Global Characterization of Time Series Using Fractal Dimension of Corresponding Recurrence Plots: From Dynamical Systems to Heart Physiology

P. Babinec, M. Kučera and M. Babincová

Department of Nuclear Physics and Biophysics, Faculty of Mathematics, Physics, and Informatics, Comenius University, Mlynská dolina F1, 842 48 Bratislava, Slovakia; email: babinec@fmph.uniba.sk

Abstract

Novel method for the global characterization of time series, based on the calculation of fractal dimension of a two-dimensional recurrence plots is proposed. The method is used for the characterization of differences between regular and chaotic systems and for the analysis of human electrocardiogram.

1. Introduction

There is a long history of using image analysis to determine the morphology of a system associated with an image [1]. Recently, techniques based on concepts from the field of artificial life have been used for image analysis [2]. The measures of complexity that they use involve fractal dimension and percolation.

The concept of a fractal dimension to describe structures, which look the same at all length scales, was first proposed by Mandelbrot [3]. Although in strict terms, this is a purely mathematical concept, there are many examples in nature that closely approximate a fractal object, though only over particular ranges of scale. Such objects are usually referred to as *self-similar* to indicate their scale-invariant structure. In simple terms, the common characteristic of such fractal objects is that their length (if the object is a curve, otherwise it could be the area or volume) depends on the length scale used to measure it, and the fractal dimension tells us the precise nature of this dependence.

Our aim in this study is to apply the concepts of fractals to global characterization of time series through the fractal structure of their two-dimensional images – recurrence plots.

2. Visual recurrence analysis method

Recurrence Plots (RPs) are relatively new technique for the qualitative assessment of time series [4]. With RP, one can graphically detect hidden patterns and structural changes in data or see similarities in patterns across the time series under study. The fundamental assumption underlying the idea of the recurrence plots is that an observable time series is the realization of some dynamical process, the interaction of the relevant variables over the time. Because the effect of all the other (unobserved) variables is already reflected in the series of the observed input, one can recreate a topologically equivalent picture of the original multidimensional system behaviour by using the time series of a single observable variable [5]. Furthermore, the rules that govern the behaviour of the original system can be recovered from its output.

Recurrence plots are intricate and visually appealing [6]. They are useful for finding hidden correlations in highly complicated data. Because they make no demands on the stationarity of a data set, RPs are particularly useful in the analysis of systems whose dynamics may be changing. The use of recurrence plots in time-series analysis has become more common in recent years, particularly in the area of physiology, for instance, they have been used to discern between “quiet“ and “active“ breathing in laboratory rats [7] or to study neuronal spike trains in cats [8]. RPs have been also used in mathematical problems primarily to identify transition points in non-stationary data sets [9], and in the area of molecular dynamics simulations as a tool for the detection of transients and both linear and nonlinear state changes.

An RP is a two-dimensional representation of a single trajectory. The time series spans both ordinate and abscissa and each point (i, j) on the plane is shaded according to the distance between the two corresponding trajectory points y_i and y_j. In an unthresholded RP (UTRP) the pixel lying at (i, j) is color-coded according to the distance, while in a thresholded RP (TRP) the pixel lying at (i, j) is black if the distance falls within a specified threshold corridor and white otherwise. RPs are symmetrical along the $x = y$ axis, where each point is plotted against itself, and this diagonal roughly represents time [10]. Figure 1 shows UTRPs generated from four different data sets: (a) a time series electrocardiogram, (b) a time series of a Brownian motion, (c) a time series of a Dow Jones index. The colors on these plots range from white for very small spacing to dark blue for large inter-point distances, as shown on the calibration bars in the figure. With this in mind, the sine-wave RP is relatively easy to understand; each of the "blocks" of colour simply represents half a period of the signal. The RP generated from a chaotic data set is far more complicated, although it too has block-like structures resembling to what might be expected from a periodic signal. For random signal, the uniform (even) distribution of colours over the entire RP is expected and the colours on the UTRP for the time sequence "deepen" away from the main diagonal.

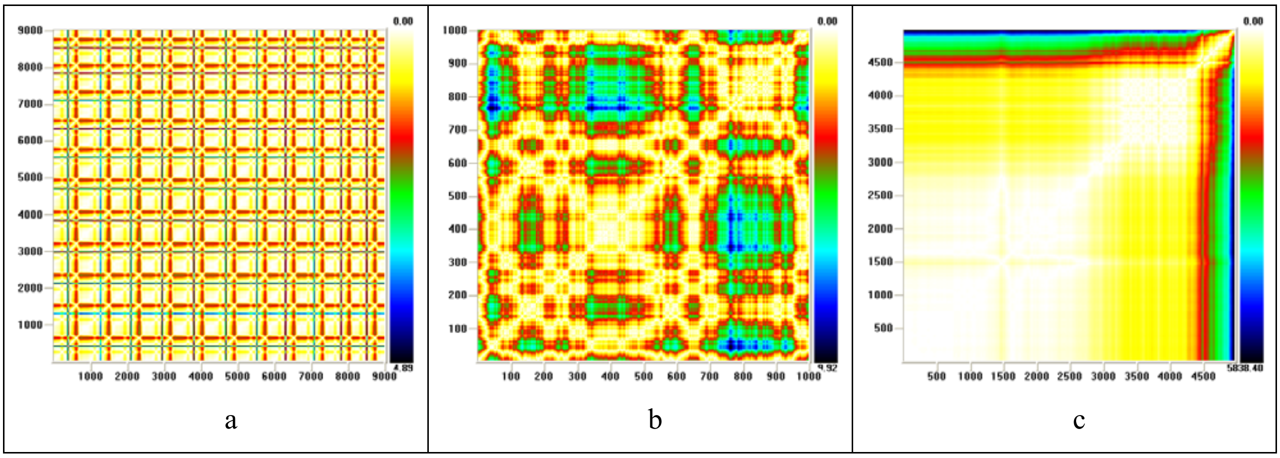


Figure 1 Recurrence plots of (a) a time series electrocardiogram, (b) a time series of a Brownian motion, (c) a time series of a Dow Jones index. On the right side of each recurrence plot is shown calibration bar showing its colour range.

The basic idea behind the interpretation of the RPs is simple: if the underlying signal is truly random and has no structure, the distribution of colours over the RP will be uniform, and so there will not be any identifiable patterns. If, on the other hand, there is some determinism in the signal generator, it can be detected by some characteristic, distinct distribution of colours. The main advantage of the recurrence plots over another widely used techniques as for example Fourier analysis, is that they preserve both temporal and spatial dependence in the time series. Even though Fourier analysis reveals the distribution of spectral frequencies, it does not show how self-similar, resonant frequencies are patterned as a function of time. Yet, RP is mostly a qualitative tool and the precise meaning of the patterns is unknown.

Recurrence plots of the obtained time series from the DRP simulations we performed, were created with the program *Visual Recurrence Analysis* (VRA) provided by E. Kononov [11]. In VRA, a one-dimensional time series from a data file is expanded into a higher-dimensional space, in which the dynamic of the underlying generator takes place. This is done using a technique called "delayed coordinate embedding", which recreates a phase space portrait of the dynamical system under study from a single (scalar) time series. To expand a one-dimensional signal into an m-dimensional phase space, one substitutes each observation in the original signal X(t) with vector

$$y(i) = \{x(i), x(i - T), x(i - 2T), ..., x(i - (d - 1)T)\} \quad (1)$$

where i is the time index, d is the embedding dimension and T represents the time delay. As a result, we have a series of vectors

$$Y = [y(1), y(2), y(3), ..., y(N - (d - 1)T)] \quad (2)$$

where N is the length of the original series. Using $T = 1$ merely returns the original time series; one-dimensional embedding is equivalent to no embedding at all. Proper choice of the time delay and the embedding dimension is said to be critical to this type of phase space reconstruction. Only correct values of these two parameters yield embeddings that are guaranteed to be topologically equivalent to the original (unobserved) phase-space dynamics [5].

Once the dynamical system is reconstructed, a recurrence plot can be used to show which vectors in the reconstructed or original space are close and far from each other. VRA calculates the Euclidean distances between all pairs of vectors and codes them as colors.

Essentially, UTRP is a color-coded matrix D, where each [i][j]th entry is calculated as the distance between vectors $Y(i)$ and $Y(j)$ in the reconstructed series

$$D_{ij} = \sqrt{(x(i) - x(j))^2 + (x(i-d) - x(j-d))^2 + ... + (x(i-(m-1)d) - x(j-m-1)d)^2} \quad (3)$$

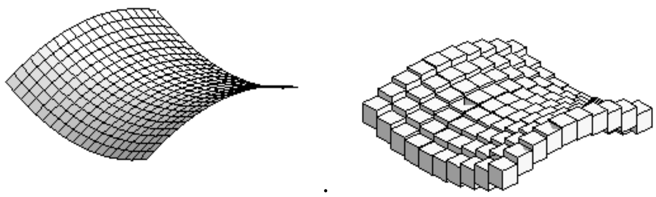
in the case of $d = 1$

$$D_{ij} = |x(i) - x(j)| \quad (4)$$

After the distances between all vectors are calculated, they are mapped to colors from the predefined colour map and are displayed as coloured pixels in their corresponding places.

3. Calculation of fractal dimension

The images were analysed using program *HarFA 5.1* provided by O. Zmeskal [12] based on the improved box counting method where binary images were covered with different grids (box length ε), and the number of boxes $N(\varepsilon)$ required to cover the structures of the images is recorded



If an object is fractal, $N(\varepsilon)$ increases according to the relation

$$N(\varepsilon) = C\varepsilon^D \quad (5)$$

where D is fractal dimension and C is a constant. From this equation the fractal dimension can be obtained as

$$D = \lim_{\varepsilon \to 0} \{-\log(N(\varepsilon)) / \log(\varepsilon)\} . \quad (6)$$

The HarFA code is based on counting of squares (black, white, and partially black) from a squared network behind the fractal figure. The difference between calculated and exact values of fractal dimensions obtained using HarFA is very small (e.g. for Sierpinsky triangle the error is less than 0.2 %).

4. Comparison of regular and chaotic systems

Determinism in the mathematical sense means that there exists an autonomous dynamical system, defined by a first order ordinary differential equation $\dot{\boldsymbol{x}} = f(\boldsymbol{x})$ in a state space $\Gamma \subset \mathrm{R}^d$, which we assume to be observed through a single measurable quantity $h(x)$. The system thus possesses d natural variables, but the measurement is usually nonlinear projection onto a scalar value. Deterministic chaos offers an interesting explanation for the emergence of aperiodicity and unpredictability. Since rather simple systems exhibit chaos, one is lead to use nonlinear time series methods to verify whether such source of unpredictability is underlying a given observation. In fact, the concept of deterministic low-dimensional chaos has proven to be fruitful in the understanding of many complex phenomena despite the fact that very few natural systems have actually been found to be low-dimensional deterministic in the sense of the theory. Deterministic chaos is not the only, and not even the most probable source of aperiodicity. The superposition of a large number of active degrees of freedom can produce extremely complicated signals, which may not be distinguishable from randomness. Stochasticity in the sense that a system is driven by processes whose dynamics are too complex to be inferred from the information stored in the observations is the most frequent source of unpredictability in open systems and field measurements.

To investigate the capabilities of this method we have used three time series, sine function – prototype of regular deterministic signal, the same signal slightly perturbed with white noise, and white noise itself. In Figure 2 are shown the series together with their recurrence plots produced using VRA program. These plots (images without the frames, bars and captions, of course) were then analyzed using HarFA program. The obtained fractal dimensions are shown in Figure 3. As can be seen the fractal dimensions in this case correlate with the complexity of the signals.

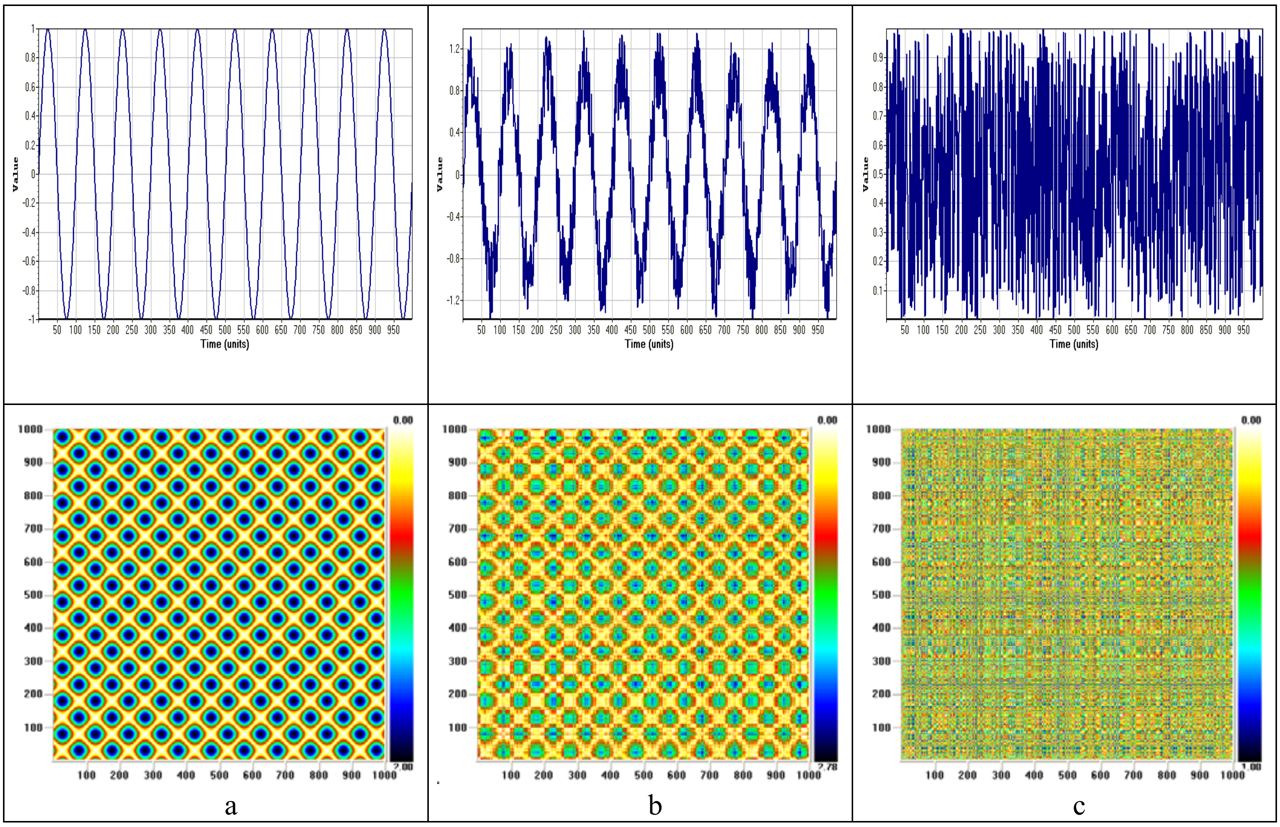


Figure 2 Time series and recurrence plots of (a) time series derived by sampling the function sin*(t), (b) a time series of a sine perturbed with white noise signal, (c) a time series of a random signal (white noise). On the right side of each recurrence plot is shown calibration bar showing its colour range.*

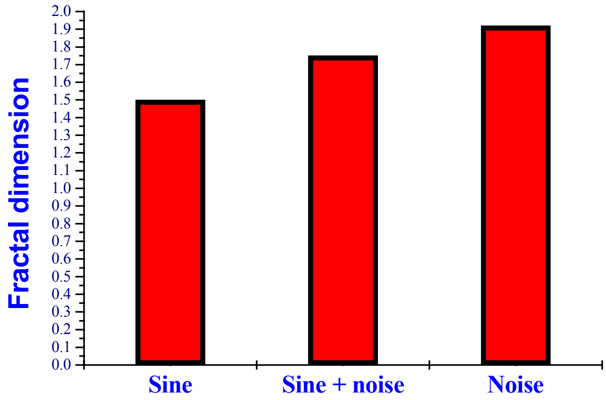


Figure 3 Fractal dimensions of a recurrence plots of the corresponding time series.

5. Applications in heart physiology

In the mid-1880's Ludwig and Waller discovered that the electrical activity of the heart could be monitored through the skin. Their device, called a capillary electrometer, used sensor electrodes and magnets to generate an electrical field. A capillary tube with fluid was placed in the field. As current passed through the electrodes, the field increased and decreased causing the fluid in the tube to bounce up and down. This device, as cool as they probably thought it was, was far too crude for clinical use. Einthoven devised a clever system for recording the same electrical activity on light-sensitive paper. Noticing a recurring pattern of movement, Einthoven named the prominent waves alphabetically, P, Q, R, S, and T the P-wave, representing the impulse across the atria to the A/V node; The QRS representing the impulse as it travels across the ventricles; the T-wave, representing the repolarisation of the ventricles (Figure 4).

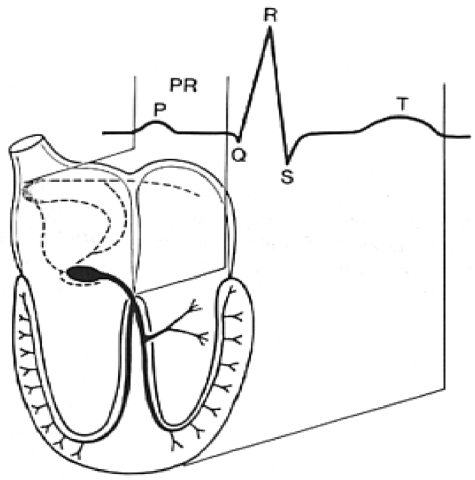


Figure 4 Anatomy of the heart with assignment of P, Q, R, S, T, and P waves.

Inter-beat intervals of a two groups of healthy subjects, young (mean age 27 yr.) as well as 5 old (mean age 74 yr.) were analysed in our recent study [13]. A. L. Goldberger, Harvard University, made these data available. Subjects lay supine for 120 min while continuous ECG signals were collected. All subjects remained in an inactive state in sinus rhythm while watching the movie ”Fantasia” (Disney) to help maintain wakefulness.

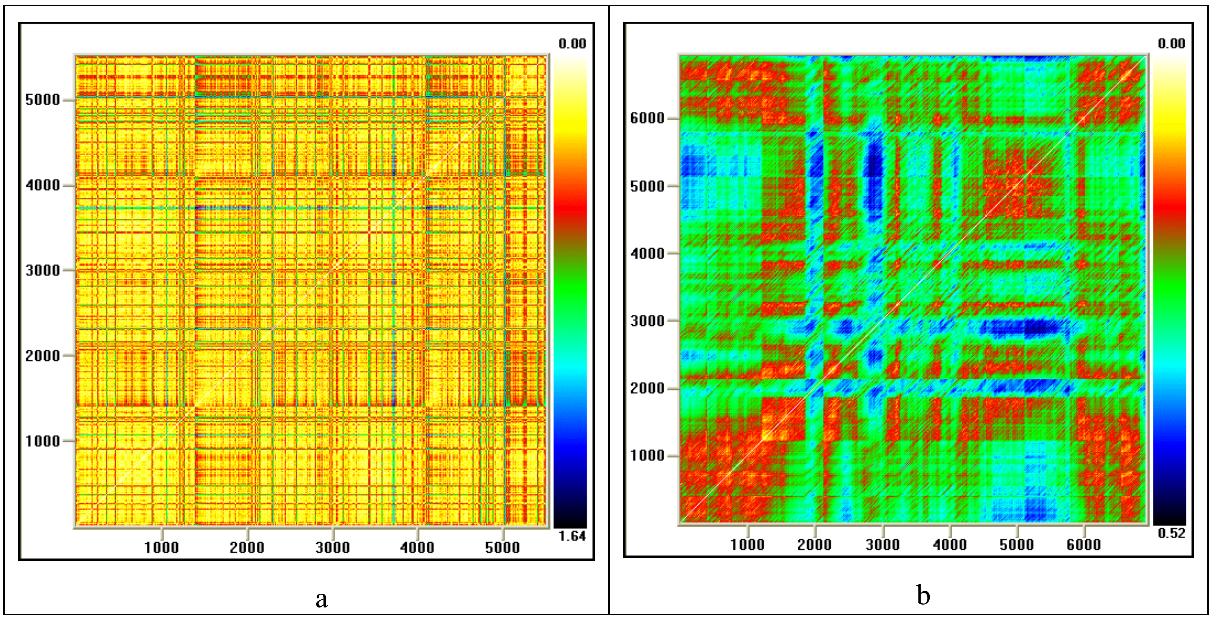


Figure 5 Recurrence plots corresponding to ECG of a young (a) and an old person (b).

The continuous ECG was digitised at 250 Hz. Each heart-beat was annotated using an automated arrhythmia detection algorithm, and each beat annotation was verified by visual inspection.

The R-R interval (inter-beat interval) time series for each subject was then computed and using them we have constructed recurrence plots as those shown in Figure 5.

Fractal dimensions of these plots are equal to 1.61 and 1.82 for old and young persons, respectively. This reflects the fact that the data obtained from old subjects have a more deterministic origin and the data obtained from young subjects are more random and complex. The complex dynamics of the healthy heart-beat arise from numerous coupled control systems and feedback loops that regulate the cardiac cycle on different time scales. Aging has a profound impact on many of the interacting neural and endocrine mechanisms that regulate heart rate, which may explain why the heart-rate time series loses much of its complex, irregular behaviour. This suggests that the distinctive patterns evident in recurrence plots of inter-beat intervals are empirically correlated with the age of the studied subjects, as they diagrammatically represent the complex dynamical interaction of the sympathetic and parasympathetic nervous systems.

6. Conclusions

In this study we have for the first time demonstrated fractal nature of recurrence plots and used fractal dimension of these images for the interpretation of various time series.

Acknowledgement

This work was supported by VEGA grant 1/2012/05.

References

[1] P. Soile, (2002) *Morphological Image Analysis*, Springer-Verlag, New York.
[2] Spillman, W.B., Jr. and Huston, D.R., Cellular automata for image analysis of damage in large structures. *Optical Engineering* 37, 898, (1998).
[3] Mandelbrot, B.B., How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science* 156, 636-638, (1967).

[4] Eckmann, J. P., Kamphorst, S. O. and Ruelle, D., Recurrence plots of dynamical systems. *Europhysics Letters, 4,* 973-977, (1987).

[5] Takens, F., Detecting strange attractors in turbulence. In*: Dynamical Systems and Turbulence*, Lecture Notes in Mathematics, No. 898, Springer-Verlag, Berlin, (1981).

[6] Marwan, N. URL: http://www.recurrence-plot.tk/

[7] Webber, C L., and Zbilut, J. P., Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology, 76,* 965-973, (1994).

[8] Kaluzny, P., and Tarnecki, R., Recurrence plots of neuronal spike trains. *Biological Cybernetics*, 68, 527-534, (1993).

[9] Trulla, L.L, A. Giuliani, J.P. Zbilut, and Webber, C.L., Recurrence quantification analysis of the logistic equation with transients, *Physics Letters A,* 223, 255-260, (1996).

[10] Zbilut, J. P. and Webber, C. L., Embeddings and delays as derived from quantification of recurrences. *Physics Letters A,* 171*,* 199-203, (1992).

[11] Visual Recurrence Analysis (*VRA*) can be downloaded from the following URL: http://home.netcom.com/~eugenek/download

[12] Zmeskal, O., Bzatek, T., and Nezadal, M., *HarFA*: Harmonic and Fractal Image Analyser. Software for determination of fractal dimension. Available at the following URL: http://www.fch.vutbr.cz/lectures/imagesci, (2005).

[13] Babinec, P., Zemanová, L. Babincová, M., Randomness and determinism in human heart-beat dynamics: Recurrence plot analysis, *Physica Medica*: *International Journal of Medical Physics*, 18, 63-68, (2002).