.98% respectively. The sensitivities for polyps $\geq 10mm$ were 100%, 100%, 92.86%, 100% and 92.86% for the phantom data acquired using the following radiation doses: 100mAs, 40mAs, 30mAs, 20mAs and 13mAs respectively. The sensitivity for polyps in the range $[5 - 10mm)$ was 100% for 100mAs, 40mAs, 20mAs and 13mAs except the 30mAs phantom data, where it was 94.74%. The sensitivity for polyps $< 5mm$ was 80% for 100mAs, 40mAs, 30mAs, 20mAs and 13mAs.

Table 5.4: Performance analysis for (100mAs) longitudinal phantom polyp data

Type	Number	FNNN		PNN	
		TP	*Sensitivity*	*TP*	*Sensitivity*
$< 5mm$	5	4	80%	3	60%
$[5 - 10)mm$	19	19	100%	18	94.74%
$\geq 10mm$	14	14	100%	13	92.86%
Flat	9	4	44.44%	2	22.22%
Total	47	41	87.23%	36	76.60%
FP			2		2

The *PNN* classifier was also employed to classify the candidate surfaces into polyps or folds. The overall sensitivity for *PNN* classifier was 65% with 5.83 false positive per dataset. The sensitivity for polyps $\geq 10mm$ was 100%, for polyps $[5 - 10mm)$ was 66.67% and for polyps $< 5mm$ was 62.82%. The sensitivities for masses and flat polyps were 72.72% and 33.33% respectively. The overall sensitivity for synthetic polyps was 72.73% and the false positive level was 6.4 per dataset (Table 5.3). The sensitivity for polyps $\geq 10mm$ and $[5 - 10mm)$ were 77.78%

Table 5.5: Performance analysis for low-dose (40 mAs) longitudinal phantom polyp data

mAs	Type	Number	FNNN		PNN	
			TP	Sensitivity	TP	Sensitivity
40	$< 5mm$	5	4	80%	3	60%
40	$[5 - 10)mm$	19	19	100%	18	94.74%
40	$\geq 10mm$	14	14	100%	14	100%
40	Flat	9	4	44.44%	1	11.11%
Total		47	41	87.23%	36	76.60%
FP				3		4

and 88.23% respectively. The sensitivity for polyps smaller than 5mm was 33.33%. Results of the *PNN* classifier are shown in Tables 5.1 5.2, 5.3, 5.4, 5.5, 5.6,5.7,5.8 and these results demonstrate that the *FNNN* classifier outperforms the *PNN* classifier.

Table 5.6: Performance analysis for low-dose (30mAs) longitudinal phantom polyp data

mAs	Type	Number	FNNN		PNN	
			TP	Sensitivity	TP	Sensitivity
30	$< 5mm$	5	4	80%	3	60%
30	$[5 - 10)mm$	19	18	94.74%	15	78.94%
30	$\geq 10mm$	14	13	92.86%	12	85.71%
30	Flat	9	4	44.44%	1	11.11%
Total		47	39	82.98%	31	65.96%
FP				4		4

To determine whether a polyp was correctly detected by the proposed algorithm, we compared the polyp location with the *CTC* reports performed by the radiologists. Also we compared the result with the colonoscopy reports. In our tests we used both supine and prone views for polyp detection.

5.1.1　Discussion on the performance of the Geometrical Fitting Approach

The proposed CAD system for colorectal polyp detection provides high sensitivity for medium and large polyps, while maintaining a low false positive incidence per dataset. Also in these experiments two different classifiers were evaluated in order to determine the optimal classification scheme that minimizes the false positive incidence while keeping the sensitivity higher than 90% for polyps larger than 5mm.

Table 5.7: Performance analysis for low-dose (20mAs) longitudinal phantom polyp data

mAs	Type	Number	FNNN		PNN	
			TP	Sensitivity	TP	Sensitivity
20	$< 5mm$	5	4	80%	3	60%
20	$[5-10)mm$	19	19	100%	16	84.21%
20	$\geq 10mm$	14	14	100%	12	85.71%
20	Flat	9	4	44.44%	2	22.22%
Total		47	41	87.23	33	70.21%
FP				4		2

Table 5.8: Performance analysis for low-dose (13mAs) longitudinal phantom polyp data

mAs	Type	Number	FNNN		PNN	
			TP	Sensitivity	TP	Sensitivity
13	$< 5mm$	5	4	80%	3	60%
13	$[5-10)mm$	19	19	100%	17	89.47%
13	$\geq 10mm$	14	13	92.85%	12	85.71%
13	Flat	9	3	33.33%	2	22.22%
Total		47	39	82.98%	34	72.34%
FP				3		4

The developed *CAD-CTC* system shows a relative low sensitivity for small polyps (67.95%). Since the proposed *CAD-CTC* system used data with 3mm collimation and 1.5mm reconstruction interval, the number of surface voxels that belong to polyps smaller than 5mm is small, and this was the reason why the sensitivity for small polyps was drastically reduced. Another reason for missing small polyps was the condition where the polyp was adjacent to folds. Therefore, the features derived from small polyps when positioned adjacent to folds show similar characteristics as generic folds, and the classifier detected them as folds. When the CAD system was applied to real datasets, 16% (4 out of 25) of the undetected small polyps were placed adjacently to folds and the classifier failed to identify them correctly.

The developed *CAD-CTC* method presents better results for the detection of small and medium size polyps when applied to lower resolution data (reconstruction interval (RI) 1.5mm) compared to the high resolution *CT* data used to evaluate the methods developed by Kiss et al. [96, 38] (0.8mm RI), Summers et al. [91] (1.0mm RI), Acar et al. [39] (1.0-1.50mm RI), and Kiraly et al. [97] (1.0mm RI). Also it is

worth mentioning that the developed *CAD-CTC* algorithm exhibits a remarkable robustness to noise. To demonstrate this, the developed algorithm was applied to low-dose phantom datasets (clinical investigations in Ireland typically use 100mAs as a standard dose) and numerical results are depicted in Tables 5.4 to 5.8.

One particular advantage of the developed *CAD-CTC* system method is its low computational overhead and more importantly it shows high sensitivity for medium $[5 - 10mm)$ and large $(\geq 10mm)$ polyps while the false positive rate is maintained at low levels.

5.2   Results of Statistical Feature Based Approach

In order to evaluate the performance of the statistical feature-based *CAD-CTC* method, 50 patients' (80 supine and prone) datasets with 127 polyps, 11 low-dose patients datasets (13-40mAs) with 2 polyps, five patients' datasets with 33 synthetic polyps [134], 25 WRMC patient (47 supine and prone) datasets with 54 polyps [135] of various sizes were used for experimentation. The overall sensitivity of the polyp detection using the *FNN* classifier for real patient data was 71.65% and the false positive level per dataset was 4.01 (Table 5.9). The sensitivity of the *CAD-CTC* technique for polyps $\geq 10mm$ was 100%, for polyps $[5 - 10mm)$ was 92% and for polyps $< 5mm$ was 64.25%. The sensitivities for mass and flat polyps were 72.72% and 66.67% respectively. In total 323 folds were detected in 80 supine and prone real patient data (Table 5.10). 53% of the false positives were generated by folds, 14% were caused by the residual material, 27% were spurious convex surfaces and 3.41% were surfaces generated by the rectal tube. The overall sensitivity in polyp detection using the *PNN* classifier for the real patient data was 61.417% (Table 5.9) and the false positive per dataset was 2.41. The sensitivity for polyps $\geq 10mm$ was 100%, for polyps $[5 - 10mm)$ was 84% and for polyps $< 5mm$ was 58.33%. The sensitivities for flat polyps and masses were 0% and 36.36% respectively.

The overall sensitivity of the system using the *FNN* classifier when applied to the detection of synthetic polyps was 84.85% and the rate of false positives was 2.8 per dataset (Table 5.11). The sensitivity for polyps larger than 5mm was 100.00% and the sensitivity for polyps smaller than 5mm was 33.33%. The overall sensitivity of the system using the *PNN* classifier when applied to the detection of synthetic polyps was 84.85% (Table 5.11) with a rate of false positives of 2.8 per dataset.

Table 5.9: Performance analysis for real polyp data

Type	Number	FNNN		PNN	
		TP	Sensitivity.	TP	Sensitivity
$< 5mm$	84	54	64.28%	49	58.33%
$[5 - 10)mm$	25	23	92%	21	84%
$\geq 10mm$	4	4	100%	4	100%
Flat	3	2	66.66%	1	0.0%
Mass	11	8	72.73%	8	36.36%
Total	127	91	71.65%	78	61.417%
FP			4.01		2.84

Table 5.10: Statistics for false positives - *FNNN* classifier

Type	Number	
Fold	174	53.87%
Convex Surface	90	27.86%
Residual Material	48	14.87%
Tube	11	3.41%
Total	323	-

The sensitivity for polyp $\geq 5mm$ was 84.85%. The overall sensitivity in polyp detection using *FNNN* classifier for WRMC data was 87.04% (Table 5.12) with a false positive rate of 2.17. The sensitivities for polyps $\geq 10mm$, $[5 - 10)mm$ and $< 5mm$ were 100%, 90.32% and 60% respectively. The sensitivity for flat polyps was 33.33%. The WRMC patient datasets are contrast enhanced and require electronic cleansing. Since, the proposed *CAD-CTC* method did not implemented the option for electronic cleansing, the features associated with the contrast enhanced material were ignored.

Table 5.11: Performance analysis for synthetic polyp data

Type	Number	FNNN		PNN	
		TP	Sensitivity	TP	Sensitivity
$< 5mm$	6	2	33.33%	2	33.33%
$[5 - 10)mm$	17	17	100%	17	100%
$\geq 10mm$	9	9	100%	9	100%
Flat	1	0	0%	0	0%
Total	33	28	84.85%	28	84.85%
FP			2.8		2.8

The developed CAD-CTC system was also evaluated on five 13mAs patient datasets, three 30mAs patient datasets, two 20mAs patient datasets and one 40mAs

Table 5.12: Performance analysis for WRMC patients data

Type	Number	FNNN		PNN	
		TP	Sens.	TP	Sens.
$< 5mm$	5	3	60%	2	40%
$[5-10)mm$	31	28	90.32%	24	77.42%
$\geq 10mm$	15	15	100%	14	93.33%
Flat	3	1	33.33%	0	0%
Total	54	47	87.04%	41	74.07%
FP			2.17		1.44

patient dataset. The overall sensitivity for both the *FNNN* and *PNN* classifiers were 50% 5.15. The sensitivity for polyps smaller than 5mm was 50%. The false positives per dataset were 2.18 and 1.27 when the system employed the *FNNN* and *PNN* classifiers.

Table 5.13: Performance analysis for low-dose patients data

Type	Number	FNNN		PNN	
		TP	Sens.	TP	Sens.
$< 5mm$	2	1	50%	1	50%
Total	2	1	50%	1	50%
FP			2.18		1.27

The average size of a typical interpolated CT dataset was 300MB for each view. The average time required for processing each volume of data was approximately 3.60min (see Table 5.14) on a Pentium-IV 1.6 GHz processor machine with 1GB memory.

Table 5.14: Average computation time (in seconds) for polyp detection

Candidate surface generation time	193
Feature extraction time	21.5
Classification	0.7
overall time (seconds)	215.2

5.2.1 Discussion of Statistical Features

The proposed statistical feature-based CAD system for colonic polyp detection provides high sensitivity yet maintaining a low false positive incidence per dataset. The developed polyp detection scheme was not able to correctly classify the polyps that are adjacent to fold or on fold, elongated and flat. Figure 5.1 illustrates a small

polyp situated on a fold that is missed by the *CAD-CTC* system. Since the polyp was adjacent to a fold surface, the candidate surface includes large parts of the fold and the features derived from the candidate surface shows similar value as the surfaces generated by folds. As a result the classifier assigned the candidate surfaces as being generated by a fold. Figure 5.1.b shows an elongated 7mm polyp which was also missed by the developed *CAD-CTC* technique. Since the shape of the polyp was elongated, the features calculated from the candidate surface have strong similarities with those calculated from surfaces generated by folds. Figure 5.1.c shows a flat polyp missed by the developed *CAD-CTC* system. The flat polyps have distinct geometrical shapes when compared to the sessile and pedunculated polyps and the developed polyp *CAD-CTC* system performed poorly for the detection of flat polyps.

(a)   (b)   (c)
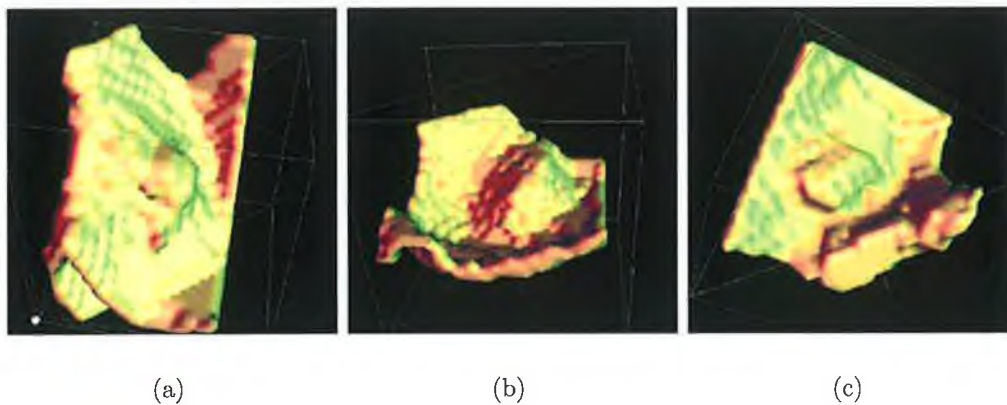
Figure 5.1: Missed polyps and their 3D surfaces.

The statistical feature-based system detected 323 false positive in 80 patient datasets. 53.87% of the false positives were generated by folds and Figure 5.2 shows a number of surfaces that were incorrectly classified by the developed *CAD-CTC* system. 14.87% of the false positives were generated by the residual material and three surfaces (FPs) generated by residual material are illustrated in Figure 5.3. It is useful to note that these false positives can be eliminated by applying texture analysis [40] as they have different CT densities than the colon tissue.

The experimental data indicates that the statistical feature-based polyp detection technique shows robustness in detection of polyps when applied to low dose data (13mAs) with a very low rate of false positives (2.18) per dataset. Another advantage of the *CAD-CTC* method detailed in this section is its low computational
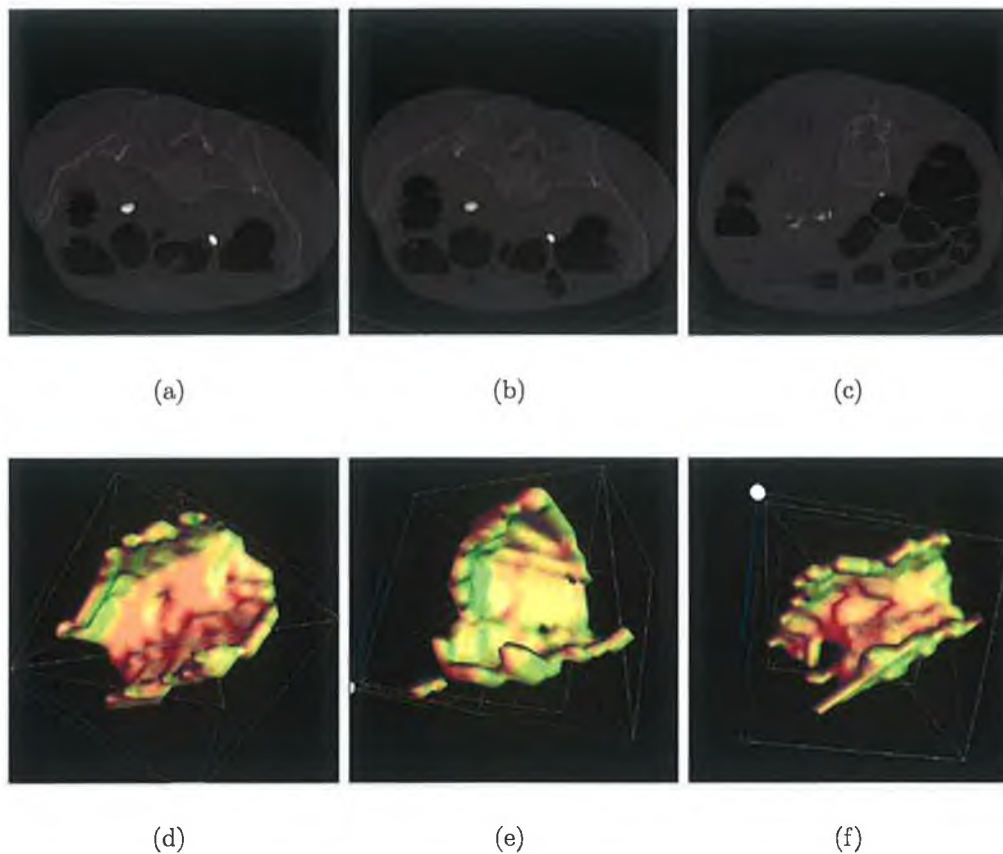
Figure 5.2: Detected false positive and their 3D surfaces.

overhead and more importantly it shows high sensitivity for medium $[5 - 10mm)$ and large $(\geq 10mm)$ polyps that are clinically significant.

5.3   Results of 3dB Feature-Based Approach

In order to evaluate the 3dB feature-based *CAD-CTC* technique, 50 patient (80 supine and prone) datasets with 127 polyps, 11 low-dose (13-40mAs) patient datasets with 2 polyps, five patients datasets with 33 synthetic polyps [134], a phantom with 47 polyps and 44 WRMC patient datasets (82 supine and prone) with 78 polyps of various sizes were used for experimentation. The overall sensitivity of the developed technique for the *FNNN* classifier when applied to standard dose (100mAs) real patient data was 71.65% and the false positive level per dataset was 5.15 (Table 5.15). Sensitivity for polyps $\geq 10mm$ was 100%, for polyps $[5 - 10mm)$ was 92% and for polyps $< 5mm$ was 65.47%. For mass and flat polyps the sensitivities were 63.64% and 66.67% respectively. In total 408 folds were detected in 80 patients data (Table 5.16). The overall sensitivity of the *CAD-CTC* method using the *PNN* and
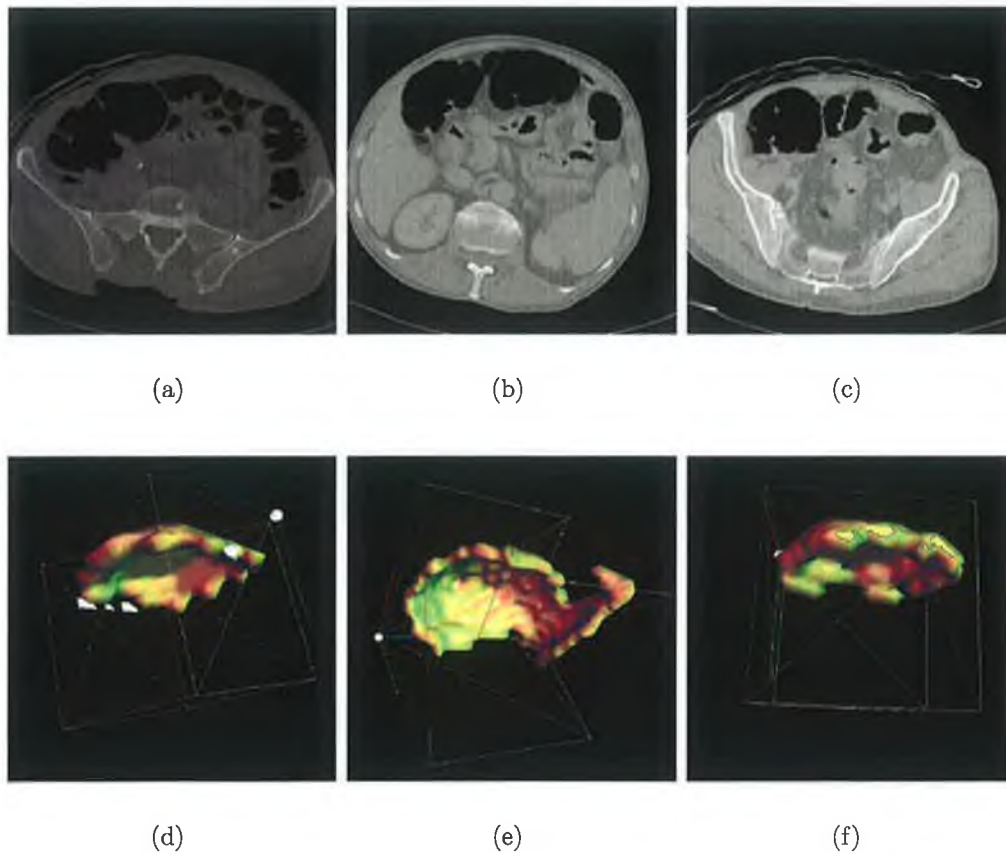
Figure 5.3: Detected residual material and their 3D surfaces.

SVM classifiers when applied to the standard dose real patient data were 67.72% and 70.08% respectively. The sensitivity for polyps $\geq 10mm$ was 100% for both *PNN* and *SVM* classifiers. The *PNN* and *SVM* classifiers show 84% and 88% sensitivities for polyps with sizes between $[5 - 10)mm$. The sensitivities for polyps $< 5mm$, flat polyps and masses were 59.52%, 66.66% and 81.82% when the algorithm was applied with *PNN* classifier. The sensitivity for polyps $< 5mm$, flat polyps and masses were 64.29%, 66.67% and 63.64% when the algorithm was applied with *SVM* classifier. The false positive rate for *PNN* and *SVM* classifiers were 4.19 and 3.58 respectively.

The overall sensitivities of the system when applied to synthetic datasets were 84.85%, 87.88% and 87.88% for *FNNN*, *PNN* and *SVM* classifiers (Table 5.17). The rate of false positives was 2.8, 2.6 and 2.8 per dataset for *FNNN*, *PNN* and *SVM* respectively. The sensitivity for polyps greater than 5mm was 100.00% for all three classifiers. Sensitivities for polyps smaller than 5mm was 33.33%, 50% and 50% for *FNNN*, *PNN* and *SVM* respectively.

Table 5.15: Performance analysis of 3dB feature approach for real polyp data

Type	Number	FNNN		PNN		SVM	
		TP	Sens.	TP	Sens.	TP	Sens.
$< 5mm$	84	55	65.47%	50	59.52%	54	64.29%
$[5 - 10)mm$	25	23	92%	21	84%	22	88%
$\geq 10mm$	4	4	100%	4	100%	4	100%
Flat	3	2	66.67%	2	66.67%	2	66.67%
Mass	11	7	63.64%	9	81.82%	7	63.64%
Total	127	91	71.65%	86	67.72%	89	70.08%
FP			5.15		4.19		3.58

Table 5.16: Statistics for false positives - *FNNN* classifier

Type	Number	
Fold	222	54.41%
Convex Surface	118	28.92%
Residual Material	55	13.48%
Tube	13	3.18%
Total	408	-

The overall sensitivity in polyp detection for WRMC data was 91.02%, 78.21% and 83.83% for *FNNN*, *PNN* and *SVM* respectively (Table 5.18). The false positive rate per dataset is 2.68, 2.36 and 2.16 for *FNNN*, *PNN* and *SVM* respectively. The sensitivities for polyps $\geq 10mm$, $[5 - 10)mm$ and $< 5mm$ were 100%, 93.75% and 70% respectively ($(($FNNN$)$ classifier). The sensitivities for the *PNN* classifier for polyps $\geq 10mm$, $[5 - 10)mm$ and $< 5mm$ were 100%, 79.17% and 70% respectively. The sensitivities of the *SVM* classifier for polyps $\geq 10mm$, $[5 - 10)mm$ and $< 5mm$ were 100%, 85.41% and 70% respectively. The sensitivities for flat polyps were 75%, 0% and 25% for *FNNN*, *PNN* and *SVM* respectively.

The algorithm was also applied for automatic polyp detection to 100mAs, 70mAs,

Table 5.17: Performance analysis of 3dB feature approach for synthetic polyp data

Type	Number	FNNN		PNN		SVM	
		TP	Sens.	TP	Sens.	TP	Sens.
$< 5mm$	6	2	33.33%	3	50%	3	50%
$[5 - 10)mm$	17	17	100%	17	100%	17	100%
$\geq 10mm$	9	9	100%	9	100%	9	100%
Flat	1	0	0%	0	0%	0	0%
Total	33	28	84.85%	29	87.88%	29	87.88%
FP			2.8		2.6		2.8

Table 5.18: Performance analysis of 3dB feature approach when applied to the WRMC polyp data

Type	Number	FNNN		PNN		SVM	
		TP	Sens.	TP	Sens.	TP	Sens.
$< 5mm$	10	7	70%	7	70%	7	70%
$[5-10)mm$	48	45	93.75%	38	79.17%	41	85.41%
$\geq 10mm$	16	16	100%	16	100%	16	100%
Flat	4	3	75%	0	0%	1	25%
Total	78	71	91.02%	61	78.21%	65	83.83%
FP			2.68		2.36		2.16

30mAs, 20mAs and 13mAs phantom data. The overall sensitivities with the *FNNN* classifier for 100mAs, 70mAs, 40mas, 30mAs, 20mAs and 13mAs were 89.36%, 89.36%, 87.23%, 85.11%, 82.98% and 85.11% respectively (see Tables 5.19 to 5.24). The sensitivities of the *PNN* classifier for 100mAs, 70mAs, 40mAs, 30mAs, 20mAs and 13mAs were 87.23%, 80.85%, 80.85%, 85.11%, 80.85% and 78.72% respectively. The sensitivities of the *SVM* classifier for 100mAs, 70mAs, 40mAs, 30mAs, 20mAs and 13mAs were 89.36%, 87.23%, 85.11%, 87.23, 82.98% and 82.98% respectively. The sensitivities for polyps $\geq 10mm$ were 100% for *FNNN*, *PNN* and *SVM* classifiers. The *FNNN* classifier shows 100% sensitivity for polyps between $[5-10)mm$ at 100mAs, 70mAs, 40mAs, 20mAs, and 13mAs radiation doses and 94.74% sensitivity at 30mAs radiation dose. The sensitivities of the *PNN* classifier for polyps between $[5-10)mm$ at 100mAs, 70mAs, 40mAs, 30mAs, 20mAs and 13mAs are 100%, 89.47%, 89.47%, 94.73%, 94.73% and 89.47% respectively. The sensitivities of the *SVM* classifier for $[5-10)mm$ polyps at 100mAs, 70mAs, 40mAs, 30mAs, 20mAs and 13mAs are 100%, 94.73%, 100%, 100%, 100% and 94.73% respectively. The method also shows 50% sensitivity for small polyps when it was applied to low-dose patient data with a false positive level smaller than 3.0 per dataset (see Table 5.25).

5.3.1   Discussion of 3dB Features

The proposed 3dB feature-based *CAD-CTC* system for colonic polyp detection provides high sensitivity for both *FNNN* and *SVM* classifiers. All three classifiers (FNNN, PNN, SVM) show 100% sensitivity for polyps $\geq 10mm$. The experimental data indicates that the *FNNN* shows better performance in polyp detection when compared to the performance of the *PNN* classifier especially for the detection of

Table 5.19: Performance analysis of 3dB feature approach for 100mAs phantom data

Type	Number	FNNN		PNN		SVM	
		TP	Sens.	TP	Sens.	TP	Sens.
< 5mm	5	5	100%	5	100%	5	100%
[5 − 10)mm	19	19	100%	19	100%	19	100%
≥ 10mm	14	14	100%	14	100%	14	100%
Flat	9	4	44.44%	3	33.33%	3	33.33%
Total	47	42	89.36%	41	87.23%	41	89.36%
FP			4		2		2

Table 5.20: Performance analysis of 3dB feature approach for 70mAs phantom data

Type	Number	FNNN		PNN		SVM	
		TP	Sens.	TP	Sens.	TP	Sens.
< 5mm	5	5	100%	4	80%	5	100%
[5 − 10)mm	19	19	100%	17	89.47%	18	94.73%
≥ 10mm	14	14	100%	14	100%	14	100%
Flat	9	4	44.44%	3	33.33%	4	44.44%
Total	47	42	89.36%	38	80.85%	41	87.23%
FP			4		1		2

medium sized polyps ([5 − 10)mm). The *SVM* classifier shows best performance with respect to the false positives level per dataset when compared to *FNNN* and *PNN* classifiers. The developed polyp detection scheme when used with the *FNNN* classifier missed 5 polyps with sizes between 5-8mm in Mater and WRMC patient datasets. On the other hand the *SVM* classifier missed 10 polyps with size between 5-8mm. The *SVM* classifier demonstrates lower detection rate for polyps with sizes between 5-8mm when compared to the polyp detection rate offered by the *FNNN* classifier. All three classifiers failed to detect the polyps when they are adjacent to folds.

Based on the reported results we conclude that the developed *CAD-CTC* system

Table 5.21: Performance analysis of 3dB feature approach for 40mAs phantom data

Type	Number	FNNN		PNN		SVM	
		TP	Sens.	TP	Sens.	TP	Sens.
< 5mm	5	5	100%	5	100%	5	100%
[5 − 10)mm	19	19	100%	17	89.47%	19	100%
≥ 10mm	14	14	100%	14	100%	14	100%
Flat	9	3	33.33%	2	22.22%	2	22.22%
Total	47	41	87.23%	38	80.85%	40	85.11%
FP			4		2		2

Table 5.22: Performance analysis of 3dB feature approach for 30mAs phantom data

Type	Number	FNNN		PNN		SVM	
		TP	Sens.	TP	Sens.	TP	Sens.
$< 5mm$	5	5	100%	5	100%	5	100%
$[5 - 10)mm$	19	18	94.74%	18	94.73%	19	100%
$\geq 10mm$	14	14	100%	14	100%	14	100%
Flat	9	3	33.33%	3	33.33%	3	33.33%
Total	47	40	85.11%	38	85.11%	41	87.23%
FP			4		4		1

Table 5.23: Performance analysis of 3dB feature approach for 20mAs phantom data

Type	Number	FNNN		PNN		SVM	
		TP	Sens.	TP	Sens.	TP	Sens.
$< 5mm$	5	5	100%	5	100%	5	100%
$[5 - 10)mm$	19	19	100%	18	94.73%	19	100%
$\geq 10mm$	14	14	100%	14	92.86%	14	100%
Flat	9	1	11.11%	2	22.22%	1	11.11%
Total	47	39	82.98%	38	80.85%	39	82.98%
FP			4		3		2

show similar or better performance in polyp detection when compared to the performance of the *CAD-CTC* methods reported in [35, 37, 38, 40, 41, 90, 91, 96, 98, 99] for polyps $\geq 10mm$ (sensitivity of 100%). The proposed system outperforms the methods proposed in [35, 37, 38, 40, 41, 90, 91, 96, 98, 99] when applied to the detection of small and medium size polyps. The rate of false positives for *FNNN* classifier (5.15 per dataset) is lower than the rate achieved by the methods proposed by [35, 40, 41, 90, 91, 98, 99] but shows a higher level of false positive when compared to the *CAD-CTC* techniques developed by [37, 38, 96]. The false positives level reported by Yoshida et al. and Kiss et al. are 2 to 3.5 and 2.48 per dataset respectively. On the other hand, the false positives level achieved by the

Table 5.24: Performance analysis of 3dB feature approach for 13mAs phantom data

Type	Number	FNNN		PNN		SVM	
		TP	Sens.	TP	Sens.	TP	Sens.
$< 5mm$	5	5	100%	4	80%	5	100%
$[5 - 10)mm$	19	19	100%	17	89.47%	18	94.73%
$\geq 10mm$	14	13	100%	14	100%	14	100%
Flat	9	3	33.33%	2	22.22%	2	22.22%
Total	47	40	85.11%	37	78.72%	39	82.98%
FP			2		3		2

Table 5.25: Performance analysis of 3dB feature approach for low-dose patient data

Type	Number	FNNN		PNN		SVM	
		TP	*Sens.*	*TP*	*Sens.*	*TP*	*Sens.*
$< 5mm$	2	1	50%	1	50%	1	50%
Total	2	1	50%	1	50%	1	50%
FP			2.54		2.81		2.27

3dB feature-based technique when using the SVM classifier is 3.57 when applied to the Mater Hospital datasets and 2.16 when applied to the WRMC datasets. The experimental results also indicate that the developed *CAD-CTC* method shows robustness in automatic polyp detection when applied to low-dose patient data while maintaining a low level of false positives per dataset.

5.4 Conclusion

In this chapter results of three different polyp detection schemes are presented. All the three methods show 100% sensitivity for polyps $\geq 10mm$ and 92% sensitivity for polyps $5 - 10mm$ and 64% or higher sensitivity for polyps $< 5mm$. It is also useful to note that all three techniques are affected by similar false positive such as those generated by residual material, rectal tubes, and small folds. The first method which uses geometrical features is optimised for the non-isotropic CT datasets used by our clinical partners. The second method based on statistical features and the third method based on 3dB and surface variation features are robust techniques suitable for all kinds of datasets and are tested on publicly available WRMC datasets. The experimental results indicate that the developed *CAD-CTC* system shows similar results to both manual *CTC* and the traditional optical colonoscopy approach, thus making it suitable tool to be used in clinical studies [51, 52, 53, 54, 55, 56, 57].

Chapter 6

Conclusion and Future Works

The main aim of this research was the development of a fully automatic *CAD-CTC* system that can be applied for the robust identification of colorectal polyps in *CTC* datasets acquired using standard and low dose radiation. The developed system consists of a number of distinct stages where the focus of this thesis is in the automatic colon segmentation, candidate surface extraction, classification, and the analysis of the effects of low-dose on automatic polyp detection in *CAD-CTC*. This research work has generated a number of significant novel theoretical and experimental contributions in all of these processing stages. The theoretical contributions generated from this research work are as follows:

- Development of an algorithm for automatic segmentation of collapsed colon in *CTC*

- Development of three distinct polyp detection schemes in *CTC*

- Development of a gradient operator for robust polyp detection in low-dose *CAD-CTC*.

The experimental contributions of this research are as follows:

- Design and construction of a synthetic phantom

- Detection of optimum scanning parameters for *CAD-CTC* systems

- Development of standard testing datasets that can be used in the development of the *CAD-CTC*.

The first step of the developed *CAD-CTC* system addresses the problem of automatic colon segmentation. As discussed in Chapter 2, the manually assisted colon segmentation technique based on standard region growing is feasible only

when applied to CT data acquired in ideal conditions. In many clinical studies, the colon data is often collapsed (approximately 50% of the total examinations) due to either insufficient colon insufflation or to blockages caused by water and residual material. Thus, one of the main objectives of this research work was the development of an automatic colon segmentation technique that is able to perform robust colon segmentation when applied to well-distended and collapsed CT data. In this sense, an automatic colon segmentation technique has been developed and it is explained in detail in Chapter 2. The developed technique is a multi-stage approach where the colon is reconstructed based on morphological measurements associated with the candidate colon segments (V/L analysis) and by enforcing the geometrical constraints imposed by a generic model of a well-distended colon. The automatic segmentation of colons in collapsed CT data is a relatively unexplored research topic and this is illustrated by the literature review presented in the introductory part of Chapter 2. When the automatic colon segmentation technique was applied to 88 standard dose and 8 low dose datasets the experimental data indicates that this fully automatic method is able to detect 99.68% of the total colon wall when applied to standard dose collapsed CT data and 96.52% when applied to low dose collapsed CT data. The performance of the colon segmentation technique detailed in this thesis is superior than the performance of other published automatic colon segmentation techniques [65, 66].

The second step of the *CAD-CTC* system deals with the problem of polyp detection. This is the main subject of this research work detailed discussion is provided in Chapter 3. The approach described in this thesis is based on a geometrical analysis of the colon wall and it has several well-defined stages including the extraction of polyp candidate surfaces, feature extraction and classification. The first step of the polyp detection algorithm consists of identification of the colon surfaces that have strong convex characteristics. In order to achieve this goal, a technique that evaluates the intersection of the normal vectors has been developed. To eliminate polyp candidate surfaces with low convex properties, the algorithm applies further processing including the calculation of the Gaussian distribution of the Hough points and a convexity test. In order to extract the optimal features that can be used for polyp detection, three feature extraction techniques have been developed and examined. The aim of all techniques investigated in this thesis was the extraction of the optimal features that can be used for robust classification based on the observation

that polyps have spherical/ellipsoidal shapes whereas folds have cylindrical shapes. In this regard, the first technique analysis the geometrical features that measure in the least square sense the similarity between the candidate surface and the ellipsoidal and spherical surfaces. The sensitivities in polyp detection achieved by this technique for polyps $> 10mm$ and polyps between $[5 - 10mm)$ are 100% and 91.67% respectively and the rate of false positives per dataset is 3.90. The second polyp detection technique attempts to sample the surface variation using statistical features that are calculated using the least square approximation between candidate surfaces and ellipsoidal/spherical surfaces. The sensitivities in polyp detection achieved by the statistical feature-based technique in the detection of polyps $> 10mm$ and polyps between $[5 - 10mm)$ are 100% and 92% respectively and the rate of false positives per dataset is 4.01. The third polyp detection technique evaluated in this thesis evaluates the surface variation sampled by the 3dB attenuation point on the surface change curve, surface normal concentration and maximum distance constraint. This techniques has been evaluated using three classification schemes, the *FNNN*, *PNN* and *SVM* and the overall performance in polyp detection for polyps $> 10mm$ and polyps between $[5 - 10mm)$ are 100% and 92% respectively (*FNNN* classifier). The lowest false positives level (3.58) has been obtained when the polyp detection technique has been used in conjunction with the *SVM* classifier. The overall performance of the polyp detection methods evaluated in this thesis compares well with the performances achieved by the most advanced *CAD-CTC* techniques evaluated in Chapter 5.

A distinct part of this research work was the development of a synthetic phantom that is used to investigate the influence of the scanning parameters on the overall performance of the developed *CAD-CTC* techniques. As indicated in Chapter 1, the major concern associated with *CTC* is the patient exposure to ionising radiation and this is a major deterring factor in using *CTC* as a mass screening technique. Many studies reviewed in Chapter 4 indicate that the risk of inducing cancer to patients that undergo CT examinations is significantly reduced if they are subjected to low radiation levels. To address this issue an important aim of this research was to analysis the effect of low radiation dose (and other scanning parameters such as reconstruction interval and table speed) on the overall performance of the developed *CAD-CTC* system. In order to perform a detailed analysis on this subject, a synthetic phantom has been developed where we were able to create synthetic polyps

with different shapes and sizes that approximate with high accuracy the real polyps encountered in clinical studies. The development of the synthetic phantom allowed us to generate a large number of datasets acquired using six different protocols that were used to evaluate the performance of the *CAD-CTC* systems. The experimental results indicate that accurate polyp detection can be performed even at radiation doses as small as 13mAs.

6.1   Future Work

The developed technique has been designed to detect the polyps in datasets that were obtained without oral enhanced patient preparation such as the WRMC data. Although the *CAD-CTC* technique described in this thesis can be applied with no restriction to oral-enhanced *CTC* datasets it is useful to mention that the overall level of false positives can be substantially reduced if the method were to incorporate an advanced electronic cleansing procedure [136, 137, 138] that will allow the identification of false positives caused by the residual material. In addition, a digital cleansing procedure will also be beneficial for the detection of polyps that are immersed in the water left in the colon at the time of the CT examination.

Another possible development would be the implementation of a new polyp candidate surface technique that will improve the detection for small polyps ($< 5mm$) situated adjacently to large folds. The experimental results indicate that 14% of the undetected small polyps were placed adjacently to folds and this condition generated candidate surfaces that include many sections of the folds. This has a negative impact on the overall performance as the system assigned these surfaces as being generated by folds. In order to address this problem we have developed an alternative surface extraction technique based on a 3D level-set implementation (see Appendix-b) and the initial results are encouraging.

The reduction of the level of false positives can be achieved by evaluating the shape of the candidate surface in the frequency domain. Recently, Miranda et al. [139] proposed a novel technique that evaluates the 3D histogram calculated from the candidate surface in the frequency domain and they demonstrated that the use of spectral information can be used for robust polyp identification.

The developed *CAD-CTC* system can be further developed in order to reduce the computational time required to process the CT datasets. Currently, the average

computational time required by the system to process one datasets is 3.6 minutes and this processing time can be considerably reduced since the developed algorithms for colon segmentation and polyp detection have not been optimised for speed.

The results of the automatic polyp detection show that the 11-16% of the false positives are generated due to the residual material in the colon. The reduction of these false positive can be done by implementing texture analysis [40] of the polyp and residual material.

The developed *CAD-CTC* system used FNNN, PNN ans SVMs classifiers (see Chapter 3) for classification of polyp candidate surface as polyp or fold. Experimental results (see Chapter 5.3) show that all the three classifiers illustrate similar sensitivity for polyp greater than 5mm. These results indicate that the features derived from the polyp candidate surface are robust for classification of the candidate surface as polyp or fold. But further implementation of advanced classification scheme like committee of SVMs [101] can increase the sensitivity of polyp detection while the false positive level can be reduced.

As indicated in Chapter 5, the developed *CAD-CTC* system was tested on 11 low-dose patient data having only 2 polyps smaller than 5mm. The amount of low-dose patient data available to test the system did not allow us to fully characterise the performance of the *CAD-CTC* techniques discussed in Chapter 3. Our future work will involve the validation of the *CAD-CTC* system on a larger number of standard and low-dose datasets including the data supplied by our clinical partners (Mater Hospital) and public available *CTC* databases like WRMC [135] and ACRIN [140] that are recently made available.

Bibliography

[1] Parker S, Tong T, Bolden S, and Wingo P. Cancer statistics 1997. *A Cancer Journal for Clinicians*, 47:5–27, 1997.

[2] Cancer Research UK. Bowel cancer factsheet april. 2003.

[3] NCRI 2000. Cancer in ireland,1997 : Incidence and mortality. *Healy and Associates*, 2000.

[4] Greenlee RT, Murray T, Bolden S, and Wingo PA. Cancer statistics, 2000. *A Cancer Journal for Clinicians*, 50:7–33, 2000.

[5] American Cancer Society: Cancer Facts and Figures. *American Cancer Society*, 1999.

[6] Ransohoff DF and Sandler RS. Screening for colorectal cancer. *The New England Journal Medicine*, 346(1), January 3 2002.

[7] National Cancer Institute. Working guidelines for early cancer detection: Rationale and supporting evidence to decrease mortality. *Bethesda: National Cancer Institute*, 1987.

[8] Robert A, Cokkinides SV, , and Eyre HJ. American cancer society guidelines for the early detection of cancer, 2003. *A Cancer Journal for Clinicians*, 53:27–43, 2003.

[9] Bond JH. Polyp guideline: diagnosis, treatment, and surveillance for patients with nonfamilial colorectal polyps. *The Practice Parameters Committee of the American College of Gastroenterology. Annals of Internal Medicine*, 119:83643, 1993.

[10] Schrock TR. Colonoscopy versus barium enema in the diagnosis of colorectal cancer and polyps. *GastroIntestinal Endoscopy Clinics of North America*, 3:585–610, 1993.

Bibliography

[11] Winawer JS, Stewart ET, Zauber AG, and Bond JH. A comparison of colonoscopy and double-contrast barium enema for surveillance after polypectomy. *The New England Journal of Medicine*, 342:1766–1772, 2000.

[12] Sieg A, Ulrike HE, and Thomas E. Prospective evaluation of complications in outpatient gi endoscopy: A survey among german gastroenterologists. *Gastrointestinal Endoscopy*, 53(6):620–627, May 2001.

[13] Hall C, Dorricott NJ, Donovan IA, and Neoptolemos JP. Colon perforation during colonoscopy: surgical versus conservative management. *The British journal of surgery*, 78(5):542–4, May 1991.

[14] Anderson JC, Gonzalez JD, Messina CR, and Pollack BJ. Factors that predict incomplete colonoscopy: thinner is not always better. *American Journal of Gastroenterology*, 95(10):2784–7, October 2000.

[15] Vining DJ, Gelfand DW, Bechtold RE, Scharling ES, Grishaw EK, and Shifrin RY. Technical feasibility of colon imaging with helical ct and virtual reality [abstract]. *American Journal of Roentgenology*, 162:104, 1994.

[16] Sato M, Lakare S, Wan M, Kaufman A, Liang Z, and Wax M. An automatic colon segmentation for 3d virtual colonoscopy. *IEICE Transaction Information and Systems*, E84-D(1):201–208, January 2001.

[17] Johnson CD, Hara AK, and Reed JE. Virtual endoscopy: what's in a name? *American Journal of Roentgenology*, 171:1201–2, 1998.

[18] Lichan H, Arie K, Yi Chih W, Ajay V, Mark W, and Zhengrong L. 3d virtual colonoscopy. *In Proceeding Biomedical Visualization, M. Loew and N. Gershon, Eds., Atlanta*, pages 26–33, 1995.

[19] Hara AK, Johnson CD, Reed JE, Ahlquist DA, Nelson H, Ehman RL, McCollough CH, and Ilstrup DM. Detection of colorectal polyps by ct colonography: Feasibility of a novel technique. *Gastroenterology*, 100:284–290, 1996.

[20] Johnson CD and Hara AK. Ct colonography: the next colon screening examination? *Radiology*, 216:331–341, 2000.

Bibliography

[21] Fenlon HM, Nunes DP, Schroy PC, Barish MA, Clarke PD, and Ferrucci JT. A comparision of virtual and convention colonoscopy for the detection of colorectal polyps. *The New England Journal Medicine*, 341:1496–1503, 1999.

[22] Fenlon HM, Mcaneny DB, Nunes DP, Clarke PD, and Ferrucci JT. Occlusive colon carcinoma: Virtual colonoscopy in the preoperative evaluation of the proximal colon. *Radiology*, 210:423–428, 1999.

[23] Fenlon HM and Ferrucci JT. Virtual colonoscopy: what will the issue be? *American Journal of Roentgenology*, 169(2):453–8, 1997.

[24] Johnson CD. Virtual colonoscopy for colorectal polyp detection. *RBM-Revue Europeenne de Technologie Biomedicale*, 19(5):143–147, 1997.

[25] Royster A, Fenlon H, Clarke P, Nunes D, and Ferrucci J. Ct colonoscopy of colorectal neoplasms: Two-dimensional and three-dimensional virtual-reality techniques with colonoscopic correlation. *American Journal of Roentgenology*, 169:1237–1242, 1997.

[26] Vos FM, van Gelder RE, Serlie IW, Florie J, Nio CY, Glas AS, Post FH, Truyen R, Gerritsen FA, and Stoker J. Three-dimensional display modes for ct colonography: Conventional 3d virtual colonoscopy versus unfolded cube projection. *Radiology*, 228(3):878–85, 2003.

[27] Macari M and Bini EJ. Ct colonography: Where have we been and where are we going? *Radiology*, 237:819–833, 2005.

[28] Halligan S, Altman DG, Taylor SA, Mallett S, Deeks JJ, Bartram CI, and Atkin W. Ct colonography in the detection of colorectal polyps and cancer: Systematic review, meta-analysis, and proposed minimum data set for study level reporting. *Radiology*, 237:893–904, 2005.

[29] Pickhardt PJ, Choi JR, Hwang I, Butler JA, Puckett ML, Hildebrandt HA, Wong RK, Nugent PA, Mysliwiec PA, and Schindler WR. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *The New England Jounal of Medicine*, 349(23):2191–2200, December 4 2003.

Bibliography

[30] Pickhardt PJ. Three-dimensional endoluminal ct colonography (virtual colonoscopy): Comparison of three commercially available systems. *American Journal of Roentgenology*, 181:1599 – 1606, 2003.

[31] Pickhardt PJ, Choi JR, Hwang I, and Schindler WR. Nonadenomatous polyps at ct colonography: Prevalence, size distribution, and detection rates. *Radiology*, 232:784–790, 2004.

[32] Yee J, Akerkar GA, Hung RK, Steinauer-Gebauer AM, Wall SD, and McQuaid KR. Colorectal neoplasia: Performance characteristics of ct colonography for detection in 300 patients. *Radiology*, 219(3):685–92, 2001.

[33] Fletcher JG, Johnson CD, Welch TJ, MacCarty RL, Ahlquist DA, Reed JE, Harmsen WS, and Wilson LA. Optimization of ct colonography technique: prospective trial in 180 patients. *Radiology*, 216(3):704711, 2000.

[34] Summers RM. Current concepts and future directions in computeraided diagnosis for ct colonography. *Proceedings of CARS, 16th International Congress and Exhibition, Paris, France*, pages 743–748, June 2629 2002.

[35] Paik DS, Beaulieu CF, Rubin GD, Acar B, Jeffrey RBJr, Yee J, Dey J, and Napel S. Surface normal overlap: A computer-aided detection algorithm with application to colonic polyps and lung nodules in helical ct. *IEEE Transactions on Medical Imaging*, 23(6):661–75, 2004.

[36] Summers RM, Johnson CD, Pusanik LM, Malley JD, Youssef AM, and Reed JE. Automated polyp detection at ct colonography: Feasibility assessment in a human population. *Radiology*, 219:51–59, 2001.

[37] Yoshida H, Masutani Y, MacEneaney P, Rubin DT, and Dachman AH. Computerized detection of colonic polyps at ct colonography on the basis of volumetric features: Pilot study. *Radiology*, 222:327–336, 2002.

[38] Kiss G, Cleynenbreugel J, Suetens P, and Marchal G. Computer aided diagnosis for ct colonography via slope density functions. *Medical Image Computing and Computer-Assisted Intervention*, pages 746–753, 2003.

[39] Acar B, Beaulieu CF, Gokturk SB, Tomasi C, Paik DS, Jeffrey RBJr, Yee J, and Napel S. Edge displacement field-based classification for improved detec-

Bibliography

tion of polyps in ct colonography. *IEEE Transactions on Medical Imaging*, 21(12):1461–7, 2002.

[40] Wang Z, Li L, Anderson J, Harrington D, and Liang Z. Colonic polyp characterization and detection based on both morphological and texture features. *International Congress Series*, 1268:1004–1009, 2004.

[41] Jerebko AK, Malley JD, Franaszek M, and Summers RM. Multi neural network classification scheme for detection of colonic polyps in ct colonography data sets. *Academic Radiology*, 10(2):154–160, 2003.

[42] Summers RM, Yao J, Pickhardt PJ, Franaszek M, Bitter I, Brickman D, Krishna V, and Choi JR. Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. *Gastroenterology*, 129(6), 2005.

[43] Roebuck DJ and Metreweli C. Radiation risk in ct for acute abdominal pain. *Radiology*, 209:287–288, 1998.

[44] Donnelly LF, Emery KH, Brody AS, Laor T, Gylys-Morin VM, Anton CG, Thomas SR, and Frush DP. Minimizing radiation dose for pediatric body applications of single-detector helical ct: Strategies at a large children's hospital. *American Journal of Roentgenology*, 176(2):303–6, February 2001.

[45] Nagel HD. Radiation exposure in computed tomography. *European Coordinate Committee of the Radiological and Electromedical Industries*, 2001.

[46] Brenner DJ and Georgsson MA. Mass screening with ct colonography: Should the radiation exposure be of concern? *Gastroenterology*, 129(1):328–37, July 2005.

[47] Fletcher JG, Johnson CD, Krueger WR, Ahlquist DA, Nelson H, Ilstrup D, Harmsen WS, and Corcoran KE. Contrast-enhanced ct colonography in recurrent colorectal carcinoma: Feasibility of simultaneous evaluation for metastatic disease, local recurrence, and metachronous neoplasia in colorectal carcinoma. *American Journal of Roentgenology*, 178(2):283–290, February 2002.

[48] Hoon J, Rolnick JA, Haker S, and Barish MA. Multislice ct colonography:

Bibliography

Current status and limitations. *European Journal of Radiology*, 47(2):123–134, 2003.

[49] Sosna J, Morrin M, Kruskal JB, Farrell RJ, Nasser I, and Raptopoulos V. Colorectal neoplasms: Role of intravenous contrast-enhanced ct colonography. *Radiology*, 228(1):152–156, July 2003.

[50] Filippone A, Ambrosini R, Fuschi M, Marinelli T, Genovesi D, and Bonomo L. Preoperative t and n staging of colorectal cancer: Accuracy of contrast-enhanced multi-detector row ct colonography-initial experience. *Radiology*, 231(1):83–90, April 2004.

[51] Chowdhury TA, Whelan PF, and Ghita O. A method for automatic segmentation of collapsed colons at ct colonography. *2nd Indian International Conference on Artificial Intelligence, Pune, India*, 20-22 December 2005.

[52] Chowdhury TA, Sadleir RJT, Whelan PF, Sezille N, Moss A, OHare A, Foley S, Fenlon H, and MacMathuna P. Automatic detection of colon at ct colonography. *Irish Society of Gastroenterology, Winter Meeting*, 2004.

[53] Chowdhury TA, Whelan PF, and Ghita O. The use of 3d surface fitting for robust polyp detection and classification in ct colonography. *Journal of Computerized Medical Imaging and Graphics (In Press)*, 2006.

[54] Chowdhury TA, Ghita O, Whelan PF, and Miranda A. A note on feature selection for polyp detection in ct colonography. *The 18th International Conference on Pattern Recognition, Hong Kong*, 2006.

[55] Chowdhury TA, Ghita O, and Whelan PF. A statistical approach for robust polyp detection in ct colonography. *27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Shanghai, China*, 2005.

[56] Sadleir RJT, Whelan PF, Sezille N, Chowdhury TA, Bruzzi J, Moss A, McMathuna P, and Fenlon H. Automatic detection of colorectal polyps at ct colonography using shape information. *Association of Physical Scientists in Medicine, 2004 Annual Scientific Meeting*, 11th June 2004.

Bibliography

[57] Sadleir RJT, Whelan PF, Sezille N, Chowdhury TA, Bruzzi J, Moss A, Mc-Mathuna P, and Fenlon H. Computer-aided detection of colorectal polyps at ct colonography. *Irish Society of Gastroenterology Winter Meeting*, December 2003.

[58] Chowdhury TA, Whelan PF, Ghita O, Sezille N, and Foley S. Development of a synthetic phantom for the selection of optimal scanning parameters in cad-ct colonography. *Medical Engineering and Physics (Accepted for publication)*, 2006.

[59] Chowdhury TA, Whelan PF, Fenlon H, and MacMathuna P. Evaluation of radiation dose on automatic polyp detection at ct colonography: Experiments with a synthetic phantom. *Association of Physical Scientists in Medicine, 2005 Annual Scientific Meeting, Galway*, 25-26February 2005.

[60] Chowdhury TA, Sadleir RJT, Whelan PF, Moss A, Varden J, Short M, Fenlon H, and MacMathuna P. The impact of radiation dose on imaged polyp characteristics at ct colonography: Experiments with a synthetic phantom. *Association of Physical Scientists in Medicine, 2004 Annual Scientific Meeting, Dublin*, 11th June 2004.

[61] Liang Z, Yang F, Wax M, Li J, You J, Kaufman A, Hong L, Li H, and Viswambharan A. Inclusion of a priori information in segmentation of colon lumen for 3d virtual colonoscopy. *Conf. IEEE NSS-MIC*, pages 1423–1427, 1997.

[62] Chen D, Wax MR, LI L, Liang Z, LI B, and Kaufman AE. A novel approach to extract colon lumen from ct images for virtual colonoscopy. *IEEE Transactions on Medical Imaging*, 19(12):1220–6, 2000.

[63] Wyatt CL, Ge Y, and Vining DJ. Automatic segmentation of the colon for virtual colonoscopy. *Computerized medical imaging and graphics*, 24:1–9, 2000.

[64] Masutani Y, Yoshida H, MacEneaney PM, and Dachman AH. Automated segmentation of colonic walls for computerized detection of polyps in ct colonography. *Journal of Computer Assisted Tomography*, 25(4):629–638, 2001.

[65] Napii J, Dachman AH, Maceneaney P, and Yoshida H. Automated knowledge-guided segmentation of colonic walls for computerized detection of polyps in

Bibliography

ct colonography. *Journal of Computer Assisted Tomography*, 26(4):493–504, July/August 2005.

[66] Iordanescu G, Pickhardt JP, Choi JR, and Summers RM. Automated seed placement for colon segmentation in computed tomography colonography. *Academic Radiology*, 12:182–190, 2005.

[67] Li H and Santago P. Automatic colon segmentation with dual scan ct colonography. *Journal of Digital Imaging*, 18(1):42–45, 2005.

[68] Frimmel H, Nappi J, and Yoshida H. Centerline-based colon segmentation for ct colonography. *Medical Physics*, 32:2665–2672, 2005.

[69] Gonzalez RC and Woods RE. *Digital image processing*. Reading MA: Addison-Wesley, 1993.

[70] Atlas of Human Anatomy. Eagle Editions.

[71] Dijkstra EW. A note on two problems in connexion with graphs. *Numerical Math*, 1:269–71, 1959.

[72] Sadleir RJT and Whelan PF. Colon centerline calculation for ct colonography using optimised 3d topological thinning. *First international symposium on 3D data processing visualisation and transmission*, 2002.

[73] Sezile N, Sadlier RJT, and Whelan PF. Fast extraction of planes normal to the centerline from ct colonography datasets. *International Conference on Visual Information Engineering*, pages 161–164, 2003.

[74] Munikrishnan V, Gillams AR, Lees WR, Vaizey CJ, and Boulos PB. Prospective study comparing multislice ct colonography with colonoscopy in the detection of colorectal cancer and polyps. *Diseases of the colon and rectum*, 46(10):1384–90, October 2003.

[75] Pineau BC, Paskett ED, Chen GJ, Espeland MA, Phillips K, Han JP, Mikulaninec C, and Vining DJ. Virtual colonoscopy using oral contrast compared with colonoscopy for the detection of patients with colorectal polyps. *Gastroenterology*, 125(2):304–310, 2003.

Bibliography

[76] Callstrom MR, Johnson CD, Fletcher JG, Reed JE, Ahlquist DA, Harmsen WS, Tait K, Wilson LA, and Corcoran KE. Ct colonography without cathartic preparation: Feasibility study. *Radiology*, 219(3):693–698, 2001.

[77] Hong L, Muraki S, Kaufman A, Bartz D, and He T. Virtual voyage: Interactive navigation in the human colon. *Proceeding ACM SIGGRAPH '97*, pages 27–34, August 1997.

[78] You S, Hong L, Wan M, Junyaprasert K, Kaufman A, Muraki S, Zhou Y, Wax M, and Liang Z. Interactive volume rendering for virtual colonscopy. *IEEE Visualization '97 Conference Proceeding, ACM/SIGGRAPH Press*, pages 433–436, October 1997.

[79] McFarland E, Brink J, Loh J, Wang G, Argiro V, Balfe D, Heiken J, and Vannier M. Visualization of colorectal polyps with spiral ct colography: Evaluation of processing parameters with perspective volume rendering. *Radiology*, 205(3):701–707, 1997.

[80] Wan M, Li W, Kreeger K, Bitter I, Kaufman A, Liang Z, Chen D, and Wax M. 3d virtual colonoscopy with real-time volume rendering. *SPIE Medical Imaging*, February 2000.

[81] Haker S, Angenent S, and Tannenbaum A. Nondistorting flattening maps and the 3-d visualization of colon ct images. *IEEE Transactions on Medical Imaging*, 19(7):665–670, July 2000.

[82] Bartroli AV, Wegenkittl R, Konig A, and Groller E. Nonlinear virtual colon unfolding. *In IEEE Visualization 2001 Conference Proceedings*, pages 411–418, October 2001.

[83] Dachman A. Diagnostic performance of virtual colonoscopy. *Abdominal Imaging*, 27:260–267, 2002.

[84] Sezille N, Sadleir RJT, and Whelan PF. Fast extraction of planes normal to the centreline from ct colonography datasets. *VIE 2003 - IEE Visual Information Engineering Conference, Guildford, UK*, pages 161–164, 2003.

[85] Frimmel H, NappiJ, and Yoshida H. Fast and robust computation of colon

centerline in ct colonography. *Medical Physics*, 31(11):3046–3056, November 2004.

[86] Nappi J, Frimmel H, and Yoshida H. Virtual endoscopic visualization of the colon by shapescale signatures. *IEEE Transactions on Information Technology in Biomedicine*, 9(1):120–131, March 2005.

[87] Sadleir RJT and Paul Whelan PF. Fast colon centreline calculation using optimised 3d topological thinning. *Computerized Medical Imaging and Graphics*, 29(4):251–258, 2005.

[88] Johnson C, Harmsen W, Wilson L, MacCarty R, Welch T, Ilstrup D, and Ahlquist D. Prospective blinded evaluation of computed tomographic colonography for screen detection of colorectal polyps. *Gastroenterology*, 125:311–319, 2003.

[89] van Gelder RE, Florie J, and Stoker J. Colorectal cancer screening and surveillance with ct colonography: current controversies and obstacles. *Abdominal Imaging*, 30:5–12, 2005.

[90] Vining DJ, Hunt GW, Ahn DK, Stelts DR, and Helmer PF. Computer-assisted detection of colon polyps and masses. *Radiology*, 219:51–59, 2001.

[91] Summers RM, Beaulieu CF, Pusanik LM, Malley JD, Jeffrey RB, Glazer DI, and Napel S. Automated polyp detector for ct colonography: Feasibility study. *Radiology*, 216:284–290, 2000.

[92] Yoshida H and Nappi J. Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. *IEEE Transactions on Medical Imaging*, 20(12):1261–1274, 2001.

[93] Paik DS, Beaulieu CF, and Rey RBJ. Computer aided detection of polyps in ct colonography: Method and free-response roc evaluation of performance. *Radiology*, 217(P):370, 2000.

[94] Kiss G, Cleynenbreugel J, Thomeer M, Suetens P, and Marchal G. Computer aided diagnosis for virtual colonography. *Medical Image Computing and Computer-Assisted Intervention*, pages 621–628, 2001.

Bibliography

[95] Kiss G, Cleynenbreugel J, Thomeer M, Suetens P, and Marchal G. Computer-aided diagnosis in virtual colonography via combination of surface normal and sphere fitting methods. *European Radiology*, 12(1):77–81, 2002.

[96] Kiss G, Cleynenbreugel J, Thomeer M, Suetens P, , and Marchal G. Computer aided detection of colonic polyps via geometric feature classication. *Vision, Modeling, and Visualization*, pages 27–34, 2002.

[97] Kiraly AP, Laks S, Macari M, Geiger B, Bogoni L, and Novak CL. A fast method for colon polyp detection in high-resolution ct data. *International Congress Series*, 1268:983–988, 2004.

[98] Acar B, Napel S, Paik D, Gokturk SB, Tomasi C, and Beaulieu CF. Using optical flow fields for polyp detection in virtual colonoscopy. *Medical Image Computing and Computer-Assisted Intervention, Utrecht, The Netherlands*, 14-17 October 2001.

[99] Gokturk SB, Tomasi C, Acar B, Beaulieu CF, Paik DS, Jeffrey RBJ, Yee J, and Napel S. A statistical 3d pattern processing method for computer aided detection of polyps in ct colonography. *IEEE Transactions on Medical Imaging*, 20(12):1251–1260, 2001.

[100] Wang Z, Liang Z, Li L, Li X, Li B, Anderson J, and Harrington D. Reduction of false positives by internal features for polyp detection in ct-based virtual colonoscopy. *Medical Physics*, 32(12):3602–16, 2005.

[101] Jerebko AK, Malley JD, Franaszek M, and Summers RM. Support vector machines committee classification method for computer-aided polyp detection in ct colonography. *Academic Radiology*, 12(4):479–86, April 2005.

[102] Iordanescu G and Summers RM. Reduction of false positives on the rectal tube in computer-aided detection for ct colonography. *Medical Physics*, 31(10):2855–2862, 2004.

[103] Zucker SW and Hummel RA. A three-dimensional edge operator. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(3):324–331, 1981.

[104] Schneider P and Eberly DH. *Geometric Tools for Computer Graphics*. Morgan Kaufmann Publishers, 2003.

Bibliography

[105] Lancaster P and Salkauskas K. *Curve and Surface Fitting: An Introduction.* London: Academic Press, 1986.

[106] Ghita O and Whelan PF. A bin picking system based on depth from focus. *Machine Vision and Application*, 13:234–244, 2003.

[107] Neural Network Toolbox. www.mathworks.com, 1992-2001.

[108] Vapnik VN. *The Nature of Statistical Learning Theory.* Springer, New York, 1995.

[109] Gunn SR. Support vector machines for classification and regression. Technical report, Image Speech and Intelligent Systems Research Group, University of Southampton, 1997.

[110] Burges CJC. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1988.

[111] Gelder REV, Venema HW, Florie J, Yung C, Serlie IWO, Schutter MP, Rijn JCV, Vos FM, Glas AS, Bossuyt PMM, Bartelsman JFW, Lameris JS, and Stoker J. Ct colonography: Feasibility of substantial dose reduction-comparison of medium to very low doses in identical patients. *Radiology*, 232:611–620, 2004.

[112] Brenner DJ, Elliston CD, and Hall EJ. Estimated risks of radiation induced fatal cancer from pediatric ct. *American Journal of Roentgenology*, 176(5):289–296, February 2001.

[113] Commission on Life Sciences. *Health effects of exposure to low levels of ionizing radiation: BEIR V (1990).* National Academy Press,Washington, DC, 1990.

[114] Cohen BL. Cancer risk from low-level radiation. *American Journal of Roentgenology*, 179(5):1137–1143, 2002.

[115] Gelder REV, Venema HW, Serlie IWO, Nio CY, Determann RM, Tipker CA, Vos FM, Glas AS, Bartelsman JFW, Bossuyt PMM, Lameris JS, and Stoker J. Ct colonography at different radiation dose levels: Feasibility of dose reduction. *Radiology*, 224:25–33, 2002.

Bibliography

[116] Iannaccone R, Laghi A, Catalano C, Brink JA, Mangiapane F, Trenna S, Pia-centini F, and Passariello R. Detection of colorectal lesions: lower-dose multi-detector row helical ct colonography compared with conventional colonoscopy. *Radiology*, 229:775–781, 2003.

[117] Capunay CM, Carrascosa PM, Bou-Khair A, Castagnino N, Ninomiya I, and Carrascosa JM. Low radiation dose multislice ct colonography in children: Experience after 100 studies. *European Journal of Radiology, In Press, Available online*, 5 July 2005.

[118] Vogt C, Cohnen M, Beck A, Dahl SV, Aurich V, Modder U, and Haussinger D. Detection of colorectal polyps by multislice ct colonography with ultra-low-dose technique: Comparison with high-resolution videocolonoscopy. *Gastrointestinal Endoscopy*, 60(2):201–209, August 2004.

[119] Beaulieu CF, Napel S, Daniel BL, Chen IY, Rubin GD, Johnstone IM, and JB Jeffrey. Detection of colonic polyps in a phantom model: Implications for virtual colonoscopy data acquisition. *Journal of Computer Assisted Tomography*, 22(4):656–663, 1998.

[120] Dachman AH, Lieberman J, Osnis RB, Chen SY, Hoffmann KR, Chen CT, Newmark GM, and McGill J. Small simulated polyps in pig colon: Sensitivity of ct virtual colography. *Radiology*, 203:427–430, 1997.

[121] Taylor SA, Halligan S, Bartram CI, Morgan PR, Talbot IC, Fry N, Saunders BP, Khosraviani K, and Atkin W. Multi-detector row ct colonography: Effect of collimation, pitch, and orientation on polyp detection in human colectomy specimen. *Radiology*, 229:109–118, 2003.

[122] Springer P, Stohr B, Giacomuzzi SM, Bodner G, Klingler A, Jaschke W, and Zur Nedden D. Virtual computed tomography colonoscopy: Artifacts, image quality and radiation dose load in a cadaver study. *European radiology*, 10(1):183–7, 2000.

[123] Whiting BR, McFarland EG, and Brink JA. Influence of image acquisition parameters on ct artifacts and polyp detection in spiral ct colonography: In vitro evalution. *Radiolgy*, 217:165–172, 2002.

Bibliography

[124] Power NP, Pryor MD, Martin A, Horrocks J, McLean AM, and Reznek RH. Optimization of scanning parameters for ct colonography. *The British Journal of Radiology*, 75:401–408, 2002.

[125] Wessling J, Fischbach R, Meier N, Allkemper T, Klusmeier J, Ludwig K, and Heindel W. Ct colongraphy: Protocol optimization with muti-detector row ct-study in an anthropomorphic colon phantom. *Radiology*, 228:753–759, 2003.

[126] Laghi A, Iannaccone R, Mangiapane F, Piacentini F, Iori S, and Passariello R. Experimental colonic phantom for the evaluation of the optimal scanning technique for ct colonography using a multi-detector spiral ct equipment. *European radiology*, 13(3):459–466, 2003.

[127] Embleton KV, Nicholson DA, Hufton AP, and Jackson A. Optimization of scanning parameters for multi-slice ct colonography: experiments with synthetic and animal phantoms. *Clinical radiology*, 58(12):955–63, December 2003.

[128] Ozgun A, Rollven E, Blomqvist L, Bremmer S, Odh R, and Fransson A. Polyp detection with mdct: A phantom-based evaluation of the impact of dose and spiral resolution. *American Journal of Roentgenology*, 184(4):1181–8, 2005.

[129] Ling SH, Summers RM, Loew MH, McCollough CH, and Johnson CD. Computer-aided detection of polyps in a colon phantom: Effects of scan orientation, polyps size, coliimation and dose. *Journal of Computer Assisted Tomography*, 26(6):1013–1018, 2002.

[130] Sundaram P, Beaulieu CF, Paik DS, Schraedley-Desmond P, and Napel S. Ct colonography: Does improved z resolution help computer-aided polyp detection? *Medical Physics*, 30(10):2663–74, October 2003.

[131] NHS Purchasing and Supply Agency. Report 05071, siemens somatom sensation open ct scanner technical evaluation. Technical report, December 2005.

[132] Park SH, Ha HK, Kim AY, Kim KW, Lee MG, Kim PN, Shin YM, Byeon JS, Yang SK, Kim JH, and Min YI. Flat polyps of the colon: Detection with 16-mdct colonographypreliminary results. *American Journal of Roentology*, 186:1611–1617, 2006.

Bibliography

[133] Kiss G, Cleynenbreugel JV, Drisis S, Bielen D, and Suetens P. Computer aided detection for low-dose ct colonography. *Lecture notes in computer science*, 3749:859–867, 2005.

[134] Sezille N, Sadleir RJT, and Whelan PF. Automated synthesis, insertion and detection of polyps for ct colonography. *Opto-Ireland - SPIE's Irish Meeting on Optoelectronics, Photonics and Imaging, Galway*, September 5th - 6th 2002.

[135] Choi R. Virtual colonoscopy. *Virtual Colonoscopy Center, Walter Reed Army Medical Center*.

[136] Lakare S, Chen D, Li L, Kaufman AE, Wax M, and Liang Z. Electronic colon cleansing using segmentation rays for virtual colonoscopy. *SPIE Medical Imaging, Physiology and Function from Multidimensional Images, San Diego, CA, USA*, 4683:412–418, 2002.

[137] Lihong Li, Chen D, Lakare S, Kreeger K, Bitter I, Kaufman AE, Wax M, Djuric PM, and Liang Z. An image segmentation approach to extract colon lumen through colonic material tagging and hidden markov random field model for virtual colonoscopy. *SPIE Medical Imaging, Physiology and Function from Multidimensional Images, San Diego, CA, USA*, 4683:406–411, 2002.

[138] Lakare S, Wan M, Sato M, and Kaufman A. 3d digital cleansing using segmentation rays. *IEEE Visualization 2000 Proceedings, Salt Lake City, Utah, USA*, pages 37–44, 2000.

[139] Miranda A, Chowdhury T, Ghita O, and Whelan P. Shape filtering for false positive reduction at computed tomography colonography. *MICCAI 2006 - 9th International Conference on Medical Image Computing and Computer Assisted Intervention, Copenhagen, Denmark*, 1-6 October 2006.

[140] Aberle DR, Chiles C, Gatsonis C, Hillman BJ, Johnson CD, McClennan BL, Mitchell DG, Pisano ED, Schnall MD, and Sorensen AG. Imaging and cancer: research strategy of the american college of radiology imaging network. *Radiology*, 235(3):741–51, January 2005.

[141] Farid H and Simoncelli EP. Differentiation of multi-dimensional signals. *IEEE Transactions on Image Processing*, 13(4):496–508, 2004.

Bibliography

[142] Simoncelli EP. Design of multi-dimensional derivative filters. *IEEE International Conference on Image Processing*, 1:790–793, 1994.

[143] Jahne B, Scharr H, and Korkel S. *Principles of filter design, Handbook of Computer Vision and Applications*. Academic Press, 1999.

[144] Freeman WH and Adelson EH. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:891–906, 1991.

[145] Osher S and Sethian JA. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1998.

[146] Sethian JA. A marching level set method for monotonically advancing fronts. *In Proceedings of the National Academy of Sciences*, 93, 1996.

[147] Neskovic P and Kimia BB. Threedimensional shape representation from curvature dependent surface evolution. *In IEEE Proceedings for the Internaltional Conference of Image Processing*, pages 6–10, 1994.

[148] Michael L, Chowdhury T, Ghita O, and Whelan PF. Determining candidate polyp morphology from ct colonography using a level-set method. *European Medical and Biological Engineering Conference EMBEC 2005, Prague, Czech Republic*, November 2005.

Appendix A

The extraction of the gradient information from 3D surfaces plays an important role for many applications including 3D graphics and medical imaging. The extraction of the 3D gradient information is performed by filtering the input data with high pass filters that are typically implemented using $3 \times 3 \times 3$ masks. Since these filters extract the gradient information in a small neighborhood, the estimated gradient information will be very sensitive to image noise. The aim of this Appendix is to detail the implementation of an optimized 3D gradient operator that is applied to sample the local curvature of the colon wall in CT data and evaluate its influence on the overall performance of the developed *CAD-CTC* method. The developed 3D gradient operator has been applied to extract the local curvature of the colon wall in a large number CT datasets captured with different radiation doses and the experimental results are presented and discussed.

A-1 Mathematical background of gradient detection

In image processing the gradient operators are widely used to identify strong data features such as edges or the local orientation of the curves and surfaces. The extraction of the local derivative from a continuous signal can be done by applying directly the well-known derivative formula:

$$Def(f(x)) = lim_{\alpha \to 0} \frac{f(x + \alpha) - f(x)}{\alpha} \qquad \text{(A-1.1)}$$

When designing a gradient operator one should bear in mind that the image data is discrete and the finite differences cannot be applied without compromising the accuracy of the gradient approximation [141, 142, 143, 144]. Thus, it is assumed that the original continuous optical signal that generates the image has been uniformly sampled at a rate of T samples per length. Using the Nyquist sampling theorem the continuous signal can be reconstructed from these discrete samples as follows:

$$f(x) = \sum_{k=1}^{N} f[k]s(x - KT), s(x) = \frac{\sin(x)}{x} \qquad \text{(A-1.2)}$$

In equation B-1.2 the term $f[k]$ represents the discrete sampled signal and $s(x)$ defines the sampling function that can be approximated with the *sinc* function. Hence, to obtain the gradient of the discrete signal derivative was required for the reconstructed signal $f(x)$ that is depicted in equation A-1.2.

$$Der(f(x)) = \sum_{k=1}^{N} f[k]der(s(x - KT)) = \sum_{k=1}^{N} f[k]s'(x - KT) \qquad \text{(A-1.3)}$$

where $s\prime$ represents the derivative of the *sinc* function. As the derivative of the *sinc* function is dependent on the sampling frequency, it is worth noting that the spectrum of the discrete signal is bounded by $\frac{2\pi}{T}$ that is in agreement with the sampling theorem. It was noted that the derivative of *sinc* signal decays relatively slowly and the implementation of an optimal gradient filter would require large filters that are not feasible to be applied in practice due to the onerous computational cost required to extract the partial derivatives. Next, a practical method will be introduced to design one-dimensional (1-D) gradient filters whereas the generalization to multiple dimensions is a relatively simple task.

In order to design gradient operators that are to be applied to discrete signals several constraints have been considered. The vision literature indicates that the gradient filters are anti-symmetric and usually have an odd order. Thus, the 1-D gradient filter can be represented in the following generic form:

$$d(k) = [d_{-N}, d_{-N+1}, ..., d_{-1}, 0, d_1, ..., d_{N-1}, d_N], d_{-k} = -d_k, k = 1, ..., N \qquad \text{(A-1.4)}$$

In order to design 1-D derivative filters several constraints were imposed for parameters d_k as illustrated in the following expressions [143],

$$\sum_{k=-N}^{N} d_k = 0 \qquad \text{(A-1.5)}$$

$$\sum_{k=-N}^{N} d_k K = 0 \qquad \text{(A-1.6)}$$

In this way, the equation 5 translates in the requirement that the derivative filter should have the sum of the coefficients equal to 0, while equation A-1.6 can be used

to select the values for d_k coefficients. The derivative operator has to fulfill the condition illustrated in equation A-1.5 to achieve insensitivity to DC signals. Since the derivative filters are anti-symmetric the first coefficient of the operator can be determined using the following relationship:

$$d_1 = \frac{1}{2} - \sum_{k=2}^{N} d_k \qquad (A\text{-}1.7)$$

Using the formulas illustrated in equations 4 to 7, a $5 \times 5 \times 5$ derivative filter that is applied to extract the gradient in the x direction has the following mask $[\text{-}1\ 8\ 0\ \text{-}8\ 1]/12 \bullet [1\ 4\ 6\ 4\ 1]/16$, where $\bullet$ defines the convolution operator. To extract the gradient for other directions the $5 \times 5 \times 5$ mask need to be rotated in the direction required for a particular axis. It can be noted that this operator, as expected, represents the direct extension of the 5×5 Sobel operator to the 3D case. Using equations 5 to 7, a new 5×5 gradient operator can be developed that implements a two-peak operator illustrated in Figure A-1. This gradient operator shows two peaks in the frequency domain and it will provide improved performance when applied to data with step discontinuities or 3D CT datasets defined by a low signal to noise ratio such as the low-dose CT data.

```
-.125  -.25   -.5   -.25  -.125          -.25   -.5   -1.0   -.5   -.25
 -.25   -.5  -1.0   -.5   -.25           -.5  -1.0   -2.0  -1.0   -.5
  -.5  -1.0  -2.0  -1.0   -.5           -1.0  -2.0   -4.0  -2.0  -1.0      0₅ₓ₅
 -.25   -.5  -1.0   -.5   -.25           -.5  -1.0   -2.0  -1.0   -.5
-.125  -.25   -.5   -.25  -.125          -.25   -.5   -1.0   -.5   -.25


         .25   .5   1.0   .5   .25       .125  .25   .5   .25  .125
          .5  1.0   2.0  1.0   .5         .25   .5  1.0   .5   .25
         1.0  2.0   4.0  2.0  1.0          .5  1.0  2.0  1.0   .5
          .5  1.0   2.0  1.0   .5         .25   .5  1.0   .5   .25
         .25   .5   1.0   .5   .25       .125  .25   .5   .25  .125
```
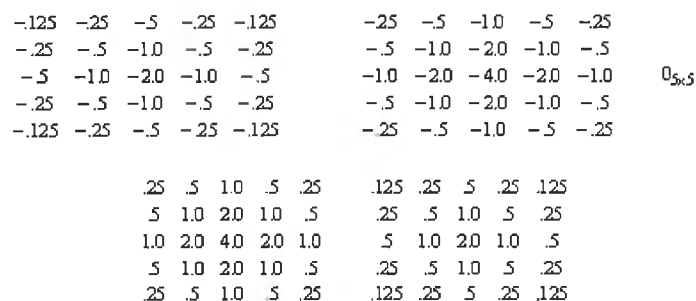
Figure A-1: The masks of the $5 \times 5 \times 5$ 3D OptDer operator to extract the gradient in the z axis (the mask $0_{5 \times 5}$ indicates a 5×5 mask where all elements are zero)

In these experiments the effect of using several filters including the $3 \times 3 \times 3$ Zucker-Hummel operator, $5 \times 5 \times 5$ Sobel operator and $5 \times 5 \times 5$ optimized operator - OptDer filter (for more details about the implementation of optimal derivative filters refer to [141, 142]) on the overall performance of the CAD-CTC system has been evaluated. A particular interest we had in assessing the performance of these gradient operators when applied to CTC datasets that have been acquired with different radiation doses. In these experiments it has became clear that the $3 \times 3 \times 3$

gradient masks are inefficient in sampling the correct curvature of the colon wall when dealing with irregular surfaces while the optimized $5 \times 5 \times 5$ gradient operator was able to return improved performance (this operator has been designed using the masks illustrated in Figure A-1). The experiments were performed on *CTC* prone and supine views where the reconstruction interval was set to 1.5mm. The tests were conducted on phantom (synthetic) data and on real patient data. Of particular interest was the evaluation of the level of false positives detected by the automated *CAD-CTC* system and a detailed performance of the developed system is illustrated in Tables 1 to 5 where different gradient operators are evaluated.

A-2   Results and Discussions

The statistical feature-based method discussed in Chapter 3 was used for automatic polyp detection *CAD-CTC*. The efficiency of the derivative operators was evaluated on 52 standard dose (100mAs) patient datasets (prone and supine views) with 75 polyps, 9 low dose (13-50mAs) patient data with 2 small polyps and phantom data (low-dose and standard dose) with 48 polyps of various sizes and shapes.

Table A-1: Sensitivity for synthetic phantom data (polyps $>= 10mm$).

mAs	Total	Sensitivity		
		Zuker	*Sobel*	*OptDer*
100	14	100%	100%	100%
40	14	100%	100%	100%
30	14	100%	92.85%	100%
20	14	100%	100%	100%
13	14	92.85%	92.85%	100%

Table A-2: Sensitivity for synthetic phantom data (polyps $[5 - 10)mm$).

mAs	Total	Sensitivity		
		Zuker	*Sobel*	*OptDer*
100	20	100%	100%	100%
40	20	100%	100%	100%
30	20	95%	90%	95%
20	20	100%	100%	95%
13	20	95%	95%	100%

When the *CAD-CTC* system was applied on phantom data the OptDer operator shows 100% sensitivity for polyps $>= 10mm$ for datasets acquired with radiation

Table A-3: Sensitivity for synthetic flat polyps.

mAs	Total	Sensitivity			False Positive		
		Zuker	*Sobel*	*OptDer*	*Zuker*	*Sobel*	*OptDer*
100	9	55%	55%	44.44%	1	1	1
40	9	33.33%	33.33%	44.44%	2	1	1
30	9	44.44%	44.44%	55%	0	2	2
20	9	11.11%	33.33%	44.44%	2	2	2
13	9	22.22%	22.22%	44.44%	2	2	3

Table A-4: Sensitivity for polyps $>= 5mm$ in real patient standard dose (100mAs) data.

mAs	Total	Sensitivity			False Positive		
		Zuker	*Sobel*	*OptDer*	*Zuker*	*Sobel*	*OptDer*
100	18	88.89%	88.89%	88.89%	4.32	4.69	4.71

doses in the range 100-13mAs where the Zucker-Hummel and Sobel operators shows 92.85% sensitivity at 30mAs and 13mAs radiation doses (see Table A-1). Figure A-2(a) illustrates the 3D surface extraction for a 12mm polyp when the Zucker-Hummel operator was applied to compute the surface normal vectors and Figure A-2(b) shows the surface extraction using the OptDer operator. Figure A-3 illustrates the surface extraction for an 8 mm phantom polyp from a dataset scanned with 13mAs. It can be noted that in both cases the *CAD-CTC* system achieved a more accurate surface extraction when the OptDer operator was employed. Due to incomplete surface segmentation the developed *CAD-CTC* system missed the polyp illustrated in Figure 2 when the Zucker-Hummel operator was used to extract the surface normal vectors (see Table A-1), whereas the polyp was correctly detected when the OptDer operator was applied. In Figures 2 and 3 it can be also observed that the OptDer operator generates better surface normal concentration than the Zucker-Hummel operator. The application of the OptDer operator to extract the surface normal vectors offers better detection for polyps in the range 5-10mm than the

Table A-5: Sensitivity for polyps $< 5mm$ in real patient's standard and low dose data.

mAs	Total	Sensitivity		
		Zuker	*Sobel*	*OptDer*
100	48	60.41%	60.41%	68.75%
13-40	2	50%	100%	100%

Sobel operator (see Table A-2). It also provides a better detection of flat polyps when compared to the performance of the Zucker-Hummel and Sobel operators (see Table A-3). When the Zuker-Hummel, Sobel and OptDer operators were used to calculate the surface normals of the colon wall for standard dose real patient datasets, the sensitivities for the detection of polyps $>= 5mm$ were 88.89% (see Table A-4) for all operators, but the OptDer operator provides higher sensitivity (see Table A-5) in the detection of small polyps ($< 5mm$) than the Zucker-Hummel and Sobel operators. Table A-5 indicates that the overall sensitivity for polyp detection was highest when the OptDer operator was used and the experimental data indicates that this operator outperformed the Zucker-Hummel and the Sobel operators especially when the system is applied to low-dose datasets.

(a)   (b)

Figure A-2: 3D surface extraction of a 12mm phantom polyp (radiation dose 13mAs). (a) The 3D surface extracted by the CAD-CTC system using the Zucker-Hummel operator. (b) The 3D surface extracted by the CAD-CTC system using the OptDer operator.

A-3   Conclusions

The main objective of this Appendix was to address the problem of robust calculation of the surface curvature in 3D CT data. As numerous automated *CAD-CTC* systems identify the colorectal polyps based on analysing the local convexity of the colon surface, one of the most important steps in this analysis is the precise calculation of the normal vectors. In this regard, a number of 3D gradient operators were investigated and the experiments were conducted on a large number of synthetic and real patient datasets. Experimental data indicated that the commonly used 3D
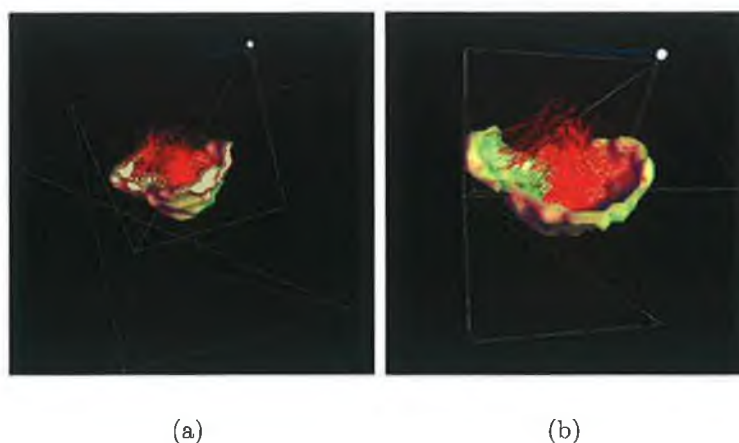
(a)   (b)
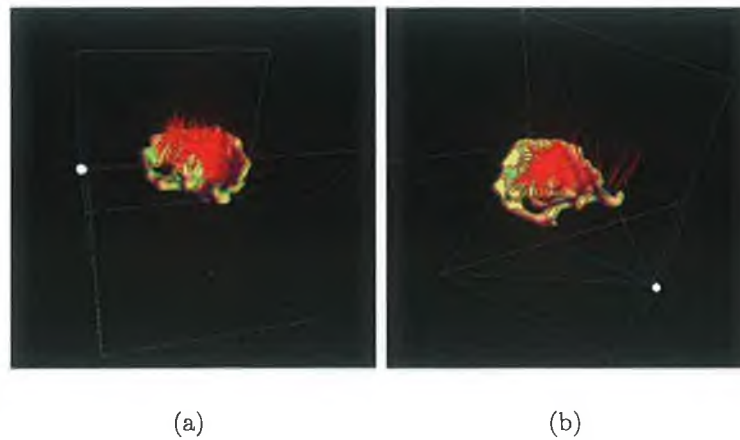
Figure A-3: 3D surface extraction of a 12mm phantom polyp (radiation dose 13mAs). (a) The 3D surface extracted by the CAD-CTC system using the Zucker-Hummel operator. (b) The 3D surface extracted by the CAD-CTC system using the OptDer operator.

gradient operators such as Zucker-Hummel and Sobel fail to accurately determine the normal vector when dealing with datasets characterized by a low signal to noise ratio. To address this problem a new gradient operator was proposed that was able to return better performance when applied to CT data that is acquired with different radiation dose levels.

Appendix B

This section describes a method for the accurate segmentation of polyp candidate surface using a level-set segmentation method. The level set is a deformable surface that evolves under a force that includes gradient and curvature. The curvature property was exploited in the evolution to extract only the surface of the candidate polyp to avoid over segmentation of the colon wall.

B-1 Level-Set Initialisation. Fast-Marching Algorithm

The formulation of the level-set formulation is conceptually simple. The evolving curve or front Γ, evolves as the zero levelset of a higher dimensional function ϕ. This function deforms with a force F that is dependent on both curvature of the front and external forces in the image. The force acts in the direction of the normal to the front.

$$\phi_t + F|\nabla\phi| = 0 \qquad \phi(x, y, t = 0) = given \qquad \text{(B-1.1)}$$

The proposed implementation is a standard two step approach which includes a fast-marching initial step to speed up the segmentation. Fast marching is a special case of the above equation where $F(x, y) > 0$. Let $T(x, y)$ be the time when the front Γ crosses the point (x, y). The function $T(x, y)$ then satisfies the equation;

$$|\nabla T|F = 1 \qquad \text{(B-1.2)}$$

which simply says that the gradient of the arrival time is inversely proportional to the speed of the surface. The T function is evaluated using the diffusion and attraction to pixels within the front. This forces the front to grow out from its initial position to points with the smallest value of $T(x, y)$. The $T(x, y)$ function is then updated until the front converges to a stable state.

B-2   Level-Set Analysis

The theory behind level-set segmentation is largely based on work in partial differential equations and the propagation of fronts under intrinsic properties such as curvature [145, 146]. Representing the boundary as the zero level set instance of a higher dimensional function ϕ, the effects of curvature can be easily incorporated. ϕ is represented by the continuous Lipschitz function $\phi(s, t = 0) = \Gamma \pm d$, where d is the signed distance from position s to the initial interface $\Gamma 0$ (see Equation B-2.1). The distance is given a positive sign outside the initial boundary ($D\Omega$), a negative sign inside the boundary ($|\Omega \setminus \partial\Omega|$) and zero on the boundary ($\partial\Omega$).

$$\phi(s) = \begin{pmatrix} -d & \forall s \in \Omega \setminus \partial\Omega \\ 0 & \forall s \in \partial\Omega \\ +d & \forall s \in R^n \setminus \partial\Omega \end{pmatrix}. \tag{B-2.1}$$

From this definition of ϕ, intrinsic properties of the front can be easily determined, like the normal $n = \pm\frac{\nabla\phi}{|\nabla\phi|}$.

Since curvature of the polyp is an important factor in the segmentation evolution, particular emphasis is given to this measure. The mean curvature (H), is connected to the physical evolution of soap bubbles and the heat equation as follows:

$$H = \nabla.\frac{\nabla\phi}{|\nabla\phi|} \tag{B-2.2}$$

Gaussian curvature (K), has also being used to model physical problems and can be calculated using the following expression:

$$K = \frac{\nabla\phi^T Adj(H(\phi))\nabla\phi}{|\nabla\phi|^2} \tag{B-2.3}$$

where $H(\phi)$ is the Hessian matrix of ϕ, and $Adj(H)$ is the adjoint of the matrix H.

The proposed method used the Neskovic and Kimias [147] measure of curvature which involves both mean and Gaussian. In this approach, the direction of flow is obtained from the Mean curvature while the magnitude of the flow is dictated by the Gaussian curvature. This is appropriate as the Mean curvature alone can cause singularities and extracts the strictly convex surface of the polyp candidate.

$$k = sign(H)\sqrt{K + |K|} \tag{B-2.4}$$

Using this value for k, the level set is iteratively updated within a defined narrow band around the segmented boundary to increase the computational efficiency. The

following equation details the update parameters

$$\phi_{t+1} = \phi_t + K_I(1 - \varepsilon\kappa)|\nabla| + \beta\nabla I.\nabla\phi \qquad \text{(B-2.5)}$$

where ε and beta are user defined parameters (see Table 1), κ is the curvature term defined in Equation B-2.4and K_I is the gradient dependent speed term and is given by $\frac{1}{1+\nabla I}$. The third term, $\nabla I.\nabla\phi$ represents the attraction force vector normal to the front.

Possible polyp candidate centres are determined over the entire data set by calculating the normal vectors at each voxel on the colon wall. Polyp candidates are defined as regions of high convexity, therefore the centres for possible polyp candidates are located at points that contain high concentration of normal intersections (see Chapter 3).

The level set is initialised at the polyp candidate centres and grows outwards until a stable boundary is encountered. The convex surface is maintained by placing a high influence on the curvature parameter. Once the level-set has converged the surface of the polyp candidate is taken as all boundary points that have an associated gradient in order to ensure that only the colon surface is extracted.

Table B-1: Control parameters used in the level-set segmentation [148].

Index	Control Parameters	Values
1	Fast-Marching Iterations	3
2	Level-set Iterations	10
3	Level-set ε	0.5
4	Level-set β	0.08
5	Level-set Narrow bandwidth	10

Once the true surface of the polyp candidates has being extracted, they are passed to a classifier to determine whether they are polyps or folds. The statistical features that are discussed in Chapter 3 are used to classify the candidate polyp surfaces into polyps or folds using the *FNNN* classifier.

B-3 Results

In total 181 polyp candidates were tested through the volume. Visual representations of the segmentation polyp are shown in Figure B-1. Table B-1 lists the user defined parameters used in the level-set algorithm. From this table it can be seen

that curvature is given a large influence to preserve the convexity of the polyp candidate surface. The narrow bandwidth is given a small value of 10 to increase the efficiency of the update.
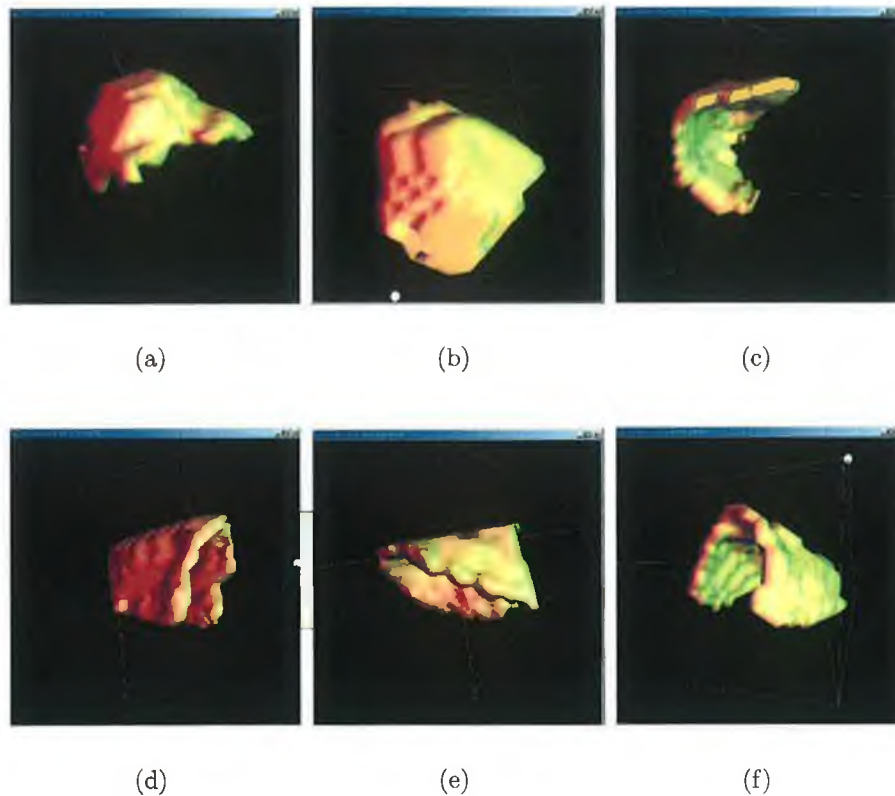

(a) (b) (c)

(d) (e) (f)

Figure B-1: Images above show the polyp candidate renderings of the extracted surface. Figures (a)-(c) show correctly classified polyps, where Figures (d)-(f) show correctly classified folds.

Table B-2 shows the measured point-to-curve error between the automatic segmentation results against those found from a manual segmentation of the small number of polyp candidates. Indicated in the table are the average error, standard deviation of the error and the rootmean - square (RMS) of the error. This error is measured in voxels.

Table B-2: Control parameters used in the level-set segmentation.

Average	Standard Deviation	RMS
0.298	0.587	0.661

Table B-3 gives the results on two real patient supine data sets. From the high number of polyp surface candidates(181 and 191), a relatively low number are detected (6 and 3). The results show a sensitivity of 100% for all polyps larger than

5mm. In current clinical studies the polyps below 5mm are discarded in the classification. One cause that generated the low sensitivity for detection of polyps smaller than 5mm is the low curvature difference between the polyp and the colon wall, therefore parts of the colon wall is taken into the candidate surface (see Figure B-2). One particular advantage of this surface extraction technique is the low number of false positives present in the analysed data.

Table B-3: Performance analysis for automatic polyp detection.

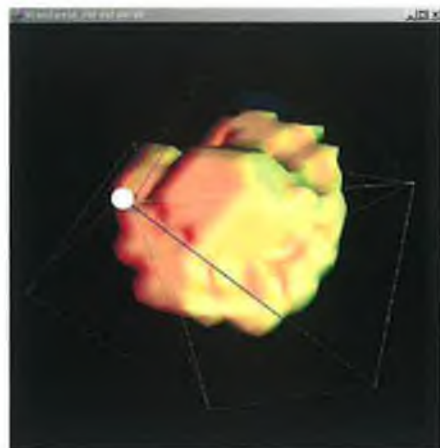| Data	Size	Detected	TP	FP	Missed
Data 1 Supine	$> 5mm$	6	3	3	0
(181 surf.)	$\leq 5mm$	0	0	0	2
Data 2 Supine	$> 5mm$	3	2	1	0
(191 surf.)	$\leq 5mm$	0	0	0	2
Total		9	5	4	4

Figure B-2: One of the $\leq 5mm$ polyps misclassified due to the inclusion of colon wall in the surface extraction.

In conclusion, the accurate segmentation described in this Appendix is the first important step in the classification of polyp candidates into polyp and fold. This Appendix describes a method for the extraction of accurate polyp candidate surfaces using a level-set segmentation. The level-set is initialised using the distribution of surface normal vectors and the resulting surfaces are classified into polyp and non-polyp. The level-set evolution is constrained by the image gradients and by the curvature of the boundary and is able to perform robust polyp segmentation when applied to standard and low dose datasets.

Author's Publication

Journal Publication

- Tarik A. Chowdhury, Paul F. Whelan, Ovidiu Ghita, Nicholas Sezille, Shane Foley, Development of a synthetic phantom for the selection of optimal scanning parameters in CAD-CT colonography, Journal of Medical physics and Engineering, (Accepted for publication).

- Tarik A. Chowdhury, Paul F. Whelan and Ovidiu Ghita, The use of 3D surface fitting for robust polyp detection and classification in CT colonography, Journal of Computerized Medical Imaging and Graphics (In Press).

Peer Reviewed Conference Papers

- Tarik Chowdhury, Ovidiu Ghita, Paul Whelan, Evaluation of 3D gradient filters for estimation of the surface orientation in CTC, Irish Machine Vision and Image Processing Conference, 30th august - 1st September, 2006, Dublin, Ireland.

- Tarik A. Chowdhury, Ovidiu Ghita, Paul F. Whelan and Abhilash Miranda, A Note on Feature Selection for Polyp Detection in CT Colonography, The 18th International Conference on Pattern Recognition, Hong Kong , 20-24 August, 2006.

- Tarik A. Chowdhury, Paul F. Whelan, and Ovidiu Ghita, A Method for Automatic Segmentation of Collapsed Colons at CT Colonography, 2nd Indian International Conference on Artificial Intelligence, December 20-22, 2005, Pune, India.

- Tarik A. Chowdhury, Ovidiu Ghita and Paul F. Whelan, A statistical approach for robust polyp detection in CT colonography, 27th Annual International

Conference of the IEEE Engineering in Medicine and Biology Society, 1-4 September 2005, Shanghai, China.

- Michael Lynch, Tarik Chowdhury, Ovidiu Ghita and Paul F. Whelan (2005), Determining Candidate Polyp Morphology from CT Colonography using a Level-Set Method, European Medical and Biological Engineering Conference EMBEC 2005, November 2005, Prague, Czech Republic.

Abstracts and Posters

- T.A. Chowdhury, P.F. Whelan, H. Fenlon, P. MacMathuna, Evaluation of radiation dose on automatic polyp detection at CT colonography: Experiments with a synthetic phantom, Association of Physical Scientists in Medicine, 2005 Annual Scientific Meeting, Galway, 25-26February, 2005 (Poster Presentation).

- T.A. Chowdhury, R.J.T. Sadleir, P.F. Whelan, N. Sezille, A. Moss, A. O Hare, S. Foley, H. Fenlon, P. MacMathuna (2004), Automatic Detection of Colon at CT Colonography, Irish Society of Gastroenterology, Winter Meeting 2004 (Abstract / Presentation).

- R.J.T. Sadleir P.F. Whelan, N. Sezille, T.A. Chowdhury, J. Bruzzi, A. Moss, P. McMathuna, H. Fenlon (2004), Automatic detection of colorectal polyps at CT colonography using shape information, Association of Physical Scientists in Medicine, 2004 Annual Scientific Meeting, Dublin,11th June 2004, (Presentation).

- T.A. Chowdhury, R.J.T. Sadleir, P.F. Whelan, A. Moss, J. Varden, M. Short, H. Fenlon, P. MacMathuna (2004), The impact of radiation dose on imaged polyp characteristics at CT colonography: Experiments with a synthetic phantom, Association of Physical Scientists in Medicine, 2004 Annual Scientific Meeting, Dublin,11th June 2004, (Presentation).

- R.J.T. Sadleir, P.F. Whelan, N. Sezille, T.A. Chowdhury, A. Moss, J. Bruzzi, H. Fenlon, P. MacMathuna, Computer-Aided Detection of Colorectal Polyps at CT Colonography, Irish Society of Gastroenterology Winter 2003. (Poster Presentation).

Under review

- Tarik A. Chowdhury, Paul F. Whelan and Ovidiu Ghita, A Method for Automatic Segmentation of Collapsed Colons in CT Data, International Journal of Tomography and Statistics, under review.