Analysis of Epileptic and Normal EEG Signals based on Random Walk

Jun MIN[1,a], Jun WANG[2,b*]

[1]School of Telecommunications and Information Engineering,Nanjing University of Posts and Telecommunications,Nanjing 210003,Jiangsu,China

[2]School of Geography and Biological Information, Nanjing University of Posts and Telecommunications,Nanjing 210023, Jiangsu,China

[a]1396705817@qq.com, [b]wangj@njupt.edu.cn

*Jun Wang

Abstract. In this paper, a random walk model was established for the EEG signal. The characteristic parameters qt and at were extracted from the original EEG signal by Bayesian inference method. By comparing the global maximum of the autocorrelation function of the characteristic parameter qt between patients with epilepsy and normal subjects, this paper obtained the following conclusion, the global maximum of the autocorrelation function of the characteristic parameter qt of EEG signal extracted in epilepsy patients is larger than the normal human in the general trend; and the fluctuation of the global maximum of the autocorrelation function in epilepsy patients is also greater, which suggests that the use of random walk model to analyze the EEG signal can be found the difference in patients with epilepsy and normal subjects; therefore, random walk model can be used to analyze the difference between epileptic EEG signal and normal human EEG signals.

Introduction

Epilepsy is a chronic neurological disorders, and it is due to sudden onset of brain neurons abnormal discharge, so that the central nervous system activity is abnormal excitement or inhibition, which has a great impact on the psychological and physical aspects of epilepsy patients. EEG is the basic tool for the diagnosis of epilepsy. It is very important for physiology research and medical application and it is the response of the cerebral cortical or scalp surface brain tissue to the total electrical activity. It contains a lot of brain activity and function information. At present, the common methods of EEG signal processing are: time-frequency analysis[1], artificial neural network[2], nonlinear dynamics[3], independent component analysis[4].

Random walk[5] is an important part of the random process. It is an irregular form of change, in this process of change, every step of the changes are purely random, in our lives, the random walk associated with the natural phenomena, such as the movement of gas molecules, Drops of ink into water, the spread of smell and so on. At present, random walks have been used in many areas, financial analysis, image segmentation and data mining and other aspects.

In this paper, we construct a random walk model[6] for the EEG signal of epilepsy patients and normal people, and use the Bayesian inference[7] to get the characteristic parameter from the EEG signal, and compare the difference between epilepsy patients and normal subjects. The experimental result show that: compare with normal subjects, the global maximum of the autocorrelation function of the characteristic parameter qt is greater, the amplitude of the fluctuation is greater and is more unstable, this show the difference of EEG signal between patient with epilepsy and normal subjects, which suggest that random walk can be used as a method to analyze EEG signal in patients with epilepsy and normal subjects.

Basic Principle

Random walk

Random walk is a mathematical statistical model, known as a stochastic or random process, which is composed of a series of trajectories, each of which is random. It can be used to express irregular changes in the form of random, the core concept of random walk is that the conserved quantity of any random walker corresponds to a law of diffusion, it is similar to the Brownian movement, and it is an ideal mathematical state of Brownian movement. The random walk formula used in this paper is as follows:

$$u_t = q_t u_{t-1} + a_t n_t \tag{1}$$

In Formula (1), u_t is a two-dimensional vector $u_t=(u_{xt}, u_{yt})$, and in this paper, u_{xt} is the first guide of EEG signal, and u_{yt} is the second guide of EEG signal. n_t is also a two-dimensional vector $n_t=(n_{xt}, n_{yt})$, it is a normal distributed, uncorrelated random noise with unit variance. The Parameter (q_t, a_t) represent the characteristics of the two-dimensional vector u_t. In order to get the parameter (q_t, a_t), we use the method of Bayesian inference to get the parameter from u_t.

Bayesian inference

Bayesian inference is a method of deducing statistics. This approach uses the Bayesian theorem, and when there is more evidence and information, update the probability of a particular hypothesis, Bayesian inference is one of the most important techniques in statistics. Bayesian updates are particularly important in sequence analysis. Bayesian inference is used in many fields, including science, engineering, medicine, and so on.

In this paper, in order to extract the parameter q_t and a_t from u_t in formula (1), we first get the time-dependent joint probability density $p(q_{t+1}, a_{t+1})$ of parameter q_t and a_t using sequential Bayesian updating, and then get the q_t and a_t by calculate the mean of the joint probability density $p(q_t, a_t)$. we start at time t = 0 with a flat prior distribution $p_0(q, a)$, which can be interpreted as a "first guess" about the parameter values, in this paper, we give the first guess value $p_0(q, a) = 1/40000$. From the measured successive vector u_0 and u_1, we calculate the likelihood distribution $L_1(q, a)$, which provides a first information about probable parameter values. And obtain the posterior distribution $p_0 L_1$, and continue on this step, we get the posterior distributions in formula (2):

$$p_t = p_{t-1}L_t \tag{2}$$

K is a transformation that both gradual and abrupt parameter changes can be identified. And at last, we perform the same sequential parameter inference in the reverse time direction and combine both distributions. With the iterative multiplication of the formula(2), the posterior distribution would yield an increasingly accurate estimate result.

In formula (2), we use the discretized probability distributions to describe posterior distribution p_t, based on equally spaced parameter values $q_i, a_j (i \in \{1,2,\ldots, N_q\}, j \in \{1,2,\ldots, N_a\})$, a distribution $p(q, a)$ can be approximated by $N_q \times N_a$-dimensional matrix: $(p(q, a))_{ij} = p(q = q_i, a = a_j)$. the multiplication of two distributions is thus reduced to the element-wise multiplication of two matrices, in the paper, we set $N_q = 200$ and $N_a=200$, so the discretized probability distribution is the 200×200-dimensional matrix.

As for L_t in formula (2), we use the likelihood function, we set $L_{t+1} = p(u_{t+1} \mid q_{t+1}, a_{t+1}; u_t)$ that describes the probability of observing a certain measurement u_{t+1}, given the value of the latent parameters. The likelihood function is given by formula (3):

$$L_{t+1} = \frac{1}{(2\pi a_{t+1}{}^2)^{d/2}} \exp\left(-\frac{(u_{t+1}-q_{t+1}u_t)^2}{2a_{t+1}{}^2}\right) \tag{3}$$

in formula (3), d is the number of the dimension of the u_t(d is two in this paper).

Analysis of EEG Signals Based on Random Walking

Experimental data

In this paper, the EEG signals of normal subjects and patients with epilepsy were collected from clinical diagnosis by the General hospital of Nanjing Military Region. The recorded length was more than 1 min and the sampling frequency was 512Hz. The library records multi-parameter EEG data, including 16 guide signals. EEG data of 21 healthy subjects and EEG data of 21 patients with epilepsy were selected in this paper. And in the experiment, the EEG data of the first guide and the second guide were select for analyzing.

Data preprocessing

In this paper , we take the first 10000 points of the EEG data, and then divide the EEG data into 10 segment, sub-section followed by [1-1000], [1001-2000], [2001-3000], [3001-4000], [4001-5000], [5001-6000], [6001-7000], [7001-8000], [8001-9000], [9001-10000], and then process each segment of the data.

Experimental process

In this paper , we firstly construct the random walk model with the formula like equation (1), and for the vector $u_t=(u_{xt}, u_{yt})$, u_{xt} is the first guide of EEG , u_{yt} is the second guide of the EEG. Then using Bayesian updating to get the characteristic parameter (q_t, a_t). because we divided the EEG data into 10 segments, so we obtain 10 parameter sequences of $\{q_{ti}, 1 \leq i \leq 10\}$ and $\{a_{ti}, 1 \leq i \leq 10\}$. In order to decrease the abrupt change of q_t,we use the method to of average q_t, firstly, we use the first q_{t1}, calculate the autocorrelation function of q_{t1}, and get the global maximum of the autocorrelation function. And then get the mean q_t' of the first three $\{q_{t1}, q_{t2}, q_{t3}\}$, calculate the autocorrelation function of q_t', and obtain the global maximum of the maximum. third, get the mean q_t' of the five $\{q_{t1}, q_{t2}, q_{t3}, q_{t4}, q_{t5}\}$, calculate the autocorrelation function of q_t', and obtain the global maximum of the autocorrelation. Finally, using the ten result of q_t, get the means q_t' of the q_t, and calculate the global maximum of the autocorrelation function of q_t'. From this four result , We can find more and more accurate experimental results.

Experimental results and analysis

Using the above four steps of calculation, we calculate the result of all EEG data, which contains 21 normal subjects and 21 patients with epilepsy. And compare the maximum of their autocorrelation functions of q_t'. the experimental results are shown in Figure 1- Figure 4.
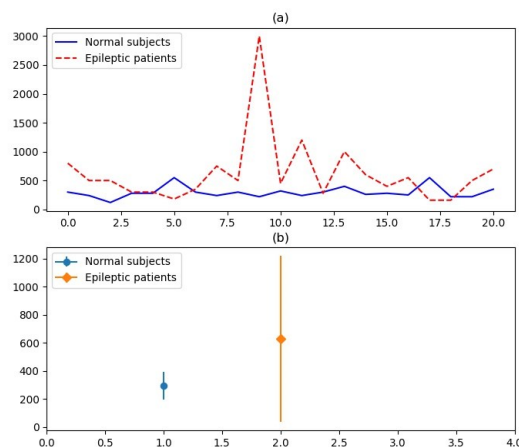

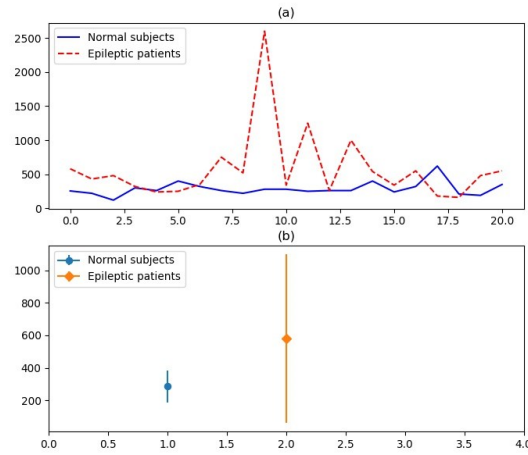Fig. 1. Use the first q_t, calculate global maximum of the autocorrelation function of q_t.

Fig. 2. Use the first three q_t, get the mean q'_t and calculate global maximum of the autocorrelation function of q'_t.
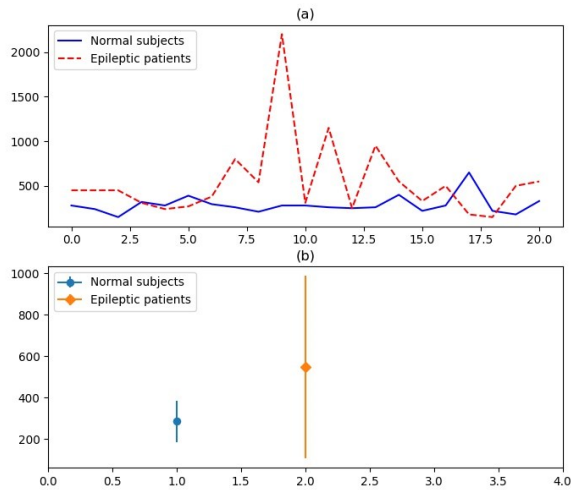

Fig. 3. Use the first five q_t, get the mean q'_t and calculate global maximum of the autocorrelation function of q'_t.
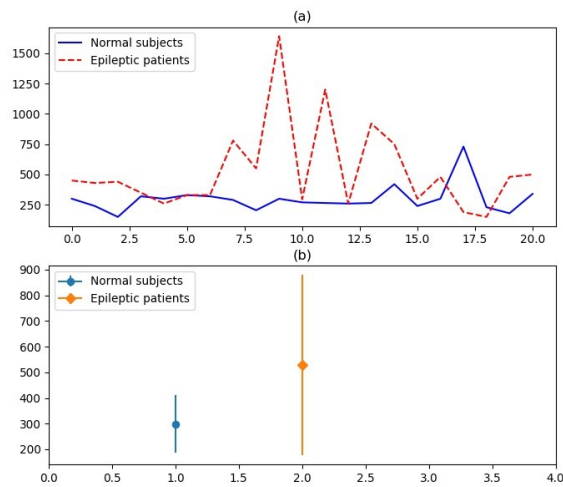

Fig. 4. Use the first ten q_t, get the mean q'_t and calculate global maximum of the autocorrelation function of q'_t.

From Fig1 to Fig(4), the subgraph (a) represent the comparison of autocorrelation function of q_t' between normal subject and patient with epilepsy. And the subgraph (b) is the mean and standard deviation of the autocorrelation fuction of q_t'.

From Fig1(a) – Fig4(a), we can see that in the general trend, the global maximum of the autocorrelation function of q_t' in epilepsy patients is always larger than that of the normal person, whether or not average the q_t, although the method of average decrease the maximum of autocorrelation function of q_t'. and we can also see that the maximal autocorrelation function of q_t in patient with epilepsy is more volatile and more unstable. This show that the patient with epilepsy have more unstable EEG signals.

In order to show the above conclusion clearly, we draw the mean square deviation graph from fig 1(b) – fig4(b). in fig1(b) – fig4(b), we can see that the mean square deviation graph of normal people range from $200 – 400$. But the mean square deviation graph of patient with epilepsy is changing, with the increasing in the number of q_t, the range of mean square deviation graph is decreasing, when average first ten q_t, the range of epilepsy patient is about $150 – 1000$, this completely cover the range of normal people. And this picture also show the means of the maximum value of the q_t' autocorrelation function. The normal people is about 300, but the epilepsy is about 550.

According to fig1-fig4, we can conclude that, for the global maximum of the autocorrelation function of q_t', the value of normal subjects is fluctuate near 300, and the volatility is not very large, but the patients is near 550, and the volatility is very large, its fluctuation range contains the range of normal people's fluctuation range. This reflects the instability of epileptic signals on the side, we deduce that the patient with epilepsy can get high value of the global maximum of the autocorrelation function of q_t' is due to the abnormal discharge of brain.

Conclusion

In this paper, we construct a random walk model to analyze EEG signal, and use the Bayesian inference to extract the characteristic parameter of EEG signal, and obtain the characteristic parameter (q_t, a_t), and calculate the autocorrelation function of q_t'. by comparing the global maximum of autocorrelation function of q_t' between normal subjects and patients with epilepsy, we find that in general trend, for the global maximum of autocorrelation function of q_t', the value in patients with epilepsy is higher than normal people, and the fluctuation in normal is more stable, but the patient with epilepsy is not stable, it contains all the scope which the normal people covers. This result reflect the instability of EEG in epileptic patient. The experimental result show that random walk model can be used to analyze the EEG signal and find the difference between normal people and patient with epilepsy.

References

[1] Dliou A, Latif R, Laaboubi M. Arrhythmia ECG signal analysis using non parametric time-frequency techniques[J]. International Journal of Computer Applications, 2012, 41(4).

[2] Reference to a chapter in an edited book: Rai H M, Trivedi A, Shukla S. ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier[J]. Measurement, 2013, 46(9): 3238-3246.

[3] Owis M I, Abou-Zied A H, Youssef A B M. Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification[J]. IEEE transactions on Biomedical Engineering, 2002, 49(7): 733-736.

[4] Iriarte J, Urrestarazu E, Valencia M. Independent component analysis as a tool to eliminate artifacts in EEG: a quantitative study[J]. Journal of clinical neurophysiology, 2003, 20(4): 249-257.

[5] Spitzer F. Principles of random walk[M]. Springer Science & Business Media, 2013.

[6] Metzner C, Mark C, Steinwachs J, et al. Superstatistical analysis and modelling of heterogeneous random walks.[J]. Nature Communications, 2015, 6:7516.

[7] Box G E P, Tiao G C. Bayesian Inference in Statistical Analysis[M]// Bayesian inference in statistical analysis /. Wiley, 1992:478-479.