

## MIT Open Access Articles

*Scale-Specific Multifractal Medical Image Analysis*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Braverman, Boris, and Mauro Tambasco. "Scale-Specific Multifractal Medical Image Analysis." Computational and Mathematical Methods in Medicine 2013 (2013): 1–11.

**As Published:** <http://dx.doi.org/10.1155/2013/262931>

**Publisher:** Hindawi Publishing Corporation

**Persistent URL:** <http://hdl.handle.net/1721.1/96102>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of use:** Creative Commons Attribution



## Research Article

# Scale-Specific Multifractal Medical Image Analysis

**Boris Braverman<sup>1</sup> and Mauro Tambasco<sup>2</sup>**

<sup>1</sup> *Department of Physics, MIT-Harvard Center for Ultracold Atoms and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

<sup>2</sup> *Department of Physics, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-1233, USA*

Correspondence should be addressed to Mauro Tambasco; [mtambasco@mail.sdsu.edu](mailto:mtambasco@mail.sdsu.edu)

Received 10 June 2013; Accepted 17 July 2013

Academic Editor: Ricardo Femat

Copyright © 2013 B. Braverman and M. Tambasco. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fractal geometry has been applied widely in the analysis of medical images to characterize the irregular complex tissue structures that do not lend themselves to straightforward analysis with traditional Euclidean geometry. In this study, we treat the nonfractal behaviour of medical images over large-scale ranges by considering their box-counting fractal dimension as a scale-dependent parameter rather than a single number. We describe this approach in the context of the more generalized Rényi entropy, in which we can also compute the information and correlation dimensions of images. In addition, we describe and validate a computational improvement to box-counting fractal analysis. This improvement is based on integral images, which allows the speedup of any box-counting or similar fractal analysis algorithm, including estimation of scale-dependent dimensions. Finally, we applied our technique to images of invasive breast cancer tissue from 157 patients to show a relationship between the fractal analysis of these images over certain scale ranges and pathologic tumour grade (a standard prognosticator for breast cancer). Our approach is general and can be applied to any medical imaging application in which the complexity of pathological image structures may have clinical value.

## 1. Introduction

Many biological phenomena exhibit chaotic or fractal-like behavior, and these features have been studied extensively [1–5]. One area of widespread use has been the application of fractal geometry in the analysis of medical images as it lends itself naturally to the pragmatic characterization of irregular non-Euclidean structures found in medical imaging [6–9]. However, care must be taken when applying fractal analysis to natural objects [10]. The quintessential requirement for an object to be a fractal is for the object to exhibit a form of self-similarity to arbitrarily small scales. As such, actual fractals do not exist in nature, since there is a fundamental natural limitation to the scaling behaviour of natural objects when their substructures approach the atomic scale. Even real renderings of mathematical fractals cannot be truly fractal because of the finite resolution of the rendering. In general, the scaling behaviour of natural objects depends on the scale at which the objects are considered. Despite this limitation, most fractal analysis techniques have focused on characterizing the fractal behaviour of natural objects by

finding an interval of scales in which these objects have an approximately constant scaling behaviour [3].

Some authors have noted the danger in making this interpretation because of the possibility of illusory “fractal” behaviour where none is actually present [11], as well as the difficulty in defining the interval of scales over which the object has a consistent scaling behaviour [12]. To limit these difficulties, our approach in recent studies was to manually pick the interval of scales based on physical considerations (i.e., range of the sizes of the histological structures of interest) and the linearity of the dependence of entropy on scale [13–16].

Another approach used by others to resolve this difficulty has been to model the scale dependence of the fractal dimension in a rendering of a mathematical fractal by fitting a functional model to the dimension as a function of scale and interpreting the fit parameters as meaningful fractal dimensions for the renderings [17, 18]. The conceptual arguments behind these models are derived for renderings of mathematical fractals, and they may not necessarily hold for images of natural objects. However, the authors of these

studies are aware that natural objects do not fit the mold of real mathematical fractals.

In this study, we adopt the view expressed by others that the fractal dimension of a natural object changes with scale. Assuming that physically meaningful information in an image is contained in the scaling behaviour of the image in a certain interval of spatial dimension, then a thorough and objective approach to extract this information would be to assess the clinical relevance of different scale ranges and determine an absolute scale at which the fractal dimension is relevant for a given medical imaging application. That is, the spatial scale interval chosen for the analysis of medical images would be selected on the basis of patient classification performance and patient outcome for diagnostic and prognostic applications, respectively, rather than the resemblance of the image structures to a fractal at these image scales.

In order to address the shortcomings of existing techniques, we propose an approach to analyzing the scaling behaviour as a function of scale itself, providing an estimate of the fractal dimension as a scale-dependent parameter rather than a single fixed value. We present this fractal analysis approach in the context of the more generalized Rényi entropy [19], in which we can also compute the information and correlation dimensions of images [20].

In addition, we describe a novel fractal analysis algorithm based on integral images [21] which speeds up the computation of these generalized dimensions by orders of magnitude, and we validate our algorithm and approach by applying it to increasingly complex data, producing meaningful results throughout. We also illustrate our method with the analysis of a set of histology images acquired from tissue samples of breast tumours with modified Bloom-Richardson grades of 1, 2, and 3, in which we determine the spatial scale intervals that exhibit statistically significant differences in fractal dimensions between the different tumour grades. Looking forward, this approach provides a rapid tool to determine an absolute scale at which the generalized fractal dimensions are relevant and may allow for objective interimage comparison of medical images acquired at different resolutions.

## 2. Fractal Dimension

Fractal dimension (FD) is a generalization of the intuitive notion of topological dimension and allows for noninteger dimensions; for example, a point, a line, and a plane have both topological and fractal dimensions of 0, 1, and 2, respectively. However, fractal objects such as the Koch snowflake and Sierpinski carpet both have a topological dimension of 1 but have fractal dimensions of 1.26 and 1.89, respectively. The fractal dimension of an object is fundamentally a reflection of its scaling behaviour. Hence, to provide an estimate of the fractal dimension of a real image, we first need to define a scale-dependent function  $f(A, \epsilon)$ , with pixel intensities  $A(x, y)$  defined for  $(x, y) \in S$ , where  $S$  is the area contained in the image and  $\epsilon$  is the analysis scale. For mathematical fractals, the fractal dimension is usually defined as

$$D \equiv -\lim_{\epsilon \rightarrow 0} \frac{f(A, \epsilon)}{\log(\epsilon)}. \quad (1)$$

However, for a real image it is meaningless to analyze an infinitely small scale by letting  $\epsilon \rightarrow 0$ , and so at any given scale we can define the fractal dimension as

$$D(\epsilon) \equiv -\frac{\partial f(A, \epsilon)}{\partial \log(\epsilon)}. \quad (2)$$

The function  $f_b(A, \epsilon)$  corresponding to the box-counting dimension is commonly used in medical imaging because of its conceptual and algorithmic simplicity. In this method, the parameter  $\epsilon$  is the side length of the square boxes into which the image is partitioned, and  $f_b(A, \epsilon)$  is the logarithm of the number of squares of this size in the image that contains at least one pixel that is a part of a structure of interest.

**2.1. Generalized Dimensions.** The box-counting function  $f_b(A, \epsilon)$  is a special case of a more generalized class of functions, the Rényi entropies [19], which also generalize Shannon entropy. The Rényi entropy  $H_\alpha$  of order  $\alpha$  of a probability distribution  $\{\mu_i\}$  is given by

$$H_\alpha \equiv \frac{1}{1-\alpha} \log \left( \sum_i \mu_i^\alpha \right). \quad (3)$$

The values  $\{\mu_i\}$  are the set of natural measures of the probability distribution. For an image data set  $I(x, y)$  and a scale  $\epsilon$ , the values  $\{\mu_i\}$  are found by dividing the image into  $\epsilon \times \epsilon$  squares, and for each square by finding the proportion  $\mu_i$  of the total image intensity that is contained in the  $i$ th square:

$$\mu_i \equiv \frac{\int_{S_i} I(x, y)}{\int_S I(x, y)}, \quad (4)$$

where  $S_i$  is the domain of the image contained in the  $i$ th square. Definition (4) implies  $\sum_i \mu_i = 1$  because the set  $S$  is entirely covered by the subsquares  $S_i$ .

The fractal dimensions  $D_\alpha = -\lim_{\epsilon \rightarrow 0} H_\alpha / \log(\epsilon)$  obtained by using different orders  $\alpha$  generate the *multifractal spectrum* [6] of an image. There are three values of  $\alpha$  that have clear physical significance [22]. When  $\alpha = 0$ , we treat the term  $\mu_i^\alpha$  in (3) as the limit of  $\mu_i^\alpha$  as  $\alpha \rightarrow 0$ . Hence, for  $\mu_i \neq 0$ ,  $\mu_i^0 = \lim_{\alpha \rightarrow 0} \mu_i^\alpha = 1$ , while for  $\mu_i = 0$ ,  $\mu_i^0 = \lim_{\alpha \rightarrow 0} \mu_i^\alpha = 0$ . The entropy  $H_0$  is therefore the logarithm of the number of nonzero values of  $\mu_i$  in the image, which is the same as logarithm of the number of nonempty squares of size  $\epsilon$  in the image, equaling the box-counting function  $f_b$  from before. The Rényi entropy of order 0 will thus yield the *box-counting dimension*  $D_0$  of the image.

If  $\alpha = 1$ , we again treat (3) as a limit when  $\alpha \rightarrow 1$ . In this case, by using L'Hôpital's rule for the limit as  $\alpha \rightarrow 1$ , the Rényi entropy reduces to the Shannon entropy as follows:

$$H_1(\epsilon) = -\sum_i \mu_i \log(\mu_i). \quad (5)$$

The dimension  $D_1$  obtained from the Shannon entropy is known as the *information dimension* of the image. This measure of the fractal dimension gives the rate at which information is gained about the structure of the image as the resolution of the image increases.

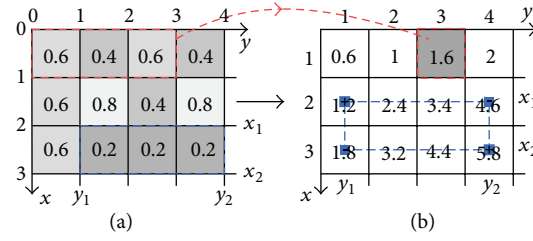


FIGURE 1: An illustration of an intensity image (a) being summed to produce an integral image (b). The sum of the elements of (a) in the dotted red box gives the corresponding element of (b). After (b) is computed, the sum of the elements of (a) in the dashed blue box, with  $2 < x < 3$ ,  $1 < y < 2$ , can be found by using (11). In this particular example, the sum equals 0.6, while  $B_{3,4} - B_{3,1} - B_{2,4} + B_{2,1} = 5.8 - 4.6 - 1.8 + 1.2$  also equals 0.6.

The third value of  $\alpha$  with a clear physical meaning is  $\alpha = 2$ , which gives rise to  $D_2$ , the *correlation dimension* of the image, addressing the number of neighbours a point of the structure has as a function of scale. That is,  $D_2$  gives the power law that relates the number of other image pixels that are within a range  $\epsilon$  of a given pixel to the value of  $\epsilon$ .

In essence, larger  $\alpha$  values assign a greater weight to the brighter parts of the image being analyzed. This is particularly useful for the analysis of medical images in which both the spatial structure and relative intensity of edge structures may carry useful information about the image.

### 3. Methods and Materials

**3.1. Scale Dependence of Fractal Dimension.** As mentioned in the introduction, natural objects do not exhibit true fractal behaviour, as their scaling behaviour depends on the scale at which the objects are considered. Hence, we propose an objective approach to determine the scale or scale interval in which the different orders of fractal dimensions may have physical relevance (e.g., diagnostic or prognostic value in medical imaging).

To estimate the fractal dimension  $D_\alpha(\epsilon)$  as a function of scale, we use definition (2) and measure the entropies  $H_\alpha(\epsilon)$  given by (3). To differentiate the entropy with respect to the analysis scale, we use the locally weighted regression and smoothing scatterplots (LOWESS) method [23], which is widely used in situations where a good theoretical model for the observed data does not exist. A low-order polynomial is fitted to a weighted subset of the data around each data point, and then all parameters (such as the derivatives) of the fitted curve can be extracted from this polynomial. In our case, we are interested in the first derivative of the entropy with respect to scale, which immediately gives a scale-dependent estimate of the fractal dimension of the image. The advantage of this method over direct numerical differentiation is its vastly enhanced robustness to noise, while its advantage over a fit to the entropy or numerical derivative data, as used by others [17, 18], is in its greater flexibility and scale resolution, as well as a lack of assumptions that may not be valid for images of natural objects.

**3.2. LOWESS Method.** To estimate parameters of a scatterplot of a data set  $Y = \{y(x)\}$  given as a function of  $X = \{x\}$

using the LOWESS method [23], for each  $x_0 \in X$ , we fit a polynomial  $p$  to the data in such a way as to minimize the weighted sum of the squared residuals  $R$  given by  $R = \sum_{x \in X} (y(x) - p(x))^2 \cdot w_{x_0}(x)$ , where  $w_{x_0}(x)$  is a weighting function. The fitted curve  $f(x)$  is then approximated around  $x_0$  by the polynomial  $p$ . In particular, for  $x \approx x_0$ , we estimate the  $n$ th derivatives of  $f$  and  $p$  as equal for  $n \leq \deg(p)$ .

The function  $w$  is central to the LOWESS method, since it makes the regression *locally* weighted. The simplest weighting function  $w$  that can be used makes the fit around a point  $x_0$  local a Gaussian with a standard deviation  $\sigma$ :

$$w_{x_0}(x) = \exp \left[ -\left( \frac{x - x_0}{\sigma} \right)^2 \right]. \quad (6)$$

Larger values of  $\sigma$  will smooth out the fit more than smaller values, producing a fit that is more resistant to noise at the cost of resolution in  $x$ . In the case of scale-dependent fractal analysis, the input data  $x$  and  $y$  to the LOWESS model are  $x = \log(\epsilon)$  and  $y(x) = H_\alpha(\epsilon)$ , from which we can obtain  $D_\alpha(\epsilon) = -f'(\epsilon)$ .

**3.3. Rapid Computation of the Natural Measures.** The most common operation encountered in the box-counting approach to fractal image analysis is the determination of the natural measure  $\mu_i$  given in (4) of a certain rectangular (usually square) subset  $S_i$  of the image. If the subset  $S_i$  is bounded by  $x_1 < x < x_2$  and  $y_1 < y < y_2$ , while the entire image is bounded by  $0 < x < x_M$  and  $0 < y < y_M$ , finding the natural measure is equivalent to calculating

$$p_i = \int_{S_i} I(x, y) dx dy = \sum_{\substack{x_1 < a \leq x_2 \\ y_1 < b \leq y_2}} A_{a,b}, \quad (7)$$

$$P = \int_S I(x, y) dx dy = \sum_{\substack{0 < a \leq x_M \\ 0 < b \leq y_M}} A_{a,b}, \quad (8)$$

$$\mu_i = \frac{p_i}{P}. \quad (9)$$

For example, in Figure 1, the sum of the image intensity  $p_i$  in image  $A$  over the dotted (red) rectangle bounded by  $0 < x < 1$  and  $0 < y < 1$  is equal to  $p_i = A_{1,1} + A_{1,2} + A_{1,3} = 0.6 + 0.4 + 0.6 = 1.6$ .

Direct summation of the image intensity over a rectangle has computational cost proportional to the area of the rectangle. A method to speed up the summation is to first compute the following *integral image* (also known as the *summed area table*) [21]:

$$B_{x,y} = \sum_{\substack{0 \leq a \leq x \\ 0 \leq b \leq y}} A_{a,b}. \quad (10)$$

This computation reduces the subsequent summation in (7) to a simple arithmetic operation (e.g., see Figure 1):

$$p_i = \sum_{\substack{x_1 < a \leq x_2 \\ y_1 < b \leq y_2}} A_{a,b} = B_{x_2,y_2} - B_{x_1,y_2} - B_{x_2,y_1} + B_{x_1,y_1}, \quad (11)$$

where we define  $B_{x0} = 0$  and  $B_{0y} = 0$  for all  $x$  and  $y$ .

The usefulness of this algorithm lies in its ability to speed up natural measure computation. The computational complexity of a naïve algorithm, in which for each scale  $\epsilon$  we determine the natural measures, (4), by directly summing the image intensity is of order  $O(A \cdot s)$ , where  $A$  is the area of the image and  $s$  is the number of different values of  $\epsilon$  considered. However, by using the integral image algorithm approach, the computational complexity falls to  $O(A)$  since the cost of computing the natural measures at a scale  $\epsilon$  is approximately equal to  $A/\epsilon^2$  and  $\sum_{\epsilon=1}^{\infty} 1/\epsilon^2 = \pi^2/6 \approx 1.6$  is small. The integral image approach is particularly advantageous for large values of  $\epsilon$  where a large number of summations are replaced with only a few subtractions. For a typical fractal scale analysis, the integral image approach is 10 to 100 times faster than the naïve algorithm.

**3.4. Algorithm.** Our method was implemented in MATLAB Version 7 (The MathWorks, Inc., Natick, MA, USA). The steps of the method are as follows.

- (i) Select the analysis scales  $\epsilon$  and entropy order  $\alpha$  values. The choices made are based on purely physical considerations, for example, the range of sizes of the structures of interest in the image being analyzed. To be general, one can begin by analyzing a wide range of  $\epsilon$  and  $\alpha$  values and select the physically meaningful subset of scales and orders.
- (ii) Normalize the original image  $A$  by dividing every element by the sum of the pixel intensities of the entire image, giving the natural measure contained in each image pixel.
- (iii) Calculate the integral matrix  $B$  for the normalized image  $A$ . For each scale  $\epsilon$ , rescale the image using the integral matrix. That is, calculate the matrix

$$R_{x,y} = \sum_{\substack{i=x\epsilon \\ j=y\epsilon \\ i=(x-1)\epsilon+1 \\ j=(y-1)\epsilon+1}}^{i=x\epsilon \\ j=y\epsilon} A_{ij}. \quad (12)$$

- (iv) For each value of  $\alpha$ , raise the natural measure matrix  $R$  to the power  $\alpha$  except for  $\alpha = 1$  obtaining  $M_{ij} = R_{ij}^\alpha$ ; for  $\alpha = 1$ , we calculate the matrix  $M_{ij} = -R_{ij} \cdot \log(R_{ij})$ .

- (v) Use (3) and (5) to give  $H_\alpha = (1/(1-\alpha)) \log(\sum M_{ij})$  for all  $\alpha \neq 1$ , while for  $\alpha = 1$ ,  $H_1 = \sum M_{ij}$ .

Our method was tested on the following data sets: exact renderings of deterministic strictly self-similar fractals, randomized renderings of deterministic fractals, randomized renderings of statistically self-similar fractals, and a set of breast histological tissue samples.

**3.5. Breast Cancer Tissue Specimens.** In a previous study [16], we had retrospectively selected 408 patients with primary invasive ductal carcinoma (IDC) of the breast from Calgary regional hospitals after appropriate ethics approval from the institutional review board. The breast tissue from these patients was used to construct tissue microarray (TMA) cores (each with a 600  $\mu\text{m}$  diameter) stained with pan-cytokeratin to highlight the morphology of epithelial architecture. The number of cores per patient ranged from one to three. Images of the cores were acquired with an effective magnification of 6.3x using an AxioCam HR digital camera (Carl Zeiss, Inc.) mounted on an optical microscope (Zeiss Axioscope). The images were saved at the camera's native resolution of 1300  $\times$  1030 pixels in tagged image file format (tiff). In this previous study [16], we found the box-counting fractal dimension of the breast cancer TMA core images to be an independent and statistically significant prognosticator. However, the study did not include an explicit examination of the role of the scale range on the fractal dimensions computed from the images. Instead, fractal dimension was computed from plots of the slope of  $f_b(A, \epsilon)$  versus  $\log(\epsilon)$  over a scale range of ( $\epsilon = 10\text{--}50 \mu\text{m}$ ), which was chosen based on a visual assessment of the range of the linear region of a small random sample of plots taken from the whole image set.

In this study, we selected all the cases from our previous study set of 408 patients [16] that satisfied the conditions of having exactly three evaluable TMA core images and contained pathologic grade information. This selection resulted in a set consisting of a total of 157 patients in the following tumour grade categories: 56 grade 1, 84 grade 2, and 17 grade 3 tumours. We applied our method to these cases, and the analysis was used to demonstrate the capabilities of the algorithm and to check that our previous choice for the scale range was a judicious one. The grayscale images of the tissues were converted into black-and-white outline images by thresholding. An example of the overall analysis process for a breast cancer tissue core is shown in Figure 2. To determine the optimal thresholding level for the edge detection, we varied the threshold level for each image to maximize the fractal dimension in the 10–50  $\mu\text{m}$  scale interval.

The motivation for using three cores per patient (chosen from different tumour regions) was to ensure that we had a representative sample of a heterogeneous tumour. For the final analysis, we selected the one core from each patient that had the greatest average fractal dimension in the 10–50  $\mu\text{m}$  scale range. The rationale for choosing the core with the greatest fractal dimension is that it is likely representative of the portion of a possibly heterogeneous tumour that has deviated most from normal epithelial breast morphology, and



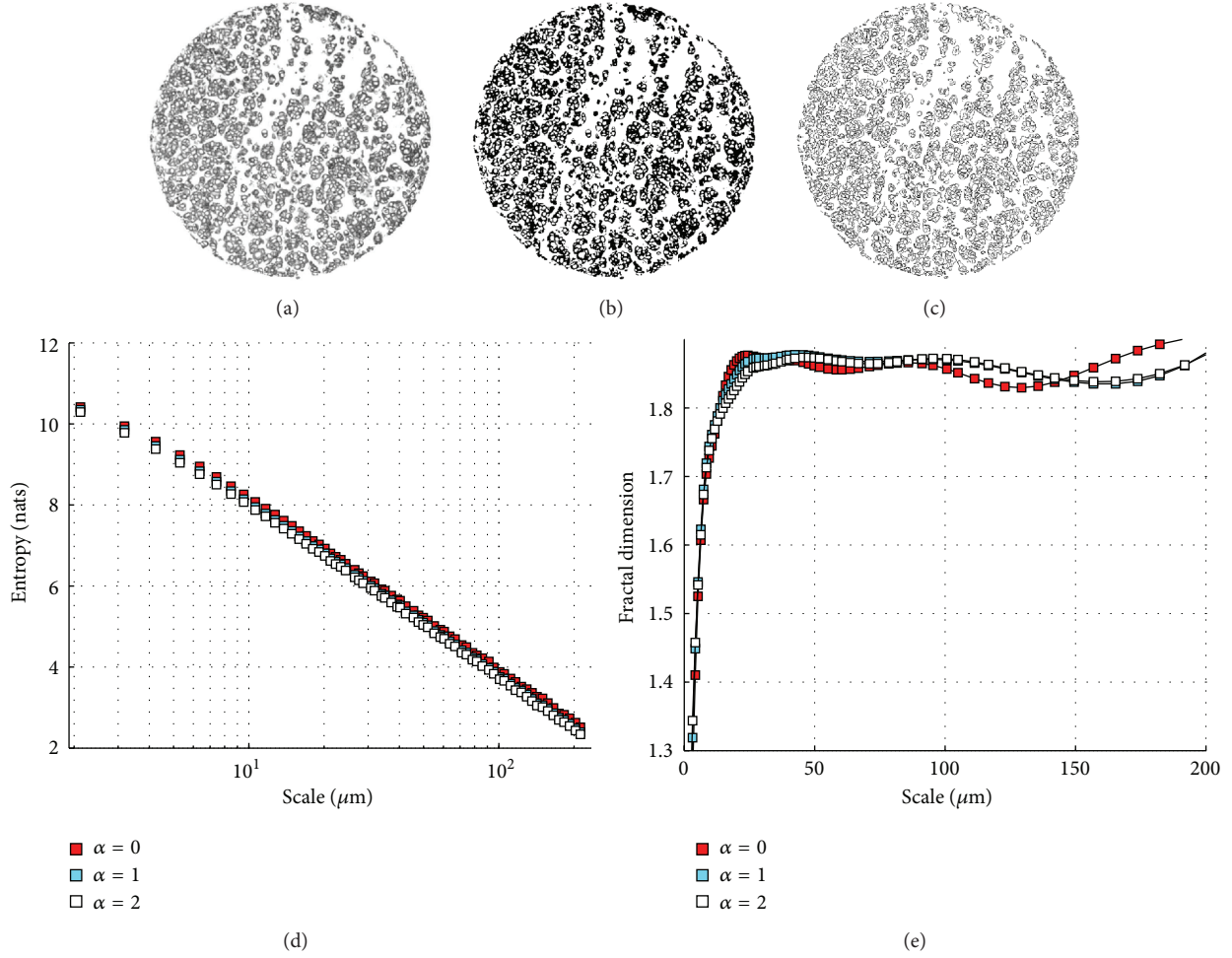


FIGURE 2: An illustration of the overall analysis process. (a) Grayscale image of a breast tissue sample (600  $\mu\text{m}$  in diameter). (b) Black and white thresholded version of (a). (c) Outlines of (b). (d) Image entropies determined from (c). (e) Scale-dependent fractal dimensions for this tissue sample.

therefore it is the most probable indicator of abnormal and/or aggressive tumour growth with metastatic potential. For the analysis, we produced a single curve of fractal dimension as a function of scale  $\epsilon$  for each patient. We averaged the fractal dimensions within each tumour grade category.

**3.6. Statistics.** We performed the statistical analysis using the MATLAB Statistics Toolbox 7.4 (The MathWorks, Inc., Natick, MA, USA). We quantified the differences between the three tumour grade categories using the nonparametric Kruskal-Wallis analysis and a multiple comparison test (MATLAB functions `kruskalwallis` and `multcompare`, resp.).

## 4. Results

**4.1. Assessment of Scale-Dependent Algorithm on Renderings of Strictly Self-Similar Fractals.** Figure 3 shows the results of our approach applied to renderings of strictly self-similar

mathematical fractals. For the Koch snowflake, Pascal triangle, and Sierpinski carpet, the estimated fractal dimensions are consistently within 0.1 of the Hausdorff dimension of the mathematical fractal being rendered throughout the scale range of 10 to 250 pixels. Note that both the mean and maximum deviations of the measured fractal dimension from the Hausdorff dimension decrease with increasing fractal dimension, increasing value of  $\sigma$  in (6), and have a minimum in the range of 20 to 50 pixels, where neither the small nor large-scale granularities of the image affect the estimate of the fractal dimension. In this smaller interval of scales, the measured fractal dimensions have a root-mean-squared deviation of 0.074, 0.040, 0.086, and 0.029 for  $\sigma = 0.3$  and 0.045, 0.018, 0.052, and 0.034 for  $\sigma = 1.0$  from the Hausdorff dimension, respectively, for the four fractals presented.

For the Koch island boundary (Figure 3(d)), a jump of fractal dimension from 1 to 1.5 (the Hausdorff dimension) can be seen at  $\epsilon \approx 4$  pixels. This behaviour is consistent with the real structure of the rendering, which at small scales consists of straight line segments of 4–8 pixels long. Hence, below

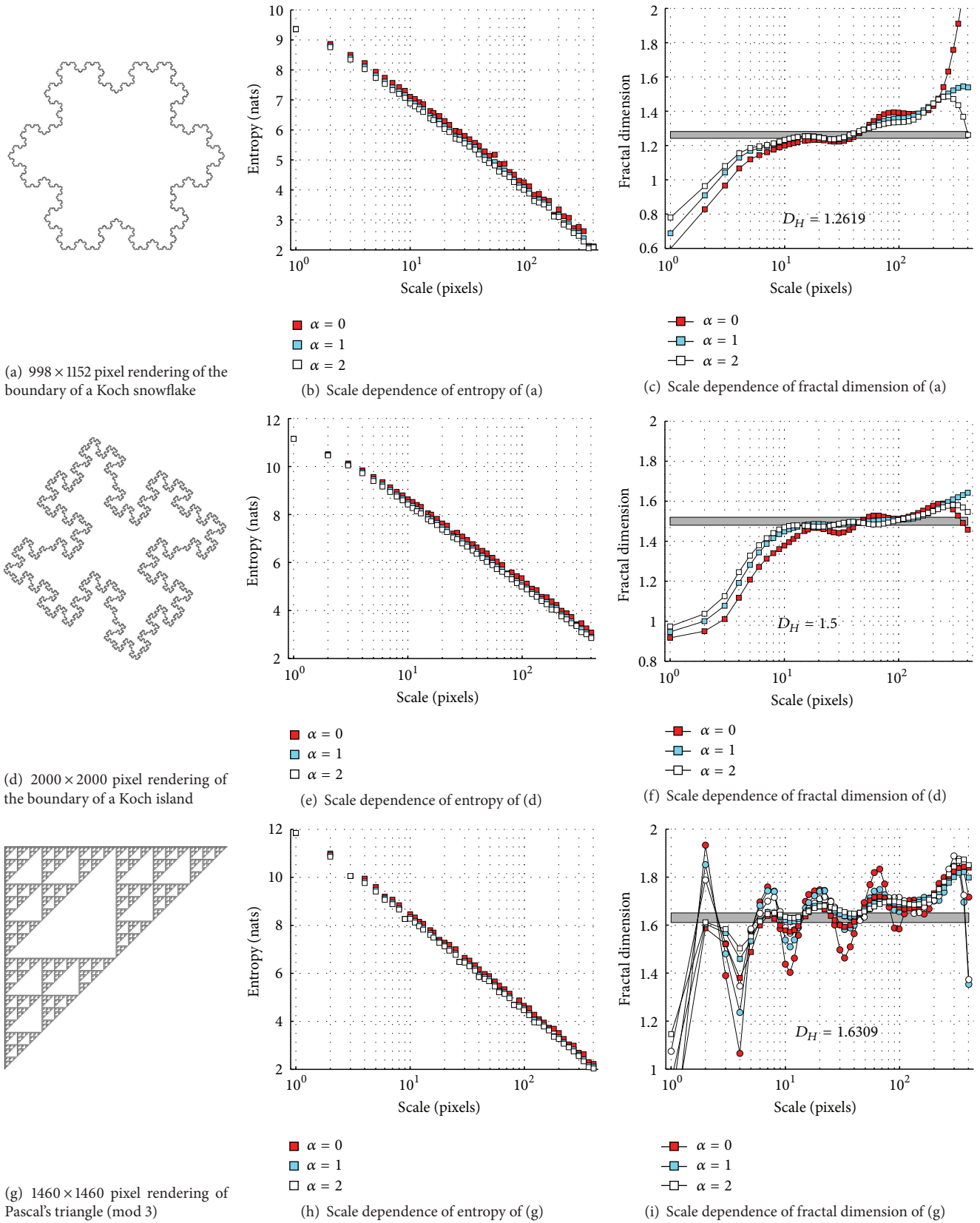


FIGURE 3: Continued.

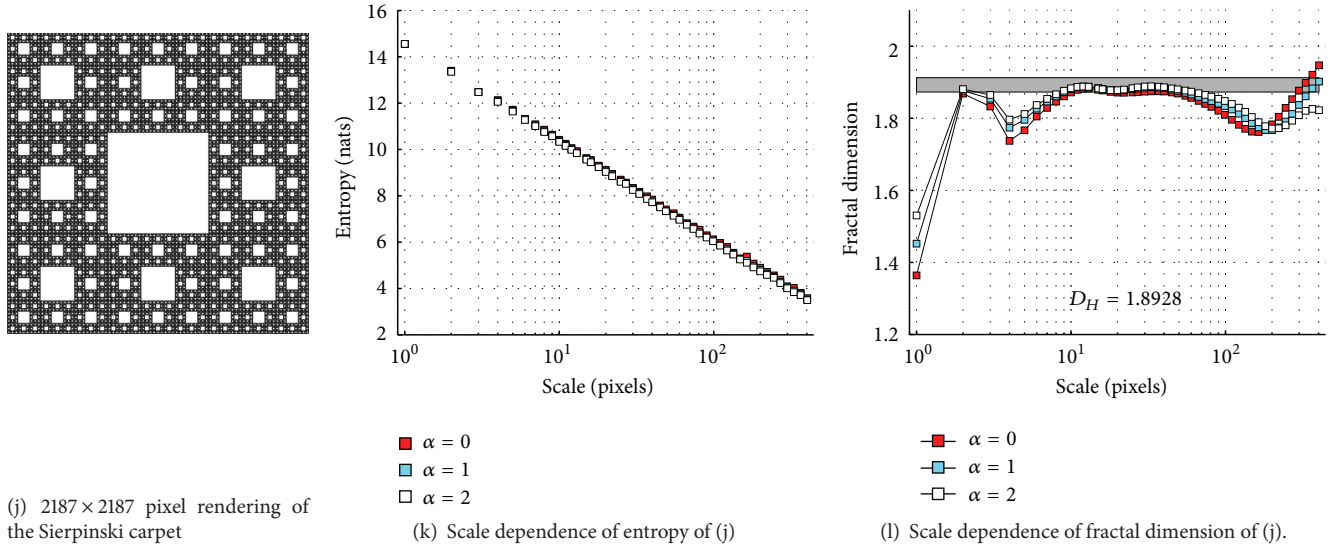


FIGURE 3: Results of applying our algorithms to four renderings of mathematical fractals. The entropies in (b), (e), (h), and (k) are in nats, which are the natural units for information and entropy, with base  $e$  rather than 2: 1 nat  $\approx$  1.44 bits, and are plotted for  $\alpha = 0, 1, 2$ . All the fractal dimension plots (c), (f), (i), and (l) use  $\sigma = 0.5$  in (6), except for (i) where  $\sigma = 0.3$  for the circles. Horizontal bar indicates Hausdorff dimension  $D_H$  of each mathematical fractal.

a scale of 4 pixels, the rendering is really linear and hence has a fractal dimension of 1 at this scale.

In Figure 3(i), we compare the estimate of fractal dimensions produced by using the values of 0.3 and 0.5 for  $\sigma$  in (6). The larger value of  $\sigma$  produces a smoother dependence of fractal dimension on scale, which is close to the Hausdorff dimension of Pascal's triangle. On the other hand, the smaller value gives an estimate which is more locally accurate, showing the nonuniform scaling behaviour of the rendering, which can be seen in the oscillations in the fractal dimension as a function of scale. In fact, this inhomogeneous scaling behaviour is seen in all four of the sample renderings. These oscillations occur because the real renderings of the fractals have discrete characteristic scales. For example, the Pascal's triangle mod 3 rendering (Figure 3(g)) consists of black and white triangles of several discrete scales: 3, 9, 27, 81, 243, and 729 pixels. Around each of these scale values, the fractal dimension experiences a large drop. This drop occurs because when the analysis scale grows through each of these scales, the white triangles in the rendering become "invisible" to the box-counting algorithm, causing a smaller than expected drop in the entropy  $H_\alpha$  and consequently causing a dip in the box-counting dimension. However, between these characteristic scales, the image becomes nearly 2-dimensional, just like the plane in which the image is contained, because the analysis cannot "see" any change in the image features.

**4.2. Assessment of Scale-Dependent Algorithm on Renderings of Statistical Fractals.** We further tested our algorithm on two kinds of statistical fractals, which are generated by random processes, but nonetheless possess a statistical form of self-similarity.

**4.2.1. Randomized Sierpinski Triangle.** An approximation to a Sierpinski triangle can be generated by an iterative random process known as the "chaos game" [24]. In each step of the game, one new point is added to the rendering of the triangle. We employed this method with a variable number of iterations (from  $n = 10$  to  $n = 10^6$ ) to generate several Sierpinski triangle approximations, with an example for 10000 points shown in Figure 4(a). When the structure is rendered with only a few points, the structure of the randomized triangle is essentially point-like or 0-dimensional at small scales. For example, with only 1000 points in a  $2000 \times 2000$  pixel image, the expected mean distance between points is  $2000/\sqrt{8n} = 23$  pixels; indeed, the fractal dimension measured for  $n = 1000$  begins to grow around that scale, but already has the Hausdorff dimension  $\log(3/2)$  scaling behaviour at larger scales above 100 pixels.

In this case, we can see that our algorithm correctly identifies the scaling behaviour of the fractal renderings: the large-scale behaviour of the fractal quickly reaches the correct scaling power law, while the smaller-scale features are really 0-dimensional. However, as more and more points are added to the approximation, the fine-scale structure of the fractal is filled in, and the fractal dimension approaches its ideal value at all scales.

**4.2.2. Brownian Surfaces.** The fractal dimension of nine  $1025 \times 1025$  pixel renderings of Brownian surfaces with dimensions 2.1 through 2.9 were evaluated by finding cross-sections through the images (which should have a dimension that is one dimension less than the surfaces themselves). The results of this analysis are shown in Figure 5. Several features can immediately be seen from Figure 5(c). All of the outlines are inherently 1-dimensional at small scales because



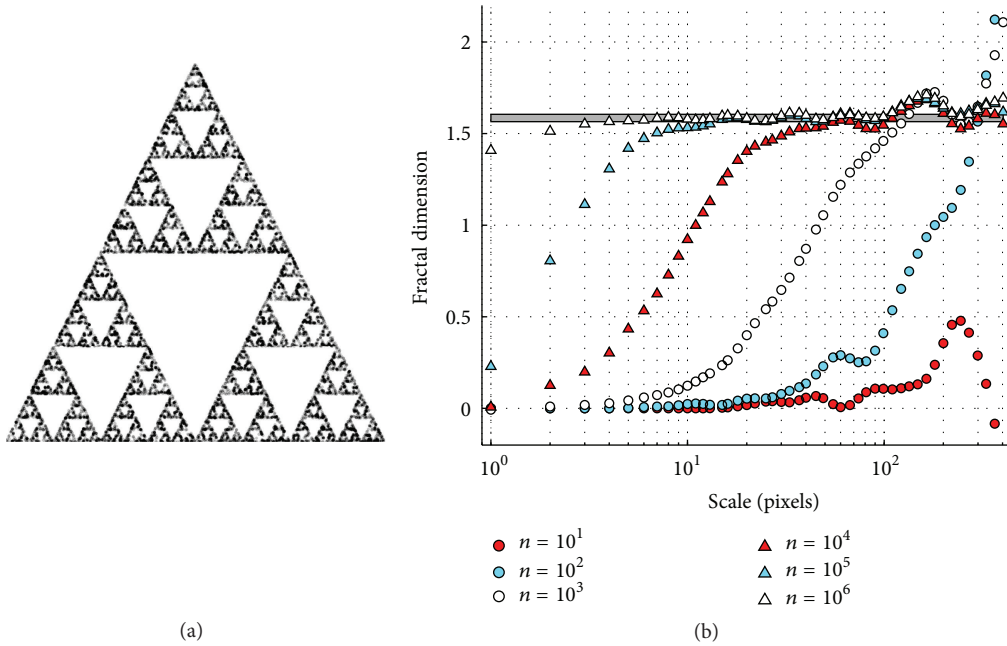


FIGURE 4: (a) 2000 pixel  $\times$  2000 pixel rendering of a statistical Sierpinski triangle ( $10^4$  points). Each point in the statistical fractal is rendered as a small grey disk for clarity. (b) Scale dependence of (box counting) fractal dimension of statistical Sierpinski triangles rendered with the indicated number of points (legend). The value  $\sigma = 0.3$  is used in (6). Horizontal bar shows Hausdorff dimension  $D_H = \log(3/2) \approx 1.585$ .

of their linear structure. The Brownian surface outlines (i.e., Figure 5(b)) pack the image more densely as the scale of the image increases, causing the fractal dimensions to approach 2 for all the cross-sections. These features are *not* due to a flaw in the algorithm, but rather reflect the true behaviour of the curves obtained by slicing through the Brownian surfaces. The behaviour of other measures of fractal dimension, such as other components of the multifractal spectrum, is similar to the box-counting dimension shown in the figure. The fractal dimensions match the expected values most closely in the 10–40 pixel range, with a root-mean-square error of less than 0.06 in this interval. This example illustrates the great sensitivity of fractal dimension measures to the scale at which they are computed and the consequent need for a scale-dependent measure of the fractal dimension to quantitatively estimate this sensitivity and choose an appropriate scale or scale range for analysis.

**4.3. Application of Scale-Dependent Method to Breast Tissue Samples.** It is readily apparent that the curves of the averaged fractal dimensions within each tumour grade category are similar at small scales below  $5\mu\text{m}$  but rapidly differentiate at larger scales Figure 6(a). In addition, there seems to be a larger difference between grades 1 and 2 at smaller scales and a larger difference between grades 2 and 3 at larger scales.

The Kruskal-Wallis analysis showed that statistically significant differences exist between at least 2 of the 3 tumour grade groups for fractal dimensions averaged over scale

ranges of  $15\text{--}50\mu\text{m}$  ( $P < 0.0001$ ) and  $100\text{--}150\mu\text{m}$  ( $P < 0.008$ ). The multicomparison test showed that a statistically significant difference ( $P < 0.0005$ ) exists between grades 1 and 2 and 1 and 3 in the  $15\text{--}50\mu\text{m}$  scale range and a statistically significant difference ( $P < 0.05$ ) between grades 1 and 3 and 2 and 3 in the  $100\text{--}150\mu\text{m}$  scale range (Figure 6(a)). Figures 6(a), 6(b), and 6(c) also show the fractal dimension distributions for the different grades and scale ranges in the form of boxplots. For these plots, the middle 50% of the data lie within the boxes, the lines within the boxes are the median values, the lines above and below the boxes show the upper and lower 25 percent of the data, respectively, and the crosses outside the boxes show outliers. The results illustrate the way scale range can affect the results and how different ranges can be useful for distinguishing different groups (e.g., grades 2 and 3 were not significantly different in the smaller  $15\text{--}50\mu\text{m}$  scale range, but differences become more significant in the  $100\text{--}150\mu\text{m}$  larger-scale range).

In a previous study using a scale range of 10 to  $50\mu\text{m}$ , it was found that  $\text{FD} < 1.56$ ,  $1.56 < \text{FD} < 1.75$ , and  $\text{FD} > 1.75$ , correlated to high, intermediate, and low survival from breast cancer [16]. The results of this study are consistent with the previous finding, as higher tumour grade is also correlated to poorer survival, and grades 1, 2, and 3 correspond to similar FD ranges (Figure 6). It is important to note that any comparisons with other results reported in the literature will only be meaningful if a similar scale range is used for finding the fractal dimension and if the tissues' specimens are stained using the same stain and tissue preparation technique.

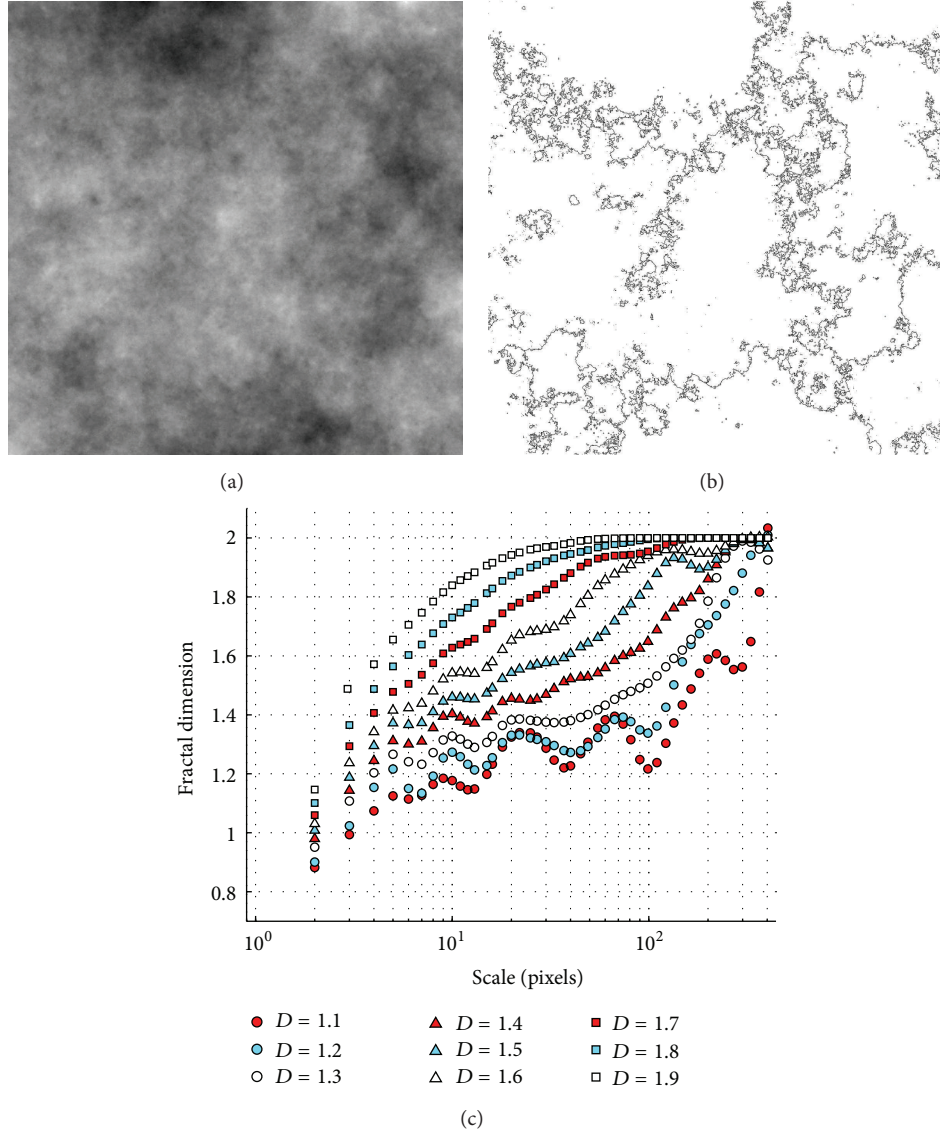


FIGURE 5: (a)  $1025 \times 1025$  pixel rendering of a Brownian surface of dimension 2.5. (b) Outlines of (a), which have a theoretical fractal dimension of  $2.5 - 1 = 1.5$ . (c) Scale dependence of fractal dimension of (a) along with 8 other Brownian surfaces of different fractal dimensions. The theoretical dimensions of the outline images are indicated in the legend.

## 5. Conclusion

In this study, we have described two novel ideas for the application of fractal analysis to medical images: fractal dimension as a scale-dependent scaling parameter of a statistical distribution and an application of the integral image method for rapid evaluation of fractal dimensions. The notion of considering fractal dimension as a model-free scale-dependent parameter is a fundamental shift in perspective on investigations of fractal image analysis. By forgoing a specific model for how an image should behave, we allow ourselves to extract as much information as possible from the image. Hence, the true power of our method lies in its ability to determine the appropriate scale range or ranges that need to be analyzed using fractal methods for

any particular application. In testing our algorithms on both real fractal structures and medical images, we showed the algorithm's reliability in measuring fractal dimensions and in picking up subtle scale-dependent features in the fractal dimensions. More specifically, our analysis of invasive breast cancer tissue cores from 157 patients has shown that the ability to differentiate images of different grades of cancer depends on the scale at which images are analyzed. As tumour grade is a prognosticator for breast cancer survival, it is evident that the analysis scale has an impact on the prognostic value of a fractal analysis approach, a point which has not been systematically studied or appreciated previously. Future studies are needed to further validate and extend the breast cancer results using an independent set of tissue images.

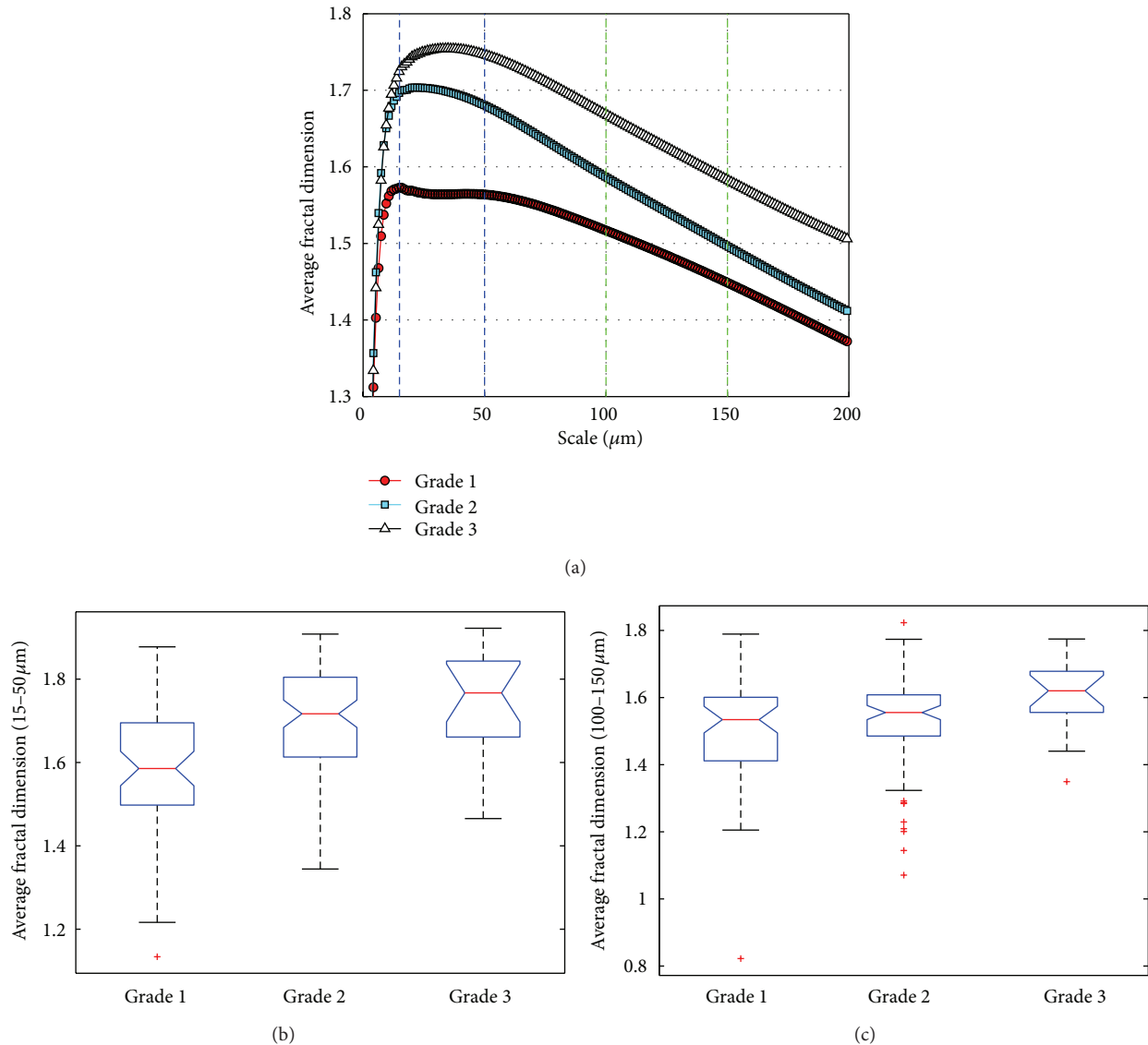


FIGURE 6: (a) Fractal dimensions of breast histology images (averaged over all images in each grade) as a function of the image scale. The two scale intervals (15–50  $\mu\text{m}$  and 100–150  $\mu\text{m}$ ) used as examples are indicated by the vertical dashed lines. (b), (c) Boxplots of fractal dimensions in the scale interval 15–50  $\mu\text{m}$  and 100–150  $\mu\text{m}$ , respectively. See main text for statistical analysis.

## Conflict of Interests

The authors of this paper do not have a financial relation with the commercial identities mentioned in this paper. As such, there is no conflict of interests to report.

## Acknowledgments

This study was supported by a grant from the Alberta Cancer Foundation and Alberta Innovates-Health Solutions.

## References

- [1] B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, New York, NY, USA, 1977.
- [2] D. S. Coffey, "Self-organization, complexity and chaos: the new biology for medicine," *Nature Medicine*, vol. 4, no. 8, pp. 882–885, 1998.
- [3] D. Avnir, O. Biham, D. Lidar, and O. Malcai, "Is the geometry of nature fractal?" *Science*, vol. 279, no. 5347, pp. 39–40, 1998.
- [4] R. Femat, J. Alvarez-Ramirez, and M. Zarazua, "Chaotic behavior from a human biological signal," *Physics Letters A*, vol. 214, no. 3–4, pp. 175–179, 1996.
- [5] A. Espinoza-Valdez, F. C. Ordaz-Salazar, E. Ugalde, and R. Femat, "Analysis of a model for the morphological structure of renal arterial tree: fractal structure," *Journal of Applied Mathematics*, vol. 2013, Article ID 396486, 6 pages, 2013.
- [6] R. Lopes and N. Betrouni, "Fractal and multifractal analysis: a review," *Medical Image Analysis*, vol. 13, no. 4, pp. 634–649, 2009.
- [7] O. Heymans, J. Fissette, P. Vico, S. Blacher, D. Masset, and F. Brouers, "Is fractal geometry useful in medicine and biomedical sciences?" *Medical Hypotheses*, vol. 54, no. 3, pp. 360–366, 2000.

- [8] J. F. Veenland, J. L. Grashuis, F. van der Meer, A. L. D. Beckers, and E. S. Gelsema, "Estimation of fractal dimension in radiographs," *Medical Physics*, vol. 23, no. 4, pp. 585–594, 1996.
- [9] C. Fortin, R. Kumaresan, W. Ohley, and S. Hoefer, "Fractal dimension in the analysis of medical images," *IEEE Engineering in Medicine and Biology Magazine*, vol. 11, no. 2, pp. 65–71, 1992.
- [10] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley & Sons, New York, NY, USA, 2nd edition, 2003.
- [11] M. Ciccotti and F. Mulargia, "Pernicious effect of physical cutoffs in fractal analysis," *Physical Review E*, vol. 65, no. 3, Article ID 037201, 2002.
- [12] H.-W. Chung and H.-J. Chung, "Correspondence re: J. W. Baish and R. K. Jain, Fractals and Cancer. Cancer Research, 60: 3683–3688, 2000," *Cancer Research*, vol. 61, no. 22, pp. 8347–8350, 2001.
- [13] M. Tambasco and A. M. Magliocco, "Relationship between tumor grade and computed architectural complexity in breast cancer specimens," *Human Pathology*, vol. 39, no. 5, pp. 740–746, 2008.
- [14] M. Tambasco, "Scale issue in fractal analysis of histological specimens-reply," *Human Pathology*, vol. 39, no. 12, pp. 1860–1861, 2008.
- [15] M. Tambasco, B. M. Costello, A. Kouznetsov, A. Yau, and A. M. Magliocco, "Quantifying the architectural complexity of microscopic images of histology specimens," *Micron*, vol. 40, no. 4, pp. 486–494, 2009.
- [16] M. Tambasco, M. Eliasziw, and A. M. Magliocco, "Morphologic complexity of epithelial architecture for predicting invasive breast cancer survival," *Journal of Translational Medicine*, vol. 8, no. 1, article 140, 2010.
- [17] G. Landini and J. Paul Rigaut, "A method for estimating the dimension of asymptotic fractal sets," *Bioimaging*, vol. 5, pp. 65–70, 1997.
- [18] J. W. Dollinger, R. Metzler, and T. F. Nonnenmacher, "Bi-asymptotic fractals: fractals between lower and upper bounds," *Journal of Physics A*, vol. 31, no. 16, pp. 3839–3847, 1998.
- [19] R. Alfréd Rényi, "On measures of entropy and information," in *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, vol. 1, pp. 547–561, 1960.
- [20] M. J. Turner, J. M. Blackledge, and P. R. Andrews, *Fractal Geometry in Digital Imaging*, Academic Press, 1998.
- [21] F. C. Crow, "Summed-area tables for texture mapping," *Computer Graphics*, vol. 18, no. 3, pp. 207–212, 1984.
- [22] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Physical Review Letters*, vol. 50, no. 5, pp. 346–349, 1983.
- [23] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829–836, 1979.
- [24] H.-O. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals: New Frontiers of Science*, Springer, New York, NY, USA, 2nd edition, 2004.