

Computer recognition of cancer in the urinary bladder using optical coherence tomography and texture analysis

Colleen A. Lingley-Papadopoulos

The George Washington University
Department of Electrical and Computer Engineering
Staughton 107, 707 22nd Street NW
Washington, DC 20052

Murray H. Loew

The George Washington University
Department of Electrical and Computer Engineering
Staughton 101, 707 22nd Street NW
Washington, DC 20052

Michael J. Manyak

Cytogen Corporation
Suite 3100, 650 College Road East
Princeton, New Jersey 08540

Jason M. Zara

The George Washington University
Department of Electrical and Computer Engineering
Staughton 107, 707 22nd Street NW
Washington, DC 20052

Abstract. The vast majority of bladder cancers originate within 600 μm of the tissue surface, making optical coherence tomography (OCT) a potentially powerful tool for recognizing cancers that are not easily visible with current techniques. OCT is a new technology, however, and surgeons are not familiar with the resulting images. Technology able to analyze and provide diagnoses based on OCT images would improve the clinical utility of OCT systems. We present an automated algorithm that uses texture analysis to detect bladder cancer from OCT images. Our algorithm was applied to 182 OCT images of bladder tissue, taken from 68 distinct areas and 21 patients, to classify the images as noncancerous, dysplasia, carcinoma *in situ* (CIS), or papillary lesions, and to determine tumor invasion. The results, when compared with the corresponding pathology, indicate that the algorithm is effective at differentiating cancerous from noncancerous tissue with a sensitivity of 92% and a specificity of 62%. With further research to improve discrimination between cancer types and recognition of false positives, it may be possible to use OCT to guide endoscopic biopsies toward tissue likely to contain cancer and to avoid unnecessary biopsies of normal tissue. © 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2904987]

Keywords: bladder cancer; optical coherence tomography; texture analysis; computer-aided diagnosis.

Paper 07179RR received May 18, 2007; revised manuscript received Nov. 2, 2007; accepted for publication Nov. 10, 2007; published online Apr. 14, 2008.

1 Introduction

The American Cancer Society estimates that in 2007 there will be 67,160 new cases and an estimated 13,750 deaths from bladder cancer in the United States. Bladder cancer is the fourth-most-common cancer in American men.¹ The treatment outcomes and survivability of bladder cancer improve significantly with early detection. Currently, bladder cancer is diagnosed and treated through the use of endoscopic visualization techniques, frequently video based, that direct the urologist only to visible surface abnormalities. However, some cancerous and precancerous epithelial lesions are either not visible with conventional visual inspection or the diagnosis is inconclusive.² Biopsies of suspicious areas of tissue are necessary to diagnose these conditions accurately, but with current methods the afflicted areas may be overlooked and biopsies performed in the wrong location. Furthermore, multiple biopsies of normal appearing areas are frequently necessary to rule out the presence of an occult high grade malignancy known as carcinoma *in situ*³ (CIS). This adds to the expense and potential for complications from the evaluation.

The ability to visualize subsurface structures at high resolution and to evaluate changes in optical properties of mucosal tissues would assist in diagnosing conditions unidentifiable with conventional visualization techniques. Optical coherence tomography (OCT) has the potential to provide this crucial information. OCT is an optical imaging technique analogous to ultrasound that uses partially coherent near-IR light to interrogate a target and create images of subsurface microscopic structures with a resolution of 15 μm or less.⁴ The origin of the received backscattered light is detected with low-coherence interferometry, so a map of reflectivity versus optical depth and lateral position can be created. OCT has been shown to produce images with high spatial resolution but, due to the high level of scattering of near-IR light in biological tissues, penetration depths range from only 1 to 2 mm.⁵ Mucosal cancers, however, such as cancer of the bladder, tend to arise in the urothelium within 600 μm of the tissue surface, which is an ideal imaging depth for endoscopic OCT imaging systems.

Because OCT has the potential to produce high-resolution images of subsurface structures at near-video rates, one of its most promising possibilities is to serve as a guide when determining the correct location to perform a biopsy. For this possibility to become a reality, it is essential to develop meth-

Address all correspondence to Colleen Lingley-Papadopoulos, Electrical and Computer Engineering, George Washington University, Staughton 107-707, 22nd St. NW Washington, DC 20052; Tel: 703-944-7772; Fax: 202-994-0280; E-mail: alingley@gwu.edu

ods of detecting cancerous and precancerous conditions with OCT images during the time of endoscopic inspection. This will enable the system to indicate in real time the presence of suspect tissue at or beneath the surface. In this paper, we introduce an algorithm to distinguish OCT images of cancerous tissue from noncancerous tissue. The algorithm, based on texture analysis, in the future could be integrated with an OCT system to provide real-time guidance to surgeons.

While this paper introduces an algorithm for the detection of cancers in the mucosa of the urinary bladder, the techniques might also be applicable to other epithelial tissues, including the lining of the gastrointestinal tract, the oral cavity, the tracheobronchial system, and the genitourinary tract.

2 Background

2.1 Bladder Anatomy

The urinary bladder consists of three layers: the epithelium (commonly called the urothelium) made up of transitional cells, a connective tissue layer known as the lamina propria, and the muscularis comprised of three concentric layers of muscle. In healthy tissue, these layers are well organized and well defined. The urothelium is the very thin innermost layer of the bladder wall, bound by the lamina propria with muscle the outermost tissue.

The primary growth patterns for urothelial neoplasms, which account for 98% of primary tumors of the bladder,⁶ are papillary lesions, CIS, and flat lesions. Papillary lesions are exophytic growths resembling a sea anemone, projecting into the bladder with multiple fronds. CIS is a diffuse, high-grade, intraepithelial malignancy. The much more uncommon flat lesions are heaped up epithelial tumors without the characteristic stalk of papillary lesions. Any of these tumors can become invasive by infiltrating the lamina propria and more deeply into the muscle.⁷ Of these lesions, only papillary lesions are readily visible endoscopically. In addition, with visual inspection it is not possible to determine if any of these lesions have become invasive.

Thickening of the urothelial layer by additional normal cells is known as hyperplasia and is usually associated with previous trauma or infection. Inflammation is classified as exudative or infiltrative, depending on the anatomic depth and type of cells involved in the inflammatory process. Cellular characteristics distinguish normal from malignant tissue with a spectrum of abnormal to precancerous conditions known as dysplasia between the two extremes. Dysplasia ranges from mild to severe with severe dysplasia often indistinguishable from CIS. While a thickened bladder epithelium may signify cancer, the cellular characteristics define its biological activity.⁸

2.2 OCT Imaging of the Bladder

Feldchtein et al.⁹ demonstrated endoscopic use of OCT to examine the mucosal membranes of several internal organs including the bladder. They were able to image the discrete layers of the bladder, along with blood vessels and cysts located within these layers. In healthy tissue, the urothelium (U) appears in OCT images as an area of low intensity, while the lamina propria (LP) appears as an area of high intensity and the muscle (MS) appears as an area of low intensity thereby

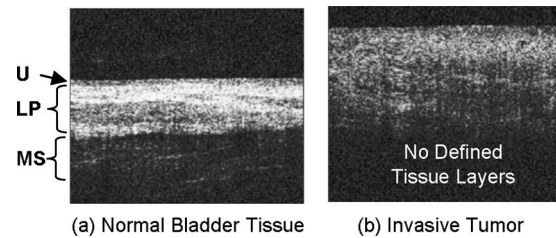


Fig. 1 OCT images of normal and cancerous bladder tissue, dimensions 1.8 by 1.5 mm, showing characteristics recognized by (a) Feldchtein et al.⁹ and (b) Jesser et al.¹⁰ Organized layers can be seen in the image of normal bladder tissue, while clearly defined layers cannot be seen in the image of the invasive tumor. The images were taken during a study at the George Washington University Medical Center¹¹ (GWUMC).

providing contrast between layers. Figure 1(a) shows an OCT image of the lining of a healthy bladder with well defined layers.

Jesser et al.¹⁰ used *in vitro* specimens to demonstrate the possibility of using OCT to distinguish between normal human bladders and those with invasive transitional cell carcinoma (TCC). Jesser noted that, unlike OCT images of the normal bladder, images of invasive TCC did not contain distinct layers or boundaries, and concluded that malignant invasion disturbed the normal well-defined strata seen in OCT images. Figure 1(b) shows an endoscopic OCT image of the bladder with an invasive tumor in which the layers cannot be distinguished.

In a study by Zagayanova et al.,¹² OCT scans from 63 patients were evaluated using OCT and compared with the pathological diagnosis obtained from biopsy. The OCT images were evaluated as being either malignant or benign, with the absence of a layered structure being the primary indicator of malignancy. The overall results had a sensitivity of 98% and a specificity of 72%. During a similar study conducted at our institution which also evaluated invasion,¹¹ 24 patients underwent cystoscopic examination of the bladder followed by OCT scan, photography, and biopsy from at least six mucosal areas. The OCT scans were reviewed and declared to be healthy, abnormal but not invasive, or invasive, and compared later with the pathological biopsy results. The results indicated 100% sensitivity and an overall specificity of 89%, with a negative predictive value of 100%. The algorithm introduced in this paper was tested on a subset of the images obtained during that study.

The preceding studies used human observers to analyze OCT images and assign diagnoses, which make the results highly dependent on the training and ability of the observers. A computer algorithm capable of analyzing and diagnosing the images would be operator-independent and make OCT diagnostics interpretable without significant training.

2.3 Texture Analysis of OCT Images

Texture analysis is an image-processing technique that describes an image or portion of an image by characterizing its structure and pattern. Texture is made up of texture primitives. To describe texture it is necessary to describe the tone, or pixel intensity properties within the primitives, as well as the structure, which describes the spatial relationship between

primitives. There are three principal approaches to texture analysis: statistical, structural, and spectral. Statistical analysis describes information about the content of the primitives, while structural describes the arrangement of the primitives, and spectral analysis describes the periodicity of the primitives.¹³

Gossage and Tkaczyk¹⁴ demonstrated that texture analysis can be used to analyze OCT images in an attempt to classify different tissue types. The study, which was able to distinguish between *in vitro* images of mouse skin (correct classification rate of 98.5%) and testicular fat (97.3%), as well as normal lung (88.6%) and abnormal lung (64.0%), analyzed the texture due to speckle in OCT images.

Although texture analysis has not been previously applied to the bladder, there have been studies that have used texture analysis to recognize dysplasia or cancer in other tissues. In two separate studies Qi et al.^{15,16} applied texture analysis along with other image analysis techniques to OCT images of the esophagus to diagnose dysplasia. The resulting sensitivities were 87 and 82%, and the specificities were 69 and 74%. Likewise, image analysis was applied to the recognition of breast cancer by Zysk and Boppart.¹⁷ The results of the study indicated that their algorithm had a tumor tissue sensitivity of 97% and a specificity of 68%.

These studies indicate that texture analysis of OCT images is useful in distinguishing between types of tissue, but none of the studies involved images of the bladder. As far as we know, our study is the first in which texture analysis has been applied to OCT images of the lining of the urinary bladder.

3 Method

Deidentified data from 21 patients at high risk of having TCC of the urinary bladder, who underwent cystoscopic examination with the OCT protocol in our previous study, were used in this work.¹¹ During the previous study, scanning was performed with a 980-nm, 10-mW superluminescent diode using a 2.7-mm (OD) optical fiber positioned through a cystoscope sheath. Patients underwent a standard cystoscopic examination. Visually suspect lesions, as well as normal-appearing urothelial tissue, were photographed, scanned with OCT, and biopsied. Multiple scans were taken in each area, but at different sites within the area. The scans of 1.5-s duration, which generated 200×200 -pixel images, were performed at 1-mm intervals on the lesions and at their junctions with the bladder epithelium. All scans were obtained by placing the end-firing OCT probe on the desired site perpendicular to the wall of the bladder. Each patient had at least one apparently normal area that was photographed, scanned with OCT, and biopsied. Biopsy specimens were preserved in formalin for standard histopathologic analysis. The endoscopic scanning probe had a depth range of 2.2 mm in air, a lateral scanning range of 1.6 to 2.4 mm, and a working distance from the probe surface of 0.5 mm. The system, coupled with the 2.7-mm-diameter probe, had a lateral resolution of $\leq 50 \mu\text{m}$ (focused beam waist diameter $25 \mu\text{m}$) and axial (depth) resolution of 10 to $20 \mu\text{m}$ in air.

A total of 182 OCT images from the study, along with their corresponding pathology results, were used as the training set for algorithm development. The 182 images include scans diagnosed as healthy, exudative inflammation, infiltrative in-

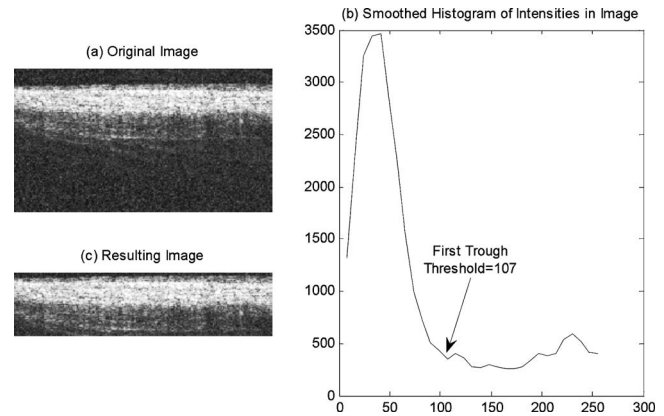


Fig. 2 To recognize the portion of the image (a) containing the most information, a smoothed histogram of the intensities in the image was created (b), and the image was thresholded (c).

flammation, dysplasia, CIS, papillary lesion, or invasive tumor. Lesions of any type that had become invasive were diagnosed as invasive tumors regardless of the type of lesion. Only diagnoses for which at least six images were available were included in this study.

The algorithm developed here has three stages: (1) preprocessing, in which the portion of the image containing the bladder lining is identified and in which the dc bias is removed; (2) processing, in which texture analysis is used to determine a set of texture features for the portion of the image containing information; and (3) classification, in which the texture features are used to traverse a decision tree and produce a diagnosis.

The 182 scans were used to identify the features used in the decision tree, and then as the training set for the algorithm. Due to the limited data currently available, the algorithm was tested using leave-one-out cross validation, rather than on an independent set.

3.1 Image Preprocessing

Our algorithm was designed to characterize images based on a training set to identify features that will be useful in making a diagnosis. To remove as much variation as possible between the images, the dc bias and any existing background area were removed from the images before analysis.

Histogram analysis was used to identify the background in each image. Due to the low intensities associated with the lumen and bottom portion of the OCT image, the histograms produced a bimodal distribution, as seen in Fig. 2(b). The first trough after the large low-intensity area was used to identify a threshold intensity to separate the background from the portion of the image containing information about the bladder lining. Once the threshold was determined, the portions of the image at the top and bottom of the image containing only intensities that were entirely below the threshold were removed from the image.

Once the background was removed from the image, the mean intensity of the removed background was calculated. Since the background would theoretically be of regions in which there should be no signal (such as the lumen), or at the bottom of the image where little or no signal was measured,

the mean intensity of the background would be a reliable estimate of the dc bias of the system. Consequently, the mean intensity of the removed background was used as an estimate of the dc bias and subtracted from the portion of the image containing information. Since the portion of the image containing information had higher intensities than the background, subtracting the mean intensity of the background did not and will not cause intensity levels to drop below zero.

3.2 Texture Analysis

To develop our classification algorithm, we used five methods representing the three approaches to texture analysis (statistical, structural, and spectral) to obtain 74 texture features for each image in our training set.

The first method used cooccurrence matrices to analyze texture, which is a statistical method based on the repetitive nature of various intensity levels. The method requires that a neighbor be defined and a cooccurrence matrix created using that definition. The resulting cooccurrence matrix can be used to calculate six texture characteristics, where $P_{\phi,d}(a,b)$ is the probability that a pixel with intensity a will have a neighbor of intensity b ; where neighbor is defined as the pixel at distance d in the direction ϕ ; and μ_x , μ_y , σ_x , and σ_y are the means and standard deviations of the cooccurrence matrix¹³:

$$\text{energy} = \sum_{a,b} P_{\phi,d}^2(a,b), \tag{1}$$

$$\text{entropy} = \sum_{a,b} P_{\phi,d}(a,b) \log_2 P_{\phi,d}(a,b), \tag{2}$$

$$\text{maximum probability} = \max_{a,b} P_{\phi,d}(a,b), \tag{3}$$

$$\text{contrast} = \sum_{a,b} |a-b|^k P_{\phi,d}^\lambda(a,b) \quad \text{with } k=2, \lambda=1, \tag{4}$$

inverse difference moment

$$= \sum_{a,b:a \neq b} (P_{\phi,d}^\lambda(a,b) / |a-b|^k) \quad \text{with } k=2, \lambda=1, \tag{5}$$

$$\text{correlation} = \left(\sum_{a,b} [(ab)P_{\phi,d}(a,b)] - \mu_x \mu_y \right) / \sigma_x \sigma_y. \tag{6}$$

The amount of energy along the diagonal of the cooccurrence matrix can also be used as an additional texture measure. We used five definitions of neighbor: 1, 2, and 4 pixels to the right; 1 pixel down; and 1 pixel to the southeast.

The second statistical method calculates a number of features based on the gray level histogram of the image.¹⁸ The first few features include the mean and second through fourth moments of the histogram. If the histogram is defined as $p(z_i)$, where z is the gray level, L is the number of gray levels, and i goes from 0 to $L-1$, the remaining histogram features¹⁸ are

$$\text{relative smoothness} = 1 - \{1/[1 + \sigma^2(z)]\}, \tag{7}$$

$$\text{uniformity} = \sum_i p^2(z_i), \tag{8}$$

$$\text{entropy} = \sum_i p(z_i) \log_2 p(z_i). \tag{9}$$

The third method used basic statistics to define four additional features for the image: the mean, standard deviation, range, and median.

The fourth method used a structural approach to texture analysis, Laws' texture measures,¹⁹ to calculate 14 texture features. Laws's method begins with three 1-D convolution kernels, representing averaging [$L3=\{1 \ 2 \ 1\}$], edges [$E3=\{-1 \ 0 \ 1\}$], and spots [$S3=\{-1 \ 2 \ -1\}$], and eventually creates 14 texture features.

The final method used the spectral approach to texture analysis by taking the 2-D Fourier transform of the image and calculating the average energy values in several regions of the Fourier spectrum. We calculated the energy within six rings around the origin, within two horizontal bands close to the origin, and within two vertical bands close to the origin. We also calculated the normalized autocorrelation function, defined¹³ as $C_{ff}=F^{-1}\{|F|^2\}$, where F represents the 2-D Fourier transform of the image. The resulting autocorrelation decreases slowly with increasing distance if the primitives in the texture are large, and decreases more quickly as the primitives decrease in size. We used the maximum autocorrelation value and the relative number of correlation values above three different thresholds as four additional texture features.

3.3 Feature Selection

Once feature vectors were created for the training set, redundant features were removed, and the remaining features were compared for their ability to distinguish between images of non-cancerous and cancerous tissue.

To remove redundant features, we calculated the correlation matrix for the feature set, normalized it over the interval $[0, 1]$, and removed all but one feature from each set of highly correlated features (correlation >0.85). After the correlated features were removed, there were 18 remaining features in the feature vectors, which were then evaluated together and in groups for their ability to discriminate between tissue types. The 18 uncorrelated features are listed in Table 1.

3.3.1 Scatter plots of features

After the feature set reduction, the resulting feature vectors were normalized over the interval $[0, 255]$ and grouped according to their associated pathology results. Since research suggests that cancerous tissue disrupts the normally well structured, layered appearance of bladder tissue and appears more homogeneous than normal tissue, we hypothesized that the second moment of the histogram would be one of the strongest remaining features in distinguishing cancerous from noncancerous tissue, and used this as the second feature to produce 2-D plots for each pair of features.

The purpose of the plots was to identify whether the cancerous and noncancerous cases were indeed separable from

*The variance used in the relative smoothness equation should be normalized to the interval $[0, 1]$ by dividing it by $(L-1)^2$ before using the equation.

Table 1 Features remaining after removal of uncorrelated features.

Cooccurrence matrix: energy, neighbor defined as 1 east	Histogram: mean
Cooccurrence matrix: contrast, neighbor defined as 1 east	Histogram: second moment
Cooccurrence matrix: inverse difference moment, neighbor defined as 1 east	Histogram: third moment
Cooccurrence matrix: correlation, neighbor defined as 1 east	Basic statistics: range
Cooccurrence matrix: energy along the diagonal, neighbor defined as 1 east	Laws': level-edge texture measure
Cooccurrence matrix: contrast, neighbor defined as 1 south	Autocorrelation: maximum value
Cooccurrence matrix: inverse difference moment, neighbor defined as 1 south	Fourier transform: energy in vertical band around origin
Cooccurrence matrix: energy along the diagonal, neighbor defined as 1 south	Fourier transform: energy in horizontal band around origin
Cooccurrence matrix: inverse difference moment, neighbor defined as 1 southeast	Fourier transform: energy in ring around origin

one another, and whether they clustered tightly. Normal and exudative inflammation did tend to cluster together somewhat separately from dysplasia and CIS, which also tended to cluster together. Infiltrative inflammation, invasive tumors, and papillary lesions did not cluster well with themselves or each other. Figure 3 shows an example plot in which these gener-

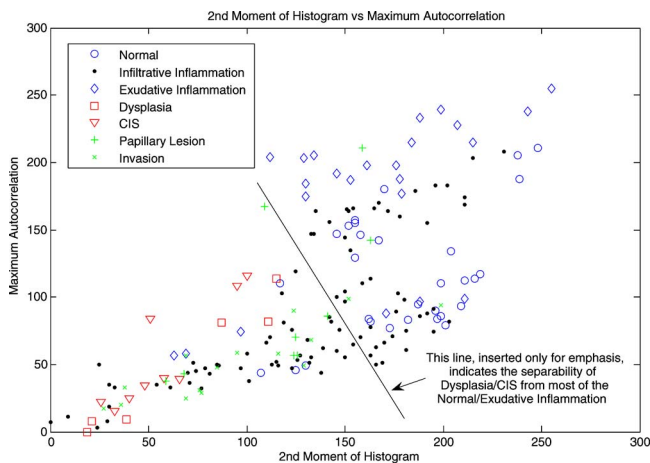


Fig. 3 Example scatter plot showing data set using two features: the second moment of the histogram and the maximum autocorrelation. For emphasis, a line was manually added to the plot to show the separability of dysplasia and CIS from normal tissue and exudative inflammation. This line was, however, not used to classify the data points. The line only serves to emphasize the separation evident in the image. All cases of dysplasia and CIS fall on one side of the line shown, while most cases of normal tissue and exudative inflammation fall on the other. The cases of infiltrative inflammation, papillary lesions, and invasive tumors are not separated by the two features.

alizations are evident.

Based on our observations from the feature plots, we excluded infiltrative inflammation from the training set for non-cancerous tissue and began the decision process by assigning the data to one of two classes: normal/exudative inflammation and dysplasia/CIS. Each class then would be checked separately for instances of papillary growth and for invasion, as well as for infiltrative inflammation.

3.3.2 Criterion to select feature subsets

It is well known that two features acting together may perform well even though the features acting alone perform poorly.²⁰ Consequently, when determining which feature subset produces the best separation between two classes of data, all combinations of features must be considered. The problem becomes one of determining which subset produces the best results. When dealing with clusters of data, a criterion can be used to quantify how well a certain set of features separate the data. One such criterion is the trace of the ratio of between-class scatter S_B to within-class scatter²⁰ S_W .

The within-class scatter S_i for cluster i represents the spread of the data points within cluster i , and can be calculated²⁰ as $S_i = \sum_j (x_j - m_i)(x_j - m_i)^t$, where m_i is the mean vector for cluster i , and x is the set of j points in cluster i . The total within-class scatter S_W for a group of clusters is the sum of the within-class scatter across all clusters $S_W = \sum_i S_i$.

The between-class scatter S_B for a group of classes represents the distance separating the clusters, and may be calculated as $S_B = \sum_i n_i (m_i - m)(m_i - m)^t$, where m is the mean vector for the entire data set, m_i is the mean vector for cluster i , and n_i is the number of points in class i . The larger the trace is of the ratio of between-class scatter to within-class scatter,

the more distinct are the clusters. When searching for the best possible subset, the subsets with the largest traces should be considered first.

To maintain the generality of our algorithm, we limited our search to feature subsets consisting of four or fewer features, since theory states that for finite data sets there is an optimal (in the sense of probability of error) number of features; larger or smaller values lead to decreased performance.²⁰ The trace of the ratio of between-class scatter S_B to within-class scatter S_W was calculated for all combinations of one, two, three, and four features for the following class comparisons:

1. normal/exudative versus dysplasia/CIS
2. normal/exudative versus papillary lesion \pm invasion
3. dysplasia/CIS versus papillary lesion \pm invasion
4. dysplasia/CIS versus infiltrative inflammation
5. infiltrative inflammation versus papillary lesion \pm invasion

In the preceding comparisons, normal and exudative inflammation are grouped into a single category, dysplasia and CIS are grouped into a single category, and all types of invasive tumors are grouped together with papillary lesions. The papillary lesions do not have to be invasive to be included in the papillary \pm invasion group, nor do the invasive tumors have to be papillary.

In the next step of our algorithm development, the subsets producing the largest trace values using one, two, three, and four features for each comparison were tested on their ability to accurately separate the training set for the classes in question.

3.3.3 Data classification using a discriminant function

One method of classifying data, when the probability distribution for the classes is available, is to use discriminant functions to assign the data to one of the classes. Each possible class ($i=1, \dots, c$) has a discriminant function $g_i(x)$, which can be calculated and then compared to the discriminant functions of the other classes for each piece of sample data x . The sample x is assigned to the class for which the discriminant function is the greatest. If the classes have a normal density with mean μ_i and covariance matrix Σ_i , the discriminant function $g_i(x)$ can be defined as $g_i(x) = (-0.5)(x - \mu_i)' \Sigma_i^{-1} (x - \mu_i) - (d/2) \ln(2\Pi) - (1/2) \ln |\Sigma_i| + \ln P(\omega_i)$ (Ref. 20). Here $P(\omega_i)$ is the prior probability of x being in class i , and d is the dimensionality of the feature vectors.

For each comparison listed previously, the feature subsets selected by the trace comparisons were used to create feature subsets for the two classes being compared. The mean vectors and covariance matrices were calculated using the training data for each class. For this study, we assumed that the prior probabilities were the same for all classes being considered and kept them at 0.5; they can be adjusted if known or estimated values become available. For each image in both classes, the discriminant functions were calculated and compared, and the image was assigned to one of the two classes. The feature subset that produced the highest number of correct classifications was selected for use in the final decision process.

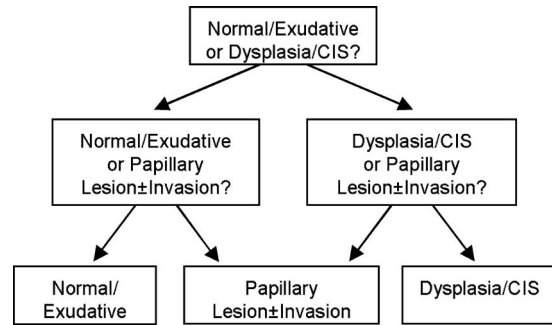


Fig. 4 Decision tree used in classification of images. Images are first subdivided into normal/exudative or dysplasia/CIS, after which each class is checked for cases of papillary lesion \pm invasion, resulting in three possible end diagnoses: normal/exudative, dysplasia/CIS, or papillary lesion \pm invasion.

3.4 Decision Tree

Based on the results of the feature selection process, a decision tree was created to classify any image into one of three classes: normal/exudative inflammation, dysplasia/CIS, or papillary lesion \pm invasion. The decision tree shown in Fig. 4 uses various feature subsets at each decision point in the tree. The decisions are made using a discriminant function and the training data for the specific diagnoses included in each class. When distinguishing between cancerous and noncancerous tissue, the dysplasia/CIS and papillary lesion \pm invasion classes are combined to form the more general cancerous class.

Infiltrative inflammation was not included in the decision tree since none of the feature subsets were able to reliably distinguish infiltrative inflammation from either dysplasia/CIS or papillary lesion \pm invasion. Cases of infiltrative inflammation that were diagnosed in the normal/exudative class were considered to be correct, so no attempt was made to distinguish between normal/exudative and infiltrative inflammation.

Once the decision tree was created, leave-one-out cross validation was used to determine the accuracy of the algorithm. Leave-one-out cross validation is a method of estimating classifier performance that does not require the data set to be divided into two separate training and test sets, but maintains independence between the test and training sets.²¹ The method requires that the following steps be repeated for each individual data point in the data set: remove the data point from the data set, use the remaining data set as the training data, test the removed sample using the mean vector and covariance matrix calculated from the training data, and return the removed data point to the data set. When this method is followed, each data point is tested on a training data set of which it was not a member, thereby maintaining the independence of the test and training data sets.

4 Results

4.1 Selected Features

While the trace values were useful when differentiating between subsets containing the same number of features, trace values tended to be higher when fewer features were in the subset. In all cases, a single feature was selected as the subset with the highest trace even though the selected feature never

Table 2 Algorithm performance: cancerous versus noncancerous.

Pathology Results	Decision		Specificity(%)	Sensitivity(%)
	Noncancerous	Cancerous		
Normal	26	4	87	
Exudative inflammation	21	3	88	
Infiltrative inflammation	42	47	47	
Dysplasia	0	6		100
CIS	0	9		100
Invasive tumor	1	14		93
Papillary lesion	2	7		78
Total noncancerous	89	54	62	
Total cancerous	3	36		92

provided the best separation between classes. In general, however, the higher the trace value, the better the separation of the classes.

The first step in the decision tree separates the data into two classes based on the training data representing normal/exudative inflammation and dysplasia/CIS. The set of features selected for this step was the set of four features with the highest trace (0.98), and included the following features: (1) the correlation of the cooccurrence matrix with “neighbor” defined as the pixel immediately to the right, (2) the mean of the histogram, (3) the second moment of the histogram, and (4) the range of intensities in the image. When the discriminant function was used to separate the data using the preceding subset, 51 out of 54 normal/exudative images were classified correctly, and all images of dysplasia/CIS were classified correctly. While the set of four features was selected as the best, the subsets of two and three features, with trace values of 1.54 and 1.39, respectively, were also able to separate the classes effectively.

The second step of the decision tree consisted of two parts, one that attempted to distinguish images of papillary lesion \pm invasion from normal/exudative, and one that attempted to distinguish them from dysplasia/CIS. None of the subsets had particularly high trace values when comparing normal/exudative images to papillary lesion \pm invasion, indicating that it would be difficult to accurately separate the classes. The subset eventually chosen was the subset containing (1) the second moment of the histogram and (2) the maximum autocorrelation. The trace for this subset was only 0.59. When the discriminant function was used to classify the images, 47 out of the 51 normal/exudative images remained in the class normal/exudative while 10 of the 13 papillary lesion \pm invasion images previously in the normal/exudative class were reassigned to the papillary lesion \pm invasion class.

As with the previous comparison, none of the subsets had particularly high traces when comparing dysplasia/CIS with papillary lesion \pm invasion. The subset chosen for this branch

of the tree had a trace value of 1.08, and included the following three features: (1) the second moment of the histogram, (2) the third moment of the histogram, and (3) the amount of energy in a horizontal band around the origin of the Fourier transform of the image. The discriminant function left 10 of the 15 dysplasia/CIS images classified as dysplasia/CIS and moved 6 of the 11 papillary lesion \pm invasion images previously in the dysplasia/CIS set to the correct class.

4.2 Classification Results

Using the decision tree described previously, we classified (as either cancerous or noncancerous) each of the 182 images in the test set using leave-one-out cross validation. The results are shown in Table 2. As expected from the scatter plots, the specificities for normal tissue and exudative inflammation are fairly high, 87 and 88% respectively, while the specificity for infiltrative inflammation is only 47%, giving an overall specificity of 62%. The sensitivities were 100, 93, and 78% for dysplasia and CIS, invasive tumors and papillary lesions, respectively.

The results for correctly distinguishing dysplasia/CIS and papillary lesion \pm invasion are shown in Table 3, along with the sensitivity for each specific diagnosis. The sensitivity for papillary lesions is 86%, but the others are lower, with dysplasia having a sensitivity of only 50%.

If the results of the first step in the decision tree are considered separately, it becomes apparent that the specificity was significantly reduced by the second set of decisions. The specificity after the first branch of the decision tree was 100% for normal images, 88% for exudative inflammation, and 80% for infiltrative images, producing an overall specificity of 85%. The sensitivity for the cancerous classes involved in the first step, dysplasia and CIS, was 100%. The implications of this observation are discussed later.

Table 3 Algorithm performance: dysplasia/CIS versus papillary lesion \pm invasion.

Pathology Results	Decision			Specificity	Sensitivity (%)
	Noncancerous	Dysplasia/CIS	Papillary Lesion \pm Invasion		
Dysplasia	0	3	3		50
CIS	0	7	2		78
Invasive tumor	1	4	10		71
Papillary lesion	2	1	6		86

5 Discussion

While a number of studies have attempted to characterize the appearance of various types of bladder pathology in OCT images, there have been only a few studies that have attempted to use OCT images to differentiate pathological states within the bladder mucosal surface, and none that has attempted to automate the process. The studies that used human observers to analyze OCT images of the bladder produced results slightly better than the results presented here, but those results may have been affected by some outside knowledge such as the visual classification of the image as suspicious or the appearance of the other OCT images taken in the same area. Moreover, those results are highly dependent on the particular observer. The results presented here, on the other hand, are based entirely on the features in the image in question along with the features present in the training set. Inter- and intra-operator error is eliminated by this process.

The results of our study are presented two ways, one in which the images are classified as being either cancerous or noncancerous, and one in which the images are classified as being noncancerous, dysplasia/CIS, or papillary lesion \pm invasion. While the second instance is more informative, simple recognition of cancerous tissue is a significant step in guiding for biopsy. In fact, one of the studies that used human observers to classify OCT images of the bladder epithelium did not even attempt to distinguish between the types of cancerous tissue.¹²

Note that the cancerous condition with the worst sensitivity in our study, a papillary lesion, is the condition that is most readily visible with video endoscopy. One contributing reason for this anomaly is that when papillary growths were present, the surgeon obtained OCT images at the base of the lesion, which often resulted in images that did not include features specific to papillary lesions, but rather features of the neighboring tissue. It is reasonable to expect that recognition of a papillary lesion with OCT is not of high priority because it will be recognized visually during the procedure, although OCT imaging of these lesions is still necessary to determine the penetration depth. The sensitivity of our algorithm increases from 92 to 97% if papillary growths are excluded from the sensitivity calculation.

The specificity of our algorithm suffers significantly due to difficulty recognizing infiltrative inflammation. If the cases of infiltrative inflammation are removed from the study, the specificity increases to 87%. The study by Zagayanova et al.¹²

noted a similar problem with nonproliferative cystitis being misdiagnosed 30% of the time. The study by Manyak et al.¹¹ indicated that inflammatory conditions contributed to the rate of false positives. Infiltrative inflammation can be common in the epithelial lining of the bladder when cancer is present, so to improve the specificity of diagnoses based on OCT images, new techniques must be found to distinguish infiltrative inflammation from cancerous conditions.

In regard to the texture features selected for the decision tree, it appears that the second moment of the histogram is indeed well suited for distinguishing between tissue classes, in that it was included in each of the subsets used for the decision tree. On the other hand, the fact that some of the images of superficial papillary growths and invasive tumors were not recognized by our algorithm, while all cases of dysplasia and CIS were recognized, could indicate that the description of a homogeneous image may apply better to cases of dysplasia and CIS than to invasive tumors or papillary lesions. The nature of the texture in those cases requires further explication. In addition to the second moment of the histogram, the mean of the histogram, the range of intensities, and the correlation between a pixel and the pixel immediately to the right were selected as features when distinguishing between normal or exudative tissue and dysplasia or CIS. Images of dysplasia or CIS appear to have both a lower mean intensity and a smaller intensity range than images of normal tissue or tissue with exudative inflammation, while having a higher correlation between neighboring pixels. A possible explanation for this is that the growth of cancerous cells disrupts the normal structure of the bladder wall. Larger optical reflections occur at boundaries between tissue types, especially when these boundaries are smooth and flat (acting as specular reflectors). Consequently, less heterogeneous, less structured cancerous tissues would be expected to exhibit lower mean intensities and ranges of intensities. In addition, the locally homogeneous nature of the cancerous tissue would increase the possible correlation between a pixel and its neighbor. We are unsure, however, why not all the images of invasive tissue or papillary lesions appeared to have similar characteristics, since they too should have become more homogeneous and unstructured. As mentioned previously, the nature of the images of papillary lesions or invasive tumors needs to be further studied.

6 Conclusions

Our algorithm, which was tested on 182 OCT images of bladder tissue, had a sensitivity of 92% and a specificity of 62% when classifying tissue as either cancerous or noncancerous. The results indicate that texture analysis is a viable method for analyzing OCT images of the bladder. Other researchers have used computer analysis of OCT images to recognize dysplasia in the esophagus^{15,16} and tumors in the breast¹⁷ with similar performance, indicating that computer-aided diagnosis of OCT images is realistic.

Our study, as well as studies using human observers, suffers from a common problem, namely, the difficulty in distinguishing infiltrative inflammation from cancerous tissue when analyzing OCT images of the lining of the bladder. To improve the specificity of diagnoses provided by OCT, new methods must be found to make this distinction. Once a method of distinguishing infiltrative inflammation from cancerous tissue is found, computer diagnosis of OCT images of the lining of the bladder should be much more specific, enabling physicians to begin treatment at the time and place of the identification of the cancer, rather than having to attempt to return to that site in a subsequent procedure.

While our algorithm did produce a sensitivity of 92%, this must be improved, as does the algorithm's ability to distinguish between the types of cancerous tissue. Further research into other texture and image analysis methods must be conducted to see if methods can be found that provide more reliable recognition of invasive tumors and papillary lesions. In this study, we did not differentiate between invasive tumors and papillary lesions since these cases seemed to have similar texture features for the features considered. However, this information is critical when determining the course of treatment and must be addressed. Invasive lesions are much more dangerous and require significant alterations in treatment.

As mentioned previously, the specificity after the first step of our algorithm was 85%, while producing a sensitivity of 100% for dysplasia and CIS. If the goal of our study had been to recognize dysplasia, and distinguish it from noncancerous tissue, as in the studies by Qi et al.^{15,16} our algorithm would have been very successful. It is therefore realistic to assume that the features selected for the first step of our algorithm are sufficient to distinguish between dysplasia/CIS and noncancerous tissue.

Further testing on an independent data set will strengthen the clinical value of this approach.

Acknowledgments

This work was supported in part by funding provided by The Wallace H. Coulter Foundation and the ARCS Foundation. In addition, we would like to thank N. Gladkova, J. Makari, A. Schwartz, E. Zagaynova, L. Zohlfaghari, R. Iksanov, and F. Feldchtein for their assistance in obtaining the clinical data used in this work.

References

1. The American Cancer Society, "Cancer Facts and Figures 2007," www.cancer.org.
2. H. B. Grossman, L. Gomella, Y. Fradet, A. Morales, J. Presti, C. Ritenour, U. Nsevo, and M. J. Droller, "PC B302/01 study group. A phase III multi center comparison of hexaminolevulinatate fluorescence cystoscopy and white light cystoscopy for the detection of superficial papillary lesions in patients with bladder cancer," *J. Urol. (Baltimore)* **178**(10), 62–67 (2007).
3. S. F. Shariat, G. S. Palapattu, G. E. Amiel, P. I. Karakiewicz, C. G. Rogers, A. Vazina, M. P. Schoenberg, S. P. Lerner, A. I. Sagalowsky, and Y. Lotan, "Characteristics and outcomes of patients with carcinoma in situ only at radical cystectomy," *Urology* **68**(3), 538–542 (2006).
4. P. H. Tomlins and R. K. Wang, "Theory, developments and applications of optical coherence tomography," *J. Phys. D* **38**, 2519–2535 (2005).
5. J. G. Fujimoto, M. E. Brezinski, G. J. Tearney, S. A. Boppart, B. Bouma, M. R. Hee, J. F. Southern, and E. A. Swanson, "Optical biopsy and imaging using optical coherence tomography," *Nat. Med. (N.Y.)* **1**(9), 970–972 (1995).
6. University of Virginia Health System, "Urinary tract pathology tutorials," www.med-ed.virginia.edu/courses/path/urinary.
7. F. K. Mostofi, C. J. Davis, and I. A. Sesterhenn, *Histological Typing of Urinary Bladder Tumours*, 2nd ed., World Health Organization International Histological Classification of Tumours, Springer-Verlag, Berlin Heidelberg (1999).
8. R. B. Bracken and B. Mackay, *Atlas of Bladder Pathology*, Igaku-Shoin Medical Publishers, Inc., New York (1991).
9. F. I. Feldchtein, G. V. Gelikonov, V. M. Gelikonov, R. V. Kuranov, and A. M. Sergeev, "Endoscopic applications of optical coherence tomography," *Opt. Express* **3**(6), 257–270 (1998).
10. C. A. Jesser, S. A. Boppart, C. Pitris, D. L. Stamper, G. P. Nielsen, M. E. Brezinski, and J. G. Fujimoto, "High resolution imaging of transitional cell carcinoma with optical coherence tomography: feasibility for the evaluation of bladder pathology," *Br. J. Urol.* **72**, 1170–1176 (1999).
11. M. J. Manyak, N. D. Gladkova, J. H. Makari, A. Shwartz, E. V. Zagaynova, L. Zolghari, J. M. Zara, R. Iksanov, and F. I. Feldchtein, "Evaluation of superficial bladder transitional-cell carcinoma by optical coherence tomography," *J. Endourol.* **19**(5), 570–574 (2005).
12. E. V. Zagaynova, O. S. Streltsova, N. D. Gladkova, N. M. Shakova, F. I. Feldchtein, V. A. Kamensky, G. V. Gelikonov, L. B. Snopova, and E. V. Donchenko, "Optical coherence tomography in diagnostics of precancer and cancer of human bladder," *Proc. SPIE* **5312**, 75–81 (2004).
13. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, 2nd ed., PWS Publishing, Pacific Grove, CA (1999).
14. K. W. Gossage and T. S. Tkaczyk, "Texture analysis of optical coherence tomography images: feasibility for tissue classification," *J. Biomed. Opt.* **8**(3), 570–575 (2003).
15. X. Qi, D. Y. Rowland, M. V. Sivak, and A. M. Rollins, "Computer-aided diagnosis of dysplasia in Barrett's esophagus using multiple endoscopic OCT images," *Proc. SPIE* **6079**, 60790I (2006).
16. X. Qi, M. V. Sivak, G. Isenberg, J. E. Willis, and A. M. Rollins, "Computer-aided diagnosis of dysplasia in Barrett's esophagus using endoscopic optical coherence tomography," *J. Biomed. Opt.* **11**(4), 044010 (2006).
17. A. M. Zysk and S. A. Boppart, "Computational methods for analysis of human breast tumor tissue in optical coherence tomography images," *J. Biomed. Opt.* **11**(5), 054015 (2006).
18. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed., Prentice-Hall, Upper Saddle River, NJ (2002).
19. K. I. Laws, "Texture energy measures," in *Proc. DARPA Image Understanding Workshop*, pp. 47–51, IEEE, Los Altos, CA (1979).
20. R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed., Wiley, New York (2001).
21. K. Fukunaga and R. R. Hayes, "Estimation of classifier performance," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(10), 1087–1101 (1989).