Early Detection of Breast Cancer using Self Similar Fractal Method

Bhagwati Charan Patel
Associate Professor (IT)
Shri Shankaracharya College of Engg. & Tech.,
Bhilai, India

Dr. G.R.Sinha
Professor & Head (IT)*)*
Shri Shankaracharya College of Engg. & Tech.,
Bhilai, India

ABSTRACT

Breast cancer is one of the major causes of death among women. Small clusters of micro calcifications appearing as collection of white spots on mammograms show an early warning of breast cancer. Early detection performed on X-ray mammography is the key to improve breast cancer diagnosis. Image segmentation consists in finding the characteristic entities of an image, either by their contours (edges) or by the region they lie in. Our aim in this paper is to present a method for medical image enhancement based on the well established concept of fractal derivatives and selecting image processing techniques like segmentation of an image with self similar properties. The concept of a fractal is most often associated with geometrical objects satisfying two criteria: *self-similarity* and *fractional dimensionality*. The method was tested over several images of image databases taken from BSR APPOLO for cancer research and diagnosis, India.

Keywords

Breast cancer, mammography, segmentation, fractal.

1. INTRODUCTION

Breast cancer is the type of cancer with highest incidence rates in women. It is also the most common cause of cancer death in women in many countries, only exceeded by lung cancer in Asian countries and recently in the United States [1]. The early detection of breast cancer is vital to improve its prognosis. Moreover, it is well known that screening mammography is the best tool available for detecting cancerous before clinical symptoms appear. the shape and arrangement of microcalci-fications help the radiologist to judge the likelihood of cancer being present. Malignant calcifications are typically very numerous, clustered, small, dot-like or elongated, variable in size, shape and density. Benign calcifications are generally larger, more rounded, smaller in number, more diffusely distributed, and more homogeneous in size and shape[2]. In the literature, several techniques have been proposed to detect the presence of micro calcifications using various methodologies. Concerning image segmentation and specification of regions of interest (ROIs), several methods have been proposed such as classical image filtering and local thresholding [3]. Stochastic fractal models [4], wavelet analysis [5,6]. Furthermore, various classification methodologies have been reported for the characterization of ROI such as, fuzzy logic systems [7]. Nevertheless, the most work reported in the literature employs neural networks for cluster characterization [8-10]. In this study, we present system, aiding

radiologist for breast cancer diagnosis and identification of micro calcification clusters in digitized mammographic images. As the micro calcifications correspond to high-frequency components of the image spectrum, detection of micro calcifications is achieved by decomposing the mammograms breast image into sub image by using fractal method.

2. FRACTAL AND FRACTAL DIMENTION

Fractals are of rough or fragmented geometric shape that can be subdivided in parts, each of which is (at least approximately) a reduced similar of the whole. They are creased objects that defy conventional measures, such as length and are most often characterized by their fractal dimension. A fractal dataset is known by its characteristic of being self-similar. The dataset has roughly the same properties for a wide variation in scale or size i.e., parts of any size of the fractal are almost similar to the whole fractal [11]. Intuitively, a set of points which exhibit self similarity over all scales fractals are creased objects that defy conventional measures, such as length and area, and are most often characterized by their fractional dimension [12 and 13].

The fractal dimension is an important characteristic of fractal because it contains information about their geometric structure. Usually we think that the dimension of a point is zero; the dimension of a line is one; the dimension of a surface is two; and the body's is three, and that the dimension of the object wouldn't change whatever the object does any transformation. This kind of dimension called as: topology dimension, defined as: d. Mandelbrot, who introduce the concept of fractal and self-similarity to describe realistic objects, believe that, in the fractal world. The fractal dimension is a statistical quantity which gives the indication of how completely the fractal appears as it is zoomed down to finer and finer scales. Fractal dimension measures the degree of fractal boundary fragmentation or irregularity over multiple scales. It determines how fractal differs from Euclidean objects (point, line, plane, circle etc.). Fractal dimension is an effective measure for complex objects. It is widely applied in the fields of image segmentation and shape recognition. The fractal dimension can be any non-negative real number and can be estimated in several ways [14].

2.1 Fractal dimension and measurement

The following gives a brief explanation of the salient features of the method used for computing the fractal dimensions using Box Counting Algorithm. It incurs only O(N) computational cost [15]. This algorithm derives its name from the imposition of nested hypercube grids over the data, followed by counting the occupancy of each grid cell [16], thus the focus will been the individual points instead of on pairs. By substituting these counts

for the pair counts in the same power law relationship which estimates, the fractal dimension. Given sufficient space to store all the counters simultaneously, all counts can be computed in a single pass over the data. Thus the algorithm provides a greater efficiency.

Considering the solid fractal object only a finite resolution is available so the limit $r \to 0$ cannot be taken into consideration. The breast image is considered as the fractal object. A natural and direct approximation is just to apply directly but with the smallest r available. That is,

Here the concept of Self Similarity across scales is used. The fractal dimension is found to be a measure of roughness and hence is used to model the texture.

$$D = \frac{logN}{\log \left(^1/r\right)} \qquad (1)$$

D – is the fractal dimension
N – number of copies of a self similar set, which has been scaled down by ratio 'r'.
r – scaled ratio of the self similar set.

Instead, log $n(r)$ versus log r is usually plotted for better results. The negative slope of this curve will give Do for small r

The principle for choosing the best slope for the fractal estimation for breast recognition is thus done by which it saves a lot of computational cost and time [17].
Hausdorff [18] suggests one way to generalize the notion of dimension. The idea of generalization involves measuring the same object with different units of measurement. The measure is called the topological dimension of a space. A line segment has a topological dimension of one. In each step, we reduce the image size by r in each spatial direction, thus, its measure will increase to $N = (1/r)^D$ times the original, where r is the magnification factor, N is the number of self-similar pieces, and D_0 is the fractal dimension as given in Eq. 1. The box-counting procedure is mainly defined by two parameters: the selection of r and the range of r. Since a digitized image consists of a finite set of points, we have an upper limit (image size) and lower limit (pixel unit). The box size and the number of boxes counted can only be an integer Some research [19] recommends using 2, 4, 8, 16. . . 2^n pixels as box sizes, r, to give a uniform spread of observations on the independent variable during the log-log least square regression.

3. METHOD
For the image analysis of mammograms, the measure of a region is defined as a function of the gray levels of the points belonging to the region. With the fractal approach, instead of one quantity or measure ω. describing the phenomenon in all scales (as in case of fractals), a set of measures, $\sum \omega_i$ (weight factors) depicting statistically the same phenomenon in different scales, has to be used for characterizing such structures. The signal value within the box or region, is embedded into the process of signal characterization. At the first step, the quantity called roughness exponent δ is derived as
$$\delta = \frac{\log \omega(box)}{\log \epsilon} \qquad (2)$$
where δ quantifies the strength of the singularities of the measure, describing the local regularity of the object, with the determined

measure of the box $\omega(box)$ and size of the box ϵ . The roughness exponent δ corresponds to the fractal dimension of the measure, and it is close to the corresponding fractal dimension of the structure under observation (close to 2 for two-dimensional images). Fractal analyses exploit both local regularity of a given measure and the global distribution of the regularity, being possible to derive a spectrum that characterizes this measure. Therefore, both local and global information of the image regularity can be derived. A single roughness exponent denotes number of self similar fractal, while in the self similar fractal case the different parts of the structure are characterized by different values of δ, leading to the existence of the spectrum [20]

$$f(\delta) = \frac{-logN_\epsilon(\delta)}{\log \epsilon} \qquad , \qquad (3)$$

Where $N_\epsilon(\delta)$ is the number of boxes of size ϵ having the common roughness exponent equal to δ. The value of $f(\delta)$ may be seen as the fractal dimension of the image region that corresponds to a singularity δ. It is also known as the roughness dimension of the distribution of δ. The spectrum is a monotonous decreasing curve with its maximum showing at which δ is positioned the most statistically significant part of the image.

Step 1: Consider a self-similar surface $X \subset R^2$ denoted by a two dimensional array $X(i, j)$, where $i = 1,2,3....M$ and $j = 1,2,3....N$. The surface is partitioned into $Ms \times Ns$ boxes of lateral size s, Each segment can be denoted by $X_{v,w}$, such that

$X_{v,w}(i, j) = X(l1 + i + 1, l2 + j + 1)$ for $1 \leq i, j \leq s$, where $l1 = (v - 1)s$ and $l2 = (w - 1)s$.

Srep2: For each segment $X_{v,w}$ identified by v and w, the cumulative sum $u_{v,w}(i, j)$ is calculated as

$$u_{v,w}(i, j) = \sum_{k1=1}^{i} \sum_{k2=1}^{j} X_{v,w}(l1, l2)$$

where $1 \leq i, j \leq s$ and $u_{v,w}$ is a surface itself

Step 3: Local distance function $F(v, w)$ are calculated for each segment as

$$F(v, w) = \sqrt{\frac{1}{s^2} \sum_{i=1}^{s} \sum_{j=1}^{s} u_{v,w}(i, j) - \tilde{u}_{v,w}(i, j)},$$

where many fitting procedures (*m*-order two-dimensional polynomials) $\tilde{u}_{v,w}(i, j)$ can be used.

Step 4: The overall change in distance is calculated by averaging over all the segments, that is,

$$F_q(s) = \frac{1}{M_s N_s} \sum_{v=1}^{M_s} \sum_{w=1}^{N_s} [F(v, w]^q$$

where q can take any real value except zero.

Step 5: The key property of $Fq(s)$ is that for an image with self-similarity properties, a presence of a power-law scaling is revealed with a linear relationship on a double log plot within a significant range of s .One can relate the roughness exponent δ and the fractal spectrum $f(\delta)$ via Legendre transform [21], deriving these fractal parameters as $f(\delta) = q[\delta - h(q)] + 1$.

4. EXPERIMENTAL RESULT

Image segmentation consists in finding the characteristic entities of an image, either by their contours (edges) or by the region they lie in. The edge detection approach and the region extraction approach often vary much both in algorithms and segmentation results. In the classical methods for edge detection, edges are usually considered to correspond to local extreme of the gradient of the gray levels in the image. The difficulty arising is the computation of the derivative of a most of the time noisy, discrete signal. These methods therefore involve smoothing of the discrete image data and the gradient is then computed by differentiating the smoothed signal [22]. The self similar fractal approach, on the contrary, uses the initial discrete image data directly. Based on the idea that the underlying continuous process might not be possible to recover, the relevant information is extracted directly from the singularities. The advantage is that no information is lost or introduced by the smoothing process, which is an important feature in applications like edge detection, segmentation and texture classification, particularly in medical diagnosis [23],[24]. The drawback is that this approach might be more sensitive to noise. The idea in self similar fractal segmentation is to extract image regions based on particular values of δ and/or f(δ). By appropriate choice of the ordered pair (δ, f(δ)), different features may be recognized, extracted and even classified, both in geometric and probalistic sense. The value of δ will hold information about the local behavior of the measure, The values of f(δ) gives the global information of the image.

The proposed method of extraction has been applied to regions of interest in mammograms from the medical image database. the acquired 8-bit grey images were cropped and various regions of each mammogram were analyzed. The box size ranges varies between 8 and 64, for each 256×256 pixel region. The extraction results obtained and the spectral analysis are illustrated on the following figures.
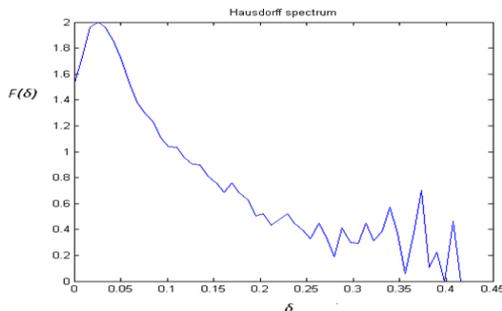
Figure 1. shows the self similar fractal spectrum

Figure 1 shows the self similar fractal spectrums with a cropped image size of 128×128 (square pointed) and 256×256, of a region from the mammogram containing micro calcifications. It can be seen that the roughness spectrum plots approximately fit, specially for $f(\delta) < 0.4$ and $\delta > 0.24$. clearly confirms that micro

calcifications are small light local anomalies. From the multifractal standpoint they are characterized by both high δ and low $f(\delta)$ values, because they represent sharp local changes of contrast and rare events in global sense. This property was verified for all the micro calcifications detected in the analyzed data set. The extractions in Fig. 2(b) were obtained precisely with Hausdorff $f(\delta) < 0.4$. In addition, both crop sizes of this mammogram lead to the same extraction result because of the similar multifractality seen in Fig. 2(b).

(a)   Original image   (b) Extraction of micro-
                                      calcification.
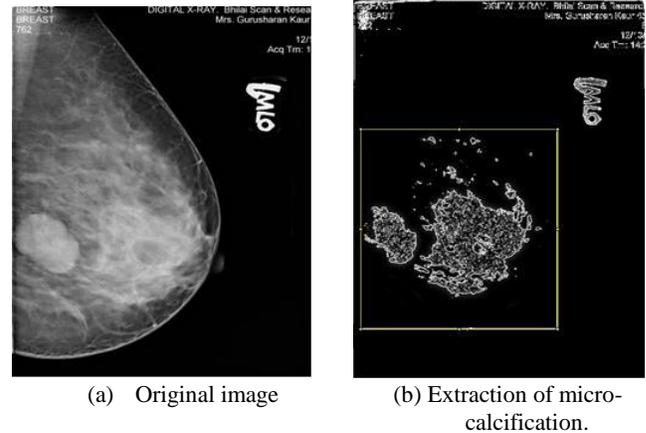
Figure2. Original region from the mammogram (a) and the extraction of micro calcifications with crop sizes (b).

We test the extraction for Hausdorff $f(\delta) < 0.4$ and, once again, we found this to be around the optimal value for the detection of micro calcifications with no significant noise visible, as it is represented in Fig. 3(a).

(a)Detection of micro-calcification   (b) extraction of masses
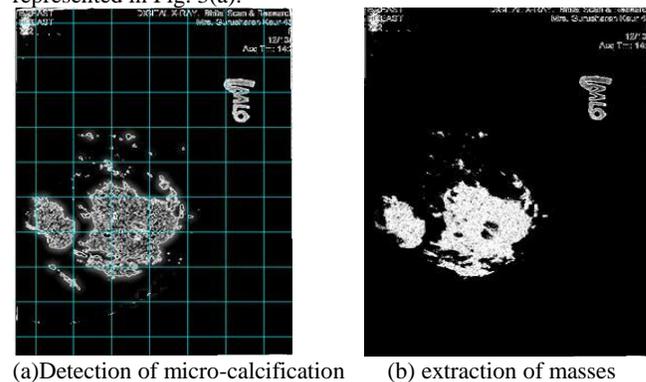
Fig.3 Detection of micro calcification (a) and the extraction of masses (b).

In figure 3(b) all of them extracted with Hausdorff $f(\delta) < 1.1$, which was the value producing the most efficient results. Masses are not completely limited but correctly marked.

(a) Gray level value in image
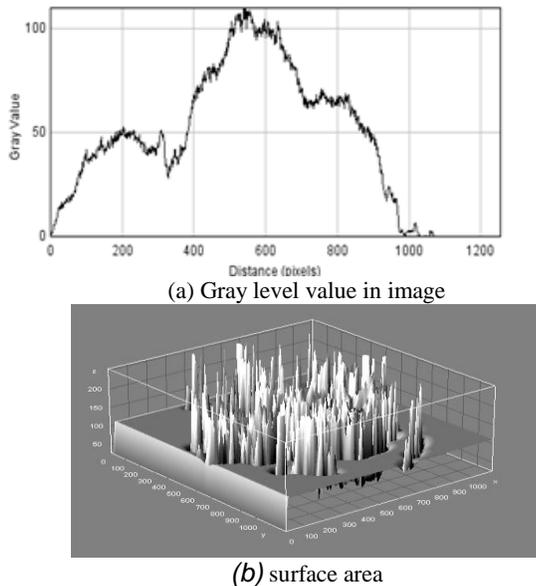
(b) surface area

Figure 4: Gray level value in image and its surface area.

Fig. 4(a) shows the gray level distribution over the area that is affected by micro-calcification and corresponding effective surface covered by **micro**-calcification has been projected in Fig. 4(b). The gray value can be considerably seen in Fig. 4(a) that makes the affected area easily predictable for diagnosis of breast cancer.

5. CONCLUSION AND FUTURE SCOPE

We have presented a method for detection of cancer tumors/ lesions in breast at an early stage by digital mammogram image analysis. The paper deals with detailed results of the automatic detection of breast cancer mass using self similar fractal based segmentation. Different alternatives for the algorithm were tested. The outcomes of these experiments were significantly different from one another. This helps in segmentation of a mammogram image and can be used for developing an expert system for an early detection of breast cancer. This new self similar fractal feature can be applied to many other applications including segmentation of natural scene and detection, segmentation and classification of white blood cells. One possible future work is to develop learning methods to achieve the optimal weight in case that the actual or desired fractal dimension of image is known.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] CancerNet, A service of the National Cancer Institute http://cancernet.nci.nih.gov.

[2] Dengler J, Behrens S, Desage JF.1993. Segmentation of micro-calcifications in mammograms, IEEE Trans Med Image; 12:634-42.

[3] Moti Melloul, Leo Joskowicz, 2002. Segmentation of microcalci fication in xray mammograms using entropy thresholding, Technical Report, Hebrew University, Leibniz Center.

[4] Li H, Liu KJR, Lo SCB, 1997. Fractal modelling and segmentation for the enhancement of micro-calcifications in digital mammograms", IEEE Trans Med Imaging, 16(6):785-98.

[5] Yoshida H, Doi K, Nishikawa RM, Giger ML,Schmidt RA, 1996. An Improved CAD scheme using wavelet transform for detect- Ion of clustered micro-calcifications in digital mammograms, Academic Radiology, 3:621-7.

[6] Lado MJ, Tahoces PG, Mendez AJ, Souto M, Vidal JJ.1999 A wavelet-based algorithm for detecting clustered micro-calcifications in digital mammograms. Med Phys 26(7):1294-305.

[7] B. Verma and J. Zakos, 2001. A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques, Information Technology in biomedicine IEEE 5, pp. 46,54.

[8] P. Sajda and C. Spence. , 2002. Learning Contextual Relationships in Mammograms using a Hierarchical Pyramid Neural Network" *IEEE Transactions on MedicalImaging* 21 (3)

[9] Zheng L, Chan A.2001. An artificial intelligent system for tumor detection in screening mammogram. IEEE Trans Med;20(7):559-67.

[10] Z. R. Yang and R. G. Harrison,2000. Detecting false benign in Breast cancer diagnosis,Neural Networks, IEEE 3, pp. 655,658.

[11] Pentland, A., and Choudhury, T.2000. Face recognition for smart environments". Computer 33(2): 50-55.

[12] Daugman, J., and Downing, C.2001.Epigenetic randomness, complexity, and singularity of human face patterns. Proceedings of the Royal Society, Biological Sciences 268: 1737-1740.

[13] Kronfeld, P.2002 Gross anatomy and embryology of the e ye. In: The Eye (H. Davson, Ed.) Academic Press: London.

[14] Daugman J.2001 Statistical richness of visual phase information: Update on recognizing persons by their retina patterns. International Journal of Computer Vision.

[15] H. S. Greenside, A. Wolf, J. Swift, and T. Pignataro,2002. Impracticality of a box-counting algorithm for calculating the dimensionality of strange attractors, Phys. Rev. 25, 3453.

[16] Simon, A., Worthen, D.M., and Mitas, J.A.2007. An evaluation of iridology. Journal of the American Medical Association 242: 1385-1387.

[17] Guido Gonzato. Dipartimento di Fisica, Settore di Geofisica,2008. "Practical implementation of the box counting algorithm", Università di Bologna, Viale Berti Pichat 8, 40100 Bologna, Italy.

[18] Penn AI, Thompson SF, Schnall SF, Loew MH, Bolinger L (2000). Fractal discrimination of MRI breast masses using multiple segmentations. In: Hanson KM (ed) Medical imaging: Image processing. SPIE 3979: 959–966.

[19] Falconer KJ Hausdorff, 1990. Measure and dimension. In: Fractal geometry mathematical foundations and applications John Wiley, Cambridge.

[20] MandelbrotBB,1983 "The fractal geometry of nature". Freeman, San Francisco.

[21] H. Peitgen, H. J˙urgens, and D. Saupe,1992. "Chaos and Fractals: New Frontiers of Sciences.

[22] Jacques Levy-Vehel and Jean-Paul Berroir.1993. "Image analysis through multifractal description.", Technical Report.

[23] Irini Reljin, Branimir Reljin, Ivan Pavlovi´c, and Ivana Rakoˇcevi´c.2000. Multifractal analysis of gray-scale images. In MEleCon, 490–493.

[24] Irini S. Reljin and Branimir D. Reljin.2002. Fractal geometry and multifractals in analyzing and processing medical data and images. Archive of Oncology, 10(4):283–293.