

基于改进的符号相对熵的脑电信号时间不可逆性研究*

张梅 王俊[†]

(南京邮电大学, 图像处理与图像通信江苏省重点实验室, 南京 210003)

(2012年7月25日收到; 2012年9月1日收到修改稿)

提出了一种新的使用过程的前向概率和后向概率计算符号相对熵, 并利用符号相对熵来估计熵产的方法. 该方法是基于熵增和过程不可逆特性关系的, 同时证明脑电信号具有时间不可逆特性, 而且该不可逆特性可以提供脑电信号的熵增信息. 最后应用该方法对青老年脑电信号进行数值计算及对比, 结果是老年人的平均能量损耗显著高于年轻人, 证明符号相对熵可以作为一个物理过程不可逆程度的度量参数, 这对脑电信号是否处于积极或平衡状态的诊断治疗具有积极的作用.

关键词: 符号相对熵, 平均能量耗散, 脑电信号, 时间不可逆性

PACS: 87.85-d, 05.45.-a

DOI: 10.7498/aps.62.038701

1 引言

生命过程的不可逆性和熵产之间的关系是统计力学和物理学的核心^[1]. 以物理学角度来理解生命, 生命就是正向熵流和负向熵流的同时作用. 自然界的一切自发进行的过程都是朝熵增的方向进行的. 从统计力学角度来讲, 一个生理机理产生的静态时间序列的具有时间逆不对称性, 这种不对称性可以提供该生理机理的熵产信息^[2]. 在熵增原理的前提下, 物理学和统计力学紧密联系起来, 实现了利用统计力学来量化研究生命现象.

熵的本质是变化的方向性和时间的方向性, 熵指明了不可逆过程的方向性. 熵值较小对应于比较有序的状态, 熵值较大对应于比较无序的状态. 可逆过程的熵值为零, 现实中是不存在可逆过程的. 过程的不可逆程度越大, 熵产越大, 即能量耗散越大^[3].

本文在熵增原理的基础上, 引入一种分析脑电信号时间不可逆特性的新方法——符号相对熵^[4]. 相对熵量化了熵增和不可逆性之间的关系, 将物理

学和统计力学联系起来^[5-7]. 提出了一种新的使用过程的前向概率和后向概率计算符号相对熵的方法, 利用符号相对熵来估计能量耗散. 并运用该方法分别对老年及青年脑电信号^[8-13]进行处理, 最后针对实验结果进行了分析讨论.

2 基本原理

时间序列的不可逆性和熵产之间的关系是物理学和统计力学的核心, 两个随机过程的概率分布 $p(x)$ 和 $q(x)$ 间的相对熵定义为

$$D(p||q) = \sum_x p(x) \log \frac{p(x)}{q(x)}. \quad (1)$$

相对熵是非负的, 当且仅当 $p(x)$ 和 $q(x)$ 相同时才为零. 但是 $D(q||p) \neq D(p||q)$, 即相对熵不具有对称性.

熵产可用来衡量这两个随机过程间的可区别性. 不同情况下的熵产和相对熵间的关系已经推导处理, 也经过验证.

* 国家自然科学基金(批准号: 61271082, 61201029, 61102094)和江苏省自然科学基金(批准号: BK2011759, BK2011565)资助的课题.

[†] 通讯作者. E-mail: wangj@njupt.edu.cn

对于非平衡态, 每个时间单位的熵产为

$$\frac{\langle \dot{s} \rangle}{k} = \lim_{t \rightarrow \infty} \frac{1}{t} D[p(\{x(\tau)\}_{\tau=0}^t) \| p(\{x(t-\tau)\}_{\tau=0}^t)], \quad (2)$$

其中 k 是波尔兹曼常数, $p(\{x(\tau)\}_{\tau=0}^t)$ 是一个给定轨迹 $\{x(\tau)\}_{\tau=0}^t$ 的概率分布. 对于静态轨迹, $\bar{x}(\tau)$ 唯一确定熵产 s .

对于离散静态曲线 x_1, \dots, x_n , 本文提出了一种新的符号相对熵的计算方法, 即定义长度为 n 的序列的符号相对熵为

$$D_n(p_F \| p_B) \equiv \sum_{l=1}^n p(x_1, \dots, x_l) \log \frac{p(x_1, \dots, x_l)}{p(x_l, \dots, x_n)}. \quad (3)$$

我们对序列逐点向前向后求其概率分布, 其中 p_F 为序列每一点的向前概率分布, p_B 为序列每点的向后概率分布, $D_n(p_F \| p_B)$ 为长度单位为 n 的符号相对熵, 代表了序列的不可逆程度. 由生理机理驱动产生一个随机过程, 那么该生理活动的平均熵产就定义为描述这个随机过程的时间前向概率分布和后向概率分布间的符号相对熵. 符号相对熵反映了随机过程前向分布和后向分布的差异性. 相对熵越大, 差异性越大, 序列的不可逆性也就越大.

推导可得熵产:

$$\frac{\langle \dot{s} \rangle}{k} \geq d(p_F \| p_B) \equiv \lim_{n \rightarrow \infty} \frac{1}{n} D_n(p_F \| p_B). \quad (4)$$

这个公式揭示出物理机理和时间序列的统计学信息有密切联系. 公式左侧是一个纯物理量, 而右侧是仅与序列 x_1, \dots, x_n 有关的纯统计量. 这种联系是 Landauer's 定律的泛化, Landauer's 定律将熵产和逻辑不可逆性联系起来. 根据 (4) 式, 通过计算一个时间序列前向分布和后向分布间的相对熵, 来得到产生该序列的生理机理的熵产信息, 据此进而可以判断该生理过程是活跃的还是消极的.

3 基于平均能量耗散的脑电信号分析

熵增原理将熵产和逻辑不可逆性联系起来, (4) 式是对熵增原理的扩展. 根据本文提出的算法, 我们首先对脑电信号进行相应的符号化及编码处理后, 再利用 (3) 式计算脑电序列的相对熵, 然后根据 (4) 式计算出脑电信号的熵产生率, 即平均能量耗散. 据此我们可以判断该脑电信号是积极的还是消极的, 这对于脑类疾病的诊断和资料有着重大意义. 计算过程包括符号化、编码、逐点求序列前向概率及后向概率、计算符号相对熵和熵产生率, 最后

使用 spss 统计分析软件对计算结果进行了假设检验, 验证了该算法的有效性.

3.1 实验数据

我们从 Neuroelectric and Myoelectric Databases 数据库中随机提取十位健康年轻人及十位健康老年人的脑电信号. 其中年轻人数据分别取自数据库中的“yng”文件 (包括 fMT1419, fMT1420, fMT1415, fMT1416, fMT1427, fMT1421, fMT1422, fMT1423, fMT1424, fMT1428), 记为样本 YngEEG; 老年人数据取自数据库中的“eh”文件 (包括 fMT1502, fMT1503, fMT1504, fMT1506, fMT1523, fMT1510, fMT1509, fMT1512, fMT1519, fMT1520), 记为样本 OldEEG.

3.2 实验方法

对受训者样本 YngEEG 和 OldEEG 脑电信号进行相对熵分析, 首先对这两组非线性数据序列进行符号化, 编码, 相空间重构后, 选择数据长度为 2000 分别计算其相对熵, 并进行显著性检测, 然后进行多样本验证.

3.3 实验结果及分析

3.3.1 研究平均能量耗散与数据长度 L 的关系

对受训者样本 YngEEG 和 OldEEG 中的每组脑电信号序列分别取数据长度 $L = 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000$, 计算每个个体的能量耗散, 并分别对青年样本和老年样本进行平均, 得到平均能量耗散与数据长度 L 的关系如图 1 所示.

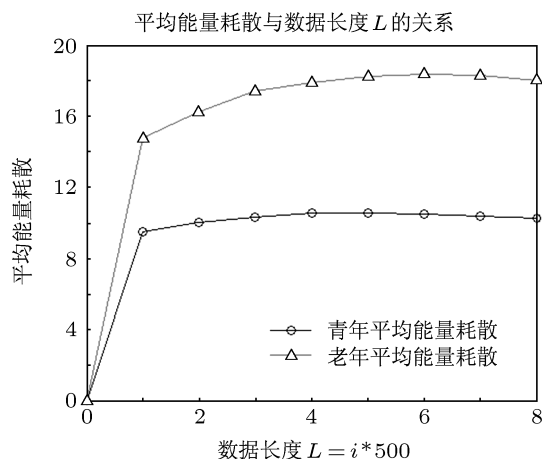


图 1 平均能量耗散与数据长度 L 的关系

分析图 1 青老年能量耗散曲线可知:

1) 在数据长度相同的情况下, 老年人的平均能量耗散要高于青年人;

2) 当数据长度由 0 增至 2000 过程中, 青老年的平均能量耗散值也呈递增趋势, 数据长度大于 2000 时, 平均能量耗散值趋于平稳, 基本不变.

从数据的精准性考虑, 数据长度越大, 其统计概率分布越接近实际分布, 相应的准确性越高. 但从计算量以及计算速度来讲, 数据长度越小, 那么算法速度也越快, 综合我们的实验结果, 取数据长度 $L = 2000$ 时, 即能兼顾处理速度, 又可以保证实验精度.

3.3.2 研究平均能量耗散与编码长度 N 的关系

对受训者样本 YngEEG 和 OldEEG 中的每组脑电信号序列分别取编码长度 $N = 1, 2, \dots, 20$, 计算每个个体的能量耗散, 并分别对青年样本和老年样本进行平均, 得到平均能量耗散与编码长度 N 的关系如图 2 所示.

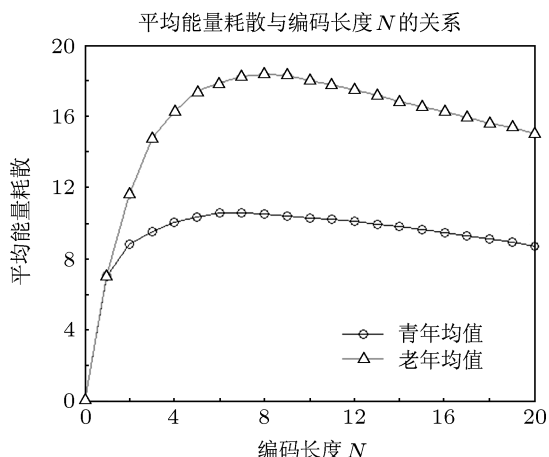


图 2 平均能量耗散与编码长度 N 的关系

分析图 2 平均能量耗散与编码长度 N 的关系可知:

1) 在编码长度 N 从 1 到 8 的递增过程中, 青老年的平均能量耗散都成迅速增大趋势, 从 N 取 7 之后平均能量耗散缓慢减小;

2) 老年人的平均能量耗散明显高于青年人的, 且在 $N = 8$ 时青老年差异最显著.

考虑到实验效果的明显性, n 应该取大些; 考虑到算法的复杂度会影响实验处理速度以及在临床应用上的实时性, n 应该取小一些. 综合考虑上述因素, n 取 8.

3.3.3 研究在数据长度 $L = 2000$, 编码长度 $N = 8$ 条件下青老年能量耗散的差异

对受训者样本 YngEEG 和 OldEEG 中的每组脑电信号序列分别取编码长度 $N = 8$ 及数据长度 $L = 2000$, 计算每个个体的能量耗散, 并分别对青年样本和老年样本进行平均, 得到平均能量耗散与年龄的关系如图 3 所示.

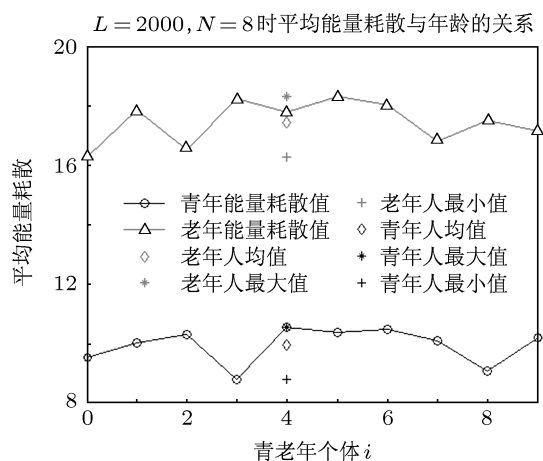


图 3 平均能量耗散与年龄的关系

分析图 3 平均能量耗散与年龄的关系可知: 老年人的能量耗散明显高于青年人, 老年人的能量耗散值分布在 $[8, 10]$, 青年人的能量耗散值分布在 $[16, 19]$.

3.3.4 统计分析与假设检验

为进一步验证本文计算结果的准确性及算法的有效性, 使用 spss 统计分析软件对计算结果进行了假设检验. 主要方法如下:

分别将青老年的平均能量耗散值记作样本 Y 和 E , 分别计算其单样本均值 T 检测, 结果如表 1 所示.

本文关注的是准确地对青老年脑电信号进行判断和辨别, 从表 1 分析结果可知:

1) 青年人平均能量耗散均值为 9.94, 老年人平均能量耗散均值为 17.43;

2) 青年人平均能量耗散在显著性水平 α 取 0.01 时的置信区间为 $[9.31, 10.56]$, 即将 $[9.31, 10.56]$ 作为青年人平均能量耗散判断区间时, 误判概率小于 0.01, 同理将 $[16.68, 18.18]$ 作为老年人的判断区间, 也能保证误判概率小于 0.01. 证明可以很准确地对青老年脑电信号进行判断.

对青老年平均能量耗散值的差异显著性进行假设检验, 使用 spss 对两组样本数据 Y, T 进行独立样本 T 检验, 结果如表 2 所示.

表1 青老年平均能量耗散的单个样本 T 检测

群体类型	检验值 = 0					
	t	df	Sig.(双侧)	均值	差分的 99% 置信区间	
					下限	上限
Y(青年)	51.619	9	0.000	9.940611601700	9.31477241676	10.56645078664
E(老年)	76.166	9	0.000	17.43261396100	16.6888064629	18.1764214591

表2 青老年平均能量耗散值差异(均值之差)显著性的独立样本 T 检验

青老年平均能量耗散值差异		均值方程的 T 检验			
		均值差值	标准误差值	差分的 95% 置信区间	
				下限	上限
V1	假设方差相等	7.492002359300	0.299114261377	6.863586615023	8.120418103577
	假设方差不相等	7.492002359300	0.299114261377	6.862266992728	8.121737725872

根据表2分析青老年平均能量耗散差异的显著性,可以得到以下结论:

1) 在假设青老年平均能量耗散方差相等和不等两种情况下,青老年差异显著性基本相同,均值差值均为 7.49;

2) 青老年平均能量耗散平均值的差值在显著性水平 α 取 0.05 时的置信区间为 [6.86, 8.12], 证明青老年平均能量耗散的差异性非常显著,该算法可以有效地对青老年脑电信号进行区分。

由前文中讨论可知,在一个孤立系统内部自然发生的过程总使系统不可逆地趋于熵取极大值的平衡态。将人的生命视作一个过程时,随着年龄的增长,生命的不可逆程度越大,熵产越大,即能量耗散越大。熵值较小对应于比较有序的状态,即年轻状态,也就是脑电活动的积极状态;熵值较大对应于比较无序的状态,也就是老年人脑电活动的消极状态。当生命体达到无序平衡高度混乱状态,熵增到最大值,生命衰老终结,那就是死亡。这与我们的实验结果也是一致的,证明使用相对熵来计算

过程不可逆程度的算法是可行的,我们可以以平均能量耗散作为特征值来衡量脑电信号是否处于积极状态。

4 结论

1) 基于过程不可逆特性提出了一种新的符号相对熵的计算方法,即使用过程的前向概率和后向概率来计算相对熵,从而获得过程的不可逆程度信息。

2) 根据熵增原理,通过相对熵得到过程的能量耗散,将物理学和统计学紧密联系起来,实现了利用统计学来量化研究生命现象。

3) 基于该算法对青老年的脑电信号进行了分析,结果表明,老年人的能量耗散明显高于青年人,即老年人脑电活动的无序性也就是不可逆性显著高于青年人。该算法的计算结果与实际相符,证明该算法是正确的,表明能量耗散可以作为衡量一个过程是否处于积极状态的参数。

[1] Roldan E, Parrondo J M R 2010 *Phys. Rev. Lett.* **105** 150607
 [2] Costa M, Goldberger A L, Peng C K 2010 *Phys. Rev. Lett.* **95** 198102
 [3] Costa M, Peng C K, Goldberger A L 2008 *Cardiovasc Eng.* **8** 88
 [4] Shen W, Wang J 2011 *Acta Phys. Sin.* **60** 118702 [沈韡, 王俊 2011 物理学报 **60** 118702]
 [5] Porporato A, Rigby J R, Daly E 2007 *Phys. Rev. Lett.* **98** 094101
 [6] Huang J H, Liu N H, Liu J T, Yu T B, He X 2010 *Chin. Phys. B* **19** 110312
 [7] Lu H X, Zhao B 2006 *Chin. Phys. B* **15** 1914
 [8] Meng Q F, Zhou W D, Chen Y H, Peng Y H 2010 *Acta Phys. Sin.* **59** 123 (in Chinese) [孟庆芳, 周卫东, 陈月辉, 彭玉华 2010 物理学报

59 123]
 [9] Bian H R, Wang J, Han C X, Deng B, Wei X L, Che Y Q 2011 *Acta Phys. Sin.* **60** 118701 (in Chinese) [边洪瑞, 王江, 韩春晓, 邓斌, 魏熙乐, 车艳秋 2011 物理学报 **60** 118701]
 [10] Fang X L, Jiang Z L 2007 *Acta Phys. Sin.* **56** 7330 (in Chinese) [方小玲, 姜宗来 2011 物理学报 **56** 7330]
 [11] Ma Q L, Bian C H, Wang J 2010 *Acta Phys. Sin.* **59** 4480 (in Chinese) [马千里, 卞春华, 王俊 2010 物理学报 **59** 4480]
 [12] Xie Y, Xu J X 2002 *Acta Phys. Sin.* **51** 205 (in Chinese) [谢勇, 徐健学 2002 物理学报 **51** 205]
 [13] You R Y, Chen Z, Xu S C, Wu B X 2004 *Acta Phys. Sin.* **53** 2882 (in Chinese) [游荣义, 陈忠, 徐慎初, 吴伯儒 2004 物理学报 **53** 2882]

Modified symbolic relative entropy based electroencephalogram time irreversibility analysis*

Zhang Mei Wang Jun[†]

(Image Processing and Image Communications Key Laboratory, Nanjing University of Posts and Telecomm., Nanjing 210003, China)

(Received 25 July 2012; revised manuscript received 1 September 2012)

Abstract

In this letter, we show that the trajectory of electroencephalogram (EEG) possesses the character of time reversal asymmetry, which can provide information about the entropy production of EEG. We develop a kind of new method to estimate symbolic relative entropy and entropy production using forward and backward trajectories. Finally, we use this method to dispose and analyse the EEGs of younger and elder people, and is turned out that this method works and the average energy dissipation can be used as a parameter to detect nonequilibrium.

Keywords: symbolic relative entropy, average energy dissipation, EEG, time reversal asymmetry

PACS: 87.85-d, 05.45.-a

DOI: 10.7498/aps.62.038701

* Project supported by the National Natural Science Foundation of China (Grant Nos. 61271082, 61201029, 61102094), the Natural Science Foundation of Jiangsu Province, China (Grant Nos. BK2011759, BK2011565).

[†] Corresponding author. E-mail: wangj@njupt.edu.cn