

Fractal analysis of scatter imaging signatures to distinguish breast pathologies

Alma Eguizabal^a, Ashley M. Laughney^{bc}, Venkataramanan Krishnaswamy^c, Wendy A. Wells^c, Keith D. Paulsen^d, Brian W. Pogue^c, Jose M. Lopez-Higuera^a, Olga M. Conde^a

^aUniv. de Cantabria (Spain);

^bMassachusetts General Hospital/Harvard Medical School (United States);

^cThayer School of Engineering at Dartmouth (United States);

^dDartmouth Hitchcock Medical Ctr. (United States);

ABSTRACT

Fractal analysis combined with a label-free scattering technique is proposed for describing the pathological architecture of tumors. Clinicians and pathologists are conventionally trained to classify abnormal features such as structural irregularities or high indices of mitosis. The potential of fractal analysis lies in the fact of being a morphometric measure of the irregular structures providing a measure of the object's complexity and self-similarity. As cancer is characterized by disorder and irregularity in tissues, this measure could be related to tumor growth. Fractal analysis has been probed in the understanding of the tumor vasculature network. This work addresses the feasibility of applying fractal analysis to the scattering power map (as a physical modeling) and principal components (as a statistical modeling) provided by a localized reflectance spectroscopic system. Disorder, irregularity and cell size variation in tissue samples is translated into the scattering power and principal components magnitude and its fractal dimension is correlated with the pathologist assessment of the samples. The fractal dimension is computed applying the box-counting technique. Results show that fractal analysis of ex-vivo fresh tissue samples exhibits separated ranges of fractal dimension that could help classifier combining the fractal results with other morphological features. This contrast trend would help in the discrimination of tissues in the intraoperative context and may serve as a useful adjunct to surgeons.

Keywords: breast tumor; localized backscattering; scattering parameters; principal component analysis; fractal dimension; box counting.

1. INTRODUCTION

Tumor margin in breast conserving surgery continues being a handicap in operating rooms. Breast Conserving Therapy (BCT) is the standard of care for patients with early invasive breast cancers [1]. However, BCT requires a very accurate delineation of tumor, as residual disease decreases considerably the survival rate of patients. This is sometimes difficult to achieve with the current techniques [2]. Here, scatter-imaging signatures were used to detect and discriminate pathologies to improve the resection precision. Topological features related to image shape are then searched. To this aim fractal analysis has been considered being based on box-counting to evaluate its efficiency for malignancy detection. The fractal dimension is extracted on model-based parameters and statistical-based parameters and results are compared.

Several studies show that fractal dimension can be an interesting feature for describing the pathological architecture of tumors, an even tumor growth and its irregular shape [3]. A fractal approach also could lead to a model of tissue that could help to extract optical properties, such as local refractive index variation and size distribution [4]; and it should also be possible to identify changes in size/volume concentration of the tissue from diffuse reflectance measurements employing a fractal model of tissue [5]. A higher fractal dimension is generally associated with malignancy [6], and fractal analysis improves automatic classification of histopathology H&E images of cancer [7] as tumor samples present higher cell disorder. However, this relation it is not definitive. Fractals of breast cancer carcinoma have also been used in classification on optical coherence tomography [8], whereas stroma had higher dimension than invasive carcinoma, while adipose tissue resulted to have the lowest fractal dimension.

We propose fractal analysis of scatter imaging signatures to clarify the detection of malignancy regions. The potential of fractal analysis lies in the fact of being a morphometric measure of the irregular structures providing a measure of the

object's complexity and self-similarity. Two different images extracted from localized reflectance are used to the study. The process of analysis is summarized on Figure 1.

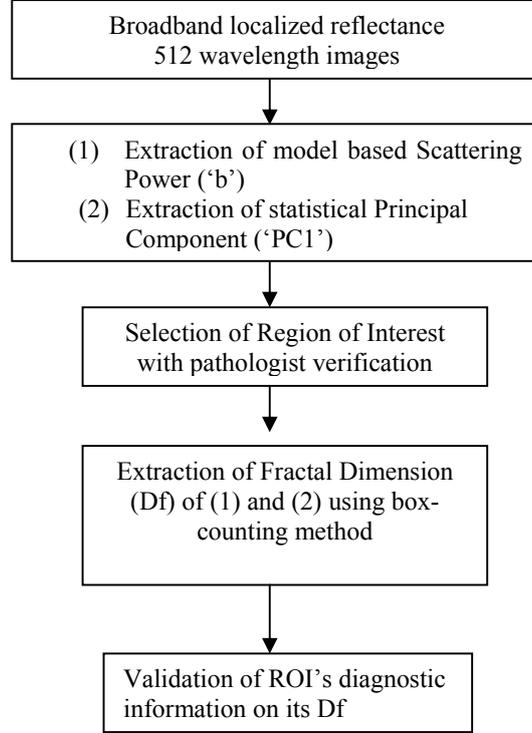


Figure 1. Process of analysis to extract the fractal dimension of scatter images.

2. MATERIALS AND METHODS

2.1 Localized reflectance and scattering model

Surgical breast tissue specimens were imaged with a custom-built micro-sampling reflectance system [9] consisting of a confocal spectroscopic set-up and a raster-scanning sampling system. Tissue samples are hydrated with a phosphate buffer solution during the measurement procedure. The system employs a quasi-confocal illumination and detection to constrain the overlapping illumination and detection spot sizes within approximately one scattering distance in tissue ($\sim 100 \mu\text{m}$ in the visible). The optical and electromechanical subsystems are integrated via a custom developed LabVIEW interface. The background response, $R_{bkgrd}(\lambda)$, is subtracted from the measured spectra, $R_{acquired}(\lambda)$, and the data is normalized, $R(\lambda)$, with respect to a diffuse reflectance standard (Spectralon, Labsphere, Inc., North Sutton, New Hampshire), as shown in Eq. 1.

$$R(\lambda) = \frac{R_{acquired}(\lambda) - R_{bkgrd}(\lambda)}{R_{spectralon}(\lambda) - R_{bkgrd}(\lambda)} \quad (1)$$

We used a combination of an empirical approximation to Mie theory and a Beer's Law attenuation factor to describe the reflectance, $R(\lambda)$ [10], as shown on (Eq.2).

$$R(\lambda) = A\lambda^{-b} \exp^{-\Gamma^* [HbT] \{StO_2^* \epsilon_{HbO_2}(\lambda) + (1 - StO_2)^* \epsilon_{Hb}(\lambda)\}} \quad (2)$$

where Γ refers to the mean optical pathlength (dependent on the illumination and detection geometry), $[HbT]$ is the total hemoglobin concentration, StO_2 is the oxygen saturation factor (ratio of oxygenated to total hemoglobin), ε_{HbO_2} and ε_{Hb} refer to the molar extinction coefficients of these two chromophores, respectively (Oregon Medical Laser Center Database, [11]). A and b are scattering amplitude and scattering power and both depend on the size and number density of scattering centers in the tissue. The scattering power parameter, b , provides better tissue discrimination capability [12] than all other scattering parameters. Consequently, this will be the employed model parameter on study.

Table 1 shows the distribution of the Regions of Interest (ROIs) imaged [10] with 42 different ROIs corresponding to 3 different diagnosis categories: non-malignant, malignant and adipose.

Table 1. Distribution of the analyzed categories of breast tissue.

Tissue type	No of ROIs
Non-Malignant	19
Malignant	12
Adipose	11
Total ROI	42

2.2 Principal Component Analysis

The Principal Component Analysis is employed to reduce the number of variables in the set of data without losing interesting information [13]. These variables are initially correlated and, after PCA, they become projected in a new space where the new projections are uncorrelated. PCA is a second order statistics method because it only requires the information contained in the covariance matrix of the input data. This fact makes the algorithm simple and easy to execute, as it only needs to compute matrix algebra equations. PCA assumes linear mixture of data, so a matrix of mixture \mathbf{W} , size $M \times M$, can be defined:

$$\mathbf{y} = \mathbf{W}\mathbf{x} \quad (3)$$

where \mathbf{y} is the $M \times N$ variable containing the uncorrelated components and \mathbf{x} the $M \times N$ input data to the algorithm. In this study M is the number of wavelengths, 512, and N the number of pixels (observations). Resulting components on \mathbf{x} are ordered by its variance, so the first will be the one with more spectral variance of the data. That is the reason why PC1 is chosen as the scattering statistical feature on this study. Previous studies [14] have suggested that principal components can extract interesting features about scatterers and malignancy of tissue.

2.3 Fractal dimension using box counting

A fractal is a shape that has the same structure at all scales [15]. Fractal analysis tries to quantify a property of roughness or natural irregularity in the intensity of an image. This variation is collected on the fractal dimension, a parameter that summarizes the fractal behavior described on power law on equation 4:

$$N = N_0 \varepsilon^{-D_F} \quad (4)$$

On this equation N expresses how many replicas of the fractal structure, scaled down by factor ε , can fit in. Both parameters are exponentially related by the fractal dimension, D_F . Box counting method [16] and a fitting to the power law is applied to estimate the D_F . The box-counting algorithm describes how many boxes of dimension ε are required to cover image objects as a function of the box size N . The process is summarized and illustrated on Figure 2.

The fractal dimension within each region of interest is calculated, after a binary consideration, using the two-dimensional box-counting method. The fractal dimension is expected to have a value between a line ($D_F = 1$) and a plane ($D_F = 2$) [6] depending on the morphology of the image.

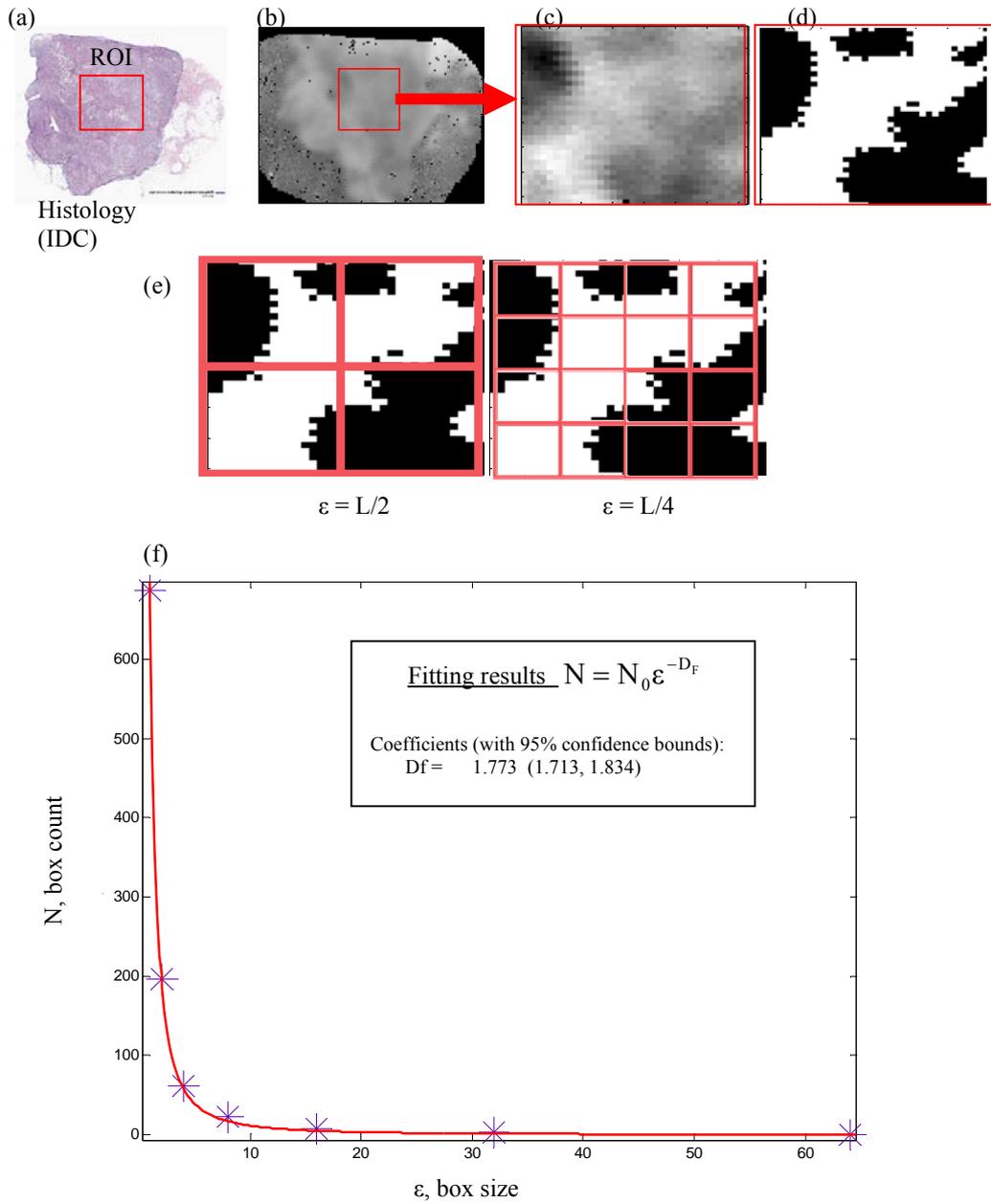


Figure 2. Summary of the image analysis process to obtain the fractal dimension through box counting. (a) H&E histopathology of the sample. (b) Scattering parameter image ('b' or first PC). (c) Region of interest and (d) Binary region of interest. (e) Different box sizes of the box counting method, where L is the longitude of the region of interest. (f) Final fitting of the box counting results to the power law that describes the fractal properties. Fractal dimension, D_f , is finally extracted from this fitting.

3. RESULTS AND DISCUSSION

Fractal analysis is proposed to define the morphology of region-based scatter signatures. The initial data consists of 512 images of localized reflectance. This spectrum is relatively featureless on each image, so we extract other parameters to explore its fractal dimension. Two different options are evaluated: (1) The power of scattering map, obtained from Mie’s empirical approach; (2) First Principal Component, obtained from correlation analysis of covariance matrix. Box-counting method obtains an estimation of the D_F of regions of interest in (1) and (2), with different diagnosis. These regions have been selected by pathologists to validate the results.

In general PCA-based fractal dimension of reflectance presents better results than scattering power map ones. Figure 3 shows that the ‘b’ scattering power region with larger fractal dimension tends to be non-malignant, which is contrary to what is expected based upon the increasing disorder of malignancy [3]; while on PCA-based regions this tendency is opposite, although is not concluding. The range of variation of the fractal dimension in Malignant and Non-malignant groups overlapped in both cases, which indicates that with fractal dimension alone high sensitivity and specificity would not be probable. However Adipose presents a quite differentiate range from Malignancy on PCA case. Table 2 shows the means and standard deviations of the fractal dimensions obtained on each process are collected, where adipose tissue is definitely the more different of the set. This fact is clearly demonstrated on the Receiving Operating Characteristic (ROC) that illustrates the performance of a binary linear classifier and its threshold variation. The pairs “Malignant/Non-Malignant” and “Malignant/Adipose” are evaluated, being always the best option the one based on the PCA analysis. In both cases the best ROC is obtained when classifying adipose tissue, which shows this is the better diagnosis to be classified by its region-based fractal analysis. Moreover, the results of the ROC evaluating the pair “Malignant/Non-Malignant” suggests that fractal dimension could help as a supplementary feature to a region on evaluation but would be too weak to detect malignancy itself. The pair- classification results for “Malignant/Adipose” are: for D_F on ‘b’ 83% probability of detection and 9% of false alarm; and for D_F on ‘PCA’ 100% and 0% respectively. Nevertheless, when classifying “Malignant/Non-Malignant” D_F on ‘b’ obtains 75% on detection and 55% on false alarm, while D_F on ‘PCA’ 75% and 35% respectively.

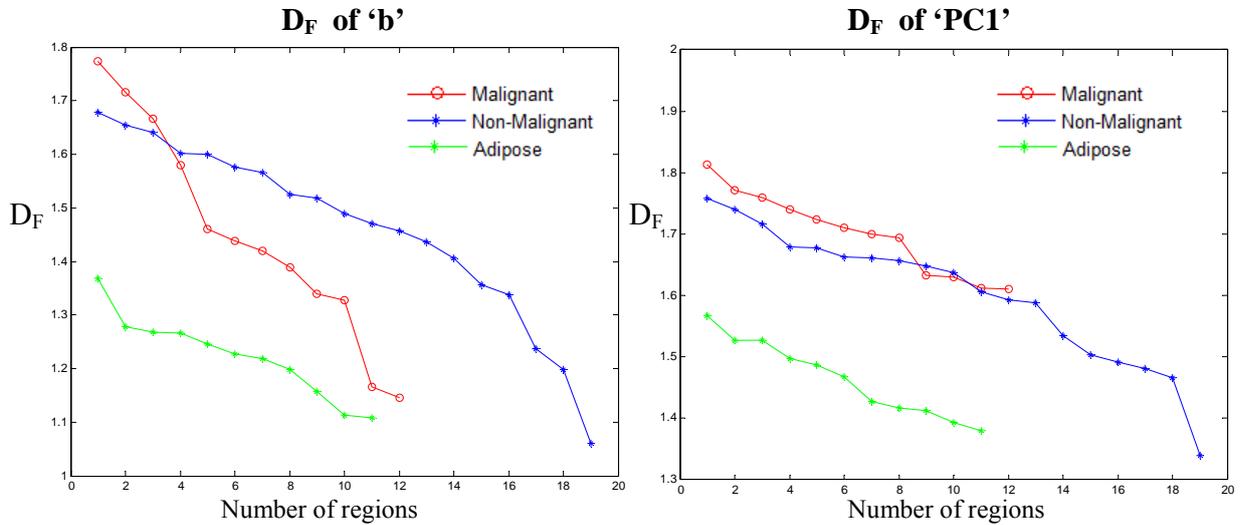


Figure 3. Distribution of the fractal dimension on the malignant, non-malignant and adipose ROIs

Table 2. Mean±STD values of the fractal dimensions on each diagnosis.

Tissue type	D_F of ‘b’	D_F of ‘PC1’
Non-Malignant	1.46±0.18	1.60±0.10
Malignant	1.45±0.16	1.67±0.06
Adipose	1.22 ±0.07	1.46±0.06

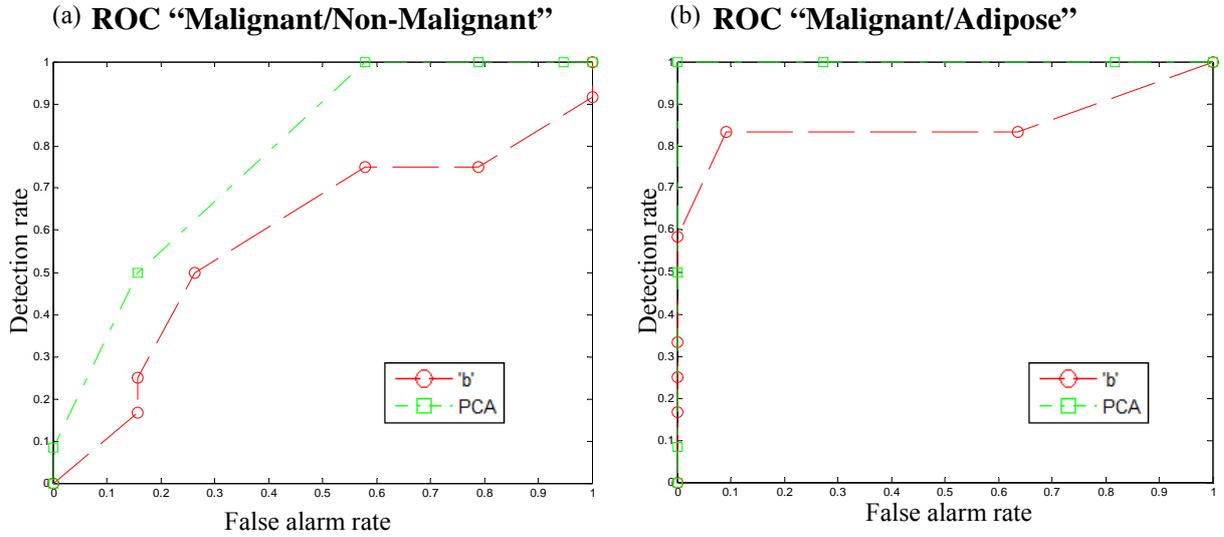


Figure 3. ROCs of classification of pairs (a) "Malignant/Non-Malignant" and (b) "Malignant/Adipose", using just the fractal dimension as a classifying feature.

4 CONCLUSIONS

Fractal analysis combined with a label-free scattering technique is proposed for describing the pathological architecture of tumors. The potential of fractal analysis lies in the fact of being a morphometric measure of the irregular structures providing a measure of the object's complexity and self-similarity. Two different groups of scatter images have been analyzed: (1) model based scattering power map, and (2) first principal component statistical map. The fractal dimension estimation is computed for both sets with a 2D box counting approach, and then examine whether this feature is concluding to classify diagnosis of the regions of interest on the images.

Results conclude that adipose tissue tends to always have a lower fractal dimension, while the difference between malignant and not-malignant samples is not as clear. Nonetheless, some very good probabilities of detection and false alarm are obtained when classifying by pairs with a binary linear classifier. Adipose tissue is accurately classified with its fractal dimension. Therefore the fractal dimension could help classifiers to identify the diagnosis of a region, what would facilitate a pixel diagnosis, combined with other features, such as texture analysis or statistical ones.

While adipose tissue shows the lowest fractal dimension on both analyses, it is not understandable why malignant and not-malignant present different behaviors on the scattering power 'b' and 'PCA' sets, and why the malignant has not always a higher dimension, as expected on references. Deeper research is still needed to understand how the morphology and the scatter concentration definitely affect the fractal shape of the scatter imaging, and why a model based approach as 'b' and a correlation based one, as PC, may present differences on its fractal characteristics. Further analysis is also required to resolve whether the resolution is correct to obtain an accurate estimation of the fractal dimension, and how size of the region of analysis could lead to erroneous estimations. Binary transformation and/or gray-levels considered may also spoil the estimation, so its influence should be as well studied on future work.

ACKNOWLEDGEMENTS

This work has been supported by CYCIT projects DA2TOI (FIS2010-19860) and TFS (TEC2010-20224-C02-02), as well as FPU PhD Scholarship (FPU12/04130), all funded by the Spanish Government.

REFERENCES

- [1] Veronesi, U., Cascinelli, N., Mariani, L., Greco, M., Saccozzi, R., Luini, A., Aguilar, M., Marubini, E., "Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer," *The New England Journal of Medicine*, 347, 1227-1232 (2002).
- [2] Fisher, B., Anderson, S., Bryant, J., Margolese, R.G., Deutsch, M., Fisher, E.R., Jeong, J-H., Wolmark, N., "Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer," *The New England Journal of Medicine*, 347, 1233–1241 (2002).
- [3] Baish R. K. and Jain R. K., "Fractals and cancer," *Cancer Res.* 60, 3683–88 (2000).
- [4] Wang, R. K. , "Modelling optical properties of soft tissue by fractal distribution of scatterers," *Journal of Modern Optics* 47(1), 103-120 (2000).
- [5] Sharma S.K., Banerjee S., "Volume concentration and size dependence of diffuse reflectance in a fractal soft tissue model". *Med Phys*, 32(6), 1767-74 (2005)
- [6] Laughney, A.M., Krishnaswamy, Rizzo J. Elisabeth, et al. "Scatter Spectroscopic Imaging Distinguishes between Breast Pathologies in Tissues Relevant to Surgical Margin Assessment" *Clinical Cancer Research* 18(22), 6315 (2012).
- [7] Esgiar, A. N., Naguib, R. N. G., Sharif, B. S., Bennett, M. K. and Murray, A. , "Fractal analysis in the detection of colonic cancer images," *IEEE Transactions on Information Technology in Biomedicine* 6(1), 54-8 (2002).
- [8] Sullivan, A. C., Hunt, J. P. and Oldenburg, A. L. , "Fractal analysis for classification of breast carcinoma in optical coherence tomography," *J.Biomed.Opt.* 16(6), 066010 (6 pp.) (2011).
- [9] Krishnaswamy, V., Hoopes, P.J., Samkoe, K.S., O'Hara, J.A., Hasan T., Pogue, B.W., "Quantitative imaging of scattering changes associated with epithelial proliferation, necrosis, and fibrosis in tumors using microsampling reflectance spectroscopy," *Journal of Biomedical Optics* 14(1), 014004 (2009).
- [10] Laughney, A.M., Krishnaswamy, V., Garcia-Allende, P.B., Conde, O.M., Wells, W.A., Paulsen, K.D., Pogue, B.W., "Automated classification of breast pathology using local measures of broadband reflectance," *Journal of Biomedical Optics* 15(6), 066019 (2010).

- [11] Jacques, S.L., Pahl, S., Oregon Medical Laser Center. 2010.
- [12] Garcia-Allende, P.B., Krishnaswamy, V., Hoopes, P.J., Samkoe, K.S., Conde, O.M., Pogue, B.W., "Automated identification of tumor microscopic morphology based on macroscopically measured scatter signatures," *Journal of Biomedical Optics* 14(3), 034034 (2009).
- [13] Semmlow, J.L., [Biosignal and Biomedical Image Processing. MATLAB-Based Applications], CRC Press (2004).

- [14] Eguizabal, A., Laughney A. M., Allende P. B. G. et al. "Enhanced tumor contrast during breast lumpectomy provided by independent component analysis of localized reflectance measures," *Biomedical Applications of Light Scattering VI*, January 21, 2012 - January 22, The Society of Photo-Optical Instrumentation Engineers (SPIE) (2012).
- [15] Petrou, M. and Garcia Sevilla, P., [Image Processing : Dealing with Texture], John Wiley & Sons, Chichester; Hoboken, NJ (2006).

- [16] Moisy, F., *boxcount*, 2008, Matlab Central File Exchange.