Investigation of Dynamical Systems in Pulse Oximetry Time Series

Jun Liang

***Abstract* –** Chaotic behavior of human physiology is a problem that can be investigated through various measurements. One of the most noninvasive and easy measurement of the cardiovascular and respiratory systems is blood oxygen saturation through a pulse oximeter. In this investigation, the level of chaos and complexity of the system underlying the pulse ox time series are quantified through calculation of largest Lyapunov exponent and correlation dimension. The data is also examined qualitatively by time-delay embedding. The results show that blood oxygen saturation is a physiological process that exhibits a low dimensional chaotic behavior. The finding is consistent with previous research and sparks directions for future studies.

I. Introduction

The human physiological system is vastly complex. There are many descriptors for the behavior of the system. One popular measure is the electrocardiogram (ECG), which captures the electrical activity generated by the heart. Another measure is blood oxygen saturation through pulse oximeter. The latter has more to do with the respiratory cycle while the ECG is completely associated with cardiovascular variations. However, the time series generated by both descriptors are very similar. There has been prior research done on ECG and pulse ox data that have discovered nonlinear dynamical behavior in the signals [3][4][7]. Practical applications of the investigation of the chaotic behaviors of pulse ox signals include determination and prediction of patient illness, such as detection of histologic chorioamnionitis in infants [1][2].

The purpose of this investigation is to test prior methods used in analyzing chaotic properties of time series on our data. In this paper, the dynamics of the set of pulse ox data is quantified through calculation of largest Lyapunov exponent and correlation dimension. Qualitative observation of chaos is done through time delayed embedding. The paper is organized as follows: Section II describes prior research and results on chaos using human physiological data; Section III describes the data used in this study; Section IV details the method used to calculate the parameters; Section V presents and interprets the results; Section VI concludes the investigation and recommends directions for future studies.

II. Prior Research

Blood oxygen saturation and ECG involve similar human physiology. A study on nonlinearity of ECG concludes that the time series is consistent with the behavior of a low dimensional, nonlinear chaotic system [9]. Correlation dimensions and entropy were calculated for three different ECG signals: normal, ventricular tachycardia (VT) and ventricular fibrillation (VF). The two descriptors are calculated using surrogate data, which are be generated from the original time series using windowed Fast Fourier Transform at frequencies f=0, 1/N, 2/N, …1/2, where N is the number of data points in the original time series [6][10]. The surrogate data retains the same power spectra as the original data. The null hypothesis on which the surrogate data is generated is that the ECG process is linear and thus not chaotic. By finding the correlation dimension of the surrogate data, mean and standard deviation of the correlation dimension for a supposedly linear ECG process are found. For 52 of the 81 data sets, at the 98% confidence level, the correlation dimension for the original time series rejects the null, which means that there's sufficient evidence that the ECG is a deterministically chaotic process [9]. The author concludes normal, VT, and VF recordings demonstrate a chaotic system underlying the ECG signal. However, comparison in correlation dimension and entropy are inconclusive in distinguishing the abnormal ECG from the normal ECG. Another method using the temporal distance between R peaks of ECG as discrete data points has also been used with similar results.

For dynamical analysis of pulse ox data, research has built on the nonlinearity analysis of blood oxygen saturation and has computed parameters that distinguish between healthy subjects and those with serious illnesses. A patent on the assessment of human physiology analyzes pulse ox data in real time and computes probability distribution, power spectrum, Lyapunov exponent, and other measures of the fractal dimension [1]. The basis for the patent is that a healthy physiology has complex dynamics while diseases can decrease the complexity of the system. Therefore, parameters that measure complexity of a dynamical system, i.e. correlation dimension, can reach a diagnosis. A similar study seeks to distinguish pulse ox data between normal infants and infants affected by histologic chorioamnionitis (HCA), a disease often fatal to newborns. Lempel-Ziv, largest Lyapunov exponent, correlation dimension, and Hurst values are calculated with the Lempel-Ziv complexity being the most prominent indicator in HCA cases [2]. One reason for pulse ox data being favored in clinical applications is that it is an easy and noninvasive measurement. Due to the limitation in data and subjects, this paper will calculate two of the parameters described in the literature that provide a good description of the dynamics of the physiological system that is measured by the pulse ox.

III. Data

The data used in this investigation is collected by a finger cuff pulse oximeter as a part of the Biopac® Student Lab equipment. The data is discrete with a sampling frequency of 1000 Hz. The data is filtered by a band-pass filter preset on the data collection software. It is 200 seconds long and collected from a healthy male with no known cardiovascular or respiratory illnesses. Due to the length of the data (200,000 points), it is re-sampled at every 50th point, resulting in a new frequency of 20 Hz. Another way for shortening the data is to use the first 40 seconds of the data. It was tested and yielded similar computational results. However, the method of chaos analysis used here prefers the first method [6].

IV. Methodology

A. Time-delay embedding

Time-delay plots can qualitatively characterize the dynamics of the data in 1 and 2 dimensions. The delay is chosen at the time lag when the autocorrelation function drops to $1-\frac{1}{e}$ of its initial value [7].

B. Lyapunov exponent

To calculate the largest Lyapunov exponents and correlation dimensions of the data, a low computational cost method is used as described by Rosenstein et al (1992) [7]. If we assume that the pulse ox time series is chaotic, then two initial conditions close to each other will diverge exponentially at the rate of the largest Lyapunov exponent.

$$d(t) = Ce^{\lambda_1 t} \tag{1}$$

where d is the distance between nearest neighbors, C is a constant, and λ_1 is the largest Lyapunov exponent.

We reconstruct the time series into M time state vectors, each representing the state of the system at discrete time i. Starting with data set $\{x_1, x_2, \ldots x_N\}$, construct a matrix $\mathbf{X}$ of size $M \times m$ such that:

$$\begin{aligned} \mathbf{X} &= [\mathbf{X}_1\ \mathbf{X}_2 \ldots \mathbf{X}_M]^T \\ \mathbf{X}_i &= [x_i, x_{i+J}, \ldots x_{i+(m-1)J}] \end{aligned} \tag{2}$$

J is the reconstruction lag chosen to be the time-delay found through the autocorrelation function in time-delay embedding; m is the embedding dimension, which should be in accordance with Takens' theorem such that $m > 2n$ where n is the dimension of the underlying attractor of the time series (m is chosen to be 6 from literature [3]). This results in a $M \times m$ matrix $\mathbf{X}$ where each row vector denotes a discrete time state.

(3)

The nearest neighbors are then found by:

$$C_j = d_j(0) = \min_{X_{\hat{j}}} \| \mathbf{X}_j - \mathbf{X}_{\hat{j}} \|$$

where $\|..\|$ represent the Euclidean norm and the neighbors conform to the temporal constraint

$$j - \hat{j} > mean\ period \tag{4}$$

where the mean period is estimated by the reciprocal of the mean frequency of the power spectrum calculated using Welch's method with no windowing. Thus $d_j(i)$ is the separation between the j^{th} nearest pair of neighbors after i time steps. From (1), the largest Lyapunov λ_1 exponent can be represented as:

$$\ln(d_j(i)) \approx \ln(C_j) + \lambda_1(i \cdot \Delta t) \tag{5}$$

It represents M parallel lines, with $j = 1, 2, 3, \ldots M$. The least square fit to the "average" of these lines results in the curve

$$y(i) = \frac{1}{\Delta t}\left\langle \ln(d_j(i)) \right\rangle \tag{6}$$

where $\left\langle \ln(d_j(i)) \right\rangle$ is the average for all j. The curve is plotted with $i \cdot \Delta t$ on the x-axis and $\left\langle \ln(d_j(i)) \right\rangle$ on the y-axis, and the slope of the linear regression to the curve is the largest Lyapunov exponent for the time series.

C. Correlation dimension

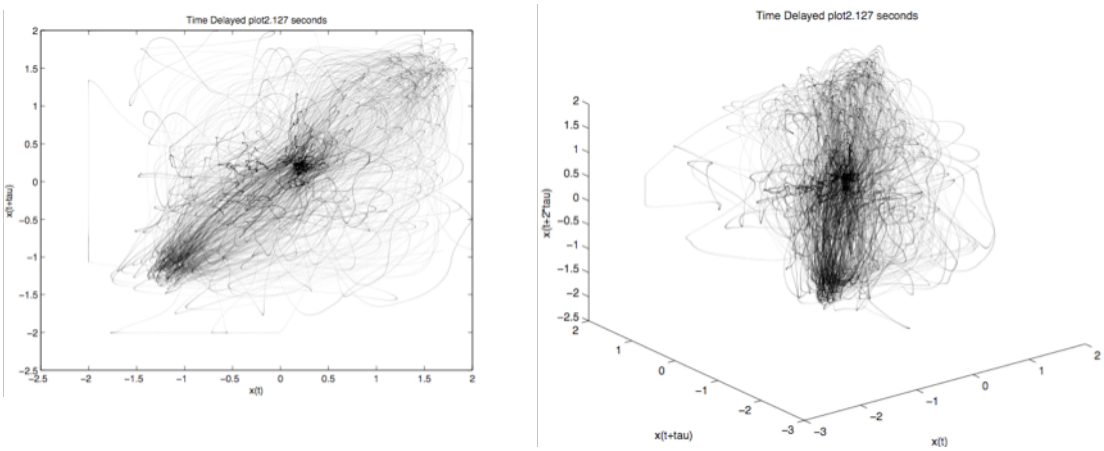
In calculating the correlation dimension, the results from the Lyapunov exponent calculations are used. The correlation sum for embedding dimension m is computed as:

$$C_m(r) = \frac{2}{M(M-1)} \sum_{i \neq k} \theta\left[r - \| \mathbf{X}_i - \mathbf{X}_k \|\right] \tag{7}$$

where r is the distance between time states; θ is the Heavy side function, which equals to 1 if $r - \| \mathbf{X}_i - \mathbf{X}_k \| > 0$ and 0 if $r - \| \mathbf{X}_i - \mathbf{X}_k \| < 0$. By choosing a few values of r, the linear fit to $\ln(C_m(r))$ plot against $\ln(r)$ has a slope that is equal to the correlation dimension.

V. RESULTS

Figure 1 shows the time delayed plots of the time series. There seems to be an unclear attractor in the plots, but it is difficult to tell qualitatively. This is consistent with ECG time-delay plots by Mehta and Miller [5]. Previous experiments have shown that human physiological data tend to have a low dimensional attractor. Another set of plots is done for a different lag (7 seconds) [8] with similar results.



(a) (b)

Figure 1: (a) is time-delay plot with lag J=2.127 seconds in 1-D; (b) is time-delay plot with same lag in 2-D

For calculating the Lyapunov exponent, 40 discrete times are used for the divergence plot (Figure 2a). The result shows a slightly positive exponent (0.001). However, the results of this experiment are similar to the plots of a noisy signal in Rosenstein et al (1992) [7] (Figure 2b). The actual Lyapunov exponent is the slope of the portion of the divergence graph with small values of $i \cdot \Delta t$. A plot of the divergence using a filtered version of the original time series shows a stronger initial linear trend (Figure 2d). The filtered data consists of a segment of data that shows prominent peaks and troughs. Under the assumption that the pulse ox time series is also noisy, the Lyapunov exponent is recalculated as 0.012. The positive exponent is evident that the process underlying the pulse ox signal is chaotic. However, the small value is indicative that the initial conditions diverge slowly.

The correlation dimension is calculated with 5 values of r (Figure 3). The best-fit line estimates the correlation dimension as 2.7 (literature for ECG recordings show correlation dimension of 2.3±0.36 [9]). This shows that the dimension of the chaotic attractor in blood oxygen saturation variation over time is fairly low. This is consistent with previous research on ECG signals that have discovered a low dimensional dynamic system. This also validates our selection of the embedding dimension $m = 6$. Repeated calculations with $m = 7$ and $m = 8$, the range of discovered embedding dimension in ECG time series [3], result in correlation dimensions of 2.3 and 1.9, respectively. The consistency is acceptable given that the data is likely to be noisy.

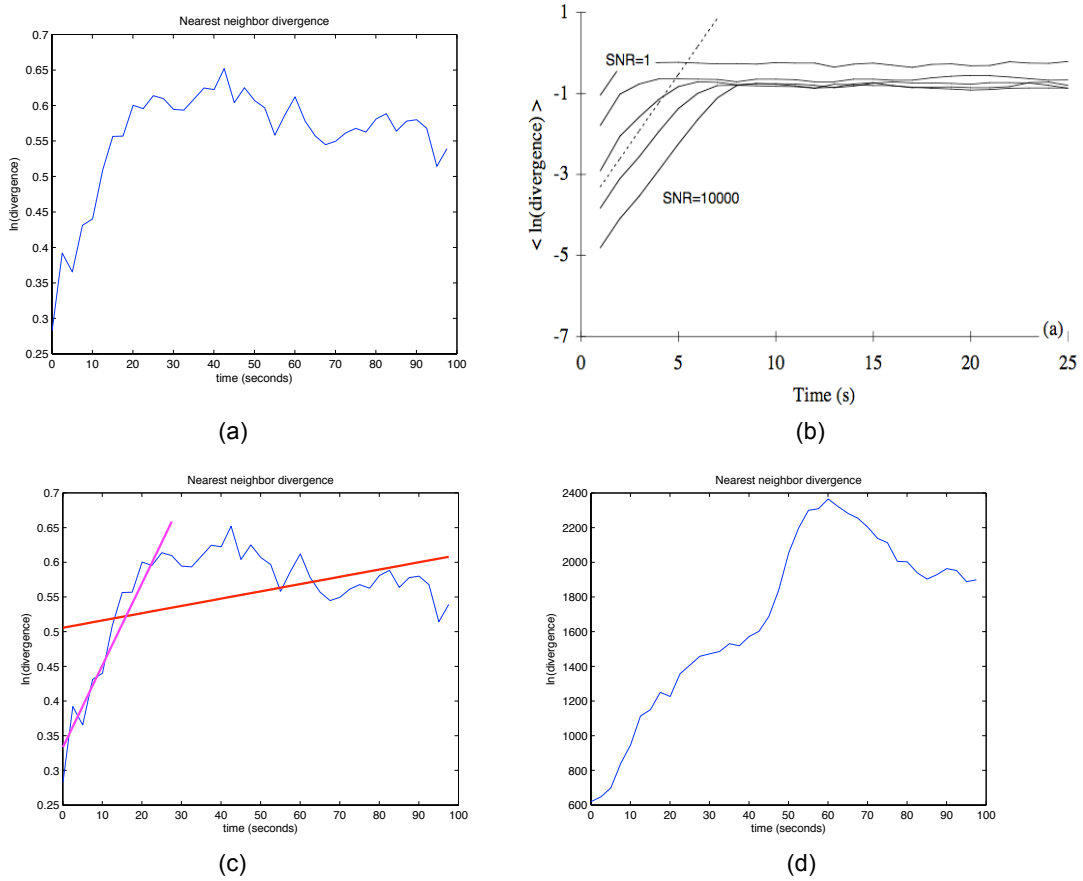


Figure 2: (a) shows the divergence of nearest time states over time; (b) is a similar plot produced by Rosenstein et al (1992), the solid line is the divergence of a noisy timeseries and the dotted line has a slope equal to the largest Lyapunov exponent of the underlying attractor; (c) shows the lines of best fit: red line is the overall fit, magenta line assumes noisy data; (d) shows the divergence of nearest neighbors for time series after removal of some noise.

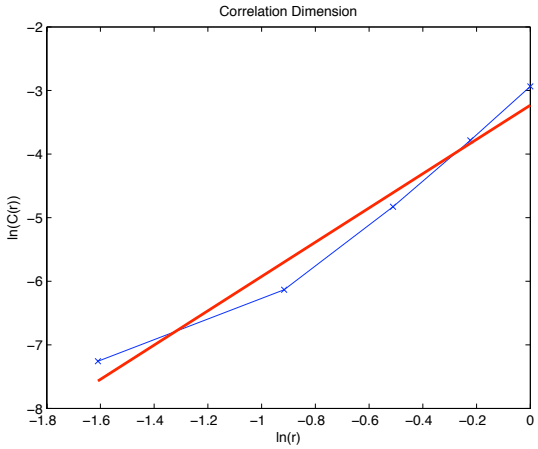


Figure 3: natural log of the correlation sum plotted against the natural log of r. The red line is the linear regression with a slope approximately equal to the correlation dimension. The linear characteristic of the curve is strong

VI. Conclusion

Using the methods described by Rosenstein et al (1993), blood oxygen saturation data collected from a pulse oximeter is analyzed and its chaotic properties qualitatively and quantitatively observed. The two parameters, Lyapunov exponent and correlation dimension, show that the pulse ox measures a process that is chaotic but not very complex. This is in opposition to the patent that claims to uncover serious illness in patients based on the decreased complexity in their physiology based on pulse ox data. Based on our study, any difference would be too small to discern or attribute to the sensitivity of the device.

However, keep in mind that it is possible that the data used for this investigation is too noisy, which will bias the findings presented here. Therefore, there are more questions raised by this investigation than answered. Is the pulse ox a reliable measurement tool? Are there statistically significant differences between the parameters measured from healthy and ill subjects? What about other measures of chaos, i.e. Lempel-Ziv complexity? The limitations on the current project do not allow for adequate exploration of these questions, which should uncover interesting findings.

References

[1] De Felice, Claudio, Mitchell Goldstein, and Giuseppe Latini. Physiological assessment system. Masimo Corporation, assignee. Patent 7292883. 6 Nov. 2007. Print.

[2] De Felice, Claudio, Mitchell R. Goldstein, Stefano Parrini, Alberto Verrotti, Mario Criscuolo, and Giuseppe Latini. "Early dynamic changes in pulse oximetry signals in preterm newborns with histologic chorioamnionitis." *Pediatric Critical Care Medicine* 7.2 (2006): 138-42. Web. 26 Nov. 2009.

[3] Fell, J., K. Mann, J. Roschke, and M. S. Gopinathan. "Nonlinear analysis of continuous ECG during sleep I. Reconstruction." *Biological Cybernetics* 82.6 (2000): 477-83. *SpringerLink*. Web. 22 Nov. 2009.

[4] Goranitou, Georgia S., Dimitrios N. Chrissos, Haralampos V. Harissis, Dimitra D. Kontoyianni, and Panagiotis G. Antonatos. "Nonlinear Analysis of ECG in Normal Subjects and Acute Myocardial Infarction Patients." *Annals of Noninvasive Electrocardiology* 3.2 (2008): 139-46. 28 June 2008. Web. 27 Nov. 2009.

[5] Mehta, Chetan, and Matthew Miller. *Chaos Analysis for EKG Time Series Data*. Tech. Hanover, NH, 2007. Print.

[6] Prichard, Dean, and James Theiler. "Generating Surrogate Data for Time Series with Several Simultaneously Measured Variables." *Physical Review Letters* 73.7 (1994): 951-54. Print.

[7] Rosenstein, Michael T., James J. Collins, and Carlo J. De Luca. *A practical method for calculating largest Lyapunov exponents from small data sets*. Tech. no. 05.45.+b, 02.50.+s, 02.60.+y. Print

[8] Small, Michael. *Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance*. Vol. 52. Singapore: World Scientific. Mar. 2005. Web. 1 Dec. 2009. <http://www.worldscibooks.com/chaos/5722.html>.

[9] Small, Michael, Dejin Yu, Jennifer Simonotto, Robert G. Harrison, Neil Grubb, and K.A.A. Fox. "Uncovering non-linear structure in human ECG recordings." *Chaos, Solitons & Fractals* 13 (2002): 1755-762. Print.

[10] Theiler, James, Stephen Eubank, Andre Longtin, Bryan Galdrikian, and J. Doyne Farmer. "Testing for Nonlinearity in time series: the method of surrogate data." *Physica D* 58 (1992): 77-94. Print.

[11] Yalamova, Rossita, Liu Qi, Lin Wang, Yudong Chen, Spencer Mathews, and Guannan Gong. *Detecting chaos in financial time series*. Tech. Print.