Reduction of false positives by internal features for polyp detection in CT-based virtual colonoscopy

Zigang Wang
Department of Radiology, State University of New York, Stony Brook, New York 11794

Zhengrong Liang[a)]
Departments of Radiology, Computer Science, and Physics and Astronomy, State University of New York, Stony Brook, New York 11794

Lihong Li
Department of Radiology, State University of New York, Stony Brook, New York 11794, and Department of Engineering Science and Physics, College of Staten Island of the City University of New York, New York 10314

Xiang Li
Department of Radiology, State University of New York, Stony Brook, New York 11794, and Department of Radiation Oncology, Columbia University, New York, New York 10027

Bin Li
Department of Radiology, State University of New York, Stony Brook, New York 11794

Joseph Anderson
Department of Gastroenterology, State University of New York, Stony Brook, New York 11794

Donald Harrington
Department of Radiology, State University of New York, Stony Brook, New York 11794

(Received 14 March 2005; revised 9 August 2005; accepted for publication 15 September 2005; published 16 November 2005)

In this paper, we present a computer-aided detection (CAD) method to extract and use internal features to reduce false positive (FP) rate generated by surface-based measures on the inner colon wall in computed tomographic (CT) colonography. Firstly, a new shape description *global curvature*, which can provide an overall shape description of the colon wall, is introduced to improve the detection of suspicious patches on the colon wall whose geometrical features are similar to that of the colonic polyps. By a ray-driven edge finder, the volume of each detected patch is extracted as a fitted ellipsoid model. Within the ellipsoid model, CT image density distribution is analyzed. Three types of (geometrical, morphological, and textural) internal features are extracted and applied to eliminate the FPs from the detected patches. The presented CAD method was tested by a total of 153 patient datasets in which 45 patients were found with 61 polyps of sizes 4–30 mm by optical colonoscopy. For a 100% detection sensitivity (on polyps), the presented CAD method had an average FPs of 2.68 per patient dataset and eliminated 93.1% of FPs generated by the surface-based measures. The presented CAD method was also evaluated by different polyp sizes. For polyp sizes of 10–30 mm, the method achieved mean number of FPs per dataset of 2.0 with 100% sensitivity. For polyp sizes of 4–10 mm, the method achieved 3.44 FP per dataset with 100% sensitivity. [DOI: 10.1118/1.2122447]

Key words: CT-based virtual colonoscopy, colonic polyps, computer-aided detection, shape index, texture feature

I. INTRODUCTION

Colonic polyps have a high probability of greater than 90% developing into colon cancer, which is the third most common human malignancy and the second leading cause of cancer-related deaths in the United States in 2004.[1] The early detection and removal of the polyps will dramatically reduce the risk of death.[2] Currently available detection methods consist of the fecal occult blood test, sigmoidoscopy, barium enema, and fiber optic colonoscopy (OC), where OC is currently the gold standard. Computed tomographic colonography (CTC) or CT-based virtual colonoscopy (VC) is an emerging method for the polyp detection. VC utilizes advanced medical imaging and computer technologies to mimic the OC procedure, looking for polyps by navigating through a virtual colon-lumen model that is constructed from the patient abdominal images.[3–8] Research findings show that both VC and OC share a similar procedure of high accuracy for the detection of clinically significant polyps.[9] Compared to OC, VC has shown the potential to become a mass screening modality in terms of safety, cost, and patient compliance.[9–12] Thus, the positive findings by VC screening will be examined by OC follow-up.[13]

Although it has several advantages as a minimally invasive screening modality, VC is a time-consuming procedure.

Even with a well-polished commercial VC navigation system (Viatronix, Inc., Stony Brook, NY),[9,14] it takes more than 15 min for a trained radiologist to mimic both forward and backward navigations of the OC procedure. The time can be longer if some suspicious locations need more attention. The anticipated large amount interpretation effort in VC screening procedure demands a computer-aided detection (CAD) scheme. A CAD scheme that automatically detects potential polyp candidates could substantially reduce the radiologists' interpretation time and increase their diagnostic performance with reduced false positives (FPs) and false negatives.

Applying CAD concepts for the detection of colonic polyps is a very challenging task, because the polyps can have various sizes and shapes. Furthermore, there are numerous colon folds and residual leftover colonic materials on the colon wall that mimic the polyps. Therefore, CAD should have the ability to eliminate the FPs and identify the most likely candidates or the true polyps. Up to now, there have been several CAD methods reported in the literature with variable success for polyp detection. Vining *et al.*[15] utilized the measure of abnormal colon-wall thickness to detect polyp suspects. Summers *et al.*[16,17] employed many of the geometrical features on the colon wall inner surface, such as the mean, Gaussian, and principal curvatures, etc., to find the polyp candidates. Yoshida *et al.*[18–20] and Nappi *et al.*[21,22] further characterized the curvature measures by shape index and curvedness to distinguish the polyp candidates from the normal colon wall tissues. Paik *et al.*[23–25] and Kiss *et al.*[26,27] presented another solution for the polyp detection where they utilized the normal (rather than the curvature) and sphere fitting as the references to extract some geometrical features on the polyp surfaces. By analyzing the shape change of the colon wall, polyp candidates can be detected efficiently using all these CAD methods. Since the traditional shape descriptions (such as the curvature- and normal-based shape measures) are sensitive to the irregularity of the colon wall, these surface shape-based CAD methods share a relatively high FP rate. Great efforts are needed to reduce the high FP rate.

Gorturk *et al.*[28–30] explored a pattern recognition method for the polyp detection, which utilizes the orthogonal triple planes to extract some features from the polyp candidates and then applies three-dimensional (3-D) pattern recognition on the features to eliminate FPs. Acar *et al.*[31,32] proposed a CAD scheme, which uses an edge-displacement field to analyze and improve the polyp detection. Nappi *et al.*[33] employed a conditional morphological dilation strategy to extract the suspects region. By analyzing three features from the region, they claimed the reduction of FPs. Recently Yao *et al.*[34] explored image segmentation methods to reduce the FPs.

In this paper, we propose an alternative CAD scheme to reduce the FP rate. From each of the detected suspects on the colon wall inner surface by the surface shape-based geometrical measures, we search its inner border beyond the surface inside the colon wall. An ellipsoid model is fitted to both the detected portion inside the lumen and the searched inner border inside the wall. Within the ellipsoid model, geometrical and CT image density (morphological and textural) features are extracted for FP reduction.

The contents of this paper are organized as follows. In Sec. II we present the proposed CAD scheme, i.e., finding initial candidates, fitting an ellipsoid model, extracting internal features, and eliminating FPs. All these four stages are described in Secs. II A–II E, respectively. Validation on patient datasets is reported in Sec. III, followed by a discussion and conclusions in Sec. IV.

II. METHODS

A. Global curvature

Up until now, principal curvature and corresponding curvature measures, such as the mean curvature[17] and Gaussian curvature,[18,35,36] are widely investigated for polyp detection. Since the curvatures reflect the shape "tendency" or trend among local neighborhoods, they are very sensitive to the shape change of the iso-surface at a given position.[37] Therefore, curvature-based shape measures can detect any specific shape-based section of the colon wall efficiently.[16–22] However, the locality property of the curvatures will sometimes mislead the shape detection due to noise and other distortions, resulting in a high FP rate. In order to overcome this limitation of the curvatures, Wijk *et al.*[38] presented a more accurate calculation method of the curvatures to reduce the noise-induced artifacts. In this paper, we introduce a smoothed principal curvature, which is based on the Gaussian curvature, to reflect a more general "tendency" or trend, which can provide an overall shape description of a wider surrounding region. For presentation purposes, the traditional Gaussian curvature is referred to as the "local curvature" and its associated direction is called the "local principal direction," while the smoothed curvature is called the "global curvature" in the following.

Given a nonumbilic point x_0 in a segmented 3-D colon mucosa layer,[39,40] there exist two orthogonal local principal directions. Along each local principal direction, a 3-D *convolution curve* from point x_0 is generated. A convolution curve l_c is defined as a curve starting from point x_0 and going both forward and backward in the 3-D principal direction field. For each point x_n on l_c, the gradient direction of l_c at x_n is parallel to the local CT density-based principal direction at x_n. The curvature of l_c at x_n is equal to the corresponding local CT density-based principal curvature at x_n. The concept of a "convolution line" was first introduced by Cabral *et al.*[41] in the LIC (Line Integral Convolution) algorithm for rendering of a 2-D vector field. We adapt this concept for the generation of a 3-D convolution curve.

Along each (a total of two) convolution curve starting from x_0, a smoothed or global curvature C^{new} is calculated by a convolution along this convolution curve:

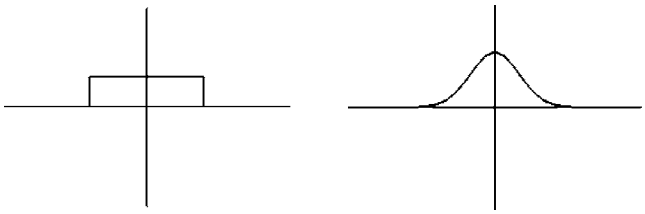

FIG. 1. An illustration of two different kernel functions—Uniform kernel (left) and Gaussian kernel (right).

$$C^{\text{new}} = \frac{\int_{x=x_0-L}^{x_0+L} k(x)\langle g_0 \cdot g_x\rangle C_x \, dx}{\int_{x=x_0-L}^{x_0+L} k(x)\langle g_0 \cdot g_x\rangle dx}, \tag{1}$$

where L is a half-curve length of the convolution curve, $k(x)$ represents the convolution kernel function, g_x is the gradient vector at point x, g_0 is the gradient vector at point x_0, C_x represents the corresponding local curvature at point x, and $\langle . \rangle$ indicates the inner product of two vectors.

The convolution kernel function plays an important role in the generation of the global curvature. By different convolution kernel functions, the global curvature can provide different shape information for different purposes. In this paper, we introduce and discuss two typical kernel functions: the uniform kernel function and the Gaussian kernel function, as shown in Fig. 1.

The uniform kernel function is a simple and widely used convolution kernel function. This kernel function has one parameter: the line length. With a short line length, the uniform kernel is usually more suitable for the detection of small polyps than with a long line length. With a longer line length, the global curvature with uniform kernel is less sensitive to the shape change of the colon wall and, therefore, more suitable to the detection of larger polyps, but it may overlook smaller polyps. Given a polyp size threshold, an appropriate line length can be determined.

Similar to the uniform kernel function, the Gaussian kernel function is also controlled by one parameter: the alpha value. The unique property of the Gaussian kernel is its capability to retain some of the "original" shape information. Comparing to the uniform kernel, the global curvature using the Gaussian kernel can retain more detectable shape information of small polyps that are helpful for the detection of the small polyps. However, the remaining too many shape details in the global curvature may lower down the efficiency of a CAD method.

Equation (1) defines the *global curvature* along the corresponding principal direction. For each voxel in the segmented colon mucosa layer,[39,40] there exist two global curvatures along the two principal directions, respectively. Applying these two global curvatures to the curvature-based measures, such as shape index, curvedness, sphericity rate, *etc.*,[17,18] we obtain the corresponding global curvature-based shape measures.

B. Finding initial candidates by curvature-based measures

As assumed in the previous CAD reports,[16,18,19,21,26,27] the colonic polyps usually have an "elliptic curvature of the peak subtype," i.e., the shape at the top section of a regular polyp (toward the colon wall) is more likely a "spherical cup" or "trough" shape. Correspondingly, the local shape-index values of the image voxels will increase smoothly from the top section to the bottom section of the polyp on the colon wall inner surface.[18–22]

For some irregular polyps without smooth surface, the shape-index values vary from the top to the bottom sections in a significantly unsmooth manner as compared to that of the regular polyps. It is difficult, based only on the local geometrical shape information,[16,18,26] to pick a complete protuberance section from the colon wall. By including a modified shape-index measure,[42] which is derived from a smoothed version of the local curvatures, as described above, the difficulty can be mitigated, i.e., a complete protuberance section of an irregular polyp candidate can be detected. Based on both the traditional and the modified local shape-index measures, a clustering algorithm is developed to find suspicious areas or patches on the segmented colon mucosa layer.[39,40] It is a growing-and-merging algorithm. Taking the advantage of space connectivity of the voxels, it clusters all the concerned voxels into several groups, as detailed below.

Prior to the clustering step, the colon lumen is first segmented and the mucosa layer is extracted.[39,40] Then all voxels in the mucosa layer are labeled into nine basic classes according to their traditional and modified shape-index values. (Nine basic classes are sufficient to cover the whole range of the shape index values. Class 1 corresponds to the peak type and class 9 to the valley type.[18]) If one voxel is labeled into class i, where $i \in (1,9)$ is referred to as the class number of this voxel, then this voxel is called an *i-class* voxel. The clustering step for growing and merging obeys the following three rules.

(i) ***Rule 1***: A suspicious patch group starts to grow at an *i-class* voxel, where i is the smallest class number among the class numbers of all the voxels in that group.

(ii) ***Rule 2***: If an *i-class* voxel is clustered into a suspicious patch group, only its nonclustered adjacent voxels, whose class numbers are equal or greater than i but less than or equal to the *max_class* number, can be clustered into this group in the next clustering step, where the *max_class* number is chosen based on the polyp size threshold.

(iii) ***Rule 3***: If two suspicious patch groups meet each other in space, they can merge into a larger suspicious patch if they satisfy the following two criteria:

(1) The number of the bordering voxels between these two groups is not too small (e.g., not less than 10% of the total voxel number in that candidate).

(2) The maximum class number of the bordering voxels is close to the class number of one group's starting-growing voxel.

Rule 1 ensures that each suspicious patch would have a somewhat spherical top section. ***Rule 2*** ensures that each suspicious patch contains as many available voxels as possible under the *max_class* number threshold (which is corresponding to a shape index threshold). By ***Rule 3***, each final suspicious patch can contain the protuberance section as completely as possible.

The clustering algorithm is very sensitive in detecting small changes on the colon mucosa layer and usually generates over a hundred suspicious patches in a colon dataset, concurring with the previous reports.[16,18,26] In general, these suspicious patches can be classified into three basic categories: (1) the true polyps; (2) the patches due to "noise;" and (3) the patches due to colon folds and residual colonic materials. The patches due to "noise" occur because of the system scan protocol (such as limited number of x rays, finite spatial resolution, patient motion, etc). The patches due to colon folds and residual colonic materials occur because the folds and colonic residues mimic the polyps. Both the noise candidates and the mimicking suspicious patches are called misclassifications. In order to improve the classification, a series of simple filters are employed to remove the misclassifications.

By setting the clinically relevant colonic polyps (e.g., larger than 4 mm in diameter) as the threshold and because the suspicious patches due to noise usually have a smaller size or smaller spherical top section, our first detecting filter is stated as follows.

(1) ***Filter 1***: If the total surface area of a suspicious patch is smaller than a given threshold, this suspicious patch is a misclassification. If the ratio of areas of the continuous spherical top section by both the traditional and the modified local geometrical measures is smaller than a given threshold, this suspicious patch is a misclassification.

Since the sizes and spherical top sections of the mimicking candidates are somewhat similar to those of the true polyps, it is difficult to erase them by ***Filter 1***. To erase this kind of misclassification, we introduce a new measure of *General Shape (GS)*. Given a polyp candidate $B\{voxel_i|i=1\ldots|B|\}$, its GS is defined as

$$\mathrm{GS}=\frac{1}{2}-\frac{1}{\pi}\arctan\frac{K^{\mathrm{mean}}}{K^{\mathrm{differ}}},\quad \mathrm{GN}=\frac{\sum_{i=1}^{|B|}g_i\cdot|K_i^1+K_i^2|}{\sum_{i=1}^{|B|}|K_i^1+K_i^2|},$$

$$K^{\mathrm{mean}}=\sum_{i=1}^{|B|}(K_i^1+K_i^2)\cdot\langle g_i\cdot GN\rangle, \tag{2}$$

$$K^{\mathrm{differ}}=\sum_{i=1}^{|B|}(K_i^1-K_i^2)\cdot\langle g_i\cdot GN\rangle,$$

where g_i is the gradient at voxel i, K_i^1 and K_i^2 are the principal curvatures (with $K_i^1\geq K_i^2$), and $\langle.\rangle$ represents the inner product of two vectors.

If the local curvature definition (for K_i^1 and K_i^2) is used for Eq. (2), we obtain a local GS measure, which provides the information of what the candidate "looks like." If the smoothed curvature definition of Eq. (1) is used, a "global" GS measure is obtained, which gives an overall shape description of the candidate around its surroundings. Based on both the local and the global GS measures, our second detecting filter is stated below.

(2) ***Filter 2***: A classified suspicious patch, whose local and global GS measures do not reflect a spherical cup or trough shape, is a misclassification.

It is noted that both the traditional and the smoothed local curvatures have their complementary properties, as described above.[42] A typical polyp shall show some geometrical attributes by the local and global GS measures. Therefore, the combination of both the traditional and the modified local shape measures in these filters is expected to reduce the misclassifications. Both ***Filter 1*** and ***Filter 2*** are very simple and work very well.[42] Experimental studies in this work show that both ***Filter 1*** and ***Filter 2*** can eliminate up to 80% of misclassifications. All the suspicious patches survived these two filtering operations are called the initial candidates. A further reduction of the FPs in the initial candidates is achieved by a much more complex process of texture analysis in the following sections.

C. Constructing an ellipsoid model for each initial candidate

As shown by the previous CAD reports,[15–27] it is difficult to eliminate some polyp-like false suspects by the surface shape-based measures only. A further reduction needs more information beyond the colon wall inner surface. Improvement has been shown by the inner features beyond the surface.[28–34] Given the detected patch or a portion of a suspect inside the colon lumen, our goal is to find the inner border of the suspect inside the colon wall so that the whole 3D volume of the suspect can be extracted for a quantitative analysis.

1. Finding the inner border of each initial candidate

Based on the polyp pathology in general,[43] as shown in Fig. 2(a), and the assumption that the detected initial candidates have some kinds of "elliptical" volume shape, as shown in Figs. 2(b) and Fig. 2(c), we aim to construct an ellipsoid that matches the suspect volume as close as possible. Usually the whole border of the ellipsoid consists of two parts: the outer part in the colon lumen and the inner portion behind the colon wall inner surface. The outer border in the mucosa layer has been detected as the suspicious patch, as described above. The inner border lies between the suspect and its adjacent normal tissues, as shown in Fig. 2(b) and Fig. 2(c). It is very challenging to determine the inner border because of the small image contrast between the suspect and its surroundings. One way to construct an ellipsoid is to grow the detected outer portion into the wall until some thresholds are satisfied. Another way is to find the inner bor-

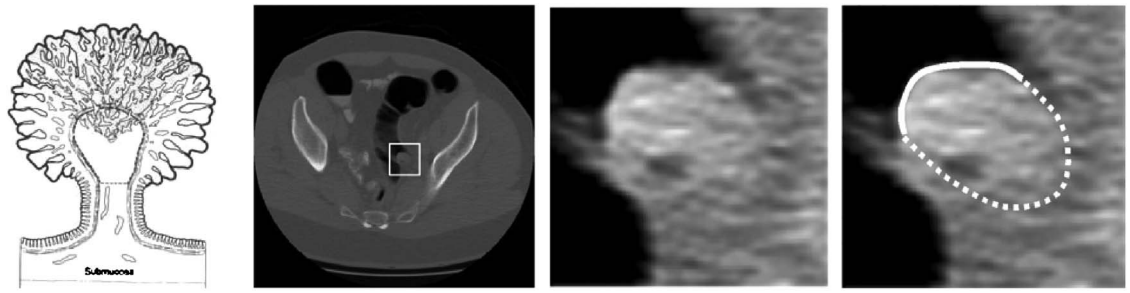

FIG. 2. (a)An illustration of a colonic polyp that is fully inside the lumen and has a typical elliptical shape. (b) In a CT image, the geometrical feature of a polyp (indicated by the white rectangle) becomes blurred. (c) The zoomed display of the polyp that shows some kinds of "elliptical" shape for both the detected portion inside the lumen and the inner portion inside the wall. (d) In a CT image, a polyp only shows part of its whole border inside the lumen. The solid curve indicates the outer border in the lumen; the dashed curve indicates the inner border behind the surface that cannot be detected by the shape analysis on the surface directly.

der points and fit the inner points together with the outer portion into an ellipsoid. We explore the later approach as follows.

Based on the 3-D convex ellipsoid model, a ray emitted from a point on the outer border will intersect with the inner border at least once in most cases. Taking advantage of this geometrical attribute of the border points, we propose a ray-driven technique to search for the inner border points in the CT image. Given a voxel v in an initial candidate, the image density gradient at that voxel is computed as (g_x^v, g_y^v, g_z^v). From this voxel, up to four rays are emitted whose directions are defined as [an example shown in Fig. 3(a)

$$\begin{aligned}
&\mathrm{Ray}_x = [-\mathrm{SIGN}(g_x^v), 0, 0],\\
&\mathrm{Ray}_y = [0, -\mathrm{SIGN}(g_y^v), 0],\\
&\mathrm{Ray}_z = (0, 0, -\mathrm{SIGN}(g_z^v)),\\
&\mathrm{Ray}_{\mathrm{grad}} = (-g_x^v, -g_y^v, -g_z^v)\\
&\text{where}\quad \mathrm{SIGN}(t) = \begin{cases} 1, & t > 0,\\ 0, & t = 0,\\ -1, & t < 0. \end{cases}
\end{aligned} \tag{3}$$

According to the elliptical geometrical attribute, there exists another border point along each ray. To identify this border point, we introduce a wavelet-based edge detector. First, a CT data profile along this ray is generated. By the Harr wavelet transformation[44] on the CT profile, a series of wavelet coefficients under different scales are extracted. In our method, the length of the CT data profile is chosen as 128 voxel units (to cover a relatively long range, ensuring the coverage of the inner border point), so the highest wavelet scale is 7.[44] After removing the high-scale (high-frequency) coefficients, e.g., 5 to 7, and performing inverse transformation, the original CT profile is transformed in a stepwise-like profile, as shown in Fig. 3(b) and Fig. 3(c). A more detailed operation between Fig. 3(b) and Fig. 3(c) is shown by Fig. 4.

We employ numbers 1 to 4 to represent the four-step status in the new profile, such as 1 represents the short plane, 2 represents a long plane, 3 represents a jump-up, and 4 represents a jump-down. Through a merge of the smaller steps, the whole profile is transformed into a number series; see Fig. 3(d). Since a typical border point has a specific variance pattern that can be represented by a number pattern, such as "423", "2413," and so on, it is easy to identify this pattern from the profile's number series. Usually the transformed profile can only provide an approximated location, instead of an exact position. We employ the first- and second-order derivatives of the original profile to identify the final position of the border point around the approximated location.

Not infrequently, due to image noise and other artifacts, some of the detected border points are not the real points on or near the inner border. To avoid these false border points, we define a search distance range for each ray, as shown in Fig. 3(e). This search range is defined quantitatively by the curvedness at the starting voxel v. Only those border points identified by the edge finder within this search range or curvedness can be treated as the inner border points. Figure 3(f) shows an example of the founded border points.

2. Constructing an ellipsoid model

Given the identified inner and outer border points, a 3-D ellipsoid region of interest (eROI) can be generated using the minimum algebraic distance fitting category[45] of the form

$$x^T A x + b^T x + c = 0, \quad A \in R^{3x3}, \quad x, b \in R^3, \quad c \in R \tag{4}$$

where the mathematical conventional notations have been used.

Figure 5 shows examples of constructing the ellipsoid model by Eq. (4) given the inner and outer border points. In most cases, a single solid curve can cover all outer border points of a polyp candidate through shape analysis on the detected patch. However, there are some cases where the whole outer border is divided into several separated parts due to noise and other artifacts, as shown in Fig. 5(b-1). These parts will likely lead to several different eROIs. To solve this problem, a "Merge" operation is employed. If two ellipsoids intersect each other and the intersecting region consists of at least 50% of the total volume in one ellipsoid, then all the outer border points and the inner border points of these two candidates will be merged. A new ellipsoid will then be generated using Eq. (4) for a new "merged' candidate, as shown

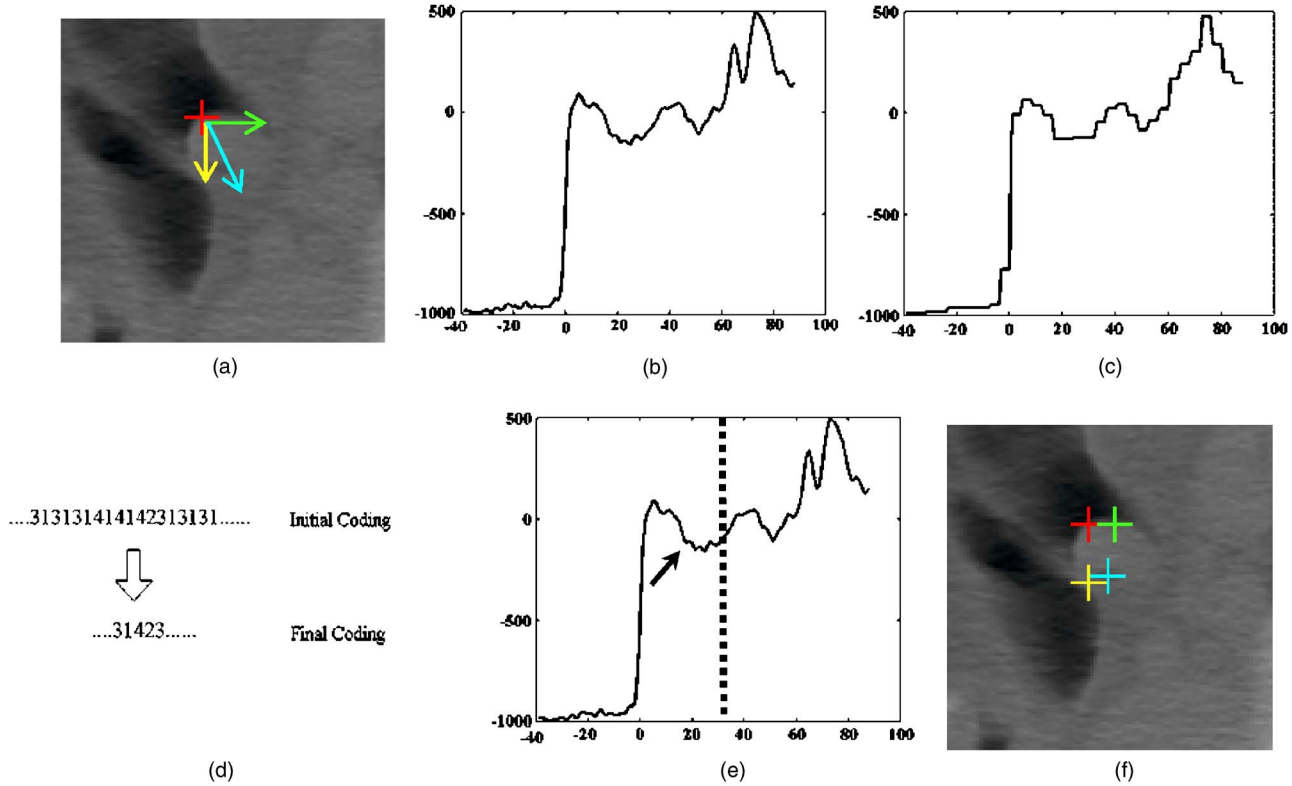

FIG. 3. (a) 2D illustration of three rays emitted from a given voxel (indicated by the cross) in a candidate. (b) The CT density profile along Ray_{grad}, i.e., the middle arrow toward the bottom right in (a). (c) The step-like profile after the Harr wavelet transformation and filtering. (d) The coding sequence of the step-like profile. The above is the initial coding sequence. After merging some small parts, we get the final coding as shown in the bottom. (e) Some special variance pattern can be detected using the coding sequence and the exact position is identified by the first- and second-order derivatives of the original profile. The detected point is indicated by the arrow. Meanwhile, we define a search range (indicated by the dotted vertical line) to reduce the possibility of finding a wrong edge. (f) The detected three new border points (the crosses) along the three rays, respectively. Among them, two are the outer border points and one is the inner border point.

in Fig. 5(b-3). Our experimental studies show that the ray-driven edge finder is a powerful tool to "dig out" the whole suspect volume from the surrounding tissues. It works very well even for some candidates whose outer borders are smaller than 4 mm, as shown in the bottom row of Fig. 5.

D. Extracting features from the constructed eROI

Based on the eROI model, a total of three types of features are extracted for a further reduction of the FPs in the detected initial candidates after the surface shape-based mea-

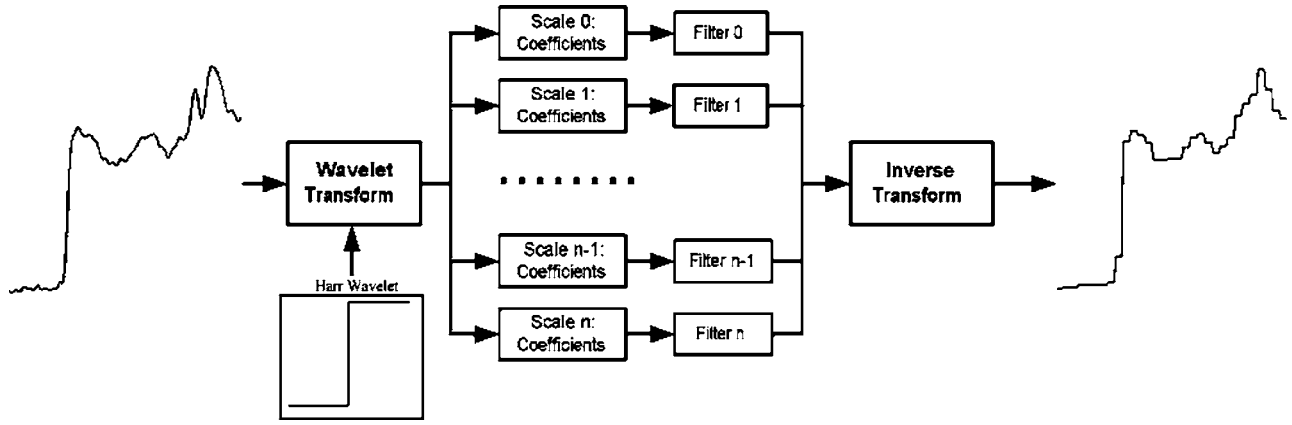

FIG. 4. An illustration of Harr wavelet transformation in the edge detector.

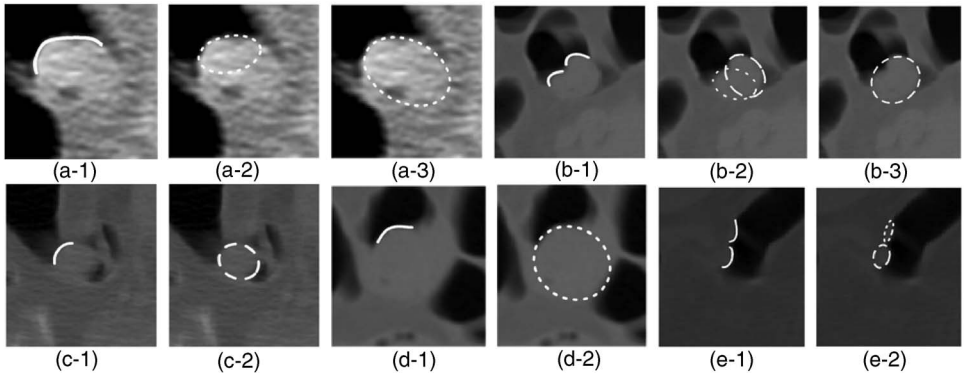

FIG. 5. (a-1) A polyp (6 mm in diameter) in a CT image. The solid curve indicates all the outer border points that are grouped into the initial candidate. (a-2) The ellipsoid (dashed line) constructed using only the outer border points in the left image. This ellipsoid may only cover a portion of the polyp. (a-3) The ellipsoid generated using both the outer and the inner border points. All the inner border points are detected using the ray-driven edge finder. (b-1) Another polyp candidate (10 mm in diameter) with an irregular surface. Its outer border points are grouped into two different candidates. (b-2) Two different ellipsoids are generated. Either of them covers the whole polyp volume. (b-3) A new ellipsoid generated is generated by merging the two candidates. (c) Another example for a candidate of 7 mm in diameter and its corresponding ellipsoid. (d) Another candidate of 15 mm in diameter that covers only a small section of the whole border of the polyp, and its ellipsoid generated from the small solid curve using the ray-driven edge finder. (e) An example of two polyp-like colon folds that cannot be eliminated by the shape analysis and become polyp candidates, resulting in two ellipsoids. However, due to the lack of inner border information, the upper ellipsoid is spindly. Although the lower ellipsoid shows a similar shape as a true polyp, it has much less inner border information than a true polyp has and, therefore, is likely to be eliminated by the texture information derived from the ellipsoid, as discussed in the next section.

sures or filters: geometrical, CT density distribution or texture, and morphological features. Each feature type is detailed in the following sections.

1. Geometrical features

A polyp has at least two typical geometrical attributes in the CT images, which are the shape change on the colon mucosa layer and the elliptical-like volume in 3-D space.[43] The shape change on the mucosa layer has been described before for the detection of the initial candidates. From a constructed eROI for each initial candidate, we extract two geometrical features: *Volume* and *Axis_Ratio*. Suppose the three radius of the eROI are $axis_1$, $axis_2$, and $axis_3$ (where $axis_1 >= axis_2 >= axis_3$), the definition of the *Volume* and *Axis_Ratio* are

$$Volume = \tfrac{4}{3}\pi \cdot axis_1 \cdot axis_2 \cdot axis_3,$$
$$Axis\_Ratio = axis_3/axis_1. \tag{5}$$

The *Volume* and *Axis Ratio* are two important geometrical features for a description of the whole shape of the eROI. Since our CAD scheme only focuses on the polyps with size greater than 4 mm in diameter, an eROI with too small *Volume* shows low probability to be a true polyp. The *Axis_Ratio* provides another shape description of the eROI. As noted by previous studies that a typical polyp may have a sphere-like shape,[16,18,26] although more polyps will deform the shape for many reasons.[43] However, the deformation may not change the shape dramatically. Therefore, it is highly likely a polyp will have a larger *Axis_Ratio* value, while the FPs from the colon folds and residue colonic materials will have a small *Axis_Ratio* value in their corresponding eROIs. Thus an eROI with larger *Axis_Ratio* indicates a higher probability of being a true polyp.

2. CT density distribution or texture features

Besides the eROI geometrical features, the CT density distribution within the eROI reflects another kind of important feature. A polyp shall have less image-density uniformity than the normal colon tissues. Furthermore, the image-density variation within the polyps shows a specific pattern,[9] which can be utilized as an indicator for polyp identification. In the following, we present a 3-D texture measure for the density variation pattern.

Due to the subtle change of CT density values from the polyp region to its neighborhood, it is desired to minimize the effect from the adjacent tissues. To that end, we enlarge and shrink the extracted eROI with a fixed scale to obtain two borders. (Enlarged Border and Shrunk Border, as shown in Fig. 6). Apparently, the derived density or texture features from voxels within the shrunk border would show more stability because of less effect from the adjacent tissues. In this study, all the texture features are derived from the voxels within the shrunk border.

Given a voxel v within the shrunk border of an eROI, three eigenvalues from its Hessian matrix[35,36] can be obtained. Without loss of generality, the three eigenvalues are λ_1, λ_2, and λ_3 (with $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$). For each pair of the eigenvalues (λ_i, λ_j), the corresponding pattern parameters $PA_{i,j}$ are calculated by

$$PA_{i,j} = -\frac{2}{\pi}\arctan\left(\frac{\lambda_i + \lambda_j}{|\lambda_i - \lambda_j|}\right), \quad \{i,j | i,j \in \{1,2,3\}, i \neq j\}. \tag{6}$$

Thus we obtain, for each voxel, a triple-element vector $\langle PA_{1,2}, PA_{1,3}, PA_{2,3}\rangle$ that represents the density variation pattern around that voxel. By plotting the triple-element vectors in 2-D/3-D space, we observe that the vector from each

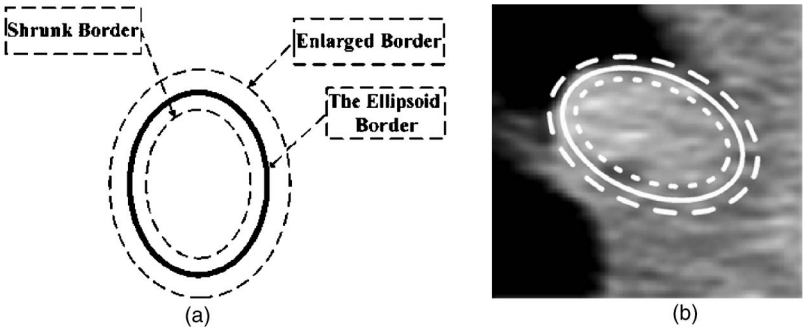

FIG. 6. (a) An illustration and definition of the enlarged and shrunk borders for an eROI. (b) The three borders in a CT image. The solid line represents the original border of the eROI; and the two dashed lines represent the shrunk border and the enlarged border of the eROI, respectively.

polyp voxel shows a different distribution pattern from that of a nonpolyp voxel, as shown in Fig. 7. The polyp voxels show a converging attribute toward the top right in the plots (by the circles), while the voxels of FPs from the colon folds and residue materials do not have this converging attribute (by the crosses).

Based on the observation,[9,42,46] the density values within a polyp change gradually and smoothly from the center to its border. This attribute is reflected by the convergence of the triple-element vectors toward the corner (1.0, 1.0, 1.0) in the 3-D presentation of Fig. 7(d). Based on the observed converging attribute, we introduce a texture feature of the

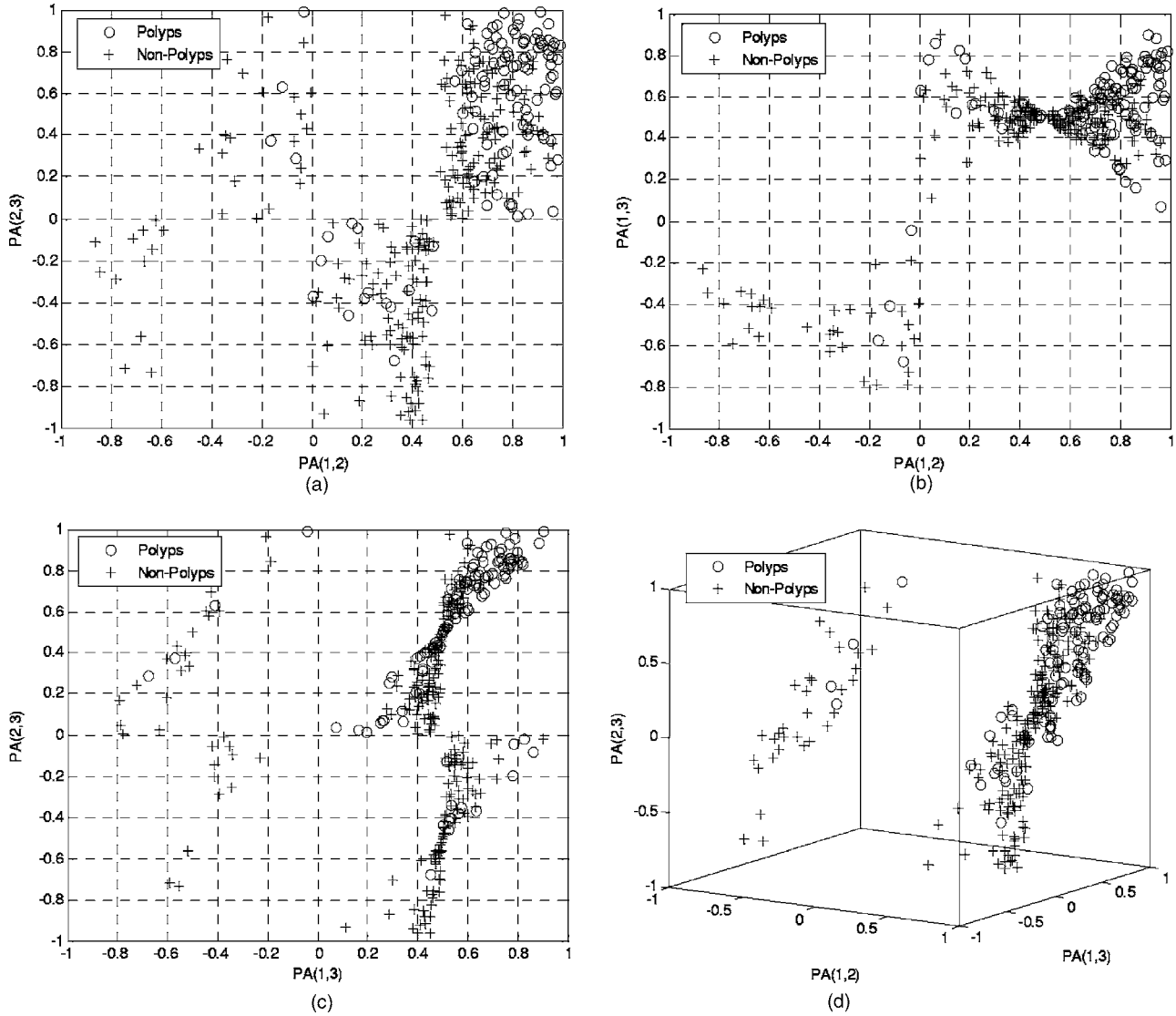

FIG. 7. The 2D/3D plots of the triple-element vectors of voxels from polyps and nonpolyps (or FPs). All voxels are selected by the computer randomly. Pictures (a)–(c) show the 2D plots of the triple-element vectors. Picture (d) shows the 3D plot. The triple-element vectors of polyp voxels show a convergence toward the top right corner (by the circles).

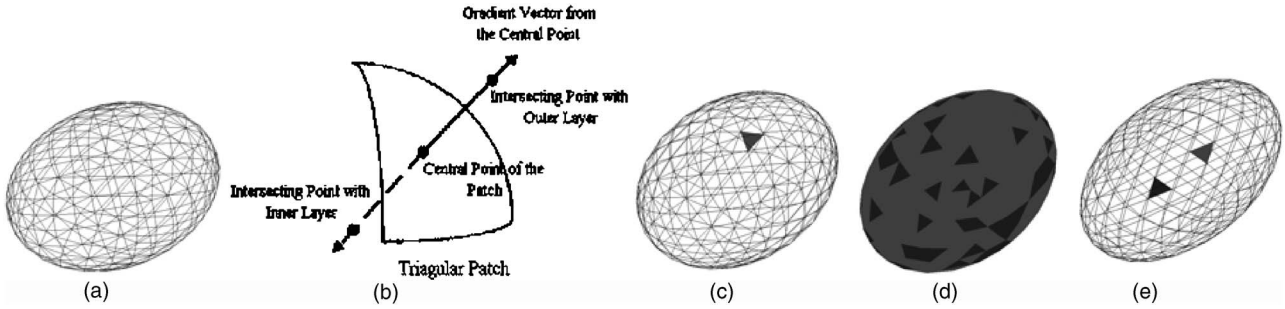

FIG. 8. An illustration of the mapping procedure of octsphere parametrization. Picture (a) shows the parametrization of the ellipsoid surface or border. The whole surface is divided into many triangular patches. In picture (b), a CT density profile is drawn, for each patch, along the gradient ray emitted from its central point. The search range is defined by the intersection points of the ray with the enlarged and shrunk borders. If there exists a border in the given search range, this patch is marked [or colored by the red in picture (c)]. Picture (d) shows the mapped image. The gray patches indicate all marked patches, and the black (darker) patches are the unmarked patches. Picture (e) shows a patch pair.

Growth_Ratio:

$$Growth\_Ratio_i = \frac{|S_i^g|}{|S_i|}, \tag{7}$$

where S_i={voxel $v\,|\,v$ is located within the shrunk border of eROI i}; S_i^g={voxel $v\,|\,v$ is located within the shrunk border of eROI i, and its triple-element vector is located at a 3-D boundary, as defined by, e.g., [0.5:1.0; 0.5:1.0; 0.5:1.0] in Fig. 7(d)}; and symbol $|\ |$ indicates the number of voxels in the set.

For a polyp candidate, the *Growth_Ratio* reflects the whole density distribution pattern within its eROI. If the *Growth_Ratio* reaches 1.0, the density variation pattern of this candidate indicates a good match to the typical pattern of the true polyps. The lower the *Growth_Ratio*, the less likely this candidate will be a true polyp. Besides the *Growth_Ratio*, the CT density–mean value may be another internal feature. Although the mean value cannot provide a detailed description of the density information, it reflects a feature to differentiate some FPs caused by the tagged or enhanced residue colonic materials by oral contrast solutions.

3. Morphological features

As discussed above, a typical polyp has a relatively complete border in the CT image. This border is resulted from the difference between polyp cells and the surrounding normal tissue cells.[43] In contrary, the colon folds and/or other normal colon tissues seldom show a relatively complete border due to the similarity between their CT densities. By this attribute, we introduce two morphological features to provide a quantitative measure of the border for each eROI.

First, the whole eROI border is divided into several regular patches by parametrization of the Octsphere.[47] For each patch on the eROI border, a ray crossing its center along the normal direction will intersect its shrunk and enlarged borders, respectively, as shown in Fig. 4, and Figs. 8(a) and 8(b). Similar to the ray-driven edge finder described before, a CT density profile along this ray is generated. If a border point is detected between the shrunk and enlarged borders, this patch is marked, as shown in Fig. 8(c).

Given a patch on the eROI border, there must be another patch where the line between these two patches' center points must cross the center of the eROI. These two patches are called a patch pair, as shown in Fig. 8(e). If two patches in a patch pair are both marked, this pair is called a marked patch pair. Given an eROI, let PP and PP_{pair} be the set including all patches and all patch pairs, respectively, two morphological features of this eROI are defined as

$$Coverage\_Ratio = \frac{|PP^{\text{marked}}|}{|PP|},$$
$$Radiation\_Ratio = \frac{|PP_{\text{pair}}^{\text{marked}}|}{|PP_{\text{pair}}|}, \tag{8}$$

where PP^{marked} and $PP_{\text{pair}}^{\text{market}}$ are in a marked patch set; and $|\ |$ indicates the number of voxels in the set.

The *Coverage_Ratio* provides a quantitative measure for the border coverage information of the eROI. An eROI with a larger *Coverage_Ratio* must have a more complete border. The *Radiation_Ratio* reflects mainly the border distribution information. For example, if an eROI only has a half-contiguous border, its *Radiation_Ratio* will be 0 while its *Coverage_Ratio* remains 50%.

E. Eliminating false positives by the extracted internal features from the eROI

In the presented CAD scheme above, there are in total six internal features extracted from each eROI: *Volume, Axis_Ratio, Growth_Ratio, Density_Mean, Coverage_Ratio* and *Radiation_Ratio*. Based on these features, we employ a two-level classifier to reduce the FPs in the initial candidates. This classifier consists of two levels. Each feature will go through a transformation function at the first level to enter a linear discrimination at the second level, as shown in Fig. 9. Among all these features, the *Axis_Ratio, Growth_Ratio, Coverage_Ratio*, and *Radiation_Ratio* are four "normalized" features, i.e., their feature values are normalized to the range of [0, 1] so that they can pass through the first level of the transformation function and directly go into the second level of the linear discrimination.

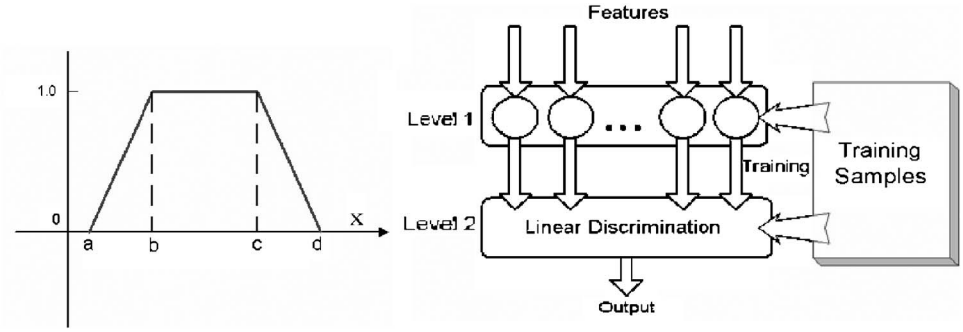

FIG. 9. Illustrations of a transformation function (left) and a two-level classifier for the final polyp identification (right).

In contrast, the *Volume* and *Density_Mean* features are two "non-normalized" features, whose transformation functions are specially designed as follows, as shown in Fig. 9 (left):

$$\phi_i(t) = \begin{cases} 0, & t \in (-\infty, a), \\ t-a/b-a, & t \in [a,b), \\ 1, & t \in [b,c], \\ d-t/d-c, & t \in (c,d], \\ 0, & t \in (d,+\infty). \end{cases} \tag{9}$$

Apparently the transformation function of Eq. (9) has four parameters to be determined for the *Volume* and *Density_Mean* features: a, b, c, and d. Two basic approaches can be taken to determine these parameters. One uses a computer learning or fitting strategy. By this strategy, the computer can automatically determine an optimal selection of these four parameters by training samples. The other is a manual approach. The user can find by error and trial or define the parameter values according to his/her experience and expectation. In this study, we employ the first approach for the *Density_Mean* feature and the second approach for the *Volume* feature. After the transformation, both the *Volume* and the *Density_Mean* features are "normalized" in the range [0, 1].

The classifier function for the six internal features in the linear discrimination can be written as follows:

$$F = \sum_i w_i \cdot \phi_i(f_i) + \eta, \tag{10}$$

where $\phi_i(\cdot)$ is the transformation function for feature f_i, w_i is a weight factor for this feature, η is a constant factor, and i indexes the features. For the four "normalized" features, $\phi_i(\cdot)=f_i$. The weight factors f $\{w_i\}$ and constant factor η for all the six internal features are determined by computer learning or fitting strategy using training datasets.

For each feature vector (i.e., the extracted six from an eROI) from a polyp candidate, the linear two-level classifier will output a likelihood or probability value F that is normalized between 0.0 and 1.0. The more close this value reaches to 1.0, the more likely this candidate will be a true polyp. Using a likelihood threshold, all the candidates can be classified and identified according to their likelihood values from the classifier.

III. EXPERIMENTAL RESULTS

The efficacy of the proposed CAD scheme was evaluated by a database of 153 patients' abdominal CT images, which were acquired using a single-slice spiral CT scanner (Hi-Speed CT/i, GE Medical Systems, Milwaukee, WI). Prior to obtaining the CT images, the patients underwent a one- or two-day bowel preparation of a low-residue diet and mild laxatives.[13] In order to enhance the CT density of the residual colonic materials, the patients also ingested three to four (depending on one- or two-day preparation) 250 cm^3 doses of 2.1% w/v barium sulfate suspension with meals before the CT procedure, as well as two doses of 60 cm^3 of gastroview (diatrizoate meglumine and diatrizoate sodium solution) given during the night before and the morning of the CT procedure. After the patients' colons were inflated with CO_2 or room air (2–3 L) given through a small rectal tube, the CT images were obtained using routine clinical CT protocols for the VC procedure. The imaging protocol parameters were 120 kVp, 100–200 mA (depending on body size), 512 $\times$512 array size for the field of view (FOV) (completely covering the body), 1.5–2.0:1.0 pitch, 5 mm collimation (completely covering the entire colon in a single breathhold), and 1 mm image reconstruction. The 5 mm collimation sets the upper resolution limitation. By a pitch in the range of [1.5, 2.0], the image resolution is limited to 4–5 mm.

For each patient, two CT images were acquired from the supine and prone positions, respectively. In the same day following the CT scans, an OC procedure was performed for the same patient. The position and size of each polyp found by the OC procedure were recorded as the ground truth or gold standard of the findings. There are a total of 45 patients with 61 OC verified polyps larger than 4 mm in diameter. Among these polyps, 50 polyps were measured with a size between 4 to 10 mm in diameter and the other 11 polyps were measured with a size larger than 10 mm. The largest one was measured with a size of 30 mm.

All the 153 patient datasets were first segmented and their corresponding colon lumens and mucosa layers were extracted.[39,40] By applying the three clustering rules on all image voxels in each mucosa layer with a *max_class* number of 3 [correspondingly the shape index is between 0.0 and 0.25 (Ref. 18] without missing true polyp or no false negative, more than 200 suspicious patches per dataset were de-

TABLE I. Experimental results from the 153 patient datasets after the surface shape-based filtering operation.

Suspicious patch number per dataset after the three clustering rules	Suspicious patch number per dataset after ***Filter 1*** and ***Filter 2***	Eliminated suspicious patch percentage by ***Filter 1*** and ***Filter 2***
221.8	42.9	80.6%

tected, as shown in Table I. The number of suspicious patches is too high. The suspicious patch number was dramatically reduced by the two surface shape-based filters. In ***Filter 1***, the smallest size of the patches was chosen as 15 mm^2, and the minimum sphere ratio of the traditional and the smoothed local curvature measures on the detected patches was 25% (ensuring no false negative). In ***Filter 2***, the maximum GS value was 0.25 for both the local and global GS measures (also ensuring no false negative). Up to an 80.6% reduction of the suspicious patches was achieved by the use of all the surface shape-based curvature measures without missing polyps of greater than 4 mm. All the remained suspicious patches are collected as our basic samples (or initial polyp candidates) for a further FP reduction by the morphological and internal density information.

A. FP reduction analysis

For each initial polyp candidate after the surface shape-based filtering, its eROI model was generated by the algorithm described in Sec. II B. Every eROI model was scaled by a factor of 0.7 for the shrunk border and 1.3 for the enlarged border (or an erosion and dilation operation). The boundary for the *Growth_Ratio* feature in three dimensions was defined by [0.5:1.0; 0.5:1.0; 0.5:1.0], as shown in Fig. 7(d). These choices for the scale factors and the 3-D boundary were based on the model construction process, as shown in Fig. 6 and Fig. 7. Among the extracted six features, the *Volume* and *Density_Mean* features underwent the transformation of Eq. (9) and normalization. For both the *Volume* feature and the *Density_Mean* feature, the parameter vector (a,b,c,d) was determined automatically by the training datasets below, and so were the weight factors $\{w_i\}$ and constant factor η for all the six features.

We trained and tested all the initial polyp candidates using the round-robin (leave-one-out) strategy.[21,22] For each patient dataset, all initial polyp candidates from other patient datasets were selected as the training samples. If an OC verified polyp (and location) was detected in either the supine or prone scan, the detection of this polyp was scored successfully. The experimental results are shown in Table II. For a polyp detection sensitivity of 100%, the presented CAD scheme had an average FP number of 2.68 per patient dataset and eliminated 93.1% of the initial polyp candidates. The mean number of FPs per polyp is nearly 2.8. We also analyzed the performance of the presented CAD scheme on various polyp sizes. For the detection of polyps larger than 10 mm in diameter, improved performance was observed, as expected. For 100% detection of all the larger polyps, the CAD scheme achieved an average FP number of 2.0 with an elimination rate of 93.5% on the initial polyp candidates. For 100% detection of polyps smaller than 10 mm in diameter, the CAD scheme still achieved an average FP number of 3.44 with an elimination rate of 92.8% on the initial polyp candidates.

In the presented CAD scheme, there are a total of six internal features extracted and used for the reduction of FPs. In order to compare the contribution of different features, we divided all six features into two groups: the Geometrical/Morphological (GM) group and the Intensity/Texture (IT) group. The GM group contains four features: *Volume*, *Axis_Ratio, Coverage_Ratio*, and *Radiation_Ratio*. And the IT group contains two features: *Growth_Ratio* and *Density_Mean*. According to our definition, the GM features provide the general shape description on the surface of eROI while the IT features provide the detailed image density description inside the eROI. We trained and tested each group by the 153 patients' datasets using the same parameters, as described before. The experimental results are shown in Table III.

For a more aggressive detection (sensitivity) rate, the increase of FP number and decrease of the eliminating rate are usually quantified by a free-response operating characteristic (FROC) analysis.[50] The FROC curve illustrates the tradeoff between the detection sensitivity for the true positives and the false positives. Therefore a FROC analysis was per-

TABLE II. Experimental results of the presented CAD scheme after using the internal features.

Size of polyps in (in diameter)	Lesion located fraction (by polyp)		
	100%	90%	80%
	Mean number of FP per dataset/Mean number of FP per polyp–Percentage of eliminated candidates		
All polyps (4 mm–3 cm)	2.68/2.82–93.1%	1.21/1.61–96.9%	0.58/0.90–98.5%
Small/median polyps (4 mm–10 mm)	3.44/3.10–92.8%	1.56/2.00–97.1%	1.00/1.14–97.9%
Large polyps (>10 mm)	2.00/2.43–93.5%	0.90/1.17–96.2%	0.20/0.50–99.2%

TABLE III. Experimental results of the presented CAD scheme using different feature groups.

	Lesion Located Fraction (by polyp)		
	100%	90%	80%
	Mean number of FP per dataset/Mean number of FP per polyp–Percentage of eliminated candidates		
All features used	2.68/2.82–93.1%	1.21/1.61–96.9%	0.58/0.90–98.5%
GM features only	6.05/6.47–84.5%	3.31/3.64–91.5%	1.37/1.90–96.5%
IT feature only	8.05/7.11–79.5%	5.73/4.76–85.3%	5.52/4.23–85.9%

formed to demonstrate the detection capability of the presented CAD scheme, as shown by Fig. 10. These curves in the figure reflect their relative performances.

B. Performance of the global curvature versus the local curvature

By Eq. (1), the difference between the local and global curvatures is shown by the difference of the convoluted curvatures between the starting point and its neighboring points on the same isodensity surface along the convolution curve. Comparing to the local curvature, the global curvature has several advantages. (1) Uniformity: the distribution of shape measures, such as the shape index, becomes more uniform using the global curvature. This uniformity of shape measures benefits the polyp detection. (2) Robustness: the global curvature is the convolution of the curvature from the neighborhood region and reflects a general trend of the shape change in the neighborhood region. So it is expected to be less sensitive to the noise as compared to the local curvature, while remaining the same shape information of the colonic polyps. (3) Flexibility: by different convolution kernel functions, the global shape measures can provide not only the exact shape information at a given position, but also the different "overall" shape description of its neighborhood region.

The performance of the global curvature versus the local curvature was measured in terms of detection of the initial polyp candidates using the database of 153 patient studies. From the database, 20 short colon segments, each containing an OC proved colonic polyp, were selected randomly as the true positive samples. Among these polyps, 13 polyps were measured with a size between 4 to 10 mm in diameter and the other 7 polyps were measured with a size larger than 10 mm. In addition, another 30 short colon segments, each containing a suspicious region, were selected randomly from the database as the FP samples. Some of the FP samples are the colon folds, and the others are the causes due to the noise and residuals in the colons. Among all the FP samples, 18 have a size smaller than 10 mm in diameter, and the others are larger than 10 mm. In this experiment, five different curvatures were compared: local curvature, two global curvatures using a uniform kernel function with a half-line length of 5 and 10 mm, respectively; and two global curvatures using a Gaussian kernel function with an alpha value of 10 (the half-line length is approximately 10 mm) and 20 (the half-line length is approximately 15 mm), respectively.

For each curvature from the five, a ROC (receiver operating characteristics) study was performed for the detection of the polyps from the 50 samples. The results are shown in Fig. 11(a). The global curvatures showed an equivalent performance with Gaussian-20 kernel function and some improved performance with other three kernels as compared to the local curvature. Both the local and global curvatures have their advantages and limitations. It is advantageous to use a

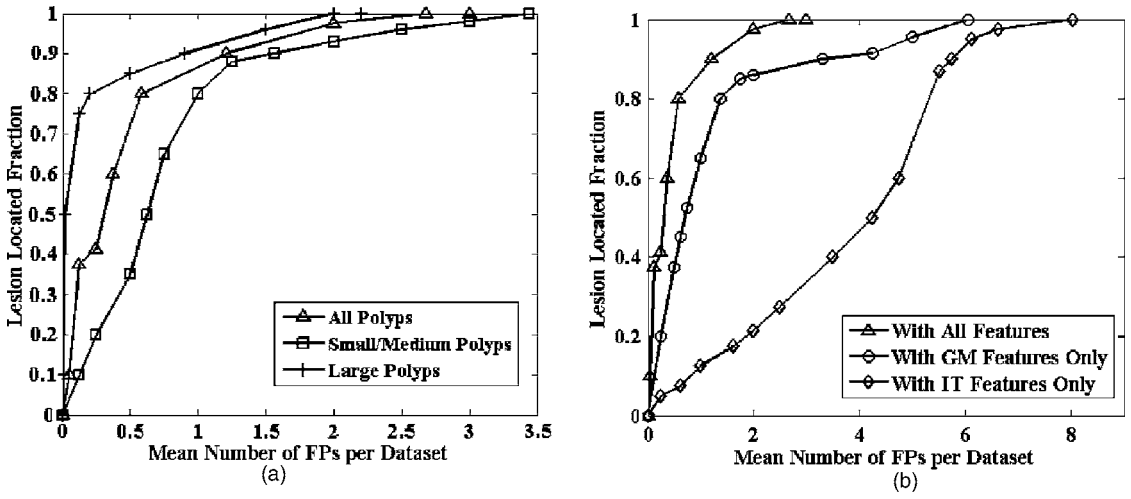

FIG. 10. The FROC curve of the presented CAD scheme for the 153 patient datasets. (a) The FROC curves of the presented CAD scheme for different polyp sizes. (b) The FROC curves for the CAD scheme with different groups of the internal features from all datasets.

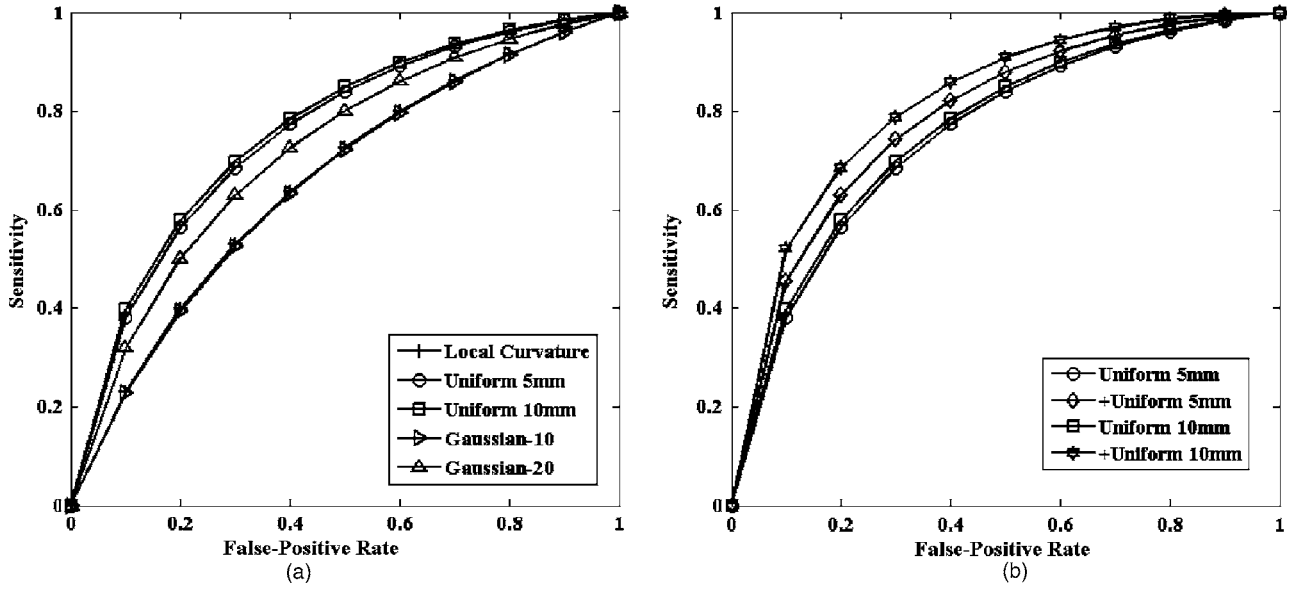

FIG. 11. The ROC curves resulted from polyp detection using different curvatures. Notation—"Uniform [x] mm" indicates the result of the global curvature with 2 uniform kernel function and a half-curve length of x mm; and "Gaussian-[x]" is the result of the global curvature with Gaussian kernel function and an alpha value of "x" mm. Notation "+Uniform" shows the result of a combination of the local and global curvatures. (a) The ROC curves of local and global curvature. (b) The ROC curves from polyp results using the combination of local and global curvature.

combination of the local and global curvatures for the generation of the initial polyp candidates. To see this, we tested the combination of the local and the global curvatures in the surface shape-based filters by the following experiment. In this experiment, only those candidates who passed the shape-based filters with both local curvature and corresponding global curvature will be retained for a further reduction. The results are shown in Fig. 11(b).

The area under the curve (AUC) of the ROC curves is utilized to evaluate the effectiveness of the above different curvatures. From Fig. 11(a), the local curvature (0.656 AUC) is not distinguishable from that of the global curvature with Gaussian-10 (0.653 AUC), both of which have the worst performance. The ROC curves of the global curvature with Uniform-5 mm (0.751 AUC) and 10 mm (0.76 AUC) are almost the same, showing an improved performance over the local curvature. The global curvature with Gaussina-20 (0.717 AUC) performs better than the global curvature with Gaussian-10 (and the local curvature), but worse than the uniform kernels. With a smaller alpha value, the global curvature with Gaussian kernel has less information from its neighborhood. Thus, its difference from the local curvature becomes smaller. Our experiments showed that the global curvature with more neighborhood information can perform better than the curvature with less neighborhood information. This can be seen from the performance between the Gaussina-10 kernel and the Gaussina-20 kernel. The gain by the combination of the local and the global curvatures in the surface shape-based filters is clearly seen in Fig. 11(b). The combination of uniform 5 mm (0.787 AUC) and 10 mm (0.816 AUC) show better performance than those methods using global or local curvature individually.

IV. DISCUSSION AND CONCLUSION

In this paper, we presented a CAD scheme for a CT-based VC screening purpose. It utilized surface shape-based geometrical measures to generate initial polyp candidates (or suspicious patches) on the colon wall inner surface. The FP rate was high, concurring with the previous CAD reports.[16,18,24,26] Even though a smoothed version of the most sensitive curvature measures was considered to minimize a small and rapid shape change, the FP rate remained relatively high, as shown in Table I. Then we turned to utilize the internal features inside the polyp candidates for a further FP reduction.

Six internal features were extracted and used for a further reduction of FPs in the initial candidates. By a simple linear-discrimination classifier, a noticeable gain in eliminating the FPs was demonstrated. In our study, we collected all the 153 cases using a single-slice CT scanner and variable bowel preparations over more than eight years. The presented CAD scheme can detect all polyps (larger than 4 mm in diameter) with a mean FP number of 2.8 per dataset. An eliminating rate of 93.1% on the initial candidates can be achieved for 100% detection sensitivity by the internal features, as shown in Table II. The gain by the internal features is clearly seen, also observed by others.[29,31,33] Nappi, *et al.*[33] achieved a performance of 1.22 average FP per patient with 95% detection sensitivity using 27 CTC datasets, in which there are 21 polyps of 5–25 mm diameter size. Acar *et al.*[31] achieved a performance of 3.0 FP per dataset with 100% sensitivity using 8 patient datasets, in which there are 19 polyps of 5–23 mm diameter size. Goturk *et al.*[29] tested their statistical pattern method by 48 patient datasets. In their experiments, they

reduced 62% FPs with 100% sensitivity in a training set of 30 polyp and 1997 nonpolyp samples (or colon segments). Compared to these research findings, our presented CAD method showed an excellent performance and demonstrated the potential to be one of the efficient methods for the reduction of FPs without sacrificing the sensitivity.

Although the presented CAD scheme showed the potential to improve the efficacy on colonic polyps, several improvements are needed. The threshold for polyp size equal to or larger than 4 mm would be reduced. This will be helpful to follow-up management of the smaller polyps. One solution to improve the detection of smaller polyps is to employ multislice CT scanners, which can now provide isotropic image resolution of less than 1 mm in three dimensions. This image resolution would theoretically enable the detection of polyps as small as 1 mm possible. Another solution is to improve the surface shape-based measures and the internal features. The explored modification on the principal curvatures by the convolution in this paper is an approach to improve the surface shape-based measures. A further investigation on the convolution is needed with other kernels in any direction in three dimensions. In extracting the internal features, a simple eROI model was assumed in this paper. This model may fit the outer portion of the polyps (inside the lumen), but might not be accurate enough for the inner portion (inside the wall). Alternative models for the inner portion shall be explored and are under investigation. Correspondingly, the internal features would be altered.

The use of the linear discrimination for the internal features in this paper has served the purpose of showing the gain of these features for FP reduction. However, it cannot automatically identify which feature is more important and should be weighted more than other less important features. It has been recognized that a poorly designed feature may provide misleading information to the final decision. A good machine learning scheme would have the ability to distinguish the useful features from those less useful features automatically.[48,49] Apparently the linear discrimination is too simple to distinguish these features. Further research and a quantitative analysis on this track is under progress.

ACKNOWLEDGMENTS

This work was partly supported by NIH Grant No. CA82402 of the National Cancer Institute. The authors would like to acknowledge the use of the Viatronix VC navigation system and assistance from Dr. Mark Wax and Dr. Kevin Kreeger on the acquisition of the patient CT datasets. The authors are grateful to Ms. Maura Chodkowski for editing this paper.

[a]Corresponding author. Telephone: 631-444-7837. E-mail: jzl@mil.sunysb.edu

[1]"Cancer facts & figures 2004," American Cancer Society Annual Report, 2004.
[2]M. O'Brien, S. Winawer, A. Zauber, L. Gottlieb, S. Sternberg, B. Diaz, G. Dickersin, S. Ewing, S. Geller, and D. Kasimian, "The national polyp study: patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas," Gastroenterology **98**, 371–379 (1990).
[3]C. Coin, F. Wollett, J. Coin, M. Rowland, R. Deramos, and R. Dandrea, "Computerized radiology of the colon: A potential screening technique," Comput. Radiol. **7**, pp. 215–221 (1983).
[4]D. Vining, D. Gelfand, R. Bechtold, E. Scharling, E. Grishaw, and R. Shifrin, "Technical feasibility of colon imaging with helical CT and virtual reality," *Annual Meeting of American Roentgen Ray Society*, New Orleans, 1994, pp. 104.
[5]W. Lorensen, F. Jolesz, and R. Kikinis, "The exploration of cross-sectional data with a virtual endoscope," in R. Satava and K. Morgan, in *Interactive Tech and New Med Paradigm for Health Care* (IOS Press, Washington, DC, 1995), pp. 221–230.
[6]L. Hong, A. Kaufman, Y. Wei, A. Viswambharan, M. Wax, and Z. Liang, "3D virtual colonoscopy," *IEEE Symposium on Frontier in Biomedical Visualization* (IEEE CS Press, Los Alamitos, 1995), pp. 26–32.
[7]L. Hong, Z. Liang, A. Viswambharant, A. Kaufman, and M. Wax, "Reconstruction and visualization of 3D models of colonic surface," IEEE Trans. Nucl. Sci. **44**, pp. 1297–1302 (1997).
[8]L. Hong, S. Muraki, A. Kaufman, D. Bartz, and T. He, "Virtual voyage: Interactive navigation in the human colon," *Computer Graphics Proceedings*, Annual Conference Series, SIGGRAPH 97, 1997, pp. 27–34.
[9]P. Pickharder, J. Choi, I. Hwang, J. Butler, M. Puckett, H. Hildebrandt, R. Wong, P. Nugent, P. Mysliwiec, and W. Schindler, "Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults," N. Engl. J. Med. **349**, 2191–2200 (2003).
[10]J. Yee, G. Akerkar, R. Hung, A. Steinauer-Gebauer, S. Wall, and K. McQuaid, "Colorectal neoplasia: performance characteristics of CT colonoscopy for detection in 300 patients," Radiology **219**, 685–692 (2001).
[11]D. Johnson, W. Harmsen, L. Wilson, R. MacCarty, T. Welch, D. Ilstrup, and D. Ahlqist, "Prospective blinded evaluation of CT colonoscopy for screen detection of colorectal polyps," Gastroenterology **125**, 311–319 (2003).
[12]J. Ferrucci, "Colon cancer screening with virtual colonoscopy: promise, polyps, politics," Am. J. Roentgenol. **177**, 975–988 (2001).
[13]Z. Liang, S. Lakare, M. Wax, D. Chen, J. Anderson, A. Kaufman, and D. Harrington, "A pilot study on less-stressful bowel preparation for virtual colonoscopy screening with follow-up biopsy by optical colonoscopy," SPIE Medical Imaging **5746**, 810–816 (2005).
[14]Z. Liang, "Virtual colonoscopy: An alternative approach to examination of the entire colon," INNERVISION, 2001, Vol. **16**, No. 1, pp. 40–44.
[15]D. Vining, Y. Ge, D. Ahn, and D. Stelts, "Virtual colonoscopy with computer-assisted polyps detection," in *Computer-Aided Diagnosis in Medical Imaging* (Elsevier Science B. V., Amsterdam, Netherlands, 1999), pp. 445–452.
[16]R. Summers, C. Beaulieu, L. Pusanik, J. Malley, R. Jeffrey, D. Glazer, and S. Napel, "Automated polyp detection at CT Colonography: Feasibility study," Radiology **216**, 284–290 (2000).
[17]R. Summers, C. Johnson, L. Pusanik, J. Malley, A. Youssef, and J. Reed, "Automated polyp detection at CT Colonography: Feasibility assessment in a human population," Radiology **219**, 51–59 (2001).
[18]H. Yoshida and J. Nappi, "Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps," IEEE Trans. Med. Imaging **20**, 1261–1274 (2001).
[19]H. Yoshida, Y. Masutani, P. Maceneaney, D. Rubin, and A. Dachman, "Computerized detection of colonic polyps at CT colonography on the basis of volumetric features: Pilot study," Radiology **222**, 327–336 (2002).
[20]H. Yoshida, J. Nappi, P. MacEneaney, D. Rubin, and A. Dachman, "Computer-aided diagnosis scheme for detection of polyps at CT colonography," Radiographics **22**, 963–979 (2002).
[21]J. Nappi and H. Yoshida, "Automated detection of polyps with CT Colonography: Evaluation of volumetric features for reduction of false-positive findings," Acad. Radiol. **9**, 386–397 (2002).
[22]J. Nappi, H. Frimmel, A. Dachman, and H. Yoshida, "Computerized detection of colorectal masses in CT colonography based on fuzzy merging and wall-thickening analysis," Med. Phys. **31**, 860–872 (2004).
[23]D. Paik, C. Beaulieu, R. Jeffrey, C. Karadi, and S. Napel, "Detection of polyps in CT colonography: A comparison of a computer-aided detection algorithm to 3D visualization methods," Radiology **213** 197 (1999).
[24]D. Paik, C. Beaulieu, R. Jeffrey, J. Yee, A. Steinauer-Gebauer, and S. Napel, "Computer aided detection of polyps in CT colonography: Method and free-response ROC evaluation of performance," Radiology **217**, 370 (2000).

[25]D. Paik, C. Beaulieu, A. Mani, R. Prokesch, J. Lee, and S. Napel, "Evaluation of computer-aided detection in CT colonography: Potential applicability to a screening population," Radiology **221**, 332 (2001).

[26]G. Kiss, J. Cleynenbreugel, M. Thomeer, G. Marchal, and P. Suetens, "Computer aided diagnosis for virtual colonography by geometrical model fitting," *Proc. of 7th International Workshop on Vision, Modeling, and Visualization*, Germany, 2002, pp. 27–34.

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

[28]S. Gokturk, C. Tomasi, B. Acar, D. Paik, C. Beaulieu, and S. Napel, "A new 3-D volume processing method for polyps detection," *Proc. of the 23rd Annual Intl Conference of the IEEE Engineering in Medicine and Biology Society*, 2001, pp. 2522–2525.

[29]S. Gokturk, C. Tomasi, B. Acar, C. Beaulieu, D. Paik, R. Jeffrey, J. Yee, and S. Napel, "A statistical 3-D pattern processing method for computer-aided detection of polyps in CT Colonoscopy," IEEE Trans. Med. Imaging **20**, 1251–1260 (2001).

[30]S. Gokurk and C. Tomasi, "A new 3D pattern recognition technique with application to computer aided colonoscopy," *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, pp. I93–T100.

[31]B. Acar, C. Beaulieu, D. Paik, S. Gokturk, C. Tomasi, J. Yee, and S. Napel, "Assessment of an optical flow field-based polyp detector for CT Colonography," *Proc. of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2001, pp. 2774–2777.

[32]B. Acar, C. Beaulieu, S. Gokturk, C. Tomasi, D. Paik, R. Jeffrey, J. Yee, and S. Napel, "Edge displacement field-based classification for improved detection of polyps in CT Colonoscopy," IEEE Trans. Med. Imaging **21**, 1461–1467 (2002).

[33]J. Nappi and H. Yoshida, "Feature-guided analysis for reduction of false positives in CAD of polyps for CT Colonography," Med. Phys. **30**, 1592–1601 (2003).

[34]J. Yao, M. Miller, M. Franaszek, and R. Summers, "Colonic polyp segmentation in CT colonoscopy-based on fuzzy clustering and deformable models," IEEE Trans. Med. Imaging **23**, 1344–1352 (2004).

[35]O. Monga, S. Benayoun, and O. Faugeras, "From partial derivative of 3D density images to ridge lines," *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1992, pp. 354–359.

[36]O. Monga, R. Lengagne, and R. Deriche, "Crest lines extraction in volume 3D medical images: A multi-scale approach," *Proc. of the 12th IAPR International Conference on Computer Vision & Image Processing*, 1994, pp. 553–555.

[37]C. Dorai and A. Jain, "COSMOS—A representation scheme for 3D free-form objects," IEEE Trans. Pattern Anal. Mach. Intell. **45**, 2132–2138 (1997).

[38]C. Wijk, R. Truyen, R. Gelder, L. Vliet, and F. Vos, "On normalized convolution to measure curvature Features for automatics polyps detection," *Proceedings of MICCAI 2004*, 2004, pp. 200–208.

[39]Z. Liang, F. Yang, M. Wax, J. Li, J. You, A. Kaufman, L. Hong, H. Li, and A. Viswambharan, "Inclusion of a priori information in segmentation of colon lumen for 3D virtual colonoscopy," Conf. Record IEEE NSS-MIC, in CD-ROM, 1997.

[40]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 Trans. Med. Imaging **19**, 1220–1226 (2000).

[41]B. Cabral and L. Leedom, "Imaging vector fields using line integral convolution," ACM SIGGRAPH Computer Graphics, *Proc. of the 20th Annual Conference on Computer Graphics*, 1993, pp. 263–270.

[42]Z. Wang, L. Li, and Z. Liang, "Skeleton based 3D computer aided diagnosis for detection of the colon polyps," SPIE Medical Imaging **5032**, 843–853 (2003).

[43]R. Haggitt, R. Glotzbach, E. Soffer, and L. Wruble, "Prognostic factors in colorectal carcinomas arising from adenomas: Implications from lesions removed by endoscopic polypectomy," Gastroenterology **89**, 328–336 (1985).

[44]G. Strang and T. Nguyen, *Wavelets and Filter Banks* (Cambridge Press, Wellesley, 1996).

[45]W. Gander, G. Golub, and R. Strebel, "Least-squares fitting of circles and ellipses," BIT Numerical Mathematics **34**, 558–578 (1994).

[46]Z. Wang, L. Li, J. Anderson, D. Harrington and Z. Liang, "Computer aided detection and diagnosis of colon polyps with morphological and texture features," SPIE Medical Imaging **5370**, 972–979 (2004).

[47]Z. Wang and Z. Liang, "Sphere light field rendering," SPIE Medical Imaging **4681**, 357–365 (2002).

[48]A. Jerebko, J. Malley, M. Franaszek, and R. Summers, "Multiple neural network classification scheme for detection of colonic polyps in CT Colongraphy data sets," Acad. Radiol. **10**, 154–160 (2003).

[49]A. Jerebko, R Summers, J. Malley, M. Franaszek, and C. Johnson, "Computer-assisted detection of colonic polyps with CT colonography using neural networks and binary classification trees," Med. Phys. **30**, 52–60 (2003).

[50]D. Chakraborty, "The FROC, AFROC and DFROC variants of the ROC analysis," *Handbook of Medical Imaging, Vol. 1, Physic and Psychophysics* (SPIE Press, Bellingham, WA, 2000), pp. 771–796.