Global Journal of researches in engineering
Electronic and electronics Engineering
Volume 11 Issue 8 Version 1.0 December 2011
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-4588 & Print ISSN: 0975-5861

Statistical Investigation of ECG Signal of Sleep Apnea Patient

By Chandan Das,Mofazzal H. Khondekar
Dr.B.C. Roy Engineering College, Jemua Road, Fuljhore, Durgapu

Abstract - The Hurst Exponent of the time series of a normal patient and apneal patient suggest that they are anti-persistent and the later has more self similarity compared to the former. It has been established that they are AR process and nonstationary. The Semblance analysis suggests strong correlation both positive and negative between them. Tentative mathematical models of the normal an apneal patient has also been suggested using Yule Walker method.

Keywords:: Hurst exponent, ECG, autocorrelation, partial autocorrelation, Wavelet, Semblance.

GJRE-F Classification : FOR Code: 090609

STATISTICAL INVESTIGATION OF ECG SIGNAL OF SLEEP APNEA PATIENT

Strictly as per the compliance and regulations of:

Statistical Investigation of ECG Signal of Sleep Apnea Patient

Chandan Das [α] , Mofazzal H. Khondekar [β]

Abstract -The Hurst Exponent of the time series of a normal patient and apneal patient suggest that they are anti-persistent and the later has more self similarity compared to the former. It has been established that they are AR process and nonstationary. The Semblance analysis suggests strong correlation both positive and negative between them. Tentative mathematical models of the normal an apneal patient has also been suggested using Yule Walker method.

Keywords: Hurst exponent, ECG, autocorrelation, partial autocorrelation, Wavelet, Semblance.

I. Introduction

Sleep apnea is the occurrences of interrupted breathing during sleep. Obstructive sleep apnea is a well-known disorder in which relaxation of muscles in the throat repeatedly close off the airway during sleep; the person wakes just enough to take a gasping breath. This process is repeated many times during sleep and usually is not remembered the next day. Those suffering from severe obstructive sleep apnea typically complain of sleepiness, irritability, forgetfulness, and difficulty in concentrating. They may have difficulties in their occupational or social lives and be prone to motor vehicle accidents. The disorder has been medically linked to hypertension, which in turn puts people at greater risk of heart failure and stroke.

An electrocardiogram (ECG or EKG, abbreviated from the German Elektrokardiogramm) is a graphic produced by an electrocardiograph, which records the electrical activity of the heart over time [1]. Its name is made of different parts: electro, because it is related to electronics, cardio, Greek for heart, gram, a Greek roots meaning "to write". Specific waveforms within the ECG represent the electrical activity associated with mechanical events such as ventricular contraction and relaxation (systole and diastole). Analysis of the various waves and normal vectors of depolarization and re-polarization yields important diagnostic information [2].

ECG signals of the normal patient and apnea patient being taken for a period of 15minutes [3, 4] with the sampling interval of 4 msec. In this paper we will try to find out the nature of variability of the above two ECG signals using Finite Variance Scaling Method (FVSM). But before we proceed for the above action we have to consider that in practical cases all the observed data involve some amount of circumstantial errors which may creep in due change in environment, or systematic error which is due to factors inherent in the manufacture of the measuring instrument arising out of tolerances in the components of the instruments. Study of such data in presence of error may often not succeed to give true information. There is the need to remove these errors up to a satisfactory level. For these purpose we frequently use different methods of filtration in the time-dependent data. Here Simple Exponential Smoothing technique has been used for the filtration purpose.

The Hurst Exponent obtained from FVSM quantifies the relative affinity of a time series either to regress strongly to the mean or to cluster in a direction. Autocorrelation plots are used for checking randomness in a data set. This randomness is estimated by computing autocorrelations for data values at varying time lags. For random time series, such autocorrelations are near zero value for every time-lag, whereas for deterministic series, one or more of the autocorrelations will have notably non-zero values.

Partial autocorrelation plots are used here for model identification in Box-Jenkins models of the time series.

Semblance Analysis using the continuous wavelet transform has been done to investigate the similarity of the phase relationship locally between the two signals which is a function of frequency and time of the signals.

II. Theory

a) Simple Exponential Smoothing

Exponential Smoothing helps to produce a smoothed Time Series by assigning exponentially decreasing weights as the observation in the time series get older. Simple Exponential Smoothing [5] the prescribed model for a time series data $\{x_i\}; where\ i = 1,2,3,\ldots\ldots\ldots,n$ after being exponentially smoothed is $y_1 = x_1$ And $y_i = \alpha x_i + (1-\alpha) y_{i-1}\ ;\ i = 1,2,3,\ldots\ldots.,n$ where y_i is the smoothed data at the i-th position and α ($0 < \alpha < 1$) is a parameter. This is equivalent to $y_1 = x_1$ and

Author : Department of Electronics and Instrumentation Engineering, Dr. B.C. Roy Engineering College, Jemua Road, Fuljhore, Durgapur-713 206, India

$$y_i = \alpha x_i + \alpha(1-\alpha)x_{i-1} + \alpha(1-\alpha)^2 x_{i-2} + \dots$$

$$\dots + \alpha(1-\alpha)^{i-2} x_2 + \alpha(1-\alpha)^{i-1} x_1 \quad for\ i = 1,2,3,\dots\dots,n$$

where the sum of the corresponding weights α, α (1-α), α (1-α) ^{2},α (1-α) $^{i-2}$ and (1-α) $^{i-1}$ is equal to unity. Thus in effect, each smoothed value is a convex linear combination of all the previous observations as well as the current observation.

III. Finite Variance Scaling Method

A familiar version of Finite Variance Scaling Method (FVSM) is the Standard Deviation Analysis (SDA) [6, 7, 8], which is based on the assessment of the standard deviation D (t) of the variable x (t).

In a time series {x (t_i)} observed at the instants t_i for i=1, 2…, n it yields

$$D(t_i) = \left[\left\{\frac{\sum_{i=1}^{n} x^2(t_i)}{i}\right\} - \left\{\frac{\sum_{i=1}^{n} x(t_i)}{i}\right\}\right]^{\frac{1}{2}} \qquad 1$$

For n=1, 2, 3………j

Eventually it is observed [6, 7 and 8]

$$D(t) = t^H \qquad 2$$

The exponent H is known as the Hurst exponent. It is evaluated from the gradient of the best fitted straight line in the log-log plot of D (t) against t. The value of the Hurst exponent ranges between 0 and 1. A value of 0.5 indicates a true random walk (a Brownian time series). In a random walk there is no correlation between any element and future element. A Hurst exponent value 0<H<0.5 will exist for a time series with anti-persistent behavior (negative autocorrelation) [9]. If the Hurst exponent is 0.5<H<1.0, the process will be a long memory process. A Hurst exponent value in this range indicates persistent behavior (or, a positive autocorrelation).

IV. Autocorrelation and Partial Autocorrelation

Autocorrelation is a statistical method used for time series analysis. It refers to the correlation of a time series with its own past and future values. The values of the autocorrelation coefficients serve two purposes. It can detect non-randomness in a data set. If the values in the data set are not random, then autocorrelation can help the analyst chose an appropriate time series model.

The set of autocorrelation coefficients arranged as a function of separation in time is the sample autocorrelation function (acf). If x_i be signal of length N and $\bar{x}$ be its overall mean i.e. $\bar{x} = \sum_{t=1}^{N} x_t$

The autocorrelation coefficient at lag k is given by:

$$r_k = \frac{\sum_{i=1}^{N-k}(x_i - \bar{x})(x_{i+1} - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \qquad 3$$

The plot of the autocorrelation coefficients as a function of lag is called the correlogram.

Positive autocorrelation signifies the persistent trend in the series where the system likes to remain in the same state from one observation to the next. Whereas negative autocorrelation is distinguished by an inclination for positive departures from the overall mean $\bar{x}$ to follow a negative departure, and vice versa.

In order to find the connection between x_i and x_{i+k} partial autocorrelation is used where linear influence of the random variables lying between $x_{i+1},\dots,x_{i+k-1}$ is filtered out of the x_i and x_{i+k} then the correlation of the transformed random variables is calculated. If we define a function $P(k)$ as

$$P(k) = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{pmatrix} \qquad 4$$

Then partial autocorrelation can be expressed as, [10]

$$\phi_{kk} = \frac{\Delta(P(k)^*)}{\Delta(P(k))} \qquad 5$$

Where $P(k)^*$ is same as $P(k)$ except the kth column in equation 4 is replaced by ρ_k. ρ_k Is the autocorrelation function at lag k.

Partial autocorrelation is a commonly used tool for model identification. If the sample autocorrelation plot indicates that an AR model may be appropriate, then the sample partial autocorrelation plot is examined in order to identify the order. We look for the lag on the partial autocorrelation plot beyond which its values essentially become zero, more specifically where the values of the coefficients are considerably less than a 95% confidence level i.e. $\pm 2/\sqrt{N}$.

V. The continuous wavelet transform and semblance analysis

The continuous wavelet transform (CWT) [11] of a signal $x(t)$ is given by

$$CWT(u,s) = \int_{-\infty}^{\infty} x(t) \frac{1}{|s|^{0.5}} \psi^{*}\left(\frac{t-u}{s}\right) dt \qquad 6$$

Where Ψ is the mother wavelet, and Ψ^{*} is complex conjugate of Ψ, s allows the wavelet to be stretched to various scales and u allows the wavelet to be translated to by various displacements. The CWT basically is the convolution of the signal with scaled version of the mother wavelet. Here, the complex Morlet wavelet has been used, which is defined as [11, 12]

$$\psi(x) = \frac{1}{\pi f_b} e^{j2\pi f_c x} e^{-x^2/f_b} \qquad 7$$

Where f_b tunes the wavelet bandwidth and f_c is the wavelet centre frequency. For $f_c = 1.0$, scale becomes equivalent to wavelength. The behaviour of the signal on different scales can be revealed by varying the scale s (in Eq. (6)). When the mother wavelet chosen here is complex and hence its real and imaginary parts generate a Hilbert transform pair, to order to have orthogonality. Since the mother wavelet is complex, the CWT will also be complex which has a phase at every time and scale. The cross-wavelet transform [13, 14] defined as:

$$CWT_{1,2} = CWT_1 \times CWT_2^{*} \qquad 8$$

CWT_1 & CWT_2 are the continuous wavelet transforms of two signals $x(t)$ and $y(t)$. $CWT_{1,2}$ is a complex quantity having an amplitude given by

$$A = |CWT_{1,2}| \qquad 9$$

and local phase θ given by:

$$\theta = \tan^{-1}\left(\Im(CWT_{1,2}) / \Re(CWT_{1,2})\right) \qquad 10$$

θ varies between $-\pi$ and $+\pi$.

The Semblance S is defined as: [15]

$$S = \cos^{n}(\theta) \qquad 11$$

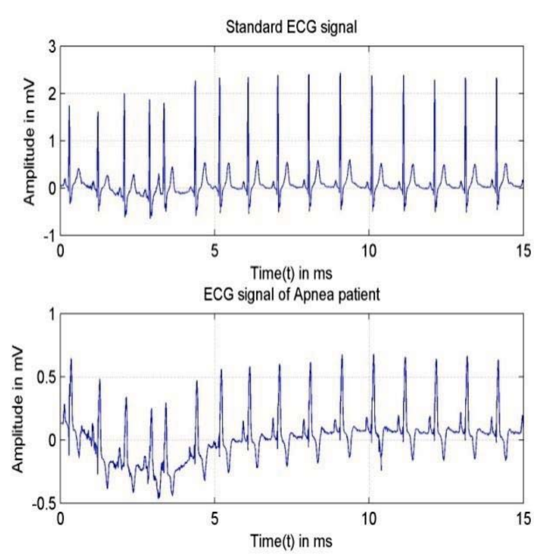
For every odd values of. $n \geq 0$

S may take a value between -1 and +1. The value of S gives the degree of correlation between the two signals as below:

$$S = \begin{cases} -1 & x \text{ and } y \text{ has negative correlation} \\ 0 & x \text{ and } y \text{ has NO correlation} \\ +1 & x \text{ and } y \text{ has positive correlation} \end{cases}$$

So if the value of S is close to -1, it implies that x and y has high negative correlation whereas its value close to +1 implies a high positive correlation between the signals. The value close to 0 indicates a poor correlation between the two signals.

VI. Results



The ECG signal for the normal patient and the apnea patient are shown in fig.1

Fig. 1: ECG signals

The $D(t)\ vs.\ t$ curve for the two signals are shown in fig.2

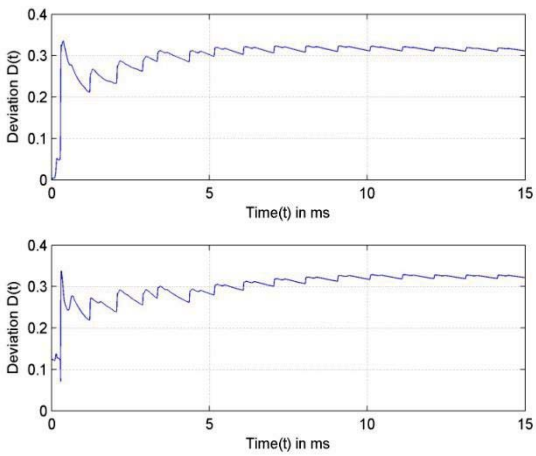


Fig.2: Deviation versus Time curve of ECG signal

We have applied the FVSM to obtain the Hurst exponent H for the time series ECG signals of the normal patient and apnea patient being taken for a period of 15minutes [3,4] with the sampling interval of 4 msec. The values of H obtained are given in table 1.

Table: 1

Signals	Hurst Exponent H
Normal Patient	0.2779
Apnea Patient	0.1300

The autocorrelation coefficients of the two signals for various lags up to 20 are given in fig.3

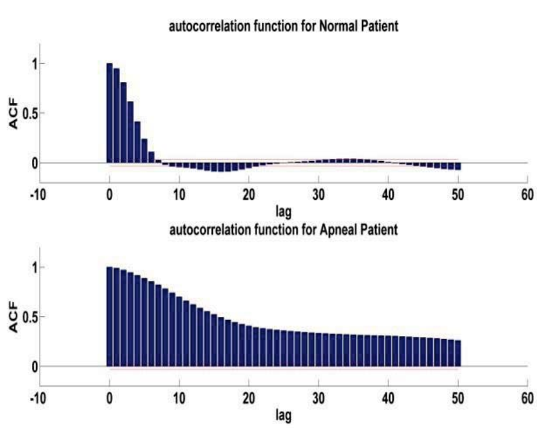


Fig.3: Autocorrelation of two ECG lead signal

The partial autocorrelation coefficients are given in fig.4

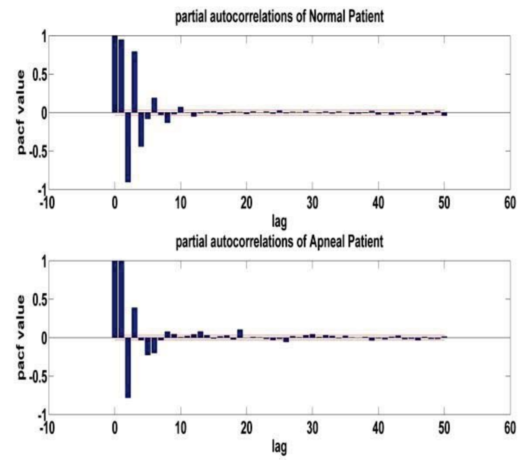


Fig.4: Partial Autocorrelation of two ECG lead signal

The Semblance analysis results shown in fig.5

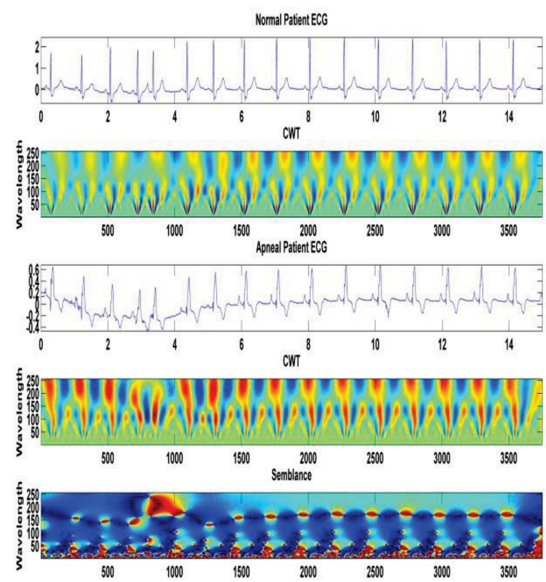


Fig.5: Semblance analysis of two signals

Bright red corresponds to a semblance of +1, 50% green to a semblance of zero, and dark blue to a semblance of -1.

VII. Discussion

The Hurst exponent that we have obtained for both the normal and apneal patient are less than 0.5 which suggest that the signals are having anti-persistent behavior i.e. there are trends of a decrement in values followed by an increment and vice versa and it is more pronounced in case of the apneal patient.

The Fractal Dimension (D) is related to the Hurst exponent by the equation of D=2-H. Hence the D for the normal patient is 1.7221 and for the apneal patient it is 1.87. These values of D suggest that the apneal patient signal has more self similarity than that of a normal patient.

From the auto-correlogram as shown in fig.3 we find that the autocorrelation coefficients die down to zero more rapidly than that of the apneal patient. The autocorrelation coefficients for apneal patient seem not to die down to zero except for large values of the lag. It signifies that the apneal patient's time series has a stronger trend compared to that of a normal patient. The auto-correlogram also suggests that both the systems from which the signals originated are Autoregressive (Markov) process (AR). The tendency of the autocorrelation coefficients of the apneal patient not to die quickly as compared to those of the normal patient can be taken as an indication of stronger nonstationarity of the former signal with respect to the later.

From the partial auto-correlogram as in fig.4 we can claim that normal signal is auto-regressive process of order 9 i.e. AR (9) but the patient signal is auto-regressive process of order 4 i.e. AR (4). Using the Yule Walker Equation [10], the model of the two data series can be estimated as

$$x_t = \delta + \sum_{i=1}^{9} \varphi_i x_{t-i} + e_t$$ for normal patient

And

$$x_t = \delta + \sum_{i=1}^{4} \varphi_i x_{t-i} + e_t$$ for apneal patient

Where $$\delta = \left(1 - \sum_i \varphi_i\right)\mu$$

For normal patient $\delta = 0.0792$ and for apneal patient $\delta = 0.006$.

The values of the coefficients φ_i as calculated by Yule walker Equation are given in the table 2.

Table: 2

i	φ_i for normal patient	φ_i for apneal patient
1	2.841	2.0254
2	-3.1143	-1.381
3	1.2009	0.3446
4	0.5002	0.0023
5	-0.5196	
6	-0.1257	
7	0.2836	
8	-0.0776	
9	-0.0163	

Solving the equation of the polynomial $\phi(z) = 1 - \sum_{i=1}^{p} \varphi_i z^i = 0$ we have obtained the complex roots which are found to be less than unit circle, which establish the nonstationarity of the signals. $p = 9$ for the normal patient and $p = 4$ for apneal patient.

Semblance analysis gives the phase relationship between the two signals. It is found that the two signals are highly negatively correlated at lower scales at regular intervals of time where as at higher scales these are highly positively correlated at regular intervals of time. At even higher scales (more than 200) the signals are mostly negatively correlated except at the time between 3-4 minutes.

References Références Referencias

1. Braunwald E. (Editor), Heart Disease: A Textbook of Cardiovascular Medicine, Fifth Edition, p. 108, Philadelphia, W.B. Saunders Co., 1997. ISBN 0-7216-5666-8.
2. American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care - Part 8: Stabilization of the Patient With Acute Coronary Syndromes." Circulation 2005; 112: IV-89 - IV-110.
3. http://ecg.mit.edu/timeseries/index.html
4. www.physionet.org/
5. R.L. Winkler. and S. Makridakis, Journal of the Royal Statistical Society, Series A (General), vol. 146 (2), 1983, p. 150.
6. A. Sarkar, P. Barat, P. Mukherjee and S. K. Bandyopadhyay, Proceedings of National Conference on Nonlinear Systems and Dynamics (held at A.M.U., Aligarh during February 24-26, 2005) 2005, p.155.
7. N. Scafetta and P. Grigolini, Phys. Rev. E, Vol. 66, 2002,p. 036130.
8. B. B. Mandelbrot, The Fractal Geometry of Nature, Freeman, New York, 1983.
9. K. Mofazzal Hossain, Dipendra N. Ghosh and Koushik Ghosh, "Scaling Analysis Of National Stock Exchange Index"
10. . Box G.E.P, Jenkins G.M, Reinsel G.C, 'Time Series Analysis-Forecasting and Control', 3rd Edition.Pearson Education Ltd.
11. Mallat, S., 1998. A Wavelet Tour of Signal Processing. Academic Press, New York, 577pp.
12. Teolis, A., 1998. Computational Signal Processing with Wavelets. Birkha¨ user Boston Inc., Boston, MA, 352pp.
13. M. Alfaouri and K. Daqrouq, ECG Signal Denoising By Wavelet Transform Thresholding, American Journal of Applied Sciences 5 (3): 276-281, 2008, ISSN 1546-923.

14. Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological Society 79, 61–78.
15. G.R.J. Coopera,_, D.R. Cowan, 'Comparing time series using wavelet-based semblance analysis', Computers & Geosciences 34 (2008) 95–1