

ENHANCED AUTOMATIC COLON SEGMENTATION FOR BETTER CANCER DIAGNOSIS

Marwa Ismail¹, Aly A Farag¹, Robert Falk², and Gerald W Dryden³

¹University of Louisville, Louisville, USA

²Medical Imaging, Jewish Hospital, Abraham Flexner Way, Louisville, USA

³Division of Gastroenterology/ Hepatology, University of Louisville, USA

ABSTRACT

Colon segmentation is the first stage towards polyp detection, the main cause of colon cancer. Due to the immense importance of colon cancer diagnosis which is the second leading cause of death in the world, the segmentation phase must guarantee that no polyps are missed, especially the flat ones that are usually hard to detect. This work validates the 3D automated colon segmentation approach using the convex contour model previously proposed in literature. It also adds improvements to its pre-processing stage in order to better capture the colon walls and to enhance the results of the subsequent phases of the segmentation process. Experiments were conducted on 27 colon data sets that include 30 polyps. Moreover, 30 synthesized polyps with various shapes and sizes were placed at challenging areas of the colon's complex structure. Experiments conducted show a significant improvement in the construction of colon walls and the rate of polyp detection over that provided by the original technique.

Index Terms— *air-filled colon, fluid-filled colon, non-colonic attachments- polyps-convex snake model*

1. INTRODUCTION

Virtual Colonoscopy (VC) has become one of the most reliable techniques for polyp detection, which is the primary cause of colon cancer. It has a growing interest among physicians with colon cancer being the second leading cause of death in the world.

The first stage of a VC procedure is colon segmentation from abdominal CT scans that will significantly affect subsequent VC stages if it is not accurately employed. Poor colon segmentation would dramatically affect the rate of polyp detection, especially for flat ones that grow directly to the colon wall. Colon segmentation itself is a highly challenging problem due to many reasons including: 1) the presence of other structures in CT scans with the same intensity as of air-filled colon segments, such as lungs, and small bowels. 2) The tortuosity of the colon (caused by haustral folds). 3) The partial volume (PV) effect caused by residing of the contrast agent, the colon is injected by to insufflate it, in the lower concave parts of colon segments. An air-fluid boundary (AFB) region will then develop that will disconnect the colon into two parts, fluid-filled (similar to bone in intensity), and air-filled. 4) Poorly distended colon sets are collapsed at some areas, which will make false interpretation of colon segments as small bowels [1].

Colon segmentation has been widely addressed in literature. In [2], pre-processing algorithms, followed by 3D region growing were employed. In [3], a multistage approach was developed with two levels of classification based on vector quantization and region

growing. A post-processing method was proposed in [4] to repair the gaps in the segmented colon.

In this paper, we improve our work proposed in [5] for automatic colon segmentation that is based on the convex formulation of the active contour model. Pre-processing stage is of great importance, as it deals with the partial volume effect that might be the cause of missing significant colon parts if not well handled. We improve the pre-processing stage of the previously proposed algorithm and validate it with an extended number of real and synthetic polyps, which was not provided at our previous work. Results show improvement in the rate of polyp detection from the original algorithm, especially for flat ones that were missed before. The rate of polyp detection is up to $98 \pm 1.02\%$.

2. METHODS

Our work in [1] had 3 stages: 1) digital cleansing for detection of AFB and merging air and fluid colon parts. 2) air-filled parts segmentation based on [6]. 3) post processing to remove structures with the same intensity of that of the colon. In this paper, we enhance phase 1 of this framework for the sake of detecting all colon walls and better detection of polyps.

Figure 1 (a) shows a CT slice with PV that disconnects the colon into air-filled and fluid-filled segments. We used anatomical information in [5] to detect the AFB and its intensity was then set to that of air. Then all white structures (including fluid-filled colon parts) were set to the intensity of air as well. This resulted in setting the intensity of the bony areas (that have the same intensity of the fluid such as the spine, the ribs, and the pelvis) to the same intensity of the colon, to be all then discarded in the post processing stage. The main limitation of this cleansing procedure is that significant parts of the colon walls were missed, figure 1(b), along with the presence of gap-like artifacts, as they have the intensity between air and fluid which were neglected in this phase. This eventually led to missing polyps as will be shown later.

The threshold applied above was that of the fluid (T_{fluid}) which is high (255) and increases the chances of missing colon walls. If it is lowered to be between air and fluid in order to capture the walls, this will result in having non-colonic attachments that provide a lot of false positives and could not even be discarded in the post processing stage, figure 2 (b), (d).

The proposed change to the pre-processing phase in [5] will thus be as follows:

1-Detect all connected components (clusters) in the volume, denoted as C , along with the centroid CNT and the mean intensity I_{mean} of each.

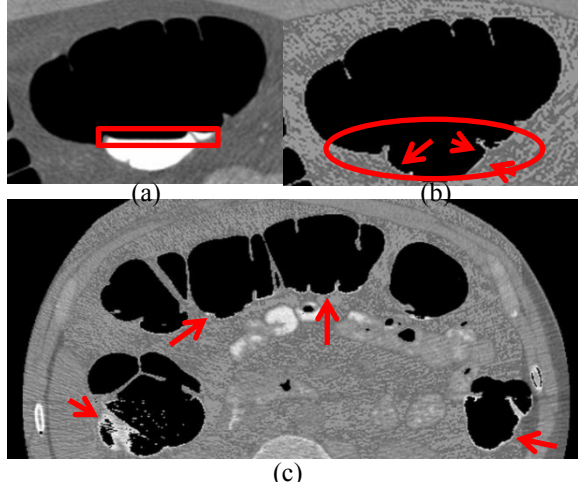


Fig. 1 a) AFB that is between air-filled and fluid-filled colon parts. b) results of digital cleansing from [1], where parts of the colon walls are missing. c) a CT slice cleansed with [1] and arrows show some boundary points that were not captured.

2-Get AFB from anatomical information using T_{air} and T_{fluid} and set it to air intensity.

3-Apply T_{fluid} for initial detection of fluid-filled parts, and set their intensities to T_{air} .

4-Apply an intermediate threshold intensity value T_{interm} that falls between T_{air} and T_{fluid} . T_{interm} should not be less than the smallest possible intensity a colon point might have. For each point satisfying this criterion, calculate the normalized distances $Dist$ between it and all the clusters' centroids.

5- Search within the N closest clusters to each point for its own C_i ; the one closest in distance. For a colon surface point, its C_i should have $I_{mean} = T_{air}$. If the search stops without finding any clusters, it assumes that this is not a colon point.

Small bowels are usually found with air intensity, yet leakage of the contrast agent from colon segments sometimes occurs and affects their intensity. Step 5 is significant in this case, because some colon points on the boundaries might be closer to a small bowel segment that has leaked contrast agent, but because of the intensity criterion, the algorithm keeps searching for the correct cluster.

We might also have the case that $T_{fluid} < I_{small\ bowel} < T_{air}$ for a small bowel with leakage, and thus get involved in step 5 for non-colonic points. Those candidates will have 2 scenarios: 1) the mean intensity of their closest cluster will not be zero (it is $I_{small\ bowel}$). 2) If this small bowel is closer to the centroid of a colon segment, only its boundary points will get affected by the average intensity of colon. Both scenarios lead to the results of significantly limiting applying T_{interm} directly on all the volume points that gets the non-colonic attachments as in figure 2.

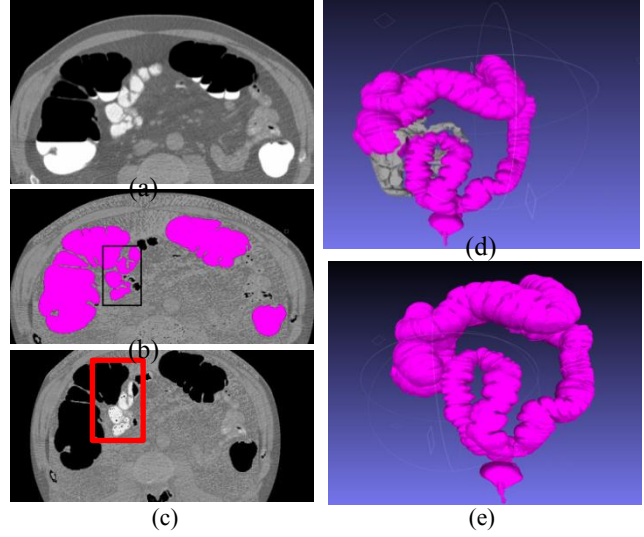


Fig. 2 a) CT colon slice. b) Result of Applying low threshold in efforts to capture all colon walls, but a small bowel gets attached to a colon segment. c) Applying T_{interm} that successfully captured colon walls and could discard the small bowel as well. d) 3D reconstructed colon of (b) with non-colonic attachments in gray. e) 3D reconstructed colon of (c) from the proposed enhancement.

It is worth mentioning here that T_{interm} can be readily selected at any value between 0, and 255 (air, and fluid) without the constraint of not considering relatively low values as in figure 2 (b). Also all threshold values selected were based on the average of intensity values for air and fluid from the 27 colon sets used. A pseudo code of the algorithm is present in Figure [3].

The core of the framework is then used to segment all air-filled parts. It is based on optimizing the following convex energy function:

$$E_1(u, c_1, c_2, \lambda) = TV_g(u) + \lambda \int r_1(x, c_1, c_2)u \, dx, (1)$$

where

$$r_1(x, c_1, c_2) = ((c_1 - f(x))^2 - (c_2 - f(x))^2)$$

f is the given image, λ is a positive parameter controlling the tradeoff between regularization process and fidelity of solution with respect to f and $c_1, c_2 \in R$. $TV_g(u)$ is the weighted total variation that convexifies the energy function. Please refer to our work in [5] for further details. In [5], the algorithm converges at 50 iterations, but due to the enhancements in the pre-processing stage, it got reduced to 10 iterations with accuracy up to 99%.

```

For all volume  $V(x, y, z)$ 
  Find all connected components  $C = [C_1, C_2, \dots, C_k]$ , their
  centroids  $CNT = [CNT_1, CNT_2, \dots, CNT_k]$ , and their mean
  intensities  $I_{mean} = [I_{mean_{C_1}}, I_{mean_{C_2}}, \dots, I_{mean_{C_k}}]$ .
End

For each point  $(x, y, z)$ 
  If  $T_{air} < I(x, y, z) < T_{fluid}$  &&  $I(x, y - 2, z) = T_{air}$  &&
 $I(x, y + 2, z) = T_{fluid}$ 
    Label the point as an AFB:  $I(x, y, z) = T_{air}$ ;
  End

  If  $I(x, y, z) = T_{fluid}$ 
    Label the point as a fluid-filled colon one, and set to air
    intensity:  $I(x, y, z) = T_{air}$ ;
  End

  If  $I(x, y, z) \geq T_{interm}$ 
    For each cluster  $C_i$ 

       $Dist(x, y, z) =$ 
      normalized Euclidean distance ( $V(x, y, z), CNT_i$ );
    End
    Sort  $C_k$  according to  $Dist$  from  $(x, y, z)$  :  $C_s =$ 
 $[C_{1s}, C_{2s}, \dots, C_{ks}]$ 
    For each  $C_{is} \in [C_{1s}, C_{ks}]$ 
      If  $I_{mean}(C_{is}) = T_{air}$ ;
         $I(x, y, z) = T_{air}$  and STOP the search.
      End
    End
  End
End
End

```

Fig.3 Pseudo code of the proposed pre-processing algorithm.

The average run time of the algorithm thus reduced from 5.5 ± 2 minutes to 3 ± 1 minutes, yet the modified cleansing procedure compensated that change.

Post-processing stage finalizes the process, by discarding all non-colonic structures and keeping only the colon. It mainly relies on the anatomical feature that the colon is the largest connected component in abdominal CT scans, but this condition will not suffice for the cases of poorly distended colons where further information is employed [5]. Modified pre-processing significantly improved final segmentation results, where non-colonic attachments are removed as we will show in the coming section.

3. VALIDATION AND RESULTS

The enhanced segmentation technique was tested on 27 colonography sets received from Walter Reed Army Medical Center, Washington, DC. The patients underwent standard 24-hour colonic preparation by oral administration of 90 ml of sodium phosphate and 10 mg of bisacodyl; then consumed 500 ml of barium for solid-stool tagging and 120 ml of Gastrografin to opacify luminal fluid. Each dataset contains ~500 slices. The sets included 30 polyps. 30 synthesized polyps with various shapes and sizes were also placed at challenging areas recommended by physicians [7].

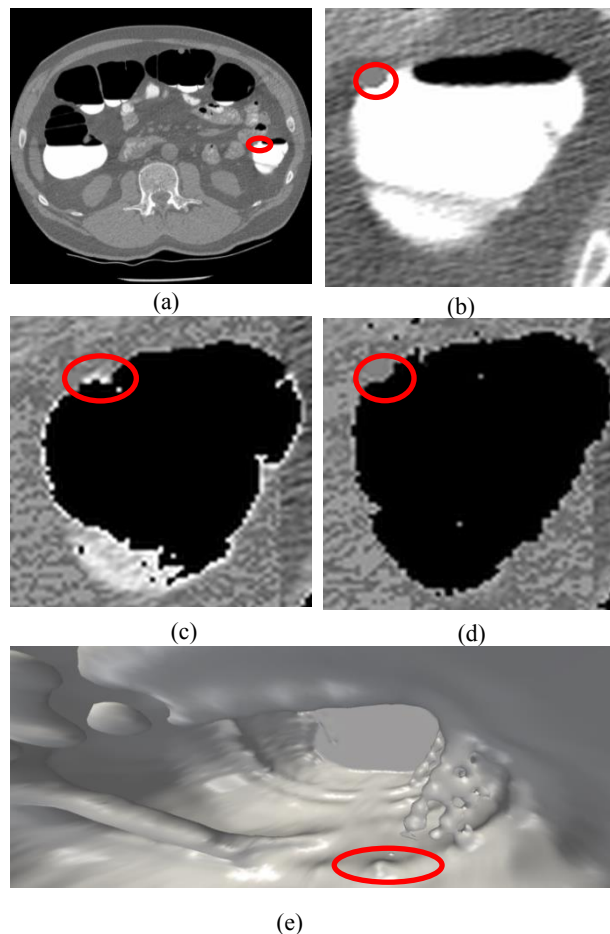


Fig.4 a) A synthesized 4 mm-polyp placed at a fluid-filled colon segment and zoomed in (b). c) The polyp was completely missed with the cleansing procedure in [1]. d) the polyp is detected with modified pre-processing algorithm, with its geometry preserved. e) Visualizing the polyp in (d) with fly-through technique.

Segmentation was first qualitatively validated by experienced physicians who read the results (from both the original algorithm and modified one) independently in a blinded manner and searched for polyps. Any suspected findings were to be confirmed by ground truth. The modified framework was able to detect flat polyps (<5mm) that were completely missed before due to poor construction of the walls, figure [4], [7]. Physicians also used our work in [7] to detect and visualize the findings and measure distortion factor of detected polyps [7].

Quantitative assessment was conducted by applying geometric properties such as shape index [SI], and sphericity (SP), [7], on all surface points. All points that satisfied the criteria ($SI > 0.7$, $SP < 1.2$) were clustered to identify polyps from segmented colon sets. This was done in addition to the other measures for the segmentation framework itself, Table [1].

Figure [5] shows more comparison results of the 2 segmentation frameworks highlighting the differences between original and enhanced techniques. Table [1] also shows comparison results that included the rate of polyp detection, sensitivity, specificity, and accuracy of the segmentation process.

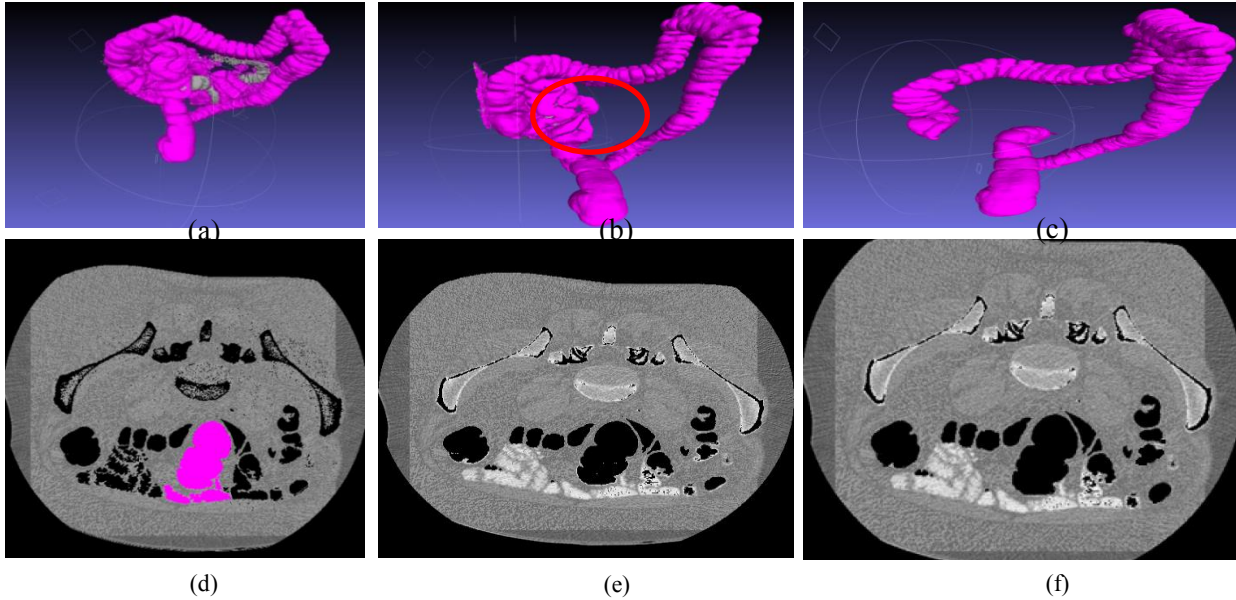


Fig. 5 a) 3D reconstructed colon with non-colonic attachments due to applying low threshold, and a corresponding 2D slice in (d) that shows the attachments. b) 3D poorly reconstructed colon due to applying T_{fluid} and its corresponding 2D slice in (e) where colon boundaries are missed. c) 3D reconstructed colon as a result of applying the proposed pre-processing and its corresponding slice in (f) where colon walls are captured and the small bowels are detached.

Table 1. Comparison results of original and modified segmentation approaches

		Original Segmentation Approach	Modified Segmentation Approach
% of Polyp Detection	Flat	50%	95%
	Pedunculated	95%	98%
Accuracy		94%	99%
Sensitivity		93.5%	98.4%
Specificity		92%	99.01%
Time (min.)		6-7	6-7

4. CONCLUSION

Colon segmentation is a crucial stage in a VC system. This work enhances our previously proposed convex model through its first stage of digital cleansing. It also provides extended validation of it on a wide number of polyps that are located at areas of the colon that are hard to capture from the segmentation phase.

5. REFERENCES

[1] J. Yao, M. Miller, M. Franazek, and R. Summers, "Colonic polyp segmentation in CT Colonography-based on fuzzy clustering and deformable models," IEEE TMI, vol. 23, no. 11, pp. 1344-1352, November 2004.
 [2] H. Li, P. Santago, "Automatic colon segmentation with dual scan CT Colonography," J. Dig. Imag., Vol. 18, 2005, pp. 42-54.

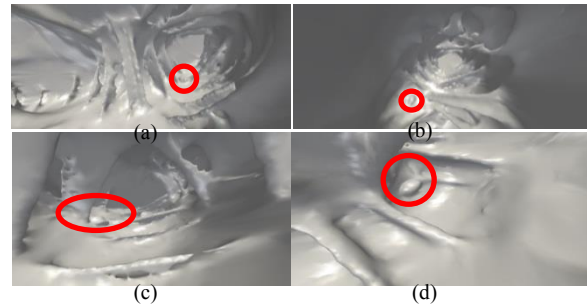


Fig. 6 (a) colon walls resulted from [5], where they are poorly constructed, and folds are missing. A polyp was detected, but with 60% distortion. Same polyp is zoomed in in (c). (b) Results of the enhanced framework where colon walls are better constructed and the polyp in (a) is detected with distortion <1%.

[3] D. Chen, Z. Liang, M. Wax, L. Li, B. Li, and A. Kaufman, "A novel approach to extract colon lumen from CT images for Virtual Colonoscopy," IEEE TMI, vol. 19, 2000, pp. 1220-1226.
 [4] L. Lu, Zhao, "An improved method of automatic colon segmentation for virtual colon unfolding," Computer Methods and Programs in Biomedicine, 2013.
 [5] M. Ismail, S. Elhabian, A. Farag, G. Dryden, and A. Seow. "Fully Automated 3D Colon Segmentation for Early Detection of Colorectal Cancer based on Convex Formulation of the Active Contour Model," CVPR MCV workshop, 2012.
 [6] X. Bresson, S. Esedoglu, P. Vanderghyest, J. Thiran, and S. Osher, "Fast Global Minimization of the Active Contour/ Snake Model," Journal of Math Imaging, vol. 28, 2007, pp. 151-167.
 [7] M. Ismail, A. Farag, M. Hassouna, G. Dryden, and R. Falk, "Improved Colon Navigation for Efficient Polyp Detection in Virtual Colonoscopy", fifth MICCAI Workshop on Abdominal Imaging, 2013.