Sleep Staging Based on Energy and Complexity Characteristics of Sleep EEG Signals

Xiaobo Li¹, Chunhui Ren¹ and Xiao Zou*¹

¹University of Electronic Science and Technology of China
*Corresponding author’s e-mail: 429051041@qq.com

Abstract. Nowadays, people's sleep problems become more and more serious with the increase of pressure in social life. Sleep staging is the key step in the diagnosis of sleep diseases. The traditional manual staging has the drawbacks of consuming manpower and time, and the standard is difficult to unify. In order to effectively realize automatic sleep staging, a sleep staging method based on EEG energy characteristics and complexity is proposed in this paper. Because the complexity of EEG signal and the energy of each characteristic wave change regularly with the different sleep period, sleep period can be distinguished effectively based on these features. Firstly, FIR band-pass filter is used to extract the characteristic wave from EEG signal and calculate its energy. Then the complexity of each sleep period is calculated. Finally, the improved support vector machine is used to classify the sleep stage. The experimental results show that the method proposed in this paper is simple and effective, and has good clinical application value.

1. Introduction

With the accelerating pace of modern life, people's life pressure is increasing. Sleep-related diseases have to be noticed[1].

Sleep staging is an important step in the study of sleep disease. Sleep staging has manual staging and automatic staging. Manual staging consumes a lot of manpower and is susceptible to personal experience, and its reliability is difficult to measure. Therefore, automatic staging research is very necessary.

There are many ways to automate sleep staging, such as neural networks and support vector machines. Support vector machine(SVM) is a commonly used effective classification in many training classification methods of automatic sleep staging. The method has the advantages of complete theoretical foundation, good robustness and simple algorithm. In 2002, D.Gorur applied SVM to sleep spindle waves (the extraction of characteristic waveforms in N2 stage), the correct rate is 7% higher than that of the neural network method. It proves that SVM has a good effect on sleep staging[2]. This provides a theoretical basis for the use of SVM as a classification tool in this paper.

2. Data and analysis

Since EEG signals indicate brain activity, EEG signals are used frequently during sleep staging. The latest standard of the American Academy of Sleep Medicine(AASM) divides sleep into two phases, the Non-rapid Eye Movement (NREM) and the Rapid Eye Movement (REM). NREM can be further divided into stage one (stage I, N1), stage two (stage II, N2), stage three (stage III, N3)[3]. During the entire sleep process, changes in the EEG correspond to changes in depth of sleep.

The greatest difference of EEG signals in different sleep periods is that they contain different
characteristic waves, which are $\alpha$ (8-12Hz), $\beta$ (14-35Hz), $\theta$ (4-8Hz), $\delta$ (2-4Hz), sleep spindle(12-14Hz), and k-complex(0.5-2Hz). Otherwise, $\beta$ wave consists of $\beta_1$ (14-22Hz) and $\beta_2$ (22-35Hz). Table 1 shows the corresponding relationship between different sleep periods and characteristic waves under EEG signals[4].

Table 1.EEG basic rhythms in different sleep stages

| Sleep stage | Rhythms                  |
|-------------|--------------------------|
| W           | $\alpha, \beta$         |
| N1          | $\theta$                |
| N2          | sleep piddle,k-complex   |
| N3          | $\delta$                |
| REM         | $\alpha, \beta, \theta$ |

The EEG data used in this paper is obtained from the Sleep-EDF Sleep Database in the MIT-BIH Physiology Information Base. All recorded data includes two EEG signals, Fpz-Cz and Pz-Oz, with a sampling frequency of 100 Hz. According to the experimental results, the treatment of Fpz-Cz makes the staging accuracy higher[5]. In this paper, the Fpz-Cz channel of EEG signal is used for analysis. The EEG signal in the database has been manually labeled by sleep experts and can be used as a reference for the sleep staging results.

In this paper, EEG signals of 13 subjects were selected. Among them, 12 subjects were used as a training set, and the remaining 1 subject was a test sample. The samples included in the training set are sc4021, sc4031, sc4041, sc4102, sc4151, sc4181, sc4171, sc4162, sc4141, sc4081, sc4091 and sc4092, for every 30 seconds of data to be regarded as an epoch, so the training data includes W 1040 epochs, N1 617 epochs, N2 1040 epochs, N3 965 epochs, REM 1040 epochs. The test data sc4142 has W 100 epochs, N1 27 epochs, N2 100 epochs, N3 100 epochs, REM 100 epochs.

3. Feature Selection

The brain is actually a complex nonlinear system, and the EEG signal is also a kind of nonlinear signal. Chaos is the main manifestation of nonlinear systems. Chaos theory is a new discipline for the study of nonlinear systems. Compared to linear system description methods, it is more appropriate to describe brain activity with a nonlinear system[6].

In summary, according to the characteristic rhythm and nonlinear characteristics of EEG signals, this paper selects EEG signal energy and a nonlinear feature as the classification basis.

3.1 Energy feature extraction

The FIR bandpass filter has a strict linear phase characteristic, and the acquisition of the band range is not limited by the sampling frequency, which is important for extracting a specific characteristic wave. This article uses a kaiser window to design an FIR bandpass filter. The design parameters of the filter are: Bandpass ripple tolerance is 0.1, The resistance of the stop band ripple is 0.02db, Bandpass amplitude is 1. The above seven characteristic waves have different effects on sleep staging, and all the characteristic waves will be used in this paper. The characteristic wave of the time domain can be obtained by the FIR bandpass filter, and then the characteristic wave of the time domain is squared to obtain the energy characteristic of the brain electrical energy. In order to reduce the influence of noise on the signal, the relative energy of each characteristic wave and the energy ratio of one band ($E_\alpha / E_\delta$)[7] are used. For a 30s data(3000 data points) of sc4021, extract the respective characteristic waves with the corresponding bandpass filter. As shown in Figure 1.
3.2 Nonlinear feature extraction

Lempel-Ziv complexity (LZC), is also called Kc complexity, presented by Kolmogorov in 1965, Lempel and Ziv proposed a computer implementation of the algorithm until 1987. They define the complexity of a symbol sequence as the shortest length of a computer programming algorithm that can reproduce this sequence[8]. It is an representative indicator of the complexity of sleep EEG signals and an effective response to the nonlinear dynamics of each stage of sleep. The specific process algorithm is as follows[8]:

Suppose we get the EEG signal time series as sequence \( S(S_1, S_2, \ldots, S_m) \), the first step is to calculate the average of this sequence. If \( S_i > \text{average} \), let \( S_i' = 1 \), otherwise \( S_i' = 0 \). Then we can obtain the new string \( S'(S'_1, S'_2, \ldots, S'_n) \) according to the above rules, which is denoted as \( S(S_1, S_2, \ldots, S_n) \) and consisted of 0 and 1. A new symbolic series \( SQ = (s_1, s_2, s_{m+1}, \ldots, s_{m+k}) \) would generated by adding a signal character \( Q = s_{m+1} \), or a string characters \( Q = (s_{m+1}, s_{m+2}, \ldots, s_{m+k}) \). If the sequences of \( SQV = (s_1, s_2, \ldots, s_{m+k}) \) obaining by delete the last character of \( SQ \) contains \( Q \),the copy operation adding \( Q \) to the last of \( SQV \) would be carried. Then repeat the above steps until \( Q \) disappears in \( SQV \). The insert operation could be executed when the above all steps completed. If \( SQV \) does not consist \( Q \), insert "." into behind of \( SQV \) and turn all previous characters of "." into "". Repeat all above steps until the end of the sequences. We use "." to divide the string into segments, the total number of segments is the complexity \( c(n) \). It has been verified that the complexity of all symbol sequences will approach a fixed value:

\[
\lim_{n \to \infty} c(n) = b(n) = \frac{n}{\log_L n}
\]

Where \( L \) is the number of coarse granulation, in this paper, \( L=2 \). Then \( c(n) \) can be normalized,
The complexity of sleep EEG in each sleep period of the six subjects selected in this paper are shown in Table 2.

Table 2. Kc complexity in each sleep period of each sample

| Subject | W   | N1  | N2  | N3  | REM |
|---------|-----|-----|-----|-----|-----|
| sc4021  | 0.7084 | 0.5927 | 0.4163 | 0.2770 | 0.5271 |
| sc4031  | 0.8969 | 0.5693 | 0.4519 | 0.3180 | 0.5531 |
| sc4041  | 0.8223 | 0.6460 | 0.5683 | 0.3692 | 0.5819 |
| sc4102  | 0.7081 | 0.5554 | 0.3664 | 0.2406 | 0.4471 |
| sc4151  | 0.6814 | 0.4224 | 0.3939 | 0.2672 | 0.3379 |
| sc4181  | 0.9040 | 0.6133 | 0.4683 | 0.3200 | 0.5072 |

3.3 Classification

It is a key issue to select the classifier with excellent performance to automatically classify the characteristics of EEG signals during sleep staging. According to the literature[9], LS-SVM performs well in the sleep staging process among multiple classifier algorithms. Based on the principle of LS-SVM[10], the classification function is

\[
y(x) = \text{sign} \left( \sum_{i=1}^{n} \alpha_i y_i k(x_i, x) + b \right)
\]

(3)

Where \( y \) is the output variable and \( x_i \) is the input variable. \( \alpha_i \) is the Lagrange multiplier, \( b \) is deviation and \( k \) is the kernel function. In order to solve the problem of nonlinear classification, try to avoid the "dimensionality disaster" that may be encountered in the analysis process, we need to use a kernel function for this purpose. In this paper, we choose the more commonly used Radial basis function (RBF) as the kernel function, the expression is as follows:

\[
K(x, x') = \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right)
\]

(4)

Where \( \sigma \) is the width of kernel function. Now, there are two parameters in LS-SVM, which are penalty factor \( C > 0 \) and kernel parameter \( \sigma > 0 \). In order to further improve the accuracy of staging, these two parameters need to be optimized. In reference [11], a more suitable result for sleep staging experiment was obtained by particle swarm optimization. Here we choose \( C = 87.65 \) and \( \sigma = 0.725 \).

4. Results

Based on the selected data set, the energy characteristics and Kc complexity mentioned above are calculated separately. Then, the training set data is input into the LS-SVM to obtain a classification model, and then the test set data is used for testing. The obtained result is shown in Table 3.

Otherwise, since the amount of data for the N1 phase used for testing herein is small, according to the AASM guidelines mentioned above, N1 and N2 can be combined into light sleep period (LS) and N3 as a deep sleep period (SWS).

Table 3. Classification result

| Sleep stage | W   | LS | SWS | REM | Accuracy |
|-------------|-----|----|-----|-----|----------|
| W           | 88  | 7  | 1   | 4   | 88.0%    |
| LS          | 4   | 107| 2   | 14  | 84.3%    |
| SWS         | 0   | 10 | 86  | 4   | 86.0%    |
| REM         | 3   | 13 | 1   | 83  | 83.0%    |
| Average     |     |    |     |     | 85.2%    |
5. Conclusion
In this paper, the energy characteristics of the EEG signal and the kc complexity are selected to realize the automatic staging experiment, and the staging accuracy is 85.2%. Literature [12] gives an accuracy of 83.44% for sleep staging experiments using only energy characteristics. It shows that the kc complexity feature introduced in this paper and the improved support vector machine effectively improve the accuracy. In addition, the FIR bandpass filter is used to extract the characteristic wave of EEG signal, which is convenient and easy to operate. It eliminates the cumbersome layer analysis using common wavelet decomposition and can accurately obtain the required frequency band. The experimental results show that the energy and Kc complexity features selected in this paper are simple and effective, and can realize the automatic staging of sleep.

References
[1] Prasad B., Saxena R., Goel N, et al (2018) Genetic Ancestry for Sleep Research: Leveraging Health Inequalities to Identify Causal Genetic Variants. J. Chest,:S0012369218304719.
[2] Gorur D, Halici U, Aydin H, et al (2002) Sleep spindles detection using short time Fourier transform and neural networks. C. International Joint Conference on Neural Networks.
[3] Berry R B, Brooks R, Gamaldo C E, et al (2012) The AASM manual for the scoring of sleep and associated events:rules, terminology and technical specifications, version2.0. J. American Academy of Sleep Medicine.
[4] Acharya U R, Bhat S, Faust O, et al (2015) Nonlinear Dynamics Measures for Automated EEG-Based Sleep Stage Detection. J. European Neurology, 74(5-6):268-287.
[5] Hsu Y L, Yang Y T, Wang J S, et al (2013) Automatic sleep stage recurrent neural classifier using energy features of EEG signals. J. Neurocomputing, 104:105-114.
[6] Kuo C E, Liang S F (2011) Automatic stage scoring of single-channel sleep EEG based on multiscale permutation entropy. C. Biomedical Circuits & Systems Conference.
[7] Radha M, Garcia-Molina G, Poel M, et al (2014) Comparison of feature and classifier algorithms for online automatic sleep staging based on a single EEG signal. J. Conf Proc IEEE Eng Med Biol Soc, 2014(1):1876-1880.
[8] Lempel A , Ziv J (1976) On the Complexity of Finite Sequences. J. IEEE Transactions on Information Theory, 22(1):75-81.
[9] José Manuel Sánchez Pascualvaca, Fernandes C , Alberto Guillén, et al (2013) Sleep Stage Classification Using Advanced Intelligent Methods. M. Advances in Computational Intelligence. Springer Berlin Heidelberg.
[10] Miano S, Paolino M C, Castaldo R, et al (2010) Visual scoring of sleep: A comparison between the Rechtschaffen and Kales criteria and the American Academy of Sleep Medicine criteria in a pediatric population with obstructive sleep apnea syndrome. J. Clinical Neurophysiology, 121(1):39-42.
[11] Xin Liu (2017) Research on Signal Processing and Sleep Staging Based on EEG. D.
[12] Fang Yang, Tao Li (2017) Automatic sleep staging based on energy features and fuzzy entropy. J. Chinese Journal of Medical Physics.