

Sleep EEG analysis Based on the Multiscale Jenson-Shannon Divergence

Zhengxia Zhang¹, Jia-Fei Dai^{2,a}, JunWang^{1,b} and Feng-Zhen Hou^{3,c}

¹Image Processing and Image Communications Key Lab. , Nanjing Univ. of Posts & Telecomm. , Nanjing 210003, China

²Nanjing General Hospital of Nanjing Military Command, Nanjing, 210002, China

³School of Science, China Pharmaceutical University, Nanjing 210009, China

^aE-mail: 37045613@qq.com, ^bwangj@njupt.edu.cn, ^cE-mail: houfz@cpu.edu.cn

Keywords: multi-scale; JSD; sleep EEG signals.

Abstract. Sleep EEG signals analysis is a hotspot of research recently, this paper, by using nonlinear dynamics theory knowledge, JSD algorithm and multi-scale JSD algorithm is proposed for some individual conscious period and NREM sleep stage I analyzed the research of EEG signals, and the use of SPSS statistical software to verify the veracity and reliability of the experiment, at the same time, with the error bar graph method to analysis the two different states of sleep EEG signals, the results show that both the JSD algorithm and the multi-scale JSD algorithm can effectively distinguish between awake and NREM sleep stage I of EEG signals, these two conditions' EEG signals exist significant differences, The algorithm we proposed can be further used in the study of sleep EEG in installment, which can also provide all kinds of disease diagnosis and treatment of sleep with effective auxiliary function, the research has important practical significance in the future.

Introduction

As we all known, each person needs to sleep every day, a third time of people's life is spent in sleep. As a necessary process of life, sleep is an important part of the body rehabilitation, integration and consolidate memories, is an integral part of health. Through sleep, the fatigue nerve cells will be in the normal physiological function, and the body will be in the state of mental and physical restoration. Pituitary growth hormone secretion during sleep increased significantly, is beneficial to promote the growth of the body, and increase the nucleoprotein synthesis, is advantageous to the memory storage [1, 2]. In general, the R&K is used as the rule of sleep EEG in installment internationally, according to the performance of sleep EEG signals, the sleep period is divided into: awakened period, non-rem sleep (Non - rapid eye movement, NREM) sleep and REM sleep (the rapid eye movement, REM). The NREM can be divided into four stages, such as stage I, II, III, and IV period. Sleep staging studies, the analysis of sleep and sleep quality of scientific evaluation, which all have important application values.

At present, many researchers are committed to the research work of sleep EEG, different research methods, with the development of the nonlinear dynamics theory constantly perfect, has been gradually applied to the analysis of the medical signal processing, including the main research methods used in sleep EEG are permutation and combination entropy, correlation dimension, approximate entropy, neural networks etc. Since Jensen-Shannon Divergence (JSD) [1] (which was used to measure the difference between the probability distribution of random variables) was proposed in 1991, it was widely applied to the symbol sequence analysis and characterization [2], such as pattern recognition [3], DNA sequence segmentation [4]. JSD is the result of symmetrizing and smoothing the Kullback-Leibler Divergence (KLD). The non-negativity, symmetry, continuity and boundness features of JSD have been widely used in the analysis of time series. Electrocardiogram (ECG) can also be seen as a time series, so we considered to use the analysis method which is based on the complexity of the JSD to study coupling degree between the sleep

EEG time series.

The principle

Jensen-Shannon Divergence(JSD)[3,4]

We denoted p_1, p_2 as two probability distribution of discrete random variables X while (1) stands for the KULLBACK direct difference I .

$$I(p_1, p_2) = \sum_{x \in X} p_1(x) \log \frac{p_1(x)}{p_2(x)} \quad (1)$$

From (1) we can see that KLD [5] has non-negativity and incremental feature, but it does not have symmetry. Different symmetric versions of the KL divergence measure exist, according to various ways to meet this point J-divergence which was introduced as:

$$\begin{aligned} J(p_1, p_2) &= I(p_1, p_2) + I(p_2, p_1) \\ &= \sum_{x \in X} (p_1(x) - p_2(x)) \log \frac{p_1(x)}{p_2(x)} \end{aligned} \quad (2)$$

The distance change with the metric nature was defined as follows:

$$V(p_1, p_2) = \sum_{x \in X} |p_1(x) - p_2(x)| \quad (3)$$

The minimum value of the direct differences I had been found, since long years of study of the relationship between the direct differences I and distance changes V

$$I(p_1, p_2) \geq \max\{L_1(V(p_1, p_2)), L_2(V(p_1, p_2))\} \quad (4)$$

Where,

$$L_1(V(p_1, p_2)) = \log \frac{2 + V(p_1, p_2)}{2 - V(p_1, p_2)} + \frac{2V(p_1, p_2)}{2 + V(p_1, p_2)}, 0 \leq V(p_1, p_2) \leq 2 \quad (5)$$

$$L_2(V(p_1, p_2)) = \frac{V^2(p_1, p_2)}{2} + \frac{V^4(p_1, p_2)}{36} + \frac{V^6(p_1, p_2)}{288}, 0 \leq V(p_1, p_2) \leq 2 \quad (6)$$

But there are some limitations when measuring the difference between the probability distribution, as we can't come out a general representation of the maximum value for direct differences I and differences J based on distance changes V . In order to generally describe the maximum limit of the difference, the amended definition of I and J which was introduced in [6] accordingly:

$$\begin{aligned} K(p_1, p_2) &= \\ \sum_{x \in X} p_1(x) \log \frac{p_1(x)}{\frac{1}{2} p_1(x) + \frac{1}{2} p_2(x)} &= I(p_1, \frac{1}{2} p_1 + \frac{1}{2} p_2) \end{aligned} \quad (7)$$

The corresponding symmetrical form difference was as follows:

$$L(p_1, p_2) = K(p_1, p_2) + K(p_2, p_1) \quad (8)$$

Now the difference K and L is non-negativity[7], incremental, and have a statistical feature of limitation and semi-bounded. That is:

$$\begin{aligned} K(p_1, p_2) &< +\infty, K(p_1, p_2) \geq K(p_1, p_1) \\ L(p_1, p_2) &< +\infty, L(p_1, p_2) \geq L(p_1, p_1) \end{aligned} \quad (9)$$

$$L(p_1, p_2) = 2H\left(\frac{p_1 + p_2}{2}\right) - H(p_1) - H(p_2) \quad (10)$$

Now setting the weight of two probability distributions (p_1, p_2) as π_1 and π_2 accordingly, which

meeting the statistical condition of constraint $\pi_1, \pi_2 > 0, \pi_1 + \pi_2 = 1$, Jensen-Shannon Divergence (JSD) [8] was defined as:

$$\begin{aligned} JS_{\pi}(p_1, p_2) \\ = H(\pi_1 p_1 + \pi_2 p_2) - \pi_1 H(p_1) - \pi_2 H(p_2) \end{aligned} \quad (11)$$

Associating to the probability distribution $p = \{P_j; j=1, 2, \dots, n\}$ the corresponding weight were $\pi_1, \pi_2, \dots, \pi_n$ respectively, the statistical complexity measure JSD was defined as:

$$JS_{\pi}(p_1, p_2, \dots, p_n) = H\left(\sum_{i=1}^n \pi_i p_i\right) - \sum_{i=1}^n \pi_i H(p_i) \quad (12)$$

Data processing and analysis

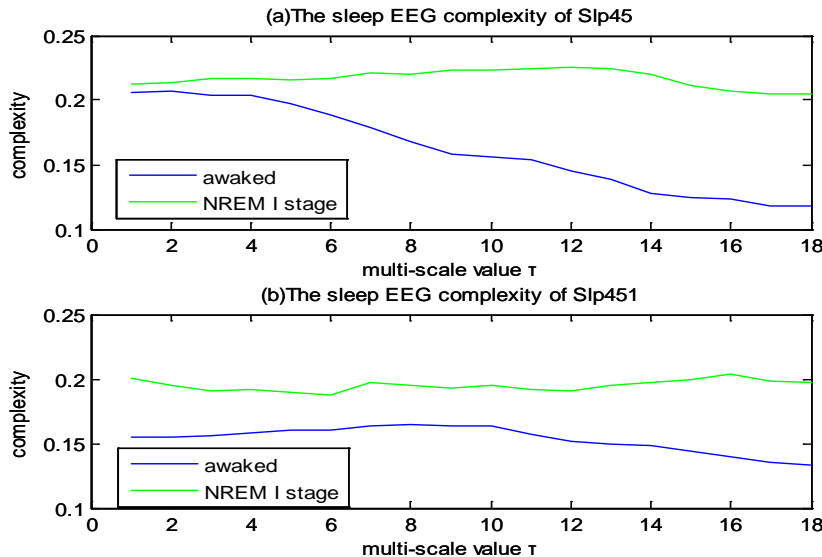
The experimental data

The data of sleep EEG signals used in this paper from the PhysioBank MIT- BIH Polysomnographic Database [9]. Data records of sleep EEG signals in the library have many parameters, including one EEG signals, one ECG signals and so on, the data sampling frequency is 250 Hz. In this paper, we adopts the subjects Slp45, Slp59, Slp48, Slp451, Slp481, Slp591, using one guide sleep EEG (C3 - A1) signals of every person, and extracting several groups sleep EEG data from the awaked periods and NREM I stage for analysis research.

Sleep EEG analysis based on the JSD

(1) Firstly, we calculate the statistical complexity of sleep EEG signals based on Jensen-Shannon Divergence, hoping to verify that the method can well distinguish between awaked period and NREM I stage of sleep EEG signals, and through the study process tell us that the two different sleep EEG time series exist significant differences. This method also demonstrated that the JSD algorithm can be used to analyze sleep EEG signals, the statistical complexity can be used as a measure of brain function parameters, the research can be applied to sleep EEG signals staging, help people to treat the various sleep diseases and get a good sleep in the future.

(2) Secondly, the data processing by the tool of MATLAB software, the concrete process is: calculating the statistical complexity of sleep EEG signals based on Jensen - Shannon Divergence, the data is extracted from the above database. Thinking about the differences of individual EEG signals, we just calculate the statistical complexity of subjects Slp45, Slp48, Slp451, and Slp481 their sleep EEG signals in awaked period and NREM I stage, the length of the data is 5000, the result is shown in figure 1 (a), (b), (c), (d) as below:



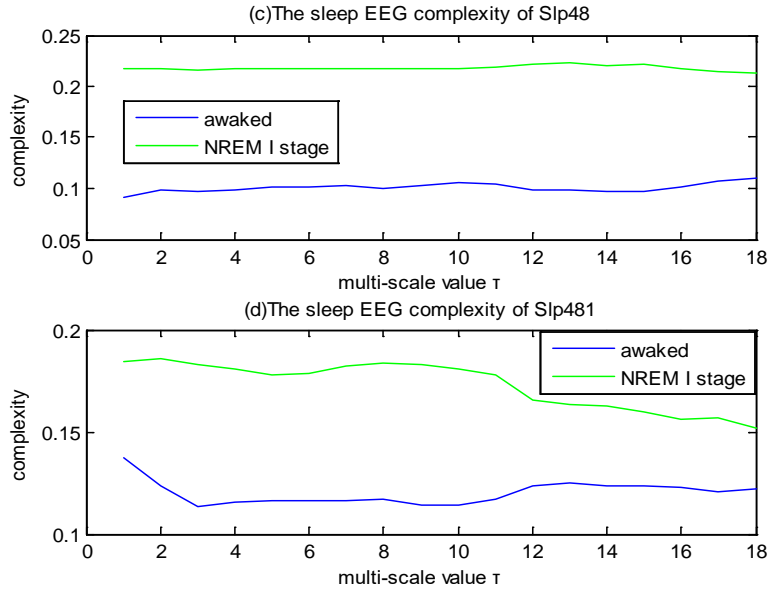


Figure.1. (a), (b),(c),(d) are the graph of the statistical complexity of two subjects , the sleep EEG signals both in awaked period and NREM I stage

The results from the above we can see that although individual EEG signals exist significant difference, we can clearly distinguish the difference of the same individual sleep EEG signals in awaked period and NREM sleep I by JSD algorithm, and find that the sleep EEG under different period exists significant difference. The JSD algorithm can be applied to sleep EEG signals' research.

(3) Statistical analysis and hypothesis testing. In order to further verify the reliability of the above research results and its accuracy, for Slp45, Slp451, Slp48, Slp481 subjects statistical complexity results for Slp45, Slp451, Slp48, Slp481 subjects statistical complexity results, we make an independent sample t-test by the SPSS statistical software , the results are got. The first sig value is less than 0.05, which means the sleep EEG signals of different periods does not meet the homogeneous variance. The second sig in the second row is also less than 0.05, shows two sets of data exists significant difference; The situation shows that two groups of data exist significant difference; The result of independent samples t test tell us that the JSD algorithm could distinguish two different period sleep EEG signals well, the algorithm is used to distinguish the sleep EEG signals are valid.

Sleep EEG analysis based on the multi-scale JSD

On the basis of the above JSD algorithm we proposed the multi-scale JSD, so-called multi-scale method is developed on the basis of the average method of a kind of approximate analytical methods. In order to improve the calculation accuracy of average method, the time scale can be divided into more refined, thereby. Using multi-scale JSD algorithm to calculate the statistical complexity results will be more precise, considering the differences of individual EEG signals, for subjects Slp45, Slp59, Slp451, and Slp591, we calculate the statistical complexity of sleep EEG signals in awaked period and NREM stage I, the length of the data is 5000, The statistical complexity is presented with error bar graph in figure 2 (a), (b), (c), (d) shown below:

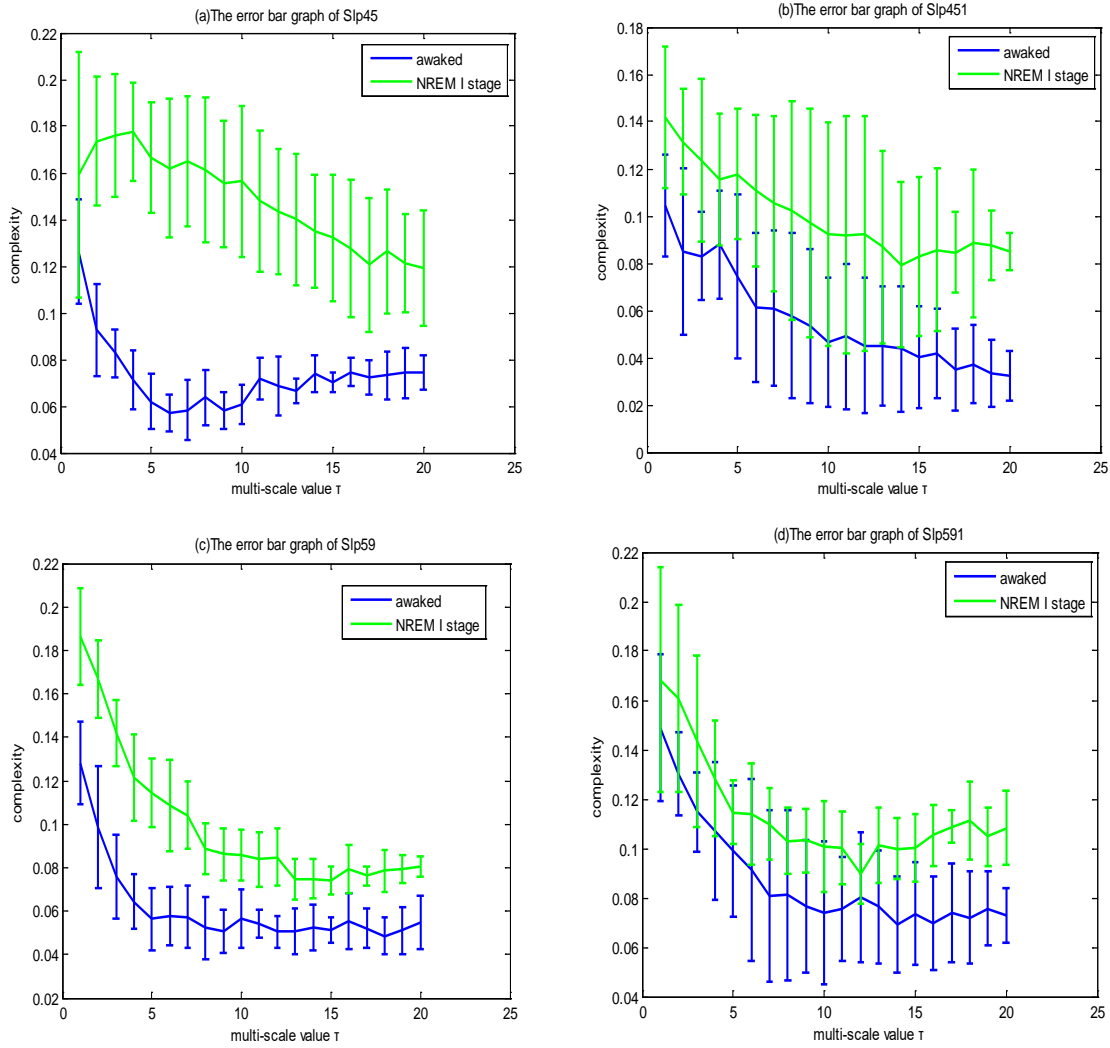


Figure.2. (a), (b),(c),(d) are the error bar graph of the statistical complexity of two subjects , the sleep EEG signals both in awaked period and NREM I stage

As can be seen from the above results, the multi-scale JSD algorithm can distinguish the sleep EEG signals between awaked period and NREM I stage. Thus there exists significant differences between the sleep EEG in awaked period and the sleep EEG in also awake period and NREM I stage .The algorithm of multi-scale JSD in distinguish between EEG signals has certain reliability and accuracy.

Conclusions

The JSD algorithm and the multi-scale JSD algorithm proposed in this paper can distinguish the sleep EEG signals between awaked period and NREM sleep stage I. There exists significant differences between the sleep EEG in awaked period and the sleep EEG in also awake period and NREM I stage .We hope that this method can be used in the study of sleep EEG in installment, which can also provide all kinds of disease diagnosis and treatment of sleep with effective auxiliary function, help more people have a good sleep.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant Nos. 61271082, 61401518), Jiangsu Provincial Key R & D Program (Social Development) (Grant No. BE2015700), the Natural Science Foundation of Jiangsu Province (Grant No. BK20141432), the Foundation of Nanjing General Hospital of Nanjing Military Command (Grant No. 2014019), Nanjing Medical

Science and Technology Project (Grant No. 201503009) and the Fundamental Research Funds for the Central Universities (Grant No. FY2014LX0039).

References

- [1] <http://www.trhos.com/sleep/shuimianxue.htm>.
- [2] <http://www.windrug.com/pic/25/11/10/31/003>.
- [3] Lin J. Divergence Measures Based on the Shannon Entropy[J]. IEEE transaction on information theory, 1991, 37 (1):145-151.
- [4] Lamberti P W, Majtey A P. Non-logarithmic Jensen–Shannon divergence[J]. Physica A, 2003, 329(1-2):81-90.
- [5] M.T.Martin, A.plastino, O.A.Rosso. Generalized statistical complexity measures: Geometrical and analytical properties[J]. Physica A, 2006, 369:439-462
- [6] O.A.Rosso, H.A.Larrondo, M.T.Martin, et al. Distinguishing Noise from Chaos[J]. Physical review letters, 2007, 99, 154102.
- [7] Plastino A, Lamberti PW, Martin MT, et al. Intensive entropic non-triviality measure[J]. Physica A, 2004, 334 (1-2):119-131.
- [8] Qi J X, Bian Z Z, Yang Q. A microcomputer-based realtime automatic analysis system for ECG signal[J]. Beijing biomedical engineering, 1997, 16(3):157-160.
- [9] SHHS Polysomnography Database [DB/OL]. <http://www.physionet.org/physiobank/database/shhpsgdb/2012.6.10.base>.