

基于相位同步的癫痫信号识别与分析

周梦妮^a, 牛 焱^a, 曹 锐^b, 阎鹏飞^a, 相 洁^a

(太原理工大学 a. 信息与计算机学院; b. 软件学院, 太原 030024)

摘 要: 针对临床人工诊断癫痫信号效率低下的问题, 建立一种基于相位同步的癫痫信号自动诊断模型。使用相位锁定值衡量各脑区间不同状态下的同步化程度, 构建对应的脑功能网络连接矩阵, 提取聚类系数和特征路径长度 2 种全局属性作为输入支持向量机的训练特征, 使用六折交叉验证的方式对发作间期及发作期的信号进行分类识别。实验结果表明, 加权网络的分类效果优于二值网络, 其平均准确率为 83.4%, 单一属性难以全面反映癫痫患者 2 种状态下的功能网络连接差异, 多数患者在 gamma 和 beta 频段取得较好的分类效果。

关键词: 癫痫; 相位锁定值; 同步; 聚类系数; 特征路径长度; 支持向量机

中文引用格式: 周梦妮, 牛焱, 曹锐, 等. 基于相位同步的癫痫信号识别与分析[J]. 计算机工程, 2019, 45(7): 291-295, 302.

英文引用格式: ZHOU Mengni, NIU Yan, CAO Rui, et al. Recognition and analysis of epileptic signal based on phase synchronization[J]. Computer Engineering, 2019, 45(7): 291-295, 302.

Recognition and Analysis of Epileptic Signal Based on Phase Synchronization

ZHOU Mengni^a, NIU Yan^a, CAO Rui^b, YAN Pengfei^a, XIANG Jie^a

(a. College of Information and Computer; b. College of Software, Taiyuan University of Technology, Taiyuan 030024, China)

[Abstract] To address the low efficiency of clinical manual diagnosis of epilepsy, this paper establishes an automatic diagnosis model of epilepsy signal based on phase synchronization. First, the model use Phase Locking Value (PLV) to measure the degree of synchronization of brain regions in different states and constructs a corresponding brain function network connection matrix. Then, the two global attributes, clustering coefficient and characteristic path length, are extracted as training features to input onto Support Vector Machine (SVM). Finally, 6-fold cross-validation method is used for the classification and identification of interictal and ictal signals. Experimental results show that the classification effect of the weighted network is better than that of the binary network. The average accuracy of the weighted network is 83.4%. Single attribute is not enough to fully reflect the difference in functional network connections in two states of epilepsy, and most patients achieve better classification results in gamma and beta bands.

[Key words] epilepsy; Phase Locking Value (PLV); synchronization; clustering coefficient; characteristic path length; Support Vector Machine (SVM)

DOI: 10.19678/j.issn.1000-3428.0051201

0 概述

癫痫是一种常见的脑部疾病, 越来越多的人在长期忍受着其带来的巨大伤害。癫痫的发作具有突发性、反复性, 给患者及家属造成强烈的精神痛苦, 降低其生活质量^[1-2]。当癫痫患者的大脑活动出现异常时, 在记录的脑电信号中往往会出现异常的癫痫样放电, 包括棘波、尖波、尖慢复合波、棘慢复合波及高度节律失调特征波。棘波具有尖锐的波形, 多

数会出现在大发作或者局限性发作癫痫中。尖波和棘波的机理相同, 时限长于棘波, 反映放电的同步化程度。尖慢复合波和棘慢复合波在不同部位或者时间出现时, 表明可能存在多个异常的电活动区。高度节律失调多发生在发作间期, 是一种异常的背景信号。目前, 这些异常信号的诊断仍由医生通过视觉观察, 根据长期的工作经验完成。这一工作不仅需要消耗医生大量的时间和精力, 且准确率低, 不同的医生难以达成一个共同判断的标准, 主观性强。

基金项目: 国家自然科学基金 (61503272, 61305142, 61373101, 61741212); 中国博士后科学基金 (2016M601287); 山西省自然科学基金青年项目 (2015021090, 201601D202042); 山西省回国留学人员科研项目 (2016-037)。

作者简介: 周梦妮 (1992—), 女, 硕士研究生, 主研方向为智能信息处理; 牛 焱, 博士研究生; 曹 锐, 博士; 阎鹏飞, 讲师、博士; 相 洁 (通信作者), 教授、博士、博士生导师。

收稿日期: 2018-04-13

修回日期: 2018-06-29

E-mail: xiangjie@tyut.edu.cn

因此,自动识别癫痫脑电信号可以帮助医生减轻工作负担、辅助临床治疗,具有重要的现实意义^[3]。

近年来,脑电信号的识别研究多通过提取时域、频域、时频域、线性^[4]、非线性^[5-6]等特征,来反映大脑从一种状态过渡到另一种状态的变化过程^[7]。文献[8]表明,在大脑多个脑区相互作用的过程中,发生同步现象导致癫痫发作。当癫痫即将发作时,癫痫样放电开始通过患者大脑内部各个传导通路向周围的脑区扩散,进而经过一些神经环路返回到放电开始的地方,形成一个闭合回路。这种现象不停循环,将正常状态时大脑的随机放电转变成持续有规律的节律性放电。这样的发作机制说明,在发病过程中大脑各个脑区之间有一定的关联关系,而上述特征并未充分考虑这种关联关系。因此,通过同步化分析整个大脑,更能真实反映在临床上癫痫发作时各脑区之间相互作用的变化情况。

常见的分析大脑功能连接强度的相位方法有基于相位锁定值(Phase Locking Value, PLV)方法和基于相位延迟指数(Phase Lag Index, PLI)方法。两者相比,PLV 对于频率成分更加敏感,可以更好地克服脑电产生过程中的容积效应^[9]。而癫痫发作不仅与不同脑区的神经活动有关,也与人脑不同的节律波有关^[10]。因此,PLV 更适合作为分析癫痫信号大脑功能连接强度的衡量指标。

目前已有一些研究使用 PLV 来分析常见神经退行性疾病患者不同状态时的大脑活动差异以及将病患者与正常人的大脑活动进行比较^[8,11]。2015 年,文献[12]对心因性非癫痫患者和正常对照组通过 PLV 功能连接强度建立关联矩阵,进而分析聚类系数和全局效率 2 种特征,使用线性判别分析方法区分病患和正常对照组,其准确率为 85%。2016 年,文献[9]使用 PLV 通过复杂网络可视化的方式,发现癫痫发作网络在不同频率波段中关键节点的部位和连接强度均有所差异。

这些研究使用 PLV 方法对癫痫患者不同状态下的大脑活动进行分析,但是该方法对癫痫信号的识别能力尚不清楚。本文通过提取癫痫患者脑电信号的各个波段时间序列,分析各波段下任意 2 个通道之间的相位同步性,构建癫痫患者在发作间期与发作期 2 种不同状态下的功能连接网络。提取全局聚类系数和特征路径长度作为分类特征,输入支持向量机(Support Vector Machine, SVM)进行模型训练,进而识别患者在不同状态下的脑电信号。对比二值网络与加权网络对信号识别的影响,比较各频

段信号在分类效果上的差异,以此分析利用 PLV 构建复杂网络识别癫痫信号的可能性。

1 网络构建及特征提取

1.1 数据及预处理

本文使用来自美国波士顿儿童医院的癫痫数据库,共包含 24 名患者(编号 1 ~ 编号 24)的数据,这些患者在实验前都停止使用抗癫痫药物。头皮电极采用国际通用的 10/20 系统放置法,采样频率为 256 Hz。收集的每段数据都经过专业医生依据经验标注癫痫发作开始与结束时间。在信号采集过程中易受噪声干扰,患者也会出现肌肉运动、眼球运动和眨眼等行为,导致在原始信号中出现部分干扰幅值。因此,使用 50 Hz 陷波滤波器去除工频噪声,采用截止频率为 1 Hz ~ 60 Hz 的带通滤波器去除高频信号。每个患者分别选取长度为 1 s 的若干发作期和发作间期片段,通过 EEGLAB 工具箱从预处理过的信号中依次提取 delta (1 Hz ~ 4 Hz)、theta (4 Hz ~ 8 Hz)、alpha (8 Hz ~ 13 Hz)、beta (13 Hz ~ 30 Hz)、gamma (30 Hz ~ 60 Hz) 5 个频段的成分。

1.2 相位同步性分析

相位同步可以用来描述 2 个时间序列之间的关系,通过分析相位同步可以发现大脑不同脑区间的同步化程度^[13]。PLV 是常见的度量网络功能连接强度的指标之一,其值计算如下:

$$PLV(t) = \frac{1}{N} \left| \sum_{n=1}^N \exp(j(\Delta\varphi_n(t))) \right| \quad (1)$$

其中, $\Delta\varphi_n(t) = \varphi_x(t) - \varphi_y(t)$ 表示 2 个时间序列在 t 时刻的瞬时相位差, N 为时间序列的时间点个数。由式(1)可知, PLV 取值范围为 $[0, 1]$, 并且其值越大, 则 2 个信号间的同步化程度越高。

1.3 功能网络构建

构建二值网络需要确定节点和边,而构建加权功能网络需要确定节点、边以及边的权重。本文将各个通道作为节点,用 PLV 计算 2 个通道的功能连接,若 PLV 值不为 0,则这 2 个通道之间存在一条边,且 PLV 值就是其权重^[14]。若是二进制网络,则需要选取一定的稀疏度,将 PLV 关联矩阵转换成二值邻接矩阵。本文二进制网络的稀疏度选取原则是在保证网络连通性的情况下,每个频段选取其最大阈值(即大于该值,则会出现孤立点,此时特征路径长度将会出现无穷大),然后在所有阈值中取最小值,即为整个网络的连通阈值。

1.4 网络属性提取

本文将聚类系数 C 、特征路径长度 L 2 种全局属性作为衡量网络的指标(训练特征),利用 SVM 使用六折交叉验证的方式进行分类。

聚类系数 C 指网络中各个节点间的聚集程度,用来衡量网络的信息传输能力^[15]。对于二值网络,节点 i 的聚类系数 C_i 定义如下:

$$C_i = \frac{E_i}{(k_i(k_i - 1))/2} \quad (2)$$

其中, k_i 是节点 i 的度, E_i 是节点 i 的 k_i 个相邻节点之间实际存在的边数。则式(2)的分母是所有节点间可能存在的最大边数。该网络的聚类系数 C 为所有节点聚类系数的平均值。

对于加权无向网络,需要考虑节点 i 与其相邻节点之间边的权值影响,其聚类系数 \tilde{C}_i 定义如下:

$$\tilde{C}_i = \frac{1}{k_i(k_i - 1)} \sum_{j,k} \frac{w_{ij} + w_{ik}}{2} a_{ij} a_{ik} a_{jk} \quad (3)$$

$$\langle w_i \rangle = \frac{\sum_j w_{ij}}{k_i} \quad (4)$$

其中,节点 j 和 k 是节点 i 的邻接节点, $a_{ij} = 1$ 代表节点 i 和节点 j 之间有边,否则 $a_{ij} = 0$ 。 a_{ik} 、 a_{jk} 的含义以此类推, w_{ij} 表示节点 i 和 j 之间边的权值。

特征路径长度 L 定义为任意 2 个节点之间的距离的平均值^[16],对于无权网络其计算如下:

$$L = \frac{1}{\frac{1}{2}N(N-1)} \sum_{i \geq j} d_{ij} \quad (5)$$

其中, d_{ij} 表示节点 i 和 j 之间的最短路径上的边的数目。若是加权网络, d_{ij} 表示节点 i 和 j 之间最短路径对应的权值。

2 实验结果与分析

2.1 PLV 关联矩阵及属性分布

根据计算得到的 PLV,对每个样本构建功能连接的关联矩阵。以 1 号患者被测试的 gamma 频段为例,其包含 23 个通道的信号,因此可以得到 23×23 的 PLV 关联矩阵,其矩阵图如图 1 所示。从图 1 可以看出,发作间期的脑功能网络连接强度明显要高于发作期。对关联矩阵构成的加权网络进行 t 检验,结果为 $p = 0.009$ ($p < 0.05$),因此,可以判定发作间期和发作期的加权网络存在显著差异。其他频段也具有类似的结果。

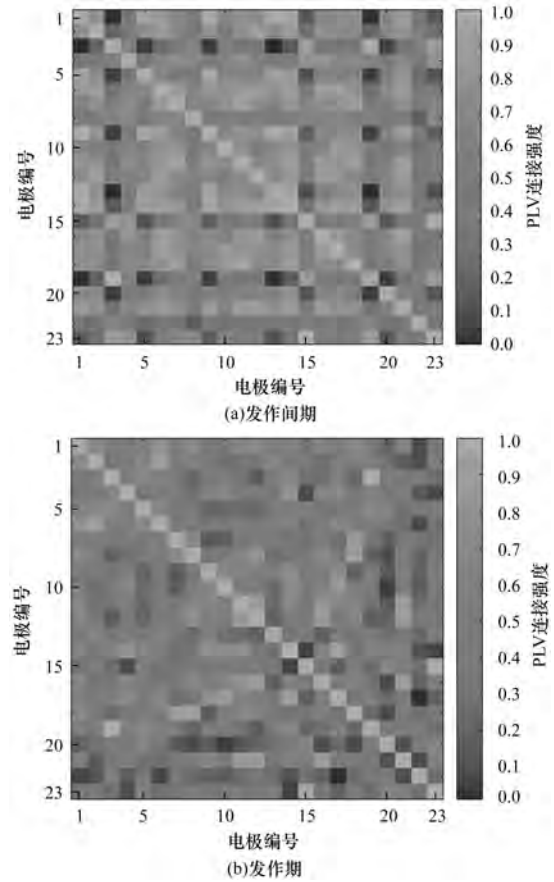


图1 gamma 频段 PLV 关联矩阵图

图 2 为在加权网络中各个频段 2 种属性分布的盒形图。

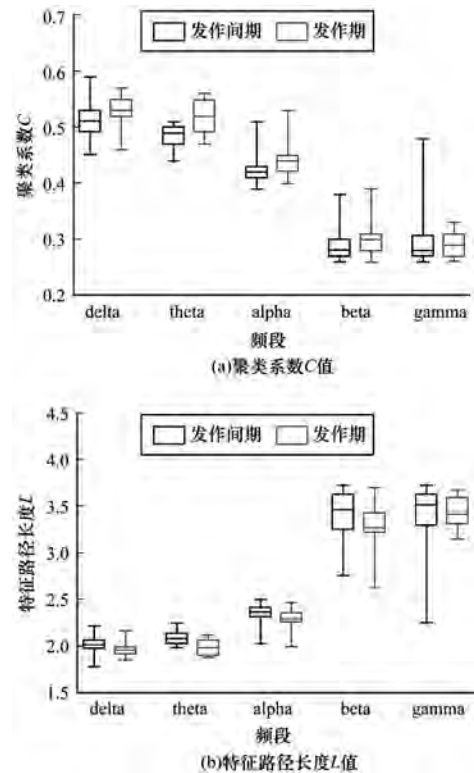


图2 各频段属性值分布

从图 2 可以看出, delta、theta 和 alpha 3 个频段的属性值分布比较接近, beta 和 gamma 2 个频段的属性范围比较接近。无论是发作间期还是发作期, 聚类系数 C 都会随着频段范围的增高而降低, 特征路径长度 L 则呈现相反的变化趋势。

2.2 分类结果对比

图 3 为对各个患者使用所有频段的 C 和 L 进行六折交叉验证后得到的准确率。从图 3 可以看出, 就单个患者而言, 其最高准确率达到 96.18%, 并且 83% 患者的准确率都超过 80%, 其中二值网络的平均准确率是 77.15%, 而加权网络的平均准确率却达到 83.4%。几乎所有患者的加权网络准确率都高于二值网络的准确率, 并且在部分患者上表现出较大的差距(如 14 号、23 号)。上述结果表明, PLV 加权网络能更好地反映癫痫患者各脑区间在不同状态下的相位同步差异。因此, 之后的实验选择加权网络来进行研究。

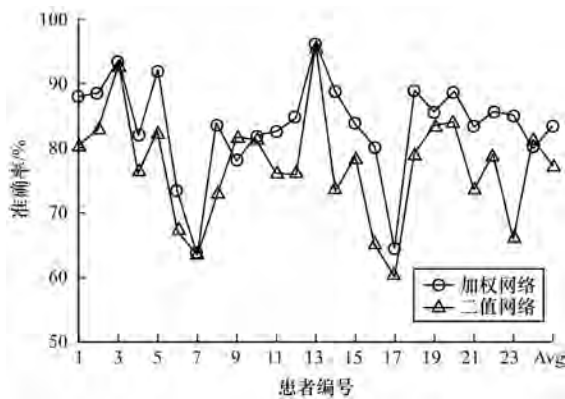


图 3 加权网络与二值网络准确率对比

2.3 属性分类结果对比

图 4 为分别使用聚类系数 C 和特征路径长度 L 进行六折交叉验证后所得准确率。

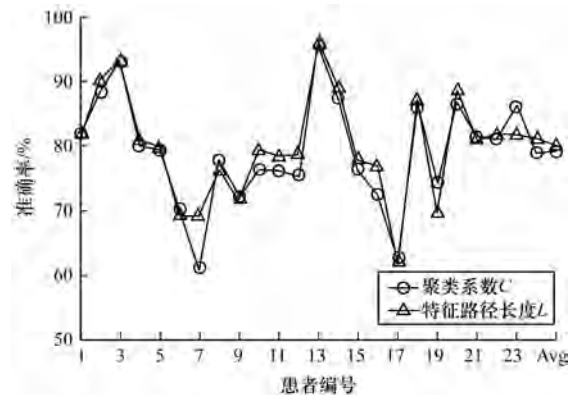


图 4 2 种属性准确率对比

聚类系数 C 的平均分类准确率是 79.22%，而特征路径长度 L 的分类准确率为 80.07%。显然, 这 2 种属性在发作间期与发作期的识别效果上表现几乎一致。但与图 3 中同时使用 2 种属性作为训练特征相比, 其效果还要差一些。这说明单一属性难以全面反映癫痫患者 2 种状态下的功能网络连接差异。

2.4 各波段分类结果对比

表 1 是将聚类系数和特征路径长度同时作为训练特征, 对各频段单独分类所得的结果。

表 1 各频段分类准确率

患者编号	delta	theta	alpha	beta	gamma	最优准确率	最优频段
1	69.83	65.00	61.50	66.33	79.83	79.83	gamma
2	50.58	77.91	77.33	85.76	75.87	85.76	beta
3	56.67	68.33	67.00	81.83	75.00	81.83	beta
4	46.08	63.24	51.96	81.37	57.84	81.37	beta
5	67.50	51.67	68.50	88.83	83.00	88.83	beta
6	58.47	51.99	62.82	56.25	69.97	69.97	gamma
7	51.04	60.94	47.40	66.15	59.90	66.15	beta
8	62.50	62.00	58.33	66.33	78.33	78.33	gamma
9	74.71	55.02	47.50	46.81	71.44	74.71	delta
10	67.25	59.03	70.91	77.41	42.94	77.41	beta
11	55.83	72.50	56.17	69.17	74.17	74.17	gamma
12	61.00	64.67	63.17	78.17	69.50	78.17	beta
13	66.49	50.35	55.56	74.48	96.70	96.70	gamma
14	60.96	67.47	61.24	69.23	82.84	82.84	gamma
15	56.17	68.33	54.17	74.00	74.17	74.17	gamma
16	66.64	57.35	50.04	49.35	60.06	66.64	delta
17	58.54	48.81	51.53	54.78	53.24	58.54	delta
18	66.00	72.67	56.33	71.83	77.33	77.33	gamma
19	61.86	55.30	44.07	71.19	84.32	84.32	gamma
20	77.55	52.04	61.39	75.68	57.82	77.55	delta
21	53.26	54.28	73.62	76.89	66.34	76.89	beta
22	68.14	65.93	61.76	67.16	79.41	79.41	gamma
23	65.94	66.98	58.79	68.61	76.96	76.96	gamma
24	51.00	72.50	61.00	69.00	65.83	72.50	theta
平均值	61.42	61.85	59.25	70.28	71.37	77.52	gamma

从表 1 可以看出,同一患者在不同频段上分类表现差异较大,不同患者在同一频段上也表现出不同的效果。其中最高准确率为 96.7%,最低却只有 58.54%。从最优准确率所分布的频段来看,有 11 位患者在 gamma 频段分类效果最佳,8 位患者在 beta 频段分类效果最佳,4 位患者在 delta 频段分类效果最佳,1 位患者在 theta 频段分类效果最佳。上述结果表明,在癫痫发作时各患者、各频段大脑活动都存在明显差异,且多数患者在 gamma 和 beta 频段获得较高的分类准确率。这与发病时患者所处的状态(如睡眠、压力)、癫痫类型、病灶位置等有较大的关系。

2.5 对比实验

在数据集中包含完整头皮脑电信号数据的共有 17 名患者,本文使用排列模糊熵和模糊熵在相同的实验条件下对其进行了发作间期与发作期的信号识别实验^[17]。本文使用 AUC 指标来对比 SVM 分类结果的优劣,AUC 为 ROC 曲线下面积,其值越大,表示分类表现越优。通过表 2 将本文实验所得 AUC 值与之前实验相比。从表 2 可知,17 名患者的模糊熵和排列模糊熵的平均 AUC 分别是 0.914、0.928,而本文实验平均 AUC 是 0.925。大量的研究已经证明,熵在癫痫的信号识别上已经表现出较高的稳定性和区分能力,而本文实验结果说明,复杂网络的方法在区分癫痫发作期与发作间期的能力上,并不亚于熵这一类非线性方法,且这种分析方法更符合癫痫的发病机理。

表 2 不同模型 AUC 值对比

患者编号	模糊熵	排列模糊熵	本文模型
1	0.988	0.988	0.944
2	1.000	0.986	0.969
3	0.924	0.957	0.978
5	0.931	0.937	0.971
8	0.937	0.970	0.907
11	0.939	0.978	0.899
12	0.908	0.939	0.927
13	0.889	0.919	0.971
14	0.866	0.884	0.945
16	0.869	0.896	0.882
17	0.883	0.891	0.761
18	0.899	0.906	0.938
19	0.884	0.900	0.932
20	0.902	0.879	0.955
21	0.883	0.886	0.899
22	0.895	0.909	0.930
24	0.945	0.959	0.914
平均值	0.914	0.928	0.925

3 结束语

本文构建脑功能网络连接矩阵,提取合理的网络属性作为训练特征,利用 SVM 进行六折交叉验证,达到信号识别的目的。从网络的权重、属性、信号频段 3 个角度依次分析其在分类效果上的影响。实验结果表明,加权网络的分类效果较优,聚类系数和特征路径长度在分类表现上差异不大,多数患者在 gamma 和 beta 频段取得比其他频段更好的分类效果。下一步将分析患者的发病类型、病灶位置等对分类结果的影响,提高分类精度。

参考文献

- [1] ZHANG Zisheng, PARHI K K. Low-complexity seizure prediction from IIEG/sEEG using spectral power and ratios of spectral power [J]. IEEE Transactions on Biomedical Circuits and Systems, 2016, 10(3): 693-706.
- [2] SHARIF B, JAFARI A H. Prediction of epileptic seizures from EEG using analysis of ictal rules on Poincaré plane [J]. Computer Methods and Programs in Biomedicine, 2017, 145: 11-22.
- [3] AARUNI V C, HARSHA A, JOSEPH L A. Classification of EEG signals using fractional calculus and wavelet support vector machine [C]//Proceedings of IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems. Washington D. C., USA: IEEE Press, 2015: 1-5.
- [4] RASHID M M, AHMAD M. Epileptic seizure classification using statistical features of EEG signal [C]//Proceedings of International Conference on Electrical, Computer and Communication Engineering. Washington D. C., USA: IEEE Press, 2017: 308-312.
- [5] XIANG Jie, LI Conggai, LI Haifang, et al. The detection of epileptic seizure signals based on fuzzy entropy [J]. Journal of Neuroscience Methods, 2015, 243: 18-25.
- [6] LI Jing, YAN Jiaqing, LIU Xianzeng, et al. Using permutation entropy to measure the changes in EEG signals during absence seizures [J]. Entropy, 2014, 16(6): 3049-3061.
- [7] ASSI E B, NGUYEN D K, RIHANA S, et al. Towards accurate prediction of epileptic seizures: a review [J]. Biomedical Signal Processing and Control, 2017, 34: 144-157.
- [8] 董国亚, 陈小刚, 史姗姗, 等. 相位同步性分析方法在癫痫研究中的应用 [J]. 北京生物医学工程, 2011, 30(3): 274-277.
- [9] 何艳, 于云莉, 杨帆, 基于复杂网络可视化的癫痫患者大脑状态研究 [J]. 中国医疗设备, 2016, 31(9): 39-42.
- [10] 裘嘉恒, 李雅堂, 许坤涵, 等. 癫痫发作间期 alpha 波的窄带相位同步分析 [J]. 生物物理学报, 2008, 24(3): 221-226.
- [11] 李红利, 王江, 邓斌. 癫痫脑电的互信息和同步性分析 [J]. 计算机工程与应用, 2013, 49(6): 19-22.
- [12] 王振宇, 薛青, 熊秀春. 基于静息态脑电的心因性非癫痫性发作患者脑功能网络分析及分类识别研究 [J]. 生物医学工程学杂志, 2015(1): 8-12.

(下转第 302 页)

(上接第 295 页)

- [13] ELAHIAN B, YEASIN M, MUDIGOUDAR B, et al. Identifying seizure onset zone from electrocorticographic recordings: a machine learning approach based on phase locking value[J]. *Seizure*, 2017, 51: 35-42.
- [14] SUPRIYA S, SIULY S, WANG Hua, et al. Weighted visibility graph with complex network features in the detection of epilepsy [J]. *IEEE Access*, 2016, 4: 6554-6566.
- [15] GAO Zhongke, CAI Qing, YANG Yuxuan, et al. Visibility graph from adaptive optimal kernel time - frequency representation for classification of epileptiform EEG[J]. *International Journal of Neural Systems*, 2017, 27(4).
- [16] 刘海红. 基于复杂网络统计特性的非线性时间序列分析方法研究[D]. 济南: 济南大学, 2017.
- [17] 王慧云, 窦大庆, 曹锐, 等. 用于癫痫 EEG 分析的排列模糊熵新算法[J]. *太原理工大学学报*, 2017, 48(1): 91-96.

编辑 刘盛龄