ONLINE EXTRACTION OF LUMEN REGION AND BOUNDARY FROM ENDOSCOPIC IMAGES USING A QUAD STRUCTURE

S Kumar1, K V Asari2 and D Radhakrishnan2

1National University of Singapore, 2Nanyang Technological University, Singapore

ABSTRACT
A new approach for the automatic extraction of the lumen region and its boundary for the gastrointestinal (GI) endoscopic images is presented. At first, a quasi region of interest representing the darker regions of the image is segmented by using a region splitting scheme named progressive thresholding. Then the lumen region is obtained by using a region growing technique called integrated neighbourhood search (INS). A new quad-structure based technique is introduced to enhance the speed of INS significantly. A back projection algorithm is suggested to optimise the search for the pixels belonging to the lumen region and boundary. A boundary-thinning algorithm is also proposed to remove the redundant pixels from the lumen boundary and to generate a connected single pixel width boundary. The proposed approach does not need a priori knowledge about the image characteristics. The efficiency of this method in terms of speed and accuracy is validated by various GI images and the results of the experiments are presented. The main advantage of the proposed technique is its high-speed response that facilitates real time analysis of the endoscopic images.

INTRODUCTION
Video-endoscopes consist of a miniature CCD camera that captures the images of the GI tract for the purpose of observation and diagnosis. Significant advancements have been made to automate the process of endoscopy as reported in Kumar et al (1) and Dai et al (2). The automated endoscopy system involves a vision-guided microrobotic device that navigates inside the body with the help of the images captured by the onboard camera. The lumen region and boundary in intestinal images form the preliminary basis of the features used for navigation and automatic guidance. The high speed and accurate extraction of these features is essential for robotic navigation. The area of lumen region and its boundary are used to define a number of important quantitative parameters that are used by the artificial intelligence based algorithms to render active advice to the endoscopist.

Extraction of lumen region from the endoscopic images is essentially a region segmentation process. Segmentation of the region of interest from a given medical image using edge information has been described by Deklerck et al (3). A comprehensive survey of various edge detection algorithms is given in Heath et al (4). But the edge information alone is not sufficient to segment the region accurately. To enhance the reliability of the segmented region integrated region growing and edge detection techniques have been proposed to merge the local edge information into the region information by Gambotto (5). The above techniques are either not very accurate or computationally expensive for real-time implementation. Also some of the above mentioned segmentation techniques need a fixed range of parameters which precludes their use in endoscopic image processing where no a priori knowledge can be used due to highly unpredictable image structure.

In this paper, a dual-step technique for lumen extraction is proposed. Firstly, a quasi-region of interest approximating the lumen area is segmented by employing a progressive thresholding approach based on threshold optimisation. Then, the actual boundary is obtained by applying integrated neighbourhood search using a quad structure. The proposed technique gives a high-speed response due to the use of quad structure that optimises the search of the lumen region and boundary.

ESTIMATION OF QUASI-REGION OF INTEREST AND QUASI CENTRE
A quasi region of interest (qROI) is segmented from the original gray level image using a histogram thresholding technique. To obtain a qROI we modify the thresholding technique presented in Otsu (6) in a similar way as suggested by Cerit et al. (7). Otsu’s thresholding technique is based on a discriminant analysis which partitions the image into two classes G0 and G1 at gray level t such that G0 = {0, 1, 2, ..., t} and G1 = {t+1, t+2, ..., L-1}, where L is the total number of gray levels in the image. Let $\sigma_B^2$ and $\sigma_T^2$ be the between-class variance and total variance respectively. An appropriate threshold t' can be obtained by maximising the between-class variance. Hence

$$t' = \arg \max_{0 \leq i < L} \left( \frac{\sigma_B^2}{\sigma_T^2} \right)$$

(1)

where

$$\sigma_B^2 = \frac{1}{N} \sum_{i=0}^{N} \left( \frac{1}{i} - \frac{1}{N} \right)^2 (\mu_i - \mu_0)^2$$

(2)

$$\sigma_T^2 = \frac{1}{N} \sum_{i=0}^{N} (i - \mu_0)^2$$

(3)
\( n_i \) is the number of pixels on the \( i^{th} \) gray level, \( N \) is the total number of pixels in the image, \( \mu_0 \) and \( \mu_1 \) represent the class means for \( G_0 \) and \( G_1 \) respectively, and \( \mu_T \) is the total mean. The mean values are calculated as

\[
\mu_0 = \frac{\sum_{i=0}^{1} n_i}{\sum_{i=0}^{1} N} \tag{4a}
\]

\[
\mu_1 = \frac{\sum_{i=0}^{1} n_i \cdot T}{\sum_{i=0}^{1} N} \tag{4b}
\]

\[
\mu_T = \frac{\sum_{i=0}^{1} n_i \cdot T}{\sum_{i=0}^{1} N} \tag{4c}
\]

This procedure requires iterative computation to find the best threshold for a given image. Otsu’s method of thresholding gray level images is efficient for separating an image into two classes where two types of fairly distinct classes exist in the image. In the present case, after finding the first appropriate threshold \( t_1^{*} \), a new histogram called progressive histogram is made using the pixels having intensity lower than the \( t_1^{*} \) while excluding the higher intensity pixels. Again Otsu’s procedure defined by eqns. 1 to 3 is applied on the progressive histogram and a second appropriate threshold \( t_2^{*} \) is obtained. It has been found that \( t_2^{*} \) serves as an appropriate measure for segmenting the darker regions from an endoscopic image. This is because the first thresholding removes the regions of high intensity due to mucosal reflections and the second one segments the darker regions from the rest of the image.

After finding the appropriate threshold, the intensity image is binarised. The region represented by the black pixels in the binarised image represents qROI. The spatial centre of qROI, named as quasi centre, qC is calculated by using all the black pixels contained in the binarised image as

\[
x_q = \frac{L}{K} \sum_{i=1}^{k} x_i \quad \text{and} \quad y_q = \frac{L}{K} \sum_{i=1}^{k} y_i \tag{5}
\]

where \((x_i, y_i)\) are the co-ordinates of the \( i^{th} \) pixel and \( K \) is the total number of black pixels contained in the binarised image. This centre is an obvious choice for being the seed to obtain the actual lumen region through region growing as the largest connected dark region is the most probable case for lumen.

**FORMATION OF A PYRAMIDAL QUAD STRUCTURE**

After binarising the image, a pyramidal quad structure is developed by generating a hierarchy of smaller images called daughter images from the parent image. The parent image of size \( M \times N \) pixels forms the base of the pyramid. To make a truncated pyramid having \( P \) levels, the image width and the image height at every successive level from the base are reduced by half. The total number of pixels in the image at the \( p^{th} \) level of the pyramid can be computed as

\[
G_p = \frac{G_0}{2^{2p}} \quad \text{for } p = 0, 1, 2, \ldots, (P-1) \tag{6}
\]

where \( G_0 \) represents the number of pixels in the parent image. The pixel intensity information of the image on the \( p^{th} \) level is obtained recursively as

\[
P^p(x, y) = \frac{1}{4} \sum_{i=0}^{1} \sum_{j=0}^{1} P^{p-1}(x+i, y+j) \tag{7}
\]

for \( p = 1, 2, \ldots, (P-1) \)

where \( P^{p-1}(x, y) \) represents the gray level of pixel at the location \((x, y)\) in the \((p-1)^{th}\) image. It may be noted that in the pyramidal structure pertaining to an image, the resolution of the image decreases while moving from the parent image to the youngest daughter. As the new level of image is generated using the prior level, a new quasi centre for that particular level is computed. The location of the quasi centre at \( p^{th} \) level of the pyramid can be computed as

\[
\left( x_q^p, y_q^p \right) = \left( \left\lfloor \frac{x_q}{2^{p-1}} \right\rfloor, \left\lfloor \frac{y_q}{2^{p-1}} \right\rfloor \right) \tag{8}
\]

where \( \left\lfloor \cdot \right\rfloor \) represents the upper integer value.

**LUMEN REGION AND BOUNDARY EXTRACTION**

The lumen region and boundary are extracted using a two step process. Firstly, a region is grown by integrated neighbourhood search (INS) in the youngest daughter image using the corresponding quasi centre as the seed. Then, this grown region is projected back onto the parent image through the intermediate levels to get the original lumen region and boundary.

**Integrated Neighbourhood Search**

Image segmentation based on INS is a region analysis procedure in which similarities between the immediate neighbourhood pixels are examined as in Asari et al. (8). The quasi centre of the youngest daughter is used as a seed for growing a region to identify all the pixels in the neighbourhood. In a daughter image, each pixel may have three states viz. black pixel, white pixel or intermediate gray pixel. If a neighbouring pixel is black, it is included in an array called Black Pixel Array (BPA) and if it is gray, it is included in another array named as Gray Pixel Array (GPA). The white pixels are discarded. Further, for each pixel contained in BPA and GPA surrounding pixels are checked and the pixels which are found black and gray in the new search are again stored in BPA and GPA respectively if they are not already present in them. This sequence is continued until all the pixels in the neighbourhood region of the
quasi centre are exhausted. This implies that all the black and gray pixels belonging to the connected region surrounding the quasi centre in the youngest daughter would be stored in BPA and GPA respectively at the end of this procedure.

**Back Projection**

To obtain the actual lumen region and its boundary, the pixels contained in BPA and GPA are projected back onto the images at prior levels. Each pixel contained in BPA is projected back onto the parent image and all the pixels in the parent image corresponding to that pixel are stored in an array named as Neighbourhood Pixel Array (NPA). Each pixel in GPA is projected back onto the immediate prior level. In this new search, if a pixel is black, it is mapped onto the parent image and all the corresponding pixels in the parent image are stored in NPA. If the pixel is white, it is discarded. In case, the pixel has a gray value, it is projected back onto the next immediate prior level again and a similar search is made on the four pixels on the new level. This process is continued until all the pixels contained in the GPA are exhausted. The pixels contained in GPA represent the final lumen region. The back projection procedure for \( P = 4 \) is illustrated in Fig. 1.

![Back projection procedure](image)

Fig. 1 Back projection procedure for \( P = 4 \)

To obtain the boundary of the lumen region, a rule base is used which gives a fairly reasonable estimate of the lumen boundary. The rule base consists of:

1. While mapping a gray pixel from the image on the 1\(^{st} \) level onto the parent image, all the white pixels contained in the parent image corresponding to that gray pixel are considered as boundary pixels.

2. For the pixel projections made onto the parent image of all the black pixels found in images at \((P-1)^{th}\) to 1\(^{st} \) level, their outer layer pixels are checked. If any pixel has a white pixel in its neighbourhood, it is considered as a boundary pixel.

The pixels that satisfy any one of the above two rules are included in a boundary pixel array.

**BOUNDARY THINNING**

Due to mucosal reflection, the lumen consists of small sub-regions having relatively higher intensities, which form white pockets embedded in the main lumen region in the binarised parent image. The boundary pixel identification algorithm described in Section 4.2 finds the boundary pixels for these pockets too, which are redundant and should be eliminated. In general, mathematical morphology provides a combination of erosion and dilation techniques for plugging the internal holes but it does not yield good results in the present application due to the large variation in the size of the pockets. In addition, the morphological procedure also creates an aberration in the actual boundary. The lumen boundary found using the quad structure based procedure is of non-uniform thickness and is discontinuous due to the strict boundary rule base. Most of the quantitative parameterisation and analysis techniques require a connected single pixel width boundary. This entails for further processing on the boundary obtained from Section 4.2. A knowledge-based technique of thinning the alphabet boundaries with the help of various templates is presented in Ahmed and Ward (9). But this technique does not give satisfactory results for the lumen boundary. In this section, we propose a simple procedure of thinning the boundary and subsequently to obtain a connected boundary.

**Thinning**

The thinning procedure is divided into four parts on the basis of the spatial location of the pixels. Initially, the boundary pixel locations having minimum and maximum \( y \) co-ordinates are segregated from the main boundary pixel array. The boundary thinning procedure starts with the top layer of boundary (pixels with minimum \( y \)) and among these, the pixel with the maximum \( x \) co-ordinate is chosen as the initial one. Then a preferential search is conducted using a sequence in the immediate neighbourhood of the initial pixel as shown in Fig. 2a. The numbers in the figure represent the preferential sequence in which neighbour pixels are searched. In the preferential search, if a boundary pixel is found in the neighbourhood, it is included into a new boundary array and all other pixels contained in that neighbourhood are eliminated. This elimination also ensures that there is no back propagation of the algorithm. In case, the immediate neighbour does not yield the next boundary pixel due to discontinuity in the boundary, the size of the neighbourhood is increased by one pixel. It is continued till the time next boundary pixel is found. Then the new boundary pixel is checked in the same preferential sequence. This procedure is continued until the pixel location having minimum \( x \) for the top row is
reached. After that a similar search is made using a preferential scheme given in Fig. 2b. This search is continued until the pixel location having maximum y is reached. In the same way a search is made to obtain a single pixel boundary from bottom row and bottom-up pixels using the preferential sequences given in Figs. 2c and 2d respectively. By this thinning procedure, the redundant boundary pixels embedded inside the lumen region due to internal reflection are also discarded.

![Fig. 2 Boundary thinning preferential sequence](image)

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

**Connecting**

Once a thinned single pixel width boundary is obtained, the gaps between consecutive boundary pixels are filled. The boundary is said to be connected if for every boundary pixel, there exists another boundary pixel in its immediate neighbourhood. To ensure connectivity, the two adjacent boundary pixels that do not satisfy the connectivity criterion are connected using an intermediate pixel optimisation scheme. Let the two consecutive disjoint pixels on the thinned boundary have the locations at \((x_i, y_i)\) and \((x_{i+1}, y_{i+1})\) respectively. The optimum locations \((x_{int}, y_{int})\) of the intermediate pixels can be obtained by

**Case 1:** If \(|x_{i+1} - x_i| \geq |y_{i+1} - y_i|\)

\[
y_{int} = \text{int} \left( y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i} (x_{int} - x_i) \right)
\]

where, \(x_{int}\) is an integer assuming all possible values in the closed interval \([x_i, x_{i+1}]\).

**Case 2:** If \(|x_{i+1} - x_i| < |y_{i+1} - y_i|\)

\[
x_{int} = \text{int} \left( x_i + \frac{x_{i+1} - x_i}{y_{i+1} - y_i} (y_{int} - y_i) \right)
\]

where \(y_{int}\) is an integer assuming all possible values in the closed interval \([y_i, y_{i+1}]\).

**RESULTS AND DISCUSSION**

An electronic endoscopy system, comprising a CCD camera with 200,000 pixel resolution and an RGB light source, was employed in capturing the image of the GI tract. The size of a typical captured image as shown in Fig. 3a was 256 × 256 pixels. The image was thresholded using progressive thresholding and the optimum threshold was found to be 23. The corresponding binarised image is shown in Fig. 3b. In the binarised image three discrete clusters of black pixels can be seen which constitute the qROI. In general, the largest cluster represents the GI lumen. The quasi center was computed as the centre of mass of the qROI, which was found to be at (57, 105). A pyramidal quad structure with \(P = 4\) was generated from the binarised image to facilitate fast computation of the lumen region and boundary without introducing ambiguous and false regions in the actual lumen. It can be seen that the image resolution decreases for each successive image level due to image compression. A region is grown in the youngest daughter using INS taking the corresponding quasi centre as the seed. The lumen boundary and the region obtained after back projection are shown in Fig. 4. The corresponding boundary length was found to be 739 pixels. The boundary obtained by the above procedure has many redundant pixels due to false internal boundaries. After thinning the boundary using the different templates on the basis of spatial location of the boundary pixels, a single pixel width boundary is obtained as shown in Fig. 5a. Then the boundary length was reduced to 371 pixels. But the thinned boundary is not a connected boundary. The intermediate missing pixels are calculated on the basis of optimum connecting scheme. The final single pixel width connected boundary of length 385 pixels is shown in Fig. 5b.

One of the main advantages of the above procedure is its high-speed response. The region is grown in the youngest daughter which is of considerably smaller size (1/64 times smaller in the present case) than the parent image, hence making the effective search area very small. But getting the original information back from the encoded image requires an intermediate process, referred to as back projection. The comparison between the INS technique directly applied to the parent image and that applied using a quad structure to find the lumen region and boundary is given in Table 1 for some typical cases. It can be seen that there is substantial improvement in the speed of the INS procedure. As the lumen area increases, the time reduction by using quad structure becomes more significant. During recent times, images with larger dimensions such as 1024 × 1024 pixels are used. In these cases, the new quad structure based algorithm will provide much faster response than the conventional INS procedure. This fast and reasonably accurate lumen region and boundary computation procedure facilitates online feature extraction and analysis.
CONCLUSION

A new technique for the extraction of lumen region and boundary from the GI images has been presented. The speed of INS procedure was enhanced significantly by using a quad structure, which makes the real-time extraction of the lumen feasible. This technique provides a connected single pixel width boundary by using a specific thinning and connecting algorithm. Investigations are being carried out to incorporate the cases when lumen is occluded or a bend is present in the GI tract. Further research is in progress to find a more accurate lumen region and boundary based on shock and perturbation analysis by using the boundary obtained by the proposed technique as input. The high-speed response of the proposed technique makes the concept of online microrobotic navigation possible and also provides an automatic solution to the problem of finding preliminary boundary for active contours.

REFERENCES


| Table 1 | Speed response of the quad based INS in comparison to the conventional methods |
|---------|-----------------|------------|------------|------------|
| INS on parent (mS) | Case 1 | Case 2 | Case 3 | Case 4 |
| 110 | 90 | 67 | 49 |
| INS through quad (mS) | 60 | 52 | 43 | 35 |
| Percentage reduction | 45.45 | 42.22 | 35.82 | 28.57 |
| Lumen area (pixels) | 22709 | 12178 | 6408 | 4947 |
| Perimeter (pixels) | 568 | 418 | 385 | 301 |