World Scientific Research Journal Volume 7 Issue 8, 2021
ISSN: 2472-3703 DOI: 10.6911/WSRJ.202108_7(8).0008

Research on Fatigue Driving Based on EEG Signals

Gao Wei, Xuebin Qin

Xi'an University of Science and Technology, Xi'an, China

Abstract

The frequent occurrence of traffic accidents has aroused widespread concern in the society. The study shows that fatigue driving is the main factor causing traffic accidents. Therefore, it is very important to detect the fatigue of the driver. In this paper, the EEG signals are preprocessed by independent component analysis, and the eye electricity signals are separated. Then the wavelet transform is used to extract the features, and the Relevance vector machine is used as the initial value of the fatigue test, RANSAC method is used to select the fatigue signal, and the fatigue state of the driver is compared with KSS standard by calculating the mahalanobis distance. The experimental results show that the precision of this method is 91% higher than that of the traditional method, and it has wide application value.

Keywords

EEG; Fatigue State; Wavelet transform.

1. INTRODUCTION

Fatigue driving is an important cause of traffic death. According to the National Transportation Safety Administration, about 35 percent of traffic accidents each year are caused by driver fatigue, and 20 percent to 30 percent of fatal traffic accidents are caused by driver drowsiness or fatigue [1]. About 50 per cent of people who drive fatigued are driving a major aspect of road safety. If the driver can receive the alarm when he is in a state of fatigue, the occurrence of these incidents will be greatly reduced.

At present, fatigue detection methods are divided into subjective detection and objective detection methods. The subjective fatigue detection method is a subjective evaluation, which is influenced by the subjective judgment ability of drivers and researchers. It has certain limitations in detecting fatigue state. In practical application, it can only be evaluated subjectively after the event, and is not convenient for real-time monitoring, therefore, it is rarely widely promoted. Fatigue detection method based on physiological signals has become a trend in recent years. It collects physiological signals of the corresponding parts of the driver, such as electroencephalogram (EEG), electromyogram (EMG), etc.a specific algorithm model is designed to determine whether the driver is in a fatigue state in real time. Electroencephalogram (EEG), which has a high time resolution and precision, can accurately reflect the physiological state of human body, is regarded as the "Gold standard" to detect human fatigue. Electroencephalogram (EEG) is the most traditional physiological method to study fatigue. In this paper, driver's eye and electromyogram (EMG) signals are used to detect fatigue. Establish a "Bridge"between the use of eye and EMG signals and fatigue [2].

2. TRESEARCH METHOD

At first, two physiological signals of driver's eye electricity and EMG were obtained under different fatigue state, and driver and driver were investigated and analyzed by subjective questionnaire. Then, the EEG signals obtained from the experiment are preprocessed, and the

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202108_7(8).0008

changes of EEG signals in different frequency bands during the process of fatigue generation are analyzed, the change trend of EEG Fatigue Index with driving time was analyzed, and the correctness of EEG fatigue index to fatigue reaction was proved [3].

2.1. Prtereatment

EEG signal is only a few dozens of micro-amplitude signal, and its composition is more complex, even through the collection of professional EEG collection devices, it is inevitable to mix a large number of noise signals. Even if the EEG signal is collected in the laboratory [4], it will be influenced by power frequency signal and electrode drift. For this reason, we must restore the EEG signal as much as possible by preprocessing the original EEG signal to reduce the influence of noise, so that the signal can be kept changing at a baseline. Figure 1 is the original EEG signal with noise [5].

EEG signal is only a few dozens of micro-amplitude signal, and its composition is more complex, even through the collection of professional EEG collection devices, it is inevitable to mix a large number of noise signals. Even if the EEG signal is collected in the laboratory [4], it will be influenced by power frequency signal and electrode drift. For this reason, we must restore the EEG signal as much as possible by preprocessing the original EEG signal to reduce the influence of noise, so that the signal can be kept changing at a baseline. Figure 1 is the original EEG signal with noise [5].

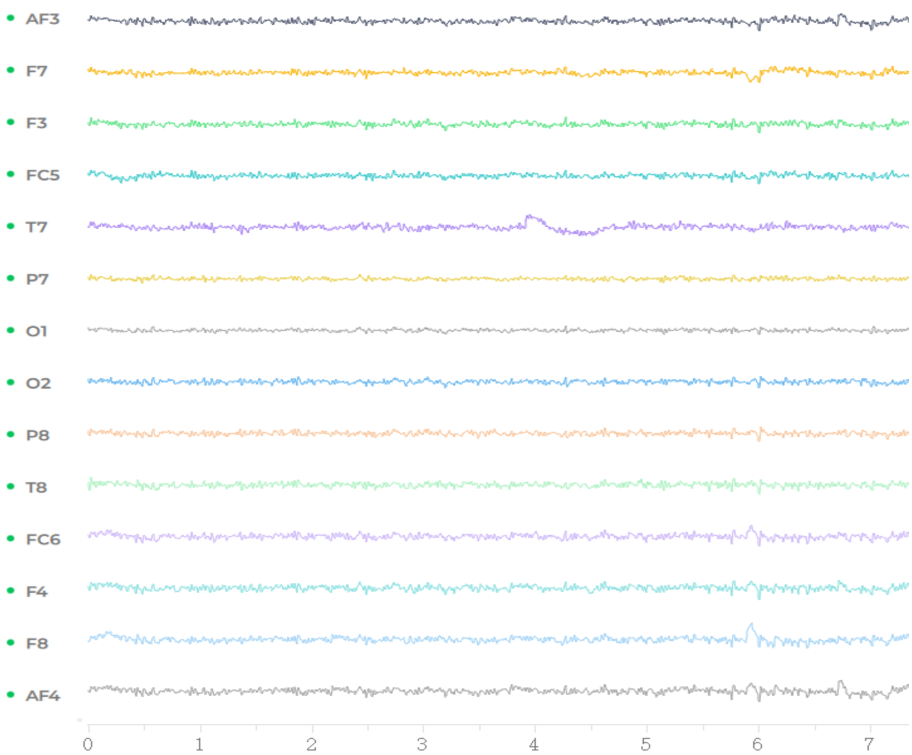

Figure 1. Primitive electroencephalogram

ICA (Independent Component Correlation Algorithm) is a linear analysis method to analyze multidimensional data. It tries to represent a group of random variables as a linear combination of statistical independent variables.

Let X(t)=[$x_1(t),x_2(t),x_3(t),..x_n(t)$] Is a set of N-dimensional observations produced by M hidden variables. Let T be the number of data samples t= 1,2,...,m, T is the length of the observed data.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202108_7(8).0008

Let S(t)=[$s_1(t), s_2(t)\ldots s_n(t)$]Is a set of random variables implicit in the m-dimension of X(t),Among them m$\leq$n, To linear problem to linear problem, that is:

$$x_i(t) = \sum_{i=1}^{n} a_{ij} s_i(t) \qquad (1)$$
$$i = 1,2,\ldots,n;\ j = 1,2\ldots,m$$

Derivation of X= AS by matrix form definition, let s_i be an independent component and $A=[a_1, a_2,\ldots a_n] \in R^{m\times n}$ is a full rank.; $a_i (i=1,2,\ldots,m)$ is the base vector of the mixed matrix, That is observation signal x_i is made by independent source x_i through different weight lines Sex-weighted a_{ij} composition. In the process of processing ICA algorithm, the system target is matched by a linear transformation matrix W.Output y_i , that is

$$y(t) = Wx(t) = WAs(t) \qquad (2)$$

First of all, a piece of the original EEG needs to be removed from the eye and EMG artifacts. ICA method is often used in the removal of EEG artifacts. Because of the interference of EEG and EMG, frequency band and information overlap, the ICA method can not be used in the removal of EEG artifacts, each noise signal has a certain period, the frequency is relatively fixed, the fluctuation is obvious in the time domain signal, and the EEG signal channel number, is convenient to observe each channel to be affected by the noise situation.

As can be seen from the processing of Figure 2, it is very important to pre-process the relatively pure EEG signals to facilitate the follow-up tracking of EEG signals, which is a very complex and weak signal, this method can greatly reduce the noise signal interference to the system.

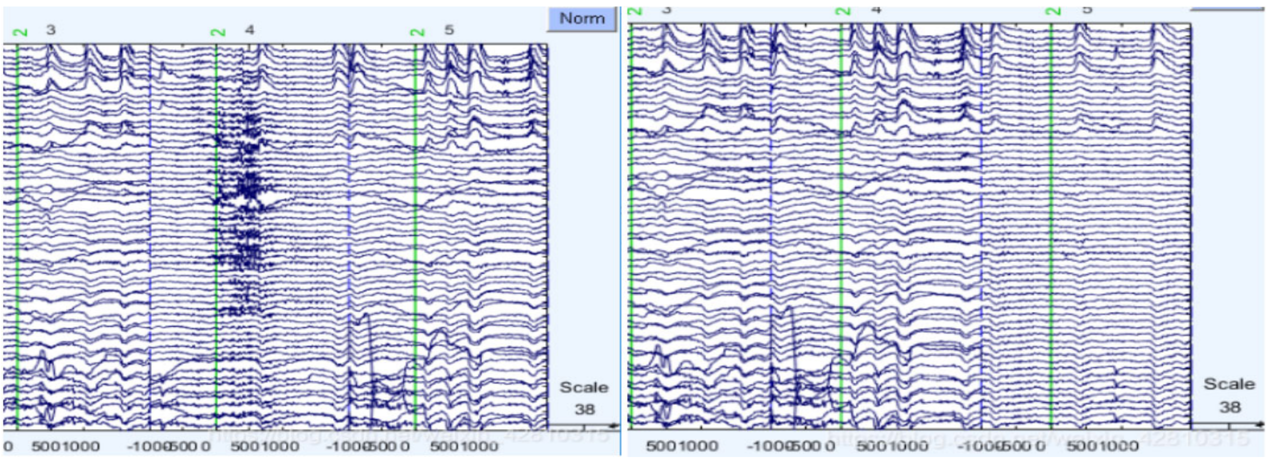

Figure 2. ICA method was used to treat the defecation before and after EEG

2.2. Wavelet Transform

Wavelet transform can choose the window that can be transformed to analyze the different frequency components of EEG signal. Therefore, this paper selects db4 wavelet to extract features. The wavelet has no explicit expression, but the conversion function $|h^2|$ is determined.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202108_7(8).0008

Suppose $P(y)=\sum_{K=0}^{N-1} C_k^{N-1+k}$,Where C_k^{N-1+k} is the coefficient of binomial, then:

$$|m_0(\omega)|^2=(\cos^2\tfrac{\omega}{2})^N p(\sin^2\tfrac{\omega}{2}) \tag{3}$$

In Formula (3)

$$m_0(\omega)=\tfrac{1}{\sqrt{2}}\sum_{k=0}^{2N-1} h_k e^{-ik\omega} \tag{4}$$

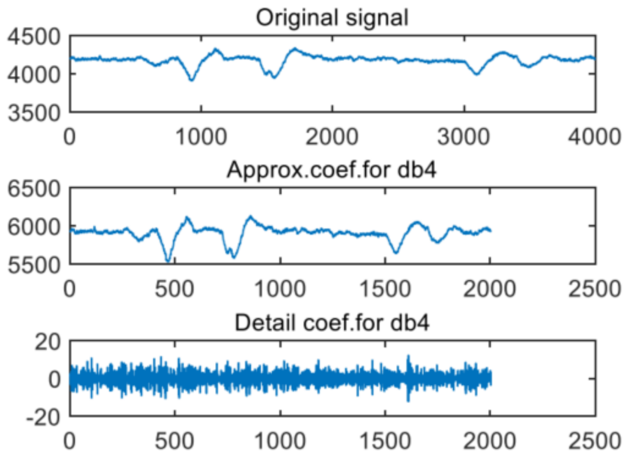

Figure 3. Original signal and high and low frequency signal after db4 wavelet decomposition

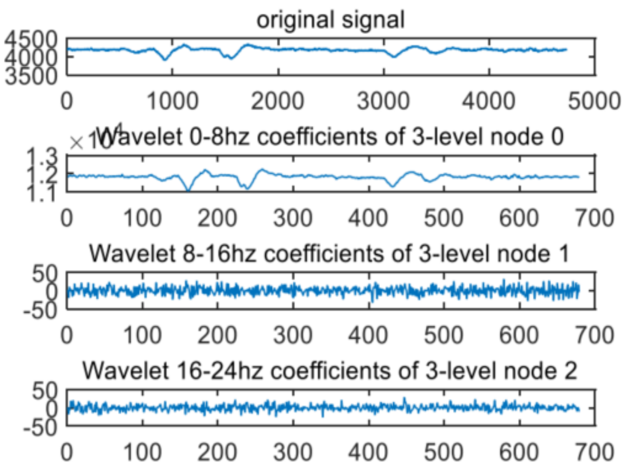

Figure 4. Wavelet coefficients of each frequency band after reconstruction

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202108_7(8).0008

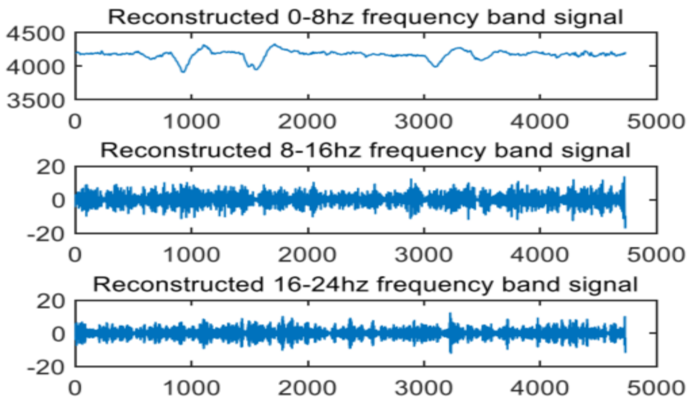

Figure 5. Signals of each frequency band after reconstruction

After decomposition and reconstruction, the required frequency band of EEG signal can be obtained. Figure 3 and Figure 4 show that the wavelet coefficients of low and high frequency signal and reconstructed frequency band respectively. It can be observed that the variation rules of original signal, low frequency band EEG signal and low frequency band wavelet coefficient are similar. The required frequency band signal can be obtained after the reconstruction as shown in Figure 5 and Figure 6.

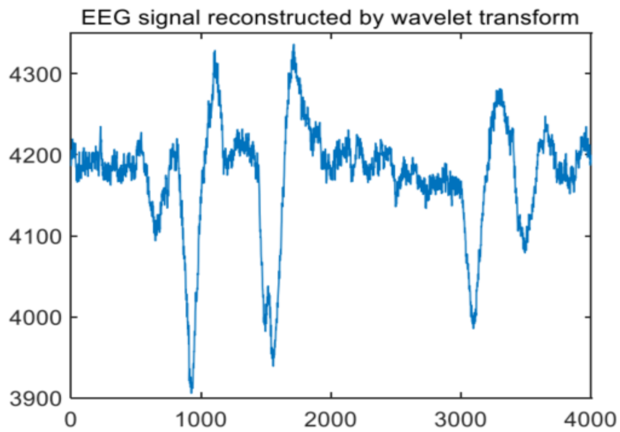

Figure 6. EEG signal after wavelet reconstruction

2.3. Relevance Vector Machine

SVM is a widely used machine learning method in the field of EEG fatigue detection, higher classification accuracy has been achieved in the study of EEG fatigue classification. SVM is used to solve nonlinear, small sample and Galway model learning problem has unique advantages, but in the process of using this model to classify emotions, there is a problem that the choice of kernel function and penalty factor is restricted by conditions, as the sample size increases, the number of support vectors also increases, which leads to high computational complexity when dealing with large data sets. RVM and SVM have the same functional form but are not restricted by kernel function and penalty factor, so RVM has high sparsity. The number of RVM correlation vectors does not increase with the increase of the number of samples. In addition, the posterior probability of RVM can estimate the classification accuracy. Compared with SVM, RVM has more advantages in EEG emotion recognition [7].

The analysis and processing technology of EEG signal is closely related to its own characteristics. First of all, the electrical signals in the brain RVM has the same function form as SVM model, so it can be used to solve the problem of linear classification. Second, EEG signals

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202108_7(8).0008

are Galway data, which is usually assumed to be sparse when modeling high dimensional data Sex. RVM is proposed on the basis of Bayesian framework, by using hyper-parameter constraint weights and using auto-correlation theorem A sparse model can be obtained by removing the sample data with less correlation. The basic principles are as follows:

In EEG fatigue recognition, training samples are assumed to be $\{x_n, \ t_n\}_{n=1}^{N}$, x_n representing the input EEG eigenvector, t_n representing the emotion class label, $t \in \{0,1\}$, 0 and 1 representing fatigue and non-fatigue, and N is the number of training samples. RVM transforms low-dimensional nonlinear inseparability problem into high-dimensional linear separability problem based on kernel function mapping. It has the same functional model form as SVM:

$$y(x;\omega) = \sum_{n=1}^{N} \omega_n K(x, x_n) + \omega_n \tag{5}$$

Where ω_n is the weight and $K(x, x_n)$ is the kernel.

In the classification problem, we need to complete the linear combination of the basis functions by sigmoid mapping function, that is $y(x;\omega) = \sigma(\omega^T \Phi(x))$. If the samples are independent and uniformly distributed, the likelihood function of the sample set can be expressed as:

$$p(t \mid w) = \prod_{n=1}^{N} \sigma\{y(x_n, w)\}^{t_n} [1 - \sigma(y(x_n, w))]^{1-t_n} \tag{6}$$

For the objective function $t_n = y(x_n, \omega) + \varepsilon_n$, the noise function ε_{n} is assumed to follow the Gauss Distribution, with a mean of 0 and a variance of σ^2. If we use maximum likelihood to find ω and σ^2 directly, it will lead to over fitting. To avoid this in RVM, a precondition is set for ω_n to obey Gauss's prior distribution, with an average of 0 and α^{-1} variance:

$$p(w \mid \alpha) = \prod_{n=0}^{N} N(\omega_n \mid 0, \alpha_n^{-1}) \tag{7}$$

Where α is a hyper parameter, a prior distribution of weight values is determined by a Gamma-common prior distribution:

$$p(\alpha) = \prod_{n=0}^{N} Gamma(\alpha_n \mid 0,0) \tag{8}$$

The model of RVM tends to infinity in the process of iterative updating, and the weight value tends to 0, which leads to the model of RVM becoming sparse.Figure 7 shows the classification classification diagram using RVM.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202108_7(8).0008

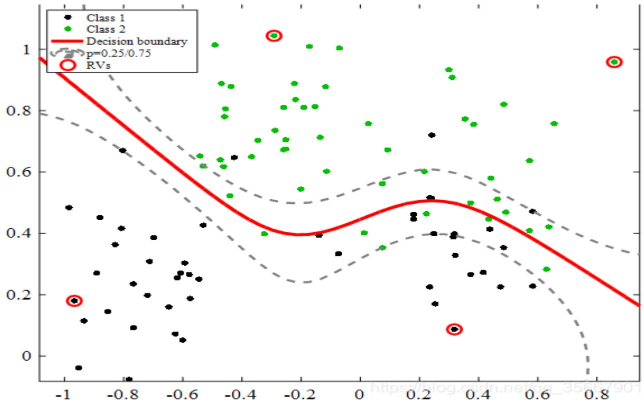

Figure 7. Using RVM classification diagrams

2.4. Markov distance

The Mahalanobis Distance is the covariance Distance of the data. It is an effective method to calculate the similarity of two unknown sample sets. Unlike the Euclidean distance, it takes into account the relationship between properties. The mahalanobis distance is independent of the dimension, and the mahalanobis distance between two points is independent of the unit of measurement of the original data The mahalanobis distance between two points calculated from standardized and centralized data (the difference between the original data and the mean) is the same. The mahalanobis distance also eliminates the correlation between variables.

For a mean of $\mu=(\mu_1,\mu_2,\mu_3,...,\mu_p)^T$ with a covariance matrix of $x=(x_1,x_2,x_3,...,x_p)^T$, the mahalanobis distance is:

$$D_M(x)=\sqrt{(x-\mu)^T S^{-1}(x-\mu)} \tag{11}$$

3. EXPERIMENTAL PROCESS

Twenty healthy people were selected to be tested, and before the EEG was recorded, the subjects were kept in a normal state (that is, awake) and filled out a Fatigue Scale. And explain to the subject to carry out the relevant tasks and notes. At first, the subjects sat in a dimly lit, muted lab. The sensor position of EEG equipment is fully coated with EEG glue to ensure the accuracy of the signal data. The boring behavior task is used in the formal experiment, which makes the subjects enter the fatigue state repeatedly. This paper introduces the KSS of Karolinska Sleep Scale 1 as the criterion of subjective fatigue. There are ten levels of KSS. The higher the score, the higher the fatigue level. This is a simple and effective method to judge the fatigue of a simple auxiliary method. As can be seen from the table, the fatigue scores of the subjects in the quiet state were far lower than those after performing boring behavioral tasks [6].

Table 1. Karolinska sleepiness scale KSS

number	Verbal description
1	extremely alert
2	alert
3	neither alert nor sleepy
4	sleepy, but no effort to keep alert
5	very sleepy, great effort to keep alert
6	fighting sleep

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202108_7(8).0008

4. RESULTS AND DISCUSSION

SPSS software was used to make statistics on the fatigue degree questionnaire of all subjects in quiet and fatigue state ($P \leq 0.05$). The results are shown in Table 2.

Table 2. Karolinska sleepiness scale KSS

state	KSS
quiet	3.10±1.20
fatigue	6.80±1.52

As can be seen from the table, the fatigue scores of the subjects in the quiet state were far lower than those after performing boring behavioral tasks. In other words, doing boring behavioral tasks can lead to a state of fatigue.

In order to improve the accuracy of classification, the electroencephalogram EEG signals of drivers are obtained by analyzing the electroencephalogram EEG signals of drivers, the above results are further classified with the artificial neural network results as the initial values. Finally, the average value of the selected data is taken as the standard value of the driver's fatigue degree. The experimental results show that the relationship between the standard value and the EEG data of the driver can be judged by calculating the geometric distance measure, this classification method is more accurate and convenient than the existing fatigue test method, and has a wide range of application value.

REFERENCES

[1] F. Zhang and F. Dai, "Research and Application of Fatigue Driving Monitoring and Emergency Treatment," 2017 International Conference on Computer Technology, Electronics and Communication (ICCTEC), Dalian, China, 2017, pp. 1072-1075.

[2] Huang Yakang. Detection of driving fatigue state based on electroencephalogram and electroophthalmic signals. Harbin University of Science and Technology, 2020.

[3] Yang Feiyu. Moving imagery EEG classification algorithm based on transfer learning [D]. South China University of Technology, 2020.

[4] Joo Won. Study on automatic removal of ocular artifacts from EEG signals [D] . Jiangsu University of Science and technology, 2020.

[5] Yao Juanjuan, Lu Kun, Ma Xingyu, Cheng Gong. Study on fatigue detection method based on EEG and eye electricity. Electronic design engineering, 2020,28(06): 115-120.

[6] Luo Jinchen, Zou renling, Jiang Yue, Hu Xiufang. Recognition of multi-class motor imagery EEG based on multi-feature fusion. Bioinformatics, 2020,18(03): 176-185.

[7] Sun Xiaomeng. Research on conic fitting algorithm based on RANSAC [J]. Western leather, 2020, 42(08): 56-57.