

基于振幅-周期二维特征的脑电样本熵分析

郭家梁 钟宁 马小萌 张明辉 周海燕

Sample entropy analysis of electroencephalogram based on the two-dimensional feature of amplitude and period

Guo Jia-Liang Zhong Ning Ma Xiao-Meng Zhang Ming-Hui Zhou Hai-Yan

引用信息 Citation: *Acta Physica Sinica*, 65, 190501 (2016) DOI: 10.7498/aps.65.190501

在线阅读 View online: <http://dx.doi.org/10.7498/aps.65.190501>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2016/V65/I19>

您可能感兴趣的其他文章

Articles you may be interested in

风力发电机自循环蒸发内冷系统稳定性的研究

Static bifurcation analysis of natural circulation inner evaporative cooling system in wind turbine

物理学报.2016, 65(3): 030501 <http://dx.doi.org/10.7498/aps.65.030501>

基于新曝光冲突性消息的网络舆论逆转研究

Newly exposed conflicting news based network opinion reversal

物理学报.2016, 65(3): 030502 <http://dx.doi.org/10.7498/aps.65.030502>

基于最大熵模型的微博传播网络中的链路预测

Link prediction in microblog retweet network based on maximum entropy model

物理学报.2016, 65(2): 020501 <http://dx.doi.org/10.7498/aps.65.020501>

考虑谣言清除过程的网络谣言传播与抑制

Propagation and inhibition of online rumor with considering rumor elimination process

物理学报.2015, 64(24): 240501 <http://dx.doi.org/10.7498/aps.64.240501>

一种基于用户相对权重的在线社交网络信息传播模型

An information spreading model based on relative weight in social network

物理学报.2015, 64(5): 050501 <http://dx.doi.org/10.7498/aps.64.050501>

基于振幅-周期二维特征的脑电样本熵分析*

郭家梁¹⁾²⁾³⁾⁴⁾ 钟宁¹⁾²⁾³⁾⁴⁾⁵⁾ 马小萌¹⁾²⁾³⁾⁴⁾ 张明辉¹⁾²⁾³⁾⁴⁾ 周海燕¹⁾²⁾³⁾⁴⁾†

1) (北京工业大学, 北京未来网络科技高精尖创新中心, 北京 100124)

2) (北京工业大学, 国际WIC研究院, 北京 100124)

3) (磁共振成像脑信息学北京市重点实验室, 北京 100124)

4) (脑信息智慧服务北京市国际科技合作基地, 北京 100124)

5) (前桥工业大学生命科学与信息工程系, 前桥 371-0816)

(2016年5月10日收到; 2016年7月10日收到修改稿)

样本熵可以有效反映一维时间序列中新模式的生成概率, 但缺乏对二维序列复杂度的表征能力. 基于对传统样本熵方法的改进, 提出一种在振幅-周期二维空间描述波形复杂度的方法, 二维样本熵反映了波形振动在振幅-周期空间中新模式的生成概率. 通过仿真实验证明了这种方法描述波形复杂度的有效性, 当波形的复杂度特征表现为振幅-周期的交互作用时, 二维样本熵对复杂度的描述比一维条件下的样本熵更加有效. 基于二维样本熵对抑郁症组和对照组的脑电复杂度进行分析, 结果表明, 抑郁症组在Alpha频段左侧顶区和左侧枕区的二维样本熵显著低于对照组, 表明在上述频段和位置, 抑郁症患者脑电中新模式的生成概率显著低于正常人, 这一特征可能成为抑郁症的潜在生物标记.

关键词: 样本熵, 二维, 脑电, 抑郁症

PACS: 05.10.-a, 89.70.Cf, 87.19.le

DOI: 10.7498/aps.65.190501

1 引言

近似熵 (approximate entropy, ApEn) 是 20 世纪 90 年代初由 Pincus 提出的, 其物理意义是当系统中的模式维数发生变化时, 时间序列产生出新模式的概率大小, 产生出新的模式概率越大, 系统复杂度越大, ApEn 值也就会越大^[1]. 为消除 ApEn 中向量自身比较造成的误差, Richman 等^[2] 对传统 ApEn 算法进行了改进, 提出计算时间序列复杂度的样本熵 (sample entropy, SampEn) 算法, SampEn 是条件概率平均数的自然对数的精确值, 因此不依赖于数据长度的影响, 对短时间序列进行分析时更加精确, 因此, SampEn 方法在脑电^[3-5]、心电图^[6] 生物信号的分析中均有广泛应用.

脑电是大量神经元突触后电位同步运动在头皮处的总体反应, 被认为是一种高度复杂的非线性系统, 因此李雅普诺夫指数^[7]、相关维数^[8]、Lempel-Ziv 复杂度^[9]、分形维数等^[10,11] 非线性方法均被引入到了脑电信号的分析中. 因为脑电中包含大量的生理和病理学信息, 样本熵、多尺度熵等分析方法也被用于癫痫^[12]、阿尔茨海默症^[13]、精神分裂症^[14]、抑郁症等^[15] 脑部疾病的诊断和病理研究中.

在以往的非线性方法中, 通常将脑电信号作为一维的时间序列进行分析, 虽然利用希尔伯特变换^[16]、小波变换等^[17] 方法可以将时间序列转换到频域进行分析, 但目前还没有对波形的振幅、频率特征综合描述的非线性指标. 最近有研究基于符号动力学理论, 提出了一种将脑电信号转化为瞬时振

* 国家重点基础研究发展计划 (批准号: 2014CB744600)、国家国际科技合作专项 (批准号: 2013DFA32180) 和国家自然科学基金 (批准号: 61420106005, 61272345) 资助的课题.

† 通信作者. E-mail: zhouhaiyan@bjut.edu.cn

幅和瞬时频率序列进行分析的方法^[18],但由于方法上的限制,上述研究只对频率序列和振幅序列的复杂度分别进行了分析,而没有对振幅-频率的总体特征进行研究.在基于样本熵的方法中,多尺度熵、多变量多尺度熵^[19]虽然可以从不同时间尺度和多变量角度对波形复杂度进行描述,但这些方法均不能有效反映不同特征间交互作用对波形复杂度的影响.

本研究基于对传统样本熵的改进,提出了一种可以从多特征交互作用角度分析波形复杂度的方法——二维样本熵,并通过仿真实验对这种方法的有效性进行了验证.基于抑郁症患者和正常人两组脑电数据,以二维样本熵为指标对两组脑电信号的复杂度进行了分析.

2 基本原理

2.1 时间序列在二维空间中的表达

对于时间序列来说,波形的振动是由两种运动构成的,即由波形极小值到极大值的单调递增运动,以及由极大值到极小值的单调递减运动,这两种运动构成了波形振动的基本单位.波形的每次单调运动包含振动幅度和持续时间两种信息,因此,我们用“瞬时振幅”表示单次单调运动中的幅值变化,用“瞬时周期”表示单次单调运动的持续时间.因此,连续的振动波形可以被抽象为瞬时振幅和瞬时周期两种参数序列,波形的振动信息可以用这两种参数序列组成的向量序列来描述.由时间序列转化为二维振动序列的步骤如下.

1) 提取极值序列和极值时间序列: 设有长度为 n 的时间序列 $S = \{s(i)\}$, 其中 $i = 1, 2, 3, \dots, n$, $n \geq 2$. 提取时间序列的极值 $e(j)$ 组成极值序列 $E = \{e(j)\}$, 极值发生的时间 $t(j)$ 组成极值时间序列 $T = \{t(j)\}$, 其中 $j = 1, 2, 3, \dots, p$, 表示极值发生的序号, p 为序列 S 中的极值个数.

2) 构造瞬时振幅序列和瞬时周期序列: 构造序列 $\{0, e(j)\}$ 和 $\{e(j), 0\}$, 设 $A = |\{0, e(j)\} - \{e(j), 0\}|$, 去除 A 的首项和尾项, 得到由单次极值运动到下次极值时的幅度变化序列 $A = \{a(k)\}$, 以同样方法得到由单次极值运动到下次极值的时间长度序列 $C = \{c(k)\}$, 其中 $k = 1, 2, 3, \dots, q$, 表示单调运动的序号, q 表示序列 S 中的单调运动个数, 序列 A 和 C 分别表征了序列 S 在一定时间内的振

幅和周期, 因此称 A 为瞬时振幅序列, C 为瞬时周期序列.

3) 构造二维振动序列: 由序列 A 和 C 的对应元素构成二维振动向量 $(a(k), c(k))$, 由振动向量组成二维振动序列 $O = \{(a(k), c(k))\}$, 其中 $k = 1, 2, 3, \dots, q$.

2.2 二维样本熵的计算方法

已知时间序列 S 的二维振动序列为 $O = \{(a(k), c(k))\}$, 振动序列 O 的元素是由多维数据组成的向量, 向量中元素的单位可能是不同的, 因此, 当计算多维向量间的差别时, 需要消除不同元素单位对分析带来的影响. 因此, 本研究引入 Jaccard 距离的概念作为向量间相似性的量度.

Jaccard 距离是由 Jaccard 相似系数计算得到的. Jaccard 相似系数的定义为: 两个集合 B, D 的交集元素在 B, D 的并集中所占的比例, 用符号 $J(B, D)$ 表示:

$$J(B, D) = \frac{|B \cap D|}{|B \cup D|}. \quad (1)$$

Jaccard 距离是与 Jaccard 相似系数相反的概念, 是由两个集合中不同元素占所有元素的比例来表示的, 其公式为

$$J_D(B, D) = 1 - J(B, D) = \frac{|B \cup D| - |B \cap D|}{|B \cup D|}. \quad (2)$$

设已知向量 $(a(u), c(u))$ 和 $(a(v), c(v))$, 两者的 Jaccard 相似系数可表示为: 在振幅-周期的欧氏空间中, 容纳 $(a(u), c(u))$ 的最小矩形 $L(u)$ 与容纳 $(a(v), c(v))$ 的最小矩形 $L(v)$ 的重合面积 ΔL , 与此时 $L(u)$ 和 $L(v)$ 共占空间面积的比值, $(a(u), c(u))$ 与 $(a(v), c(v))$ 的 Jaccard 相似系数可表示为

$$J((a(u), c(u)), (a(v), c(v))) = \frac{\Delta L}{a(u) \times c(u) + a(v) \times c(v) - \Delta L}, \quad (3)$$

其中 $\Delta L = \min(a(u), a(v)) \times \min(c(u), c(v))$, 是矩形 $L(u)$ 和 $L(v)$ 的重合面积, 如图 1 所示. 因此, 向量 $(a(u), c(u))$ 与 $(a(v), c(v))$ 的 Jaccard 距离计算公式为

$$J_D((a(u), c(u)), (a(v), c(v))) = 1 - J((a(u), c(u)), (a(v), c(v))). \quad (4)$$

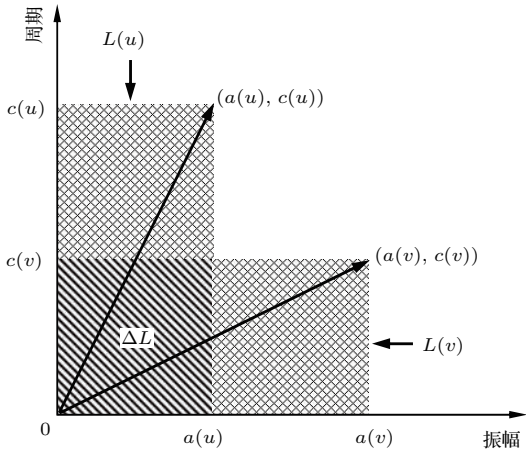


图1 由向量计算 Jaccard 距离的原理图
Fig.1. The method of calculating the Jaccard distance by the vector elements.

因此, 由二维振动序列 \mathbf{O} 计算二维样本熵方法的步骤如下.

1) 相空间重构: 设二维振动序列为 $\mathbf{O} = \{(a(k), c(k))\}$, 其中振幅序列为 $\mathbf{A} = \{a(k)\}$, 周期序列为 $\mathbf{C} = \{c(k)\}$, $k = 1, 2, 3, \dots, q$, $q \in \mathbf{N}$. 从 \mathbf{A} 中提取 $q - m + 1$ 个 m 维向量 $\mathbf{A}(x) = \{a(x), a(x + 1), a(x + 2), \dots, a(x + m - 1)\}$, 从 \mathbf{C} 中提取 $q - m + 1$ 个 m 维向量 $\mathbf{C}(x) = \{c(x), c(x + 1), c(x + 2), \dots, c(x + m - 1)\}$, $x \in \mathbf{N}$, 由 $\mathbf{A}(x)$ 和 $\mathbf{C}(x)$ 组成向量序列: $\mathbf{Q}(x) = \{(a(x), c(x)), (a(x + 1), c(x + 1)), (a(x + 2), c(x + 2)), \dots, (a(x + m - 1), c(x + m - 1))\}$.

2) 计算 $\mathbf{Q}(x)$ 与 $\mathbf{Q}(y)$ 的距离: 向量序列 $\mathbf{Q}(x)$ 与 $\mathbf{Q}(y)$ 的距离表示为: $\mathbf{Q}(x)$ 与 $\mathbf{Q}(y)$ 序列对应向量元素 Jaccard 距离的最大值:

$$D[\mathbf{Q}_m(x), \mathbf{Q}_m(y)] = \max(J_D((a(x + z), c(x + z)), (a(y + z), c(y + z))))), \quad (5)$$

其中 $0 \leq z \leq m - 1$, $1 \leq x \leq q - m + 1$, $1 \leq y \leq q - m + 1$.

3) 设相似容限为 $R(0 < R < 1)$, 对于向量 $\mathbf{X}_m(x)$, 统计出 $D[\mathbf{Q}_m(x), \mathbf{Q}_m(y)] < R$ 的个数, 计算 $\text{num}\{D[\mathbf{Q}_m(x), \mathbf{Q}_m(y)] < R\}$ 与所有距离总数的比值:

$$B_x^m(R) = \frac{1}{n - m} \text{num}\{D[\mathbf{Q}_m(x), \mathbf{Q}_m(y)] < R\}, \quad (6)$$

式中, $1 \leq y \leq n - m + 1$ 且 $y \neq x$.

4) 计算所有向量 $B_x^m(R)$ 的平均值:

$$B^m(R) = \frac{1}{n - m + 1} \sum_{x=1}^{n-m+1} B_x^m(R). \quad (7)$$

5) 将维数由 m 加到 $m + 1$, 重复步骤 1)–4), 计算 $B_x^{m+1}(R)$ 和 $B^{m+1}(R)$.

$$B_x^{m+1}(R) = \frac{1}{n - m - 1} \text{num}\{D[\mathbf{Q}_{m+1}(x), \mathbf{Q}_{m+1}(y)] < R\}, \quad (8)$$

$$B^{m+1}(R) = \frac{1}{n - m} \sum_{x=1}^{n-m} B_x^{m+1}(R). \quad (9)$$

6) 用以下公式计算二维样本熵:

$$D2SEn(m, R) = \lim_{n \rightarrow \infty} \{-\ln[B^{m+1}(R)/B^m(R)]\}. \quad (10)$$

因为在实际情况中, 时间序列单调运动的个数是有限的, 所以 $D2SEn$ 取单调运动个数为 q 的估计值进行分析:

$$D2SEn(m, R, q) = -\ln[B^{m+1}(R)/B^m(R)]. \quad (11)$$

由公式可知, $D2SEn$ 的值与参数 m, R, q 的取值有关. 本研究将主要对维数 $m = 2$ 条件下的 $D2SEn$ 值进行分析.

3 二维振动熵描述波形复杂度有效性的仿真分析

3.1 仿真波形的构建

1) 依据波形特点由正弦波构造仿真波形, 波形的子波函数为

$$s_i(t) = A_i \cdot \cos\left(\frac{2\pi}{T_i} t\right), \quad (12)$$

其中, 当 $i = 1, 3, 5, \dots, 2n - 1$ 时, $t \in [0, \pi)$, 当 $i = 2, 4, 6, \dots, 2n$ 时, $t \in [\pi, 2\pi)$, T_i 和 A_i 分别表示第 i 个正弦子波的周期和振幅.

2) 设时间序列 \mathbf{S} 的子波数为 q , 根据条件随机生成随机序列 $\mathbf{A} = \{A_1, A_2, A_3, \dots, A_q\}$ 和周期序列 $\mathbf{T} = \{T_1, T_2, T_3, \dots, T_q\}$, 由 \mathbf{A} 和 \mathbf{T} 组成振动向量序列 $\mathbf{O} = \{(A_1, T_1), (A_2, T_2), \dots, (A_q, T_q)\}$, 然后由 \mathbf{O} 中的各元素生成余弦子波序列, 将序列中的各个子波收尾相连构成仿真时间序列 \mathbf{S} .

构造 S1, S2, S3, S4, S5 五种波形进行实验, 设波形的采样频率为 1000 Hz, 五种波形的子波振幅有五种类型: 300, 375, 450, 525, 600 μV . 周期也有

五种类型: 300, 375, 450, 525, 600 ms. 在振幅、周期组成的振动向量序列中, S1 波形的振幅-周期组合模式有 5 种, S2 有 10 种, S3 有 15 种, S4 有 20 种, S5 有 25 种. 波形的子波数均为 1200, 在每种波形中, 各种子波模式平均分布. 构造每种仿真波形各 40 个进行实验.

由此可见, 五种仿真波形的振幅和周期在一维空间中均没有明显差别, 振幅-周期组和模式在二维振动空间中的数量关系为 $S1 < S2 < S3 < S4 < S5$. 分别用向量集合表示五种波形的振幅-周期组合模式, 五种波形的振幅-周期模式如表 1 所列.

表 1 仿真波形的振幅-周期模式
Table 1. Amplitude-period modes of the simulation waveforms.

波形	振幅-周期模式 (振幅单位 μV , 周期单位: ms)
S1	(300, 300), (375, 375), (450, 450), (525, 525), (600, 600)
S2	(300, 300), (375, 375), (450, 450), (525, 525), (600, 600), (375, 300), (450, 375), (525, 450), (600, 525), (300, 600).
S3	(300, 300), (375, 375), (450, 450), (525, 525), (600, 600), (375, 300), (450, 375), (525, 450), (600, 525), (300, 600), (450, 300), (525, 375), (600, 450), (300, 525), (375, 600).
S4	(300, 300), (375, 375), (450, 450), (525, 525), (600, 600), (375, 300), (450, 375), (525, 450), (600, 525), (300, 600), (450, 300), (525, 375), (600, 450), (300, 525), (375, 600), (525, 300), (600, 375), (300, 450), (375, 525), (450, 600).
S5	(300, 300), (375, 375), (450, 450), (525, 525), (600, 600), (375, 300), (450, 375), (525, 450), (600, 525), (300, 600), (450, 300), (525, 375), (600, 450), (300, 525), (375, 600), (525, 300), (600, 375), (300, 450), (375, 525), (450, 600), (600, 300), (300, 375), (375, 450), (450, 525), (525, 600).

由表 1 可知, 在上述五种波形中, S1 元素间的最小 Jaccard 距离约为 0.234, S2, S3, S4, S5 元素间的最小 Jaccard 距离为 0.125, 由此可知, 当相似容限取值为 $0 < R < 0.125$ 时, 可以对波形中的不同模式进行有效区分, 当 $R > 0.125$ 时, 相似容限 R 对波形中的不同模式的区分可能出现失真.

3.2 仿真波形的二维样本熵分析

为比较二维样本熵与振幅、周期单一维度样本熵的区别, 分别计算五种波形的振幅样本熵 (sample entropy of amplitude, $SEnA$), 周期样本熵 (sample entropy of cycle, $SEnC$), 和二维样本熵 (two-dimensional sample entropy, $D2SEn$), 其中 $SEnA$ 为一维振幅序列的样本熵, $SEnC$ 为一维周期序列的样本熵, 计算参数为 $m = 2, r = 0.25$ 倍标准差. 分别对每种波形的 $D2SEn, SEnA, SEnC$

指标进行方差分析.

为研究 $D2SEn$ 描述波形复杂度的有效性, 设 $m = 2$, 计算相似容限 $R = 0.05-0.95$ (步长为 0.5) 条件下五种波形的 $D2SEn$ 值. 因为当 $R = 0.8-0.95$ 时, 计算 $D2SEn$ 时 $B^m(R)$ 和 $B^{m+1}(R)$ 的值相等, 导致 $D2SEn$ 的值为 0, 所以不对 $R = 0.8-0.95$ 时的 $D2SEn$ 值进行分析.

方差分析结果表明: 只有在 $R = 0.05, 0.1, 0.15$ 条件下, 波形的 $D2SEn$ 值才满足 $S1 < S2 < S3 < S4 < S5$ 的关系, 在其他条件下, 五种波形 $D2SEn$ 值的大小关系均不符合预设条件. 由此可知, $D2SEn$ 的值与相似容限 R 的取值有关, 仅当 R 值满足一定条件时, $D2SEn$ 才能有效反映波形的复杂度特征. 在每种 R 值条件下, 五种波形的 $D2SEn$ 值如图 2 所示.

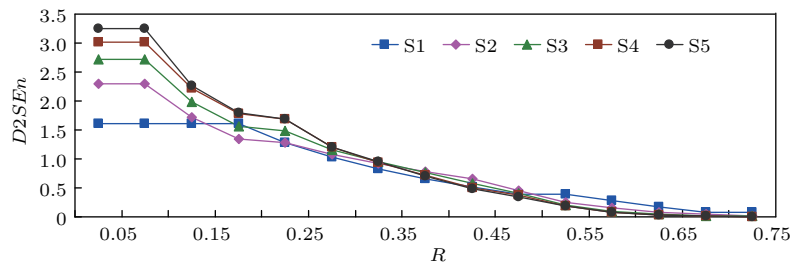


图 2 相似容限 R 取不同值时仿真波形的 $D2SEn$ 值

Fig. 2. $D2SEn$ values of simulation waveforms for different similar tolerance values.

在 $R = 0.05, 0.1, 0.15$ 条件下, 五组波形 $D2SEn$ 值的方差分析结果表明: 在上述三种条件下, $D2SEn$ 值在五种波形间的组效应显著 ($R = 0.05: F = 3245.084, P < 0.00001$; $R = 0.1: F = 3245.084, P < 0.00001$; $R = 0.15: F = 3923.475, P < 0.00001$). 组间两两比较时, 每两组波形间的 $D2SEn$ 值均存在显著差异 ($P < 0.00001$). 五组波形 $SEnA$ 和 $SEnC$ 值的方差分析表明: $SEnA$ 和 $SEnC$ 的组间效应均不显著 ($SEnA: F = 0.654, P = 0.625$; $SEnC: F = 1.667, P = 0.159$).

综上所述, $D2SEn$ 值可以有效区分不同复杂度波形间的差异. 当波形振动序列的复杂度特性体现为振幅-周期交互效应的时候, 振幅和周期样本熵均无法对不同复杂度的波形进行区分, $D2SEn$ 可以有效弥补 $SEnA$ 和 $SEnC$ 在描述波形复杂度时的不足.

3.3 选取 $D2SEn$ 相似容限 R 的方法

仿真结果表明, 两组被试间 $D2SEn$ 的差异与相似容限 R 的取值有关. 一般认为, 在一维条件下, 相似容限取 0.1 到 0.25 倍标准差时可以得到比较稳定的样本熵值, 与标准差相联系也可以有效提升相似容限的数据一致性. 为使 $D2SEn$ 相似容限的取值具有一般性, 我们根据一维元素和二维元素间分布概率的关系, 将一维条件下的相似容限方法延伸到二维条件, 提出以下方法对 $D2SEn$ 的 R 值进行估算.

首先根据脑电信号的特点提出以下两点假设:

- 1) 二维序列元素在两个维度的取值是相互独立的;
- 2) 二维序列元素在两个维度上符合二维正态分布.

因此设二维随机变量为 (G, H) , G 和 H 在二维空间中的分布概率密度为

$$f(g, h) = \frac{1}{2\pi\sigma_1\sigma_2} \exp \left[-\frac{1}{2} \left(\frac{(g - \mu_1)^2}{\sigma_1^2} + \frac{(h - \mu_2)^2}{\sigma_2^2} \right) \right]. \quad (13)$$

G 和 H 的边缘分布概率为

$$f_G(g) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp \left[-\frac{(g - \mu_1)^2}{2\sigma_1^2} \right], \quad (14)$$

$$f_H(h) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp \left[-\frac{(h - \mu_2)^2}{2\sigma_2^2} \right], \quad (15)$$

其中, μ_1, μ_2 分别为 G 和 H 的期望; σ_1, σ_2 分别为 G 和 H 的方差. 因为 G, H 相互独立, 所以 $f(g, h) = f_G(g) \cdot f_H(h)$.

计算 G 或 H 序列的一维样本熵时, 一般取序列元素 0.1 — 0.25 倍标准差作为相似容限 R 的取值. 以 G 为例, 当 R 值取 0.1 或 0.25 倍标准差时, 不同元素相似的最大概率分别为: $P_1(X) = F(\mu_1 + 0.1\sigma_1) - F(\mu_1 - 0.1\sigma_1) \approx 0.08$, $P_2(X) = F(\mu_1 + 0.25\sigma_1) - F(\mu_1 - 0.25\sigma_1) \approx 0.20$.

假设在一维序列和二维序列中, 不同元素间的最大相似概率是相等的. 设 G 和 H 的取值范围为: $g = [\mu_1 - w\sigma_1, \mu_1 + w\sigma_1], h = [\mu_2 - w\sigma_2, \mu_2 + w\sigma_2]$.

当 $P_1(G, H) = [F(\mu_1 + w\sigma_1) - F(\mu_1 - w\sigma_1)] \cdot [F(\mu_2 + w\sigma_2) - F(\mu_2 - w\sigma_2)] = 0.08$, 解得 $w_1 = 0.36$;

当 $P_1(G, H) = [F(\mu_1 + w\sigma_1) - F(\mu_1 - w\sigma_1)] \cdot [F(\mu_2 + w\sigma_2) - F(\mu_2 - w\sigma_2)] = 0.2$, 解得 $w_2 = 0.60$.

因此选取 $\{S_1(g, h) | g = [\mu_1 - 0.36\sigma_1, \mu_1 + 0.36\sigma_1], h = [\mu_2 - 0.36\sigma_2, \mu_2 + 0.36\sigma_2]\}$ 范围内的一般相似容限值作为相似容限的下限, 选取 $\{S_2(g, h) | g = [\mu_1 - 0.6\sigma_1, \mu_1 + 0.6\sigma_1], h = [\mu_2 - 0.6\sigma_2, \mu_2 + 0.6\sigma_2]\}$ 范围内的一般相似容限值作为相似容限的上限. 在 S_1 和 S_2 区域内, 一般相似容限的计算公式为

$$R_w = (\mu_1 - w\sigma_1) \cdot (\mu_2 - w\sigma_2) \times [(\mu_1 - w\sigma_1) \cdot (\mu_2 + w\sigma_2) + (\mu_1 + w\sigma_1) \cdot (\mu_2 - w\sigma_2) - (\mu_1 - w\sigma_1) \cdot (\mu_2 - w\sigma_2)]^{-1}, \quad (16)$$

其中, 对于 S_1 区域, $w = 0.36$, 对于 S_2 区域, $w = 0.6$.

4 基于二维样本熵的抑郁症脑电信号分析

4.1 实验数据

共有 24 名被试参加实验, 其中抑郁症组被试 12 人 (5 男, 年龄: 44.17 ± 14.43), 正常对照组被试 12 人 (5 男, 年龄: 44.17 ± 13.53), 共同入选标准为: 右利手, 年龄在 18—60 岁之间, 无酒精或药物滥用史. 抑郁症组被试为符合精神疾病诊断和统计手册第四版 (DSM-IV) 标准的当前单相抑郁症患

者, 无其他神经或精神系统疾病史; 对照组被试均无神经或精神疾病家族史. 两组被试的性别和年龄相互匹配. 每名被试在实验前完成抑郁自评量表(BDI), 两组被试的BDI分数差异显著(抑郁组: 22.42 ± 10.34 ; 对照组: 4 ± 2.26 ; $t = 5.563$, $P < 0.001$).

采集被试在闭眼静息状态下 2 min 的脑电信号. 采集设备为 Brain Products 公司生产的 64 通道 10—20 系统脑电采集系统, 在线参考电极为 FCz, 离线变为双侧乳突 (TP9, TP10) 平均为参考, 采样频率为 1000 Hz, 带通滤波频率为 0.01—30 Hz, 所有电极的电阻降至 5 kΩ 以下.

4.2 分析方法

用 Butterworth 滤波器(斜率为 24 db/Oct)对脑电数据进行滤波, 将数据分解为 Theta(4—7 Hz), Alpha(8—13 Hz), Beta(14—30 Hz) 三个频段的信号. 提取每个频段的振幅-周期振动序列, 用箱线图法对数据进行筛选, 标记振幅序列和周期序列中内限之外的值为异常值, 剔除振动序列中包含异常值的元素. 根据 Theta, Alpha, Beta 频段振动序列长度的最小值, 将每个频段的振动序列截至相同长度, 其中 Theta 频段振动序列的长度为 1500, Alpha 频段的长度为 2000, Beta 频段的长度为 3800. 分别对 Theta, Alpha, Beta 频段振动序列的二维样本熵进行分析. 数据分析时所用的软件为 BrainVision Analyzer 2.0 和 MATLAB 2010 b.

将电极按位置分为左侧额区 (Fp1, AF3, AF7, F1, F3, F5, F7), 右侧额区 (Fp2, AF4, AF8, F2, F4, F6, F8), 左侧中央区 (FC1, FC3, FC5, C1, C3, C5), 右侧中央区 (FC2, FC4, FC6, C2, C4, C6),

左侧顶区 (CP1, CP3, CP5, P1, P3, P5, P7), 右侧顶区 (CP2, CP4, CP6, P2, P4, P6, P8), 左侧颞区 (FT7, FT9, T7, TP7), 右侧颞区 (FT8, FT10, T8, TP8), 左侧枕区 (O1, PO3, PO7), 右侧枕区 (O2, PO4, PO8) 10 个脑区. 分别对每个脑区处的 $D2SEn$ 平均值进行分析.

4.3 实验结果

4.3.1 二维样本熵与相似容限 R 的关系

分别设 $R = 0.1—0.9$, 计算每种 R 值条件下两组被试的 $D2SEn$ 值. 因为在 $R = 0.1$ 条件下, 部分电极在计算 $D2SEn$ 过程中出现了 $B^{m+1}(r) = 0$ 的情况, 因此不对 $R = 0.1$ 时的 $D2SEn$ 值进行分析. 在 $R = 0.2—0.9$ 条件下, 分别对两组被试在各频段各脑区处的 $D2SEn$ 值进行配对样本 t 检验. 所有数据通过 SPSS20 软件进行分析.

Theta 频段分析结果表明, 在 $R = 0.2—0.9$ 条件下, 抑郁组在各脑区的 $D2SEn$ 值均小于对照组. 当 $R = 0.6—0.9$ 时, 两组被试在左侧枕区 $D2SEn$ 组间差异显著 ($P < 0.05$), t 值在 $R = 0.9$ 时达到最小, 表明当 $R = 0.9$ 时两组被试间的 $D2SEn$ 值差异最大. Theta 频段各脑区处 $D2SEn$ 的 t 检验结果如图 3 所示.

Alpha 频段的分析结果表明, 在 $R = 0.2—0.9$ 条件下, 抑郁组在各脑区的 $D2SEn$ 值均小于对照组. 当 $R = 0.7—0.9$ 时, 两组被试在双侧顶区和双侧枕区的 $D2SEn$ 值差异显著 ($P < 0.05$), 其中两组被试在右侧顶区和右侧枕区的 t 值在 $R = 0.9$ 时达到最小, 在左侧顶区和左侧枕区的 t 值在 $R = 0.8$ 时达到最小, Alpha 频段各脑区处 $D2SEn$ 值的 t 检验结果如图 4 所示.

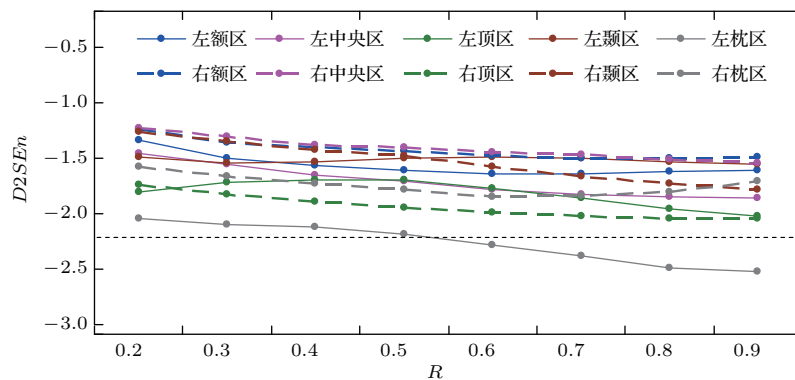


图3 (网刊彩色) Theta 频段各脑区 $D2SEn$ 值的 t 检验结果

Fig. 3. (color online) T -test results of $D2SEn$ values of each brain region in Theta band.

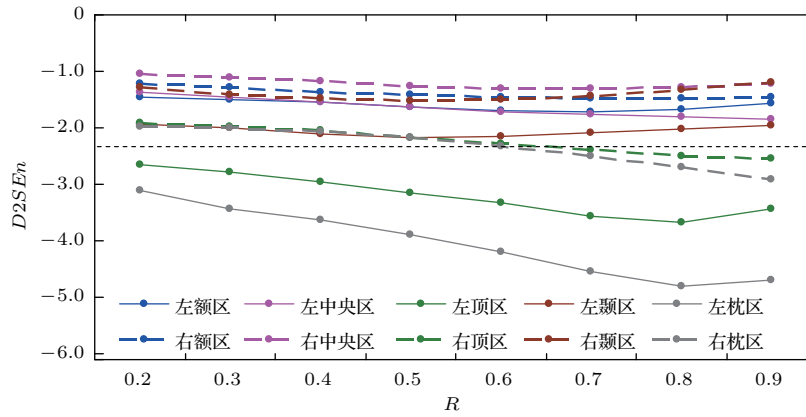


图4 (网刊彩色) Alpha频段各脑区 $D2SEn$ 值的 t 检验结果

Fig. 4. (color online) T -test results of $D2SEn$ values of each brain region in Alpha band.

在 Beta 频段, 两组被试的 $D2SEn$ 值在 $R = 0.2-0.9$ 各条件下均没有表现出显著差异 ($P > 0.05$).

4.3.2 差异显著位置 $D2SEn$ 的统计分析结果

根据 3.3 提出的方法对各通道信号在 S_1 和 S_2 区域内的相似容限进行分析, 结果如表 2 所列.

由表 2 可知, 对于 Theta, Alpha, Beta 三个频段, S_1 和 S_2 区域间 R 的取值均在 0.493 到 0.593 之间, 为简化计算, 本研究选取 $R = 0.5$ 作为二维样本熵的计算参数. 设 $R = 0.5$, 对两组被试间的 $D2SEn$ 值进行分析, 结果表明, 在 Alpha 频段, 抑郁症组在左侧顶区和左侧枕区的 $D2SEn$ 显著低于对照组, 在 Theta 和 Beta 频段, 两组被试在各脑区的 $D2SEn$ 均无显著差异. 在 Alpha 频段, 两组被

试在各脑区处的 $D2SEn$ 值统计如表 3 所列.

表 2 所有通道在 S_1 和 S_2 区域内的最大相似容限

Table 2. Maximum similar tolerance in S_1 and S_2 areas in all channels.

	Theta	Alpha	Beta
S_1 区域	0.427 ± 0.058	0.453 ± 0.057	0.493 ± 0.032
S_2 区域	0.593 ± 0.066	0.621 ± 0.065	0.665 ± 0.034

Alpha 频段分析表明, 在 $R = 0.2-0.9$ 各种条件下, 两组被试在左侧顶区和左侧枕区的 $D2SEn$ 值均满足抑郁症组显著小于对照组, 表明在对两组波形的复杂度进行分析时, Alpha 频段左侧顶区和左侧枕区的 $D2SEn$ 值具有一定的稳定性, 本文提出的选择相似容限的方法具有现实的有效性.

表 3 当 $R = 0.5$ 时, 在 Alpha 频段各脑区处的 $D2SEn$ 值

Table 3. $D2SEn$ value of each brain region in Alpha band when $R = 0.5$.

	抑郁症组 ($m \pm sd$)	对照组 ($m \pm sd$)	t	Sig(双侧)
左额区	0.502 ± 0.251	0.711 ± 0.288	-1634	0.130
右额区	0.521 ± 0.259	0.721 ± 0.323	-1.415	0.185
左中央区	0.541 ± 0.209	0.717 ± 0.261	-1.641	0.129
右中央区	0.562 ± 0.209	0.709 ± 0.289	-1.260	0.234
左顶区	0.543 ± 0.116	0.699 ± 0.221	-3.146	0.009
右顶区	0.492 ± 0.142	0.653 ± 0.239	-2.168	0.053
左颞区	0.547 ± 0.173	0.756 ± 0.249	-2.175	0.052
右颞区	0.586 ± 0.195	0.748 ± 0.277	-1.514	0.158
左枕区	0.436 ± 0.145	0.663 ± 0.254	-3.894	0.003
右枕区	0.392 ± 0.193	0.592 ± 0.279	-2.168	0.053

除此之外, 4.3.1 节结果表明, 在 Theta 频段, 当 $R = 0.6-0.9$ 时, 两组被试左侧枕区的 $D2SEn$ 值差异显著, 当 $R = 0.7-0.9$ 时, 两组被试在右侧顶区和右侧枕区的 $D2SEn$ 值差异显著, 表明上述脑区和频段也可能与抑郁症的发生和机理有关, 但因为具有对相似容限过于依赖的特性, 这些指标对抑郁症的标记作用还有待进一步研究验证.

上述结果表明, 抑郁症组和对照组间 $D2SEn$ 值的差异与相似容限 R 的取值有关. 当 R 值满足一定条件时, Alpha 频段的 $D2SEn$ 值可以有效地对两组被试的脑电波形进行区分. 抑郁症组的 $D2SEn$ 值在 Alpha 频段表现为在左侧顶区和左侧枕区显著小于对照组的特征, 表明相对于正常人, 抑郁症患者在上述脑区活动中新模式的生成概率降低. 因为抑郁症患者被认为存在明显的认知和情感缺陷^[20,21], 抑郁症脑电的这一特征可能与患者在情感和认知进程中的功能缺陷有关, 但这一推论还需要进一步的研究进行验证.

5 结 论

本文提出一种在振幅-周期二维空间描述波形复杂度的样本熵方法, 并通过数据仿真验证这种方法描述波形复杂度时的有效性. 当被描述波形的复杂度表现为振幅-周期交互作用时, 二维样本熵比单一维度的样本熵更加有效. 基于二维样本熵方法, 对抑郁症患者和正常人的脑电进行了分析, 结果表明, 在一定的相似容限条件下, 二维样本熵可以有效反映抑郁症患者和正常人间脑电差异. 在 Alpha 频段, 抑郁症患者在部分脑区的二维样本熵显著低于正常人, 表明与正常人相比, 抑郁症患者脑部活动中新模式的生成概率更低, 这一特征可能成为抑郁症的潜在生物标记.

参考文献

- [1] Pincus S M 1991 *Proc. Natl. Acad. Sci. USA* **88** 2297
- [2] Richman J S, Moorman J R 2000 *Am. J. Physiol. - Heart C.* **278** 2039
- [3] Bruce E N, Bruce M C, Vennelaganti S 2009 *J. Clin. Neurophysiol.* **26** 257
- [4] Zhu L, Deng J, Wu J H, Zhou N R 2015 *Acta Phys. Sin.* **64** 184302 (in Chinese) [朱莉, 邓娟, 吴建华, 周南润 2015 物理学报 **64** 184302]
- [5] Lei M, Meng G, Zhang W M, Sarkar N 2016 *Acta Phys. Sin.* **65** 108701 (in Chinese) [雷敏, 孟光, 张文明, Nilanjan Sarkar 2016 物理学报 **65** 108701]
- [6] Alcaraz R, Rieta J J 2010 *Biomed. Signal. Proces.* **5** 1
- [7] Kim D J, Jeong J, Chae J H, Park S, Kim S Y, Go H J, Paik I H, Kim K S, Choi B 2000 *Psychiat. Res-Neuroim.* **98** 177
- [8] Lee Y J, Zhu Y S, Xu Y H, Shen M F, Zhang H X, Thakor N V 2001 *Clin. Neurophysiol.* **112** 1288
- [9] Li Y J, Tong S B, Liu D, Gai Y, Wang X Y, Wang J J, Qiu Y H, Zhu Y S 2008 *Clin. Neurophysiol.* **119** 1232
- [10] Ahmadlou M, Adeli H, Adeli A 2012 *Int. J. Psychophysiol.* **85** 206
- [11] Bachmann M, Lass J, Suhhova A, Hinrikus H 2013 *Comput. Math. Method. M.* **2013** 251638
- [12] Zavala-Yoe R, Ramirez-Mendoza R, Cordero L M 2015 *Springerplus* **4** 437
- [13] Abasolo D, Hornero R, Espino P, Alvarez D, Poza J 2006 *Physiol. Meas.* **27** 241
- [14] Takahashi T, Cho R Y, Mizuno T, Kikuchi M, Murata T, Takahashi K, Wada Y 2010 *Neuroimage* **51** 173
- [15] Okazaki R, Takahashi T, Ueno K, Takahashi K, Higashima M, Wada Y 2013 *J. Affect. Disorders* **150** 389
- [16] Li X L, Li D, Liang Z H, Voss L J, Sleigh J W 2008 *Clin. Neurophysiol.* **119** 2465
- [17] Zhang T, Chen W Z, Li M Y 2016 *Acta Phys. Sin.* **65** 038703 (in Chinese) [张涛, 陈万忠, 李明阳 2016 物理学报 **65** 038703]
- [18] Lin P F, Tsao J, Lo M T, Lin C, Chang Y C 2015 *Entropy* **17** 560
- [19] Ahmed M U, Mandic D P 2011 *Phys. Rev. E* **84** 061918
- [20] Orzechowska A, Filip M, Galecki P 2015 *Med. Sci. Monitor* **21** 3643
- [21] Pampallona S, Bollini P, Tibaldi G, Kupelnick B, Munizza C 2004 *Arch. Gen. Psychiat.* **61** 714

Sample entropy analysis of electroencephalogram based on the two-dimensional feature of amplitude and period*

Guo Jia-Liang¹⁾²⁾³⁾⁴⁾ Zhong Ning¹⁾²⁾³⁾⁴⁾⁵⁾ Ma Xiao-Meng¹⁾²⁾³⁾⁴⁾Zhang Ming-Hui¹⁾²⁾³⁾⁴⁾ Zhou Hai-Yan^{1)2)3)4)†}

1) (Beijing Advanced Innovation Center for Future Internet Technology, Beijing University of Technology, Beijing 100124, China)

2) (International WIC Institute, Beijing University of Technology, Beijing 100124, China)

3) (Beijing Key Laboratory of Magnetic Resonance Imaging and Brain Informatics, Beijing 100124, China)

4) (Beijing International Collaboration Base on Brain Informatics and Wisdom Services, Beijing 100124, China)

5) (Dept. of Life Science and Informatics, Maebashi Institute of Technology, Maebashi 371-0816, Japan)

(Received 10 May 2016; revised manuscript received 10 July 2016)

Abstract

Sample entropy, a complexity measure that quantifies the new pattern generation rate of time series, has been widely applied to physiological signal analysis. It can effectively reflect the pattern complexity of one-dimensional sequences, such as the information contained in amplitude or period features. However, the traditional method usually ignores the interaction between amplitude and period in time series, such as electroencephalogram (EEG) signals. To address this issue, in this study, we propose a new method to describe the pattern complexity of waveform in a two-dimensional space. In this method, the local peaks of the signals are first extracted, and the variation range and the duration time between the adjacent peaks are calculated as the instantaneous amplitude and period. Then the amplitude and period sequences are combined into a two-dimensional sequence to calculate the sample entropy based on the amplitude and period information. In addition, in order to avoid the influence of the different units in the two dimensions, we use the Jaccard distance to measure the similarity of the amplitude-period bi-vectors in the waveforms, which is different from the one-dimensional method. The Jaccard distance is defined as the ratio of the different area to the combined area of two rectangles containing the amplitude-period bi-vectors in the Cartesian coordinate system. To verify the effectiveness of the method, we construct five sets of simulative waveforms in which the numbers of patterns are completely equal in one-dimensional space of amplitude or period but the numbers in two-dimensional space are significantly different ($P < 0.00001$). Simulation results show that the two-dimensional sample entropy could effectively reflect the different complexities of the five signals ($P < 0.00001$), while the sample entropy in one-dimensional space of amplitude or period cannot do. The results indicate that compared with the one-dimensional sample entropy, the two-dimensional sample entropy is very effective to describe and distinguish the complexity of interactive patterns based on amplitude and period features in waveforms. The entropy is also used to analyze the resting state EEG signals between well-matched depression patient and healthy control groups. Signals in three separated frequency bands (Theta, Alpha, Beta) and ten brain regions (bilateral: frontal, central, parietal, temporal, occipital) are analyzed. Experimental results show that in the Alpha band and in the left parietal and occipital regions, the two-dimensional sample entropy in depression is significantly lower than that in the healthy group ($P < 0.01$), indicating the disability of depression patients in generation of various EEG patterns. These features might become potential biomarkers of depressions.

Keywords: sample entropy, two-dimension, electroencephalogram, major depressive disorder

PACS: 05.10.-a, 89.70.Cf, 87.19.le

DOI: 10.7498/aps.65.190501

* Project supported by the National Basic Research Program of China (Grant No. 2014CB744600), the International Science and Technology Cooperation Program of China (Grant No. 2013DFA32180), and the National Natural Science Foundation of China (Grant Nos. 61420106005, 61272345).

† Corresponding author. E-mail: zhouhaiyan@bjut.edu.cn