An Overview of Linear and Nonlinear Analysis for Alcoholic

[1]Peng Tang, [2]Ji Li
[1,] *Changsha University of Science & Technology, cootertang@gmail.com*
[2,] *Changsha University of Science & Technology,hangliji@163.com*

Abstract

By now, many methods are introduced to analyze the electroencephalogram (EEG) signals for identifying people who are alcoholic. Linear approaches such as frequency domain analysis and time-frequency domain analysis are widely studied. Scholars have shown that employing a filter and applying Principal Component Analysis (PCA) can increase the quality of feature extraction and thus increase the classification accuracy. In contrast to linear measures, nonlinear measures are better in complexity and sometimes accuracy. These measures rely on the proper parameters. In this paper, we conclude the effectively and commonly used linear and nonlinear analyses that are used for analyzing the EEG to tell whether one is alcoholic, and show the correlation and enhancement among these algorithms.

Keywords: *EEG, Linear, Nonlinear, Alcoholic, Overview*

1. Introduction

Electroencephalogram (EEG) is a representative signal that contains useful information about the electrical activity of the brain. Since the EEG signal is noninvasive, the analysis of EEG signal is popular nowadays. One of the most important applications using EEG is analysis of brain electrical activity under various conditions for example, sleep disorders [1,2], epilepsy [3-5], depression and mania.

Drinking alcohol excessively causes a lot of health problems. For instance, some have the difficulty walking, some suffer from the blurred vision, slurred speech, slowed reaction, impaired memory and sleep [6]. One who drinks alcohol excessively over a long period of time will perhaps have a problem in recognition. Hence, it's of significance to study the effect of alcohol on EEG. Much of feature information is used to research the connection of EEG and alcohol. In recent years, the rapid development of various theories of BCI provided us with many new methods in EEG data processing.

Clinical evidence of using advanced signal processing methods has proven detecting alcoholism from the EEG can be effective [7-9]. Some Linear and nonlinear analyses can both be applied to EEG. The linear characteristics analysis is a traditional way, and it rests on the assumption that the EEG is a stationary process in a short period of time. Linear approaches include time domain analysis, frequency domain analysis, and time-frequency domain analysis. Popular methods of linear time-frequency distribution are the Short-Time Fourier Transform (STFT) and the Wavelet Transform (WT). Usually Principal Component Analysis (PCA) is applied after transforming, which reduces noise, redundancy and increase the speed and accuracy. Then the Power Density Spectral (PSD) or High Order Spectral (HOS) will be used to extract the feature. In recent years, the development of EEG using time-frequency analysis improved quickly, especially in the aspect of the combination of time-frequency features and any kinds of wavelet transform methods.

Since the EEG signals are originally dynamic and nonlinear [10,11], some nonlinear analysis techniques have been introduced for several EEG based applications such as diagnosis of Alzheimer's disease diagnosis of autistic spectrum disorder [12], and study of attention deficit/ hyperactivity disorder. Researchers studying nonlinear characteristic methods find that though there are some obstacles in applying these algorithms such as selecting the proper parameters, it has the advantages of complexity and satisfying accuracy, which could possibly used as a mode of real-time detection. For nonlinear characteristics analysis, a set of methods are classified as information theory, for instance, Permutation Entropy (PE) Singular Value Decomposition (SVD) entropy , Approximate Entropy (ApEn), Sample Entropy (SampEn). In contrast, series algorithms are in the chaos theory, namely, Correlation Dimension (CD), Largest Lyapunov Exponent (LLE), Hurst Exponent (HE).

2. Linear Analysis and Algorithm Enhancement

The linear characteristic methods are generally divided into four parts, first, time and frequency transform is applied to the original data set, then, the analysis method is applied to extract the features with the reduction of the dataset, mostly, the PCA, which also have the ability of reducing the noise [13]. Note that PCA could be applied either before the transform or after. At last, classification methods are used to perform the final result. In this paper, classification methods will not be overviewed because the performance of the classification highly relies on the dataset, besides, the specific method for classification does not change much of the final result.

2.1 Time and Frequency Transform

Using the Discrete Fourier Transform (DFT) is computationally expensive, with the complexity of $O(N^2)$, while the Fast Fourier Transform (FFT) is faster, with the complexity of $O(N \log N)$. The power spectrum of the EEG signal obtained using methods like FFT is not accurate due to a reduction in the frequency resolution as the Signal-to-Noise Ratio (SNR) of the EEG signal is low [14]. The limitation, of the FFT may be overcome by using a parametric modeling technique [15].

Instead of analyzing the signals in only frequency domain, Short-Time Fourier Transform (STFT) and WT can decompose the signal into time-frequency domain. Besides, the practical signals in engineering are non-stationary signals, usually with turbulence, and WT can resist some turbulences.

STFT maps a signal into a two-dimensional function of time and frequency. It generates a time-frequency representation of a signal by integrating the signal at each time point with a series of windows harmonics of various frequencies represents a sort of compromise between the time and frequency-based view of a signal. Given a window function with a short-time window width $\eta(t)$, and then let the window slipping, STFT of the signal $z(t)$ is defined as follows:

$$STFT_z(t, f) = \int_{-\infty}^{\infty} z(t')\eta^*(t'-t)e^{-j2\pi ft'}dt' \tag{1}$$

where $STFT_z(t, f)$ can be understood that the signal $z(t')$ is doing the Fourier transform near the time t, then called “the local spectrum”.

The wavelet is a smooth and quickly vanishing oscillating function with good localization in both frequency and time. $\psi_{ab}(t)$ is a set of elementary functions generated by dilations and translations of a unique admissible mother wavelet $\psi(t)$:

$$\psi_{ab}(t) = \frac{1}{\sqrt{|a|}}\psi(\frac{t-b}{a}) \tag{2}$$

where $a, b \in R, a \neq 0$, a, b are the scale and translation parameters respectively, and t is the time [16]. In contrast to STFT, first, WT is more applicable when the signal is non-stationary, and STFT can not do the denoising, moreover, wavelet uses a size-adjustable window more advantageous than the fixed window used by STFT [17]. Besides, the complexity of WT is $O(N)$.

2.2 Data Reduction and Noise Reduction

2.1.1 PCA

PCA is defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on a sequenced coordinate. The PCA computes as follows: Given data matrix X^T, the SVD of X is

$$X = W\Sigma V^T \tag{3}$$

where the $m \times m$ W is the matrix of eigenvectors of the covariance matrix XX^T, the matrix Σ is an $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal, and the $n \times n$ matrix V is the matrix of eigenvectors of XX^T. The PCA transformation is given by

$$Y^T = V\Sigma^T \tag{4}$$

2.2.2 Selection of the Subsets

Selecting proper subsets runs with the notion that optimal subsets are those which preserve much of the information during the task, compared with other subsets [18]. 5 steps algorithm is presented as follows:

1. Computing the values that can present the significance it contributes for each channel per subject.
2. Discarding the channels with the least significant value.
3. Analyzing between subject variability.
4. Demonstration of two localized active regions.
5. Running the correlation analysis within the localized active regions

2.2.3 Independent Component Analysis(ICA)

ICA is a method used to separate independent components from a linear mixture of the components, and a simple application of ICA is the "cocktail party problem", where the underlying speech signals are separated from a sample data consisting of people talking simultaneously in a room. The EEG waves recorded can be considered as a linear mixture of the signals from each neuron cluster in the brain and the non-related components from subject's actions such as eye movements. Hence, the ICA can be applied to EEG [19].

The components x_i of the observed random vector

$$x = (x_1, ...x_m)^T \tag{5}$$

are generated as a sum of the independent components s_k, $k = 1, ..., n$. The model is

$$x = \sum_{k=1}^{n} s_k a_k \tag{6}$$

where

$$a_k = (a_{1,k}, ...a_{m,k})^T \tag{7}$$

The basis vectors a_k form the columns of the mixing matrix

$$A = (a_1, ...a_n) \tag{8}$$

where A is a non-singular mixing matrix. Because the EEG signals are considered to be noisy, we assume the zero-mean and uncorrelated Gaussian noise $n \sim N(0, diag(\Sigma))$. The ICA takes the form

$$x = As + n \tag{9}$$

A demixing matrix, W is computed as

$$W = A^T \tag{10}$$

2.3 Feature Selection

For features extracted from the raw EEG data many methods such as time domain, frequency domain, and time–frequency domain are used. For time domain, the mean and the variance of the signal of each electrode can be the features, and then PCA is applied to reduce the features [20]. For frequency domain, scholars use the information in Gamma band after filtering [21], and often they also use the PSD [22], because the power of the EEG signal will decreased with increasing of the frequency and increased with decreasing of the frequency. For time-frequency domain, methods used include both time domain and frequency domain [23-24].

3 Nonlinear Analysis and Algorithm Enhancement

3.1 Information Theory

3.1.1 PE

Entropy is first introduced by Claude Shannon in 1948, and it is a measure of complexity or irregularity of a time series. For signal X , the definition of the entropy is

$$H(X) = -\sum_{i=1}^{n} P(s_i)\log_2[P(s_i)] \tag{11}$$

where $P(s_i)$ is the probability that s_i in signal X .

Bandt has given the PE using ordinal pattern based on Claude Shannon's entropy [25]. This measure resists the noise quite well, and can be applied to data analysis. Given signal X , with the width of the window being k , we can compare the value in the window two by two and then sequence it. After a group of swap, the values of these k numbers are arranged in a new pattern. Thus, for the sequence with a length of k, the number of the sequence pattern is $k!$. The i^{th} sequence pattern is π_i, for a sequence that contains π_i with the length of N is

$$P(\pi_i) = \frac{\{t|\ 0 \le t \le N-k, (x_{t+1},...,x_{t+k})\}}{N-k+1} \tag{12}$$

where $x_{t+1},...,x_{t+k}$ are sequence patterns. Therefore, the PE for k is

$$H(k) = -\sum_{i=1}^{k!} P(\pi_i)\log_2 P(\pi_i) \tag{13}$$

The experiment indicates that the graph of $H(k)$ to k is almost linear, hence the definition of PE for every single character is of meaning. Define that

$$h(k) = H(k)/(k-1) \tag{14}$$

In application, $3 \le k \le 7$.

3.1.2 SVD Entropy

A delay embedding theorem gives the conditions under which a chaotic dynamical system can be reconstructed from a sequence of observations of the state of a dynamical system. Takens' theorem is the delay embedding theorem of Floris Takens. Roberts calculates the SVD entropy based on the embedding theorem and the SVD. Before using the SVD algorithm, the embedding dimension m and time delay v .For signal

$$X = \{x_i, i = 1,2,...,N\} \tag{15}$$

the total amount of time delay in the vector

$$M = N-(m-1)v \tag{16}$$

the nth vector is

$$x_n = (x_n, x_{n+v},..., x_{n+(m-1)v})^T \tag{17}$$

where $n = 1,2...,M$, M column vectors constitute the original space. The embedding matrix is

$$X = (x_1, x_2,..., x_M)^T \tag{18}$$

When the m is big enough, the dynamical system can be reconstructed. Hence, we should use every point properly and make sure the information is not lost. When $v = 1$, $m = 20$, the SVD algorithm is then applied to generate a group of singular values based on equation (18). These singular values are $\{w_i\}, i = 1,2,...m$. Using regularization, it goes to

$$w_i = w_i / \sum_{j=1}^{m} w_j \tag{19}$$

The Definition of the SVD entropy is

$$H_{SVD} = -\sum_{i=1}^{m} w_i \log_2 w_i \tag{20}$$

3.1.3 ApEn

ApEn is a measure that quantifies the complexity or irregularity of a time series. Because short, noisy data sets could be applied to the algorithm, this method is widely used. The larger value of ApEn means more irregularity and complexity of the signal. The procedure of calculating the ApEn is described as follows [26-28].

The original data sets containing N data points are expressed

$$X = [x(1), x(2), ...x(N)] \tag{21}$$

m vectors $X(1), X(2), ..., X(N-m+1)$ are defined by

$$X(i) = [x(i), x(i+1), ..., x(i+m-1)] \tag{22}$$

where

$$i = 1, 2, ..., N-m+1 \tag{23}$$

These vectors represent m consecutive x values, starting with the i^{th} point. The distance between $X(i)$ and $X(j)$ is $d[X(i), X(j)]$, the maximum absolute difference between their respective scalar components. The maximum norm

$$d[X(i), X(j)] = \max_{k=1,...m} | x(i+k-1) - x(j+k-1) | \tag{24}$$

Find the number of $j(j = 1, ..., N-m+1, j \neq i)$ so the $d[X(i), X(j)] \leq r$ denoted as $M^m(i)$. Then

$$C_r^m(i) = \frac{M^m(i)}{N-m+1} \tag{25}$$

The logarithm of each $C_r^m(i)$ is computed and averaged over i.

$$\phi_r^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_r^m(i) \tag{26}$$

The ApEn is computed on the formula

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r) \tag{27}$$

m embedding dimension and a tolerance window r must be specified before executing this algorithm.

3.1.4 SampEn

SampEn examines time series for similar epochs and assigns a non-negative number to the sequence, with larger values corresponding to more irregularity or complexity in the data. This algorithm quite like the ApEn, but with some advantages that the ApEn does not have. It shuns the inconsistency in the data [29]. The SampEn is defined as follows.

m vectors $X_m(1), X_m(2), ..., X_m(N-m+1)$ defined by

$$X_m(i) = [x(i), x(i+1), ..., x(i+m-1)] \tag{28}$$

where $1 \leq i \leq N-m+1$. These vectors represent m consecutive x values, starting with the i^{th} point. The distance between $X_m(i)$ and $X_m(j)$ is $d[X_m(i), X_m(j)]$ the maximum absolute difference between their respective scalar components. The maximum norm

$$d[X_m(i), X_m(j)] = \max_{k=0,...m-1} | x(i+k) - x(j+k) | \tag{29}$$

Find the number of $j(j = 1,..., N-m, j \neq i)$, denoted as B_i, such that the distance between $X_m(i)$ and $X_m(j)$ is less than or equal to r. For $1 \leq i \leq N-m$

$$B_i^m(r) = \frac{1}{N-m+1} B_i \tag{30}$$

set $B^m(r)$ as

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r) \tag{31}$$

The SampEn is then computed as

$$SampEn(m,r) = \lim_{N \to \infty} [-\ln \frac{B^{m+1}(r)}{B^m(r)}] \tag{32}$$

When N is finite

$$SampEn(m,r) = -\ln \frac{B^{m+1}(r)}{B^m(r)} \tag{33}$$

3.2 Chaos Theory

3.2.1 CD

CD measures the complexity in attractors in the space. It exploits the correlation integral to compute the correlation for variables, and then the certainty of the signals are measured. According to Takens, we can use a time series data to reconstruct the space, and then use a tiny ball to search the closest distance between two points. Two points are correlated if the distance is smaller than the given threshold, otherwise they are not. More pairs of the points indicate a higher correlation. The CD is computed as follows.

First, the correlation integral $C(r)$ is computed.

$$C(r) = \frac{2}{(M-W)(M-W-1)} \sum_{i=1}^{M} \sum_{j=i+W}^{M} \Theta(r - || x_i - x_j ||) \tag{34}$$

where r is the threshold

$$M = N - (m-1)v \tag{35}$$

is the m dimension embedding parameter, W is the Theiler window, Θ is the Heaviside function. Then, the CD can be calculated as

$$CD = \lim_{r \to 0} \frac{d \log C(r)}{d \log r} \tag{36}$$

When r is small, CD is the slope of $\lg C(r) \sim \lg r$ in the straight line.

3.2.2 LLE

In mathematics the Lyapunov exponent or Lyapunov characteristic exponent of a dynamical system is a quantity that characterizes the rate of separation of infinitesimally close trajectories. One of the most useful approaches is the LLE. Using the method given by Kants[30].

$$S(\Delta n) = \frac{1}{N} \sum_{n_0=1}^{N} \ln[\frac{1}{|U_r(x_{n_0})|_{x_{n_0}}} \sum_{\in U_r(x_{n_0})} | x_{n_0+\Delta n} - x_{n+\Delta n} |] \tag{37}$$

$S(\Delta n)$ is the effective expansion rate. For some Δn, $S(\Delta n)$ shows a clear linear increase. The slope of $S(\Delta n) \sim \Delta n$ is the LLE.

3.2.3 Hurst Exponent

The Hurst exponent is referred to as the "index of dependence," or "index of long-range dependence." It quantifies the relative tendency of a time series either to regress strongly to the mean or to cluster in a direction. Many ways to compute the Hurst exponent, a commonly used method is R/S method, which changes the range of the value, making the data comparable. The Hurst exponent is defined as

$$H = \log(R/S)/\log(N) \tag{38}$$

where N is the length of the time sequence, R/S changes the range. $H = 0.5$ indicates that the sequence is a random one. $0.5 < H < 1$ indicates a time series with long-term positive autocorrelation while $0 < H < 0.5$ indicates time series with long-term switching between high and low values in adjacent pairs.

4. Discussion and Prospect

For linear characteristic methods, scholars have tried different domain analyses in order to obtain a better result. No final conclusion has yet been reached on this matter, because many factors need to be considered while different approaches have advantages in different aspects. Another aspect that scholars could work on is the feature selection. Until now, selected features do not show a decisive distinction between the alcoholic and the control. Selected features do not represent all the information, perhaps a good way to solve this problem is using a hybrid algorithm that considers more factors available, after which, different weights are given to different factors.

For nonlinear characteristic methods, the selection of the embedding parameter has a decisive influence on the performance. For embedding dimension, if it is too small, the attractor of the system cannot be fully unfolded, and if it is too large, it increases the complexity. For time delay, if it is too small, attractors in the reconstruction will converge near the diagonal thus lead to redundancy, and if it is too large, state-space reconstruction is diffused [31]. A good way of selecting proper parameter for nonlinear analysis is Liangyue Cao algorithm [32]. Though it does not give the most proper delay, it gives a relatively proper delay in order to obtain the embedding dimension. Since the selection of the delay affects the result sometimes more than the embedding dimensions [33].

For all these methods, one problem is that computational demands. Usually, the data for applying these methods are huge, and the time is then an important factor that should be considered. Second, noise effects also need to be addressed, because the data we collected may not be 100% pure signals. Solutions for noise effect are: data on a scale larger than most of the noise, using a low-pass filter, separating noise from deterministic effects to calculate marginal redundancy and mutual information, and PCA [34].

5. REFERENCES

[1] Walker, Jonathan E., and Gerald P. Kozlowski. "Neurofeedback treatment of epilepsy." *Child & Adolescent Psychiatric Clinics of North America* 14.1 (2005): 163-176.

[2] Weikel, Jutta C., et al. "Ghrelin promotes slow-wave sleep in humans." *American Journal of Physiology-Endocrinology And Metabolism* 284.2 (2003): E407-E415.

[3] Díaz, G. F., et al. "Generalized background qEEG abnormalities in localized symptomatic epilepsy." *Electroencephalography and clinical neurophysiology* 106.6 (1998): 501-507.

[4] Mervaala, Esa, et al. "Visual evoked potentials, brainstem auditory evoked potentials, and quantitative EEG in Baltic progressive myoclonus epilepsy." *Epilepsia* 27.5 (1986): 542-547.

[5] Morisson, Florence, et al. "Daytime sleepiness and EEG spectral analysis in apneic patients before and after treatment with continuous positive airway pressure." *CHEST Journal* 119.1 (2001): 45-52.

[6] Misulis, Karl E. *Spehlmann's evoked potential primer: visual, auditory, and somatosensory evoked potentials in clinical diagnosis*. Vol. 994. Boston:: Butterworth-Heinemann, 1994.
[7] Richman, Joshua S., and J. Randall Moorman. "Physiological time-series analysis using approximate entropy and sample entropy." *American Journal of Physiology-Heart and Circulatory Physiology* 278.6 (2000): H2039-H2049.
[8] Di, Wu, et al. "Study on human brain after consuming alcohol based on EEG signal." *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*. Vol. 5. IEEE, 2010
[9] Sun, Yuge, Ning Ye, and Xinhe Xu. "EEG analysis of alcoholics and controls based on feature extraction." Signal Processing, 2006 8th International Conference on. Vol. 1. IEEE, 2006.
[10] Adeli, Hojjat, and Samanwoy Ghosh-Dastidar. Automated EEG-based Diagnosis of Neurological Disorders: Inventing the Future of Neurology. CRC Press, 2010.
[11] Sankari, Ziad, Hojjat Adeli, and Anahita Adeli. "Intrahemispheric, interhemispheric, and distal EEG coherence in Alzheimer's disease." *Clinical Neurophysiology* 122.5 (2011): 897-906.
[12] Ahmadlou, Mehran, Hojjat Adeli, and Amir Adeli. "Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder." *Journal of Clinical Neurophysiology* 27.5 (2010): 328-333.
[13] Palaniappan, R., S. Anandan, and P. Raveendran. "Two level PCA to reduce noise and EEG from evoked potential signals." *Control, Automation, Robotics and Vision, 2002. ICARCV 2002. 7th International Conference on*. Vol. 3. IEEE, 2002
[14] Shahina, A., et al. "Autoassociative neural networks for discrimination of chronic alcoholics using Visual Evoked Potentials." *Computing, Communication and Networking, 2008. ICCCn 2008. International Conference on*. IEEE, 2008
[15] Adeli, Hojjat, Ziqin Zhou, and Nahid Dadmehr. "Analysis of EEG records in an epileptic patient using wavelet transform." *Journal of neuroscience methods* 123.1 (2003): 69.
[16] Akin, M. "Comparison of wavelet transform and FFT methods in the analysis of EEG signals." *Journal of Medical Systems* 26.3 (2002): 241-247.
[17] Giurgiutiu, Victor, and Lingyu Yu. "Comparison of short-time fourier transform and wavelet transform of transient and tone burst wave propagation signals for structural health monitoring." *Proceedings of 4th International Workshop on Structural Health Monitoring*. 2003
[18] Shooshtari, M. Alavash, and S. Kamaledin Setarehdan. "Selection of optimal EEG channels for classification of signals correlated with alcohol abusers." *Signal Processing (ICSP), 2010 IEEE 10th International Conference on*. IEEE, 2010
[19] Joshua, Lau Kah Liong, and Jagath C. Rajapakse. "Extraction of event-related potentials from EEG signals using ICA with reference." *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*. Vol. 4. IEEE, 2005
[20] Yazdani, Ashkan, and S. Kamaledin Setarehdan. "Classification of EEG signals correlated with alcohol abusers." *Signal Processing and Its Applications, 2007. ISSPA 2007. 9th International Symposium on*. IEEE, 2007
[21] Yazdani, Ashkan, et al. "Neural, Fuzzy And Neurofuzzy Approach To Classification Of Normal And Alcoholic Electroencephalograms." *Image and Signal Processing and Analysis, 2007. ISPA 2007. 5th International Symposium on*. IEEE, 2007
[22] Shri, P. T. K., and N. Sriraam. "EEG based detection of alcoholics using spectral entropy with neural network classifiers." *Biomedical Engineering (ICoBE), 2012 International Conference on*. IEEE, 2012
[23] Liu, Jihong, and Jing Xu. "Compared study of the analyzing methods for EEG data." *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*. Vol. 9. IEEE, 2010
[24] Kousarrizi, MR Nazari, et al. "Classification of Alcoholics and Non-Alcoholics via EEG Using SVM and Neural Networks." *Bioinformatics and Biomedical Engineering, 2009. ICBBE 2009. 3rd International Conference on*. IEEE, 2009
[25] Bandt, Christoph, and Bernd Pompe. "Permutation entropy: A natural complexity measure for time series." Physical Review Letters 88.17 (2002): 174102.
[26] Pincus, Steven M. "Approximate entropy as a measure of system complexity." Proceedings of the National Academy of Sciences 88.6 (1991): 2297-2301.
[27] Guo, Ling, Daniel Rivero, and Alejandro Pazos. "Epileptic seizure detection using multiwavelet

transform based approximate entropy and artificial neural networks." Journal of neuroscience methods 193.1 (2010): 156-163.

[28] Kumar, Yatindra, and M. L. Dewal. "Complexity measures for normal and epileptic EEG signals using ApEn, SampEn and SEN." Int. J. Comput. Commun. Technol.(IJCCT) 2.1 (2011): 6-12.

[29] (Chinese) Yang Z, Duoqian M, Deng W, "The research of the sample entropy between the alcoholic and the control", Chinese Journal of Biomedical Engineering, Vol. 29 No.6 0258-8021(2010)06-0939-04

[30] Camm, A. John, et al. "Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology." Circulation 93.5 (1996): 1043-1065..

[31] (Chinese)Jinfeng F, Chenxi S, Jian W, Ming Y, Zicai W, "A Comparative Evaluation of Nonlinear Characteristics of EEG for Alcoholic and Normal People" Chinese Journal of Biomedical Engineering, Vol.27 No.1, 0258-8021(2008)01-0018-05

[32] Cao, Liangyue. "Practical method for determining the minimum embedding dimension of a scalar time series." Physica D: Nonlinear Phenomena 110.1 (1997): 43-50.

[33] Xufa, SHAO Chenxi SHEN Linfeng WANG, and Y. I. N. Shijie. "Nonlinear analysis of the alcoholic's EEG." Progress in Natural Science 12 (2002): 006.

[34] Gladstone, Igor M., and Richard A. Ehrenkranz. "A regularity statistic for medical data analysis." Journal of clinical monitoring 7.4 (1991): 335-345.

[35] Gladstone, Igor M., and Richard A. Ehrenkranz. "A regularity statistic for medical data analysis." Journal of clinical monitoring 7.4 (1991): 335-345.

[36] Kumar, Yatindra, M. L. Dewal, and R. S. Anand. "Features extraction of EEG signals using approximate and sample entropy." *Electrical, Electronics and Computer Science (SCEECS), 2012 IEEE Students' Conference on*. IEEE, 2012.

[37] Palaniappan, R., and P. Raveendran. "Single trial VEP extraction using digital filter." *Statistical Signal Processing, 2001. Proceedings of the 11th IEEE Signal Processing Workshop on*. IEEE, 2001.

[38] Wang, H., et al. "Effect of alcohol-dependent EEG on the traffic signal recognition." *Information Technology and Applications in Biomedicine, 2008. ITAB 2008. International Conference on*. IEEE, 2008.

[39] Sun, Yuge, Ning Ye, and Xinhe Xu. "EEG analysis of alcoholics and controls based on feature extraction." *Signal Processing, 2006 8th International Conference on*. Vol. 1. IEEE, 2006.

[40] Ong, Kok-Meng, et al. "Selection of a Subset of EEG Channels using PCA to classify Alcoholics and Non-alcoholics." *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*. IEEE, 2006.

[41] ACHARYA, U. RAJENDRA, et al. "Automated diagnosis of normal and alcoholic EEG signals." *International Journal of Neural Systems* 22.03 (2012).

[42] Sharmilakanna, P., and Ramaswamy Palaniappan. "EEG artifact reduction in VEP using 2-stage PCA and N4 analysis of alcoholics." *Intelligent Sensing and Information Processing, 2005. ICISIP 2005. Third International Conference on*. IEEE, 2005.