POLYP DETECTION IN COLONOSCOPY VIDEO USING ELLIPTICAL SHAPE FEATURE

Sae Hwang[1]*, JungHwan Oh*[2]*, Wallapak Tavanapong*[3]*, Johnny Wong*[3]*, Piet C. de Groen*[4]

[1]Department of Computer Science and Engineering, UTA, Arlington, TX 76019
[2]Department of Computer Science and Engineering, UNT, Denton, TX 76203
[3]Computer Science Department, ISU, Ames, IA 50011
[4]Mayo Clinic College of Medicine, Rochester, MN 55905

ABSTRACT

Early detection of polyps and cancers is one of the most important goals of colonoscopy. Computer-based analysis of video files using texture features, as has been proposed for polyps of the stomach and colon, has two major limitations: this method uses a fixed size analysis window and relies heavily on a training set of images for accuracy. To overcome these limitations, we propose a new technique focusing on *shape* instead of texture in this paper. The proposed polyp region detection method is based on the elliptical shape that is common for nearly all small colon polyps.

Index Terms— Colonoscopy, Polyp Detection, Ellipse Fitting, Image Registration

1. INSTRUCTION

Early detection of polyps and cancers is one of the most important goals of colonoscopy. Recent evidence suggests that there are significant numbers of polyps and early cancers that are not detected at the time of colonoscopy [1]. Experts in endoscopy have hypothesized a number of reasons for these so called "missed" polyps and early cancers, but it is very difficult to find objective evidence explaining the occurrence of "missed" lesions since the tools able to objectively analyze what happens during colonoscopy are not available. During colonoscopy, a tiny video camera generates a video signal of the interior of the colon, which is displayed on a monitor for real time analysis by the physician. Computer-based analysis of video files using texture features, as has been proposed for polyps of the stomach and colon, has two major limitations: this method uses a fixed size analysis window, and relies heavily on a training set of images for accuracy. To overcome these limitations, we propose a new technique focusing on *shape* instead of texture in this paper. The proposed polyp region detection method is based on the elliptical shape that is common for nearly all small colon polyps. First, we determine whether elliptical shapes fit into segmented regions in a colonoscopy video frame by utilizing the watershed-based image segmentation and the ellipse fitting algorithms (Section 2). Then, we propose new techniques to distinguish the ellipses of polyp regions from those of non-polyp regions by matching curve direction, curvature, edge distance, and intensity (section 3). To improve accuracy, we develop a new method to detect polyp shots by utilizing the mutual-information-based image registration technique (Section 4). Some experimental results are reported (Section 5).

This research is partially supported by the NSF grants IIS-0513777, IIS-0513809, and IIS-0513582

2. ELLIPSE DETECTION

2.1. Region Segmentation

To generate an ellipse along a polyp boundary, we have to group the desirable edges. For this purpose, we segment an image into several regions. A typical polyp consists of various color values (or intensity values) that in part depend on the relative distance between the polyp surface and the light source. This distance varies because the shape of a polyp is 3D sphericity. Thus, a polyp could be recognized by its edge evidence not by the region uniformity. However, depending on the light condition and viewing position, only some parts of a polyp boundary have strong edge information, and others have weak edge information. Based on these observations, we use the marker-controlled watershed algorithm for polyp segmentation because it can handle the gap between broken edges properly, and place the boundaries at the most significant edges [2].

The watershed algorithm uses a topological relief function representing edge evidence as input. For the relief function, we use the gradient magnitude obtained using a set of the second derivative of the two-dimensional Gaussian (GD^2). The gradient magnitude (GM) along the orientation θ, $GM(x, y, \theta)$, is $GM(x, y, \theta) = \bar{I}(x, y) * G^2_{\sigma,\theta}(x, y)$, where $\bar{I}$ is a noise-reduced image using a median filter, and * represents convolution. The above equation is defined for the gray level images, and we extend it to color space since our colonoscopy images are in color. The gradient magnitude for a color image (GM_C) is $GM_C = \max(GM_R, GM_G, GM_B)$, where GM_R, GM_G, and GM_B are the gradient magnitudes for three

color bands (R, G, B), respectively. The watershed algorithm selects a small number of markers as the initial seeds. We select regions which have higher intensity values than those of surrounding regions as initial markers using the morphological operation called *regional maxima* because the shapes of polyps are 3D spherical or hemispherical forms so the geometric elevation level (i.e. light intensity) of a polyp is higher than that of the surrounding regions in colonoscopy images. Using the selected markers, the watershed transformation is performed. Figure 1 (a) is a gradient magnitude map. Figure 1 (b) shows the segmented regions obtained by the marker-controlled watershed algorithm.

2.2. Ellipse Fitting

Using the edges in each segmented region, we generate an ellipse using an ellipse fitting method. First, a binary edge map is constructed for each segmented region using the thresholding method as follows. Let p_i be a pixel of the gradient magnitude map (GM_c) belonging to a region i (R_i). Then the binary edge map (B_i) of the region i can be obtained by assigning '1' if p is larger than a certain threshold, otherwise '0' is assigned. For instance, Figure 1 (c) is a binary map for all regions. We use 1.0 for the threshold value.

An ellipse can be described by a second order polynomial [3] as follows:

$$F_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} \cdot \mathbf{a} = ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (1)$$

with an equality ellipse-specific constraint ($4ac - b^2 = 1$). Using the binary edge map of each region, we will find the best ellipse using the least square fitting method which was proposed in [3]. Figure 1 (d) shows the detected ellipses using the above algorithm. We note that ellipses are not detected for the regions 1 (R_1), 5 (R_5) and 6 (R_6) because there is not enough edge information in the binary edge maps corresponding to these regions.

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

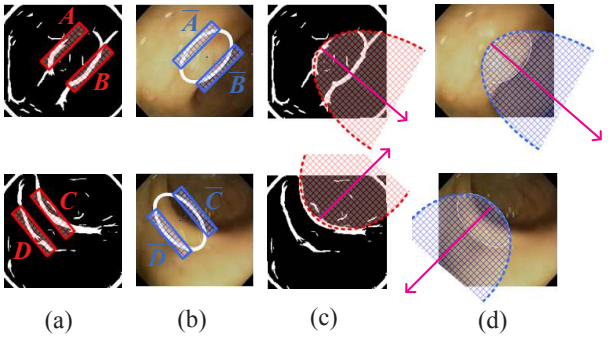
Fig. 1. (a) Gradient Magnitude (GM_c), (b) Segmented Regions, (c) Binary Edge Map, and (d) Detected Ellipses of All Regions

3. POLYP CANDIDATE SELECTION

Among the detected ellipses, we select the ellipses which have a possibility to represent actual polyps by removing the ellipses which do not represent actual polyps using three new filtering methods.

3.1. Filtering by Curve Direction and Curvature

Different edge shapes may generate similar ellipses. For instance, Figure 2 (a) shows two different edge maps in which the upper one is obtained from a polyp frame, and the lower one is obtained from a non-polyp frame. As seen in Figure 2 (b), two similar ellipses are generated from two different edge maps. Figure 2 (c) shows the best fitting parabolas obtained using the parts A and C of Figure 2 (a), which are indicated with red rectangles. Figure 2 (d) shows the best fitting parabolas obtained using the parts $\overline{A}$ and $\overline{C}$ of Figure 2 (b), which are indicated with blue rectangles. The arrows in Figure 2 (c) and (d) represent the direction of the parabolas. The basic idea distinguishing polyp ellipses from non-polyp ellipses is that if an ellipse is generated from a polyp, the direction of the parabola from any part of ellipse and the direction of the parabola from the corresponding part of edges are the same (the uppers of Figure (c) and (d)). In contrast, if there is a ceratin part in which the direction of the parabola from an ellipse and the direction of the parabola from the corresponding edges are different (the bottoms of Figure (c) and (d)), then the ellipse is not generated from a polyp.



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

Fig. 2. (a) Binary Edge Maps, (b) Detected Ellipses from (a), (c) Parabolas generated from Parts A and C in (a), and (d) Parabolas generated from $\overline{A}$ and $\overline{C}$ in (b)

Based on this observation, we propose to divide edges into four parts and calculate the curve information for each part as follows. Illustrated in Figure 3 (b), we divide an ellipse into four parts based on the two foci points (F_1 and F_2): (1)-upper side of the line between F_1 and F_2, (2)-right side of F_2, (3)-lower side of the line between F_1 and F_2, and (4)-left side of F_1. By selecting the edges in the corresponding parts, we can divide the edges into four parts. We call a part i of an ellipse as a *dismembered-ellipse* (E^i) shown in Figure 3 (c), and a part i of edges as a *dismembered-edge set* (B^i) shown in Figure 3 (d).

For each dismembered-edge set (B^i, $i = 1, \cdots, 4$) , we compute the curve direction and the maximum curvature by detecting a parabola using the polynomial curve fitting method. The second order polynomial of a parabola is the same as the second order polynomial of an ellipse (Equation (1)) with a

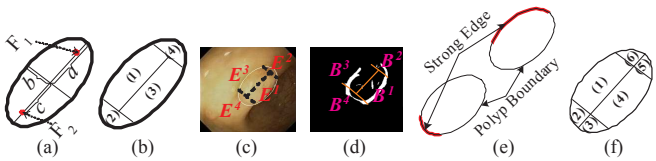


Fig. 3. (a) Ellipse, (b) Ellipse with Four Part, (c) Dismembered-Ellipse, (d) Dismembered-Edge Set, (e) Strong Edge Pattern of Polyp, and (f) Ellipse with Six Parts

different constraint ($b^2 - 4ac = 0$). It is known that the second order polynomial of a parabola cannot be solved using the least square fitting because the constraint of a parabola is $b^2 - 4ac = 0$. Therefore, we use another curve model for a parabola as follows: $f(x) = \alpha + \beta x + \gamma x^2$. However, this curve model works only if the directrix of a parabola is parallel to the x-axis. Thus, we define θ $(0 < \theta \leq \pi)$ as the counterclockwise angle from the x-axis to the major axis of an ellipse to rotate a dismembered-edge set (B^i) if B^i is not in the proper position. Based on θ, we can place each dismembered-edge set (B^i) by rotating each dismembered-edge set (B^i, $i = 1, \cdots, 4$) by $\theta + \frac{\pi(i-1)}{2}$. Figure 4 (b), (c), (d) and (e) show the rotated B^i by θ, $\theta + \frac{\pi}{2}$, $\theta + \frac{2\pi}{2}$ and $\theta + \frac{3\pi}{2}$, and the fitted parabola f^i for the corresponding B^i, respectively.

Fig. 4. (a) Original Edge, (b) Rotated by θ, (c) Rotated by $\theta + \frac{\pi}{2}$, (d) Rotated by $\theta + \frac{2\pi}{2}$, and (e) Rotated by $\theta + \frac{3\pi}{2}$

After rotating a dismembered-edge set (B^i), we can calculate the coefficients (α, β and γ) as follows. Given a set of pixels $(p(x_i, y_i), i = 1, \cdots, n)$ belonging to a dismembered-edge set (B^i), the coefficients of the second degree parabola can be obtained using the least square fitting method. By applying the above polynomial curve fitting method to four dismembered-edge sets (B^i, $i = 1, \cdots, 4$) , we can obtain up to four curves: f^1, f^2, f^3 and f^4. The curvature (K^i) at a point $t(x, y)$ which belongs to a curve f^i is $K^i(t) = \frac{\partial^2 y}{\partial x^2} / [1 + (\frac{\partial y}{\partial x})^2]^{\frac{3}{2}}$. The maximum curvature ($Kmax^i$) is the largest $K^i(t)$. Based on the coefficients of parabolas and the maximum curvatures, an ellipse (E) is declared as a polyp candidate if the ellipse satisfies both of the following two conditions. Otherwise, it is not a polyp and filtered out.

- Condition 1: If each dismembered-edge set is a part of a polyp, the coefficient γ^i of a parabola f^i should be larger than zero because the direction of f^i is turned up (see Figure 4). So, if there is a γ^i $(i = 1, \cdots, 4)$ which is less than or equal to zero, the ellipse is not a polyp candidate.
- Condition 2: If the maximum curvature $Kmax^i$ is in a certain range $(TH_{k1} \leq \max(Kmax^i | i = 1, \cdots, 4) \leq TH_{k2})$, then the ellipse is a polyp candidate. Otherwise, it is filtered out.

3.2. Filtering by Edge Distance

Even though the entire boundary of a polyp does not have strong edge information, some parts of polyp boundary must have strong edge information along the detected ellipse. Figure 3 (e) shows the typical patterns of strong edges of polyps in the colonoscopy image.

To characterize the above polyp edge patterns, we divide an ellipse into six parts as seen in Figure 3 (f) so we have six dismembered-ellipses ($E^i, i = 1, \cdot, 6$) and dismembered-edge sets ($B^i, i = 1, \cdot, 6$). By modifying the hausdorff distance [4], we define the edge distance (ED) as the sum of the distance between a dismembered ellipse (E^i) and the corresponding dismembered edge set (B^i) as follows:

$$ED^i = ED(E^i, B^i) = \sum_{a \in E^i} \min_{b \in B^i} d(a, b) \qquad (2)$$

where a and b are points of E^i and B^i respectively, and $d(a, b)$ is the Euclidian distance between a and b. ED measures how much a dismembered edge set is (dis)similar to the corresponding dismembered ellipse. The edge distance ED is asymmetric such as $ED(E^i, B^i) \neq ED(B^i, E^i)$, therefore, it can find if there are strong edges along the detected ellipse. Based on the edge distance (ED), an ellipse (E) is declared as a polyp candidate if either of the following conditions is satisfied. Otherwise, it is not a polyp.

- Condition 3: If there are strong edges close to an ellipse boundary in parts (2) and (3), or in parts (5) and (6), then the ellipse is a polyp candidate. This condition can be formulated as follows:

$$\min(ED^{(2,3)}, ED^{(5,6)}) \leq TH_{\zeta}$$

- Condition 4: If there are strong edges close to an ellipse boundary in parts (1) and (2), or in parts (1) and (6), or in parts (3) and (4), or in parts (4) and (5), then the ellipse is a polyp candidate. This condition can be formulated as follows:

$$\min(ED^{(1,2)}, ED^{(1,6)}, ED^{(3,4)}, ED^{(4,5)}) \leq TH_{\xi}$$

where $ED^{(u,v)} = ED^{(u)} + ED^{(v)}$.

3.3. Filtering by Intensity Value

Lumen areas are easily misclassified as polyps because they are elliptical shape and detected along with strong edges. Even though a lumen is similar to a polyp in shape, it is different from in color (intensity) because a lumen is relatively darker. Thus, we filter out an ellipse if its mode value is less than a certain threshold.

4. POLYP SHOT DETECTION

For a polyp candidate frame, its adjacent frame is registered based on the mutual information (MI) method [5]. The MI registration criterion states that the highest value of the MI can be obtained when the frame pair is geometrically aligned through a geometric transformation (T). We use the rigid body transformation as our geometric transformation (T) and use the simplex method [5] to maximize the MI measure under the rigid body transformation. We note that we convert the color images into the gray-level images before the image registration is performed. Figure 5 (a) is a polyp candidate frame (A), and Figure 5 (b) is an adjacent frame (B). Figure 5 (c) is obtained by registering the adjacent frame (B) into the polyp candidate frame (A). Figure 5 (d) and (e) are the corresponding binary edge map of Figure 5 (a) and (b), respectively. Figure 5 (f) is obtained by transforming Figure 5 (e) using the same parameters of Figure 5 (c)

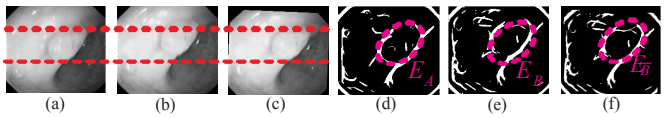


Fig. 5. (a) Polyp Candidate Frame (A), (b) Adjacent Frame (B), (c) Registered Adjacent Frame ($\overline{B}$), (d) Binary Edge Map of (a), (e) Binary Edge Map of (b), and (f) Registered Adjacent Binary Edge Map

After two frames (A and B) are registered, two edge sets (E_A and $E_{\overline{B}}$) are generated by selecting the edge pixels within the ellipse of polyp candidate frame (A) and the registered adjacent frame ($\overline{B}$). To examine if E_A and $E_{\overline{B}}$ have the similar edge pattern, we measure the distance ($Dist$) between E_A and $E_{\overline{B}}$ as follows:

$$Dist(E_A, E_{\overline{B}}) = \max(ED(E_A, E_{\overline{B}}), ED(E_{\overline{B}}, E_A)) \quad (3)$$

where ED is the edge distance which is defined in Section 3.2. If the $Dist(E_A, E_{\overline{B}})$ is less than a certain threshold TH_η, a polyp candidate frame(A) and its adjacent frame (B) have the same polyp. Otherwise, there is no polyp in the adjacent frame (B). A polyp shot is detected with four steps:
- Step 1: Let A_i be a polyp candidate frame i and A_j be its left adjacent frame ($j = i - 1$). The registered adjacent frame $\overline{A_j}$ is generated using the mutual information based image registration, and make two edge sets (E_{A_i} and $E_{\overline{A_j}}$) within an ellipse.
- Step 2: Measure a distance ($Dist$) between E_{A_i} and $E_{\overline{A_j}}$. If $Dist(E_{A_i}, E_{\overline{A_j}}) < TH_\eta$, set $i = i-1$ to replace A_i with A_j, and set $j = j - 1$ to select the left adjacent frame of A_j for new A_j. Using the new assigned A_i and A_j, repeat Step 1. If $Dist(E_{A_i}, E_{\overline{A_j}}) \geq TH_\eta$, the left-side boundary of a polyp shot is declared and move to Step 3.
- Step 3: Repeat the same procedure in Step 1 and Step 2 to detect the right-side boundary of a polyp shot with the different adjacent frame (i.e. A_j, ($j = i + 1$)).
- Step 4: Count the number of polyp candidate frames in a shot. If the number of the polyp candidate frames is larger than a certain threshold (TH_τ), the shot is declared as a polyp shot.

5. EXPERIMENTAL RESULT

For our experiment, we extracted the entire frames of the colonoscopy video with a 15 frames-per-second rate. The duration of the colonoscopy video is about 10 minutes so we have 8621 frames which consists of 815 polyp frames and 7806 normal frames.

Table 1 shows the number of total polyp shots, the number of correctly-detected polyp shots, the number of falsely-detected polyp shots, and the number of missed polyp shots. Falsely-detected polyp shots represents the shots which do not have any actual polyp but detected as having polyps by our algorithm. Missed polyp shots are the shots of actual polyps that were not detected. Among 27 polyp shots, only one shot is missed and 5 incorrect shots are detected.

Table 1. Result of Polyp Shot Detection

# of Total Polyp Shots	27
# of Correctly-detected Polyp Shots	26
# of Falsely-detected Polyp Shots	5
# of Missed Polyp Shots	1

6. REFERENCES

[1] D. Lieberman. Editorial, "Quality and colonoscopy: a new imperative," *Gastrointestinal endoscopy*, vol. 61, no. 2, 2005.

[2] A. Bieniek and A. Moga, "An efficient watershed algorithm based on connected components," *Pattern Recognition*, vol. 33, no. 6, pp. 907–916, June 2000.

[3] R Halir and J Flusser, "Numerically Stable Direct Least Squares Fitting of Ellipses," in *Int. Confonference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media*, 1998, pp. 125–132.

[4] G. Klanderman and W. Rucklidge, "Comparing Images Using the Hausdorff Distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850–863, 1993.

[5] H. Chen, P. K. Varshney, and M.A. Slamani, "On Registration of Regions of Interest (ROI) in Video Sequences," in *Proc. of IEEE International Conference on Advanced Video and Signal Based Surveillance*, Miami, FL, July 2003, pp. 313–318.