Research Article

Fractal Dimension as a Diagnostic Tool for Cardiac Diseases

Yagmur Öztürk*†

[†]Pickering College, 16945 Bayview Ave. Newmarket, Ontario, L3Y 4X2, Canada

Received 15 March 2019, Accepted 17 May 2019, Available online 20 May 2019, Vol.9, No.3 (May/June 2019)

Abstract

The discovery of fractal geometry has been one of the major developments in mathematics. Fractals, defined as selfsimilar structures, provide a new approach to the understanding of irregular structures. The dimensions of complete fractals can be easily calculated as real numbers with the fractal geometry approach. However, most of the structures in nature do not demonstrate self-similarity fully, and different approaches are needed for dimension calculations. These structures with semi-fractal properties are known as quasi fractals. Fractals derived from time series are called statistical fractals and they are an example for quasi fractals. Various methods have been developed to estimate the dimensions of statistical fractals. In this study, the methods which predict dimension in statistical fractals were investigated and the most suitable method for time series was identified. Fractal dimensions were calculated by using ECGs of 236 individuals, and the data set was divided into four disease groups and a control group of healthy individuals. Various statistical analyses were performed using the fractal dimensions of the graphs computed with MATLAB. Statistical hypothesis tests showed that the differences between the group mean of fractal dimensions are significant. Fractal dimensions of ECGs have the potential to be used as a diagnostic tool in the diagnosis of heart diseases.

Keywords: Fractals, Statistical fractals, Fractal dimension, ECG, Cardiac diseases.

1. Introduction

Fractals are non-geometric shapes that are self-similar in different scales. With his fractal geometry theory, Mandelbrot was able to create mathematical models for the irregular structures found in nature (Mandelbrot, 1982). His work brought a new perspective for the way mathematicians interpreted physical structures.

A lot of the physiological processes in the human body have fractal structures. Cancer cells can be diagnosed with fractal analysis without detailed examination. Blood flow shows fractal structure as well. Through Electrocardiography (ECG), the electrical activity of the heart is measured by placing electrodes on the patient's skin. Recent studies have shown that heartbeat, shows fractal structure as well (Lapidus, et al, 2004). Classical analysis such as the mean and standard deviation have been proven insufficient, but fractal dimension analysis can be used as a distinguishing factor alongside blood pressure and pulse for check-ups (Acharya, et al, 2005). This paper aims to find an efficient method for fractal dimension calculations of statistical fractals of ECGs. The end-goal is to investigate the possibility of using the fractal

*Corresponding author's ORCID ID: 0000-0002-4767-5692 DOI: https://doi.org/10.14741/ijcet/v.9.3.13 dimension as a diagnostic tool in the diagnosis process of some heart diseases. For this purpose, ECGs of four heart disease groups and a control group of healthy individuals were analyzed. All the data about the patients is taken from MIT Physonet Public Database. Various fractal dimension calculation approaches were examined to determine the most appropriate method that could be used for the ECG data.

2. Earlier work

Mandelbrot's fractal theory suggests that Euclid geometry is not sufficient enough to define various complex structures that exist in the nature. Euclidean theory satisfies an engineer's needs perfectly. However, nature does not always follow simple rules (Gaddis, *et al*, 1986). Benoit B. Mandelbrot came up with the Fractal theory, because "clouds are not spheres, mountains are not cones, coastlines are not circles and bark is not smooth, nor does lightning travel in a straight line..." (Mandelbrot, 1982). Mandlbrot's theory; however, is successful in building mathematical models for natural structures.

In recent years, researches have been carried out on fractal dimension analysis of the signals originating from heartbeat. According to Acharya *et al*, (2005), heartbeat signals with rapid changes in their fractal structures can be a precursor of a disease, and the fractal model is better than the other mathematical methods used to diognose such diseases. Tulppo et al, (2005) proved that there is a relationship between the changes in the fractal structure of heartbeat and some cardiac disorders. Abishekh et al (2013) asserted that in the electroconvulsive treatment of patients, a negative relationship exists between the results of psychiatric evaluation and fractal dimension. Esteban et al (2016), summarizing the applications of fractal dimension in the field of clinical neuroscience, pointed out that fractal dimension contributes significantly to the diagnosis of various nerve diseases. However, fracal dimensions were estimated from the ECG data using image based methods in these studies. As explained in Section 4, these methods are purely geometric notion and not sensitive to time series data such as ECG samples.

3. Background

In Euclidean geometry, dimensions of objects are expressed with integers. A point is zero-dimensional, a line is one-dimensional, a triangle is two-dimensional, and a cube is three-dimensional. Fractal dimensions are rational numbers. A coastline's fractal dimension is a real number between 1 and 2. If we examine a small piece of a fractal structure under magnifying glass, the shape we see will be the same as the bigger picture, because fractal structures repeat themselves in every scale. In other words, they show self-similarity.

Mandelbrot (1982) noticed that there was a relationship between the scale parameter and the number of self-similar copies in fractal structures. When a line segment is repeatedly divided into three pieces to create a shape with four line segments, the scale parameter is s=3, and the number of pieces obtained in the second step is n = 4. This curve is known as the von Koch curve (Fig.1). Each step of a fractal follows the $n=s^d$ relation. n and s will be n = 16and s = 9 in the third step, and n = 64 and s = 27 in the fourth step. In this equation fractal dimension (d), applies to every step of the fractal. If we solve for the second step where n=4 and s=3, $4 = 3^d$, we find that d =log(4)/log(3)=1,26. Therefore, von Koch curve's fractal dimension is d=1.26. Similarly, if the initial line segment is divided into three parts and only.









the middle piece is discarded, a fractal known as the Cantor set is obtained. For this case, n = 2, s = 3 fractal dimension is d = 0.631 since $2 = 3^d$. If we cut a line segment into three, and didn't fold the middle part to form a v shape, n=3, s=3 and $3=3^d$ would give us d=1, which is the topological dimension of the line. This also applies to squares and cubes. If one edge of the square is reduced by $\frac{1}{2}$, then 4 smaller squares are obtained, which would give us $4=2^d$ and d=2. If one side of the cube was reduced at the same scale, n would equal 8, and $8=2^d$ would give us d=3. The topological dimensions of such geometric shapes and their fractal dimensions are the same. However, as seen in the von Koch curve, fractal and topological dimensions differ as they are irregular.

In addition to the von Koch curve, many more fractal samples have been developed. Mandelbroth and Julia sets (Fig.2, upper row) and the Sierpinski triangle are other well-known fractals with exact smilarity properties. These fractals appear identical at different scales. Exact fractals are created through an iterative procces. However, self-similarity is observed only to a certain extent in fractal-like structures seen in nature. Fractals that have an incomplete form of self-similarity are called quasi or approximate fractals. Some other fractals, called statistical fractals, have numerical measures. Dimensions of quasi or statistical fractals can be estimated approximately (Panigrahy et al., 2017). Examples of exact, approximate and statistical fractals are presented in Fig.2.

4. Dimension Calculations of Approximate and Statistical Fractals

Fractal dimension depends on the complexity of fractals. More complex fractals have higher fractal dimensions. For example, von Koch curve is more complex than the Cantor set; therefore, von Koch's

Ozturk

fractal dimension is greater than Cantor's. Similarly, if we looked at the map, we would expect Portugal's coastline to have a smaller fractal dimension than England's coastline, because England's coastline is more indented. In other words, if we didn't have a map, we could say that the coastline that has a greater fractal dimension belongs to England.

Exact fractals' dimensions can be found with more simple methods, as it is explained for von Koch and Cantor Sets. However, we do not know the initial shapes of statistical and natural fractals. Since we don't know their initial steps, we can not make predictions about how these fractal structures reach their final form.

A well-known method called box-counting method and its modifications are used to calculate the dimensions of approximate and statistical fractals. For these fractals, there are other ways of calculating fractal dimension. Box counting method is more suitable for 2 dimensional shapes. It gives fractal dimension through counting the number of boxes the image covers. This method can be used on the twodimensional images of approximate fractal structures (Mandelbrot, 1982, Panigrahy et al, 2017). However, this method's success depends on the resolution of the images. Preliminary analyses showed that ECG graphs are not suitable for box-counting method, as this method fails to capture the differences between different ECG graphs due to the sudden changes in the graph. Box-counting can't distinguish the differences in the pixels smaller than the unit boxes it uses to analyze the image.

For time series, (Katz, 1988) had a different approach on the analysis of fractal structures. Apart from the resolution of the image, he took the total length of the line graph, and the distance between each of the two points (on the sharp ends of the graph) into consideration while coming up with his function, which he used to calculate fractal dimension. Even though the dimensions obtained through Katz method are lesser than the expected values, his method can capture the difference of fractal structures of ECG graphs (Raghavendra *et al*, 2010).

Higuchi method is another method that is frequently used for the analysis of fractal structures in ECGs.This algorithm estimates fractal dimensions more accurately compared to Katz and Box Counting methods (Higuchi, 1988). Therefore, in this study, Higuchi method was used to analyze fractal dimensions of ECG data. This method is suitable for the analysis of fractal structures in time series, and has higher accuracy compared to the other methods.

A time series sample is chosen where $x_1, x_2, ..., x_n$ are recorded at equally spaced time intervals. Subsamples $S_{m,n}$ are obtained from this series as follows:

$$S_{m,n}: x_m, x_{m+k}, x_{m+2k}, ..., x_{m+\lfloor \frac{n-m}{k} \rfloor k}$$
 (1)

where *k* is the time interval. m = 1, 2, ..., k and $\lfloor \cdot \rfloor$ is an operator to round the number to the lower integer. For k = 4 and n = 25 following sub-samples can be created

$$S_{1,4}: x_1, x_5, x_9, x_{13}, x_{17}, x_{21}, x_{25}$$

$$S_{2,4}: x_2, x_6, x_{10}, x_{14}, x_{18}, x_{22}$$

$$S_{3,4}: x_3, x_7, x_{11}, x_{15}, x_{19}, x_{23}$$

$$S_{4,4}: x_4, x_8, x_{12}, x_{16}, x_{20}, x_{24}$$
(2)

The length of the curve corresponding to sub-sample $S_{m,n}$ is defined as

$$L_{m} = c_{m}(|x_{m+1\cdot k} - x_{m+0\cdot k}| + |x_{m+2\cdot k} - x_{m+1\cdot k}| + |x_{m+3\cdot k} - x_{m+2\cdot k}| + \cdots)$$
(3)

where $c_m = (n-1)/\left(\left\lfloor \frac{n-m}{k} \right\rfloor k^2\right)$ and $|\cdot|$ denotes the absolute value operator.

The length of the curve for the time interval k is defined as the average value over k sets of L_m , (m = 1,2, ..., k) that is

$$\bar{L}_{k} = \frac{1}{k} (L_{1} + L_{2} + \dots + L_{k})$$
(4)

For all time intervals k, if $\overline{L}_k \propto k^{-d}$, the curve has a fractal dimension d. To find d, the logarithm of \overline{L}_k is plotted against the logarithm of k and the line of best fit is drawn to find the linear regression. d is the absolute value of the slope of the line of best fit.

5. Material

Using a reliable data set for fractal analysis is a major issue for this study. The data is taken from PTB Diagnostic ECG Database of MIT Computational Physiology Laboratory, which is a public database (Goldberger et *al*, 2000). To download data from this database, and draw graphs of the signal values, WFDB Toolbox application was used (WFDB Toolbox is provided by MIT Computational Physiology Laboratory's Phsio Net organization).

In this database, there are ECGs of 294 patients. These patients are divided into a total number of 9 groups: 8 groups of patients and a control group that consists of 46 healthy individuals. Some of these groups had 7 or less patients. Groups with inadequate number of patients or incomplete records, are excluded from our study. Information regarding the 5 groups is in order to have enough evidence to draw conclusions. As a result, 236 patients' ECGs were analyzed. The number of patients in each group, and their ages are given in Table1.

	Female				Male			
Groups	No. of patient	Min. age	Max. age	Mean age	No. of patient	Min. age	Max.age	Mean age
Control	7	22	67	39.5	39	17	81	45.3
Myocar.Infarc.	39	42	85	64.0	109	26	86	59.0
Cardio-Myop.	7	32	80	59.4	7	52	82	63.0
Bundle Branch	4	56	64	59.7	10	29	76	53.0
Dysrhyt	4	35	75	60.2	10	17	87	47.0

Table 1. Number of patients and their age distribution by gender and disease groups



Fig.3 Fractal dimensions calculated through Higuchi and Box-counting methods vs time intervals of the ECG graph

In an ECG, 1000 signals are recorded per second. For each patient, the signals are standardized to have a mean value of 0. 120-second-long recordings are used. Therefore about 120 000 recorded signals are used for each patient's fractal dimension calculations.



Fig 4. Images used to calculate fractal dimensions at different resolutions. Left: *Lena*, right: ECG graph for the first 30 seconds (patient 232).



Fig.5 ECG graphs for patients in the control and patient groups. Left: Patient from the control group. Right: Patient with *myocardial infarction*.

6. Results

The first step of this work was to find a suitable method to calculate the dimensions of statistical fractals. The first method used to analyze ECGs was the box-counting method. Box-counting method is widely used to calculate the fractal dimensions of exact and approximate fractals. However, after comparing the Higuchi method and box-counting methods by analyzing data with both of them, it was evident that box-counting is not suitable for analyzing ECG graphs. In Fig.3, change in fractal dimension with time is shown for patient 232. It can be inferred from the graph that Box-counting fails to capture fractal dimension, whereas Higuchi's output is consistent after 10s. These results were similar with other patients as well.

Due to their sharp ends, Box-counting can't calculate fractal dimensions of ECG graphs. Image based methods depend on the resolution of the image, and comparisons were made between the fractal dimension results of different sizes of the same picture. For Lena in Fig.4, fractal dimensions for its 512x512, 256x256 and 128x128 resolutions were calculated as d = 1.937, 1.934 and 1.929 respectively. For the ECG graphs like in Fig.4, box-counting calculates their fractal dimensions as d=1.440, 0.993 and 0.454. Thus, fractal dimensions of ECG graphs calculated through Boxcounting method is not accurate since there can't be such a big fractal dimension difference between ECGs' fractal dimensions. Higuchi algorithm; however, is not image-based, and it calculates fractal dimension directly from the data that is already on the graph.



Fig.6 Left: Plots of average curve lengths against time intervals. Right: Plots on the logarithmic scale and the corresponding fitted regression line(patient01).

Fractal dimensions regarding 46 controls and the 190 other patients in the 4 disease groups were calculated. Graphs in Figure 5 are presented to give an idea about the signal values obtained from the database. Graph on

428| International Journal of Current Engineering and Technology, Vol.9, No.3 (May/June 2019)

the left belongs to a person in the control group, and Graph on the right belongs to a patient that is diagnosed with myocardial infarction. These graphs show the signals recorded in 4 seconds.

The Higuchi algorithm in MATLAB was used for finding fractal dimensions of the ECGs. Higuchi algorithm first computes the length of the curve (\bar{L}_k) for the time interval $k=1, 2, ..., k_{max}$, then fits a regression line to the logarithm of these values. Fractal dimension is the absolute value of the slope of this line.

Table 2. Mean and standard deviations of the fractal dimension for control and patient groups

	Control	Myocar- infarc.	Cardio- myop.	Bundle branch	Dysrhy- thmia
Mean	1.696	1.7230	1.662	1.631	1.691
Standard dev.	0.047	0.078	0.106	0.090	0.062

The slope of the regression line is found by using the following formula:

$$m = \frac{\sum_{k=0}^{k_{max}} (y_k - \bar{y})(z_k - \bar{z})}{\sum_{k=0}^{k_{max}} (y_k - \bar{y})^2}$$
(5)

where $y_k = \log(k)$, $z_k = \log(\bar{L}_k)$, $k = 1, 2, ..., k_{max}$. \bar{y} and \bar{z} are the corresponding sample averages. For k_{max} =400 corresponding plots of \bar{L}_k versus k are shown in Fig.6 for patient01. In this graph, scatter plots tend to be linear on the double logarithmic scale. The slope of the regression line is m=-1.776. Fractal dimension is calculated as d=1.776.

In order to be able to make a comparison between the groups by using the Higuchi method, the optimum k value (number of sub-samples) must be determined. In this study, each patient's fractal dimensions for k=50, 100, 150, ...2000 are calculated. The change in fractal dimension according to k parameter of a patient in the control group is given in Fig.7. Fractal dimension increases until k reaches 1000, and the following values are steady. k_{max} =400, where there is a prominent increase in fractal dimension is chosen for comparison.

To compare the groups, fractal dimensions were calculated for each person and k value. For k=100, 200, 300... 1000, average dimensions for each group is shown in Fig.8. In order to calculate fractal dimension with the Higuchi method, k is determined as 400. Graphs based on fractal dimension, and statistical hypothesis tests are conducted through MATLAB. Level of significance (P) was chosen as 0.05 to test the hypotheses.



Fig.7 Graph of fractal dimensions obtained using Higuchi method, with respect to the *k*-parameter. (*patient01*)



Fig.8 Graph of group means of fractal dimensions, against *k*-values.

Fractal dimensions were calculated separately for each patient. Mean and standard deviation of fractal dimension for control and disease groups are given in Table 2. Based on the mean values in Table 2 and Fig.8, some differences are observed between the groups.

Box plots were drawn to explain the differences between the groups (Fig.9). Drawing box plots is a technique to summarize the distribution of the data. The lower side of the box stand for 25th percentile, the line segment in the middle is the median, and the upper side of the box stands for 75th percentile. The lower and upper end point of the dashed line segment represents the minimum and maximum points respectively. The values outside the region between the minimum and maximum points are considered to be extreme values or outliers. It can be inferred from these box plots that the maximum difference between the groups is in between *myocardial infection* and *bundle branch block*. The control group stands in the middle of all 5 groups. Moreover, 9 extreme values are found in the myocardial group. Regarding that these extreme values may be useful for further research, they are not discarded from this graph. t-tests were made to compare the groups in terms of observed variables.

The effect of gender and age on fractal dimension was inspected. Fractal dimension averages and standard deviation for male and female patients were $(\bar{x}_1 = 1.7057, S_1 = 0.0863)$ and $(\bar{x}_2 = 1.7071, S_2 = 0.0773)$ respectively. A two-sample *t*-test indicated that the difference between male and female group averages was not significant (*P*>0.05).

Pearson correlation coefficient (r), was computed to determine the relationship between age and fractal dimension.



Fig.9 Box plots for control and patient groups.

A *t* distribution with *n*-2 degrees of freedom, was used to test if there was a correlation between age and fractal dimension. The *t*-test indicated that there is a positive correlation between age and fractal dimension (*r*=0.171, *n* = 236, *P*<0.01). The scatterplot in Figure 10 summarizes the results. These results indicate a weak correlation between age and fractal dimension.



Fig.10 Scatter plots for fractal dimensions and ages based on 236 patients

Since gender and age do not make a significant difference on fractal dimensions, group comparisons were made without taking these factors into consideration. Then, *t*-tests were performed for group comparisons. Independent sample tests indicated that *Bundle branch block* and *Cardiomyopathy* were different from *Myocardial infarction; Dysrhythmia* was different from *Bundle branch block*, and control group was different from *Myocardial infarction* and *Bundle branch block* (*P*<0.05).

Conclusions

Various irregular structures in nature can be modeled and analyzed more effectively with new approaches offered by fractal geometry. The classical dimension concept defined in the Euclidean geometry was generalized by using the self-similarity of the fractals. Furthermore, it has been found that there is an order in which irregular structures are formed.

The methods developed for calculating the fractal dimension were generally image-based and they were not suitable to accurately calculate the dimensions of the statistical fractals based on ECG data. In this work, we preferred to use the Higuchi method, which takes into account the general characteristics of ECG graphs and calculates the fractal size more precisely for these data.

In general, ECG graphs are evaluated according to scale-based statistics such as minimum, maximum, mean and standard deviation. However, these statistics are insufficient to explain the dynamics that generate the underlying signals. Two patients with the same mean and standard deviation can exhibit different diseases. It has been pointed out in many studies that the fractal dimension plays an important role in understanding of such irregular structures.

In this study, fractal dimensions were calculated by using the test data for four different heart diseases. Based on the statistical analysis it was concluded that the differences between the certain disease groups were statistically significant.

These results suggest that the fractal dimension has the potential to be an important indicator and diagnostic tool that can be used at the first instance, such as pulse and blood pressure in patient examinations.

Acknowledgement

I am especially indebted to Professor Mehmet C. Okur, Department of Software Engineering, Yasar University for his academic guidance and for teaching me about scientific research. I am grateful to Professor Sevgi Mir, Department of Child Nephrology, Ege University for her comments on ECG. I would particularly like to thank Ms. Erin Matthews for her support and encouragement. This work would not have been possible without the test data. I would like to thank Mr. Sermet Onel, assistant professor, Department of Computer Engineering, Yasar University for his help for downloading the data from PTB Diagnostic ECG Database of MIT Computational Physiology Laboratory.

References

- U.R. Acharya, P.S.Bhat, N., Kannathal, A.Rao, C.M.Lim, (2005). Analysis of cardiac Health Using Fractal Dimension and Wavelet Transformation, *ITBM RBM, Elsewier*, Vol.26, No.2, pp.133-139.
- F.J.Esteban, L.Díaz-Beltrán, A.Di Leva, (2016). Fractal Analysis in Neurological Diseases. *The Fractal Geometry of the Brain, Springer Series in Computational Neuroscience, Springer, New York*, pp199-211.
- M.E. Gaddis, J. Michael, M.J. Zyda. (1986) The Fractal Geometry of Nature: Its Mathematical Basis and Application to Computer Graphics. *Research report, Naval Research Arlington, VA*.
- A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, H.E. Stanley, (2000). *PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals, Circulation,* 101, (23), 215-220.

- T. Higuchi, (1988). Approach to an Irregular Time Series on Basis of the Fractal Theory, *Physica D*, Vol.31, pp.277-283.
- M.J. Katz, (1988). Fractals and the Analysis of Wave Forms, *Computers in Biology and Medicine*, Vol.18, No.3, pp.145-156.
- M.L. Lapidus, M. van Frankenhuijsen, (2004). Fractal Geometry and Applications: A Jubilee of Benoit Mandelbrot. *Analysis, Number Theory, and Dynamical Systems, American Mathematical Society, Rhode Island.*
- B.B. Mandelbrot, (1982). The Fractal Geometry of Nature, *W.H. Freeman and Company, New York.*
- C. Panigrahy, A.G. Pedrero, A. Seal, D.R. Esparragón, N.K. Mahato, C.G. Martín, (2017). An Approximated Box Height for Differential-Box-Counting Method to Estimate Fractal Dimensions of Gray-Scale Images, *Entropy*, Vol.19, No.10, pp.534-550.
- B.S. Raghavendra, D.N. Dutt, (2010). Computing Fractal Dimension of Signals Using Multiresolution Box-counting Method. *International Journal of Information and Mathematical Sciences*, Vol.6, pp.50–65.
- M.P. Tulppo, A.M. Kiviniemi, A.J. Hautala, M. Kallio, T. Seppänen, T.H. Mäkikallio, H.V. Huikuri, (2005). Physiological Background of the Loss of Fractal Heart Rate Dynamics, *Circulation*, Vol.112, pp.314-319.