Characterizing breast cancer tissues through the spectral correlation properties of polarized fluorescence

Anita H. Gharekhan
Gujarat University
C. U. Shah Science College
Ahmedabad, 380 009, India

Siddharth Arora
K. B. K. Mayya
Prasanta K. Panigrahi
Physical Research Laboratory
Ahmedabad-380 009, India
and
Indian Institute of Science Education and Research
Kolkata, Salt Lake
Kolkata, 700 106, India

M. B. Sureshkumar
M. S. University of Baroda
Department of Physics
Faculty of Science
Vadodara, 390 002, India

Asima Pradhan
Indian Institute of Technology
Department of Physics and Centre for Laser Technology
Kanpur, 208 016, India

Abstract. We study the spectral correlation properties of the polarized fluorescence spectra of normal and cancerous human breast tissues, corresponding to patients belonging to diverse age groups and socio-economic backgrounds. The emission range in the visible wavelength regime of 500 to 700 nm is analyzed, with the excitation wavelength at 488 nm, where flavin is one of the active fluorophores. The correlation matrices for parallel and perpendicularly polarized fluorescence spectra reveal correlated domains, differing significantly in normal and cancerous tissues. These domains can be ascribed to different fluorophores and absorbers in the tissue medium. The spectral fluctuations in the perpendicular component of the cancerous tissue clearly reveal randomization not present in the normal channel. Random matrix-based predictions for the spectral correlations match quite well with the observed behavior. The eigenvectors of the correlation matrices corresponding to large eigenvalues clearly separate out different tissue types and identify the dominant wavelengths, which are active in cancerous tissues. © 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2997376]

Keywords: fluorescence spectroscopy; tissue; biomedical optics.

Paper 07389RR received Sep. 28, 2007; revised manuscript received Aug. 5, 2008; accepted for publication Aug. 8, 2008; published online Oct. 16, 2008.

1 Introduction

Current optical diagnosis is beginning to emerge as a viable tool for tumor detection and for monitoring the stages of growth.1–4 A number of fluorophores, starting from structural proteins to various enzymes and coenzymes, are present in human tissues that can reflect the biochemical and morphological changes taking place in cancerous tissues.5–11 Fluorescence emission can differ significantly in normal and cancerous tissues due to the differences in concentration of absorbers12,13 and scatterers, and also the scatterer sizes.14,15

The presence of scatterers leads to randomization of light, thereby generating a depolarized component in the fluorescence spectra, indicating that polarized fluorescence spectroscopy can be useful in isolating characteristic spectral features from the diffuse background. It should be noted that multiple scattering effects at the excitation and emission wavelengths are not the sole contributor to depolarization of fluorescence. Even in the absence of multiple scattering, fluorescence can be depolarized due to intrinsic causes such as random orientation of the fluorophore molecules, rotational diffusion of fluorophores, radiationless energy transfer, etc.17

It was shown previously that after individual scattering at excitation or emission wavelengths, the polarization anisotropy (ratio of the polarized component of fluorescence to the total fluorescence) of fluorescence is reduced by a constant factor.18 The value for the factor depends on the scattering properties (size, shape, refractive index of the scatterer, etc.) of the medium. Since multiple scattering is the dominant cause of depolarization of fluorescence from a turbid-medium-like tissue, for such a medium, with crossed excitation and emission polarizer (perpendicular component or the depolarized component of fluorescence), one would record fluorescence that has undergone larger scattering in the medium and has therefore traversed a larger path in the medium. Thus this component will be more sensitive to both absorption and randomization. In contrast, the parallel component will record fluorescence contributions that have not undergone significant scattering (not traveled beyond a few transport lengths in the turbid medium), and thus it bears more prominent signatures of intrinsic fluorescence of the tissue fluorophores. Hence the parallel and perpendicular components of the tissue autofluorescence are expected to reveal different aspects of tissue characteristics.19

Recently, spectral fluctuations have been studied in the wavelet domain.20,21 These differ significantly in normal and cancerous tissues. More randomization was observed for cancerous tissue fluorescence compared to normal cases. This
was particularly evident in the perpendicular channel.

The goal of this work is to extract spectral fluctuations through singular value decomposition for its characterization. Apart from providing an independent procedure to corroborate the studies through wavelets, it also removes the basis bias, inherent in wavelet transform, while extracting spectral fluctuations. Furthermore, the random and nonrandom components can be quantitatively extracted through the random matrix theory, where the eigenvalue distribution of the correlation matrix reveals these features. We study the correlation properties of polarized tissue fluorescence from normal and cancerous human breast tissues through singular value decomposition. Apart from parallel and perpendicular components, the differences in their intensities are also analyzed, since the same is relatively free of the random component. The dominant wavelengths are identified, where correlations differ significantly between diseased and normal tissues. We further study the degree of correlations in different wavelength domains and develop an understanding about the dominant fluorophores and absorbers under investigation in the visible wavelength range of 500 to 700 nm. It is expected that tissue fluorophores like flavin and its derivatives, as well as porphyrin, are active here. We carefully analyze the spectral correlation properties using the random matrix prediction with regard to the nature of fluctuations. It is found that small magnitude spectral fluctuations in the perpendicular component of cancerous tissues indeed match with the previous prediction, indicating stronger spectral randomization in cancerous tissues as compared to normal ones.

2 Experimental Procedure and Tissue Pathology

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 a triplemate monochromator (SPEX-1877E and PMT, RCA C-31034). For polarized fluorescence, a depolarizer was used after the analyzer 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. The polarized fluorescence spectra were corrected for the different sensitivities of the instrument to vertical and horizontally polarized light to isolate the specimen’s inherent polarization response from the measured. The value for $G (I_{HV}/I_{MM})$ factor, the ratio of the sensitivity of the instrument to vertical and horizontally polarized light, was measured for the entire fluorescence wavelength region 500 to 700 nm. The spectral dependence of $G$ was used to correct the polarized fluorescence spectra. The details of the measurements are provided in our previous papers.

Tissue samples were kept in moist saline and frozen (at 4 °C) immediately after biopsy. After the biopsy, a part of the tissue sample was sent for histopathology and the other part was used for fluorescence measurements. The experiment was performed within a few hours of the surgery after thawing the sample. These were then correlated with their histopathological results.

It needs to be emphasized that a single tissue sample was divided to be used for both optical and histopathological analysis. The section dimensions were approximately 5 × 5 mm in size. The spot size of the laser on the sample was 0.5 mm. The optical examination was from a single spot. Multiple samples of optical examination were conducted on a few samples, where cancerous samples were large. Since the laser spot size was small, there was no overlap between tissue areas. The same number of examinations was also conducted on the respective normal counterparts.

The patients belonged to different age groups as well as socioeconomic backgrounds. The breast cancers were also not of the same type and stage. Most of the specimens were ductal carcinoma grade 3, some were grade 2, and many of them were metastasized. A few of them were invasive lobular carcinoma.

3 Correlation Behavior of Tissue Autofluorescence

As was mentioned earlier, the fluorescence spectra were obtained from tissue samples of patients from different age groups, as well as socioeconomic backgrounds. These were supplied by Ganesh Shankar Vidyarthi Memorial Medical College, Kanpur, India. The samples were excited by 488-nm wavelength plane polarized light from an Ar-ion laser, and the parallel and perpendicularly polarized fluorescence light was measured from 500 to 700 nm. In this wavelength region flavin adenine dinucleotide (FAD) and flavin mono nucleotide (FMN) are the dominant fluorophores. Intracellular riboflavin, flavin coenzymes, and flavoproteins show slightly shifted fluorescence (540 to 560 nm) compared to flavins. The main absorber in this region is porphyrin, which absorbs at 540-nm (band) and 580-nm (band) wavelengths. Porphyrin also acts as a weak fluorophore, with an emission peak at 630 nm.

We now proceed to obtain the correlation matrix and its eigenvectors. For the construction of the empirical correlation matrix $C$, we compute $\delta \lambda (k)$ through mean subtraction of the original intensity values $I_k$. Here $i$ takes vales from 1 to 200, corresponding to the wavelength range 501 to 700 nm. Index $k$ represents the sample number. The $\delta \lambda$s have been normalized to have unit variance.

Since the patients belonged to different socioeconomic background and are of different ages, the mean subtraction has been carried out to remove the average part of the fluorescence intensity at each wavelength. The mean value was obtained at each wavelength by taking the average of the intensity over patients. This was subtracted from each patient’s intensity. Standard deviation was also computed over the patients at each wavelength. The same was then divided with the mean subtracted data to obtain the correlation matrix.

Explicitly

$$C_{ij} = \frac{1}{N} \sum_{k=1}^{N} \delta \lambda_i (k) \delta \lambda_j (k).$$

(1)

Here, $A_{ij} = \delta \lambda_i (k)$ and the correlation matrix is of the form $C = (A^T A)/N$.

If the correlation matrix is derived from random numbers, the corresponding eigenvalues have maximum and minimum values, given by:
\[ \lambda_{\text{min}} = \sigma^2(1 + 1/Q \pm 2\sqrt{1/Q}). \]  

(2)

Here, \( \sigma^2 \) is the variance (one for the present case) and \( Q = 200/N \). All the eigenvalues \( \lambda \) lie between \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \), with a density \( \rho_c(\lambda) \) of the form:

\[ \rho_c(\lambda) = \frac{Q}{2\pi\sigma^2} \left[ (\lambda_{\text{max}} - \lambda)(\lambda - \lambda_{\text{min}}) \right]^{1/2}. \]  

(3)

As is shown later, this density distribution fits well with the eigenvalues of the correlation matrix derived from the perpendicular polarized fluorescence spectra. In our analysis, the number of samples is 60, derived from 34 patients. In some cases, multiple fluorescence spectra have been collected from different tissue locations of one patient. For the possible complications arising from finite dimensional data systems, readers are referred to Refs. 27 and 28 and references therein. The finite size of datasets creates problems in finding the true area of the correlated domains, if the domain size is large. In the present case, both for normal and cancerous tissues, the domains are small compared to data size, as is seen later. When inferring about statistical characteristics, the finite size of the correlation matrix, yielding a corresponding number of eigenvalues, leads to difficulties. In the present case, this problem can affect the fits of the eigenvalue distribution in the parallel and perpendicular channels. The fit is quite robust for the perpendicular component.

4 Results and Discussion

Here we present the results of a systematic analysis of the correlation properties of the parallel and perpendicular components of the fluorescence spectra from human breast tissues for malignant and normal cases. Typical samples of the spectra are seen in Fig. 1.

It is physically clear that normal and cancer tissues have different morphological and biochemical compositions, which can give rise to different correlations in the tissue polarized fluorescence spectra. Hence the parallel and perpendicular components of the tissue autofluorescence are expected to reveal different aspects of tissue characteristics. Figures 2(a) and 2(b) depict the correlation matrices for cancer and normal fluorescence spectra in the parallel polarization component. Figures 2(c) and 2(d) depict the same for the perpendicularly polarized light. The color code has been chosen such that red denotes the highest and blue the lowest value. One clearly sees domains of high correlation differing significantly in the two tissue types. Difference of intensities are also analyzed, since the same is relatively free of random components of the spectra. It is clearly seen in Figs. 2(e) and 2(f) that the correlation domain is much bigger here compared to parallel and perpendicular ones.

All the correlation matrices show block diagonal forms. There are overlapping regions of high correlation, possibly indicating different activities, e.g., emission, absorption, etc. Three distinct domains in both cancerous and normal correlation matrices are seen for the parallel case. In cases of normal samples, the last two domains do not overlap, whereas in cancer samples there is strong overlap in the middle wavelength band, i.e., 610 to 660 nm. This is the regime where porphyrin emission takes place. We also note that the absorption activity of blood takes place in the range of 540 to 580 nm. The first domain extends from 500 to 550 nm in both tissue types. This may be due to flavin emission. We observe that in the case of normal fluorescence, this domain has partial overlap with the second domain, which may be due to the presence of intracellular riboflavin, flavin coenzymes, and flavoproteins that fluoresce in the 540- to 560-nm domain. It is worth mentioning that in the case of cancer, one observes that there is no overlap between the first two domains. The second domain in both cases corresponds to blood absorption around 580 nm. The overlap in second and third domains in cancer may be due to the porphyrin activity in this wavelength regime.

In the perpendicular channel we find two distinct domains in normal tissue types with a small overlap, whereas in cancer one finds three domains where the first two overlap strongly. Since perpendicular components travel more in the medium, it is more sensitive to absorption. We suspect that this strong overlap is reflective of this fact and the porphyrin emission at 630 nm.

We now proceed to analyze eigenvectors corresponding to the dominant eigenvalues. These are relatively free from statistical and experimental uncertainties and also reflect the correlation properties observed earlier. The elements of the eigenvectors show different characteristics. As seen in Figs. 3 and 4, eigenvectors of the correlation matrix show significant differences in normal and cancerous tissues in both parallel and perpendicular components. Large eigenvalues carry infor-
information about dominant fluorophores. The block diagonal forms seen in the correlation matrix manifests in similar structures as seen in the highest eigenvectors. It is particularly transparent for the cancerous case.

In parallel components, one observes lower contribution in the 500- to 550-nm range in the eigenvector corresponding to the highest eigenvalue, whereas in normal, there is more contribution. In this case also one sees rather small contributions from some wavelengths. These differ in cancer and normal cases. In the case of the eigenvector corresponding to the second highest eigenvalue, one sees large contributions at certain wavelengths in both tissue types. In the third eigenvector, the middle wavelength domain contributes less for cancer, as compared to normal.

The dominant eigenvector for the perpendicular case does not show a significant difference between the two tissue types. Since the dominant eigenvector relates to the largest domain in the correlation matrix, physically this implies that the large domain structures in the correlation matrix are not very different for the two tissue types. The second one shows some characteristically high contributions at certain wavelengths, implying isolated smaller domains in that spectral regime. In

Fig. 2 Correlation matrices of cancer and normal tissue samples: (a) and (b) parallel, (c) and (d) perpendicular components, and (e) and (f) difference of parallel and perpendicular components of polarized fluorescence data. The nature of correlations, as seen through different sized domains, are clearly different for cancer and normal fluorescence intensities. The difference of intensities yields much stronger correlations in the 500- to 700-nm range for normal breast samples.
the third case, the differences are not significant. Hence the parallel component retains more information about the tissue types.

The difference of intensities of parallel and perpendicular components are relatively free of the diffusive component. Hence, correlation in the intrinsic fluorescence is expected to better manifest in the corresponding correlation matrix. In Figs. 2(e) and 2(f), the correlation matrices of cancer and normal diffuse fluorescence differences indeed show very strong correlation over a much bigger wavelength regime than the cancer case. The correlation is stronger in the lower wavelength regime, with a small band around the 660-nm area. As seen in Fig. 5, for the \( I_1 - I_2 \), one sees uniform values for all the entries, showing very high degrees of correlation of all the elements with each other. One also sees, particularly for the lower eigenvalues, higher contributions from certain sections than others. There are only a few dominant eigenvalues.

The eigenvalue distributions, as seen in Fig. 6, are also quite different for parallel and perpendicular components of cancer and normal tissue correlation matrices. The eigenvalues of the perpendicular component clearly show the difference in the nature of the spectral fluctuations of the perpendicular and parallel components. The solid line fits in Fig. 6 depict the fitting of probability distribution given in Eq. (3), with the eigenvalues of the correlation matrix in the parallel and perpendicular channels. As is clear, the fit is extremely good in the perpendicular channel for cancerous tissues. The corresponding fit for the parallel case is poor, as is for the normal samples. This distribution has been derived under the assumption that the correlation matrix has been derived from Gaussian random numbers. Hence, a good fit in the perpendicular channel for the cancerous case indicates that the corresponding spectral fluctuations are more randomized. This is not the case for the parallel component for both normal and cancerous tissues.

We would like to emphasize once more that the choice of the lowest and highest eigenvalues is dictated by the random matrix theory. For randomized entries in the correlation matrix, the lowest and highest eigenvalues are given in Eq. (2) following the random matrix approach. These are determined from the number of rows and columns in the matrix \( Q = 200/N \), in the present case. Hence, they were the same in all the figures. The solid lines are fits pertaining to the distribution function given in Eq. (3), which was derived for random entries.

For smaller eigenvalues, one also sees a Gaussian distribution of the entries of the eigenvector [Fig. 4(a)] showing random characteristics. This clearly corroborates the wavelet-based analysis, which indicated that the spectral fluctuations in the perpendicular channel in cancerous tissues are much more randomized, as compared to their normal counterparts.

It is worth mentioning that the parameters of autofluorescence from tissue used for diagnosis include differences in the static spectra or decay kinetics (fluorescence lifetime). Among these, the difference in the static spectra of cancerous and normal tissues has been the more widely explored for diagnosis of cancer. The observed differences in the static fluorescence spectra between cancerous and normal tissues may arise due to several factors, like a change in concentrations, quantum efficiency of the tissue fluorophores, or a change in the modulation of the wavelength-dependent scattering and absorption properties of tissue. The usual strategy has been to develop a suitable diagnostic algorithm that can best extract the diagnostic features from the recorded bulk tissue fluorescence spectra for discriminating between cancerous and normal tissues. The present study exploits the spectral correlation properties...
properties of polarized fluorescence spectra (both the parallel and perpendicular components of fluorescence) of normal and cancerous human breast tissues for diagnosis. The results reveal interesting spectral correlation behavior of the parallel and perpendicular components of fluorescence. While the spectral fluctuations in the perpendicular component of fluorescence showed randomization effects, the parallel component revealed intrinsic information on tissue fluorophores like flavin and its derivatives, as well as porphyrins. Further, the perpendicular component also yielded complementary information on the wavelength domain where absorption takes place. The correlation properties of the polarized fluorescence spectra may thus form a useful biological metric for the diagnosis of cancer.

5 Conclusion
In conclusion, this study of spectral correlations in the polarized tissue fluorescence reveals information about wavelength ranges showing significant activity in cancer and normal tissues. Very interestingly, the spectral fluctuations in the perpendicular channel show the randomization effect of the tissue medium for the cancerous samples. This has been indicated through earlier studies involving wavelet transform. It is heartening that the fluctuations match very well with the random matrix prediction. The good fit with the probability distribution given in Eq. (2) clearly indicates the Gaussian random nature of the spectral fluctuations in the perpendicular channel for the cancerous tissue. The parallel and perpendicular components also yield other complementary information. The latter indicates the wavelength domain where absorption takes place, and the former reveals significant fluorophores in the tissue. The correlated domains clearly pinpoint the presence of fluorophores like flavin and its derivatives, as well as porphyrin. The range of this domain reveals the degree of correlation in different wavelength regimes. As is evident, the correlations are very different in parallel and perpendicular components.

Fig. 4 Plots of the entries in the eigenvectors corresponding to the three highest eigenvalues (from left to right) for (a) cancer and (b) normal perpendicular components. The x axis corresponds to the wavelength range of 500 to 700 nm. In the larger eigenvalues, one sees a domain-type structure. For the largest one, the normal sample shows much more correlation over the entire wavelength domain, since all the entries have almost identical values. For the normal component, one sees dominant contribution from certain wavelength sectors in the eigenvalues $\lambda_2$ and $\lambda_3$, as seen through the presence of peaks.
cases, for both cancer and normal tissue fluorescence. We observe that only a few eigenvalues dominate in the correlation matrix both for parallel and perpendicular tissues. The corresponding eigenvectors clearly show differences between these two tissue types.

A valuable lesson from the present analysis is that fluctuation characteristics of breast cancer tissues can be quite different from normal tissues, apart from differences in average spectral behavior between these two tissue types. This suggests that in using principal component analysis and other statistical tools like singular value decomposition to study tissue fluorescence, one should also analyze the fluctuations captured by smaller components, as compared to the principal components studied so far in the literature. Since this

Fig. 5 Plots of the entries in the eigenvectors corresponding to the three highest eigenvalues (from left to right) for (a) cancer and (b) normal tissue intensity differences of parallel and perpendicular components. The x axis corresponds to the wavelength range of 500 to 700 nm. For cancer and normal cases, we find only a few large eigenvalues. The correlation domains are much bigger, since the entries in the eigenvector are almost identical over the entire wavelength domain.

Fig. 6 Histogram plots of the level spacing distribution for the unfolded eigenvalues. Histograms (a) and (b) correspond to the eigenvalues of cancer and normal tissues, respectively, in the parallel channel. (c) and (d) give the histogram plots for the eigenvalues of cancer and normal tissues in the perpendicular channel. Solid lines depict the fit of the $\rho_s(\lambda)$ on the eigenvalues. One sees an extremely good fit for the perpendicular case.
approach is statistical, one requires a large dataset. For this purpose, one may use fine needle aspiration biopsy to go deeper into the body for accessing different breast tumor locations. We plan to study this aspect more carefully and also other types of cancer to further corroborate the observed correlation properties.

Acknowledgments

We acknowledge many useful suggestions from Amit Verma and Priyam Das, which improved the quality of the manuscript.

References


