

基于能量特征和模糊熵的睡眠自动分期

杨芳,李涛

华南师范大学物理与电信工程学院,广东 广州 510006

【摘要】目的:为实现有效睡眠自动分期,提出一种基于脑电能量特征和模糊熵的睡眠分期方法。**方法:**首先利用小波变换进行脑电信号去噪,再利用FIR带通滤波器提取脑电信号的特征波,获得能量特征,并提取脑电信号的模糊熵,最后利用支持向量机进行模式识别。**结果:**能量特征值和模糊熵值随着睡眠状态的变化而不同,睡眠各期可以根据特征值的不同而得到有效区分,通过对1 140个脑电信号样本进行睡眠分期,得到的平均准确率为88.45%。**结论:**基于脑电能量特征和模糊熵的睡眠分期方法简单、有效,具有良好的临床应用价值。

【关键词】睡眠自动分期;脑电;能量特征;模糊熵;支持向量机

【中图分类号】R388.8

【文献标志码】A

【文章编号】1005-202X(2017)09-0968-05

Automatic sleep staging based on energy features and fuzzy entropy

YANG Fang, LI Tao

School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou 510006, China

Abstract: Objective To propose a new sleep staging method based on the energy feature and fuzzy entropy of electroencephalogram (EEG) signals for achieving effectively automatic sleep staging. **Methods** The original EEG signal was denoised with wavelet transform. And then, FIR band-pass filter was used to extract the energy feature and fuzzy entropy of the denoised EEG signal. Finally, support vector machine was adopted for sleep staging. **Results** The energy feature and fuzzy entropy were different with the change of sleep stages. The sleep stages were effectively distinguished based on the different feature values, and the sleep staging performed for 1 140 EEG signals achieved an average accuracy of 88.45%. **Conclusion** The sleep staging method based on the energy feature and fuzzy entropy of EEG signals was simple and effective, with application prospects.

Keywords: automatic sleep staging; electroencephalogram; energy feature; fuzzy entropy; support vector machine

前言

睡眠是一种人体缓解疲劳、恢复精神的自然调节方式。睡眠的好坏影响着人体的心理及生理健康^[1]。睡眠分期对于疾病的监测和睡眠质量的评价具有重要意义。传统的睡眠分期通过PSG(Polysomnography)采集脑电、心电、眼电等生理信号,然后由睡眠专家进行视觉判断来实现。这种方式成本高、效率低,不具有普遍应用性。针对人工睡眠分期的不足,自动睡眠分期已成为一个有价值的研究方向。目前,国际上普遍使用2007年美国睡眠医学会(American Academy of Sleep Medicine, AASM)修改过的R&K睡眠分期标准,将睡眠状态分为5期:觉醒

期(Wake)、浅睡1期(N1)、浅睡2期(N2)、深睡期(N3)、快速眼动期(Rapid Eye Movement, REM)^[2]。

实现脑电信号的自动睡眠分期需要以下几个步骤:脑电信号采集、预处理、特征提取、模式识别分类。而睡眠分期的关键在于特征提取和模式识别分类。在特征提取方面,很多学者对脑电信号的时域特征^[3-5]、频域特征^[6-8]和非线性动力学特征^[7,9]等进行了研究。在模式识别方面,常用的方法有:线性判别、贝叶斯方法、人工神经网络、支持向量机(Support Vector Machine, SVM)等^[10-15]。这些睡眠分期方法的准确率在77.10%~95.88%不等。

脑电信号极其微弱,容易受心电、肌电等生理信号影响,同时具有非平稳性的特点。因此,在进行脑电睡眠分期前,本文使用小波变换对脑电信号进行去噪,以便后续提取的脑电特征更加有效。由于脑电信号频域特征比较突出,且特征波能很好地反映睡眠分期,所以本文利用FIR带通滤波器提取脑电信

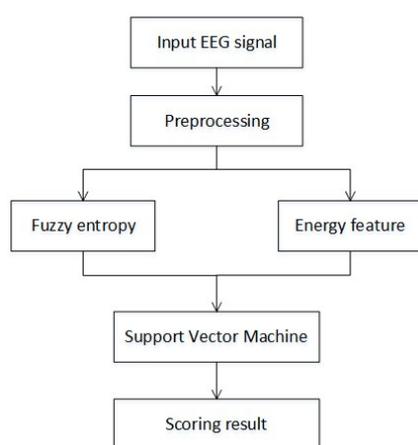
【收稿日期】2017-05-06

【作者简介】杨芳,硕士研究生,研究方向:智能仪器及生理信号处理方向, E-mail: sandrayf@126.com

号的特征波,并将各个特征波的能量特征作为分类特征之一。同时为了进一步提高睡眠分期的准确率,本文增加了非线性特征模糊熵。考虑到非线性特征运算量大,计算速度慢,本文将30 s的睡眠脑电数据分为3个短数据分别进行模糊熵计算,然后求模糊熵的平均值^[16]。最后将得到的特征值用来训练SVM,使用训练好的SVM进行睡眠自动分期。

1 方法

本文睡眠自动分期的具体步骤如下:利用小波变换对原始脑电进行去噪处理,然后通过FIR带通滤波器提取脑电的特征波,并计算相关能量特征;用模糊熵算法对去噪后的脑电信号进行计算,得到脑电信号的模糊熵值;最后将所有特征参数输入到SVM分类器中,将睡眠分成5期:清醒期(Wake)、浅睡1期(N1)、浅睡2期(N2)、深睡期(N3)、REM期。分期方法的流程图如1所示。



EEG: Electroencephalogram

图1 自动睡眠分期流程图

Fig.1 Flowchart of automatic sleep staging

1.1 数据获取和去噪处理

本文使用的脑电数据来自MIT-BIH生理信息库中的Sleep—EDF睡眠数据库^[17],所有的记录数据包括Fpz-Cz和Pz-Oz两导脑电信号,采样频率为100 Hz。根据Hsu等^[18]对两导脑电信号的分析可知,处理Fpz-Cz导联的脑电信号使分期准确率更高。本文选用Fpz-Cz导联的脑电信号进行分析。数据库中脑电信号已经被睡眠专家人工标注,可以作为本文睡眠分期结果的参考。

由于脑电信号容易受到其他生理信号和测量设备的干扰,为了更好地提取脑电信号的特征,需要对脑电信号进行去噪处理。经实验对比,本文使用db4小波基对脑电信号进行去噪处理,同时选用软阈值

中的stein无偏似然估计方法。对sc4002e0的一段30 s(即3 000个数据点)睡眠脑电数据进行去噪,结果如图2所示。

1.2 能量特征提取

脑电信号由 α 波(8~13 Hz), β 波(13~30 Hz), θ 波(4~8 Hz), δ 波(2~4 Hz),spindle波(12~14 Hz),sawtooth-wave波(2~6 Hz)和k-complex波(0.5~1.5 Hz)组成^[18],不同睡眠期出现不同特征波,其对应关系如表1所示。

根据表1分析可知: α 波, β 波只出现在清醒期和REM期,可以通过 α 波或 β 波区分N-REM期和REM期,再利用sawtooth-wave波可以区分清醒期和REM期;对于N1期和REM期,可以利用 β 波和 θ 波进行区分;spindle波和k-complex波只出现在N2期,利用这两种波可以区分出N2期;由于 δ 波只出现在N3期,所以利用它可以区分出N3期。因此利用各特征波可以简单、有效地区分睡眠各期。

FIR带通滤波器具有严格的线性相位特性,且频带范围的获取不受采样频率的限制,这对于提取特定的特征波很重要。本文使用kaiser窗设计FIR带通滤波器。其中滤波器的设计参数为:带通波纹允许误差为0.1,阻带波纹误差允许差为0.01,带通幅值为1。经实验可知,特征波的数据运算量小,且7种特征波对睡眠分期有不同程度的影响,本文将使用上述的所有特征波。通过FIR带通滤波器可以获得时域的特征波,再将时域的特征波求平方,即可得到脑电的能量特征。为了减小噪声对信号的影响,本文采用各特征波的相对能量和一个频带能量比 (E_θ/E_α) ^[19]。对sc4002e0的一段30 s数据(即3 000个数据点),用相应的带通滤波器滤出各个特征波,如图3所示。

1.3 模糊熵特征提取

熵能衡量时间序列产生新模式的概率大小,即熵能表征信号的复杂性。当人处在不同的睡眠分期下,脑电的复杂程度不一样,因此可以用熵监测睡眠深度。而模糊熵值随参数变换过渡平稳,并具备样本熵的相对一致性和短数据集处理特性,克服了样本熵定义的局限。模糊熵的算法原理如下。

(1)设 N 点的采样序列为 $\{u(i); 1 \leq i \leq N\}$;

(2)按序列顺序重构生成一组 m 维矢量:

$$X_i^m = \{u(i), u(i+1), \dots, u(i+m-1)\} - u_0(i), (i = 1, \dots, N-m),$$

其中 $\{u(i), u(i+1), \dots, u(i+m-1)\}$ 是从第 i 个点开始的

连续 m 个 u 的值, u_0 为均值, $u_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} u(i+j)$;

(3)定义矢量 X_i^m 和 X_j^m 的距离 d_{ij}^m 为两者对应元

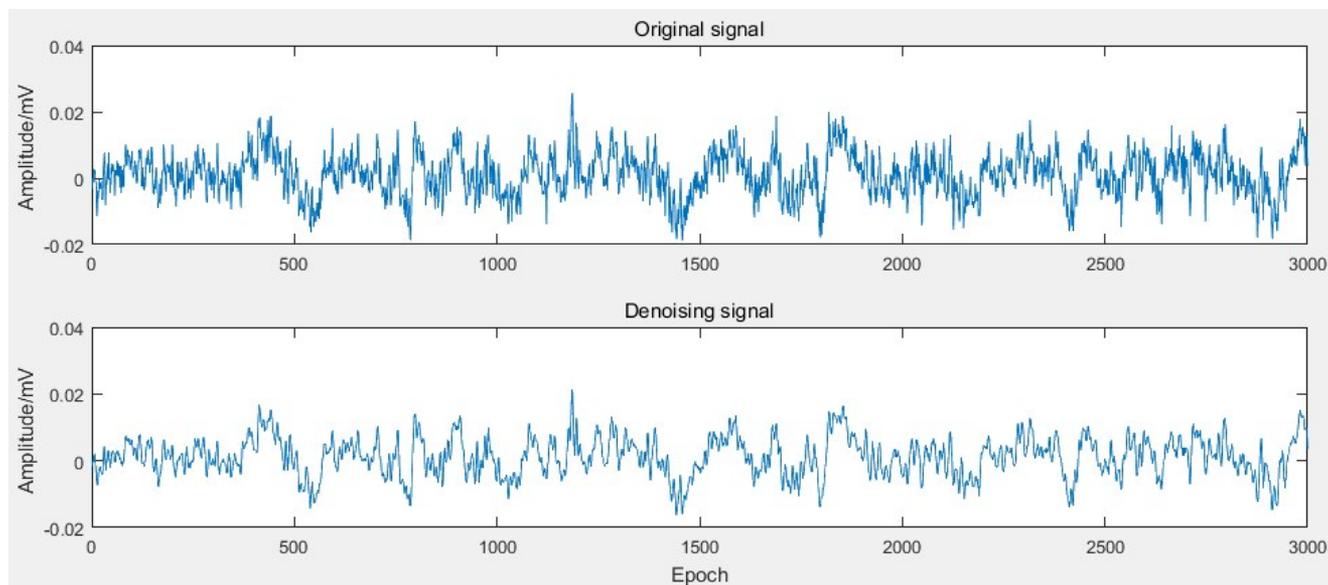


图2 小波去噪结果对比图

Fig.2 Comparison of original electroencephalogram (EEG) signals and denoised EEG signals

表1 睡眠期与特征波的对应关系

Tab.1 EEG feature waves in different sleep stages

Sleep stage	Type of activity
W	α (8-13 Hz), β (13-30 Hz)
N1	θ (4-8 Hz)
N2	spindle (12-14 Hz), k-complex (0.5-1.5 Hz)
N3	δ (2-4 Hz)
REM	α , θ , β , sawtooth-wave (2-6 Hz)

W: Awake; N1: N-REM stage 1; N2: N-REM stage 2; N3: Combination of N-REM stage 3 and N-REM stage 4; REM: Rapid eye movement

素中差值最大的一个,即:

$$d_{ij}^m = d[X_i^m, X_j^m] = \max_{k \in (0, m-1)} \{ |u(i+k) - u_0(i) - (u(j+k) - u_0(j))| \}$$

$$(i, j = 1, \dots, N-m), j \neq i;$$

(4)通过模糊函数定义矢量 X_i^m 和 X_j^m 的相似度 D_{ij}^m , 即: $D_{ij}^m = \exp(-d_{ij}^m/r)$, 其中 n 和 r 分别为指数函数边界的梯度和宽度。

(5)定义函数:

$$O^m(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left[\frac{1}{N-m-1} \sum_{j=1, j \neq i}^{N-m} D_{ij}^m \right];$$

(6)重复步骤(2)~(5),按序列顺序重新生成一组 $m+1$ 维矢量,定义函数:

$$O^{m+1}(n, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left[\frac{1}{N-m-1} \sum_{j=1, j \neq i}^{N-m} D_{ij}^{m+1} \right];$$

(7)定义模糊熵为:

$$\text{FuzzyEn}(m, n, r) = \lim_{N \rightarrow \infty} [\ln O^m(n, r) - \ln O^{m+1}(n, r)].$$

当 N 为有限值时,按上述步骤可算出模糊熵的估计值: $\text{FuzzyEn}(m, n, r, N) = \ln O^m(n, r) - \ln O^{m+1}(n, r)$ 。

计算 $\text{FuzzyEn}(m, n, r, N)$ 时,首先要对 m, n, r, N 4个参数进行选取。经实验对比,当30 s脑电睡眠片段分为3段时,既提高了运算速度,又能满足每个短数据长度的一致性。所以本文令短数据长度 $N=1000$, 嵌入维数 $m=2, n=2, r=0.25 \text{ SD}$ (SD为原始数据的标准偏差)。提取经过去噪处理的睡眠脑电的模糊熵,结果如表2所示。

由表2可见,每个受试者的睡眠各期的模糊熵值有所差异,但模糊熵值的变化规律整体上是—致的:清醒期的模糊熵值明显大于其他睡眠期;随着睡眠的加深,模糊熵值逐渐降低,因为脑细胞活动的同步有序性增加,复杂度降低;到REM期阶段,由于脑神经活动增强,模糊熵值也有所增大,并逐渐接近N1期。由此可见,根据模糊熵值的变化可以有效表现睡眠各期。

2 分类结果

SVM是在统计学习理论的基础上发展起来的学习算法,它在文本分类、手写识别、图像分类、生物信息学等领域获得较好的应用。SVM在使用训练样本分类误差极小化的前提下,尽量提高分类器的泛化能力^[20]。它不需要有样本数据的先验概率的假设,采用结构风险最小化原则代替经验风险最小化原则,较好地解决了小样本学习的问题。同时引入的核函数更好地解决了高维和非线性问题。

本文采用SVM作为分类器,选取径向基核函数

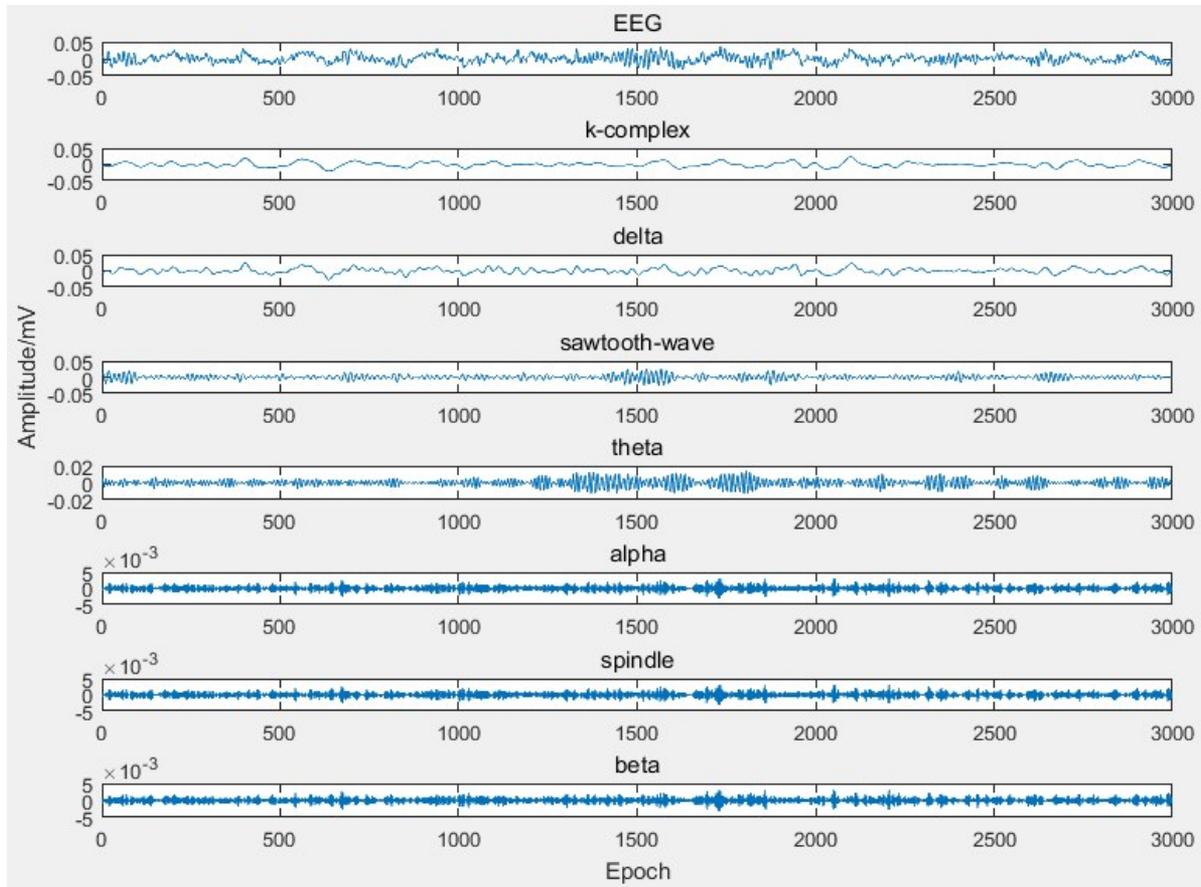


图3 时域的特征波

Fig.3 Feature waves in time domain

表2 睡眠各期的模糊熵值

Tab.2 Fuzzy entropy of sleep stages

Fuzzy entropy	Wake	N1	N2	N3	REM
sc4002e0	8.491 8	6.382 9	4.693 9	4.275 5	6.701 7
sc4012e0	9.172 8	7.340 6	5.712 7	3.890 5	7.048 4
sc4102e0	8.562 1	5.830 2	5.146 8	4.277 8	7.609 4

对睡眠脑电数据进行睡眠自动分期。由于受试者的测试时间长度将近24 h,考虑到选取的样本数据必须包含每个睡眠分期,本文选择晚上9:30到次日早上7:00的脑电睡眠数据,每个受试者包含1 140个睡眠分

期。选取2/3的睡眠数据训练SVM,剩下的睡眠数据用于测试分类。通过与睡眠专家人工标注的结果对比,可以计算出本文睡眠自动分期的准确率。具体结果如表3所示。

表3 基于SVM的分类结果(%)

Tab.3 Classification result based on support vector machine (%)

Accuracy	sc4002e0	sc4012e0	sc4102e0	Average
Energy feature	84.78	82.43	83.12	83.44
Fuzzy entropy and energy feature	89.62	88.57	87.16	88.45

从表3中可以看出,仅选择单一的能量特征作为特征值,睡眠分期准确率最高为84.78%。如果在睡

眠自动分期中,相比于睡眠分期准确率,更在乎实时的数据处理,则可以只选择特征波的能量特征作为

分类器的输入。很明显,能量特征加上模糊熵作为分类器的输入,其睡眠分期结果更为准确。同时本文将标准的30 s睡眠脑电数据分为多个短数据进行处理,在一定程度上减少了模糊熵的运算量,提高了运算速度。

3 结论

脑电特征波的提取方法有小波分解,小波包分解,EMD等。小波分解和小波包分解受采样频率的影响,只能得到粗略的特征波,EMD分解得到的IMF更是没有相应的频率范围,但本文使用的FIR带通滤波器可以提取精确的特征波。脑电信号的特征有很多,但脑电信号的频域特性最为明显,且特征波获取方便、数据处理量小,所以本文选择提取脑电信号的特征波能量特征。对于睡眠分期效果而言,分期准确率是很重要的评价指标,因此本文增加了一个非线性特征模糊熵,并将其进行分段处理,提高运算速度。实验结果表明,本文所使用的能量特征和模糊熵特征简单、有效,同时在一定程度上平衡了运算速度和分期准确率。

【参考文献】

- [1] 严由伟,刘明艳,唐向东,等.压力源及其与睡眠质量的现象学关系研究述评[J].心理科学进展,2010,18(10):1537-1547.
YAN Y W, LIU M Y, TANG X D, et al. On the stressors and its phenomenological relationship with sleep quality[J]. Advances in Psychological Science, 2010, 18(10): 1537-1547.
- [2] IBER C. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications[M]. American Academy of Sleep Medicine, 2007.
- [3] DIYKH M, LI Y, WEN P. EEG sleep stages classification based on time domain features and structural graph similarity [J]. IEEE Trans Neural Syst Rehabil Eng, 2016, 24(11):1159-1168.
- [4] HASSAN A R, BHUIYAN M I. Automatic sleep scoring using statistical features in the EMD domain and ensemble methods[J]. Biocybern Biomed Eng, 2016, 36(1): 248-255.
- [5] ŞEN B, PEKER M, ÇAVUŞOĞLU A, et al. A comparative study on classification of sleep stage based on EEG signals using feature selection and classification algorithms[J]. J Med Syst, 2014, 38(3): 18.
- [6] HUANG C S, LIN C L, KO L W, et al. A hierarchical classification system for sleep stage scoring via forehead EEG signals [C]// Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2013 IEEE Symposium on. IEEE, 2013: 1-5.
- [7] KOLEY B, DEY D. An ensemble system for automatic sleep stage classification using single channel EEG signal[J]. Comput Biol Med, 2012, 42(12): 1186-1195.
- [8] GÜNEŞ S, POLAT K, YOSUNKAY Ş. Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting[J]. Expert Syst Appl, 2010, 37(12): 7922-7928.
- [9] KUMAR Y, DEWAL M L, ANAND R S. Features extraction of EEG signals using approximate and sample entropy [C]//Electrical, Electronics and Computer Science (SCECS), 2012 IEEE Students' Conference on. IEEE, 2012: 1-5.
- [10] WEISS B, CLEMENS Z, BÓDIZS R, et al. Comparison of fractal and power spectral EEG features: effects of topography and sleep stages [J]. Brain Res Bull, 2011, 84(6): 359-375.
- [11] KHALIGHI S, SOUSA T, PIRES G, et al. Automatic sleep staging: a computer assisted approach for optimal combination of features and polysomnographic channels[J]. Expert Syst Appl, 2013, 40(17): 7046-7059.
- [12] KAYIKCIOĞLU T, MALEKI M, EROĞLU K. Fast and accurate PLS-based classification of EEG sleep using single channel data [J]. Expert Syst Appl, 2015, 42(21): 7825-7830.
- [13] HASSAN A R, BHUIYAN M I. Computer-aided sleep staging using complete ensemble empirical mode decomposition with adaptive noise and bootstrap aggregating[J]. Biomed Signal Process Control, 2016, 24: 1-10.
- [14] CORREA A G, LEBER E L. An automatic detector of drowsiness based on spectral analysis and wavelet decomposition of EEG records [C]//2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. IEEE, 2010: 1405-1408.
- [15] ABOALAYON K A, FAEZIPOUR M. Multi-class SVM based on sleep stage identification using EEG signal [C]//Healthcare Innovation Conference (HIC), 2014 IEEE. IEEE, 2014: 181-184.
- [16] PINCUS S M, GLADSTONE I M, EHRENKRANZ R A. A regularity statistic for medical data analysis[J]. J Clin Monit, 1991, 7(4): 335-345.
- [17] GOLDBERGER A L, AMARAL L A, GLASS L, et al. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals[J]. Circulation, 2000, 101(23): e215-e220.
- [18] HSU Y L, YANG Y T, WANG J S, et al. Automatic sleep stage recurrent neural classifier using energy features of EEG signals[J]. Neurocomputing, 2013, 104: 105-114.
- [19] RADHA M, GARCIA MOLINA G, POEL M, et al. Comparison of feature and classifier algorithms for online automatic sleep staging based on a single EEG signal [C]//2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2014: 1876-1880.
- [20] 李新战,赵震宇.支持向量机基础及其应用前景[J].科技信息,2009(17): 39.
LI X Z, ZHAO Z Y. Support vector machine and its application prospect[J]. Science & Technology Information, 2009(17): 39.

(编辑:薛泽玲)