HHS Public Access

Author manuscript

Int J Comput Assist Radiol Surg. Author manuscript; available in PMC 2015 October 03.

Published in final edited form as:

Int J Comput Assist Radiol Surg. 2014 November ; 9(6): 1021–1031. doi:10.1007/s11548-014-0991-2.

Volumetric texture features from higher-order images for diagnosis of colon lesions via CT colonography

Bowen Song,
Department of Radiology, Stony Brook University, Stony Brook, NY 11790, USA

Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11790, USA

Guopeng Zhang,
Department of Biomedical Engineering, Fourth Military Medical University, Xi'an 710032, Shaanxi, China

Hongbing Lu,
Department of Biomedical Engineering, Fourth Military Medical University, Xi'an 710032, Shaanxi, China

Huafeng Wang,
Department of Radiology, Stony Brook University, Stony Brook, NY 11790, USA

Wei Zhu,
Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11790, USA

Perry J. Pickhardt, and
Department of Radiology, University of Wisconsin School of Medicine and Public Health, Madison, WI 53792, USA

Zhengrong Liang
Department of Radiology, Stony Brook University, Stony Brook, NY 11790, USA

Zhengrong Liang: jerome@mil.sunysb.edu

Abstract

Purpose—Differentiation of colon lesions according to underlying pathology, e.g., neoplastic and non-neoplastic lesions, is of fundamental importance for patient management. Image intensity-based textural features have been recognized as useful biomarker for the differentiation task. In this paper, we introduce texture features from higher-order images, i.e., gradient and curvature images, beyond the intensity image, for that task.

Methods—Based on the Haralick texture analysis method, we introduce a virtual pathological model to explore the utility of texture features from high-order differentiations, i.e., gradient and

Correspondence to: Zhengrong Liang, jerome@mil.sunysb.edu.

Conflict of interest Bowen Song, Guopeng Zhang, Hongbing Lu, Huafeng Wang, Wei Zhu, Perry J. Pickhardt and Zhengrong Liang declare that they have no conflict of interest. Dr. Liang is a co-founder of Viatronix. Dr. Pickhardt has served as a consultant for Viatronix, Braintree and Mindways and is co-founder of VirtuoCTC.

curvature, of the image intensity distribution. The texture features were validated on a database consisting of 148 colon lesions, of which 35 are non-neoplastic lesions, using the support vector machine classifier and the merit of area under the curve (AUC) of the receiver operating characteristics.

Results—The AUC of classification was improved from 0.74 (using the image intensity alone) to 0.85 (by also considering the gradient and curvature images) in differentiating the neoplastic lesions from non-neoplastic ones, e.g., hyperplastic polyps from tubular adenomas, tubulovillous adenomas and adenocarcinomas.

Conclusions—The experimental results demonstrated that texture features from higher-order images can significantly improve the classification accuracy in pathological differentiation of colorectal lesions. The gain in differentiation capability shall increase the potential of computed tomography colonography for colorectal cancer screening by not only detecting polyps but also classifying them for optimal polyp management for the best outcome in personalized medicine.

Keywords

CT colonography; Colorectal lesions; Texture feature; Textural biomarker; Gradient; Curvature; Computer-aided diagnosis

Introduction

According to the up-to-date statistics from the American Cancer Society, colorectal cancer ranks second for most common occurrence and total cancer death since 2012 in the United States [1]. Fortunately, most colorectal cancers arise from colonic polyps, which may take years or even decades to become cancer, and early detection and removal of polyps prior to malignant transformation can effectively reduce the incidence of colon cancer [2,3]. Computer-aided detection (CADe) [4–8] has been introduced to improve the detection capability of CT colonography (CTC), which has shown several advantages as a potential minimally invasive screening technique comparing to the traditional optical colonography, e.g., safer and less expensive. Once a polyp of size greater than a certain threshold is detected, it will be resected regardless whether it is adenoma or not in current clinical practice. Given the knowledge that the majority of polyps are hyperplastic [9], the current practice of resecting these harmless growths may not be optimal. Therefore, differentiating hyperplastic polyps from adenomas would be useful to potentially reduce the number of unnecessary colonoscopy procedures.

The main task of computer-aided diagnosis (CADx) is to differentiate the pathological stages to which each detected lesion belongs. Typically, colorectal lesions can be categorized into three groups, i.e., hyperplastic, adenomas and malignant lesions. Hyperplastic polyps (H), usually found in the distal colon and rectum, are non-neoplastic and seldom show malignant potential [10]. Neoplastic lesions include adenomas and malignant lesions. In general, adenomas are benign in nature; however, they are at risk for malignant transformation. Adenomas are divided into three subtypes, based on histological criteria, as follows: tubular, tubulovillous and villous adenomas, with the degrees of dysplasia increasing. Tubular adenomas (TA) are usually non-advanced neoplastic lesions, while the

other two subtypes, i.e., tubulovillous adenoma (TVA) and villous adenoma (VA), are generally more advanced neoplastic lesions [11]. Colorectal adenocarcinomas (A) are malignant lesions (advanced neoplastic) which invade into the submucosa and beyond, including metastatic spread. Colorectal lesions belonging to different pathologic categories have distinct tissue patterns [12], which is a main factor based on which the pathologists make their diagnosis decision. Inspired by this fact, if we could extract useful pattern information from the detected polyps, we would achieve our CADx task goal.

Texture features are, in fact, mathematical parameters computed from the image intensity distribution of pixels/voxels in two-/three-dimensional (2D/3D) space which characterize the texture types and thus the underlying structures of the objects embedded in the image. Various texture feature extraction algorithms have been introduced [13], and among them, the co-occurrence matrix (CM)-based Haralick method [14] has been widely used as a standard way to perform the task of texture feature extraction [15–19]. Texture features have been employed for CADe, in which it usually serves as ancillary features to other features, e.g., geometric features, for classification of interested lesions (e.g., colon polyps) against false positives (FPs) or patches of normal tissues, fecal residues and image artifacts. In the study [20], the feasibility of using CT image intensity texture as biomarker to predict patient survival and the close relationship between texture and tumor staging have been shown. As reported in [21], with the assistance of CADx, radiologists' differentiation performance can be improved significantly. To the best of our knowledge, employing texture feature analysis techniques for differentiating colonic lesion neoplasm has not been widely reported in this field.

Based on our previous study [22], in which we introduced gradient–intensity CM, we take a further step and propose a novel texture feature extraction strategy by including high-order derivative images, e.g., gradient (first order) image and curvature (second order) image, in mimicking the amplification operation in pathology. Our approach first generates the 3D gradient and 3D curvature images using Sobel operator [23] and Monga's method [24], respectively. Then, the corresponding CMs are computed based on the Haralick method. The details of our texture feature extraction strategy are given in "Material and methods" section below. The diagnostic performance of different combinations of the computed texture features is evaluated and compared in "Results" section. The discussion, which includes the previous similar studies, the limitation of this study and our future work, can be found in "Discussion" section. At last, conclusion is drawn in "Conclusion" section.

Material and methods

In this section, we first describe how to generate the high-order derivative images, followed by an introduction of a 3D expansion from the original Haralick's 2D texture model. Then, our volumetric feature extraction strategy is presented. Figure 1 shows the flowchart of texture feature extraction process in our study. The database information and experiment design are discussed in the last two parts of this section.

3D gradient image

Gradient distribution reflects the intensity changing degree. The combination of intensity–gradient pair has shown some potential of differentiating pathology phases of colon lesions [22]. And here we try to study the property of texture features considering gradient only. Sobel operator [23] is a common discrete filter kernel for gradient estimation. 3D Sobel operator is employed here to generate the gradient image. A standard 3D Sobel kernel has size 3^3 for each of the three orthogonal axes. The *z*-direction 3D Sobel kernel is shown as (with 1/32 as the normalization factor):

$$h_z'(:,:,-1)=\begin{bmatrix}1&2&1\\2&4&2\\1&2&1\end{bmatrix}\quad h_z'(:,:,0)=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}\quad h_z'(:,:,1)=\begin{bmatrix}-1&-2&-1\\-2&-4&-2\\-1&-2&-1\end{bmatrix} \tag{1}$$

By rotating the *z*-direction kernel matrix, we can easily get the *x* and *y* direction kernel. For a voxel in image *I*, denoted as $I(i,j,k)$, we can compute the derivatives $G_x(i,j,k)$, $G_y(i,j,k)$ and $G_Z(i,j,k)$ in the three orthogonal directions. The corresponding gradient value for that voxel is defined as:

$$I_g(i,j,k)=\sqrt{G_x^2(i,j,k)+G_y^2(i,j,k)+G_z^2(i,j,k)} \tag{2}$$

Then, the 3D gradient image I_g is generated. The gradient image would reflect a major aspect, i.e., the rate of image intensity changing across the 3D space.

3D curvature image

Given an iso-surface, gradient is closely related to first-order derivative, considering the fact that it is one-to-one correspondence to a tangent plane for a point. We would also like to study the property of texture features based on the second-order derivative-related scalars, i.e., curvature. Intuitively, curvature is the amount by which a geometric object deviated from being flat. Curvature analysis with application to CADe has been widely reported [4–6,25]. Since curvature calculation can be considered as a 3D operator, it can be directly applied to 3D image like the Sobel operator above. If certain iso-surface-like patterns exist inside a 3D volumetric image, the curvature information can help to catch the pattern distribution. It would reflect another major aspect, i.e., the rate of image intensity gradient changing and surfacelike pattern across the 3D space. In the report [24], a voxel curvature estimation method was proposed by directly using partial derivatives of a 3D image. The key point for calculating the curvature in the 3D volume image is to construct the Hessian matrix (HM):

$$\mathrm{HM}=\begin{bmatrix}I_{xx}&I_{xy}&I_{xz}\\I_{xy}&I_{yy}&I_{yz}\\I_{xz}&I_{yz}&I_{zz}\end{bmatrix} \tag{3}$$

where I is the partial derivatives of the gray-level image function $I(x, y, z)$. The 3D Deriche filters [26] are used to compute the partial derivatives of the image data, where we set the calculation formulas as:

$$f_0(x)=c_0(1+\alpha|x|)e^{-\alpha|x|}\ f_1(x)=-c_1x\alpha^2e^{-\alpha|x|}\ f_2(x)=c_2(1-c_3\alpha|x|)e^{-\alpha|x} \quad (4)$$

The normalization coefficients c_0, c_1, c_2, c_3 are set to be:

$$c_0=\frac{(1-e^{-\alpha})^2}{1+2e^{-\alpha}\alpha-e^{-2\alpha}}\ c_1=\frac{-(1-e^{-\alpha})^3}{2\alpha^2e^{-\alpha}(1+e^{-\alpha})}\ c_2=\frac{-2(1-e^{-\alpha})^4}{1+2e^{-\alpha}-2e^{-3\alpha}-e^{-4\alpha}}\ c_3=\frac{(1-e^{-2\alpha})}{2\alpha e^{-\alpha}} \quad (5)$$

where α controls the degree of suppressing noise amplification during the high-order partial derivative calculation. Then, the partial derivatives can be calculated as following:

$$I_{xx}=(f_2(x)f_0(y)f_0(z))*I\ I_{xy}=(f_1(x)f_1(y)f_0(z))*I \quad (6)$$

where I_{yy}, I_{yz}, I_{xz}, I_{zz} can be determined by substituting the variables in the equations above. Two principal curvatures, i.e., κ_1 and κ_2, can be calculated using H, and then, Gaussian curvature is calculated based on κ_1 and κ_2 to form the curvature image:

$$I_c(x,y,z)=\kappa_1(x,y,z)*\kappa_2(x,y,z) \quad (7)$$

3D expansion of Haralick's texture analysis model

In 1973, Haralick texture model [14] was first introduced to analyze 2D gray-level image. The main idea is to construct the co-occurrence matrix (CM), which is defined over an image to reflect the distribution of co-occurring values at a given offset. By assuming all the texture information is contained in the CM, a set of 13 measures of textural features are derived, e.g., angular second moment, contrast, correlation and entropy. These features have been used with success on biological cell images, X-ray images, satellite images, aerial images, etc. [13,15–19,27–30].

While the above 2D applications generated encouraging results, 3D image processing is attracting growing amount of attention in volumetric medical imaging area, e.g., CADe/CADx studies based on 3D volumetric analysis have been reported for CTC [30–32]. Since Haralick texture analysis is based on the CM extraction, which is a counting matrix storing spatial value patterns for certain direction, its 2D model can be expanded to 3D space, as introduced in [16,22,30]. Here, we briefly introduce the 3D expansion model, very similar to [22], in this section. A comparison of 2D and 3D model can be found in "Results" section.

With a given direction θ and distance d,a $N_g \times N_g$ square gray-level co-occurrence matrix (GLCM), based on gray-level CT intensity image, can be built based on each 2D image:

$$\mathrm{GLCM}(\theta,d)=\begin{bmatrix} p(1,1) & p(1,2) & \cdots & p(1,N_g) \\ p(2,1) & p(2,2) & \cdots & p(2,N_g) \\ \vdots & \vdots & \ddots & \vdots \\ p(N_g,1) & p(N_g,2) & \cdots & p(N_g,N_g) \end{bmatrix} \quad (8)$$

where $p(\cdot,\cdot)$ is the normalized frequency. θ can take four values (directions), i.e., 0°(180°), 45°(225°), 90°(270°) and 135° (315°), and d indicates our choice of value pair from the dth neighborhood of current center point shown in Fig. 2.

As introduced in [16,22,30], the 2D model structure can be expanded to 3D case by simply adding more directions. A 3×3×3 cubic model is shown in Fig. 3 (we only consider $d = 1$ neighbor in this study, 26 neighbors). Total 13 distinct directions are defined to generate the CM. Features calculated from the CM for each direction are listed in Table 1.

By employing the 3D expansion model on different 3D images, i.e., gray-level, gradient and curvature images, we can obtain the GLCMs, gradient co-occurrence matrices (GCMs) and curvature co-occurrence matrices (CCMs). The corresponding features can be extracted from each CM. The mean and range value over the 13 directions for each feature are outputted as the final features, and there are a total of 26 features for each 3D image.

Semi-automatic operation for volume of interests (VOIs)

In implementing the 3D texture model for differentiation of polyp types, we assume a polyp candidate has been detected by a radiologist expert or a CADe pipeline via CTC. Now, our interest is to obtain a volume of interest or VOI for the polyp and extract features from the volume of interest (VOI) to determine its lesion type. Very similar to [19,20,33], a semiautomatic technique is employed here to extract the VOIs. Based on the CTC reports from radiologists with the coordinates of each polyp in the original CT images, the VOI of each lesion was firstly outlined manually on the 2D image slices by the experts using the CTC software (V3D Colon, Viatronix Inc., Stony Brook, NY, USA). After that, an automatic air cleaning algorithm was developed to remove air voxels inside the outline and an air-free 3D lesion CT image or VOI was generated [34,35]. Figure 4 shows the details of this semi-automated procedure. Validation and discussion of this procedure can be found in Results and Discussions section.

Database

Evaluation of the newly proposed high-order texture feature extraction strategy was conducted on a CTC database of 110 scans from 56 patients (most patients have a supine and prone scans; there are two patients with only one scan because of the lack of pathological report confirmation) with polyps of size from 8mm (clinically significant and computationally meaningful) to 30mm, see Fig. 5a. Total of 148 lesions are included in our study, among which 35 are H, 72 are TA, 36 are TVA and 5 are A. Since there is only one VA case, we categorized it into TVA group because of their similarity in risk (with presence of a prominent ≥25% villous component). Detailed pathological information of the dataset is shown in Fig. 5b.

It is noted that the above database, which is a part of the DoD clinical study [36], was downloaded from the online public domain (http://imaging.nci.nih.gov) where the CTC database was made by the Walter Reed Army Medical Center (WRAMC) after the clinical trial study [36]. Part of the DoD trial datasets do not have the location information of the lesions and, therefore, could not be used in this study. The rest has the locations information, but some cases may not have exact matches between VC and OC reports and certain cases

even do not have clear reports. After the removal of the uncertain cases, the remaining cases were used in this study. The information about bowel preparation, data acquisition, CT image reconstruction parameters, patient population, gold standard (optical colonoscopy and pathology) is given at the Web site and also in the full paper [36]. All examinations were performed in adherence following standards [37] with a full cathartic bowel preparation, fecal tagging, without IV contrast and with multi-detector CT scanners. The image data was acquired in helical mode with collimations of 1.25–5mm, pitch of 1–2, reconstruction intervals of 1.25–5mm, and modulated tube current–time products ranging from 50 to 200mAs and tube voltages from 80 to 120kVp. The indication for CTC was screening for colorectal cancer in all individuals. The human studies had been approved by appropriate ethical committee and have therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All subjects gave their informed consent prior to their inclusion in the study. The subjects' identities have been fully anonymized.

Experimental design

For each lesion or VOI, 26 features are calculated for each 3D image (i.e., the intensity, gradient or curvature), so we have 78 features for each lesion in total. In pathology, H lesions are considered as non-neoplastic and the other three types are considered as neoplastic lesions, which has the risk of uncontrollable abnormal growth or division of cells [9,33]. TA is considered as non-advanced neoplastic, while TVA and A are considered as advanced neoplastic lesions. We coded type H as class "0" and all the rest as class "1." We generated different feature set by combining the intensity, gradient and curvature features. A total of seven feature sets can be generated, and their classification performance was compared.

The binary SVM classifier was employed here to perform the classification task. In implementation, the widely used SVM package LIBSVM [38] with RBF kernel was employed in this study. Following the guideline of LIBSVM, the two parameters (cost: the model slack variable and gamma: the parameter in the RBF kernel) were determined by a grid search process (fivefold cross-validation) in the training step, detailed implementation can be found in [39]. The ROC analysis and the AUC merit were used as the measure for evaluation. For one experiment, we randomly split the dataset into training and test set with same size while keeping the original class proportion rate. The SVM model was trained using training set and tested using test set. Figure 6 shows the flowchart of single evaluation experiment in this study. The whole process was repeated 100 times to avoid time running error, and the average results are shown in the following section.

Results

Linear size versus volumetric size

As discussed in [9], polypoid lesions, which refer to both sessile and pedunculated polyps, account for the vast majority of lesion cases, including most advanced adenomas and cancers. Meanwhile, based on several studies, detected flat lesions are less likely to be neoplastic compared to polypoid ones, even with larger size [40–42]. For example, 8 (5.4%)

of 148 lesions in our database are flat, where 6 (75%) of which belongs to H group, with the rest of the database being polypoid ones. This indicates that the distribution of lesion morphology in our database is consistent with the previous reports and therefore is representative.

To assess our semi-automatically extracted 3D volumes, we intend to check whether there exists a linear relationship between the clinically reported linear size and our 3D volume voxel number to make sure that the extracted volume does not bias the results. As introduced in [6,7], ellipsoid or sphere-like convex ball is reasonable to model the structure the polypoid lesions, and therefore, we regress the third root of the volume of each lesion on its corresponding linear size. The regression results are shown in Fig. 7. The P value of coefficient of independent variable size is <0.001. Combining the fact that R^2 (coefficient of determination, which indicates how well the points fit the model) is 0.80, we can conclude that the sizes of our extracted 3D lesion volumes are consistent with the reported linear sizes. Even though several outliers, especially those for large size lesions, can be observed from the regression scatter plot, the result is still acceptable based on the R^2 measure. This may be explained by facts that the appearance of lesions is changing due to the scan position of patient and the changing degree is proportional to the size.

As reported in [9,43], the size of lesion is an important biomarker, which correlates with its risk of malignancy and guides its clinical treatment. The risk of malignancy increases with increasing polyp size. However, as stated in [44], one main drawback of linear polyp size measurement is that there exists 1–3mm variability among measurements in pathology, OC and CTC. And CTC polyp measurements were generally underestimated by 2–3mm compared with OC. Most studies nowadays on colon cancer detection have focused on lesions <10mm, in which case 1–3mm measure error would be non-negligible and may affect the decision accuracy. Therefore, an accurate and consistent merit on the polyp size is desired. As pointed out in a recent study [33], 3D volume size may be a more robust measurement than linear size. This point was also verified in this study. By using the reported linear size and our extracted 3D volume size as decision value, we plot their ROC curve, respectively, in Fig. 8. The AUC measures are nearly the same (0.62 level), which indicates that both size measurements contain pathology differentiation information, while the ROC curves are quite different. In Fig. 8a, there are only 10 points forming the ROC curve, while in Fig. 8b nearly each lesion case is corresponding to a single constructing point (there are more than 140 points). This is understandable because the linear lesion size is usually recorded in integer millimeter in pathology report; and comparing to the 3D volume size, i.e., the voxel number of the 3D volume, the linear lesion size carries less discriminative information. Generally speaking, a ROC curve with less constructing points can be more easily affected by outliers, e.g., error measuring cases, which usually lead to a larger change in AUC value. Based on the result of Fig. 8 and discussion above, the 3D volume size is a better choice to measure the lesion size. However, the 3D volume size alone contains limited discriminative information, which can be observed from Fig. 9, where the histogram of 3D volume size within each class group is shown. The overlap between H and the risk group renders a challenge in differentiation. By the AUC merit, an AUC of 0.62 cannot be considered as good classification accuracy. In other words, more discriminative features are needed, in addition to the size measure.

Evaluation of feature set

In addition to the agreement analysis of our database with the literature reports and the correlation study of our 3D volume measurements with the clinically reported linear measurements, we further performed an experiment to compare the original 2D Haralick model and its 3D expansion using the image intensity information. The first row of Table 2 shows the average classification results over the 100 runs of the original Haralick 2D model and its 3D expansion. An improvement of classification accuracy by the 3D expansion over its original 2D model can be observed, as expected. Based on this experimental outcome, we can conclude that our implementation of the 2D and 3D Haralick texture feature extractions is successful. Then, we moved on to perform more comparison studies on different feature sets, as shown by Table 2. A significant test on the differences over 100 runs was conducted comparing these 2D and 3D feature sets, and all the results showed that all the P values are <0.001, showing in the last column in Table 2, which indicates that the 3D texture feature extraction model is significantly better than the 2D model. Meanwhile, we can also observe that by combining the high-order texture features, i.e., gradient and curvature features, the classification results are significantly improved than that using the intensity features only. The classification performance in differentiating H from the risk group (TA, TVA, A) achieved 0.8525 of AUC when Gradient–Curvature texture features are used. From the report in Table 2, another fact can be observed that the classification performance of individual gradient or curvature features only did not show improvement, while the performance of combined features set did show improvement, which indicates that the high-order features bring new group deterministic information into the CADx paradigm. When we use all the features, i.e., intensity–gradient–curvature, its AUC (0.8399) is not the highest and even the difference to group gradient–curvature (0.8525) is significant (Table 3). It indicates that including all feature sets does not guarantee best performance and a feature selection scheme is need to optimize the differentiation performance, which would be our next work.

Following the instructions in [45], the averaged ROC curves of the 3D texture feature classification are plotted as shown in Fig. 10. The curves are consistent with the AUC results of Table 2. From both Table 2 and Fig. 10, we can see that by combining the 3D gradient and curvature texture features, we obtained the best classification result. Significance test was conducted between 3D gradient–curvature features, and the rest feature groups and results are shown in Table 3, from which we can observe that the improvement in AUC of gradient–curvature feature set is significant. Based on the averaged ROC curve, a sensitivity–specificity table is made as shown in Table 4. By the 35 non-neoplastic lesions and 113 neoplastic lesions in the database, the average classification results (i.e., the proportion of correctly classified cases within each group) with given sensitivity level can be pictured as follows. For instance, taking the gradient– curvature group as an example, when we choose 0.8 sensitivity level (sensitivity = 0.8), its corresponding specificity equals 0.7178. Therefore, the average number of correctly classified neoplastic lesions is 90.4 (of 113) with 25.1 (of 35) non-neoplastic lesions being recognized by the classifier.

Discussion

In a previous study, Suzuki et al. [46] obtained similar outcome as ours as reported above, where they achieved 0.87 neoplastic lesions classification accuracy with 0.57 on non-neoplastic lesions. The corresponding AUC of their study was reported to be 0.82. No averaged result is reported. This previous study [46] focused more on proposing CADx strategy rather than the texture feature analysis. In our texture biomarker study, when choosing gradient–curvature group and setting sensitivity to be 0.87, we have achieved a specificity of 0.595 with AUC being 0.8525. By comparing the numerical numbers, a moderate improvement can be observed over the previous study [46]. Another difference between our work and the previous study [46] is the use of average in our experiments. It is expected that averaged results would be more reliable and robust to the generalization evaluation.

Even though the microscopic pathological pattern may not be exactly reflected by the voxel-level pattern in the macro medical images, certain tissue-level texture information may be embedded in the voxel-level pattern, where the pattern distributional information can be captured by our texture analysis. This conjecture can be thought as a hypothesis, where only experiments would prove its validity. The outcome of this experimental study indicates the validity of the conjecture.

Since the Haralick method is based on the spatial CM, which records the frequency of micro-texture patterns, there-fore the sampling space could not be too small. For this reason, we only included lesions with size 8mm or larger in this study. It is expected that for lesions with different sizes, their texture features may have variable performance in differentiating the lesions. This renders one of our future research tasks.

Since gradient and curvature operators are applied to the intensity image in our feature extraction, we need to consider the effect brought by the intensity image noise. For gradient operator, we do not need to care much about the noise, because the calculation is almost linear. The consistent results of Sobel operator with d=1 (3*3*3 window) for all the datasets in this study indicate the adequacy of our assessment on the use of the gradient operator. However, higher-order derivatives, e.g., the second-order derivative, require greater care about the noise in order to achieve reasonable immunity from the noise. For example, we did a grid search over the feasible α space for the best fit in all the cases of this study, and an empirical value of $\alpha = 0.8$ was obtained. This empirical value is consistent with the one reported in [24]. A too large α value would lead to a homogeneity image with less texture, while a smaller α value would result in an unreliable calculation of curvature, and both scenarios would lead to inferior classification performance.

It is noticed that the number of our datasets is somehow small, especially for subtypes adenocarcinoma and VA. This is one limitation of the study, and we should expand our database to include more polyps for further investigation on the presented texture feature extraction strategies. Meanwhile, we also noticed that different feature subset combinations can result in different differentiation performances, and the whole feature set will not guarantee the best performance. Since we only conducted a rough feature set selection

(another limitation of the study), a finer feature subset selection is necessary and this renders another one of our future research tasks.

Theoretically, more high-order derivative images can be considered by our texture feature analysis in the absence of intensity image noise. However, the computation of high-order derivatives will amplify the noise quickly. So, a tradeoff must be made for optimal performance. Using empirical value will not guarantee a better performance. This is also the limitation of our current study. Determining the optimal point is currently an open question and would render another task of our future research. Applying the presented strategy for differentiation of other lesions, e.g., lung nodules, is also a topic of our future research.

Conclusion

In this paper, we presented a novel texture feature extraction strategy by including high-order derivative images over the original image intensity distribution, e.g., the first-order gradient and the second-order curvature images, in mimicking the amplification operation in pathology, based on the 2D Haralick texture analysis method and its 3D expansion for CADx of colon polyps via CTC. The results show that by combining the high-order texture features, the classification in differentiating neoplastic from non-neoplastic lesions can reach an AUC value of 0.8525. We can conclude that the derived new pathology-related deterministic information of the high-order gradient and curvature texture features has significantly improved the diagnosis accuracy.

Based on the differentiating capability of the extracted texture features, we conjecture that the derived pathology-related deterministic information can be a sensitive biomarker not only for diagnosis of colon polyps but also for prognosis and prediction of the lesions as well as to reveal the insight of the underline biology.

Acknowledgments

This work was partially supported by the NIH/NCI under Grant #CA082402 and #CA143111. H. Lu is partially supported by the National Natural Science Foundation of China under Grant #81230035 and #81071220 and also the National Key Technologies R&D Program of China under Grant #2011BAI12B03. P. Pickhardt was partially supported by the NIH/NCI under Grant #CA144835, #CA169331 and #CA155347. The authors would like to acknowledge the use of the Viatronix V3D-Colon Module.

References

1. American Cancer Society. Cancer facts & figures 2012. Atlanta, GA: American Cancer Society; 2012.
2. Eddy D. Screening for colorectal cancer. Ann Intern Med. 1990; 113:373–384. [PubMed: 2200321]
3. Gluecker T, Johnson C, Harmsen W, Offord K, Harris A, Wilson L, Ahlquist D. Colorectal cancer screening with CT colonography, colonoscopy, and double-contrast barium enema examination: prospective assessment of patient perceptions and preferences. Radiology. 2003; 227(2):378–384. [PubMed: 12732696]
4. Summers R, Beaulieu C, Pusanik L, Malley J, Jeffrey R, Glazer D, Napel S. Automated polyp detector for CT colonography: feasibility study. Radiology. 2000; 216(1):284–290. [PubMed: 10887263]
5. Yoshida H, Nappi J. Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. IEEE Trans Med Imaging. 2001; 20(12):1261–1274. [PubMed: 11811826]

6. Wang Z, Liang Z, Li L, Li X, Li B, Anderson J, Harrington D. Reduction of false positives by internal features for polyp detection in CT-based virtual colonoscopy. Med Phys. 2005; 32(12): 3602–3616. [PubMed: 16475759]
7. Zhu H, Fan Y, Lu H, Liang Z. Improving initial polyp candidate extraction for CT colonography. Phys Med Biol. 2010; 55:2087–2102. [PubMed: 20299733]
8. Zhu H, Fan Y, Lu H, Liang Z. Improved curvature estimation for computer-aided detection of colonic polyps in CT colonography. Acad Radiol. 2011; 18(8):1024–1034. [PubMed: 21652234]
9. Pickhardt P, Kim D. Colorectal cancer screening with CT colonography: key concepts regarding polyp prevalence, size, histology, morphology, and natural history. AJR Am J Roentgenol. 2009; 193(1):40–46. [PubMed: 19542393]
10. Kumar, V.; Abbas, A.; Fausto, N.; Aster, J.; Perkins, J. Polyps. In: SL, Robbins; RS, Cotran, editors. Pathologic basis of disease. 8th edn.. Philadelphia, PA: Saunders/Elsevier; 2010. ISBN 978-1-4160-3121-5
11. Winawer S, Zauber A. The advanced adenoma as the primary target of screening. Gastrointest Endosc Clin N Am. 2002; 12:1–9. [PubMed: 11916153]
12. Mills, S.; Carter, D.; Greenson, J.; Reuter, V.; Stoler, M., et al. Stern-berg's diagnostic surgical pathology. 5th edn.. Lippincott Williams & Wilkins; 2010.
13. Castellano G, Bonilha L, Li L, Cendes F. Texture analysis of medical images. Clin Radiol. 2004; 59(12):1061–1069. [PubMed: 15556588]
14. Haralick R, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern. 1973; 3(6):610–621.
15. Ji Q, Engel J, Craine E. Texture analysis for classification of cervix lesions. IEEE Trans Med Imaging. 2000; 19(11):1144–1149. [PubMed: 11204851]
16. Philips C, Li D, Raicu D, Furst J. Directional invariance of co-occurrence matrices within the liver. International conference on biocomputation, bioinformatics, and biomedical technologies. 2008
17. Fiori M, Musé P, Aguirre S, Sapiro G. Automatic colon polyp flagging via gometric and texture features. Conf Proc IEEE Eng Med Biol Soc. 2010; 1:3170–3173. [PubMed: 21096596]
18. Yao J, Dwyer A, Mollura D. Computer-aided diagnosis of pulmonary infections using texture analysis and support vector machine classification. Acad Radiol. 2011; 18(3):306–314. [PubMed: 21295734]
19. Rahmim A, Couhglin J, Gonzalez M, Endres C, Zhou Y, Wong D, Wahl R, Sossi V, Pomper M. Novel parametric PET image quantification using texture and shape analysis. Proceedings of the IEEE nuclear science symposium and medical imaging conference. 2012
20. Ng F, Ganeshan B, Kozarski R, Miles K, Goh V. Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. Radiology. 2013; 266(1):177–184. [PubMed: 23151829]
21. Suzuki K, et al. Computer-aided Diagnosis (CADx) for distinguishing neoplastic from non-neoplastic lesions toward CT colonography (CTC) beyond detection. Program of scientific assembly and annual meeting of radiological society of North America (RSNA), LL-GIS-TU5C. 2012 Nov.
22. Zhang G, Song B, Zhu H, Liang Z. Computer-aided diagnosis in CT colonography based on bi-labeled classifier. The 26th international congress and exhibition on computer assisted radiology and surgery (CARS). International Journal of CARS. 2012; 7(Suppl):S274.
23. Engel, K. Real-time volume graphics. Wellesley, MA: A K Peters; 2006.
24. Monga O, Benayoun S. Using partial derivatives of 3D images to extract typical surface features. Comput Vis Image Underst. 1995; 61(2):171–189.
25. van Wijk C, van Ravesteijn V, Vos F, van Vliet L. Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow. IEEE Trans Med Imaging. 2010; 29(3): 688–698. [PubMed: 20199908]
26. Deriche R. Using Canny's criteria to derive a recursively implemented optimal edge detector. Int J Comput Vis. 1987; 1(2):167–187.
27. Haralick R, Shapiro L. Computer and robot vision. Addison-Wesley Publishing Company, Reading, MA. 1993; 1:453–508.

28. Tesar L, Smutek D. Bayesian classification of sonograms of thyroid gland based on Gaussian mixtures. Proceedings of Norvegian conference on image processing and pattern recognition, NOBIM. 2004:36–40.
29. Tesar L, Smutek D, Shimizu A, Kobatake H. 3D extension of Haralick texture features for medical image analysis. SPPR 2007 proceedings of the fourth conference on IASTED international conference. 2007:350–355.
30. Tesar L, Shimizu A, Smutek D, Kobatake H, Nawano S. Medical image analysis of 3D CT images based on extension of Haralick texture features. Comput Med Imaging Graph. 2008; 32(6):513–520. [PubMed: 18614335]
31. Liang Z, Cohen H, Posniak E, Fiore E, Wang Z, Li B, Anderson J, Harrington D. Texture-based CAD improves diagnosis for low-dose CT colonography. Proceedings of the SPIE medical imaging CD-ROM. 2008
32. Showalter C, Clymer B, Richmond B, Powell K. Three-dimensional texture analysis of cancellous bone cores evaluated at clinical CT resolutions. Osteoporos Int. 2006; 17:259–266. [PubMed: 16170445]
33. Pickhardt P, Kim D, Pooler B, Hinshaw J, Barlow D, Jensen D, Reichelderfer M, Cash B. Assessment of volumetric growth rates of small colorectal polyps with CT colonography: a longitudinal study of natural history. Lancet Oncol. 2013; 14(8):711–720. [PubMed: 23746988]
34. Li L, Chen D, Lakare S, Kreeger K, Bitter I, Kaufman A, Wax M, Djuric P, Liang Z. An image segmentation approach to extract colon lumen through colonic material tagging and hidden Markov random field model for virtual colonography. Proceedings of the SPIE medical imaging. 2002; 4683:406–411.
35. Wang S, Li L, Cohen H, Mankes S, Chen J, Liang Z. An EM approach to MAP solution of segmenting tissue mixture percentages with application to CT-based virtual colonoscopy. Med Phys. 2008; 35(12):5787–5798. [PubMed: 19175136]
36. Pickhardt P, Choi J, Hwang I, Butler J, Puckett M, Hildebrandt H, Wong R, Nugent P, Mysliwiec P, Schindler W. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med. 2003; 349:2191–2200. [PubMed: 14657426]
37. American College of Radiology. ACR practice guideline for the performance of computed tomography (CT) colonography in adults. ACR Pract Guidel. 2005; 29:295–298.
38. Chang C, Lin C. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol. 2011; 2(27):1–27.
39. Song B, Zhang G, Zhu W, Liang Z. ROC operating point selection for classification of imbalanced data with application to computer-aided polyp detection in CT colonography. Int J Comput Assist Radiol Surg. 2014; 9(1):79–89. [PubMed: 23797823]
40. Fidler J, Johnson C, MacCarty R, Welch T, Hara A, Harmsen W. Detection of flat lesions in the colon with CT colonography. Abdom Imaging. 2002; 27:292–300. [PubMed: 12173360]
41. Pickhardt P, Choi J, Hwang I, Schindler W. Nonadenomatous polyps at CT colonography: prevalence, size distribution, and detection rates. Radiology. 2004; 232:784–790. [PubMed: 15247435]
42. Pickhardt P, Kim D, Robbins J. Flat (nonpolypoid) colorectal lesions identified at CT colonography in a US screening population. Acad Radiol. 2010; 17:784–90. [PubMed: 20227304]
43. Shah J, Hynan L, Rockey D. Management of small polyps detected by screening CT colonography: patient and physician preferences. Am J Med. 2009; 122(7):687. e1–9. [PubMed: 19559172]
44. Summers R. Polyp size measurement at CT Colonography: what do we know and what do we need to know? Radiology. 2010; 255(3):707–720. [PubMed: 20501711]
45. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett. 2006; 27:861–874.
46. Suzuki, K., et al. Observer performance study: effect of computer-aided diagnosis (CADx) on the performance of radiologists in distinguishing neoplastic from non-neoplastic lesions in CT colonography (CTC); Program of scientific assembly and annual meeting of radiological society of North America (RANA), SSA 07–08; 2012 Nov.

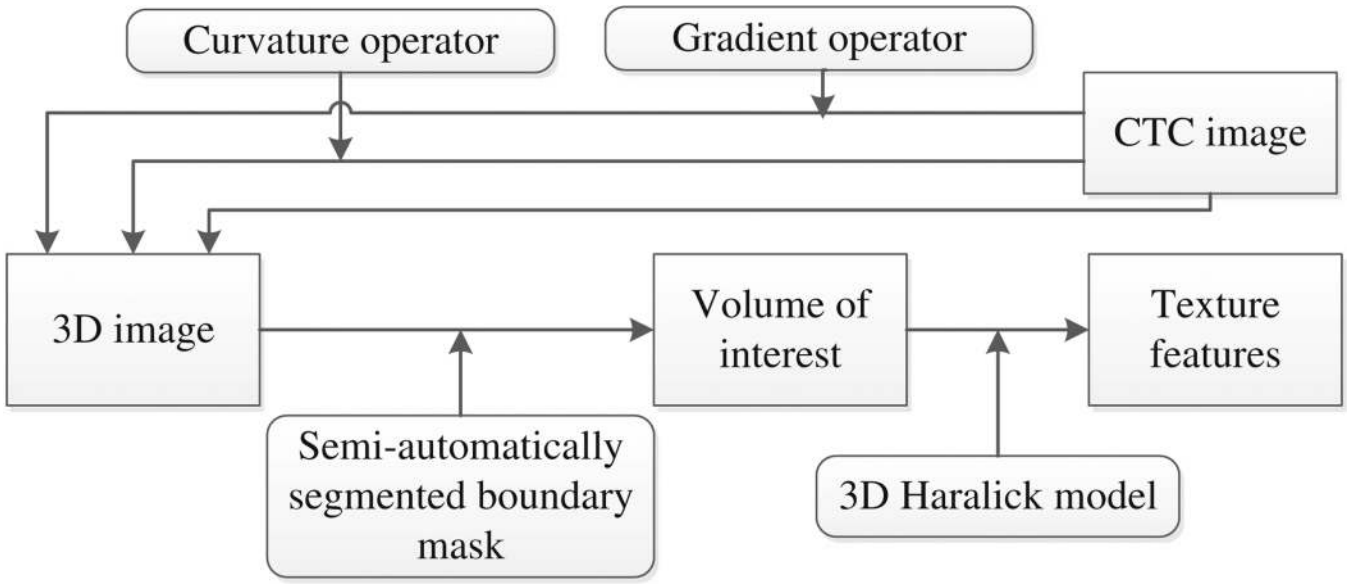

Fig. 1.
Overview of our proposed 3D texture feature extraction scheme

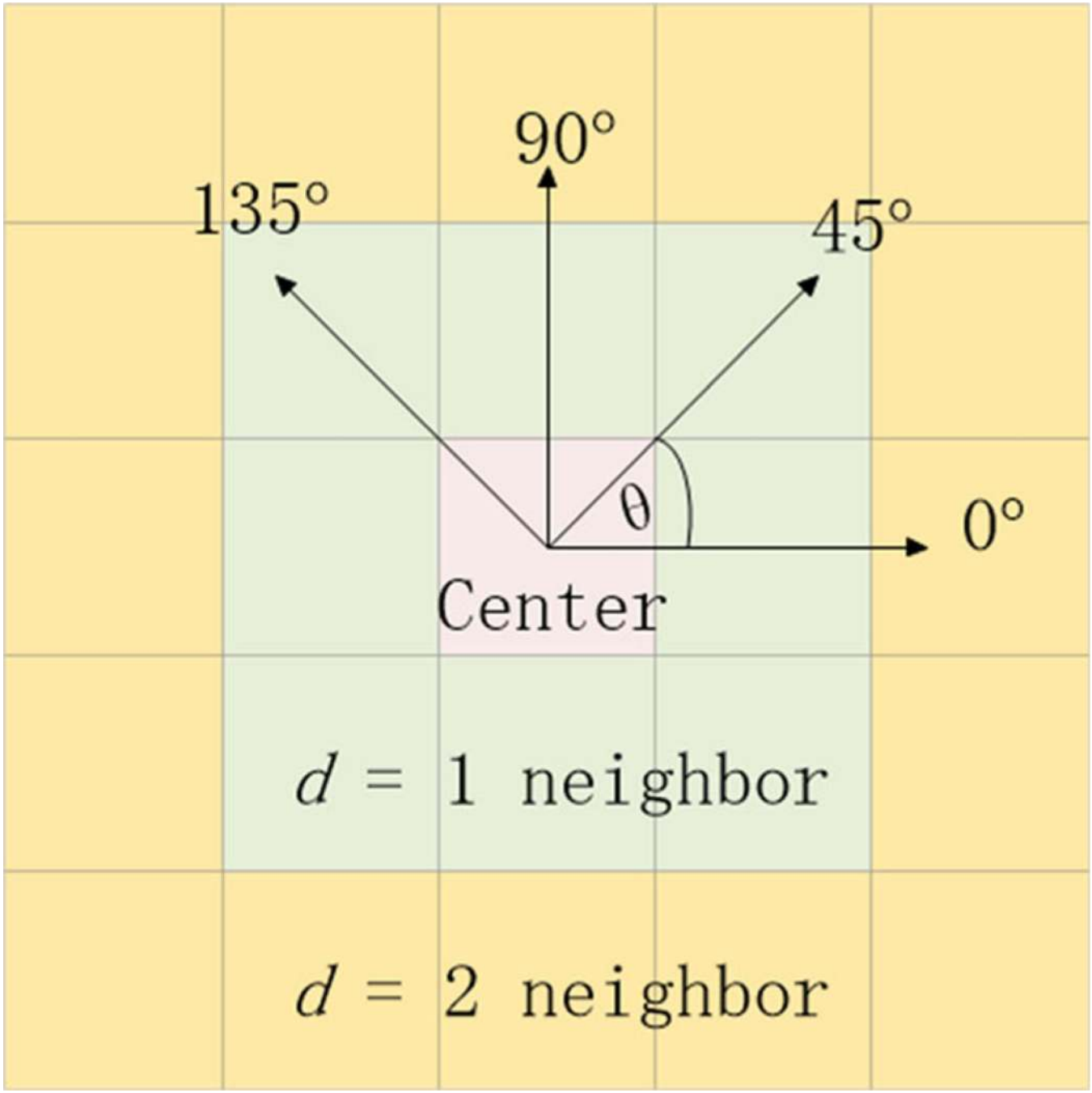

Fig. 2.
Illustration of 2D Haralick spatial information extraction model with $d = 1$ and $d = 2$ neighbor and four directions

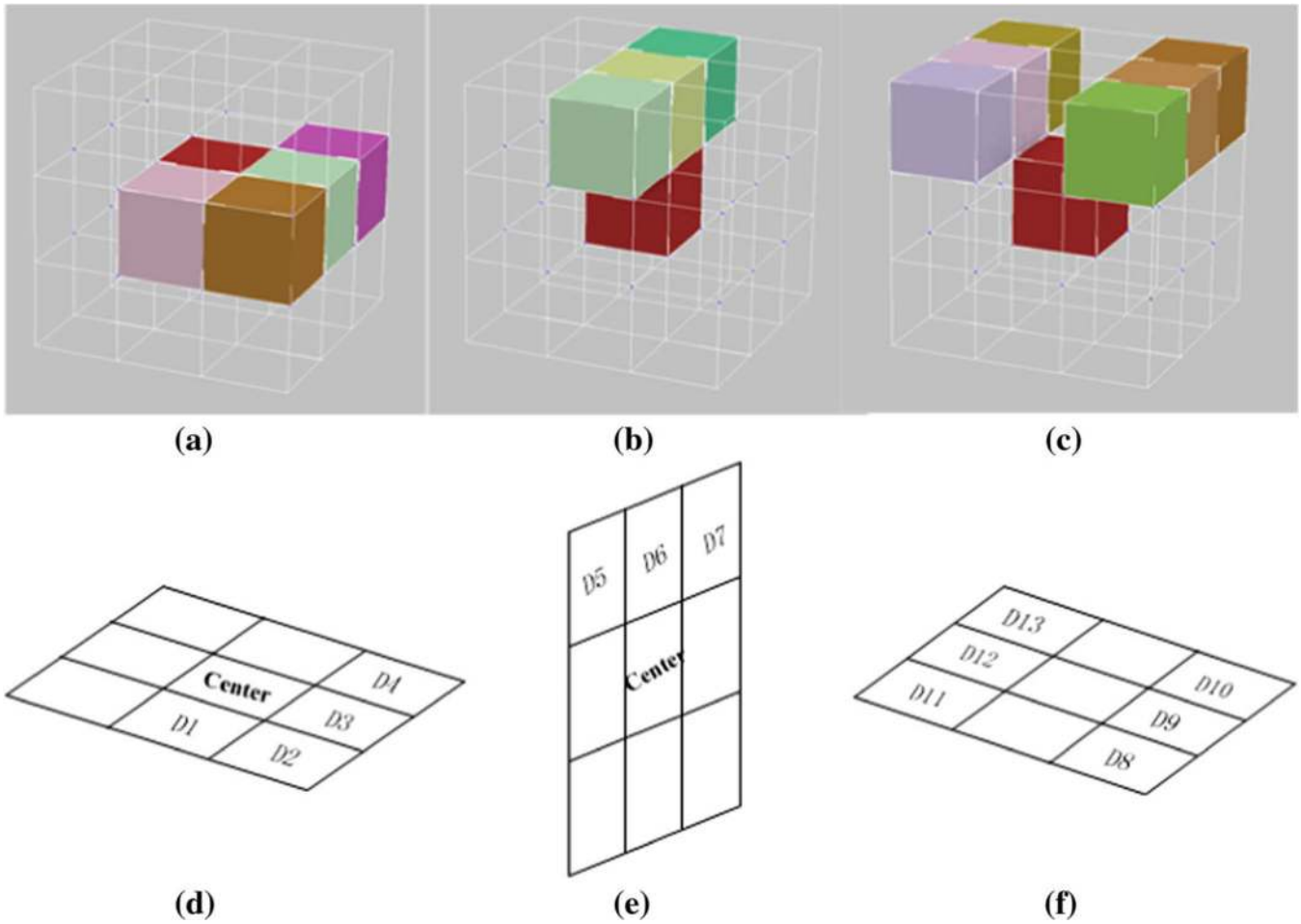

Fig. 3.

Illustration of 3D expansion of the Haralick model. The 3D frame shows the 26 neighbors (*d* = 1) of the *red cube*. Graph **a** shows the four directions in the *middle layer*. Graphs **b, c** show the rest nine directions in *top layer*. Graphs **d–f** show the direction number corresponding to **a–c**

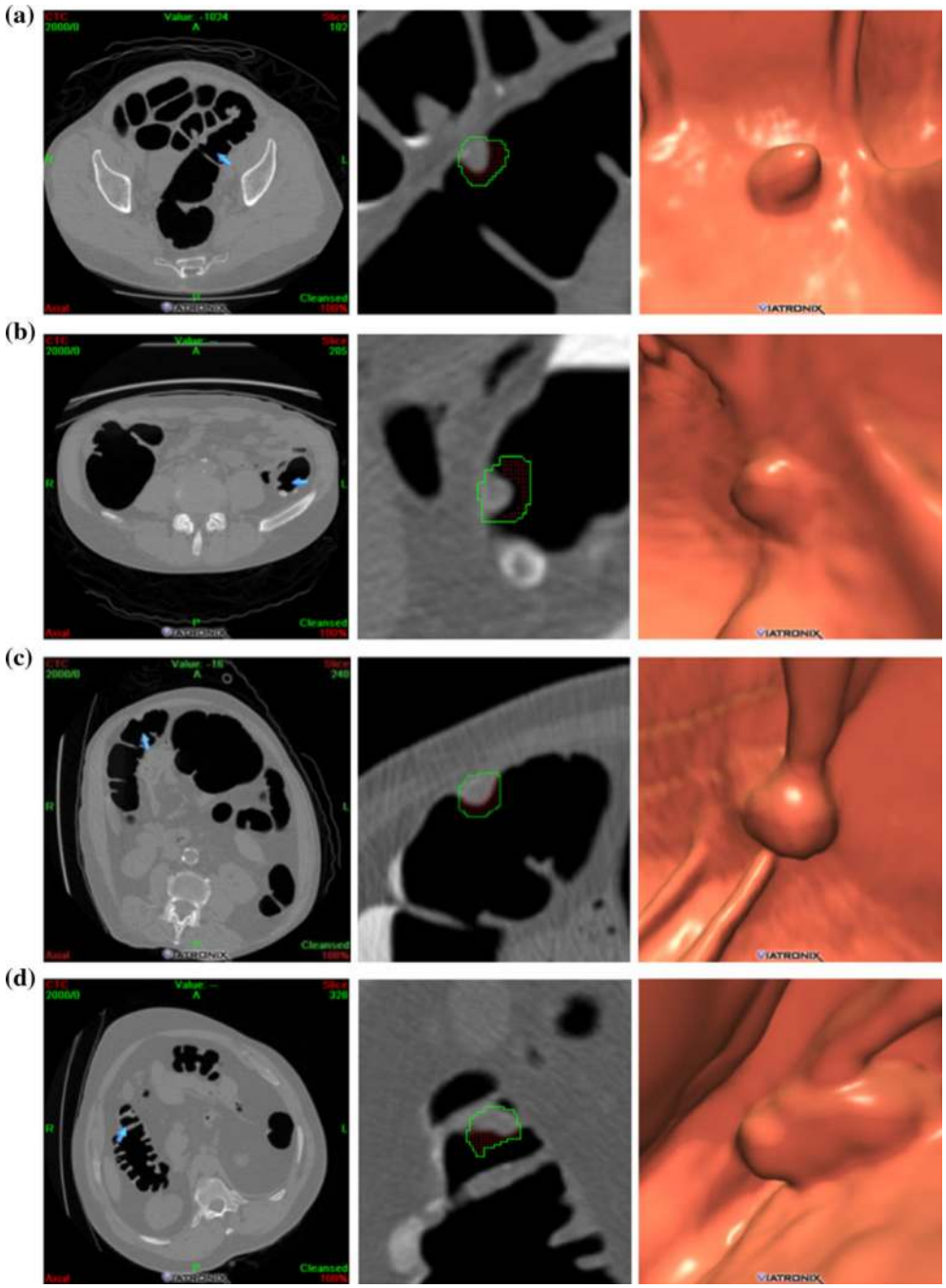

Fig. 4.

Illustration of semi-automatic volume extraction. The *first column* shows the locations of lesions in 2D slices. The *second column* shows the semi-automatically extracted region of interest in 2D view. The *green line* indicates the manually drawn boundary, and the *red border* indicates the interface between air and polyp after automatic air removal. The *third column* shows the 3D endoscopic views using the software of Viatronix. Row **a** shows a 7mm tubular adenoma located in sigmoid colon. Row **b** shows a 8mm hyperplastic polyp

located in descending colon. Row **c** shows a 9mm tubular adenoma located in hepatic flexure. Row **d** shows a 20mm tubulovillous adenoma in ascending colon

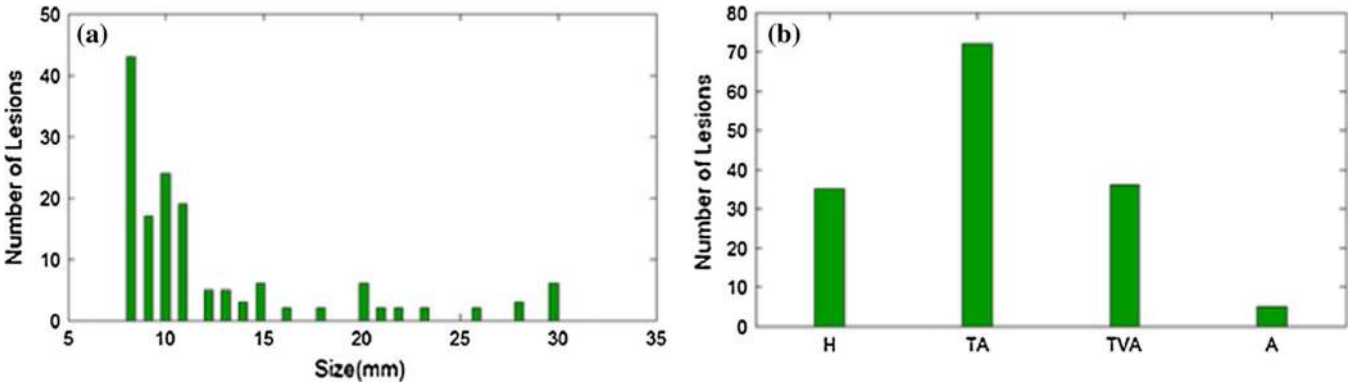

Fig. 5.

Distribution information of the database used in study. Graph **a** shows the histogram of linear size information in the pathology report. Graph **b** is the histogram of pathological phase of the 148 lesions

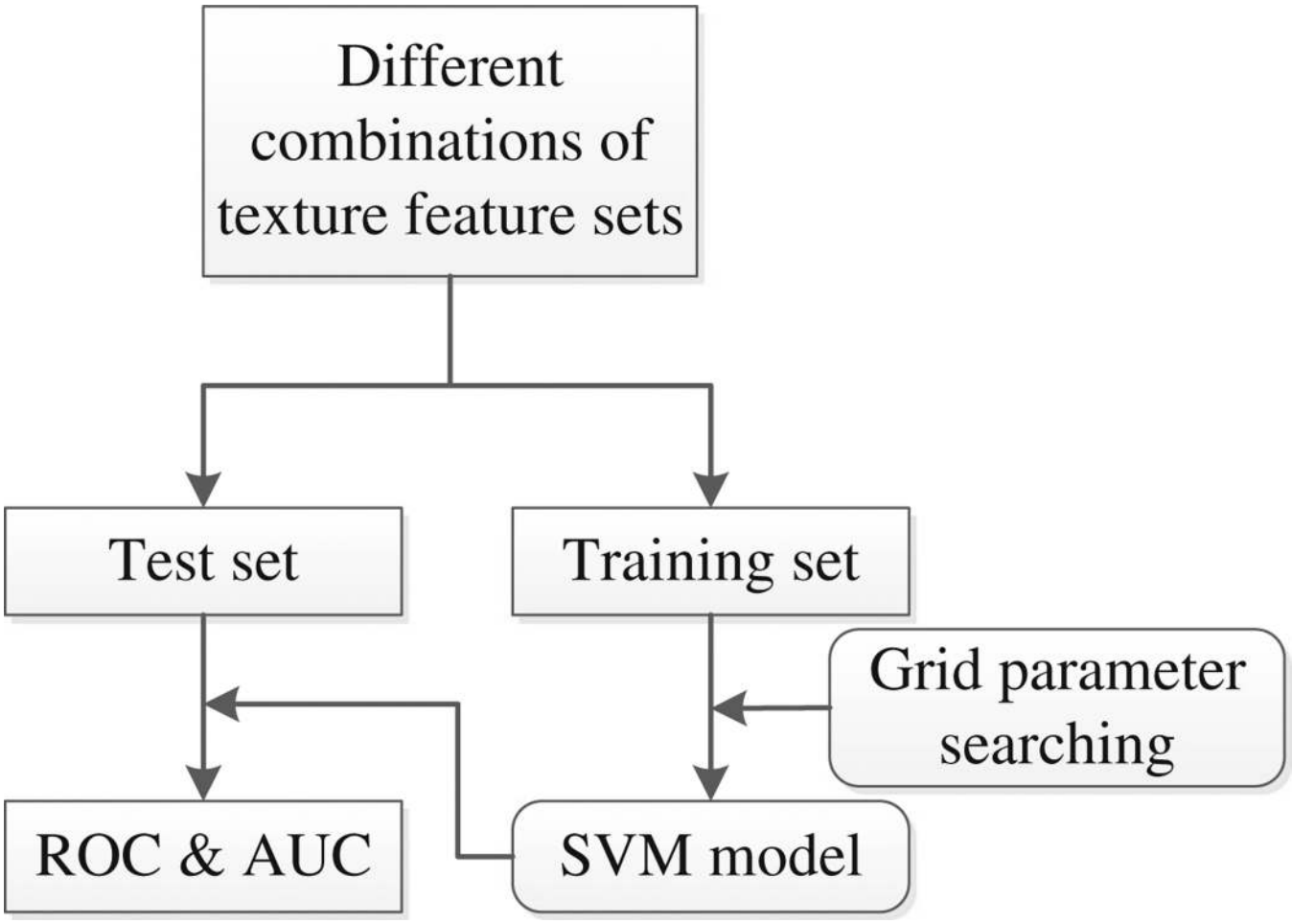

Fig. 6.
Overview of our texture feature set evaluation scheme

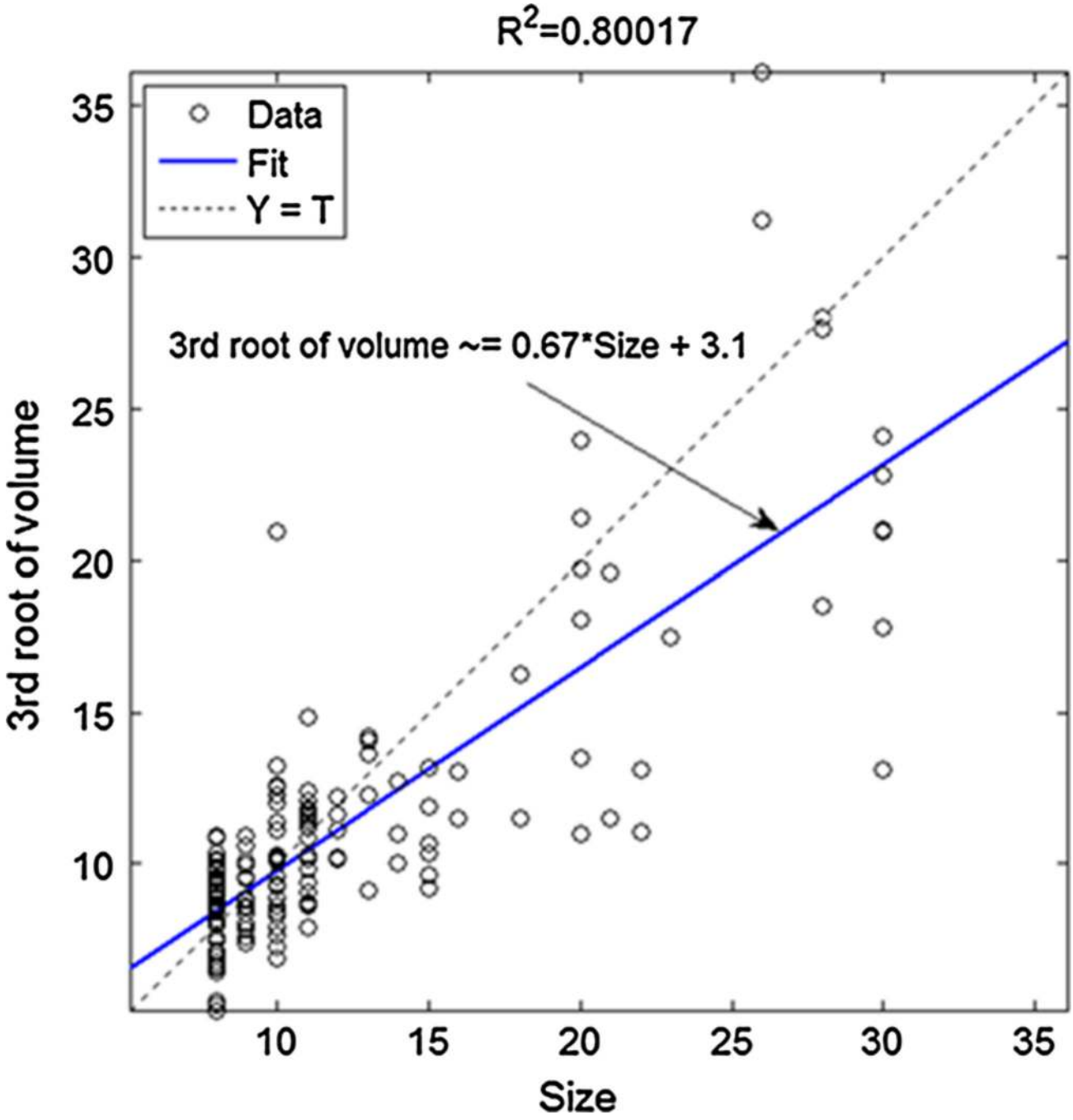

Fig. 7.

Regression scatter plot shows the linear relationship between the third root of extracted 3D volume voxel number and its corresponding linear size in the pathology report

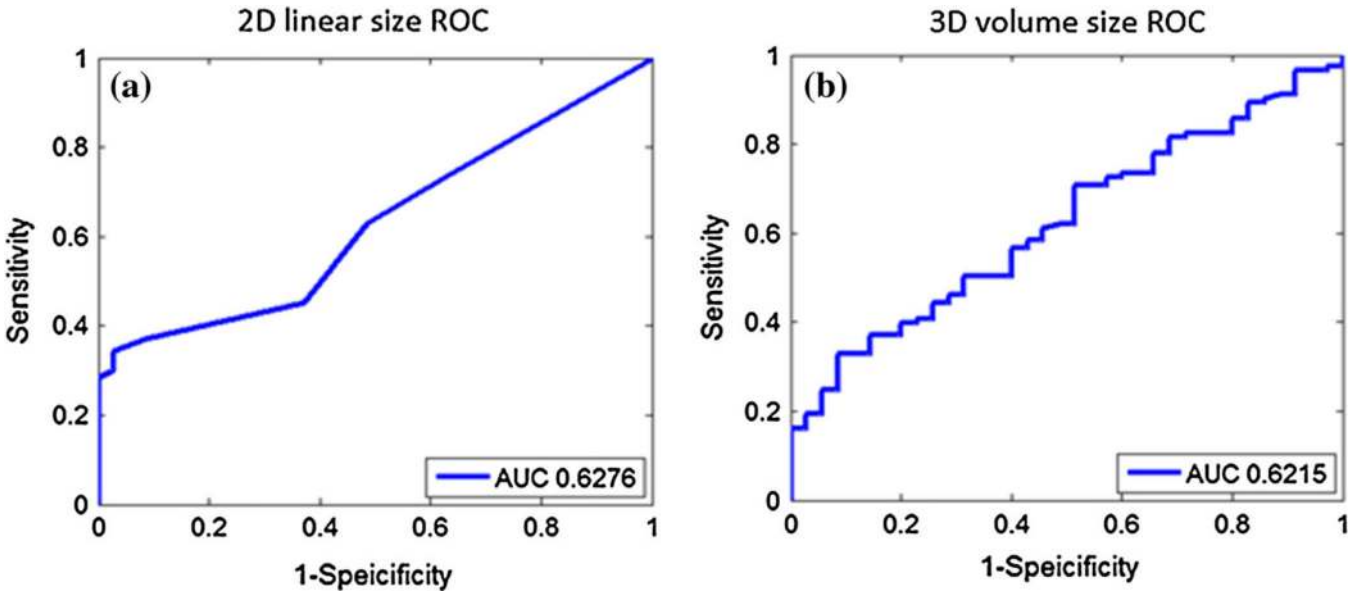

Fig. 8.

ROC curve with only lesion size information as decision value. Graph **a** shows the ROC curve based only on linear size differentiation. Graph **b** shows the ROC curve based only on 3D volume size differentiation

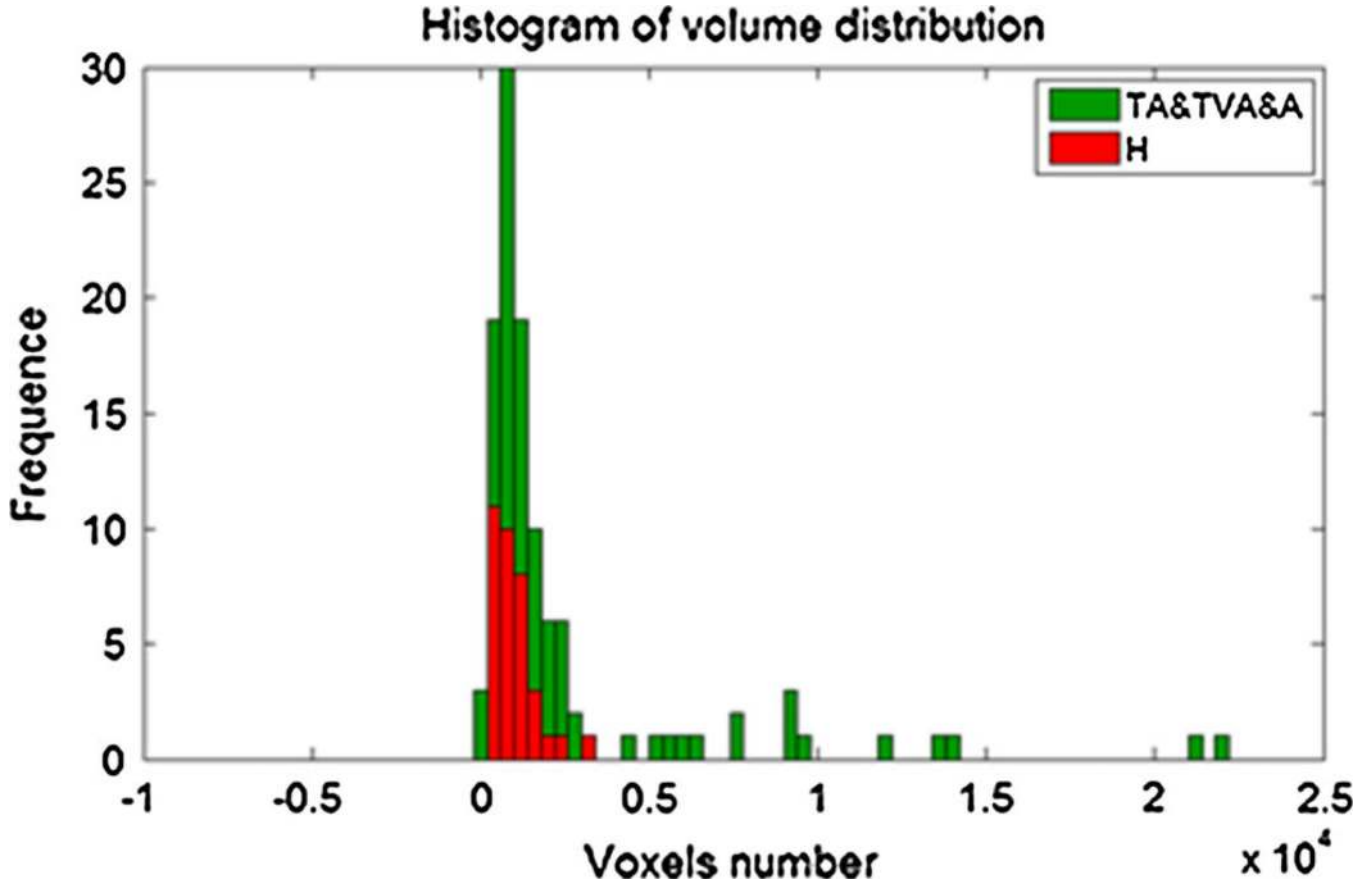

Fig. 9.

Histogram of 3D volume size distribution within different groups, i.e., H against $TA\&TVA\&A$

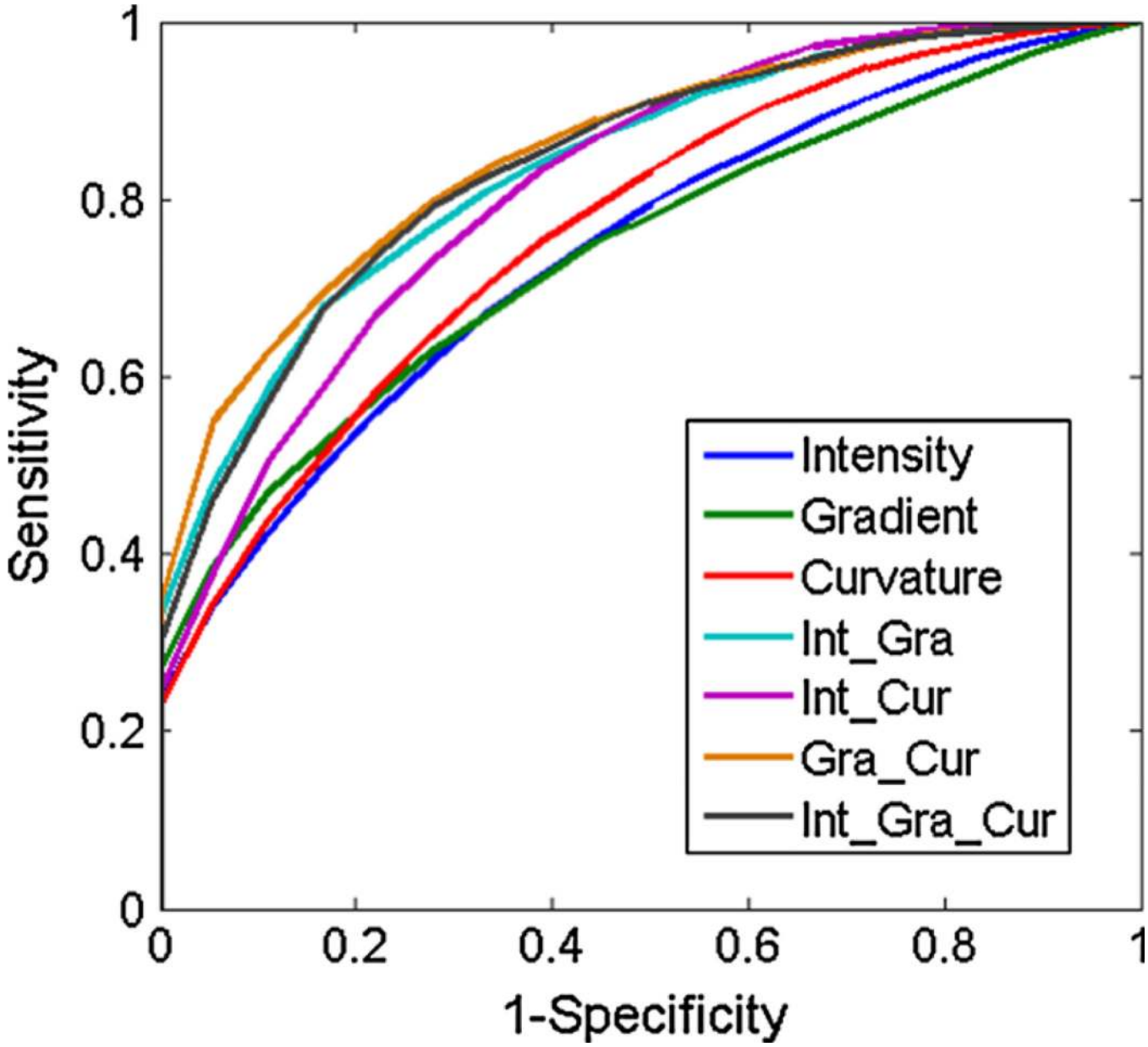

Fig. 10.

Averaged ROC curves from SVM classification results. The *curves* were conducted according to the *horizontal axis*, where the linear interpolation was employed when needed

Table 1

Haralick features extracted from co-occurrence matrix (CM)

Feature number	Description
f_1	Angular second moment
f_2	Contrast
f_3	Correlation
f_4	Sum of squares: variance
f_5	Inverse difference moment
f_6	Sum average
f_7	Sum variance
f_8	Sum entropy
f_9	Entropy
f_{10}	Difference variance
f_{11}	Difference entropy
$f_{12,13}$	Information measures of correlation

Table 2

Averaged SVM AUC information of the 100 runs for both 2D and 3D features

Group	AUC information		*P* value
	2D	3D	
Intensity	0.6210 ± 0.0950	0.7403 ± 0.0529	<<0.001 *
Gradient	0.6823 ± 0.0905	0.7407 ± 0.0493	<<0.001 *
Curvature	0.5827 ± 0.0860	0.7652 ± 0.0543	<<0.001 *
Int–Gra	0.7263 ± 0.0944	0.8368 ± 0.0392	<<0.001 *
Int–Cur	0.7187 ± 0.0668	0.8161 ± 0.0511	<<0.001
Gra–Cur	0.7318 ± 0.0500	0.8525 ± 0.0431	<<0.001
Int–Gra–Cur	0.7413 ± 0.0533	0.8399 ± 0.0427	<<0.001 *

Format: mean ± standard deviation

* indicates the normality assumption for paired *t* test is not hold, in which case we use Wilcoxon signed-rank test instead. Int, Gra and Cur are the abbreviation for intensity, gradient and curvature, respectively

Table 3

Significance test of difference between 3D feature set GraCur and the rest feature sets

	Intensity	Gradient	Curvature	Int–Gra	Int–Cur	Int–Gra–Cur
Gra–Cur	<<0.001	<<0.001	<<0.001 *	<0.001	<<0.001	0.0032

* indicates the normality assumption for paired *t* test is not hold, in which case we use Wilcoxon signed-rank test instead. Int, Gra and Cur are the abbreviation for intensity, gradient and curvature, respectively

Int J Comput Assist Radiol Surg. Author manuscript; available in PMC 2015 October 03.

Table 4

Sensitivity-specificity pairs on the averaged ROC curve

Group	Corresponding specificity with different sensitivity level			
	Sen* = 0.6	Sen=0.7	Sen=0.8	Sen=0.9
Intensity	0.7384	0.6284	0.4910	0.3043
Gradient	0.7535	0.6254	0.4630	0.2523
Curvature	0.7640	0.6681	0.5433	0.3895
Int–Gra	0.8808	0.8060	0.6773	0.4834
Int–Cur	0.8241	0.7507	0.6469	0.4980
Gra–Cur	0.9087	0.8264	0.7178	0.5251
Int–Gra–Cur	0.8732	0.8101	0.7070	0.5188

*"Sen" is the abbreviation of "sensitivity". Int, Gra and Cur are the abbreviation for intensity, gradient and curvature, respectively