Wavelet-based characterization of spectral fluctuations in normal, benign, and cancerous human breast tissues

Sharad Gupta
Maya S. Nair
Asima Pradhan
Nrusingh C. Biswal
Nidhi Agarwal
Indian Institute of Technology
Kanpur 208016
India

Asha Agarwal
Ganesh Shanker Vidhyarthi Memorial Medical College
Kanpur 208016
India

Prasanta K. Panigrahi
Physical Research Laboratory
Navarangpura, Ahmedabad 380009
India

Abstract. Fluorescence intensity fluctuations in the visible wavelength regime in normal, benign, and cancerous human breast tissue samples are studied through wavelet transform. The analyses have been carried out in unpolarized, parallel and perpendicularly polarized channels, for optimal tissue characterization. It has been observed that polarized fluorescence data, particularly the perpendicular components, differentiate various tissue types quite well. Wavelet transform, because of its ability for multiresolution analysis, provides the ideal tool to separate and characterize fluctuations in the fluorescence spectra at different scales. We quantify these differences and find that the fluctuations in the perpendicular channel of the cancerous tissues are more randomized as compared to their normal counterparts. Furthermore, for cancerous tissues, the same is very well described by the normal distribution, which is not the case for normal and benign samples. It has also been observed that, up to a certain point, fluctuations at larger scales are more sensitive to tissue types. The differences in the average, low-pass wavelet coefficients of normal, cancerous, pericanalicular, and intracanalicular benign tissues are also pointed out.
[DOI: 10.1117/1.2062404]

Keywords: fluorescence spectroscopy; polarization; lasers in medicine; wavelet transforms; biomedical optics.

Paper 04236RR received Dec. 3, 2004; revised manuscript received Apr. 5, 2005; accepted for publication Apr. 29, 2005; published online Oct. 4, 2005.

1 Introduction

Although the beneficial properties of light have been known for a long time and it has been used for therapeutic as well as diagnostics purposes, the advent of lasers in 1960s has made possible the efficient use of light for surgery and diagnosis.[1–3] Laser-based endoscopic techniques and imaging methods can be used for real-time diagnosis of diseases without biopsy.[4] In the field of biomedical applications, fluorescence spectroscopy has been used fruitfully, both for diagnosis and therapeutic purposes. In the early 1980s, Alfano and coworkers[5] first introduced laser spectroscopy for tooth decay detection and cancer diagnosis.[6] This technique was then used by Kittrel et al. for diagnosis of atherosclerotic plaque,[7] which was extended further by Deckelbaum et al.[8] Morphological and biochemical changes, due to disease, cause the fluorophores inside the tissue to fluoresce differently, as compared to their normal counterparts. Due to its sensitivity to minute variations, fluorescence spectroscopy can provide quantitative biochemical information about the state of the tissue, which may not be obtained using standard pathology. Over other light-based investigation methods, fluorescence spectroscopy is often preferred because of its high sensitivity, high speed, and safety. Over the past 15 years its diagnostic potential has been tested in different organs of the body, including the mouth, breast, esophagus, and bladder, etc.[9–17]

Diagnosis of human breast cancer through fluorescence studies has been an active area of research for quite some time. Although more than 80% of breast lumps are not cancerous, biopsy is the only way to diagnose. If diagnosed early, breast cancer is also one of the most treatable forms of cancer. In fluorescence spectroscopy, different fluorophores having excitation frequencies in the UV and visible regimes have been used as markers. To be more specific, several fluorescing compounds, such as flavins, nucleotides (NADH), tryptophan, tyrosine, elastin, collagen, etc.[18] are present in tissues. Flavins are intrinsic fluorophores emitting in the visible region. Flavin adenine dinucleotide (FAD), flavin mononucleotide (FMN), and free riboflavin are the three forms of riboflavin encountered in biological materials. These flavins may be attached to proteins or may be free. These are coenzymes that play a role in the oxidation-reduction process of the tissues.[5] Hence, these may play a major role in triggering tumor growth.[19] Our investigation is primarily concerned with the flavins and other fluorophores such as porphyrin having emission peaks in the visible region. In choosing these fluorophores one has the advantage of working in the visible wavelength region and so

Address all correspondence to Asima Pradhan, Indian Institute of Physics, Dept. of Physics, Kanpur 208016, India. Tel.: 91-512-2597691; Fax: 91-512-2590914; E-mail: asima@iitk.ac.in

1083-3668/2005/10(5)/054012/9/$22.00 © 2005 SPIE

can avoid the potentially harmful effect of ultraviolet radiation.

Tissue being a turbid medium, having different morphological and biochemical compositions for normal and tumors, makes the modeling of the same quite difficult. A number of physical[16,20–22] and statistical[23,24] methods have been employed for the purpose of identifying reliable distinguishing features between normal, benign, and cancerous tissues. The prominent statistical methods include principal component analysis (PCA),[24] singular value decomposition,[25] neural networks,[26] and other methods based on pattern recognition.[27] It is often difficult to transparently relate the distinguishing features emerging from the models with the spectral data. Furthermore, in most of the fluorescence-based studies performed so far, the characteristic tissue features are extracted from the smoothened data. For example, in PCA, one keeps a few dominant principal components for comparison, neglecting the fluctuations. The fact that the fluorescence spectrum is noisy, being affected by experimental and statistical uncertainties, also makes extraction of characteristic spectral fluctuations a difficult job.

Recently the method of wavelet transform has been employed to study intensity fluctuations in the polarized fluorescence spectra of human breast tissues.[28] A preliminary analysis indicated that fluctuations at the finest scale are much more randomized for the cancerous tissues. The widths of the Gaussian distributions characterizing fluctuations in cancerous, normal, and benign tissues also showed significant differences, indicating the possibility of using intensity fluctuations for characterization of different tissue types. From the fact that cancerous tissues possess nuclei that have irregular shapes and are much more densely packed as compared to the normal tissues, it is quite plausible to expect more randomization in the same. As has been mentioned earlier, a number of well-defined features, originating from the smoothened spectral data, have been proposed for diagnostic purposes. It is physically conceivable that, in the initial phase of the tumor growth, these features are not expected to be quite prominent. However, the differences in spectral fluctuations can manifest much earlier as compared to the above-mentioned gross features. Hence, it is imperative to study much more carefully the spectral fluctuations at different scales in cancerous, normal, and benign tissues and compare them exhaustively for differences and similarities.

The goal of the present article is to study the characteristic properties of intensity fluctuations at different scales, in the unpolarized, parallel, and perpendicularly polarized components of the fluorescence data. For this purpose, we make use of wavelet transform to isolate the variations at different scales in different tissue types. Because of its multiresolution and localization properties, this linear transform is ideally suited for disentangling variations at different scales. The impact of statistical and experimental uncertainties on the parameters derived in the wavelet domain can be easily estimated. The fact that wavelet transform relies on progressive averaging in extracting the wavelet coefficients reduces the above errors. The cancerous, benign, and normal breast tissue samples studied here have been irradiated with linearly polarized light (488 nm) as well as with unpolarized light of an argon ion laser. The unpolarized and polarized fluorescence spectra ($\parallel$ and $\perp$) were recorded and subjected to wavelet analysis. As is known by now, cancerous tissues are characterized by morphological changes such as irregular shape and size of the nuclei, random arrangement, and overcrowding, as compared to normal tissues; these can lead to altered fluorescence intensities and randomization.[29] Such features may be exploited through wavelet transform to detect differences between normal and tumor tissues. The concentration of fluorophores can also be different in the tumors giving rise to differences in fluorescence spectra of cancerous, benign, and normal breast tissues.

In the following section, we briefly describe the experimental apparatus used in the present study. Section 3 contains an abridged description of the essential properties of wavelet transform, which is the mathematical tool employed in this paper. We present the results and their analysis in Sec. 4 through a number of figures. The key features distinguishing normal, benign, and cancerous tissues in the wavelet domain are clearly pointed out. It was observed that the normal distributions describe better the intensity fluctuations in the perpendicular component of the polarized spectroscopic data in tumors. The plausible physical origin of these differences is also presented. Finally, we conclude in Sec. 5 by summarizing the essential features brought out in this study and the future directions of work.

2 Materials and Methods

2.1 *Histopathology*

In total, 45 tumor tissue samples with their normal counterparts were supplied by Ganesh Shankar Vidyarthi Memorial Medical College, Kanpur, India, after surgery. These were analyzed in close collaboration with the pathologist of the hospital. The age of patients spanned a broad range, from 16 to 85 years, coming from varied economic backgrounds.

The collected samples were analyzed on the same day, without any chemical treatment. During experiments, the sample was kept moist with isotonic saline and was placed in a quartz cuvette of size $1 \times 1 \times 5$ cm.

It should be pointed out that the human breast has 15 to 20 sections called lobes, with many smaller sections called lobules. Each section is connected by thin tubes called ducts. The most common type of breast cancer, called ductal carcinoma, is the one that affects these tubes; it is found in the cells of the ducts. Cancer that begins in the lobes or lobules is called lobular carcinoma.[30] Most of the breast cancers studied are infiltrating ductal carcinoma. The benign tumors used in this study are mostly fibroadenomas, with both intracanalicular and pericanalicular types.

2.2 *Experimental Setup*

The samples were excited by 488-nm wavelength plane polarized light from an Ar-ion laser (Spectra Physics 165, 5 W). The unpolarized and polarized fluorescence spectra were collected in right angle geometry using triplemate monochromator (SPEX-1877E) and PMT (RCA C-31034). For polarized fluorescence, a depolarizer was used after the analyzer, in order to ensure that there was no preference of the selected directions of polarized fluorescence by the detection system. The components of fluorescence light that are parallel and perpendicular to the incident polarized light were measured in the 500- to 700-nm wavelength region.[31] In this wavelength

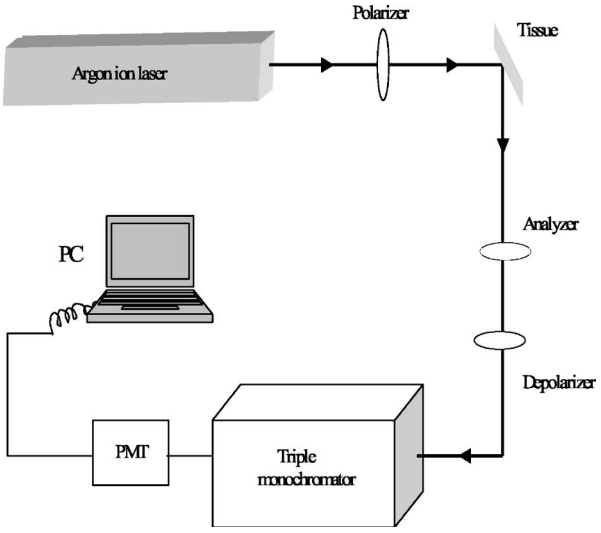

Fig. 1 Schematic of the experimental setup.

region, FAD is the dominant fluorophore in the fluorescence spectra from tissues, which has a fluorescence peak at 530 nm. The main absorber in this region is porphyrin, which absorbs at 540-nm (α band) and 580-nm (β band) wavelengths. Porphyrin also acts as a weak fluorophore, with an emission peak at 630 nm.

The experimental setup used in these measurements is shown in Fig. 1.

3 Discrete Wavelet Transform

In this section, we briefly outline the essential ideas of the wavelet transform, emphasizing those points that will be useful for the purpose of our analysis. In the last decade and a half, the wavelet transform has emerged as a powerful tool to analyze transient and time-varying phenomena. The variations at different scales are systematically separated in the wavelet basis, from the average or the trend of a given signal. The mathematical microscope nature of the wavelet basis, which enables it to perform multiresolution analysis (MRA), arises from the manner in which the basis set is constructed.[32] We will concentrate here on discrete wavelet transform (DWT), where the basis elements have a strictly finite size.

In the construction of the basis set, one starts with the scaling function $\varphi(x)$ (father wavelet) and the mother wavelet $\psi(x)$, whose height and width are arbitrary:

$$\int \varphi \mathrm{d}x = A, \quad \int \psi \mathrm{d}x = 0, \quad \int \varphi\psi \mathrm{d}x = 0,$$

$$\int |\varphi|^2 \mathrm{d}x = 1 = \int |\psi|^2 \mathrm{d}x, \tag{1}$$

where A is an arbitrary constant.

Both of these functions belong to the square integrable class. Two operations crucial to the construction of a complete orthonormal basis are translation and scaling. It can be checked that the following scaled and translated wavelets and scaling functions are square integrable:

$$\psi_{j,k} = 2^{j/2}\psi(2^j t - k),$$

$$\varphi_{j,k} = 2^{j/2}\varphi(2^j t - k). \tag{2}$$

Here, k is the translation parameter, j is the scaling parameter in the dyadic basis, and $2^{j/2}$ is the normalization factor at scale j, which takes integral values starting from zero. The original mother wavelet corresponds to $\psi_{0,0}$, whereas the father wavelet is given by $\varphi_{0,0}$. Higher values of j lead to the so-called daughter wavelets, which are of the similar form as the mother wavelet, except that they are thinner and taller by a factor of $2^{j/2}$. The translation unit $k/2^j$ is also commensurate with the thinner size of the daughter wavelet at scale j. In a given wavelet basis, only one scaling function and its translations are taken, since others are not orthogonal to the wavelets. In the above construction, the translated scaling functions are given by $\varphi_{0,k} \equiv \varphi_k = \varphi(x-k)$. In the limit $j \to \infty$ and for integral values of k, in the range $-\infty \leqslant k \leqslant \infty$, the above basis becomes a complete set.

Hence, any finite energy signal $f(t) \in L^2(R)$ (Ref. 33) can be expanded as

$$f(t) = \sum_{k=-\infty}^{\infty} c_k \varphi_k(t) + \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} d_{j,k}\psi_{j,k}(t), \tag{3}$$

where c_k and $d_{j,k}$ are the wavelet transforms of the signal $f(t)$. Although we have started with the scale value $j=0$, in principle one could have started from any finite value of j, in which case,

$$f(t) = \sum_{k=-\infty}^{\infty} c_{j,k}\varphi_{j,k}(t) + \sum_{k=-\infty}^{\infty} \sum_{j' \geqslant j}^{\infty} d_{j',k}\psi_{j',k}(t).$$

Explicitly the wavelet coefficients are given by

$$c_{j,k} = \langle f(t), \varphi_{j,k}(t) \rangle = \int f(t)\varphi_{j,k}(t)\mathrm{d}t,$$

and

$$d_{j,k} = \langle f(t), \psi_{j,k}(t) \rangle = \int f(t)\psi_{j,k}(t)\mathrm{d}t. \tag{4}$$

Since $\varphi_{j,k}$ is a scaling function located at k and having a finite window size commensurate to scale j, $c_{j,k}$ represents the average value of the signal over the same window at location k. As will become clear from later discussions, $d_{j,k}$ represent variations of the signal in the same window size. Father wavelets or scaling functions are used for extracting the low-frequency, smooth component of the signal. On the other hand, the wavelets extract the higher frequency detail component at various scales. Broadly speaking, father wavelets are used for finding the trend components and mother wavelets pick out the deviations. All wavelet basis functions satisfy the dilation equation, also known as the multiresolution analysis (MRA) equation:

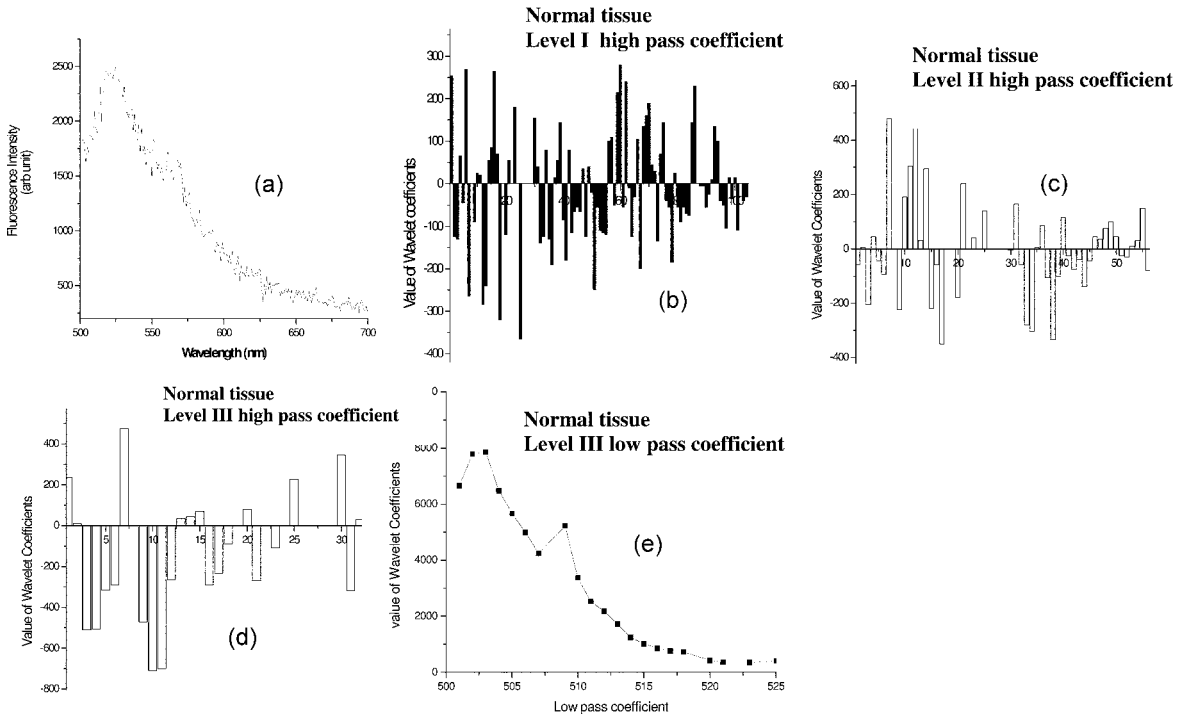

Fig. 2 (a) Fluorescence spectrum of a normal tissue, having 200 intensity values, starting from 500 nm. One can carry out three-level decompositions without discarding any point. The high-pass wavelet coefficients displayed in (b) level I, (c) level II, and (d) level III have, respectively, 100, 50, and 25 coefficients. (e) Level-III low-pass coefficients numbering 25.

$$\varphi(t) = \sum_n h(n)\sqrt{2}\varphi(2t-n),$$

and

$$\psi(t) = \sum_n \tilde{h}(n)\sqrt{2}\varphi(2t-n). \tag{5}$$

Physically, this means that scaling function and wavelet at a given scale can be constructed from the linear superposition of scaling function alone, at a higher scale. The initial scale in a given basis set is arbitrary, to be chosen keeping the application in mind. Using MRA, one can show that

$$c_{j,k} = \sum_n h(n-2k)c_{j+1,n}$$

and

$$d_{j,k} = \sum_n \tilde{h}(n-2k)c_{j+1,n}. \tag{6}$$

The coefficients $c_{i,k}$ and $d_{i,k}$ are, respectively, the low-pass and high-pass coefficients at scale j. In a given wavelet decomposition, one first fixes a high scale J and finds the coefficients, which are denoted as level-I low-pass and high-pass coefficients. Level-II coefficients correspond to the scale value J-1, level-III to J-2, and so on. It is worth noting that both scaling function coefficients (low-pass coefficients) and wavelet coefficients (high-pass coefficients) at a given scale j can be obtained from only low-pass coefficients at a higher scale. As the scale value j increases for fixed k, the scaling function becomes thinner and taller representing approximately a Dirac delta function. The corresponding low-pass coefficient is then nothing but the sample of the signal at the location k. Hence, starting from the samples of the signal at the finest resolution, one can obtain all the other scaling function and wavelet coefficients through the MRA equation. One only needs to know the filter coefficients $h(n)$ and $\tilde{h}(n)$, without explicitly knowing the forms of the wavelet functions. In this sense, wavelet transform is significantly different from Fourier transform.

There are infinite varieties of discrete wavelets; the choice of a basis set depends on the application at hand. In recent times DWT has found applications in diverse areas such as astronomy, acoustics, nuclear engineering, sub-band coding, signal and image processing, neurophysiology, magnetic resonance imaging, turbulence, earthquake-prediction, etc.[32–35] For our application, we make use of the simplest Haar wavelet, since the interpretation of the wavelet coefficients is quite transparent here; it is also free from artifacts arising due to the finite size of the data.

For the Haar wavelet

$$h(0) = h(1) = \frac{1}{\sqrt{2}} \text{ and } \tilde{h}(0) = -\tilde{h}(1) = \frac{1}{\sqrt{2}}, \tag{7}$$

these coefficients are different for different wavelet basis sets. Haar basis is special, since it is the only wavelet that is symmetric and compactly supported. In a level-I Haar wavelet decomposition, the nearest neighbor averages and differences are calculated with the normalization factor of $1/\sqrt{2}$. The alternate coefficients are thrown out in a process called downsampling or decimation, which leaves half of the data in the

form of low-pass coefficients and the other half in terms of level-I high-pass coefficients. Subsequently, the same procedure can be applied once more to the low-pass coefficients to decompose them into level-II high-pass coefficients and level-II low-pass coefficients. One can now clearly see that the level-I high-pass coefficients are the nearest neighbor differences in the Haar wavelet. The level-II high-pass coefficients are the differences of the nearest neighbor averages. In this manner, one can find out the differences of progressively larger chunks of data, which are the other higher level high-pass coefficients. The fluctuations considered here are these Haar wavelet coefficients, which are the simple differences of intensities at neighboring wavelengths. This illustrates how the high-pass coefficients at higher levels separate out the fluctuations over progressively bigger neighborhoods.

If the data set contains 2^N elements, with N being an integer, then one can have an N-level decomposition, after which one is left with only one low-pass coefficient, which is the average of all points, modulo a normalization factor. It is evident that the high-pass coefficients progressively capture differences at a broader scale starting from the nearest neighbor ones. In more sophisticated wavelets, the above averaging and differentiation is replaced by a suitable weighted averaging and differentiation. If the data is not 2^N, one needs to append the same with the required number of points by padding or other procedures. It is worth mentioning that, apart from Haar wavelets, the process of computing wavelet coefficients needs a suitable extension of data, e.g., circular or periodic, which also brings in artifacts to be clearly differentiated from the true variations. In order to avoid these complications, we have taken the Haar wavelets for our analysis. We have taken 192 points of the data starting from $\lambda = 501$ nm, since the same can be decomposed up to six levels, without encountering any artificial features. These data points are treated as the $c_{J+1,k}$ in Eq. (6) to obtain the level-I low- and high-pass coefficients, at different locations k. The level-I low-pass coefficients are then analogously used to obtain the level-II high-pass coefficients and so on.

It should be pointed out that wavelet transform satisfies the Parseval theorem implying that

$$\sum I_i^2 = \sum_k (c_k)^2 + \sum_j \sum_k (d_{jk})^2, \tag{8}$$

where I_i is the intensity at $\lambda = \lambda_i$, and c_k and d_{jk} are the low- and high-pass coefficients, respectively. It should be noted that, in a given wavelet decomposition, one fixes a scale J and takes only one scaling function. The corresponding low-pass coefficients $c_{J,k}$ are conveniently abbreviated by c_k. In case of reference to multilevel low-pass coefficients, e.g., level-I and level-II low-pass coefficients, it should be understood that these refer respectively to $c_{J,k}$ and $c_{J\text{-}1,k}$.

4 Results and Discussion

4.1 *Standard Deviation* (σ)

Figure 2 depicts the four-level wavelet transform of a typical normal tissue spectral data. The low-pass coefficients in Fig. 2 represent the average part and resemble the data itself. The level-I high-pass coefficients have 100 points, level-II high-pass coefficients have 50 points, and level-III high-pass coefficients have 25 data points. We refer the interested readers to Ref. 28 for details about the wavelet transform of cancerous and benign tissues. For quantifying the differences, we now proceed to compute a number of global and local parameters from both high- and low-pass coefficients. It is not difficult to convince oneself that instead of the spectral fluctuations, as captured by the wavelet coefficients, the percentage fluctuations, i.e., the high-pass coefficients divided by their corresponding low-pass coefficients, are much more reliable for a statistical analysis. Hence, we compute the same for various high-pass coefficients and find their mean and standard deviation values. The corresponding histograms are then fitted with normal distributions to bring out the nature of randomization in various tissue types. Preliminary studies have shown earlier that the standard deviation of the perpendicularly polarized data discriminated the various tissue types.[28] For the purpose of comparison, we have plotted (Fig. 3) the ratios of the standard deviation of cancerous and benign tissues with their corresponding normal counterparts, for unpolarized, parallel polarized, and perpendicularly polarized spectral data. It is clearly seen that, for this parameter, perpendicularly polarized light indeed brings out the best differentiation, the unpolarized spectra being the worst performer. This highlights the usefulness of polarization spectroscopy in biomedical diagnostics. We have checked that the contribution due to counting error and dark noise are extremely small as compared to the above differences. The ratio of standard deviation of intracanalicular type of benign tumor tissues to normal tissues show values greater than one. Pericanalicular type benign tumor tissues also display values less than one, similar to the malignant ones. This is expected since pericanalicular types have scattering coefficients similar to malignant tumors and hence randomization may be higher in them. Thus it appears that the standard deviation values can distinguish intracanalicular benign tumors, malignant tumors, and normal from one another in the wavelet domain.

One sees characteristic large fluctuations in the fluorescence spectra of normal tissues. Physically, the presence of fat gives rise to the Raman lines, which lead to the large fluctuations in the wavelet domain. However, some normal tissues, which are more fibrous (low fat content like the cancerous ones), do not display these characteristic Raman peaks but still have significantly large fluctuations. Hence, the differentiation between tissue types would be more robust if one can differentiate the normal and cancerous tissues after removing the large fluctuations originating from the Raman peaks. Nonremoval of the same only improves the result. We find that even after removal of these large fluctuations, normal tissue standard deviations remain larger than their cancerous counterparts.

It is worth pointing out that cancerous tissues have more scatterers of various sizes and hence one expects that the characteristic emission peaks of the fluorescence spectrum will be broadened more in the same because of multiple scattering. Therefore, the fluctuations in the wavelet domain will be less in cancerous tissues. The ratio of standard deviations of malignant to normal tissue being less than one confirms this. We notice an improved sensitivity on moving to higher levels (up to the third level) as depicted in Fig. 4. The sensitivity in the third level is 95% while in the first level it is 90%.

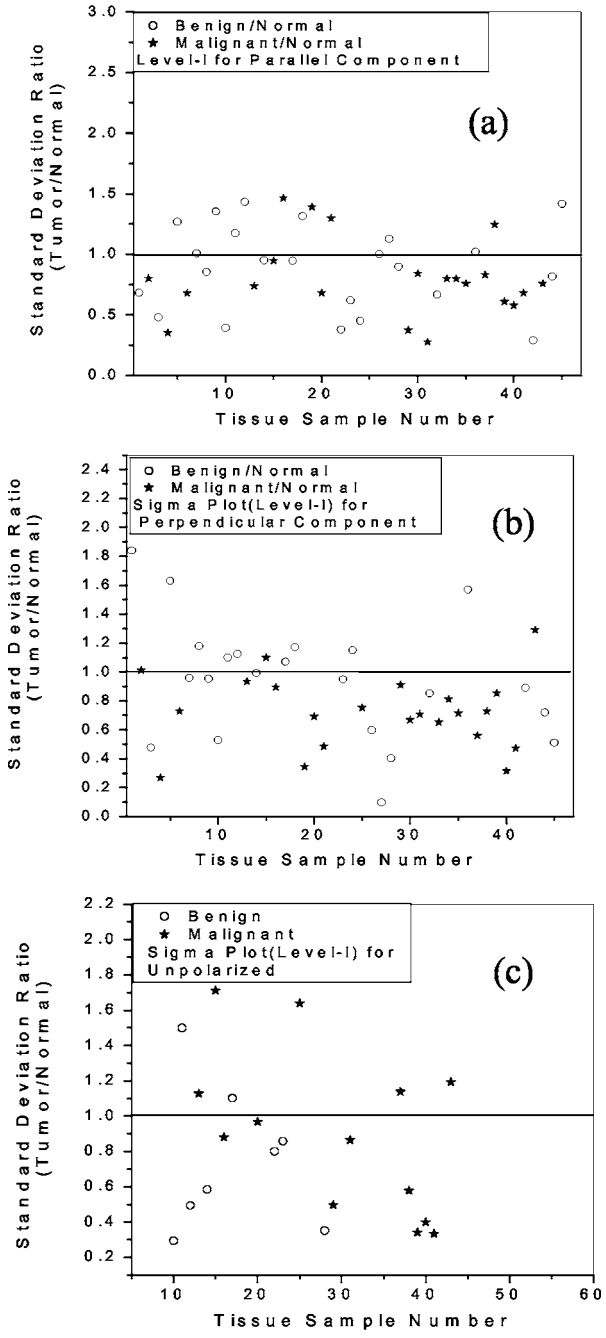

Fig. 3 Standard deviation ratios of the spectral fluctuations (the high-pass coefficients divided by their corresponding low-pass coefficients) of level I of benign (circles) and malignant (stars) tissues, with respect to the corresponding normal ones (a) for parallel, (b) for perpendicular component, and (c) for unpolarized spectra.

One also expects that randomization of fluctuations is more in malignant tumor spectra as compared to the normal ones.[29,36,37] We hence proceed to confirm the same through correlation studies. Now, having seen that the perpendicular component is the best discriminator, we focus only on the parameters of this component.

4.2 *Correlation Value* (R^2)

The normalized level-I fluctuations fit a normal distribution quite accurately for the cancerous tissues, for the perpendicularly polarized light, as seen in Fig. 5. The same is also brought out clearly by the correlation factor (R^2) during fitting and is seen in Fig. 6, where the various R^2 values have been computed. At a physical level, the rapid randomization in the cancerous tissues results from the much faster generation of the perpendicular component in tumors. The intensity variations also reflect this randomization process in the sense that they fit the Gaussian distribution much better.

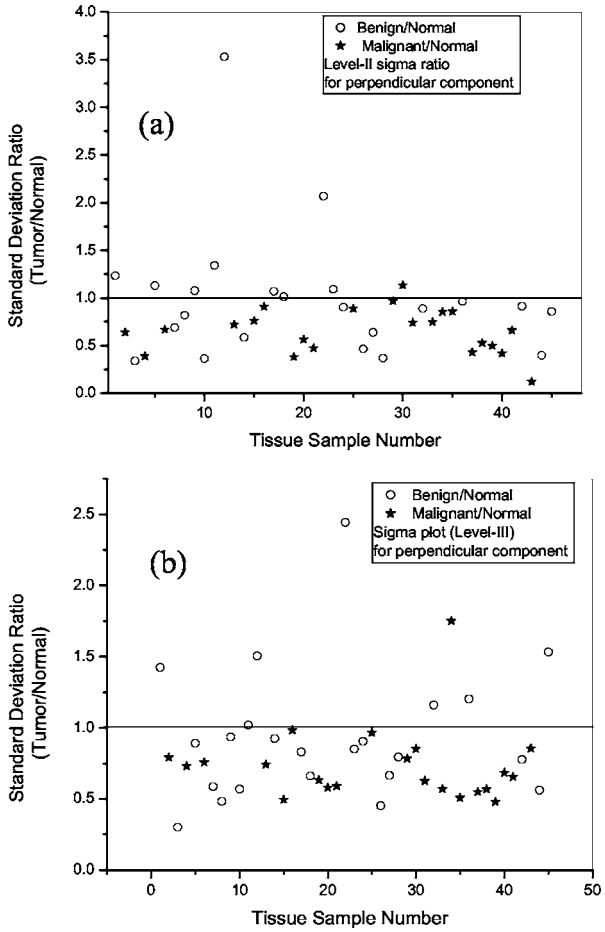

Fig. 4 Standard deviation ratios of the spectral fluctuations of (a) level II and (b) level III of benign (circles) and malignant (stars) tissues, with respect to the corresponding normal ones for perpendicular component.

From Fig. 5 it is clear that a Gaussian curve better fits the histogram of first-level high-pass coefficients of cancerous tissues in comparison to normal ones. These fits are found to be independent of the fluorescence intensity. In general, the measured fluorescence intensity in cancerous tissues was found to be higher than the normal ones; however, in some normal tissues, bulk fluorescence intensity was found to be more than their malignant counterparts. Surprisingly in all cases the Gaussian fit was better for the malignant ones. The scatter plots of the correlation values R^2 that depict this result are shown in Figs. 6(a) and 6(b) for the parallel and perpendicular components. The fits for the histograms of the first-level high-pass coefficients of perpendicular components of the polarized fluorescence data follow a Gaussian more precisely than parallel components. All this leads to the conclusion that ran-

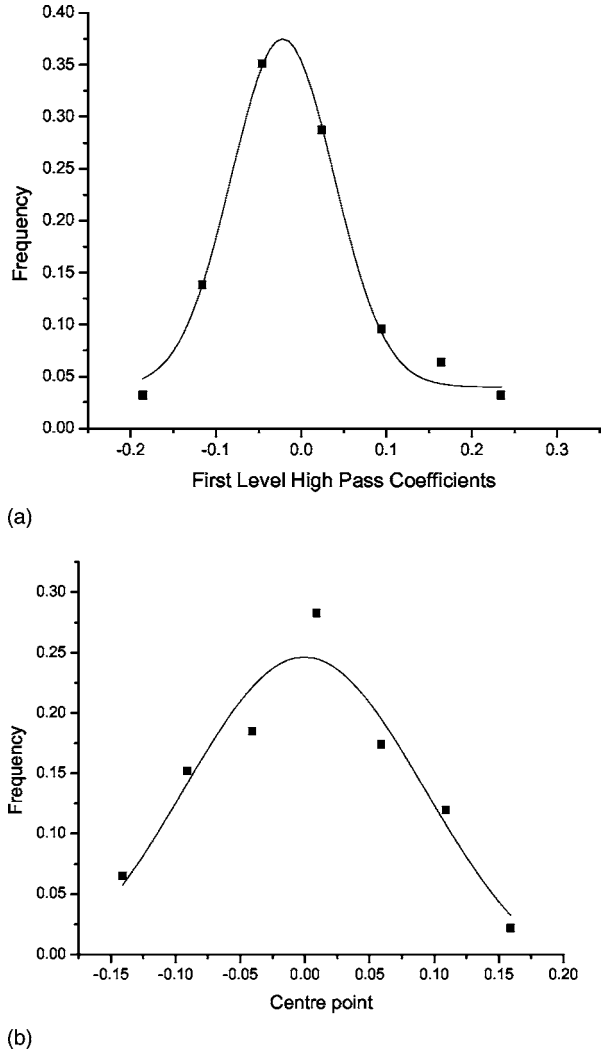

Fig. 5 (a) Histogram of normalized first-level wavelet coefficients for a typical cancer tissue (when spectral counts for cancerous tissue are greater than the corresponding normal one). Similar behavior was found when spectral counts from normal tissue were greater than its cancer counterpart. (b) Histogram of normalized first-level wavelet coefficients for a typical normal tissue (when spectral counts for cancerous tissue are greater than the corresponding normal one).

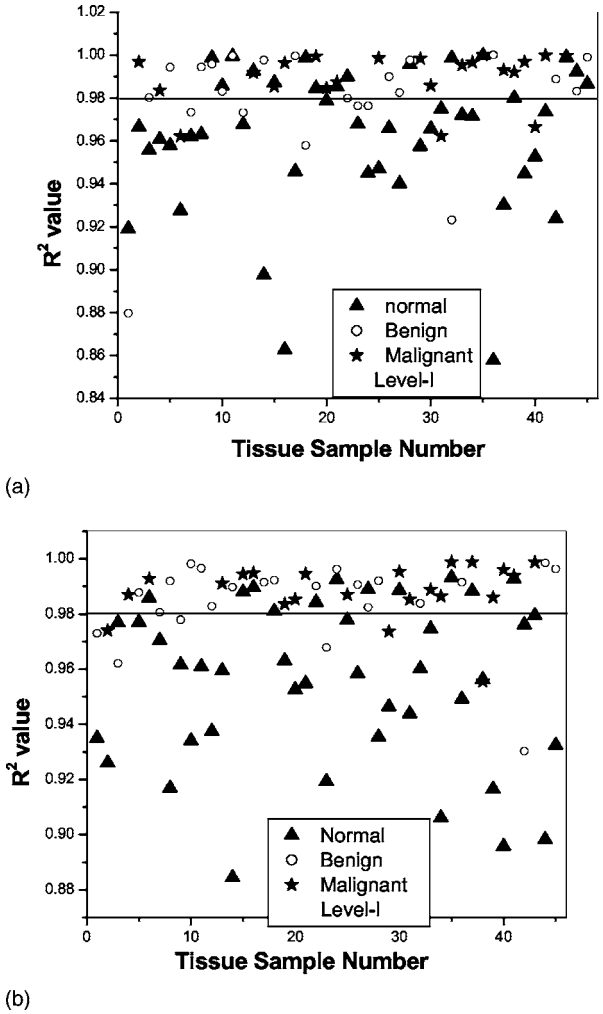

Fig. 6 R^2 value plot for (a) parallel components (level I) and (b) perpendicular components (level I).

domization is highlighted more in perpendicular components than in the parallel ones. From the scatter plot it is clear that values of R^2 for most benign and malignant tissues are greater than those of corresponding normal ones. Here, perpendicularly polarized spectral data provide better discrimination with a sensitivity of 82.2% and specificity of 75.6% (parallel polarized spectra gives sensitivity of 77.8% and specificity of 64.4%). It should be mentioned that unpolarized data showed poor discrimination and hence have not been discussed here.

4.3 *Power Plot*

It may be noted that a study of the distribution of power at various levels was performed. The power spectra at different levels are defined as the sum of the squares of high-pass coefficients at different levels. Keeping in mind that $\Sigma I_i^2 = \Sigma(c_k)^2 + \Sigma_{j,k}(d_{jk})^2$, due to Parseval's theorem, normalization of the power spectra is done by dividing it by the sum of the square of intensities at all the wavelengths. However, it does not extract very clear signatures that can differentiate between malignant and benign tumors, as done by the standard deviation.

4.4 *Low-pass Coefficients*

The low-pass coefficients represent the average part of the data. The window over which averaging has been done depends on the level. For example, at the fifth level, each low-pass coefficient represents the average of 32 data points of the original data set. The fourth- and sixth-level low-pass coefficients represent, respectively, the average of 16 and 64 data points. Characteristic fifth-level low-pass wavelet coefficients of cancerous, benign, and normal tissues averaged over all patients are shown in Fig. 7. The values of low-pass coefficients of normal tissues show an exponential behavior, not seen in cancer tissues. The error bars represent deviations in the coefficients from different samples studied. It is seen that the low-pass wavelet coefficients clearly bring out the differentiation between various tissue types.

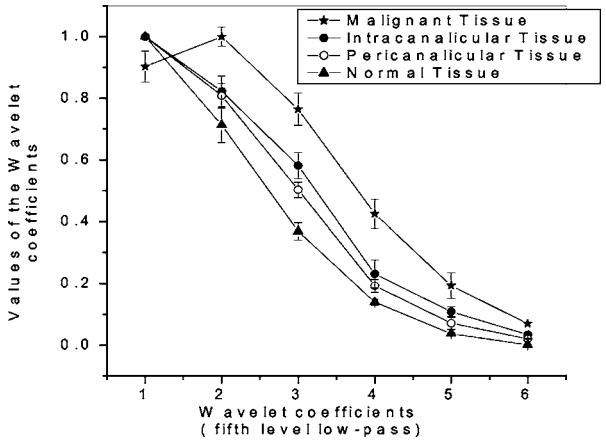

Fig. 7 Characteristic fifth-level low-pass wavelet coefficients of malignant, pericanalicular, intracanalicular, and normal tissues averaged over patients.

5 Conclusion

Wavelet analysis has enabled us to separate fluctuations at various scales in the polarized fluorescence spectra of human breast tissues. Standard deviation, power spectra, and low-pass coefficients at various levels have been computed for both parallel and perpendicular components. The above-mentioned diagnostic parameters differentiate the normal, benign, and malignant tissues better with perpendicularly polarized light. The two different types of benign tissues, intracanalicular and pericanalicular tissues, are also distinguished from one another by wavelet analysis. A significant finding of this study is the randomization in the perpendicular component of polarized fluorescence spectra of tissues. This is expected to be more in tumors than in normal tissues due to high scattering effects, and interestingly our studies have corroborated the same. We find that the fluctuations in cancerous tissues can be very accurately fitted with Gaussian distributions, which is not true for their normal counterparts. This is quantified through the correlation factor (R^2), which is higher for malignant tissues. The specificity and sensitivity of perpendicular component of fluorescence data are 75.6% and 82.2%, respectively. The wavelet-based approach also brought out features in averaged spectral profile, which differ in tissue types. This is seen in the fifth-level low-pass coefficients, which highlight distinction among tissue types. In the early stage of the tumors, when well-defined structures in the spectral profile may not have manifested, the spectral fluctuations can possibly distinguish the cancerous tissues from the normal ones.

References

1. M. Cutler, "Transillumination as an aid in the diagnosis of breast lesions," *Surg. Gynecol. Obstet.* **48**, 721–730 (1929).
2. A. Mayevsky and B. Chance, "Intracellular oxidation-reduction state measured in situ by a multichannel fiber-optic surface fluorometer," *Science* **217**, 537–540 (1982).
3. R. J. Bartrum and H. C. Crow, "Transillumination light scanning to diagnose breast cancer: a feasibility study," *AJR, Am. J. Roentgenol.* **142**, 409–414 (1984).
4. V. Backman, M. B. Wallace, L. T. Perelman, J. T. Arendt, R. Gurjar, M. G. Muller, Q. Zhang, G. Zonois, E. Kline, T. Mcgillican, S. Shapshay, T. Valdez, K. Badizadegan, J. M. Craford, M. Fitzmaurice, S. Kabani, H. S. Levin, M. Seiler, R. R. Dasari, I. Itzkan, J. Van Dam, and M. S. Feld, "Detection of preinvasive cancer cells," *Nature (London)* **406**, 35–36 (2000).
5. R. R. Alfano and S. S. Yao, "Human teeth with and without dental caries studied by visible luminescent spectroscopy," *J. Dent. Res.* **60**, 120–122 (1981).
6. R. R. Alfano, D. B. Tata, J. Cordero, P. Tomashefsky, F. W. Longo, and M. A. Alfano, "Laser induced fluorescence spectroscopy from native cancerous and normal tissue," *IEEE J. Quantum Electron.* **20**, 1507–1511 (1984).
7. C. Kittrel, R. L. Willet, C. de Los Santos Pacheo, N. B. Ratcliff, J. R. Kramer, E. G. Malk, and M. S. Feld, "Diagnosis of fibrous arterial atherosclerosis using fluorescence," *Appl. Opt.* **24**, 2280–2281 (1985).
8. L. I. Deckelbaum, J. K. Lam, H. S. Cabin, K. S. Clubb, and M. B. Long, "Discrimination of normal and atherosclerotic aorta by laser induced fluorescence," *Lasers Surg. Med.* **7**, 330–335 (1987).
9. R. R. Alfano, G. C. Tang, A. Pradhan, W. Lam, D. S. J. Choy, and E. Opher, "Fluorescence spectra from cancerous and normal human breast and lung tissues," *IEEE J. Quantum Electron.* **23**, 1806–1811 (1987).
10. G. C. Tang, A. Pradhan, and R. R. Alfano, "Spectroscopic differences between human cancer and normal lung and breast tissues," *Lasers Surg. Med.* **9**, 290–295 (1989).
11. G. C. Tang, A. Pradhan, W. Sha, J. Chen, C. H. Liu, S. J. Wahl, and R. R. Alfano, "Pulsed and CW laser fluorescence spectra from cancerous, normal, and chemically treated normal human breast and lung tissues," *Appl. Opt.* **28**, 2337–2342 (1989).
12. R. Richards-Kortum and E. Sevick-Muraca, "Quantitative optical spectroscopy for tissue diagnosis," *Annu. Rev. Phys. Chem.* **47**, 555–606 (1996).
13. G. A. Wagnieres, W. M. Star, and B. C. Wilson, "In vivo fluorescence spectroscopy and imaging for oncological applications," *Photochem. Photobiol.* **68**, 603–639 (1998).
14. Recent references on various applications of lasers in biology and medicine can be found in *IEEE J. Quantum Electron.* **9**(2) (March/April 2003).
15. S. K. Majumder and P. K. Gupta, "Synchronous luminiscence spectroscopy for cancer diagnosis," *Lasers Life Sci.* **9**, 143–152 (2000).
16. B. V. Laxmi, R. N. Panda, M. S. Nair, A. Rastogi, D. K. Mittal, A. Agarwal, and A. Pradhan, "Distinguishing normal, benign and malignant human breast tissues by visible polarized fluorescence," *Lasers Life Sci.* **9**, 229–243 (2001).
17. G. A. Wagnieres, A. Studzinski, D. Braichotte, P. Monnier, C. Depeursinge, A. Chatelain, and H. van den Bergh, "Clinical imaging fluorescence apparatus for the endoscopic photodetection of early cancers by use of Photofrin II," *Appl. Opt.* **36**, 5608–5620 (1997).
18. N. Ramanujam, *Encyclopedia of Analytical Chemistry*, Wiley, New York (2000).
19. J. W. Berg and R. V. P. Hutter, "Breast cancer," *Cancer* **75**, 257–269 (1995).
20. C. M. Gardner, S. L. Jacques, and A. J. Welch, "Fluorescence spectroscopy of tissue: recovery of intrinsic fluorescence from measured fluorescence," *Appl. Opt.* **35**, 1780–1792 (1996).
21. J. Wu, M. S. Feld, and R. P. Rava, "Analytical model for extracting intrinsic fluorescence in turbid media," *Appl. Opt.* **32**, 3585–3595 (1993).
22. A. J. Durkin and R. R. Kortum, "A comparison of methods to determine chromophore concentrations from fluorescence spectra of turbid samples," *Lasers Surg. Med.* **19**, 75–89 (1996).
23. R. W. Dillon and M. Goldstein, *Multivariation Interpretation of Clinical Laboratory Data*, Marcel Dekker, New York (1987).
24. N. Ramanujam, M. F. Mitchell, A. Mahadevan-Jansen, S. Thomsen, G. Staerkel, A. Malpica, T. Wright, A. Atkinson, and R. Richards-Kortum, "Cervical pre-cancer detection using a multivariate statistical algorithm based on laser induced fluorescence spectra at multiple excitation wavelengths," *Photochem. Photobiol.* **64**, 720–735 (1996).
25. E. N. Atkinson, M. F. Mitchell, N. Ramanujam, and R. Richards-Kortum, "Statistical techniques for diagnosing CIN using fluorescence spectroscopy: SVD and CART," *J. Cell Biochem. Suppl.* **23**, 125–130 (1995).
26. H. J. van Staveren, R. L. van Veen, O. C. Speelman, M. J. Writjes, W. M. Star, and J. L. Roodenburg, "Classification of clinical autofluorescence spectra of oral leukoplakia using an artificial neural network: a pilot study," *Oral Oncol.* **36**, 286–293 (2000).

27. S. K. Majumder, N. Ghosh, S. Kataria, and P. K. Gupta, "Nonlinear pattern recognition for laser-induced fluorescence diagnosis of cancer," *Lasers Surg. Med.* **33**, 48–56 (2003).
28. N. Agarwal, S. Gupta, Bhawna, A. Pradhan, K. Vishwanathan, and P. K. Panigrahi, "Wavelet transform of breast tissue fluorescence spectra: a technique for diagnosis of tumors," *IEEE J. Sel. Top. Quantum Electron.* **9**, 154–161 (2003).
29. D. Bicout, C. Brosseau, A. S. Martinez, and J. M. Schmitt, "Depolarization of multiply scattered waves by spherical diffusers: influence of the size parameter," *Phys. Rev. E* **49**, 1767–1770 (1994).
30. R. S. Cotran, V. Kumar, and S. L. Robbins, *Robbins Pathologic Basis of Disease*, 5th ed., Saunders, Philadelphia (1994).
31. J. Lackowicz, *Principles of Fluorescence Spectroscopy*, Plenum Press, New York (1983).
32. I. Daubechies, *Ten Lectures on Wavelets*, Vol. 64, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA (1992).
33. C. Chui, *An Introduction to Wavelets*, Academic Press, New York (1992).
34. G. Kaiser, *A Friendly Guide to Wavelets*, Birkhäuser, Boston, MA (1994).
35. G. W. Wornell, *Signal Processing with Fractals: A Wavelet Based Approach*, Prentice Hall, Englewood Cliffs, NJ (1996).
36. L. T. Perelman, V. Backman, M. Wallace, G. Zonios, R. Manoharan, A. Nusrat, S. Shields, M. Seiler, C. Lima, T. Hamano, I. Itzkan, J. Van Dam, J. M. Crawford, and M. S. Feld, "Observation of periodic fine structure in reflectance from biological tissue: a new technique for measuring nuclear size distribution," *Phys. Rev. Lett.* **80**, 627–630 (1998).
37. V. Backman, R. Gurjar, K. Badizadegan, I. Itzkan, R. R. Dasari, L. T. Perelman, and M. S. Feld, "Polarized light scattering spectroscopy for quantitative measurement of epithelial cellular structures *in situ*," *IEEE J. Sel. Top. Quantum Electron.* **5**, 1019–1026 (1999).