Use of a multiseparation fiber optic probe for the optical diagnosis of breast cancer

Changfang Zhu
University of Wisconsin-Madison
Department of Electrical and Computer Engineering
Madison, Wisconsin 53706

Gregory M. Palmer
University of Wisconsin-Madison
Department of Biomedical Engineering
Madison, Wisconsin 53706

Tara M. Breslin
University of Wisconsin School of Medicine
Department of Surgery
Madison, Wisconsin 53706

Fushen Xu
University of Wisconsin School of Medicine
Department of Pathology
Madison, Wisconsin 53706

Nirmala Ramanujam
University of Wisconsin-Madison
Department of Biomedical Engineering
#2144 Engineering Centers Building
1550 Engineering Dr.
Madison, Wisconsin 53706
E-mail: nimmi@engr.wisc.edu

Abstract. We explore the effects of the illumination and collection geometry on optical spectroscopic diagnosis of breast cancer. Fluorescence and diffuse reflectance spectroscopy in the UV-visible spectral range are made with a multiseparation probe at three illumination-collection separations of 735, 980, and 1225 μm, respectively, from 13 malignant and 34 nonmalignant breast tissues. Statistical analysis is carried out on two types of data inputs: (1) the fluorescence and diffuse reflectance spectra recorded at each of the three illumination-collection separations and (2) the integrated fluorescence (at each excitation wavelength) or diffuse reflectance over the entire spectrum at all three illumination-collection separations. The results show that using the integrated fluorescence intensities recorded at a single excitation wavelength at all three illumination-collection separations can discriminate malignant from nonmalignant breast tissues with similar classification accuracy to that using spectral data measured at several excitation wavelengths with a single illumination-collection separation. These findings have significant implications with respect to the design of an optical system for breast cancer diagnosis. Examining the intensity attenuation at a single wavelength rather than spectral intensities at multiple wavelengths can significantly reduce the measurement and data processing time in a clinical setting as well as the cost and complexity of the optical system. *© 2005 Society of Photo-Optical Instrumentation Engineers.* [DOI: 10.1117/1.1897398]

Keywords: breast cancer; fluorescence; reflectance; spectroscopy; probe geometry.

Paper 04047 received Mar. 26, 2004; revised manuscript received Aug. 23, 2004; accepted for publication Nov. 5, 2004; published online Apr. 26, 2005.

1 Introduction

Surgical biopsy and core needle biopsy are commonly used for the diagnosis of breast lesions. Compared to surgical biopsy, core needle biopsy is less invasive, less expensive, faster, and requires a shorter time for recovery. However, its sampling accuracy is limited because only a few small pieces of tissue are sampled from random locations in the suspicious mass. Consequently the needle biopsy procedure has[1] a false-negative rate of 1 to 7% when verified with follow up mammography, and repeat biopsies are required in 9 to 18% of patients.[2,3]

Optical spectroscopy has been increasingly investigated as a tool for breast cancer diagnosis.[4] In particular, several studies have demonstrated that there are significant differences in the UV-visible (UV-VIS) fluorescence and diffuse reflectance spectra of normal, benign, and malignant breast tissues.[5–14] In addition, this technique is fast, quantitative, and nondestructive. There are important benefits to be reaped by incorporating optical spectroscopy as an adjunct diagnostic modality to core needle biopsy. This technology can be deployed through fiber optic probes to quickly and nondestructively identify the tissue type (normal, benign, and malignant) at the needle tip during a breast biopsy procedure. A positive reading from the optical measurement will potentially increase the likelihood that a biopsy is being sampled from a tumor site. If the optical measurement reads negative, then the needle can be repositioned (along the needle track) to a new tissue site. Currently 6 to 24 biopsies are taken during a core needle biopsy procedure. If the optical method can maximize sampling from tissue sites that are most likely to be cancerous, and minimize unnecessary removal of many normal tissues, it could make the breast biopsy procedure more accurate, less traumatic to the patient, and also reduce the number of biopsies that must be processed to obtain a confirmatory diagnosis. Additionally, if optical spectroscopy proves to be an effective way of identifying cancerous sites, it can be incorporated into much smaller needles than the 11 gauge Mammotome needle (for example, the 21 gauge needle used for fine needle aspiration) and thus make this procedure as minimally invasive as possible.

Gupta et al.[12] and Majumder et al.[11] analyzed different spectral data sets collected from the same set of breast tissues *ex vivo* and showed that the fluorescence emission spectra at excitation wavelengths of 340 and 488 nm and excitation spectra at emission wavelengths of 390 and 460 nm exhibit

Address all correspondence to Nirmala Ramanujam, Univ. of Wisconsin/Madison, Department of Biomedical Engineering, Madison, WI 53706. Tel: 608-265-8267; Fax: 608-265-9239; E-mail: nimmi@engr.wisc.edu

1083-3668/2005/$22.00 © 2005 SPIE

differences between normal, benign, and malignant tissues. Using integrated emission intensities at 340-nm excitation in a binary classification scheme, they were able to differentiate malignant from normal and benign tissues with a sensitivity and specificity of 98%. The fluorescence was attributed to the endogenous fluorophores, reduced nicotinamide adenine dinucleotide (NADH) and collagen. Yang et al.[5,7,8] showed that fluorescence emission spectra at 300-nm excitation and excitation spectra at 340-nm emission could be used to discriminate between malignant and normal fibrous tissues, but not between malignant and normal adipose tissues. They found, for example, that the ratio of normalized intensities at 268- and 289-nm emission (of the 300-nm excitation spectra) discriminated between malignant and normal fibrous tissues with a sensitivity of 93% and specificity of 95%. The primary endogenous fluorophore in tissue probed at these excitation and emission wavelengths is tryptophan.

Nonfluorescent absorbers and scatterers in breast tissues also contribute in part to differences observed in the fluorescence spectra of normal, benign, and malignant tissues. Diffuse reflectance provides a measure of tissue absorption and scattering. Several groups have explored the utility of diffuse reflectance spectroscopy between 250 and 800 nm, for breast cancer detection *ex vivo*[6,8,9,13] and *in vivo*.[9,14] Bigio et al.[14] measured the diffuse reflectance spectra through a core biopsy needle and during breast cancer surgery and showed that this technique can differentiate malignant from normal tissues with a sensitivity of 60 to 70% and a specificity of 85 to 95%. This collection of studies shows that diffuse reflectance spectra can be used to differentiate malignant from normal tissues. Changes in diffuse reflectance of malignant tissues are likely due to increased protein and hemoglobin absorption, increased scattering and decreased β-carotene absorption.

In these previous studies, either fluorescence or diffuse reflectance was used for identification of malignancy. In the case of fluorescence, spectra were measured only at one or several excitation wavelengths. Palmer et al.[10] measured the fluorescence spectra (at a total of nine excitation wavelengths in the UV-VIS spectrum) and the UV-VIS diffuse reflectance spectra of freshly excised breast tissues. Using a multivariate statistical algorithm, they found that 4 out of the 10 measured spectra were sufficient to maximize the discrimination between malignant and nonmalignant tissues. These included fluorescence emission spectra at excitation wavelengths of 300, 400, 420, and 460 nm.

The study described in this paper builds on the work by Palmer et al.[10] The primary goal of this study is to explore the effects of the illumination and collection geometry on the optical spectroscopic diagnosis of breast cancer. The illumination and collection geometry is an important component of tissue optical spectroscopy and currently, fiber optic probes are most commonly used for this purpose. In previous studies, fiber optic probes with only one fixed separation between the illumination and collection fibers were employed for optical spectroscopy of the breast.[5–14] These probes provided a fixed optical sensing depth in tissue. In this study, fluorescence spectra at a total of eight excitation wavelengths and UV-VIS diffuse reflectance spectra were measured from breast tissues with a multiseparation fiber optic probe. This probe has three illumination and collection separations, which enable optical spectra to be measured from different depths within the tissue. The fluorescence/reflectance signal attenuates with increasing probing depth in tissue (due to the increased path length of the photons). Thus, the multiseparation fiber optic probe can measure the attenuation characteristics of the fluorescence and reflectance measured from the breast. Two different types of data sets were collected in this study for statistical analysis: (1) the fluorescence and diffuse reflectance spectra recorded at each of the three illumination-collection separations (or synonymously, different optical sensing depths) and (2) the fluorescence (at each excitation wavelength) or diffuse reflectance integrated over the entire spectrum as a function of illumination-collection separation (which reflects the attenuation of light within the tissue). The statistical analysis of both types of data sets showed that the integrated fluorescence intensities at a single excitation wavelength recorded at all three illumination-collection separations discriminates malignant from nonmalignant breast tissues with a similar classification accuracy compared to spectral data measured at several excitation wavelengths with a single illumination-collection separation. These findings have significant implications. Examining the intensity attenuation at a single wavelength rather than spectral intensities at multiple wavelengths (the entire spectrum) could significantly reduce the acquisition time as well as the cost and complexity of an optical system for breast cancer diagnosis. This study, together with our previous study,[10] provides important information on the utility of optical spectroscopy for breast cancer diagnosis during core needle biopsy.

2 Methods

2.1 *Ex Vivo Sample Collection and Handling*

The breast tissue optical spectroscopy study was approved by the Institutional Review Boards at the University of Wisconsin–Madison. Breast tissue samples were obtained from patients undergoing either a lumpectomy, mastectomy, or breast reduction surgery. Fluorescence and diffuse reflectance spectra were measured on the freshly excised breast tissues within 2 h after surgical excision. A previous study carried out by our group[15] showed that *ex vivo* spectra measured within 2 h after excision provide a relatively close approximation of *in vivo* spectra. During the tissue measurement, the fiber optic probe was placed in direct contact with the tissue surface. After each measurement, the exact site on each tissue sample where the probe was placed was inked, and transverse sections were cut and stained for histopathology. Microscopic evaluation was performed on each histological section by a board-certified pathologist (FX) and a diagnosis was established. In addition, the thickness of each sample was measured from the top (where the ink spot was located) to the bottom of the hematoxylin and eosin (H&E) stained section. Based on the histological breakdown, each sample was broadly classified as normal, benign, or malignant. A total of 47 tissue samples, including 13 malignant, 32 normal, and 2 benign specimens, were collected from a total of 18 patients. Table 1 shows the histological breakdown of the 47 samples examined in the breast tissue optical spectroscopy study. For samples obtained from lumpectomies and mastectomies, the average tissue thickness was 5.6 ± 2.0 mm. Sixteen percent of the samples had a thickness of $\sim$4 mm, 8% had a thickness of $\sim$4.5 mm, and the rest had a thickness

Table 1 The histological breakdown of the 47 samples examined in the breast tissue optical spectroscopy study.

Histological Breakdown	Subcategory	Number of Samples	
Malignant tissues	Invasive ductal carcinoma (IDC)	7	13
	Invasive lobular carcinoma (ILC)	4	
	Ductal carcinoma *in situ* (DCIS)	2	
Normal	Fibrous	12	32
	Adipose	20	
Benign		2	2
Total			47

greater than 5 mm. All samples obtained from breast reduction surgeries had a thickness of at least a centimeter.

2.2 *Optical Spectrometer and Fiber Optic Probe*

Fluorescence and diffuse reflectance spectra were measured using a fiber optic probe coupled to a multiwavelength optical spectrometer. Figure 1(a) shows the schematic of the optical spectrometer and Fig. 1(b) shows common end of the fiber optic probe that comes in contact with the tissue (the black circles correspond to the illumination fibers, the gray circles correspond to the collection fibers which form the three collection rings, and the white circles correspond to dead fibers for bundle packing). The spectrometer [Fig. 1(a)] consists of a 450-W xenon lamp (FL-1039, J.Y. Horiba), a scanning double-excitation monochromator (Gemini 180, J.Y. Horiba), a bifurcated fiber optic probe (multimode fibers), a filter wheel, an imaging spectrograph (Triax 320, J.Y. Horiba), and a CCD camera (CCD3000, J.Y. Horiba). The common end of the fiber optic probe [Fig. 1(b)], which has an outer diameter of 3 mm, consists of an 1180-μm-diam illumination core and three concentric collection rings surrounding the core. The illumination core is made up of 19 fibers, each of which has a core/cladding diameter of 200/245 μm and a numerical aperture (NA) of 0.22. Each collection ring has 12 live fibers, each with a core/cladding diameter of 200/245 μm and an NA of 0.22. The remaining fibers are dead fibers of the same size for bundle packing. The illumination diameter maximizes the coupling efficiency from the light source, and the SNR for the fluorescence measurements. The output signals from the three concentric rings of collection fibers are spatially separated on the CCD chip, thereby enabling fluorescence and diffuse reflectance spectra to be measured at three illumination-collection separations simultaneously. The illumination-collection separation is defined as the distance from the center of illumination core to the center of collection fiber in each ring. The three illumination-collection separations in this fiber-optic probe are 735, 980, and 1225 μm.

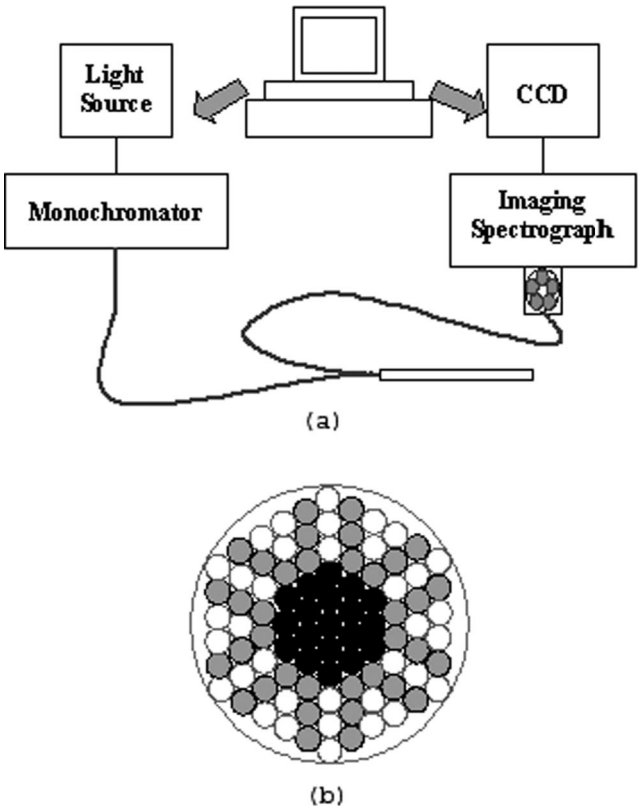

Fig. 1 (a) Schematic of the optical spectrometer and (b) common end of the fiber optic probe that comes in contact with the tissue (the black circles correspond to the illumination fibers, the gray circles correspond to the collections fibers that form the three collection rings, and the white circles correspond to dead fibers for bundle packing).

Next, Monte Carlo simulations[16] of diffuse reflectance were carried out to verify that the average thickness of tissue samples collected from patients undergoing lumpectomies and mastectomies provides a semi-infinite geometry for diffuse reflectance spectroscopy measurements with the multiseparation probe. Simulations were carried out on a homogeneous tissue model with absorption coefficients of 1 to 10 cm^{-1}, scattering coefficients of 108 to 176 cm^{-1}, and an anisotropy factor of 0.926. The range of optical properties used is within that representative of breast tissues in the UV-VIS spectrum.[17] The simulated medium was cylindrically shaped with a thickness of 4 mm, which represented the minimum thickness of the tissue samples. In each simulation, the photons were launched at random, uniformly distributed locations over a range of angles defined by an NA of 0.22 and over a circular

Table 2 The 80% probing depth and the percent fluorescence detected with the three illumination-collection pairs of the multiseparation probe from a homogeneous fluorescent medium, with (a) variable absorption coefficients (μ_a) and a fixed scattering coefficient (μ_s), and (b) variable scattering coefficients and a fixed absorption coefficient. The anisotropy factor (g) of the medium was set to 0.926.

(a)

Optical Properties								
λ_{exc} $\mu_s = 110.4\ cm^{-1}$		λ_{emm} $\mu_s = 104.3\ cm^{-1}$	80% Probing Depth (μm)			Percent Fluorescence Detected ($\times 10^{-4}$)		
$\mu_a\ (cm^{-1})$	QY	$\mu_a\ (cm^{-1})$	Inner	Middle	Outer	Inner	Middle	Outer
1.312	0.762	0.994	1200	1400	1500	14.7	10.1	7.62
2.025	0.494	1.596	1100	1300	1400	17.7	11.6	8.46
5.286	0.189	4.352	900	1000	1100	21.6	11.7	7.16
10.316	0.097	8.580	700	800	900	18.5	7.92	4.00
14.427	0.069	12.012	650	700	800	15.4	5.28	2.30
19.191	0.052	16.043	550	600	700	11.9	3.54	1.26
21.236	0.047	17.754	550	550	550	10.6	2.72	1.03
31.812	0.031	26.517	400	500	500	6.38	1.23	0.292

(b)

Optical Properties								
λ_{exc} $\mu_a = 10.82\ cm^{-1}$		λ_{emm} $\mu_a = 9.0\ cm^{-1}$	80% Probing Depth (μm)			Percent Fluorescence Detected ($\times 10^{-4}$)		
$\mu_s\ (cm^{-1})$	QY	$\mu_s\ (cm^{-1})$	Inner	Middle	Outer	Inner	Middle	Outer
50.0	1	47.2	800	1000	1050	12.3	4.60	2.34
75.0	1	70.8	800	900	1000	15.7	5.90	2.80
100.0	1	94.4	700	750	900	17.0	6.72	3.50
125.0	1	118.0	650	700	800	19.7	8.06	4.00
150.0	1	141.6	600	650	800	21.4	8.70	4.14
175.0	1	165.3	600	600	700	22.2	9.66	4.62
200.0	1	188.9	500	550	600	24.4	9.94	5.02
225.0	1	212.5	500	550	550	25.2	10.3	4.90

Note: QY is quantum yield, λ_{exc} is the excitation wavelength, and λ_{emm} is the emission wavelength.

illumination area defined by the fiber bundle diameter. The diffuse reflectance escaping the medium was collected over a circular area defined by the fiber diameter and over a range of exit angles defined by an NA of 0.22. The refractive index above the medium was set to 1.452 to simulate an optical fiber and that below the medium was set to 1.0. The refractive index of the medium was set to 1.37. The simulation results showed that the light transmittance through the 4-mm-thick sample was less than 5% for the range of optical properties used. Specifically, the light transmittance through the sample was 0.2% when the absorption and scattering coefficient were 10 and 176 cm^{-1}, respectively, and 4.8% when the absorption and scattering coefficients were 1 and 108 cm^{-1}, respectively. These results indicate that the average tissue thickness reported in this paper provides a semi-infinite geometry for diffuse reflectance spectroscopy over the UV-VIS range. The fluorescence emission spectra measured in this study occur within the same wavelength range.

To assess the probing depth and attenuation of the fluorescence measured with the multiseparation probe, Monte Carlo simulations[16] of fluorescence were carried out on a homogeneous tissue model with a wide range of optical properties in the UV-VIS spectrum (see Table 2), an anisotropy factor of 0.926 at both the excitation and emission wavelengths, and a

quantum yield of unity. The simulated medium was cylindrically shaped with a thickness of 5 mm. Note that in real tissue, the absorbers and fluorophores have distinct absorption coefficients and the fluorescence efficiency is solely related to the product of the absorption coefficient and quantum yield of the fluorophore. However, in the Monte Carlo simulations, these two absorption coefficients are combined into a single absorption coefficient and the fluorescence efficiency is a function of the overall absorption coefficient (contributed by both absorbers and fluorophores). Thus, in our Monte Carlo simulations, when the absorption coefficient was varied at the excitation wavelength, the quantum yield was appropriately scaled to account for the mismatch in fluorescence efficiency between real tissue and the simulated medium.[16]

Table 2 shows the 80% probing depth and the percent fluorescence detected with the three illumination-collection pairs of the multiseparation probe from a homogeneous fluorescent medium, with (a) variable absorption coefficients (μ_a) and a fixed scattering coefficient (μ_s), and (b) variable scattering coefficients and a fixed absorption coefficient. The 80% probing depth is defined as the depth above which 80% of the detected fluorescent photons originate. The percent fluorescence detected, which is a measure of the collection efficiency, is the number of fluorescent photons collected with a single fiber in each collection ring, divided by the total number of incident photons (5 million). In Table 2, the 80% probing depth varies from 400 to 1500 μm. For a medium with fixed optical properties, the probing depth increases, while the percent fluorescence detected decreases with increasing illumination-collection separation. For a fixed illumination-collection separation, the probing depth decreases with increasing absorption or scattering, while the percent fluorescence detected decreases with increasing absorption and increases with increasing scattering, respectively.

In summary, for a given set of optical properties, the probing depth increases only modestly with increasing illumination-collection separation, while the percentage of fluorescence detected decreases significantly with increasing illumination-collection separation, particularly for higher absorption and scattering coefficients. Thus, the multiseparation probe geometry can be used to measure the attenuation characteristics of the fluorescence from breast tissues. It is expected that this probe geometry will measure similar trends in diffuse reflectance of the breast. One of the objectives in this paper is to explore whether the fluorescence/reflectance attenuation with increasing probing depth can be utilized for discriminating malignant from nonmalignant breast tissues.

2.3 *Optical Spectroscopy of Breast Tissues*

Fluorescence emission spectra were recorded at eight excitation wavelengths ranging from 300 to 440 nm, in 20-nm increments. The slit width of the excitation monochromator (Gemini 180) was set to provide an excitation band pass of 6.2 nm. The output power at the common end of the fiber optic probe was measured using a low power detector (Newport 818-UV, Newport) connected to a handheld optical power meter (Newport 840-C, Newport) and ranged from 10 to 20 μW over the 300- to 440-nm excitation wavelength range. At each excitation wavelength, fluorescence emission spectra were measured over a 260-nm wavelength range, with the first wavelength shifted by 20 nm from the excitation wavelength. The slit width of the imaging spectrograph was set to provide a spectral resolution of 7.9 nm. Each intensity-wavelength point in the emission spectrum was binned over 18 pixels on the CCD chip, resulting in a wavelength increment of 4.7 nm. The integration time for each fluorescence spectrum ranged from 0.12 to 2 s, depending on the excitation wavelength used. The diffuse reflectance spectra were recorded from 350 to 600 nm. The slit widths of the excitation monochromator and imaging spectrograph were chosen to provide bandpasses of 3.5 and 7.9 nm, respectively. Each intensity-wavelength point in the diffuse reflectance was binned over 18 pixels on the CCD chip, resulting in a wavelength increment of 4.7 nm. The integration time for each diffuse reflectance spectrum was 0.05 s. The overall acquisition time was approximately 1 min for a single scan of eight fluorescence spectra and one diffuse reflectance spectrum.

All fluorescence and diffuse reflectance spectra were calibrated in order to correct for the (1) background spectrum, (2) wavelength dependence, and (3) throughput of the system. The background spectrum, which was measured with the probe immersed in distilled water using the same experimental setup for optical spectroscopy of tissue, was first subtracted point-by-point from each spectrum prior to further calibration. The fluorescence emission spectral intensities at each excitation wavelength were normalized to the output excitation power, to account for the wavelength-dependent variation of the excitation light intensity. The fluorescence spectrum was then corrected for the wavelength-dependent response of the collection fibers, imaging spectrograph and CCD camera, by multiplying it point-by-point by the correction factors measured using a National Institute of Standards and Technology (NIST) tungsten calibration lamp. This procedure also corrected for differences in the collection efficiency of each collection ring. Finally, each fluorescence spectrum was divided by the peak fluorescence intensity (excitation-emission wavelength of 460 to 580 nm) measured with the probe placed on the face of a quartz cuvette containing a solution of Rhodamine B (2.14 μM) dissolved in ethylene glycol (115H3423, Sigma Chemical Co., Missouri) using the same integration time, to account for the time-dependent changes in the throughput of the instrument. The diffuse reflectance spectrum was calibrated for the wavelength-dependent response and the throughput of the system by normalizing it to the diffuse reflectance spectrum measured with the common end of the fiber optic probe inserted into an integration sphere (DRA-CA-30I, Labsphere, Inc., New Hampshire).

2.4 *Data Analysis*

Statistical analysis was carried out on two types of data inputs: (1) the fluorescence and diffuse reflectance spectra recorded at each of the three illumination-collection separations and (2) the fluorescence (at each excitation wavelength) or diffuse reflectance intensity integrated over the entire spectrum as a function of illumination-collection separation. A multivariate statistical algorithm similar to the one developed previously by our group[10] was used for input 1. All data analyses were carried out using a self-programmed Matlab code.

2.4.1 *Spectral analysis at each of the three illumination-collection separations*

Briefly, the multivariate statistical algorithm consists of four steps.[10] First, the spectral data are preprocessed to minimize interpatient variations and variations due to probe-tissue contact, by normalizing each spectrum to a specific intensity-wavelength point in the spectra or the integrated spectral intensity. Particularly, the fluorescence spectra were each normalized to their peak intensity and the diffuse reflectance spectra were normalized to their integrated intensity. Second, the spectral data are dimensionally reduced using a multivariate analysis technique, such as principal component analysis.[18] The third step is to perform feature extraction, i.e., identify the principal components (PCs) that show the statistically most significant differences between malignant and nonmalignant breast tissues using a Wilcoxon rank-sum test.[19] The fourth step is to use a support vector machine (SVM) algorithm[20] to classify each sample as malignant or nonmalignant based on their PC scores.

Data obtained in the clinical study can be divided into ''independent'' and ''dependent'' variables. The ''independent'' variables are in the spectral data matrix **X**, where each row corresponds to the preprocessed fluorescence or reflectance spectrum of a tissue sample and each column corresponding to preprocessed spectral intensity at a specific wavelength. The ''dependent'' variable **Y** is a binary variable that represents the histological diagnosis of each sample, with 1 for malignant tissues and 0 for nonmalignant (benign and normal) tissues. Two multivariate statistical analysis techniques, principal component analysis[18] (PCA) and partial least-squares (PLS) regression[21] were employed for spectral data reduction. Both methods project the set of spectra onto a subspace expanded by the principal components to represent the spectral data with a few variables, which account for most of the variance in the original spectral data set.

In PCA, the set of independent variables (**X**) were used for extraction of the PCs, that is, spectra measured from all tissue samples regardless of histological diagnoses were pooled together to extract the PCs. The PCs were calculated such that the first principal component (PC1) accounts for the largest amount of the total variance of the input data. The second PC (PC2) accounts for the second largest amount of the variance while being orthogonal to PC1, and so on.

PLS is a regression method,[21] where both the independent (**X**) and dependent (**Y**) variables are utilized in calculating the PCs. The PLS regression procedure searches for a set of components that performs a simultaneous decomposition of **X** and **Y** with the constraint that these components explain as much of the covariance between **X** and **Y**. This amounts to determining two sets of weights **w** and **c** to create a linear combination of the columns of **X** and **Y** such that their covariance is maximal. Specifically, a pair of vectors $\mathbf{t}=\mathbf{X}\mathbf{w}$ and $\mathbf{u}=\mathbf{Y}\mathbf{c}$ are obtained with the constraint that $\mathbf{t}^T\mathbf{u}$ is maximal (where the superscript T denotes the matrix transpose). When **t** has converged, it is used to compute the principal component, **p** for **X**, where $\mathbf{p}=\mathbf{X}^T\mathbf{t}$ (again the superscript T denotes the matrix transpose). When the first PC (PC1) is found, the spectral data explained by this PC is subtracted from both **X** and **Y**, and the procedure is repeated. That is, additional PCs are calculated from residuals, which account for the portion of input data that were not accounted for by the previously extracted PCs.

The use of dependent variables (histological diagnoses) in PLS gives rise to the difference between PCA and PLS. PCA provides PCs that account for most of the spectral variance; however, some of these components may not necessarily be relevant to the diagnoses of the tissue samples. PLS looks for PCs that not only account for a large amount of the variance, but also those that are most relevant to the known histology of the samples.

In this study, PCA and PLS were performed on all fluorescence spectra, one excitation wavelength at a time and separately on the diffuse reflectance spectra to generate a set of PCs for each spectral data set. Rather than retaining all of the PCs, only a subset of PCs was retained for further data processing. In the case of PCA, the PCs that account for 95% of the total spectral variance were retained for further analysis. In the case of PLS, the first five PCs of each spectrum were retained, which account for 95 to 99% variance in the spectral data. In both cases, the original spectrum could be faithfully approximated using the linear combination of the selected subset of PCs.

The scores of the selected PCs extracted from either fluorescence or diffuse reflectance spectra, were then pooled together for a Wilcoxon rank-sum test. The Wilcoxon rank-sum test was used to determine which PCs showed the statistically most significant differences between malignant and nonmalignant breast tissues. The scores of three statistically most significant PCs obtained using either PCA or PLS were retained as inputs for classification. SVMs, including the linear SVM and the nonlinear SVM [e.g., polynomial SVM and a radial basis function (RBF) SVM (Ref. 20)] were employed as classification schemes to specify a particular sample as malignant or nonmalignant. Finally, the unbiased performance of the PCA-SVM and PLS-SVM algorithms was evaluated using a ''leave-one-out'' scheme.[22]

2.4.2 *Integrated spectral intensity analysis*

Integrated spectral intensities were first obtained by integrating the intensities over the entire spectrum for each fluorescence spectrum and the diffuse reflectance spectrum at a specific collection ring. Then intensity ratios were calculated by normalizing the integrated intensity recorded from each collection ring to that from the inner collection ring (shortest illumination-collection separation). Thereby two intensity ratios were obtained for each spectrum, i.e., the intensity ratio of the middle and inner ring (middle-to-inner intensity ratio) and the intensity ratio of the outer and inner ring (outer-to-inner intensity ratio). The middle-to-inner and the outer-to-inner intensity ratios of each spectrum (two variables) were then input to the SVM classifier for discriminating between malignant and nonmalignant tissues. An unbiased evaluation on the algorithm's performance was obtained using the leave-one-out scheme.[22]

3 Results

Figure 2 shows the average excitation emission matrices (EEMs) of malignant ($n=13$), normal/benign fibrous ($n=14$), and normal adipose ($n=20$) tissues. Figures 2(a) to 2(c) correspond to the average EEMs of malignant tissues measured with the inner [Fig. 2(a)], middle [Fig. 2(b)], and

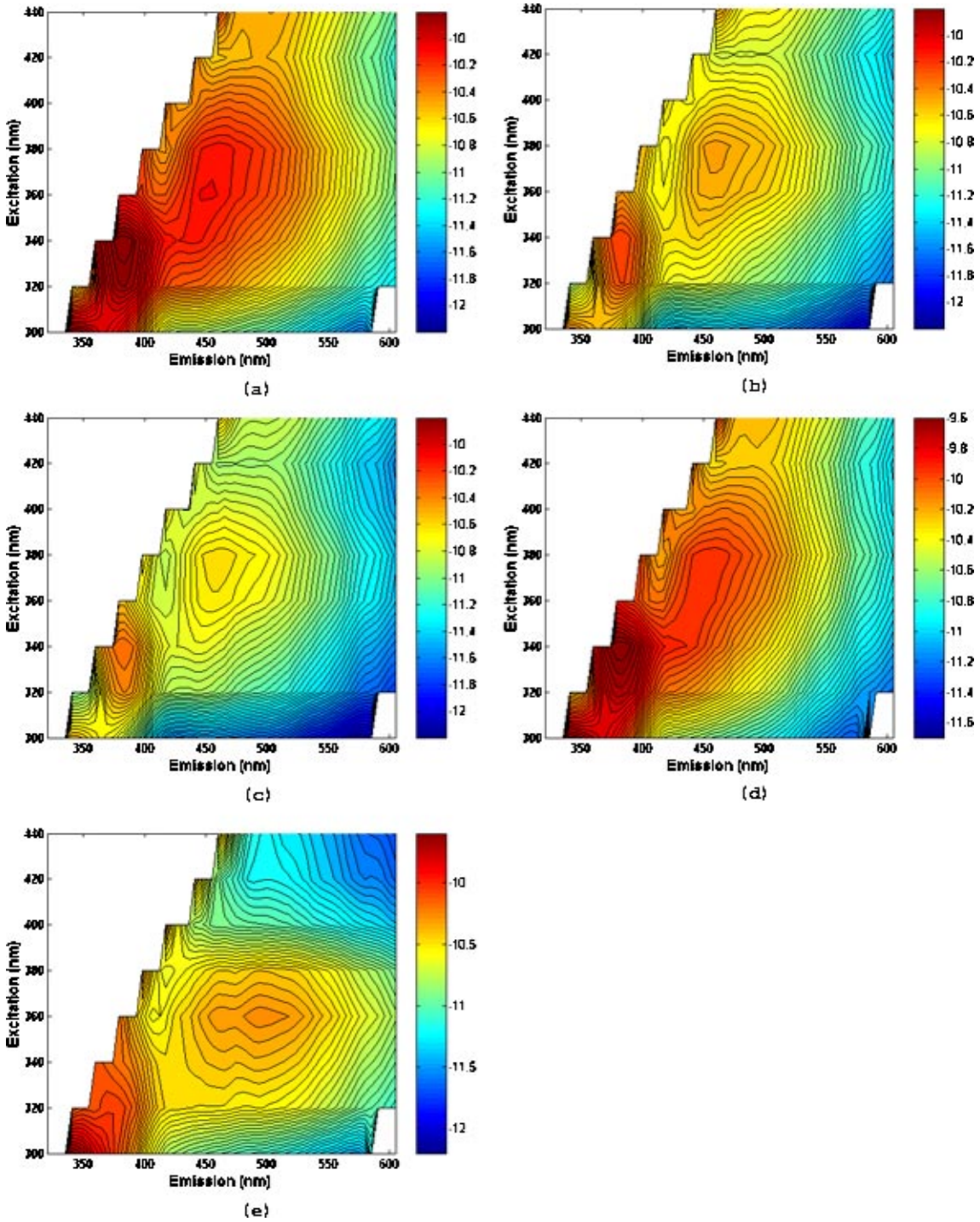

Fig. 2 Average EEMs of malignant ($n=13$), normal/benign fibrous ($n=14$), and normal adipose ($n=20$) tissues, where (a) to (c) correspond to the average EEMs of malignant tissues measured with the (a) inner, (b) middle, and (c) outer collection rings, respectively, and (d) and (e) correspond respectively to average EEMs of normal/benign fibrous and normal adipose tissues measured with the inner collection ring. All figures are plotted on a log scale.

outer [Fig. 2(c)] collection rings, respectively, and Figs. 2(d) and 2(e) correspond, respectively, to average EEMs of normal/benign fibrous and normal adipose tissues measured with the inner collection ring. All figures are plotted on a log scale. Average EEMs of malignant tissues [Figs. 2(a) to 2(c)] show that four peaks are visible and appear at similar locations in all three EEMs, i.e., at excitation-emission wavelength pairs of 300−340, 340−390, 360/380−460, and

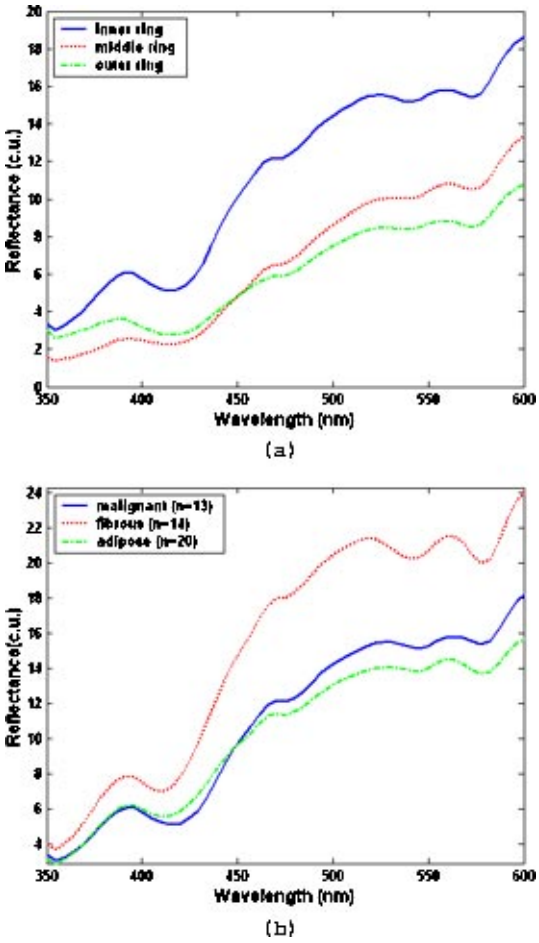

Fig. 3 (a) Average diffuse reflectance spectra of (a) malignant breast tissues ($n=13$) measured with the inner, middle, and outer rings and (b) malignant ($n=13$), normal/benign fibrous ($n=14$), and normal adipose tissues ($n=20$) measured with the inner ring.

440−520 nm. As expected, the fluorescence intensity at each peak location decreases from the inner ring to the outer ring. Such a decrease in fluorescence intensity is also observed in the EEMs of normal/benign fibrous and normal adipose tissues measured with three collection rings (not shown here). The average EEM of normal/benign fibrous tissues measured with the inner ring [Fig. 2(d)] has the same four peaks as the malignant tissue EEM measured with the inner ring [Fig. 2(a)]. However, the average adipose tissue EEM measured with the inner ring [Fig. 2(e)] displays distinct differences, with a weak presence of the 340−390-nm peak and a red shift in the peak at 360−460 nm to approximately 360−500 nm.

Figure 3(a) displays the average diffuse reflectance spectra of malignant breast tissues ($n=13$) measured with the inner, middle, and outer rings, and Fig. 3(b) shows the average diffuse reflectance spectra of malignant ($n=13$), normal/benign fibrous ($n=14$), and normal adipose tissues ($n=20$) measured with the inner ring. In each spectrum, five valleys are notable at 350, 420, 470, 540, and 570 nm. The valleys at 420, 540, and 570 nm correspond to hemoglobin absorption, the one at 350 nm corresponds to NADH, and the one at 470 nm can be assigned to β-carotene. In Fig. 3(a), the diffuse reflectance intensity measured from malignant breast tissues decreases from the inner collection ring to the outer collection ring, particularly in the wavelength range above 450 nm. Such a decrease is also observed in diffuse reflectance spectra of normal/benign fibrous and normal adipose tissues measured with the three collection rings (not shown here). A comparison of spectral intensities for the three tissue types, measured with the inner ring [Fig. 3(b)] shows that fibrous tissues have the highest average diffuse reflectance intensity, while adipose tissues have the lowest average diffuse reflectance intensity.

PCA and PLS were carried out on the fluorescence and diffuse reflectance spectra measured with each of the three collection rings to obtain a set of PCs that faithfully represent the spectral data. Table 3 shows the three statistically most significant PCs for each collection ring, identified by a Wilcoxon rank-sum test from the two sets of PCs obtained from the PCA and PLS analyses. All the PCs here display statistically significant differences between malignant and nonmalignant tissues below a significance level of $p<0.005$. The three most significant PCs obtained from PLS analysis were extracted from fluorescence spectra at 320- and 420-nm excitations and from the diffuse reflectance spectra. The three most significant PCs obtained from PCA analysis were extracted from fluorescence spectra at 320-, 420-, 300-, and 340-nm excitation and from the diffuse reflectance spectra.

Table 4 shows the overall classification rate, sensitivity, and specificity achieved with the linear SVM for discriminating between malignant and nonmalignant breast tissues using the three statistically most significant PCs (obtained from PCA and PLS analysis) for each collection ring. Note that in several cases, particularly in the case where PLS analysis was used, the three most significant PCs included those obtained from both diffuse reflectance and fluorescence spectra. The classification accuracy achieved with PCA-SVM or PLS-SVM did not differ significantly between the three collection rings. However, for spectral data acquired by a given collec-

Table 3 The three statistically most significant PCs for each collection ring, identified by a Wilcoxon rank-sum test from the two sets of PCs obtained from the PCA and PLS analyses.

PCA	→Increasing Significance			PLS	→Increasing Significance		
Inner	420 nm, PC1	320 nm, PC1	reflectance, PC4	Inner	420 nm, PC1	reflectance, PC1	320 nm, PC1
Middle	320 nm, PC1	420 nm, PC1	340 nm, PC1	Middle	420 nm, PC1	reflectance, PC1	320 nm, PC1
Outer	320 nm, PC1	300 nm, PC3	420 nm, PC1	Outer	420 nm, PC1	reflectance, PC1	320 nm, PC1

All the PCs here display statistically significant differences between malignant and nonmalignant tissues below a significance level of $p<0.005$.

Table 4 The overall classification rate, sensitivity and specificity achieved with linear SVM for discriminating between malignant and nonmalignant breast tissues using the three statistically most significant PCs (obtained from PCA and PLS analysis) for each collection ring.

	PCA+Linear SVM			PLS+Linear SVM		
	Inner	Middle	Outer	Inner	Middle	Outer
Classification rate (%)	78.72	65.96	57.45	82.98	82.98	80.85
Sensitivity (%)	38.46	30.77	15.38	61.54	69.23	69.23
Specificity (%)	94.12	79.41	73.53	91.18	88.24	85.29

tion ring, the PLS-SVM algorithm achieved consistently higher classification accuracy than the PCA-SVM algorithm. It was also found that the overall classification rate, sensitivity, and specificity obtained using nonlinear SVM classification (polynomial SVM and RBF SVM) did not differ significantly from that obtained using linear SVM classification (not shown here).

Table 5 shows the overall classification rate, sensitivity, and specificity achieved with linear SVM for discriminating between malignant and nonmalignant breast tissues using the three statistically most significant PCs extracted from PLS analysis of the diffuse reflectance spectra measured with each collection ring. The three PCs show statistically significant differences between malignant and nonmalignant tissues below a significance level of $p<0.05$. Only the results from the PLS-SVM algorithm are shown, since this approach yields superior results compared to the PCA-SVM algorithm. The overall classification rate did not differ significantly between the different collection rings. In addition, the specificity was higher than the sensitivity for all three collection rings. A comparison of the classification accuracy achieved using (1) combined fluorescence and reflectance spectra (Table 4) versus (2) reflectance spectra only (Table 5) indicates that the overall classification rates do not differ significantly for these two cases. However, the sensitivity is consistently higher for the former, while the specificity is consistently higher for the latter.

Figure 4 shows the average integrated spectral intensity

Table 5 The overall classification rate, sensitivity, and specificity achieved with linear SVM for discriminating between malignant and nonmalignant breast tissues using the three statistically most significant PCs extracted from PLS analysis of the diffuse reflectance spectra measured with each collection ring.

	PLS+Linear SVM		
	Inner	Middle	Outer
Classification rate (%)	80.85	78.72	85.11
Sensitivity (%)	53.85	30.77	46.15
Specificity (%)	91.18	97.06	100

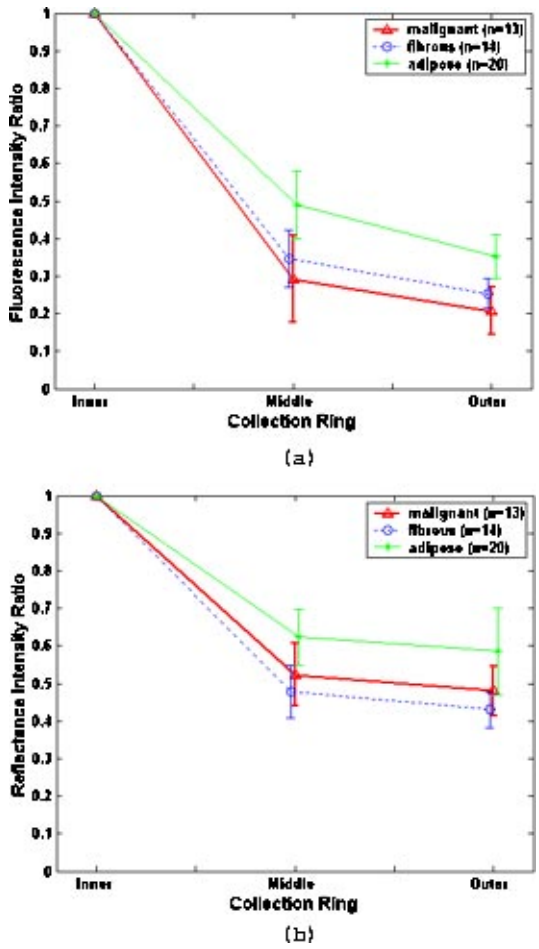

Fig. 4 Average integrated spectral intensity ratios and standard deviations for (a) fluorescence spectra at 300-nm excitation, and (b) diffuse reflectance spectra of malignant ($n=13$), normal/benign fibrous ($n=14$), and normal adipose tissues ($n=20$).

ratios and standard deviations for fluorescence spectra at 300-nm excitation [Fig. 4(a)] and diffuse reflectance spectra [Fig. 4(b)] of malignant ($n=13$), normal/benign fibrous ($n=14$), and normal adipose tissues ($n=20$). For fluorescence at 300-nm excitation, the average values of middle-to-inner ratio and outer-to-inner ratio were 0.29 ± 0.12 and 0.21 ± 0.06, respectively, for malignant tissues; 0.35 ± 0.08 and 0.25 ± 0.04 for normal/benign fibrous tissues; and 0.49 ± 0.09 and 0.35 ± 0.06 for normal adipose tissues. For the diffuse reflectance, the average values of middle-to-inner ratio and outer-to-inner ratio were 0.52 ± 0.08 and 0.48 ± 0.07, respectively, for malignant tissues; 0.48 ± 0.07 and 0.43 ± 0.05 for normal/benign fibrous tissues; and 0.62 ± 0.08 and 0.59 ± 0.12 for normal adipose tissues. As expected, the spectral intensity ratios of both fluorescence and diffuse reflectance decrease with increasing illumination-collection separation. Furthermore the decrease (attenuation) in spectral intensity is greater from inner ring to middle ring than from middle ring to outer ring. Also, the attenuation in spectral intensity is different for different tissue types. In particular, normal adipose tissues undergo less attenuation relative to normal/benign fibrous and malignant tissues. Spectral intensity ratios at the other seven excitation wavelengths are not shown here, but display a similar pattern

Table 6 The overall classification rate, sensitivity, and specificity achieved with linear SVM for discriminating between malignant and nonmalignant tissues using the middle-to-inner and outer-to-inner intensity ratios for fluorescence collected at different excitation wavelengths and for the diffuse reflectance.

	300 nm	320 nm	340 nm	360 nm	380 nm	400 nm	420 nm	440 nm	Refl.
Classification rate (%)	82.98	78.72	78.72	78.72	68.09	68.09	65.96	61.70	72.34
Sensitivity (%)	69.23	61.54	53.85	61.54	23.08	23.08	0	0	0
Specificity (%)	88.24	85.29	88.24	85.29	85.29	85.29	91.18	85.29	100

to that observed at an excitation wavelength of 300 nm. Wilcoxon rank-sum tests showed that the middle-to-inner and outer-to-inner intensity ratios of fluorescence at 300-nm excitation displayed statistically significant differences between malignant and nonmalignant tissues below a significance level of $p<0.005$. The same test indicated that the middle-to-inner and outer-to-inner intensity ratios of diffuse reflectance did not display statistically significant differences between malignant and nonmalignant tissues ($p<0.5$).

Table 6 shows the overall classification rate, sensitivity, and specificity achieved with linear SVM for discriminating between malignant and nonmalignant tissues using the middle-to-inner and outer-to-inner intensity ratios for fluorescence collected at different excitation wavelengths and for the diffuse reflectance. The results show that classification using fluorescence spectral intensity ratios at 300-nm excitation provides the highest classification accuracy, which is comparable to that obtained using the entire spectral input for a given collection ring (Table 4). Using fluorescence at 320- to 360-nm excitation provides slightly lower classification accuracy, which is still comparable to that obtained using the entire spectral input for a given collection ring. When spectral intensity ratios of either fluorescence at excitation wavelengths of 380 to 440 nm, or diffuse reflectance were used, the classification yields high specificity, but very low sensitivity.

Figure 5 shows the scatter plots of the middle-to-inner and outer-to-inner ratios for (a) fluorescence spectra at 300-nm excitation [Fig. 5(a)] and (b) diffuse reflectance spectra [Fig. 5(b)] of malignant and nonmalignant breast tissues and the corresponding hyperplanes obtained from linear SVM. In the case of fluorescence at 300-nm excitation [Fig. 5(a)], the clusters of the spectral intensity ratios for the two tissue types are separable. However, in the case of diffuse reflectance [Fig. 5(b)], the clusters of spectral intensity ratios for the two tissue types are nonseparable. Similar observations were made from scatter plots of spectral intensity ratios for fluorescence spectra at excitation wavelengths of 380 to 440 nm (not shown here).

4 Discussion

Optical spectroscopy in the UV-VIS spectral range shows distinct differences between malignant and nonmalignant breast tissues, and thus this technique has the potential to be used as a diagnostic tool for breast cancer. Statistical analysis of the spectra indicated that the fluorescence spectra at excitation wavelengths of 300, 320, 340, and 420 nm and the diffuse reflectance spectrum showed the statistically most significant differences ($p<0.005$) between malignant and nonmalignant breast tissues. A previous study by our group[10] showed that fluorescence spectra at excitation wavelengths of 300, 320, 400, 420, 440, and 460 nm showed the statistically most significant differences between malignant and nonmalignant breast tissues ($p<0.005$). The excitation wavelengths identified in this study are a subset of the wavelengths identified previously, except for the excitation wavelength of 340 nm. Another difference is that in the previous study, the diffuse reflectance spectrum was not found to show statistically significant differences between malignant and nonmalignant breast tissues.

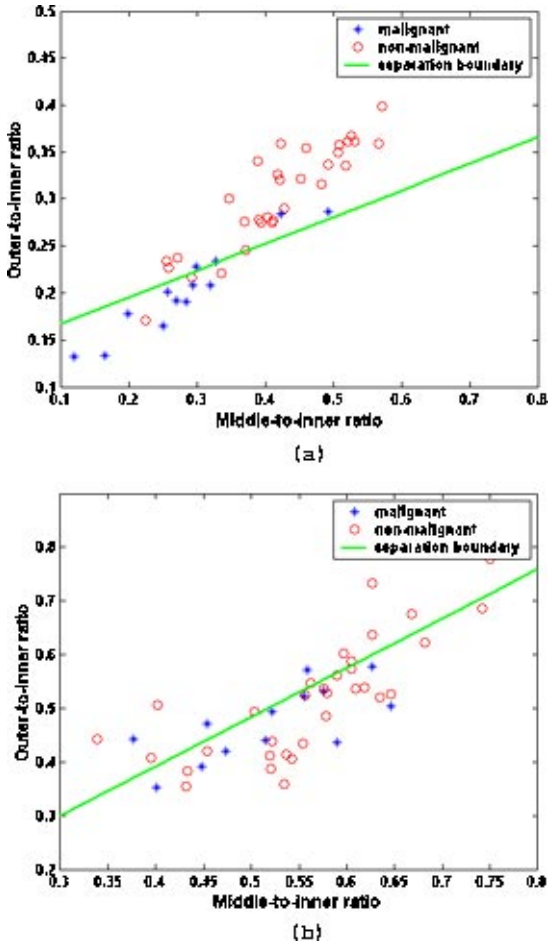

Fig. 5 Scatter plots of middle-to-inner and outer-to-inner intensity ratios for (a) fluorescence spectra at 300-nm excitation and (b) diffuse reflectance spectra of malignant and nonmalignant breast tissues and the corresponding hyperplanes obtained from the linear SVM.

Two statistical methods, PCA and PLS, were used in this study for spectral data reduction. One of the differences between PCA and PLS is that in PCA, the components are chosen so that maximal data variance is explained, while PLS looks for not only components that describe as much of the data variance as possible, but those that are also most relevant to the known group association of each sample. A linear SVM algorithm based on PCs obtained from PLS outperformed that based on PCs obtained from PCA for spectra collected at all three illumination-collection separations (Table 4). To determine how different the PCs obtained from PCA and PLS are for a given spectrum measured at a particular illumination-collection separation, those extracted from fluorescence spectra at 320- and 420-nm excitation and from the diffuse reflectance spectra measured with the inner collection ring of the fiber optic probe (see first row of Table 3) were compared. Although the line shapes of the PLS and PCA components extracted from the fluorescence and diffuse reflectance spectra were found to be generally similar, the PLS components had additional structural features that were not present in the PCA components. However, it is difficult to directly assign these features to particular fluorophores and chromophores in the tissue. This is due to the fact that these linear models do not accurately describe the nonlinear relationship between fluorescence, absorption, and scattering in turbid media such as tissue. Moreover, a component can be multiplied by -1 without affecting the decomposition model. Thus, peaks in the PCA/PLS components could correspond to peaks or valleys in a given spectrum, making it difficult to assign the spectral features to particular fluorophores or chromophores in the tissue.

Typically, the optical spectral features of different tissue types are not completely separable, and a nonlinear SVM is expected to be superior to a linear SVM classification in dealing with such nonseparable cases. However, using the same data reduction technique, PCA or PLS, the classification rates, sensitivities, and specificities provided by linear and nonlinear SVMs did not differ significantly. On the other hand, linear SVM classification associated with PLS data reduction (PLS-SVM) provided higher classification accuracy than linear SVM associated with PCA data reduction (PCA-SVM) (Table 4). This suggests that the improvement in the classification accuracy is more heavily dependent on the data reduction technique. An alternative approach to improving the classification accuracy could be to employ physical models to extract tissue optical and fluorescence properties that are diagnostically useful.[23] This will be explored in future studies.

In this study, fluorescence and diffuse reflectance spectra were measured from breast tissues with a multiseparation fiber optic probe. This probe has three illumination-collection separations, which enable optical spectra to be measured from different depths within the tissue. The classification accuracy achieved with PCA-SVM or PLS-SVM did not differ significantly between the three illumination-collection separations. This is likely due to the facts that (1) the probing depth is expected to increase only modestly with increasing illumination-collection separation (based on the Monte Carlo simulation results in Table 2) and (2) there is no apparent structural change in breast tissue with increasing depth that could be captured with different illumination-collection separations.

The fluorescence/reflectance signal attenuates with increasing probing depth in tissue (due to the increased path length of the photons) and the multiseparation fiber optic probe can be used to measure the attenuation characteristics of these signals in the breast. The results from the data analysis show that using the integrated fluorescence intensities recorded at a single excitation wavelength at all three illumination-collection separations (a measure of attenuation) can discriminate malignant from nonmalignant breast tissues with similar classification accuracy to that using spectral data measured at several excitation wavelengths at a single illumination-collection separation. The use of integrated fluorescence intensities at a single excitation wavelength is expected to (1) significantly reduce the measurement time and (2) require much less data processing. This finding has significant implications in clinical applications, where both speed and low cost are desirable. Although each spectrum takes only a few seconds to measure, hardware changes between measurements increases the time for a single scan of eight fluorescence spectra and one diffuse reflectance spectrum to a minute. If the system can be streamlined to include a subset of the wavelengths evaluated in this study (such as the integrated intensity at a single excitation wavelength), the measurement and data processing time can be significantly reduced to enable near-real-time implementation of this technology during a clinical procedure.

Note that the number of nonmalignant samples ($n=34$) is much greater than that of malignant samples ($n=13$) in this study. The principal idea of an SVM is to determine an optimal hyperplane that maximizes the margin between two classes, and equalize the distances of the misclassified samples in each class from the separation boundary. For a poorly separable case [as in Fig. 5(b)] and when the numbers of elements in the two classes are unbalanced, a large number of samples from each class become support vectors (i.e., marginal samples), and it is likely that most samples from the class that has much fewer elements (in this case, the malignant tissues) will be misclassified. This would result in a high specificity, but very low sensitivity. However, for the case where sample clusters are separable [Fig. 5(a)], it is less likely that the sensitivity and specificity are sensitive to the number of samples in each class. This also likely explains the high specificity and low sensitivity achieved when diffuse reflectance spectra alone (Table 5) rather than the combination of fluorescence and diffuse reflectance spectra (Table 4) were used for discriminating between malignant and nonmalignant breast tissues. The three PCs extracted from diffuse reflectance spectra showed statistically smaller differences between malignant and nonmalignant tissues ($p<0.05$) compared to the three PCs extracted from fluorescence and diffuse reflectance spectra ($p<0.005$). To verify this explanation, data analyses (using spectral intensity ratios only) were carried out on a subset of spectral data, which included balanced malignant and nonmalignant samples. In this data set, only half of the normal/benign fibrous samples (7) and half of normal adipose samples (10) were retained, and the number of malignant samples versus nonmalignant samples was 13:17. The classification results indicated that for the separable case (i.e., fluorescence at an excitation wavelength of 300 nm), the sensitivity and specificity achieved with the balanced data set (61.54 and 82.35%, respectively) are comparable to that achieved

with the unbalanced data set (69.23 and 88.24%, respectively). However, for the nonseparable case, there was a significant disparity in the sensitivities and specificities achieved with the balanced and unbalanced data sets (53.85 and 68.75% for the balanced data set versus 0 and 100% for unbalanced data set). In summary, when sample clusters are separable, the sensitivity and specificity do not appear to be highly sensitive to the number of samples in each class. However, when samples are poorly separable, it is likely that the specificity overwhelms the sensitivity when there are too many nonmalignant versus malignant samples. This suggests that in the poorly separable case, a balanced data set for each class would be desirable to avoid a bias toward the classification of one tissue type.

The statistical data analyses described in this paper demonstrates that malignant and nonmalignant breast tissues can be discriminated using either a subset of fluorescence and diffuse reflectance spectra measured with a single illumination-collection separation, or integrated fluorescence intensities measured at a single excitation wavelength with three illumination-collection separations. This suggests encouraging opportunities for simplifying the optical measurements and data processing procedure, to enable near-real-time implementation of the optical technology during a clinical procedure. However, in the current study, a relatively small data set of 47 samples was analyzed, of which only 13 were malignant. Moreover, the ''leave-one-out'' cross-validation scheme, which was used to provide an unbiased estimate of the algorithm's performance, is based on ''resampling,'' and thus does not provide an independent testing data set for prospectively evaluating the algorithm. In spite of the limitations in the sample size and the testing scheme, similar classification accuracies were obtained when different preprocessing techniques (spectra versus integrated spectral intensity), feature extraction methods and classifiers (linear and nonlinear SVMs) were used. These results suggest that the difference in the spectral features between malignant and nonmalignant tissues can be identified consistently using a variety of different algorithms. Note, however, that a significantly larger data set (that can be divided into independent training and testing data sets) and a training set with a balanced number of malignant and nonmalignant samples will be required in future studies to fully evaluate and validate the performance of the classification algorithms described here.

The study presented here and our previous study[10] provide the foundation for the primary focus of future work, which is to explore the utility of optical spectroscopy for breast cancer detection during core needle biopsy. The current probe geometry will be modified into a side-firing fiber optic probe and implemented for use in a vacuum-assisted core biopsy needle. If this optical technique proves to be diagnostically useful, it can potentially improve the diagnostic efficacy of breast needle biopsy and lead to fewer biopsies and follow-up procedures in patients suspected to have breast cancer. Additionally, the fiber optic probe can be made thin enough to fit through an even smaller needle than the standard 1/4-in. size, making an emotionally draining procedure less physically traumatic.

Acknowledgments

The authors would like to acknowledge the Whitaker Foundation and the National Institutes of Health (NIH) grant No. 1 R01 CA10059-01A1 for supporting this work.

References

1. R. J. Jackman, K. W. Nowels, J. Rodriguez-Soto, F. A. Marzoni, Jr., S. I. Finkelstein, and M. J. Shepard, ''Stereotactic, automated, large-core needle biopsy of nonpalpable breast lesions: false-negative and histologic underestimation rates after long-term follow-up,'' *Radiology* **210**, 799–805 (1999).
2. D. D. Dershaw, E. A. Morris, L. Liberman, and A. F. Abramson, ''Nondiagnostic stereotaxic core breast biopsy: results of rebiopsy,'' *Radiology* **198**, 323–325 (1996).
3. J. E. Meyer, D. N. Smith, S. C. Lester, P. J. DiPiro, C. M. Denison, S. C. Harvey, R. L. Christian, A. Richardson, and W. D. Ko, ''Large-needle core biopsy: nonmalignant breast abnormalities evaluated with surgical excision or repeat core biopsy,'' *Radiology* **206**, 717–720 (1998).
4. G. M. Palmer and N. Ramanujam, ''Diagnosis of breast cancer using optical spectroscopy,'' *Med. Laser Appl.* **18**, 233–248 (2003).
5. Y. Yang, E. J. Celmer, M. Zurawska Szczepaniak, and R. R. Alfano, ''Excitation spectrum of malignant and benign breast tissues: a potential optical biopsy approach,'' *Lasers Life Sci.* **7**, 249–265 (1997).
6. Y. Yang, E. J. Celmer, J. A. Koutcher, and R. R. Alfano, ''UV reflectance spectroscopy probes DNA and protein changes in human breast tissues,'' *J. Clin. Laser Med. Surg.* **19**, 35–39 (2001).
7. Y. Yang, A. Katz, E. J. Celmer, M. Zurawska-Szczepaniak, and R. R. Alfano, ''Fundamental differences of excitation spectrum between malignant and benign breast tissues,'' *Photochem. Photobiol.* **66**, 518–522 (1997).
8. Y. Yang, A. Katz, E. J. Celmer, M. Zurawska Szczepaniak, and R. R. Alfano, ''Optical spectroscopy of benign and malignant breast tissues,'' *Lasers Life Sci.* **7**, 115–127 (1996).
9. S. V. Pushkarev, S. A. Naumov, S. M. Vovk, V. A. Volovodenko, and V. V. Udut, ''Application of laser fluorescence spectroscopy and diffuse reflection spectroscopy in diagnosing the states of mammary gland tissue,'' *Optoelectron. Instrum. Data Process.* **2**, 71–76 (1999).
10. G. M. Palmer, C. Zhu, T. M. Breslin, F. Xu, K. W. Gilchrist, and N. Ramanujam, ''Comparison of multiexcitation fluorescence and diffuse reflectance spectroscopy for the diagnosis of breast cancer (March 2003),'' *IEEE Trans. Biomed. Eng.* **50**, 1233–1242 (2003).
11. S. K. Majumder, P. K. Gupta, B. Jain, and A. Uppal, ''UV excited autofluorescence spectroscopy of human breast tissues for discriminating cancerous tissue from benign tumor and normal tissue,'' *Lasers Life Sci.* **8**, 249–264 (1999).
12. P. K. Gupta, S. K. Majumder, and A. Uppal, ''Breast cancer diagnosis using N2 laser excited autofluorescence spectroscopy,'' *Lasers Surg. Med.* **21**, 417–422 (1997).
13. N. Ghosh, S. K. Mohanty, S. K. Majumder, and P. K. Gupta, ''Measurement of optical transport properties of normal and malignant human breast tissue,'' *Appl. Opt.* **40**, 176–184 (2001).
14. I. J. Bigio, S. G. Bown, G. Briggs, C. Kelley, S. Lakhani, D. Pickard, P. M. Ripley, I. G. Rose, and C. Saunders, ''Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results,'' *J. Biomed. Opt.* **5**, 221–228 (2000).
15. G. M. Palmer, C. L. Marshek, K. M. Vrotsos, and N. Ramanujam, ''Optimal methods for fluorescence and diffuse reflectance measurements of tissue biopsy samples,'' *Lasers Surg. Med.* **30**, 191–200 (2002).
16. Q. Liu, C. Zhu, and N. Ramanujam, ''Experimental validation of Monte Carlo modeling of fluorescence in tissues in the UV-visible spectrum,'' *J. Biomed. Opt.* **8**, 223–236 (2003).
17. V. G. Peters, D. R. Wyman, M. S. Patterson, and G. L. Frank, ''Optical properties of normal and diseased human breast tissues in the visible and near infrared,'' *Phys. Med. Biol.* **35**, 1317–1334 (1990).
18. R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, pp. xiii, 594, Prentice-Hall, Englewood Cliffs, NJ (1982).
19. R. M. Bethea, B. S. Duran, and T. L. Boullion, *Statistical Methods for Engineers and Scientists*, pp. xxi, 652, Marcel Dekker, New York (1995).

20. N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines: And Other Kernel-Based Learning Methods*, Cambridge University Press, Cambridge, New York (2000).
21. H. Martens, *Multivariate Calibration*, John Wiley & Sons, New York (1989).
22. J. S. U. Hjorth, *Computer Intensive Statistical Methods: Validation, Model Selection, and Bootstrap*, Chapman & Hall, London, New York (1994).
23. M. G. Muller, I. Georgakoudi, Q. Zhang, J. Wu, and M. S. Feld, ''Intrinsic fluorescence spectroscopy in turbid media: disentangling effects of scattering and absorption,'' *Appl. Opt.* **40**, 4633–4646 (2001).