

# Automated algorithm for differentiation of human breast tissue using low coherence interferometry for fine needle aspiration biopsy guidance

## Brian D. Goldberg

Massachusetts General Hospital  
Wellman Center for Photomedicine  
40 Blossom Street  
Boston, Massachusetts 02114  
and  
Harvard-MIT  
Division of Health Sciences and Technology  
77 Massachusetts Avenue, E25-519  
Cambridge, Massachusetts 02139

## Nicusor V. Iftimia\*

## Jason E. Bressner

Massachusetts General Hospital  
Wellman Center for Photomedicine  
40 Blossom Street  
Boston, Massachusetts 02114

## Martha B. Pitman

Massachusetts General Hospital  
Department of Pathology  
55 Fruit Street  
Boston, Massachusetts 02114

## Elkan Halpern

Massachusetts General Hospital  
Institute for Technology Assessment  
101 Merrimac Street  
Boston, Massachusetts 02114

## Brett E. Bouma

Massachusetts General Hospital  
Wellman Center for Photomedicine  
40 Blossom Street  
Boston, Massachusetts 02114  
and  
Harvard-MIT  
Division of Health Sciences and Technology  
77 Massachusetts Avenue, E25-519  
Cambridge, Massachusetts 02139

## Guillermo J. Tearney

Massachusetts General Hospital  
Wellman Center for Photomedicine  
40 Blossom Street  
Boston, Massachusetts 02114  
and  
Massachusetts General Hospital  
Department of Pathology  
55 Fruit Street  
Boston, Massachusetts 02114

\*Currently with Physical Sciences, Inc, 20 New England Business Ctr. Drive, Andover, MA 01810.

Address all correspondence to Brian Goldberg, Wellman Center for Photomedicine, Massachusetts General Hospital, 50 Blossom St.-BAR 710, Boston, MA 02114. Tel: 617-726-7842; Fax: 617-726-4103; E-mail: bdg@mit.edu

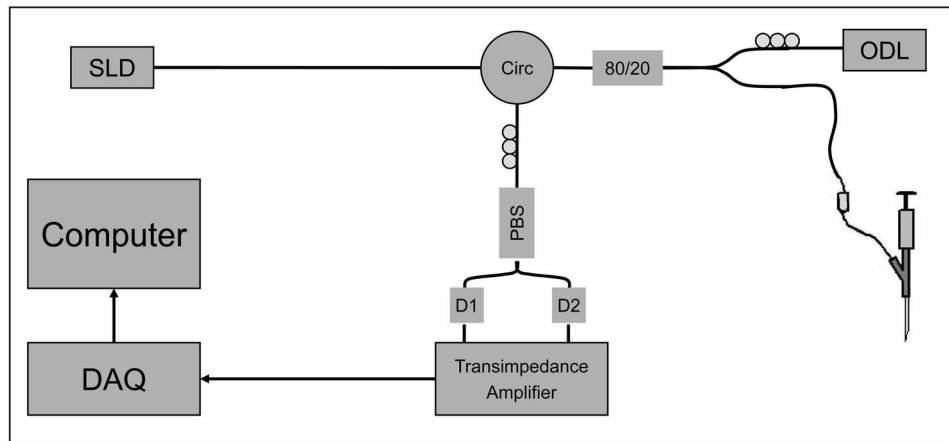
**Abstract.** Fine needle aspiration biopsy (FNAB) is a rapid and cost-effective method for obtaining a first-line diagnosis of a palpable mass of the breast. However, because it can be difficult to manually discriminate between adipose tissue and the fibroglandular tissue more likely to harbor disease, this technique is plagued by a high number of nondiagnostic tissue draws. We have developed a portable, low coherence interferometry (LCI) instrument for FNAB guidance to combat this problem. The device contains an optical fiber probe inserted within the bore of the fine gauge needle and is capable of obtaining tissue structural information with a spatial resolution of 10  $\mu\text{m}$  over a depth of approximately 1.0 mm. For such a device to be effective clinically, algorithms that use the LCI data must be developed for classifying different tissue types. We present an automated algorithm for differentiating adipose tissue from fibroglandular human breast tissue based on three parameters computed from the LCI signal (slope, standard deviation, spatial frequency content). A total of 260 breast tissue samples from 58 patients were collected from excised surgical specimens. A training set ( $N=72$ ) was used to extract parameters for each tissue type and the parameters were fit to a multivariate normal density. The model was applied to a validation set ( $N=86$ ) using likelihood ratios to classify groups. The overall accuracy of the model was 91.9% (84.0 to 96.7) with 98.1% (89.7 to 99.9) sensitivity and 82.4% (65.5 to 93.2) specificity where the numbers in parentheses represent the 95% confidence intervals. These results suggest that LCI can be used to determine tissue type and guide FNAB of the breast. © 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2837433]

**Keywords:** breast cancer; low coherence interferometry; needle guidance; fine needles aspiration; breast biopsy; automated.

Paper 07068R received Feb. 28, 2007; revised manuscript received Aug. 22, 2007; accepted for publication Sep. 25, 2007; published online Feb. 14, 2008.

## 1 Introduction

Fine needle aspiration biopsy (FNAB) is often the first line of diagnosis for a palpable mass.<sup>1-3</sup> To perform a FNAB, the mass is manually stabilized, a small diameter needle (typically 23 to 25 gauge) is inserted into the mass, and a small amount of tissue or fluid is aspirated into the needle. The aspirate within the bore of the needle is then expressed onto a slide, smeared, stained, and examined by a pathologist. Due to the small size of the needle, patient discomfort is generally limited to the initial stick of the needle. Complications including hematoma and infection are rare. The simplicity of FNAB significantly reduces the time and cost of obtaining an initial



**Fig. 1** Schematic of the LCI biopsy guidance instrument. A super luminescent diode (SLD) is sent through a circulator (Circ) and an 80/20 splitter. The reference arm optical delay line (ODL) consists of a retroreflector mounted on a galvanometer-driven lever arm. The LCI-FNA probe consists of a single mode fiber inserted through a hub connecting the syringe and the needle. An optical connector attaches the LCI-FNA probe to the sample arm of the interferometer.

diagnosis compared with core or excisional tissue biopsy and allows rapid feedback to both the clinician and patient. In addition, comparisons of the sensitivities and specificities of core needle biopsy (CNB) and FNAB for palpable masses show them to be high and similar.<sup>4,5</sup> As a result, FNAB has become a frequently used diagnostic tool for the evaluation of many superficial, palpable masses.

Manual palpation of a superficial mass is often the only cue for determining the optimal position of the needle in tissue during biopsy. As a result, FNAB can frequently be non-diagnostic, especially with an inexperienced operator.<sup>6–10</sup> Sample adequacy is graded on a sliding scale based on the degree of epithelial cellularity from which a diagnosis can be made.<sup>9,10</sup> Nondiagnostic samples are completely void of epithelial cells and consist primarily of adipose cells and cyst fluid.<sup>10</sup> When not guided by an imaging modality, breast FNABs obtain diagnostic tissue in approximately 65 to 78% of cases.<sup>6,8,9</sup> This difficulty is particularly problematic when performing FNABs in locations that are rich in adipose tissue, such as the breast and axilla. One method of increasing FNAB yield is concomitant use of noninvasive imaging devices, such as ultrasound, to guide needle placement. Radiologic guidance is almost always employed when FNAB is performed on nonpalpable masses. Although the addition of noninvasive imaging technology has been shown to increase FNAB yield, it is time-consuming, relatively expensive, and often requires additional personnel with specialized expertise.<sup>11</sup>

Recently, a portable, low-cost device based on low coherence interferometry (LCI) has been developed for fine needle aspiration (FNA) needle guidance.<sup>12</sup> LCI is an optical ranging technique that is capable of measuring depth-resolved (axial,  $z$ ) tissue structure, birefringence, flow (Doppler shift), and spectra at a micrometer-level resolution.<sup>13–15</sup> Other groups have investigated the use of needle-based optical probes for biopsy guidance based on imaging<sup>16</sup> or by direct measure of tissue optical properties such as multispectral reflection analysis,<sup>17</sup> scattering coefficient,<sup>18</sup> and refractive index<sup>19</sup> measurements. Miniature LCI needle probes have also been used

to correlate brain motion with electrocardiogram waves in a minimally invasive fashion.<sup>20</sup>

An initial feasibility study performed on excised breast surgical specimens indicated that LCI may have the potential for classifying adipose and fibroglandular breast tissue based on the slope and standard deviation of the axial depth profiles.<sup>12</sup> The sample size for this study was small and the accuracy of LCI for breast tissue type diagnosis was therefore not evaluated. Furthermore, this data was analyzed in a semi-automatic fashion that is not suitable for clinical use; the minimum and maximum boundaries over which the data were analyzed were selected manually. Here, we present an automated algorithm for classifying adipose and fibroglandular breast tissues that includes an additional, independent parameter that quantifies LCI signal spatial frequencies. The accuracy of this algorithm was determined prospectively in a blinded fashion on a cohort of 260 biopsy correlated LCI scans from 58 patients. Intrasample variability of the algorithm was also tested. Similar classification parameters were recently introduced to develop methods for computationally driven differentiation of human breast tissue.<sup>21</sup> However, to our knowledge, our study represents the first complete study to test the efficacy of such parameters for classification of human breast tissue.

## 2 Methods

### 2.1 System Description

The LCI system and probe have been described previously and are shown schematically in Fig. 1.<sup>12</sup> Briefly, the LCI system consisted of a nonreciprocal fiber optic Michelson interferometer. A broadband super luminescent diode (SLD) centered at 1310 nm with a full width at half maximum bandwidth of 50 nm (Optiphase, Inc., Van Nuys, California) was used as a light source. The axial resolution was  $\sim 15 \mu\text{m}$  in air, or  $11 \mu\text{m}$  in tissue ( $n=1.4$ ). Light from the source was transmitted through the first output port of a circulator and an 80/20 splitter, which directed approximately  $750 \mu\text{W}$  to the

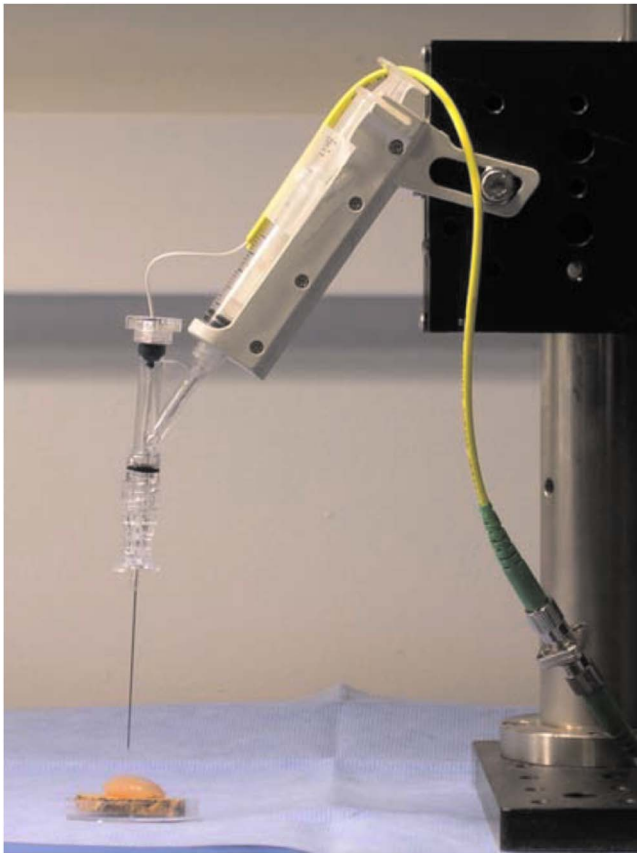


Fig. 2 FNA gun mounted with needle and positioned above sample.

sample arm. LCI depth scans (A-lines) were obtained at a rate of 40 Hz and the path length in the reference arm was scanned by illuminating a retroreflector mounted on a galvanometer-driven lever arm (Model 6220, Cambridge Technology, Lexington, Massachusetts). Light from the sample and reference arms were recombined and directed toward a polarization beamsplitter and two photodetectors, enabling polarization diverse detection. Shot noise limited detection was achieved with a maximum signal-to-noise ratio (SNR) of 101 dB.

The optical probe consisted of a single mode optical fiber inserted through the bore of a 23 gauge ( $\sim 570\text{-}\mu\text{m}$  outside diameter) FNA needle. No focusing lens was used. The needle was attached to a regular syringe through a hub (Model 54501, Inrad, Northvale, New Jersey). The syringe was held within a FNA biopsy gun. The fiber probe was designed to be simple and therefore inexpensive. Because the fiber core aperture was always at the needle tip, there was no uncertainty regarding the probe location, and the interrogated tissue was directly in front of the needle location. Although the fiber probe was housed within a FNA needle, no tissue aspirates were collected during the measurements.

## 2.2 Experimental Protocol

Excised surgical specimens were collected, stored in 10% phosphate buffered saline and data was collected at  $37^\circ\text{C}$  within 24 h of collection. The needle and FNA gun were secured onto a vertical translation stage as shown in Fig. 2.

During imaging, samples were placed flat on a piece of corrugated cardboard within a Petri dish and positioned under the needle probe. The needle was lowered onto the sample until the fiber surface came into contact with the sample. Ten consecutive A-lines were collected at each site. Following imaging, the needle was raised, and the needle location was marked with India ink. The samples were then fixed in formalin. Histologic sections were obtained and stained with hematoxylin and eosin.

Histology slides were read by a pathologist who was blinded to the LCI data. Slides were randomly ordered to avoid bias from reading samples from the same patient consecutively. Histology samples were grouped into two critical cases for this application—adipose and fibroglandular tissue types. Fibroglandular tissues included benign fibrous parenchyma, adenocarcinoma, and ductal carcinoma *in situ* (DCIS) tissue types. Only homogeneous samples classified as pure adipose or fibroglandular tissue were included for parameter extraction and algorithm development/classification. Samples with significant heterogeneity in the image field as defined by the pathologist or samples where no ink was visible on histology were excluded. Heterogeneous samples were defined as tissues where the ratio of major to minor tissue type was approximately less than 3:1 within 1 mm of the ink mark.

## 2.3 Parameter Extraction

For each sample, 10 consecutive A-lines were acquired. Signal parameters were extracted for each A-line, and the mean value for each parameter was used to represent the sample. Each parameter was calculated using an automated MATLAB script without the need for additional user input other than the sample file. Prior to parameter extraction, the raw LCI interferogram data was converted to depth-dependent reflectivity profiles in the standard fashion.<sup>12</sup> The signal was transformed using discrete Fourier transform (DFT), bandpass filtered, frequency shifted to zero, and inverse transformed. The resulting linear intensity values were then converted to decibel scale by  $20 \log_{10}$  multiplication.

### 2.3.1 Automatic LCI scan boundary extraction

At the beginning and end of the LCI scan, the signal contains data that are not representative of the tissue sample. As a result, prior to parameter extraction, the data must be automatically parsed to determine the segment of the LCI scan that contains tissue reflectivity information. The location of the fiber-sample interface was automatically determined by the following procedure. The noise floor was determined by averaging the signal within the first  $200 \mu\text{m}$  of imaging depth, which corresponded to a region proximal to the fiber-sample interface. All signal points below the threshold were set to be equal to the noise floor, using a threshold of 20 dB. Next, a first-order derivative was computed and the first peak was determined by the first zero crossing of the derivative. To avoid error from specular reflection at the fiber-sample interface, the start index was shifted an additional  $100 \mu\text{m}$  beyond the first peak. Automatic selection of the beginning location of the LCI scan in this manner allowed the effective start index to always fall within signal values representing the tissue structure. The last  $100 \mu\text{m}$  of the LCI signal were also skipped because the signal was generally low in this region.

Thus, the analyzed data consisted of the region from the effective start index to the end of the LCI scan minus the last 100  $\mu\text{m}$ . This algorithm was automated and applied to all LCI scans to determine the data range over which to compute the slope, standard deviation, and spatial frequency content parameters. The average depth over which the signal was analyzed was 936  $\mu\text{m}$  with a range from 820  $\mu\text{m}$  to 1.03 mm.

### 2.3.2 Slope

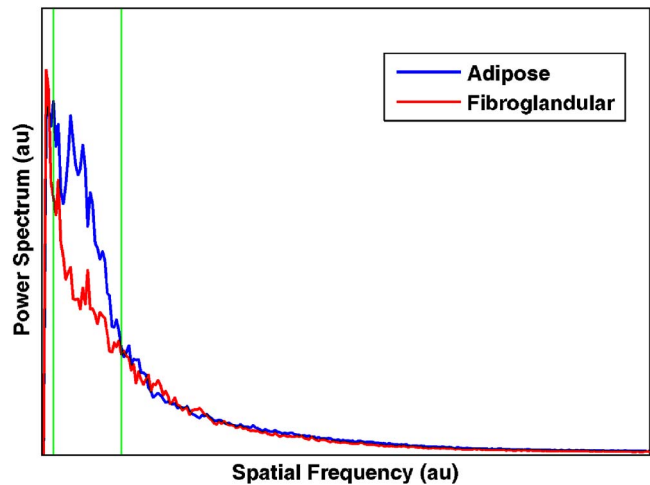
To first order, the LCI reflectivity intensity decreases in accordance with the Beer-Lambert law. At a source wavelength of 1300 nm, tissue optical properties are such that scattering dominates over absorption.<sup>22</sup> Therefore, the slope of the logarithmic axial depth profile is related to the scattering coefficient and can be used as a parameter for classifying tissue type. A higher slope indicates more attenuation and a larger scattering coefficient, whereas, a lower slope indicates a lower scattering coefficient. The slope was calculated by a first-order polynomial fit over the region of interest.

### 2.3.3 Standard deviation

The variation of scattering cross sections within a LCI depth scan can be used as another parameter for classifying tissue type. One way to assess the scattering variance is to measure the slope-subtracted standard deviation of the axial depth profile. If the scattering fluctuates significantly, the reflection profile will have peaks interspersed with periods of low signal and the standard deviation will be high. Conversely, if the scattering is relatively homogeneous, the signal will be more continuous and the standard deviation will be low. To remove the effect of the bulk averaged scattering coefficient, the residual of the linear fit was used to compute the standard deviation.

### 2.3.4 Spatial frequency content

Scattering center distribution, representing the distance between scatterers, may be evaluated by analyzing the spatial frequency components of the signal. The power spectrum of the signal can be interpreted as the signal energy within spatial frequency windows, and the unique signature from different tissue types was recently described as a method for differentiating human breast tissue.<sup>21</sup> The spatial frequency parameter was computed in the following manner. First, as with the computation of the standard deviation, the linear regression was conducted and the residual was used for subsequent processing. Next, the dc component was removed by mean subtraction. Data outside the start and end index were set to zero. The resulting signal was zero mean, with components that fluctuated with varying frequency content depending on tissue type. The DFT was then computed. The spatial frequency parameter was then defined by integrating the magnitude of the spatial frequency content over a particular window. The window was defined by calculating the average DFTs for the entire training set and observing where the adipose and fibroglandular tissue samples differed. A zoomed portion of the mean DFTs for the training set are shown in Fig. 3. The vertical lines represent the width of the integration window.



**Fig. 3** Average power spectrum data from LCI depth scans over the entire training set that show a region of difference between adipose and fibroglandular tissues types. The integration window (green lines) represents the area over which the spatial frequency content parameter is generated. (Color online only.)

## 2.4 Algorithm Model

A multivariate Gaussian model was used for classification. The data set was randomly split into training and validation sets. A pooled estimate of the covariance matrix was used for the training set. The result of the model was an equation for each class that defined the probability that any new set of parameters fell within that class. Prospective analysis was then performed on the validation set. Classification was carried out by extracting parameters for each test sample, calculating the probability of falling within a particular class, and then assigning classification based on the highest probability.

## 2.5 Intrasample Variability

To test the intrasample variability of the device and algorithm, an additional experiment was conducted using another data set. The needle probe was lowered onto each sample and a 10-A-line acquisition was performed. The needle was then raised off of the sample using the vertical translation stage and relowered back onto the sample for an additional 10-A-line acquisition. This process was repeated 10 times so that each sample had 10 data sets of 10 A-lines each all from the same exact location. After the 10 measurement, the needle was raised and the sample was marked with India ink and sent for histology sectioning and staining. Each set of 10 A-lines was processed in the same manner as the earlier experiments so that a single set of extracted parameters characterized each

**Table 1** Training set statistics.

Parameter	Adipose (N=37)	Fibroglandular (N=35)	<i>p</i>
Slope	$-1.74 \pm 1.36$	$-4.39 \pm 2.23$	$<1.7 \times 10^{-7}$
Std. Dev.	$7.23 \pm 1.54$	$5.41 \pm 0.71$	$<2.2 \times 10^{-7}$
Spat. Freq.	$1.42 \pm 0.54$	$0.97 \pm 0.27$	$<2.2 \times 10^{-9}$



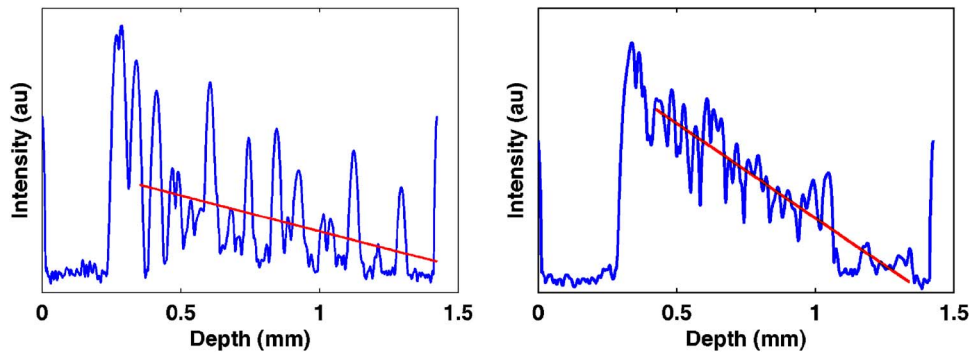


Fig. 4 Characteristic axial depth scans from adipose (left) and fibroglandular (right) human breast tissue. The red line is a first-order polynomial fit through the data. Intensity data shown in logarithmic scale. (Color online only.)

set of A-lines. The samples were then classified using the multivariate Gaussian model. The result was a set of 10 classifications from the same sample at the same location. The intrasample variability was defined as the percentage of misclassified measurements within a particular sample.

## 2.6 Statistical Analysis

The accuracy of LCI for classifying breast tissue type was assessed by comparing the predicted tissue type to the gold standard histopathologic classification. All data processing and parameter extraction were done within MATLAB. Each parameter is listed as  $\mu \pm \sigma$ , where  $\mu$  is the mean, and  $\sigma$  is the standard deviation. The  $p$ -values were calculated using a two-sided unpaired  $t$ -test to determine if the difference in sample means between parameters were statistically significant, and 95% confidence intervals (CI) are also reported.

## 3 Results

Typical LCI profiles of adipose and fibroglandular breast tissue were very different (Fig. 4). The adipose samples contained multiple reflectivity peaks, presumably representing the lipid core and cell membrane interface (Fig. 4). Human adipocytes range in size from 50 to 150  $\mu\text{m}$ , which makes the location of the reflectivity peaks highly variable.<sup>23</sup> The scattering centers in the fibroglandular tissue case are much closer together and most likely come from small changes in the refractive index from within the extracellular matrix. As a result, the LCI signal for fibroglandular tissue was smoother and more continuous.

Data was collected from a total of 260 samples from 58 patients. Of those, 34 were not analyzed due to the absence of a fiducial ink mark in the histopathologic slide, and 54 were excluded owing to the presence of heterogeneous tissue at the LCI measurement site. The set of 158 histopathology correlated LCI data sets included 71 adipose and 87 fibroglandular cases. The fibroglandular data set included 71 benign fibrous parenchyma, 13 adenocarcinoma, and 3 DCIS cases. The data sets were randomly separated into a training set ( $n=72$ ; 37 adipose, 35 fibroglandular) and a validation set ( $n=86$ ; 34 adipose, 52 fibroglandular). There were 7 (5 adenocarcinoma, 2 DCIS) and 9 tumor (8 adenocarcinoma, 1 DCIS) cases included in the fibroglandular group for the training and validation sets, respectively. The additional samples were used for intravariability testing ( $N=14$ ).

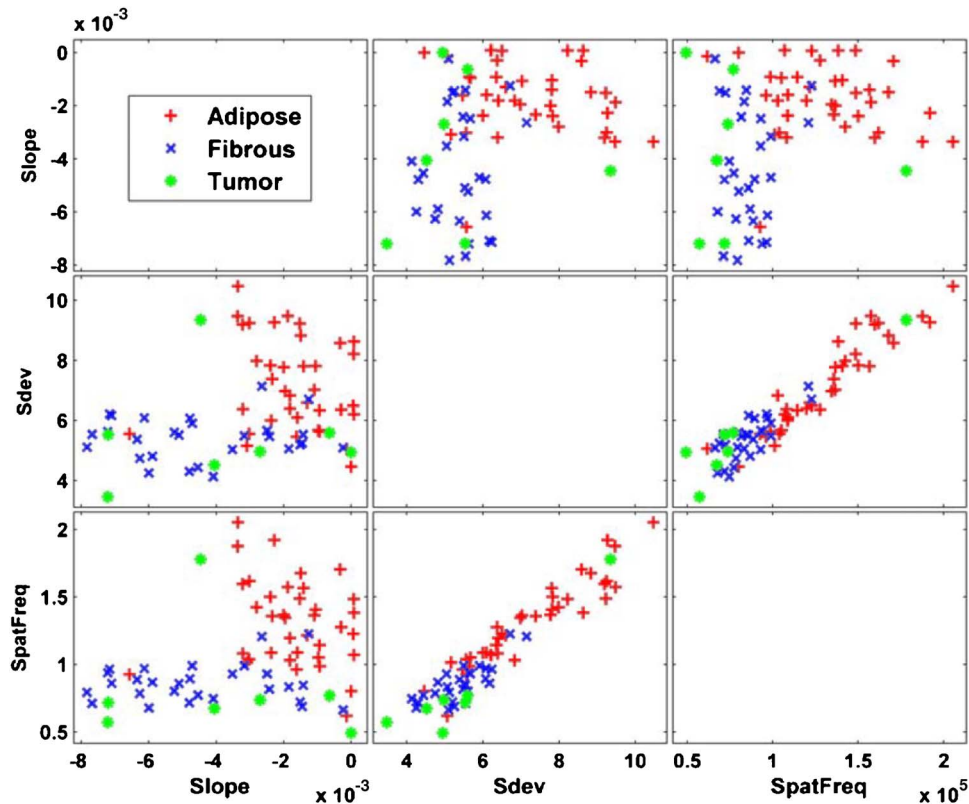
### 3.1 Training Set

The results from the training set are listed in Table 1. As the table demonstrates, each parameter has a significant  $p$ -value. The average magnitude of the slope parameter was higher for fibroglandular tissue, which indicates a higher scattering coefficient for fibroglandular breast tissue compared with adipose tissue. The mean standard deviation was higher for adipose tissue as a result of the signal variation resulting from refractive index fluctuations. The spatial frequency parameter had more energy for the adipose samples within the integrated window band. There was no spatial frequency region where the fibroglandular tissue had higher energy as was seen at higher spatial frequencies in Zysk et al.<sup>21</sup> This could be due to differences in axial resolution (10  $\mu\text{m}$  versus 2  $\mu\text{m}$ ) because the Fourier transform resolution highly depends on spatial sampling frequency.

Another way to represent the training data is through the use of a scatter matrix as shown in Fig. 5. The scatter matrix plots two-dimensional scatter plots between each set of parameters and can be used to observe correlations between classification parameters. It can be seen that there is little correlation between the slope and standard deviation as well as the slope and spatial frequency parameters. In addition, the slope–standard deviation and slope–spatial frequency scatter plots show that the adipose and fibroglandular data sets fall into separate regions, making classification based on these parameters possible. The scatter plot matrix also shows that the standard deviation and spatial frequency parameters are highly correlated. This is expected as both parameters are related to the scattering strength and scatterer distribution. The correlation is higher for the adipose ( $R=0.938$ ) than for the fibroglandular ( $R=0.780$ ) tissue samples.

### 3.2 Validation Set

The results from the validation set using all three parameters for classification are listed in Table 2. The classification parameters show the same trends as were seen in the training set data. The classification results are listed in Table 3. During a FNA procedure, the collection of adipose tissue is seen as a nondiagnostic result. Therefore, the correct classification of adipose tissue can be viewed as a true negative (TN), and the correct classification of fibroglandular tissue can be viewed as a true positive (TP). In this way, the sensitivity, as defined by  $TP/(TP+FN)$  ( $FN$ =false negative) is equivalent to the ac-



**Fig. 5** Scatter plot representation of the entire training set data that can be used to visualize relationships between parameters. There is a high degree of correlation between the standard deviation and the spatial frequency content parameters for both the adipose (red plus sign) and fibroglandular (blue cross) tissue types. Tumor tissues are labeled in green circles and were included in the fibroglandular tissue classification. (Color online only.)

curacy of detecting fibroglandular tissue. In addition, the specificity, as defined by  $TN/(TN+FP)$  ( $FP$ =false positive) is equivalent to the accuracy of detecting adipose tissue. The sensitivity and specificity of the validation set were 98.1% (95% CI: 89.7 to 99.9) and 82.4% (95% CI: 65.5 to 93.2), respectively. The overall accuracy was defined as the total number of correctly classified tissue samples regardless of tissue type. With 86 (34 adipose, 52 fibroglandular) samples in the validation set, the overall accuracy was 91.9% (95% CI: 84.0 to 96.7). CIs were calculated using the normal approximation to the binomial distribution.<sup>24</sup> The one misclassified sample from the fibroglandular validation set was an adenocarcinoma case. The other 8 of 9 tumor cases were correctly classified as fibroglandular tissues.

These results use all three classification parameters as previously described. To determine whether or not the three-

parameter model was statistically better than simply using the slope and standard deviation parameters,<sup>12</sup> it was necessary to look at a truth table describing the differences between the two models. The overall classification results using only the slope and standard deviation parameters are shown in Table 3. The sensitivity and specificity were 80.8% (67.5 to 90.4) and 82.4% (65.5 to 93.2), respectively. Using only the two-parameter model, four tumor cases, all adenocarcinoma, were misclassified as adipose tissue. A truth table to quantify the differences between the two models is shown in Table 4. In Table 4, a ++ cell indicates that both the two-parameter and the three-parameter models correctly classified the sample. A +/- cell indicates that the two-parameter model correctly classified a sample, whereas, the three-parameter model misclassified a sample. Similarly, a -/+ cell indicates that the three-parameter model classified the sample correctly when the two-parameter model misclassified the sample, and a -- cell indicates that both models misclassified the samples. The table shows that there are nine cases where the three-parameter model classified a fibroglandular sample correctly when the two-parameter model misclassified the sample, and no cases with the reverse scenario. The associated  $p$ -value is calculated using McNemar's test for correlated proportions and shows that there is a statistically significant difference between the two- and three-parameter models in terms of fibroglandular tissue classification. No statistical difference was observed for the adipose case.

**Table 2** Validation set statistics.

Parameter	Adipose ( $N=34$ )	Fibroglandular ( $N=52$ )	$p$
Slope	$-1.69 \pm 1.25$	$-5.04 \pm 2.21$	$<1.9 \times 10^{-11}$
Std. Dev.	$6.92 \pm 1.35$	$5.51 \pm 1.55$	$<7.0 \times 10^{-5}$
Spat. Freq.	$1.23 \pm 0.31$	$0.82 \pm 0.23$	$<4.6 \times 10^{-9}$

**Table 3** Classification results.

		Predicted Class			
		Two-Parameter Model Slope, Std. Dev.		Three-Parameter Model Slope, Std. Dev., Spat. Freq.	
		Adipose	Fibroglandular	Adipose	Fibroglandular
True Class	Adipose	28	6	29	5
	Fibroglandular	10	42	1	51

Because the standard deviation parameter is calculated from the slope-subtracted LCI signal, errors in the slope calculation could result in an artificially high standard deviation measurement. To test the effect this would have on our classification, we simulated errors in the slope by randomly modifying the slope parameter  $\pm 5$ , 10, and 20% of its nominal value. The standard deviation and spatial frequency content parameters were then recalculated using the modified slope value. The resulting classification was compared to the nominal value result using McNemar's test as previously described. There was no significant difference in the classification results for either the adipose or fibroglandular tissue type. For the maximum error of  $\pm 20\%$ , the sensitivity was 96.2% (95% CI: 86.8 to 99.5) and the specificity was 76.5% (95% CI: 58.8 to 89.3).

### 3.3 Intrasample Variability

Data to test intrasample variability were collected and analyzed from a separate set of 14 samples from 6 patients (6 adipose, 8 fibroglandular). The average intrasample variability was 18.3% (9.5 to 30.4) for adipose and 1.3% (0.03 to 6.8) for the fibroglandular tissue samples. The overall Cohen's  $\kappa$  statistic was 0.821 (0.725 to 0.981). The number of errors for each 10-A-line set was as follows: adipose [1 0 0 0 7 3] and fibroglandular [0 0 0 1 0 0 0]. The one outlier sample within the adipose data set (7/10 error rate) was due to low signal content, which tended to reduce the spatial frequency content parameter and shift the probability toward the fibroglandular tissue type. If the outlier were removed, the adipose intrasample variability rate would become 8.0% (2.2 to 19.2) and the  $\kappa$  value would become 0.918 (0.847 to 0.988).

**Table 4** Model comparison.

		3 Parameter Model			
		Fibroglandular		Adipose	
Two-Parameter Model		+	-	+	-
	+	42	0	26	3
	-	9	1	2	3
		$p < 0.01$		$p < 1$	

## 4 Discussion

We present an automated algorithm for differentiating *ex vivo* adipose tissue and fibroglandular human breast tissue using LCI interferometry that achieves a high sensitivity and specificity. The extracted parameters used for classification are simple and require minimal additional computation time compared with the standard postprocessing of the LCI signals. The goal of this project is to differentiate between nondiagnostic adipose tissue and the fibroglandular tissue more likely to harbor disease. The ability of LCI to differentiate between adipose and fibroglandular tissue indicates that this technology has the potential to be a useful tool in FNA procedures in an attempt to reduce nondiagnostic sampling rates. More importantly, tumor samples are correctly classified as fibroglandular, meaning that they will not be misclassified as adipose tissue, resulting in a missed diagnosis.

There remain a few challenges to taking such a system into a FNA clinic. First, in this work only homogeneous samples were used for analysis and classification. This was done to define a clear set of parameters that represent the true nature of adipose and fibroglandular tissue types. In a clinical setting, heterogeneous samples will be encountered that will decrease the accuracy of the model. Future work will focus on further defining boundaries between tissue types to provide a regional diagnosis that will account for heterogeneity. In addition, some clinical applications may require further differentiation of fibroglandular tissues into normal fibrous and tumor tissue types as well as identification of additional categories such as necrotic tissue. The classification of nondiagnostic adipose tissue samples, however, does not require this distinction, and as such, was outside the scope of this paper. Also, LCI assumes that the signal comes from single-scattering events, but the presence of multiple scattering, especially at larger depths within the sample can lead to decreased resolution, and changes in the signal profile. In particular, the slope parameter, with its connection to the Beer-Lambert law is particularly sensitive to the single-scattering assumption. It may be necessary to define the border between single and multiple scattering to improve the model and include additional tissue types. In addition, because the standard deviation parameter is calculated from the slope-subtracted LCI scan, any error in the slope fit could result in artificially high standard deviation measurements. However, we found that errors up  $\pm 20\%$  did not significantly affect the classification result.

Insertion of the LCI needle probe within the tissue struc-

ture in an *in vivo* setting, as opposed to surface-only measurements as were done in this study, may introduce additional obstacles that could limit the algorithm's effectiveness. Issues such as bleeding, tissue or optical fiber compression, and operator motion artifacts will all come into play. We plan on studying these issues through *in vivo* animal experiments to further define any limitations of the LCI needle probe. We anticipate that higher speed systems will significantly reduce any motion artifacts seen in the current device. Lastly, the ability to collect tissue aspirates directly following a LCI measurement will need to be addressed. Issues such as the collection of sufficient aspirate material as well as the development of a disposable probe are the subject of an ongoing investigation. Additional future work will focus on the development of higher speed systems based on recent advancements in LCI technology.<sup>25,26</sup> These advancements will allow for higher speed imaging, improved SNR, and greater imaging depth.

### Acknowledgments

This research was funded by the Medical Free Electron Laser program (Grant No. FA9550-04-1-0079) and a NIH Ruth L. Kirschstein individual fellowship (Grant No. 1 F31 EB005141-01A2). The authors would like to acknowledge Sven Holder and the Massachusetts General Hospital pathology lab for their generous help in obtaining tissue samples.

### References

1. M. A. S. Frable and W. J. Frable, "Fine-needle aspiration biopsy of salivary-glands," *Laryngoscope* **101**(3), 245–249 (1991).
2. S. R. S. Mandreker, N. S. Nadkarni, R. G. W. Pinto, and S. Menezes, "Role of fine-needle aspiration cytology as the initial modality in the investigation of thyroid lesions," *Acta Cytol.* **39**(5), 898–904 (1995).
3. S. Klein, "Evaluation of palpable breast masses," *Am. Fam. Physician* **71**(9), 1731–1738 (2005).
4. N. S. Michael and S. Ballo, "Can core needle biopsy replace fine-needle aspiration cytology in the diagnosis of palpable breast carcinoma: A comparative study of 124 women," *Cancer* **78**(4), 773–777 (1996).
5. C. M. Antley, E. E. Mooney, and L. J. Layfield, "A comparison of accuracy rates between open biopsy, cutting-needle biopsy, and fine-needle aspiration biopsy of the breast: A 3-year experience," *Breast J.* **4**(1), 3–8 (1998).
6. N. S. Scott Boerner, "Specimen adequacy and false-negative diagnosis rate in fine-needle aspirates of palpable breast masses," *Cancer Genet. Cytogenet* **84**(6), 344–348 (1998).
7. S. M. S. Rima Bakhos, Steven DeJong, Donald L. Gordon, Shailesh U. Pitale, Marrille Herrmann, and Eva M. Wojcik, "Fine-needle aspiration of the thyroid: Rate and causes of cytohistopathologic discordance," *Diagn. Cytopathol* **23**(4), 233–237 (2000).
8. W. H. Hindle and E. C. Chen, "Accuracy of mammographic appearances after breast fine-needle aspiration," *Am. J. Obstet. Gynecol.* **176**(6), 1286–1290 (1997).
9. A. Saxe, E. Phillips, P. Orfanou, and M. Husain, "Role of sample adequacy in fine needle aspiration biopsy of palpable breast lesions," *Am. J. Surg.* **182**(4), 369–371 (2001).
10. B.-M. Ljung, A. Drejet, N. Chiampi, J. Jeffrey, W. H. Goodson III, K. Chew, D. H. Moore II, and T. R. Miller, "Diagnostic accuracy of fine-needle aspiration biopsy is determined by physician training in sampling technique," *Cancer Genet. Cytogenet* **93**(4), 263–268 (2001).
11. E. Azavedo, G. Svane, and G. Auer, "Stereotactic fine-needle biopsy in 2594 mammographically detected non-palpable lesions," *Lancet* **333**(8646), 1033–1036 (1989).
12. N. V. Iftimia, B. E. Bouma, M. B. Pitman, B. Goldberg, J. Bressner, and G. J. Tearney, "A portable, low coherence interferometry based instrument for fine needle aspiration biopsy guidance," *Rev. Sci. Instrum.* **76**(6), 064301 (2005).
13. J. M. Schmitt, A. Knüttel, and R. F. Bonner, "Measurement of optical-properties of biological tissues by low-coherence reflectometry," *Appl. Opt.* **32**(30), 6032–6042 (1993).
14. J. M. Schmitt, A. R. Knüttel, A. H. Gandjbakhche, and R. F. Bonner, "Optical characterization of dense tissues using low-coherence interferometry," in *Holography, Interferometry, and Optical Pattern Recognition in Biomedicine III*, H. Podbielska, Ed., *Proc. SPIE* **1889**, 197–211 (1993).
15. V. V. Tuchin, "Coherent optical techniques for the analysis of tissue structure and dynamics," *J. Biomed. Opt.* **4**(1), 106–124 (1999).
16. X. Li, C. Chudoba, T. Ko, C. Pitris, and J. G. Fujimoto, "Imaging needle for optical coherence tomography," *Opt. Lett.* **25**(20), 1520–1522 (2000).
17. C. Zhu, G. M. Palmer, T. M. Breslin, F. Xu, and N. Ramanujam, "Use of a multiseperation fiber optic probe for the optical diagnosis of breast cancer," *J. Biomed. Opt.* **10**(2), 024032 (2005).
18. M. Johns, C. A. Giller, D. German, and H. Liu, "Determination of reduced scattering coefficient of biological tissue from a needle-like probe," *Opt. Express* **13**(13), 4828–4842 (2005).
19. A. M. Zysk, S. G. Adie, J. J. Armstrong, M. S. Leight, A. Paduch, D. D. Sampson, F. T. Nguyn, and S. A. Boppart, "Needle-based refractive index measurement using low-coherence interferometry," *Opt. Lett.* **32**(4), 385–387 (2007).
20. W. A. Reed, M. F. Yan, and M. J. Schnitzer, "Gradient-index fiber-optics microprobes for minimally invasive *in vivo* low-coherence interferometry," *Opt. Lett.* **27**(20), 1794–1796 (2002).
21. 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).
22. J. M. Schmitt, Z. Guan-Xiong, and J. Miller, "Measurement of blood hematocrit by dual-wavelength near-IR photoplethysmography," *Proc. SPIE* **1641**, 150–161 (1992).
23. M. P. Stern and F. Conrad, "An automated, direct method for measuring adipocyte cell size," *Clin. Chim. Acta* **65**(1), 29–37 (1975).
24. B. Rosner, *Fundamentals of Biostatistics*, 5th ed., Duxbury Press, Boston (2000).
25. S. H. Yun, G. J. Tearney, J. F. de Boer, N. Iftimia, and B. E. Bouma, "High-speed optical frequency-domain imaging," *Opt. Express* **11**(22), 2953 (2003).
26. B. R. White, M. C. Pierce, N. Nassif, B. Cense, B. H. Park, G. J. Tearney, B. E. Bouma, T. Chen, and J. F. de Boer, "In vivo dynamic human retinal blood flow imaging using ultra-high-speed spectral domain optical Doppler tomography," *Opt. Express* **11**(25), 3490–3497 (2003).