Nonlinear Methods in Heart Rate Variability: Can they Distinguish between Nonpathological and Pathological Subjects?

Stefan Hagmair[1*], Martin Bachler[1,2], Siegfried Wassertheurer[1], Christopher Mayer[1]

[1]AIT Austrian Institute of Technology GmbH, Health & Environment Department, Biomedical Systems, Donau-City- Straße 1,1220 Vienna, Austria; *stefan.hagmair.fl@ait.ac.at

[2]Department of Analysis and Scientific Computing, Vienna University of Technology, Wiedner Haupstraße 8–10, 1040 Vienna, Austria

Abstract. Heart Rate Variability (HRV), i. e., the variation of time intervals between consecutive heart beats, is a marker of the health status, since it unveils changes in beat-to-beat variation of the heart, even before there is a remarkable change in heart rate itself. HRV reflects the balance between the sympathetic and the parasympathetic nervous system. The heart rate itself is non-stationary and the structure generating the signal involves nonlinear contributions. Thus, nonlinear methods to quantify the variability of the heart rate gained interest over the last years. In this work, two nonlinear indices, i. e., Correlation Dimension (CD) and Fractal Dimension (FD), to quantify HRV derived from mathematical models are presented. The implemented methods are tested on their ability to differentiate between healthy and pathological subjects. The databases used for the test are retrieved from PhysioNet. The results show that the FD is able to differentiate between non-pathological and pathological subjects, while the other implemented method, i. e., CD, shows no significant difference. In summary, this paper shows that fractal descriptors are an appropriate support for analyzing the HRV, and therefore help to prevent or detect cardiovascular diseases. Especially Higuchi's Fractal Dimension, well established in the analysis of time series, should get more attention in analyzing biomedical signals, such as HRV.

Introduction

An epidemiological update for the Cardiovascular Diseases (CVD) Statistics in Europe 2015 reports that cardiovascular diseases are the most common cause of mortality in Europe and globally. Furthermore, CVD causes almost two times as many deaths as cancer across Europe [1].

Several papers showed the significant relationship between Autonomic Nervous System (ANS) and cardiovascular mortality [2]. Heart Rate Variability (HRV) analysis is a useful noninvasive tool for understanding and characterizing the status of the ANS. Furthermore, HRV is an early predictor of cardiac failures [3]. The term HRV refers to the variation of beat-to-beat intervals. Usually, HRV is studied in an Electrocardiogram (ECG) by considering the time duration between two R-peaks. Therefore, the beat-to-beat intervals are often named RR-intervals. A schematic representation of an ECG-signal can be seen in Figure 1.

The complex origin of biomedical signal limits the traditional linear approaches [4]. Hence, nonlinear methods gained recent interest to reveal more information embedded in the Heart Rate [5]. In this article two nonlinear indices are presented and compared.

1 Data and Test

1.1 Data

All data used for the tests have been taken from *PhysioNet* [6], a free-access, online archive of physiological signals. PhysioNet guarantees that all data have been fully deidentified (anonymized), and may be used without further institutional review board approval.
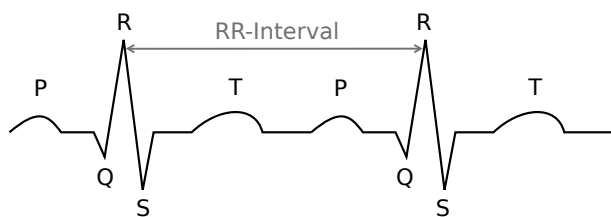
Figure 1: Illustration of an ECG-signal. The beat-to-beat interval (i. e., RR-interval) is defined between two consecutive R-peaks.

The *Normal Sinus Rhythm RR-Interval Database* [6] defines the control group. It is a composition of 54 approximately 24 hours long ECG recordings. The 30 healthy men are 28.5 to 76 years old and the 24 woman range from 58-73 years. Sample frequency of the signals is 128 Hz.

As pathological dataset, the *MIT-BIH Arrhythmia Database* [7] was used. It is a collection of 48 half-hour recordings from 47 subjects (25 men aged 32 to 89 years and 22 woman aged 23 to 89 years). 25 examples of uncommon but clinically important arrhythmias are included in this database and the remaining recordings include different pathologies. The digitization rate is 360 Hz.

Before calculation of HRV indices the data were processed. The first step of preprocessing was to remove all intervals longer than two seconds. This threshold for the deletion of very long intervals is reported in [8]. Afterwards data was clustered in order to perform the calculations only on the sinus beat cluster. The clustering method is described in detail in [9, 10]. The required parameters were set to a number of iterations $r = 20$ and the minimum number of points in one cluster $MinPts = 4$.

From all of the recordings, the middle 1500 RR-intervals were taken. After processing the data, measures were calculated on the first 1024 RR-intervals. The last four recordings, i. e., 231-234 were excluded from the pathological database, since after clustering less than 1024 RR-intervals were available.

	Nonpathological	**Pathological**
records (female)	54 (24)	48 (22)
N_{min}	77373	1512

Table 1: PhysioNet databases. N_{min} is the minimal number of RR-intervals per database.

1.2 Test

To test if the models can differentiate between healthy and unhealthy subjects, indices of the pathological database were tested against the calculated indices of the nonpathological database. Since the results follow a normal distribution, the t-test was used to determine the difference between the data sets. The test decision for a normal distribution was obtained with the Lilliefors test. All the calculations were done in MATLAB®.

A test with a p-value $p < 0.05$ was called significant. If the p-value is smaller than 0.01, the test outcome was declared as very significant.

2 Methods

Already in the 1980s, it was indicated that beat generation in a human heart is a chaotic process [11]. As the RR-time series is not constant over time, and irregular, it cannot be completely described by a linear approach [12]. Thus, there is an increasing interest in nonlinear methods and methods from chaos theory recently [13].

For all the algorithms in this section, the set of RR-intervals $\{RR_1, \ldots, RR_N\}$ of length N is denoted as

$$X := \{x(1), \ldots, x(N)\} = \{RR_1, \ldots, RR_N\}. \quad (1)$$

2.1 Correlation Dimension

The first nonlinear measure, the *Correlation Dimension (CD)* assesses the fractal dimension of the system attractor in the reconstructed phase space, i. e., measures the complexity of the chaotic system [14].

The CD index is calculated based on an algorithm proposed by Grassberger and Procaccia [15]. In the first step, a phase space reconstruction is created by

$$Y_i = \{x(i), x(i+\tau), \ldots, x(i+(m-1)\tau)\}, \\ i = 1, \ldots, N - (m-1)\tau. \quad (2)$$

Hereby, m is the embedding dimension and τ is the delay time. The concatenation Y is a $M \times m$ matrix with $M := N - (m-1)\tau$ row vectors Y_i.

The CD is obtained by considering correlations between points on this attractor. Therefore, the correlation integral is defined as

$$C(r) := \frac{2}{M(M-1)} \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \Theta\left(r - \|Y_i - Y_j\|\right), \quad (3)$$

where r is the radius and Θ denotes the Heaviside function

$$\Theta(x) := \begin{cases} 0, & x \leq 0 \\ 1, & x > 0. \end{cases} \quad (4)$$

$\left\| Y_i - Y_j \right\|$ is the Euclidean distance between a pair of points within the attractor. Hence, the correlation function counts the number of distances closer than r, over the total number of distances. Obviously $C(r)$ has values between 0 and 1. If M is sufficiently large and r is small enough, the following definition holds:

$$CD := \lim_{r \to 0} \frac{\log C(r)}{\log r}. \quad (5)$$

In this work, the delay time and the embedding dimension are chosen as $\tau = 1$ and $m = 22$, according to [16]. The last step is the design of the scaling region for the power law behavior. The approach of [16] is the following: Obtain r_{max} at $C(r_{max}) = 0.1$ and $r_{min} = r_0 + 0.25(r_{max} - r_0)$, where r_0 is the smallest radius. Since most of the subjects had nearly the same scaling region, parameters for r_{max} and r_{min}, respectively, were fixed as $r_{max} = 0.15$ and $r_{min} = 0.06$. In Figure 2, the plot of the scaling region and the linear fit for a healthy subject is shown.
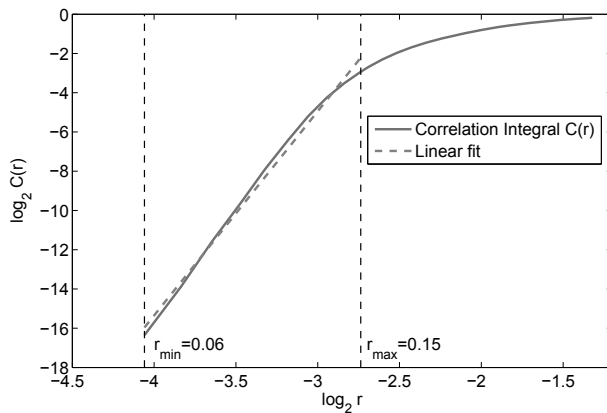

Figure 2: Correlation Dimension of a nonpathological subject. The fitting region lies between the black dashed lines.

2.2 Fractal Dimension

The term fractal is used for objects exhibiting a repeating pattern that displays at every scale. Famous examples of fractals are the Koch (Niels Fabian Helge von Koch (January 25, 1870 – March 11, 1924)) curve (also

known as Koch snowflake), or the Sierpińksi (Wacław Franciszek Sierpiński (March 14, 1882 – October 21, 1969)) Triangle. For ordinary geometric objects, the Fractal Dimension (FD) equals the well known Euclidean or topological dimension, i. e., lines are one-dimensional, surfaces have two dimensions and solids are three-dimensional. But the fractal dimension has not to be an integer.

To understand the need of the FD, the Koch curve is considered. Its topological dimension is one, but the length of the curve between two arbitrary points is infinite. One can assume that the curve can be explained as a fractal line, but it is too simple to be two-dimensional [17]. The FD of the Koch curve is defined as $FD := \frac{\log(4)}{\log(3)} \approx 1.26$, since in each step of scaling by $1/3$ the number of new sticks is equal to four. Indeed, the FD of the Koch curve lies between one and two. The first two iterations of the Koch curve are plotted in Figure 3.

Figure 3: The first two iterations of the Koch curve. In each iteration the lines are divided into three equal parts. The middle part is duplicated and then arranged like a triangle.

FD is calculated directly from the analyzed signal and has been applied to different biomedical signals, such as Electroencephalography (EEG) and ECG recordings [18].

The following algorithm was proposed by Higuchi in 1988 [19]:

1. For a fixed $k \in \{1, \ldots, k_{\max}\}$, construct k new time series

$$x_m^k = \{x(m), x(m+k), x(m+2k), \ldots, x(m+\lfloor a \rfloor k)\}$$
$$m = 1, \ldots, k, \quad (6)$$

where $\lfloor a \rfloor := \left\lfloor \frac{N-m}{k} \right\rfloor$.

2. Compute the length of each time series

$$L_m(k) = \sum_{i=1}^{\lfloor a \rfloor} |x(m+ik) - x(m+(i-1)k)|. \quad (7)$$

3. Normalize the lengths for each k

$$\widetilde{L_m}(k) = L_m(k)\frac{N-1}{\lfloor a \rfloor k}. \qquad (8)$$

4. Calculate the average length

$$L(k) := \text{mean}_m\left(\widetilde{L_m}(k)\right). \qquad (9)$$

Theses steps are repeated for $k = 1, \ldots, k_{\max}$.

The FD is defined as the slope of the linear fit (in a least squares sense) of the plot $1/\log(L(k))$ versus $\log(k)$. An example can be seen in Figure 4.

The implemented method is an adaption of Higuchi's algorithm with a moving time window of fixed window length. Hence, this method is sometimes specified as *running FD* [20].

According to [21], the parameter $k_{\max}$ was set to $k_{\max} = 10$ and the RR-intervals were divided into windows with a length of 100 samples, shifted by one sample. The FD of one entire record was obtained by averaging all FDs of all windows.
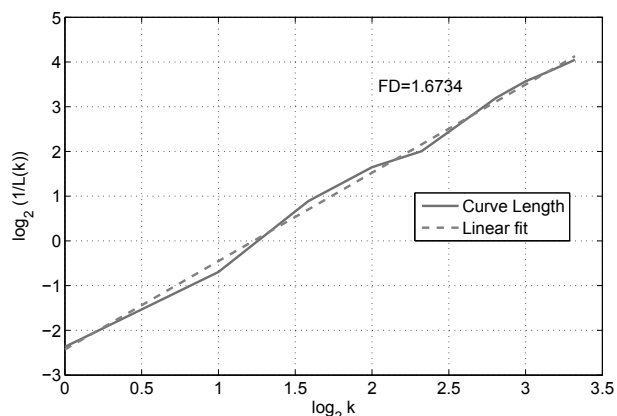

Figure 4: The Fractal Dimension calculated on a window of 100 samples of a nonpathological subject. $k_{\max} = 10$

3 Results

In table 2, the distribution parameters of both indices are presented. Since the calculated values of the indices are normally distributed, the mean value, standard deviation and the corresponding p-value are listed.

Furthermore, Figure 5 shows boxplots of the measures calculated from the nonpathological and the pathological database, respectively. The left boxplot shows the CD with no significant difference between both groups, whereas the FD shows very significant differences between the tested groups.

	Nonpath. Data	Path. Data	p-Value
CD	7.25 (2.42 SD)	8.22 (4.63 SD)	0.26
FD	1.68 (0.14 SD)	1.95 (0.09 SD)	<0.01

Table 2: Statistical parameters of the Correlation Dimension *CD* and Fractal Dimension *FD* for the test. Mean and Standard Deviation (SD) are listed since data follows a normal distribution.

4 Discussion

The computed values of the CD are similar to the reported ones in [16]. Mia et al. characterized the HRV by means of the CD for two groups, i. e., healthy subjects and hypertension patients and distinguish between values during the day and at night. Their outcome was that healthy subjects had a lower CD in daytime than at night. This relationship was not seen in hypertensive patients [16]. Since the recording time was not specified for the pathological data and for the nonpathological database, this may have affected the results. The wide scatter indicates this effect.

Acharya et al. reported in [22] that the index CD varies for different kind of cardiac abnormalities. E. g., the CD calculated of subjects with atrial fibrillation is rather high compared to the CD value of patients where the heart rate remains stable (e. g., complete heart block).

Higuchi's Fractal dimension is very close to 2 for the pathological dataset. These values are also reported in [23]. The FD of healthy subjects is significantly smaller. The higher values of FD may result during ectopic beat variation in the heart rate.

4.1 Limitations

The outcome of this work is mainly bounded by two limitations. First, the composition of the pathological database is not as homogeneous as it should be. Precisely, subjects with different pathologies, including people with and without arrhythmias, are included.

Second, the absence of the recording time may yield comparing indices calculated on different day times.
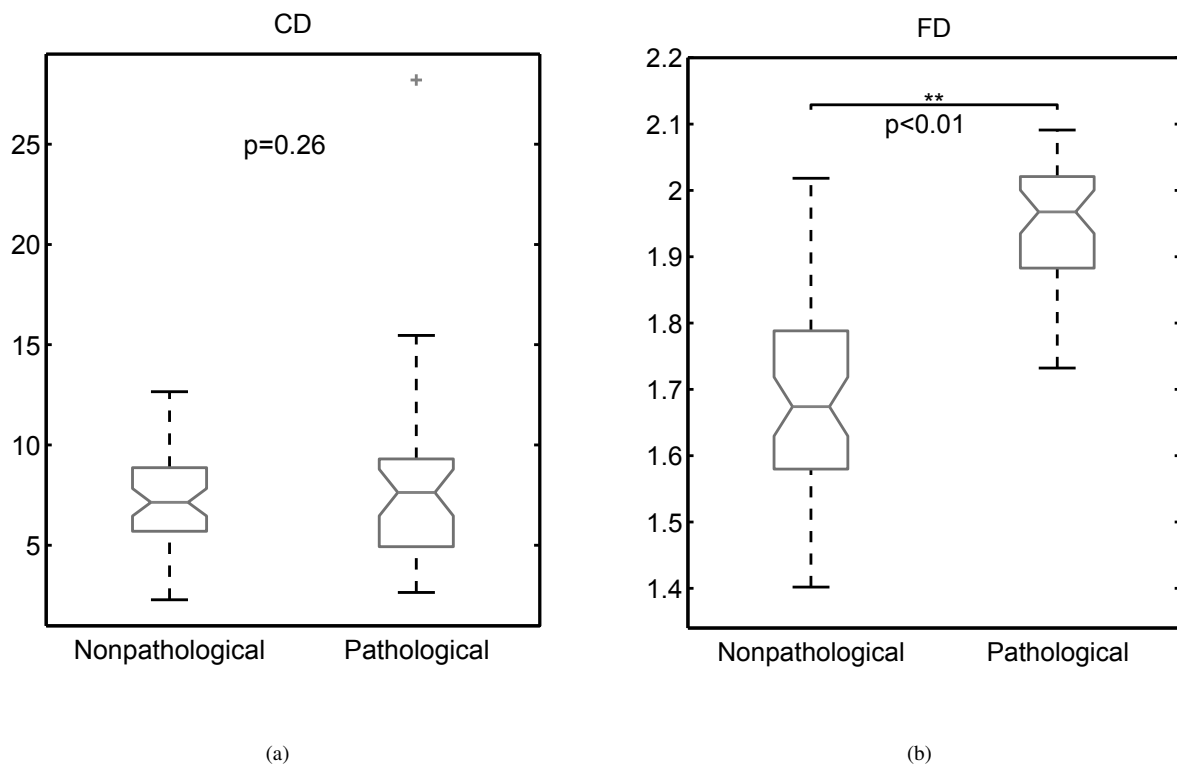
Figure 5: Boxplots for the test of the (a) Correlation Dimension index *CD*, (b) Fractal Dimension *FD*.

5 Conclusion and Outlook

Higuchi's FD is a fast and useful method to differentiate between nonpathological and pathological subjects. The CD was not able to asses, if a subject is healthy or not. Further examination of the design of the scaling region and a separation of day- and nighttime is suggested.

References

[1] Townsend N, Nichols M, Scarborough P, Rayner M. Cardiovascular disease in Europe—epidemiological update 2015. *European heart journal*. 2015;p. ehv428.

[2] Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. *Medical & biological engineering & computing*. 2006; 44(12):1031–51.

[3] Cripps TR, Malik M, Farrell TG. Prognostic value of reduced heart rate variability after myocardial infarction: clinical evaluation of a new analysis method. *British heart journal*. 1991;65(1):14–19.

[4] Signorini M, Ferrario M, Marchetti M, Marseglia A. Nonlinear analysis of heart rate variability signal for the characterization of cardiac heart failure patients. In: *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*. IEEE. 2006; pp. 3431–3434.

[5] Akay M. *Nonlinear Biomedical Signal Processing, Dynamic Analysis and Modeling*, vol. I. New York, NY, USA: Wiley-IEEE Press. 2001.

[6] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*. 2000; 101(23):e215—-e220.

[7] Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*. 2001;20(3):45–50.

[8] Bauer A, Malik M, Schmidt G, Barthel P, Bonnemeier H, Cygankiewicz I, Guzik P, Lombardi F, Müller A, Oto A, Schneider R, Watanabe M, Wichterle D, Zareba W. Heart Rate Turbulence: Standards of Measurement, Physiological Interpretation, and Clinical Use.

International Society for Holter and Noninvasive Electrophysiology Consensus. *Journal of the American College of Cardiology*. 2008;52(17):1353–1365.

[9] Xia L, Jing J. An Ensemble Density-based Clustering Method. *Proceedings on Intelligent Systems and Knowledge Engineering (ISKE2007)*. 2007;.

[10] Hörtenhuber M, Bachler M, Wassertheurer S, Mayer C. Comparison of Two Filtering Methods for Heart Rate Variability Analysis. *SNE Simulation Notes Europe*. 2014;24(3–4):137–142.

[11] Michaels DC, Chialvo DR, Matyas EP, Jalife J. Chaotic activity in a mathematical model of the vagally driven sinoatrial node. *Circulation research*. 1989; 65(5):1350–1360.

[12] Lombardi F. Chaos theory, heart rate variability, and arrhythmic mortality. *Circulation*. 2000;101(1):8–10.

[13] Huikuri HV, Mäkikallio TH, Perkiömäki J. Measurement of Heart Rate Variability by Methods Based on Nonlinear Dynamics. *Journal of Electrocardiology*. 2003;36(SUPPL.):95–99.

[14] Lombardi F, Sandrone G, Mortara A, Torzillo D, Rovere MTL, Signorini MG, Cerutti S, Malliani A. Linear and nonlinear dynamics of heart rate variability after acute myocardial infarction with normal and reduced left ventricular ejection fraction. *The American Journal of Cardiology*. 1996;77(15):1283–1288.

[15] Grassberger P, Procaccia I. Characterization of strange attractors. *Physical review letters*. 1983;50(5):346–349.

[16] Miao X, He W, Yang H, Tai HM. Heart rate variability characterization using correlation dimension. In:

Circuits and Systems, 2002. MWSCAS-2002. The 2002 45th Midwest Symposium on, vol. 1. IEEE. 2002; pp. I–447.

[17] Harte D. *Multifractals, Theory and Applications*. CRC Press. 2001.

[18] Esteller R, Member S, Vachtsevanos G, Member S, Echauz J, Litt B. A Comparison of Waveform Fractal Dimension Algorithms. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*. 2001;48(2):177–183.

[19] Higuchi T. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Nonlinear Phenomena*. 1988;31(2):277–283.

[20] Klonowski W. Application of nonlinear dynamics in biosignal analysis. *Proceedings of SPIE - The International Society for Optical Engineering*. 2005; 5975.

[21] Pierzchalski M, Stepien RA, Stepien P. New Nonlinear Methods of Heart Rate Variability Analysis in Diagnostics of Atrial Fibrillation. *International Journal of Biology and Biomedical Engineering*. 2011; 5(4):201–208.

[22] Acharya R, Lim C, Joseph P. Heart rate variability analysis using correlation dimension and detrended fluctuation analysis. *Itmb-Rbm*. 2002;23:333–339.

[23] Pierzchalski M. Assessing cardiovascular reaction to antiarrhythmic drugs using Higuchi's fractal dimension. In: *Assessing cardiovascular reaction to antiarrhythmic drugs using Higuchi's fractal dimension*. 2011; pp. 200–202.