Wavelet analysis enables system-independent texture analysis of optical coherence tomography images

Colleen A. Lingley-Papadopoulos
The George Washington University
Department of Electrical and Computer Engineering
Staughton 107
707 22nd Street Northwest
Washington, DC 20052
E-mail: apapas@aframedigital.com

Murray H. Loew
The George Washington University
Department of Electrical and Computer Engineering
Staughton 101
707 22nd Street Northwest
Washington, DC 20052

Jason M. Zara
The George Washington University
Department of Electrical and Computer Engineering
Staughton 107
707 22nd Street Northwest
Washington, DC 20052

Abstract. Texture analysis for tissue characterization is a current area of optical coherence tomography (OCT) research. We discuss some of the differences between OCT systems and the effects those differences have on the resulting images and subsequent image analysis. In addition, as an example, two algorithms for the automatic recognition of bladder cancer are compared: one that was developed on a single system with no consideration for system differences, and one that was developed to address the issues associated with system differences. The first algorithm had a sensitivity of 73% and specificity of 69% when tested using leave-one-out cross-validation on data taken from a single system. When tested on images from another system with a different central wavelength, however, the method classified all images as cancerous regardless of the true pathology. By contrast, with the use of wavelet analysis and the removal of system-dependent features, the second algorithm reported sensitivity and specificity values of 87 and 58%, respectively, when trained on images taken with one imaging system and tested on images taken with another. © *2009 Society of Photo-Optical Instrumentation Engineers.* [DOI: 10.1117/1.3171943]

1 Introduction

Optical coherence tomography (OCT) is an emerging imaging modality that provides higher resolution images than ultrasound, but is limited in its application due to its penetration depth of only $1-2$ mm in biological tissue.[1] For certain applications, however, such as imaging of the retina or analysis of epithelial tissues, OCT is an ideal imaging modality. To date, OCT has gained acceptance in the area of retinal imaging and has become a valuable tool in the diagnosis and monitoring of glaucoma.[2–5] OCT is being studied for use in: the recognition of intravascular plaque,[6–8] the recognition of cancer or pre-cancer in a number of organ systems,[9–12] and a number of other medical fields.[13–18]

Many studies have applied image analysis techniques to OCT images to automatically characterize the imaged tissue in some manner. In the case of retinal imaging, recognition of the layers of the retina and measurement of the thicknesses of those layers plays a key role in diagnosis.[2] In the case of imaging for the recognition of cancer, texture analysis has been studied as a means of recognizing differences between cancerous and noncancerous tissue.[10–12] This is due to its ability to place quantitative values on the amount of homogeneity in an image or due to the effect the size and distribution of scatterers has on certain texture features.[19] However, most of these studies have been limited to a single imaging system, which limits their applicability when used with other OCT systems.

In this paper, we discuss some of the differences between OCT imaging systems and the effect these differences have on the resulting images and on algorithms developed to classify those images. In addition, some methods are introduced to compensate for system differences, including the use of texture analysis on the approximation and detail matrices created through wavelet analysis. Finally, two algorithms to distinguish images of cancerous bladder tissue from images of noncancerous bladder tissue are compared for their ability to detect cancer in images taken with two very different OCT imaging systems. One of the algorithms was developed without concern for imaging system differences while the other was developed specifically to address those issues.

2 Background

2.1 OCT

OCT is an imaging modality similar to ultrasound that uses partially coherent near-infrared light instead of sound to create images of subsurface structures. The origin of the received backscattered light is detected with interferometry; thus, a map of reflectivity versus optical depth and lateral position

Address all correspondence to: Colleen A. Lingley-Papadopoulos, Department of Electrical and Computer Engineering, The George Washington University, Staughton 107, 707 22nd St NW, Washington, DC 20025. Tel: (703) 944-7772; Fax: (202) 994-0280; E-mail: apapas@aframedigital.com

can be created. The core of the imaging system is a Michelson interferometer. The light from the optical source is split into two paths. One of the paths, the reference arm, consists of a delay line that varies the path length to create the depth information of the image. The other path contains the scanning arm, which sweeps the beam across the sample to obtain the lateral image information. When the path lengths of the light in each of the two arms are equal, they add constructively at the detector. These signals are then processed and used as the individual axial lines for the OCT image.[20]

The axial and lateral resolutions of an OCT imaging system are decoupled. The lateral resolution is determined by the optics in the scanning arm, and the axial resolution is determined by the coherence length of the light source.[21] The coherence length is inversely proportional to the spectral bandwidth of the light source; thus, for smaller coherence lengths, larger bandwidths are needed. Depending on the bandwidth, axial resolutions as low as 4 μm can be achieved, which is up to 25 times higher than the resolution of high-frequency ultrasound.[22] When imaging, the axial and lateral resolutions are usually kept similar to maintain distance relationships in the image. However, for the lateral resolution to match the axial resolution when it drops below 10 μm, the focal length of the lens would need to be so small that the depth of focus of the system would drop. Consequently, most current catheter-based systems have axial resolutions between 10 and 20 μm.[22] The actual axial resolution depends on the medium being imaged, but is proportional to the coherence length.

Tissue penetration depth in OCT is also a function of the wavelength of light used. For optical energy, tissue preferentially absorbs some wavelengths of light, while having very little effect on others. For instance, in the near-infrared range from 800 to 1300 nm, there is very little absorption in tissue.[23] For wavelengths that are not absorbed, scattering, which decreases at greater wavelengths, is the primary limitation of penetration depth. For this reason, OCT imaging has been investigated at several near-infrared wavelengths (830, 980, 1060, and 1300 nm), depending on whether the emphasis is on resolution or tissue penetration.[1] At smaller wavelengths, resolution is higher, while at greater wavelengths tissue penetration is increased.

2.2 OCT Image Analysis

The first OCT system was approved by the Food and Drug Administration for use in ophthalmology in 1993,[24] and the first commercial OCT system was introduced into the clinic in 1996.[24] Since then, many groups have developed algorithms to automatically detect the layers of the retina.[2–5] The thickness of the retina, which can be determined from OCT images, can help diagnose certain diseases and lead to more appropriate treatment.[2] Similarly, analysis of the ganglion cells and axons in the retinal nerve fiber layer (RNFL), which can be visualized using OCT, can help determine the progress of glaucoma.[2]

More recently, differences between OCT systems and the effect those differences have on measurements, such as the thickness of the RNFL, has become an active area of study. One of the first groups to make such comparisons was Bourne et al.,[25] who imaged 139 subjects with two imaging systems and then compared the resulting RNFL thickness measure-

ments. One of the systems generally produced thicker measurements; however, a correction factor was determined that, when applied, brought the measurements within 10 μm of each other for 75% of the patients. This difference is still large enough, however, that Bourne et al.[25] suggest that the measurements taken by these systems should be compared with caution. In a similar study, Sehi et al.[26] compared the measurements obtained by two different OCT systems when measuring not only the RNFL, but also the optic disk topography and central foveal thickness. As with the Bourne study, there were significant differences between measurements taken with the two systems; although in this study, no attempt was made to compensate for these differences. A warning was simply given that when comparing values between clinics or systems, system differences should be taken into account.

Research in the area of OCT image analysis is not limited to retinal measurements. OCT is actively being studied for the characterization of intravascular plaque[6,27] as well as for the recognition of cancer. Gossage and Tkaczyk[28] may have been the first to suggest that texture analysis can be used to analyze OCT images in an attempt to classify different tissue types. They were able to analyze the texture in OCT images to differentiate 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%). In a follow-up study, Gossage et al.[19] used texture analysis and tissue phantoms to study the effect of the size and distribution of scatterers on the speckle present in OCT images. The results indicated that the change in size and distribution of scatterers did have a statistically significant effect on certain texture features, such as entropy, local homogeneity, inertia, and the low-frequency Fourier transform. In recent years, texture analysis has been used by a number of researchers in efforts to distinguish cancerous from noncancerous tissue. In two separate studies, Qi et al.[10,29] applied texture analysis along with other image analysis techniques to OCT images of the esophagus to diagnose dysplasia. The first study reported a sensitivity of 87% and a specificity of 69%, while the second study reported a sensitivity of 82% and a specificity of 74%. Zysk and Boppart created an algorithm using a combination of image and texture analysis techniques to recognize breast cancer.[11] The results of the study indicated that the combined algorithm had a tumor tissue sensitivity of 97% and a specificity of 68%. All these studies were carried out using data from only one system.

Previously, we have applied texture analysis to OCT images of the bladder to recognize the layers of the bladder,[30] and to differentiate cancerous from noncancerous tissue.[12] The study to differentiate cancerous from noncancerous bladder tissue reported a sensitivity of 92% and a specificity of 62%, using data from a single system. Unfortunately, the features selected turned out to be highly system dependent, and the algorithm overtrained due to the limited data used for algorithm development. It became apparent that for our algorithm to be applicable to multiple systems without requiring acquisition of training data on each new system, it needed to be designed with system independence in mind, causing us to begin considering methods of compensating for system differences. In addition, to avoid overtraining as we continued our research, we reverted to a simpler algorithm that required use of fewer features and would be influenced less by deviations

present in the training data set. As more data become available, it should be possible to return to more complex algorithms capable of more reliable differentiation between tissue classes. The development of methods of accounting for system differences will not only allow developed algorithms to be applicable to more than a single imaging system, it will make it possible to use data from multiple data sets, developed on different systems, to develop more robust and complex algorithms.

Chen et al.[31] conducted a study that used texture analysis to evaluate the ability of two systems having different resolutions to recognize Barrett's esophagus. *In vivo* and *in vitro* tissues from the esophagus were analyzed with both imaging systems. One system had a center wavelength of 1310 nm and a resolution between 10 and 15 μm. The high-resolution system had a center wavelength of 1260 nm and a resolution of <5 μm. The goal of the study was to determine whether the high-resolution system was able to improve the recognition of Barrett's esophagus. The results indicated that the texture features calculated with the high-resolution system were able to better discriminate between normal tissue and Barrett's esophagus. The paper does not mention, however, if the images taken with one system could be successfully analyzed using the images taken with the other as the training set.

Because of the limited availability of clinical OCT data in the area of cancer detection, recognition algorithms are being developed on single, small data sets. The reality, however, is that to develop robust algorithms capable of analyzing data from different systems, developers need multiple data sets from many different studies taken with different systems. If this is not the case, then the developed algorithms are likely to be limited to use on the system for which they were developed or to require training data gathered on each new system. To avoid having to acquire training data for every new system, we seek a method capable of compensating for system differences.

2.3 OCT System Differences

There are a few parameters of OCT systems that are apparent when differentiating between systems, including resolution and depth of penetration, which are dependent on the central wavelength of the light source used. Systems with smaller wavelengths have better resolution but less penetration depth.

Other system differences include the signal-to-noise ratio, the range of intensity values, and the pixel size. The range of intensity values is often set at the time of imaging and may vary slightly from use to use. The general range would be dependent on the average intensity of the sample and reference beams as well as the analog-to-digital converter used in the system. Likewise, the actual pixel size would be determined by probe and system parameters, but could be controlled by the system software.

To compare data sets taken with different imaging systems, or with different probes, we must compensate for these system differences. If the pixel size and intensity ranges are known, as they often are, then it would be possible to use resizing and normalization techniques to account for some of those differences. Furthermore, because the OCT signal strength decreases with depth, the penetration depth and the portion of the image analyzed will play a significant role on
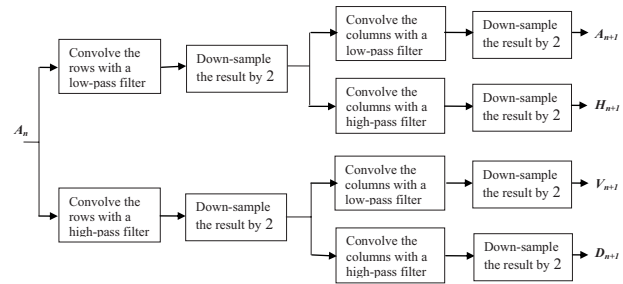

Fig. 1 Two-dimensional discrete wavelet transform.

the results of any image analysis. This paper will address the effects of system differences on OCT diagnosis algorithms, using two algorithms designed for the recognition of bladder cancer as an example.

2.4 Wavelets

An approach that could help reduce the effect of system parameters on texture features is the calculation of texture features on the output of wavelet analysis. Wavelet analysis is similar to Fourier analysis but uses waveforms of limited duration and irregular form to describe a signal, instead of infinitely long sine waves. Once the "mother wavelet" is selected, it is shifted to provide temporal information, and scaled to provide information at different scales. At each shifted location, a window of the original signal is compared to the scaled wavelet and a wavelet coefficient is calculated. After this comparison is made across the input signal, a series of wavelet coefficients is available for the given scale. The process can be repeated at any number of scales.[32]

As with the discrete Fourier transform, a discrete wavelet transform exists, that limits the wavelet analysis to scales and positions that are powers of two. There can be at most $log_2 N$ stages of analysis if the input signal is N samples long. Each stage, n, of the discrete wavelet transform has two parts. It takes the input of the stage, convolves it with a low-pass filter, and then down-samples the result by 2 to produce what are called the approximation coefficients, A_n. It also takes the input of the stage, convolves it with a high-pass filter, and then downsamples the result by 2 to produce what are called the detail coefficients, D_n. The original signal is the input to the first stage, and the approximation coefficients of the previous stage are the inputs to all subsequent stages.[32]

Discrete wavelet analysis can be performed in two dimensions, for analysis of images. On an input image of dimension m by n, the output of the first stage would be an approximation matrix **A1** of dimension $m/2$ by $n/2$, a horizontal detail matrix **H1** of dimension $m/2$ by $n/2$, a vertical detail matrix **V1** of dimension $m/2$ by $n/2$, and a diagonal detail matrix **D1** of dimension $m/2$ by $n/2$. The approximation matrix **A1** would be the input to the second stage. A diagram of the function of the discrete wavelet transform in two dimensions is shown in Figure 1.[32]

Wavelet analysis is used often for image compression and reconstruction, but can also be used to provide information about an image at several scales.[33] Texture analysis, such as the calculation of co-occurrence features, can be carried out

on the approximation or detail coefficients output from various levels of wavelet analysis.[34]

The application of wavelet analysis to OCT imaging has been predominantly in the area of image denoising[35–37] and the results have been impressive. Adler et al.[38] were able to demonstrate an improvement in signal-to-noise ratio of >7 dB using a two-dimensional wavelet filter, while Puvanathasan and Bizheva[39] report an improvement of ~10 dB in the image signal-to-noise ratio when using a denoising algorithm based on wavelet analysis.

Recently, wavelet analysis has been applied to OCT images for purposes other than image denoising. In 2005, Essock et al.[40] studied the possibility of using wavelet analysis on OCT images of the human retina to help diagnose glaucoma. During the study, 134 patients were imaged with OCT and the images analyzed. The second-level approximation coefficients and the Fourier transform of the second-level detail coefficients were normalized and used as image features. Principal components analysis was used to reduce the number of features, and linear discriminant analysis was used as a classifier. The algorithm was very successful, with an area under the receiver operating characteristic (AUC) of 0.947 for the earliest stage of glaucoma, and increasing to 0.997 for the later stages. The AUC is a measure of overall discrimination ability.[41]

Although a combination of wavelet and texture analysis has not previously been applied to OCT images for the purpose of detecting cancerous tissue, it has been applied to mammograms by Wei et al.[42] for the recognition of cancerous breast tissue. The study used 672 regions of interest extracted from mammograms and calculated co-occurrence features on both the original region and the first four approximation matrices output by the wavelet analysis. When classifying the regions as masses or normal tissue, the method yielded an AUC of 0.86.

When considering system differences, texture analysis on the output of wavelet analysis may have advantages over texture analysis of images directly. The filtering of the data at each step serves to filter out noise, while analyzing the image at courser scales reduces the effect of resolution differences. Furthermore, treating columns and rows separately allows directional properties present in the image to be emphasized or deemphasized in the resulting matrices, which may help differentiate between tissue types.

3 Methods

As an example of the effect of system differences on texture analysis algorithms for OCT, two algorithms developed to recognize bladder cancer were developed and trained on a data set provided by the George Washington University Medical Center (GWUMC) and tested on a data set provided by the Baylor College of Medicine. One algorithm, which shall be referred to as ALG1, was developed without concern for system differences, and the other, which shall be referred to as ALG2, took system differences into consideration.

3.1 Dataset Description

Both the GWUMC and Baylor studies used OCT imaging systems manufactured by the Imalux Corporation, although the central wavelength of the systems, the signal-to-noise ra-
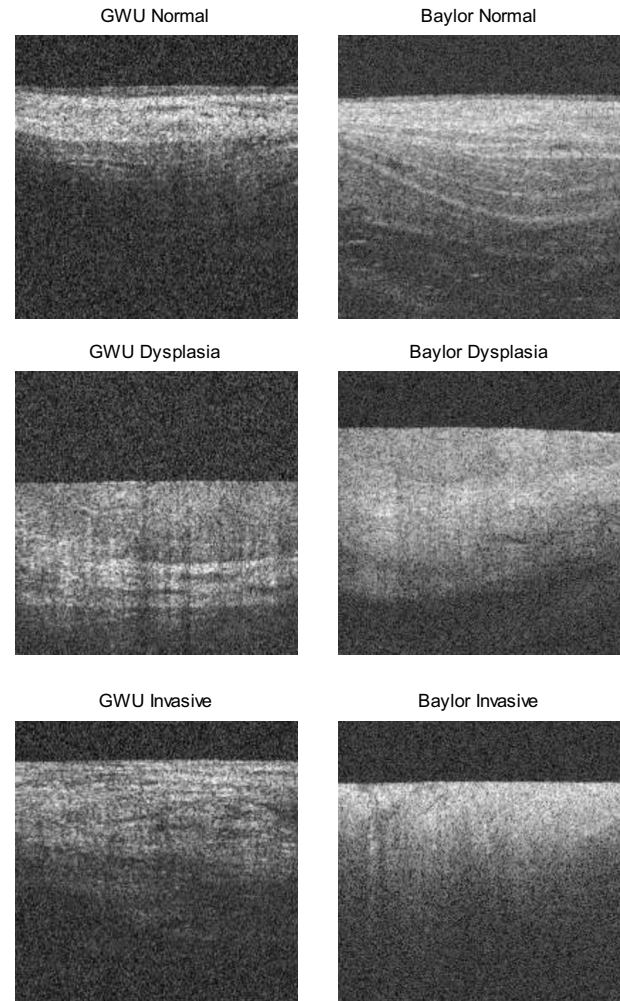

Fig. 2 Example images of healthy, dysplastic, and invasive bladder tissue from each data set.

tios, the range of intensity values, and the pixel sizes were different. During both studies, patients underwent a standard cystoscopic examination. Visually suspect lesions, as well as normal-appearing urothelial tissue, were photographed, scanned with OCT, and biopsied. The scans, which generated $200 \times 200-$pixel images, were obtained by placing the end-firing OCT probe on the desired site perpendicular to the wall of the bladder. Biopsy specimens were preserved in formalin for standard histopathologic analysis and served as the gold standard for the study.

The GWUMC study included 196 images taken from 22 patients using an OCT imaging system with a 980-nm central wavelength. Details of the original GWUMC study can be found in Ref. 43. The Baylor study included 96 images taken from 34 patients using an OCT imaging system with a 1310-nm central wavelength. Because systems with smaller wavelengths have better resolution but less penetration depth, the GWUMC system had better resolution, while the Baylor system had increased penetration. Figure 2 shows example images of healthy, dysplastic, and invasive bladder tissue taken with each system.

3.2 Preprocessing

Histogram analysis was used to locate a threshold within each image that indicated the difference between areas having very limited signal and areas representing the bladder. The threshold then was used to separate the portion of the image-containing "background" from the portion to be analyzed. No further preprocessing was done on the images before they were analyzed for training or testing of ALG1.

Prior to being analyzed for training or testing of ALG2, however, the images were normalized, resized, and cropped. The maximum and minimum intensity values for each data set were determined and used to normalize the intensity values over the range of 0–255. The image size is contained in the header of each image and was used to resize images. All images were resized to have a square pixel size of 12 × 12 μm. The width and depth of each pixel were adjusted to be the same size so that the resulting images would be isotropic. The value of 12 μm was selected as the width and depth of a pixel because it was larger than the estimated pixel size for all of the images. The largest estimated pixel size was the width associated with some of the Baylor images at 11.5 μm/pixel. Finally, the images were cropped to have a pixel width divisible by eight and to have a height of 96 pixels (864 μm). A depth of 864 μm was selected to match the system with the least penetration while maintaining a number of pixels that was divisible by eight. The images needed to have pixel dimensions divisible by eight to permit three levels of wavelet analysis.

3.3 Texture Analysis

Seventy-seven texture features, including Laws's texture features, histogram features, and co-occurrence features, were calculated for the images. The histogram features were calculated using 8, 32, and 128 bins. The co-occurrence features were calculated using 8, 32, and 128 bins, and using a neighbor defined as one pixel to the right, as well as one pixel down.

In an attempt to further reduce the effect of system differences on ALG2, we extracted additional texture features from the approximation and detail matrices that resulted from three levels of wavelet analysis using the Symlet fourth-order wavelet.[44]

To remove redundant features from the feature sets, we examined each set for correlation between features. Correlation values for each feature pair were calculated and normalized, and one feature removed from consideration if the correlation value had an absolute value of >0.9. Of the 77 features under consideration for ALG1, removing the correlated features left 22. Of the 1001 features under consideration for ALG2, removing the correlated features left 325. To allow comparison between features, the remaining features were normalized over the range of 0–255 before continuing with the analysis.

3.4 Algorithm

The algorithms were designed to use discriminant analysis to compare the distance between the texture features for each image and the means of the texture features representing the noncancerous and cancerous groups. The image would be declared cancerous if the texture features were closer to the
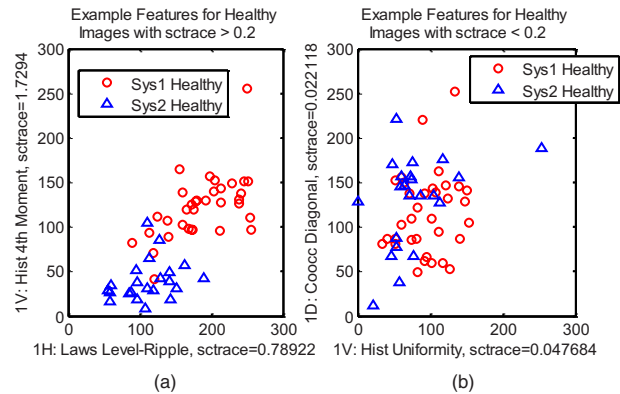

Fig. 3 Relationship of trace value to system dependence: (a) The features have *sctrace* values of >0.2, and the feature values for the two systems are separated into two well-defined clusters. (b) The features have *sctrace* values of <0.2 and the feature values for the two systems are intermixed.

means for the cancerous group, and noncancerous if they were closer to the mean for the noncancerous group. On the basis of prior research indicating that images of normal and dysplasia/carcinoma *in situ* tissue provide the best separation between classes,[12] normal images were used as representative of the noncancerous group, and dysplasia and carcinoma *in situ* (CIS) were used to represent the cancerous group. A more complex algorithm using a treelike structure of comparisons or more classes representing different pathologies would require more data to avoid overtraining.

It is known that, for finite sample sizes, there is an optimal number of features; an increase results in performance deterioration.[45] As more features are used and the algorithm becomes more tuned to the training data set, the algorithm loses its generality or ability to classify a more diverse data set. Consequently, it was necessary to select a subset of the available texture features for use in each of the algorithms.

Prior to feature selection for ALG2, however, system-dependent features were removed from consideration. The trace of the ratio of between-class scatter S_B to within-class scatter S_W (*sctrace*) was calculated for each feature using the normal images from the GWUMC and Baylor data sets as the two classes. The *sctrace* is a standard measure of the separation of two classes and is larger for classes that are well separated and lower for classes that overlap. To ensure overlap of the features from the two data sets, only features that had an *sctrace* value of <0.2 were maintained for consideration during feature selection of ALG2. The limit of 0.2 was selected because when plotted it was possible to see differences between system clusters if the *sctrace* value was >0.2, but not possible when the value was <0.2. Figure 3 shows an example of the distribution of feature values for a set of features with *sctrace* values >0.2 and for a set of features with *sctrace* values <0.2. Only images with normal pathology were included in the *sctrace* calculation to remove the effect of greatly varying pathology on the values. It was assumed that images of normal pathology would exhibit well-defined layers, as recognized by Feldchtein et al.,[46] and have less image-to-image structural variations than varying degrees and types of abnormal pathology.

For both ALG1 and ALG2, the *sctrace* of the texture features of the normal and dysplasia/CIS images from the GWUMC data set was used to determine which subset of one, two, or three features would provide the largest separation between the groups of images. The subset with the highest sc*trace* value was selected for use in the algorithm.

After selection of the features, normal and dysplasia/CIS images from the GWUMC data set were used as training data to determine the mean vectors and covariance matrices for the two classes. Each of the images in the Baylor dataset was then tested using discriminant analysis to classify the image as cancerous or noncancerous.

4 Results

The features selected for use in ALG1 were:

1. The mean of the histogram using 32 bins.

2. The amount of energy along the diagonal calculated using a co-occurrence matrix with 32 bins and neighbor defined as one pixel down.

3. The energy texture feature calculated using a co-occurrence matrix with 128 bins and neighbor defined as one pixel down.

As mentioned previously, prior to feature selection for ALG2, all features that were determined to be system dependent were removed from consideration. Interestingly, all of the features calculated directly on the image were removed from consideration during this step. However, after this step, only about half the features calculated on the output of wavelet analysis had been removed from consideration, leaving 112 features still in consideration. Furthermore, an unusually large portion of the remaining features, 32%, were Laws's texture features, whereas only 18% of all features were Laws's texture features. The following features selected for use in ALG2 were:

1. On the first-level wavelet vertical detail matrix, the amount of energy along the diagonal calculated using a co-occurrence matrix with 128 bins and neighbor defined as one pixel down.

2. On the second-level wavelet horizontal detail matrix, the Laws's edge-spot texture feature.

3. On the third-level wavelet vertical detail matrix, the Laws's level-spot texture feature.

When the data acquired during the Baylor study were tested on ALG1 using the GWUMC data as the training data, all of the test images, regardless of pathology were classified as cancerous. In comparison, when ALG1 was tested on the GWUMC data set using leave-one-out cross-validation, the resulting sensitivity was 73%, with a specificity of 69%. The poor results noted when testing the Baylor images can be directly attributed to differences between the imaging systems used in the two studies. The features selected for use in ALG1 were affected substantially by those differences and caused all of the images to fall into the "cancer" category. Figure 4 shows the distribution of the texture features used in ALG1 for the two systems. The texture features are simply too different from one another to allow the images from different data sets to be compared to one another.

When the data acquired during the Baylor study were tested on ALG2 (which had taken system differences into consideration) using the GWUMC data as the training data, the
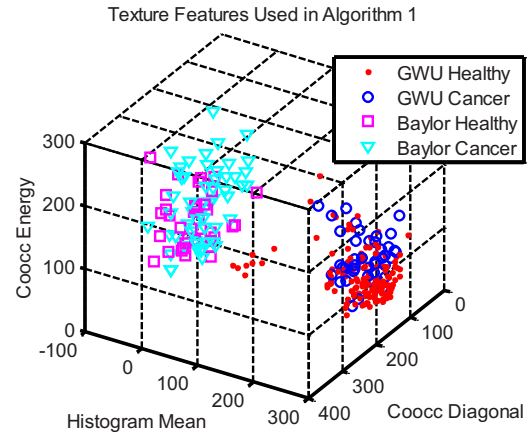

Fig. 4 Distribution by system of texture features used in algorithm 1.

results indicated a sensitivity of 87% and a specificity of 58%. Figure 5 shows the distribution of the texture features used in ALG2 for the two systems. The texture features do not cluster separately by system, but form one large cluster.

5 Discussion

Two algorithms were developed and trained on a set of images taken with one imaging system and then tested on a set of images taken with another imaging system with very different characteristics. ALG1, which was developed ignoring the existence of system differences, classified all images in the testing set as "cancer" regardless of the true pathology. When the distribution of the texture features used by the algorithm was plotted, it became evident that the selected features were highly dependent on the imaging system.

The ranges of intensity values for the two systems were very different, affecting texture features calculated using histogram analysis or Laws's method. Normalizing the images using the intensity range evident in each data set improved the correlation between some of the texture features for the two systems, but was not sufficient to compensate for all of the system differences. The pixel size used in each set of images was also quite different, affecting features calculated using co-occurrence matrices or Laws's method. Resizing the images to have a standard square pixel size improved the correlation between some of these texture features. Differences in
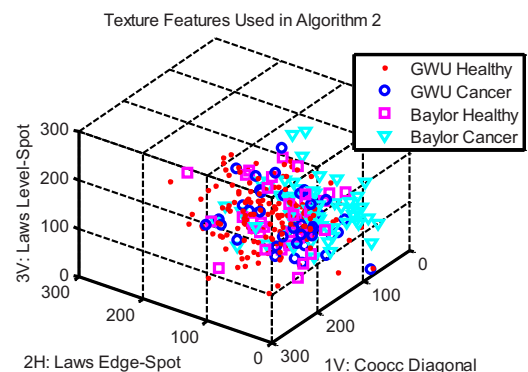

Fig. 5 Distribution by system of texture features used in algorithm 2.

the amount of depth penetration of the different systems were partially accounted for by performing texture analysis only to the penetration depth achieved by the system with the least penetration.

Other system differences that need to be considered are the system resolution and the signal-to-noise ratio of the system. Wavelet analysis offers the ability to analyze images at different scales and to filter those images to emphasize high- or low-frequency content in different directions. By evaluating an image at a larger scale, the effects of resolution differences will be reduced, and by evaluating an image that has passed through a low-pass filter, the effect of different signal-to-noise ratios will be reduced. Not only does wavelet analysis reduce the effect of system differences, it offers another advantage. It provides the opportunity to evaluate an image that has been filtered to emphasize horizontal or vertical structure, which in the case of bladder cancer detection, is very useful due to the layered structure present in healthy bladder tissue and absent in diseased tissue.

Finally, while various preprocessing techniques were implemented to reduce the effect of system differences, some differences remained. It was therefore necessary to check for system differences before feature selection. Texture features whose values for normal images taken from both data sets did not form one cluster were not considered during feature selection. All the texture features calculated on the images directly were removed from consideration, strengthening the argument that normalization and resizing alone are not sufficient when creating system-independent algorithms. On the other hand, approximately half the texture features calculated on the output of wavelet analysis passed the system differences check, supporting the suggestion that use of wavelet analysis helps circumvent problems caused by system differences. Of additional interest is the fact that Laws's texture features calculated on the output of wavelet analysis had a large representation in the set of features determined to be system independent. The robustness of those texture features in the face of system differences may be due to their structural basis as compared to the statistical approach used with co-occurrence matrices and histogram analysis.

The algorithm was tested on two sets of data taken with two different imaging systems. To confirm the system independence of the algorithm, it would be necessary to test data sets acquired with other systems. The results achieved indicate that if system differences are accounted for during algorithm development, it is possible to develop algorithms that can be trained with data from one system and used successfully on images collected with a second.

Although the results of the second algorithm were promising and demonstrate system independence, the reliability of the second algorithm needs to be improved for it to be clinically useful. More research and data are necessary to allow development of a more complex algorithm, which by taking into account different cancer and noncancer pathologies would have the potential to be more accurate. The limited amount of data currently available limits the complexity of the algorithm.

The suggestions mentioned in this paper require that images be compared at both the poorest resolution and poorest penetration depth considered. This, however, may diminish the ability to develop algorithms that take advantage of supe-rior resolution or superior depth penetration. Superior resolution may improve the ability to differentiate between pathological differences close to the surface, such as the difference between dysplasia and CIS, while superior depth penetration will be required to grade an invasive tumor. It may, therefore, be beneficial to develop algorithms that operate within certain well-described limitations. The portability of an algorithm may become another parameter to consider, just as one considers whether higher resolution or increased penetration is necessary.

6 Conclusion

For automated algorithms to be applicable to more than one OCT system, steps must be taken during algorithm development to account for system differences. Preprocessing steps to address pixel size, image depth, and intensity range reduce the effect of some system differences, but are not sufficient. It is necessary to consider system differences during feature selection and remove all features from consideration that are still affected by system differences after preprocessing. Unfortunately, removing these features from a collection of features calculated for an image will significantly reduce the selection of features available for tissue characterization. The ability of wavelet analysis to reduce the noise in images, consider images at different scales, and break images into horizontal and vertical components, provides a number of additional representations from which useful texture features can be calculated.

An algorithm was introduced that when developed and trained on one system was able to successfully classify images taken with a second imaging system that had very different characteristics. The sensitivity when testing the images taken with the second imaging system was 87% and the specificity 58%. The algorithm considered system differences when preprocessing the images, considered texture features calculated on the output of wavelet analysis, and excluded features from consideration that were classified as system dependent based on the trace of the ratio of the between-class scatter to the within-class scatter for normal images taken with the two imaging systems.

References

1. 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.* **1**(9), 970–972 (1995).
2. D. C. Fernández, H. M. Salinas, and C. A. Puliafito, "Automated detection of retinal layer structures on optical coherence tomography images," *Opt. Express* **13**(25), 10200–10216 (Dec. 2005).
3. B. Cense, T. C. Chen, B. H. Park, M. C. Pierce, and J. F. deBoer, "*In vivo* birefringence and thickness measurements of the human retinal nerve fiber layer using polarization-sensitive optical coherence tomography," *J. Biomed. Opt.* **9**(1), 121–125 (Jan./Feb. 2004).

4. M. M. Garvin, M. M. Abramoff, R. R. Kardon, S. S. Russell, X. X. Wu, and M. M. Sonka, "Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search," *IEEE Trans. Med. Imaging* **27**(10), 1495–1505 (2008).

5. M. R. Hee, J. A. Izatt, E. A. Swanson, D. Huang, J. S. Schuman, C. P. Lin, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography of the human retina," *Arch. Ophthalmol. (Chicago)* **113**(3), 325–332 (1995).

6. N. A. Patel, D. L. Stamper, and M. E. Brezinski, "Review of the ability of optical coherence tomography to characterize plaque, including a comparison with intravascular ultrasound," *Cardiovasc. Intervent Radiol.* **28**(1), 1–9 (2005).

7. J. Rogowska, N. Patel, S. Plummer, and M. E. Brezinski, "Quantitative optical coherence tomographic elastography: method for assessing arterial mechanical properties," *Br. J. Radiol.* **79**, 707–711 (2006).

8. O. C. Raffel, G. J. Tearney, D. D. Gauthier, E. F. Halpern, B. E. Bouma, and I. K. Jang, "Relationship between a systemic inflammatory marker, plaque inflammation, and plaque characteristics determined by intravascular optical coherence tomography," *Arterioscler., Thromb., Vasc. Biol.* **27**, 1820–1827 (2007).

9. J. M. Zara and C. A. Lingley-Papadopoulos, "Endoscopic OCT approaches toward cancer diagnosis," *IEEE J. Sel. Top. Quantum Electron.* **14**(1), 70–81 (Jan./Feb. 2008).

10. 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**, 81-90 (2006).

11. 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).

12. C. A. Lingley-Papadopoulos, M. H. Loew, M. J. Manyak, and J. M. Zara, "Computer recognition of cancer in the urinary bladder using optical coherence tomography and texture analysis," *J. Biomed. Opt.* **13**(2), 024003 (2008).

13. G. J. Tearney, M. E. Brezinski, J. F. Southern, B. E. Bouma, S. A. Boppart, and J. G. Fujimoto, "Optical biopsy in human urologic tissue using optical coherence tomography," *J. Urol. (Baltimore)* **157**(5), 1915–1919 (May. 1997).

14. S. A. Boppart, J. M. Hermann, C. Pitris, D. L. Stamper, M. E. Brezinski, and J. G. Fujimoto, "Real-time optical coherence tomography for minimally invasive imaging of prostate ablation," *Comput. Aided Surg.* **6**, 94–103 (2001).

15. F. Feldchtein, V. Gelikonov, R. Iksanov, G. Gelikonov, , R. Kuranov, A. Segeev, N. Gladkova, M. Ourutina, D. Reitz, and J. Warren, "*In vivo* OCT imaging of hard and soft tissue of the oral cavity," *Opt. Express* **3**(6), 239–250 (Sep. 1998).

16. J. J. Armstrong, M. S. Leigh, D. D. Sampson, J. H. Walsh, D. R. Hillman, and P. R. Eastwood, "Quantitative upper airway imaging with anatomic optical coherence tomography," *Am. J. Respir. Crit. Care Med.* **173**(2), 226–233 (2006).

17. A. M. Karamzadeh, R. Jackson, S. Guo, J. M. Ridgway, H. S. Wong, G. S. Ahuja, M. C. Chao, L. L. Liaw, Z. Chen, and B. J. F. Wong, "Characterization of submucosal lesions using optical coherence tomography in the rabbit subglottis," *Arch. Otolaryngol. Head Neck Surg.* **131**, 499–504 (2005).

18. P. Wilder-Smith, M. J. Hammer-Wilson, J. Zhang, Q. Wang, K. Osann, Z. Chen, H. Wigdor, J. Schwartz, and J. Epstein, "*In vivo* imaging of oral mucositis in an animal model using optical coherence tomography and optical Doppler tomography," *Clin. Cancer Res.* **13**(8), 2449–2454 (Apr. 2007).

19. K. W. Gossage, C. M. Smith, E. M. Kanter, L. P. Hariri, A. L. Stone, J. J. Rodriguez, S. K. Williams, and J. K. Barton, "Texture analysis of speckle in optical coherence tomography images of tissue phantoms," *Phys. Med. Biol.* **51**, 1563–1575 (2006).

20. A. F. Fercher, W. Drexler, C. K. Hitzenberger, and T. Lasser "Optical coherence tomography—principles and applications," *Rep. Prog. Phys.* **66**(2), 239–303 (2003).

21. A. F. Fercher, "Optical coherence tomography," *J. Biomed. Opt.* **1**(3), 157–173 (1996).

22. M. E. Brezinski and J. G. Fujimoto, "Optical coherence tomography: high-resolution imaging in nontransparent tissue," *IEEE J. Sel. Top. Quantum Electron.* **5**(4), 1185–1192 (July/August 1999).

23. D. Huang, E. Swanson, C. Lin, J. Schuman, W. Stinson, W. Chang, M. Hee, T. Flotte, K. Gregory, C. Puliafito, and J. Fujimoto, "Optical coherence tomography," *Science* **254**, 1178–1181 (August 1991).

24. C. Mason, J. F. Markusen, M. A. Town, P. Dunnill, and R. K. Wang, "The potential of optical coherence tomography in the engineering of living tissue," *Phys. Med. Biol.* **49**, 1097–1115 (2004).

25. R. A. Bourne, F. A. Madeiros, C. Bowd, K. Jahanbakhsh, L. M. Zangwill, and R. N. Weinreb, "Comparability of retinal nerve fiber layer thickness measurements of optical coherence tomography instruments," *Invest. Ophthalmol. Visual Sci.* **46**(4), 1280–1285 (2005).

26. M. Sehi, D. C. Guaqueta, W. J. Feuer, and D. S. Greenfeld, "A comparison of structural measurements using 2 stratus optical coherence tomography instruments," *J. Glaucoma* **16**(3), 287–292 (2007).

27. M. E. Brezinski, G. J. Tearney, B. E. Bouma, J. A. Izatt, M. R. Hee, E. A. Swanson, J. F. Southern, and J. G. Fujimoto, "Optical coherence tomography for optical biopsy. Properties and demonstration of vascular pathology," *Circulation* **93**(6), 1206–1213 (1996).

28. 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).

29. 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).

30. C. A. Lingley-Papadopoulos, M. H. Loew, and J. M. Zara, "Real-time bladder-layer recognition: an approach to optical biopsy," presented at *4th IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, pp. 1276–1279 (Apr 2007).

31. Y. Chen, A. D. Aguirre, P. Hsiung, S. Huang, H. Mashimo, J. M. Schmitt, and J. G. Fujimoto, "Effects of axial resolution improvement on optical coherence tomography (OCT) imaging of gastrointestinal tissues," *Opt. Express* **16**(4), 2469–2485 (Feb. 2008).

32. M. Misiti, Y. Misiti, G. Oppenheim, and J. Poggi, *Wavelet Toolbox 4 User's Guide*, MathWorks, Inc., Natick, MA (2008).

33. A. Laine, J. Fan, "Texture classification by wavelet packet signatures," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(11), 1186–1191 (1993).

34. S. Arivazhagan and L. Ganesan, "Texture classification using wavelet transform," *Pattern Recogn. Lett.* **24**, 1513–1521 (2003).

35. S. H. Xiang, L. Zhou, and J. M. Schmitt, "Speckle noise reduction for optical coherence tomography," presented at *Opt. Imaging Techniques for Biomonitoring III* Vol. **3196**, Jan. 13, 1998.

36. H. Choi, T. E. Milner, and A. C. Bovik, "Speckle noise reduction and segmentation on polarization sensitive optical coherence tomography images," in *Proc. of 25th Annual Int. Conf. of the IEEE EMBS*, Cancun, Mexico, Sept 17–21, 2003.

37. A. Ozcan, A. Bilenca, A. E. Desjardins, B. E. Bruma, and G. J. Tearney, "Speckle reduction in optical coherence tomography images using digital filtering," *J. Opt. Soc. Am. A* **24**(7), 1901–1910 (2007).

38. D. C. Adler, T. H. Ko, and J. G. Fujimoto, "Speckle reduction in optical coherence tomography images by us of a spatially adaptive wavelet filter," *Opt. Lett.* **29**(24), 2878–2880 (2004).

39. P. Puvanathasan and K. Bizheva, "Speckle noise reduction algorithm for optical coherence tomography based on interval type II fuzzy set," *Opt. Express* **15**(24), 15747–15758 (2007).

40. E. A. Essock, Y. Zheng, and P. Gunvant, "Analysis of GDx-VCC poloarimetry data by wavelet-fourier analysis across glaucoma stages," *Invest. Ophthalmol. Visual Sci.* **46**, 2838–2847 (2005).

41. S. Walter, "The partial area under the summary ROC curve," *Stat. Med.* **24**(13), 2025–2040 (2005).

42. D. Wei, H. P. Chan, M. A. Helvie, B. Sahiner, N. Petrick, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis," *Med. Phys.* **22**(9), 1501–1513 (1995).

43. M. J. Manyak, N. D. Gladkova, J. H. Makari, A. Shwartz, E. V. Zagaynova, L. Zolgahari, 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).

44. M. Misiti, Y. Misiti, G. Oppenheim, J. Poggi, *Wavelet Toolbox 4 User's Guide*, The MathWorks, Inc. (2008).

45. R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed., Wiley, Hoboken, NJ (2001).

46. F. I. Feldchtein, G. V. Gelikonov, V. M. Gelikonov, R. V. Kuranov, A. M. Sergeev, N. Gladkova, A. Shakhov, N. Shakhova, L. Snapova, A. Terent'eva, E. Zagainova, Y Chumakov, and I. Kuznetzova, "Endoscopic applications of optical coherence tomography," *Opt. Express* **3**(6), 257–270 (Sept. 1998).