

改进的相对转移熵的癫痫脑电分析*

王莹¹⁾ 侯凤贞²⁾ 戴加飞³⁾ 刘新峰^{3)†} 李锦⁴⁾ 王俊^{1)‡}

1)(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)

2)(中国药科大学理学院, 南京 210009)

3)(南京军区南京总医院神经内科, 南京 210002)

4)(陕西师范大学物理学与信息技术学院, 西安 710062)

(2014年5月29日收到; 2014年6月21日收到修改稿)

脑电信号是由脑神经活动产生并且始终存在于中枢神经系统的自发性电位活动, 是一种重要的生物电信号. 脑电信号是非常微弱的且是非线性的, 脑电信号也具有时间不可逆性. 本文提出了一种新的基于正向序列转移概率与逆向序列转移概率的相对熵方法即相对转移熵方法, 并应用此方法研究了正常脑电与癫痫脑电的不可逆性, 实验结果显示癫痫患者的脑电信号的不可逆性明显小于正常人的脑电信号的不可逆性. 这说明改进的相对转移熵可以作为一个物理过程不可逆程度的度量参数, 这使得应用脑电信号区分病人是否患有癫痫疾病具有积极指导意义.

关键词: 相对转移熵, 脑电信号, 符号化, 时间不可逆性

PACS: 87.85-d, 05.45.-a

DOI: 10.7498/aps.63.218701

1 引言

作为非平衡系统的一个基本属性, 目前不可逆性在生物医学上得到了很高的关注. 时间不可逆性^[1]提供了一个高效的检测生命动力系统的非线性复杂度方法. 所谓不可逆性, 指物质过程能重复但不可能返回过去的性质. 时间序列当且仅当时间反转前后的概率分布相同时才是可逆的. 系统产生的信号的时间不可逆性可以反映出系统在确定性和随机性之间的情况. 在自由运行条件下, 健康的生理系统的时间不可逆性最大, 而疾病的产生或年龄的变化会导致生理系统的时间不可逆性缺失^[2], 代表此时生理系统向随机性减弱的方向发展.

癫痫是由于脑部神经细胞群异常放电引起的过性脑功能紊乱综合症, 是一种严重且较顽固的慢

性疑难性脑部疾病, 其发作时表现为大脑神经元群兴奋性增高以及过度同步化放电, 并导致短暂性中枢神经系统功能失常, 是神经系统疾病中仅次于脑血管疾病的第二大顽症, 人群患病率为0.5%—1%. 我国约有600万癫痫患者. 每年新发病例近65—70万. 癫痫病人就诊时均在发作以后而且体检大多无异常发现. 因此如何区分患者是否患有癫痫疾病就显得很重要.

目前生物电信号采集和记录技术迅猛发展, 相应的生物电信号的分析技术却稍显滞后. 现代心脑电信号研究最常见的分析方法主要有时域分析^[3,4]、频域分析^[5,6]、神经网络^[7,8]等, 但大多方法尚未广泛运用于实践中. 针对以上问题, 本文提出了一种新的基于正向序列转移概率与逆向序列转移概率的相对转移熵方法. 该方法是基于过程的不可逆性, 通过对脑电信号(electroencephalogram,

* 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: xfliu2@gmail.com

‡ 通讯作者. E-mail: wangj@njupt.edu.cn

EEG) 的不可逆性研究, 期待以此区分健康人群与癫痫患者. 在进一步的分析中, 本文利用军区总医院提供的正常人脑电信号与癫痫患者脑电信号, 计算了他们的相对转移熵, 并对其效果进行了讨论, 发现该方法能很好的区分健康人群与癫痫患者.

2 基本原理

理解大脑不同部分之间的因果关系与连通性是非常重要的, 因为大脑的很大一部分活动都认为是大脑内部产生的, 因此仅仅量化刺激-反应关系是不能完全反应脑动力活动的. 转移熵^[9]是一个衡量两个序列的混乱程度的参数, 由于其基于转移概率且是非对称的, 故混合了定向与动态的信息, 这使得转移熵成为定量分析连通性的有效工具, 并已应用于脑部病灶的定位分析. 比较著名的转移熵定义是 Schreiber 于 2000 年提出的转移熵^[10]

$$T_{J \rightarrow I} = \sum p(i_{n+1}, i_n^{(k)}, j_n^{(l)}) \times \log \frac{p(i_{n+1} | i_n^{(k)}, j_n^{(l)})}{p(i_{n+1} | i_n^{(k)})}, \quad (1)$$

其中, i_n, j_n 分别代表了 I 序列、 J 序列在 n 时刻的状态, $i_n^{(k)}$ 指长度为 k 的字符串 $i_{n-k+1}, \dots, i_n$, 同理, $j_n^{(l)}$ 指长度为 l 的字符串 $j_{n-l+1}, \dots, j_n$.

转移熵是一个量化两个系统相关程度或混沌程度的参数, 它可用于处理有限长度的生理信号数据. 但因上述转移熵对参数之间较高的协调性要求和噪声的敏感性, 限制了在时间序列分析领域的应用.

相对熵^[11]是衡量序列不可逆性的重要方法. 若时间序列的前向概率分布与后向概率分布相同时, 其不可逆性为零, 若时间序列的前后向概率差别越大, 则其不可逆性越大. 相对熵又称为 KL 散度, 运用在生物学里的相对熵定义如下:

$$D_{kl}(p_f || p_b) = \sum_x p_f(x) \cdot \log \frac{p_f(x)}{p_b(x)}, \quad (2)$$

其中, p_f, p_b 指同一个序列的前向与后向的概率分布, 当且仅当 p_f, p_b 的概率分布相同时相对熵才为零, 且相对熵不具有对称性, 即

$$D_{kl}(p_f || p_b) \neq D_{kl}(p_b || p_f).$$

通常采集到的脑电信号背景噪声比较复杂, 有 50 Hz 的工频干扰, 电极与皮肤的接触噪声以及电

极与地之间的共模信号的干扰等等. 基于转移熵与相对熵的优缺点, 本文对二者进行了改进, 提出了一种新的计算方法即相对转移熵

$$rH_{n||m} = \sum_{X,Y} p(x_{t-n}^t, y_{t-m}^t) \times D(P(x_t | x_{t-n}^{t-1}) || P(y_t | y_{t-m}^{t-1})). \quad (3)$$

此公式中, x_{t-n}^t 指长度为 n 的序列 $x_{t-n}, x_{t-n+1}, \dots, x_t$, y_{t-m}^t 指长度为 m 的序列 $y_{t-m}, y_{t-m+1}, \dots, y_t$. X 序列是正向的, Y 是 X 序列的逆向序列.

算法的图形化显示如图 1 所示.

使用此方法所需的已知条件是: 给定原始序列长度 N (比如 $N = 500, N = 1000$), 取定 n, m 的长度 (比如 $n = 2, m = 1$). 首先将原始序列符号化得到符号序列 X (比如: $X = 0033022211320$), 将符号序列 X 逆向得 Y 序列 ($Y = 0231122203300$), 对 X 序列逐点求解前 n 个点预测下一点的概率分布 (即 x_t 点的转移概率), 对 Y 序列逐点求解前 m 个点预测下一个点的概率分布 (即 y_t 点的转移概率), 利用这正向、后向的转移概率分布求解其相对熵, 即代表了序列的不可逆程度.

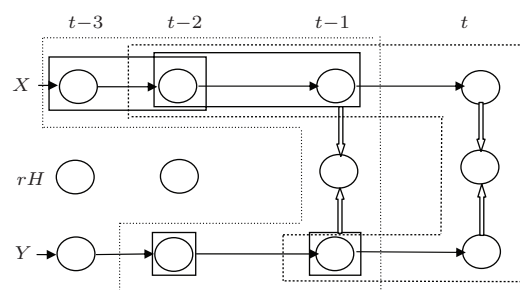


图 1 相对转移熵算法的图形描述

在上述方法介绍中, 我们提到了将原始序列符号化^[1], 这是因为原始脑电数据易受噪声的影响, 对原始脑电数据进行了符号化映射使其成为符号序列可降低噪声影响, 另外虽然在原始序列映射成为符号序列之后, 符号序列会丢失原始序列中一些细节信息, 但是保存了原始序列的动力学^[12]特性. 然后对映射后的序列做进一步的相对转移熵的研究.

原始序列的符号化方法^[13-15]定义如下:

$$s_i(\dot{x}_i) = \begin{cases} 0: & \mu_1 < x_i \leq (1+a)\mu_1 \text{ or } (1+a)\mu_2 \leq x_i < \mu_2, \\ 1: & (1+a)\mu_1 < x_i < \infty \text{ or } -\infty < x_i < (1+a)\mu_2, \\ 2: & (1-a)\mu_1 \leq x_i < \mu_1 \text{ or } \mu_2 < x_i \leq (1-a)\mu_2, \\ 3: & (1-a)\mu_2 < x_i < (1-a)\mu_1, \end{cases} \quad (4)$$

其中, $i = 1, 2, \dots, N$ 代表序列的位置索引值. μ_1 是原始序列中大于等于零的取样信号的平均值, μ_2 是小于零的平均值. a 是一个特殊常数, 通常取值范围为 $a \in [0.03, 0.07]$, 在这个区间范围内可得到相似的结果, 若取 $a < 0.04$ 或 $a > 0.07$ 则得到较不合理的值, 这是因为 a 取值过大或过小, 在原始时间序列转为符号序列时引起细节信息的丢失, 不能较好捕捉信号中的动态信息, 故通常取 $a = 0.05$ ^[16]. 符号化是一个“粗粒化”过程, 这一过程能够捕获大尺度特征, 从而降低动力学噪声和测量噪声的影响.

3 基于改进的相对转移熵的癫痫脑电分析

3.1 实验数据

本文使用了正常健康者脑电信号和癫痫患者脑电信号, 由南京军区总医院从临床诊断中采集得到, 记录长度均大于 1 min, 采样周期为 512 Hz. 该库中记录了多参数脑电数据, 包括 16 个导联信号. 本论文使用的数据分别从正常脑电数据中取十位健康者的脑电数据、从癫痫脑电数据中取十位癫痫患者的脑电数据中取其 Fp_1 导联信号, 经实验验证, 在 16 个导联信号中, 第一导联 Fp_1 信号区分度较好, 所以下面以 Fp_1 导联数据做分析对象. 我们将这两组数据分别记为样本“normal”及样本“epilepsy”.

3.2 实验方法

首先读取各组原始脑电信号中有效的原始数据, 利用 (4) 式进行符号化处理, 然后利用 (3) 式对符号化后的序列计算其相对转移熵, 最后运用 SPSS 统计分析软件对计算结果进行显著性检验, 并进行多样本验证, 确保该算法的有效性.

3.3 实验结果及分析

确定导联信号的选取. 对不同导联数据, 从样本“normal”及样本“epilepsy”中对每个个体的脑电信号序列取固定数据长度为 $N = 4000$ 的脑电信号原始序列. 针对每一个体, 以时间起点为 $t = 5, 405, 805, 1205, 1605, 2005, 2405, 2805, 3205, 3605$ 的十个不同时间点 (从 5 开始是因为本文的有效数据点即从第 5 个数据点开始, 所以参与分析的点必须 ≥ 5), 分别计算十组序列长度为 4000 的相对转移熵值, 并把每一个体中的十组转移熵值取平均作为这个个体的最终的 rH . 在本论文中, 为计算方便, 取 $n = 2, m = 1$, 计算 $rH_{2||1}$ (经实验验证, 只要满足 $m < n$, 正常人脑电信号的相对转移熵与癫痫脑电信号均可以区分, 只是随着取值的增加计算量也会增加, 此处为方便计算取 n 为 2, m 为 1). 最后分别对正常脑电信号与癫痫脑电信号的相对转移熵进行平均, 得到每一个导联信号最终的相对转移熵, 结果如表 1、表 2 所示.

表 1 基于健康人群脑电信号不同导联脑电信号的相对转移熵 $rH_{2||1}$ (4000)

导联	Fp_1	Fp_2	F_3	F_4	C_3	C_4	P_3	P_4
mean	0.2893	0.2825	0.3104	0.3089	0.3130	0.3130	0.3218	0.3239
std	0.0094	0.0100	0.0077	0.0074	0.0087	0.0062	0.0059	0.0046
导联	O_1	O_2	F_7	F_8	T_3	T_4	T_5	T_6
mean	0.3315	0.3326	0.2977	0.3032	0.3142	0.3128	0.3192	0.3173
std	0.0046	0.0052	0.0127	0.0105	0.0116	0.0103	0.0089	0.0053

表2 基于癫痫患者不同导联脑电信号的相对转移熵 $rH_{2||1}(4000)$

导联	Fp_1	Fp_2	F_3	F_4	C_3	C_4	P_3	P_4
mean	0.2262	0.2414	0.2778	0.2749	0.2866	0.2790	0.2942	0.2935
std	0.0117	0.0150	0.0137	0.0110	0.0096	0.0101	0.0107	0.0097

导联	O_1	O_2	F_7	F_8	T_3	T_4	T_5	T_6
mean	0.2980	0.3021	0.2622	0.2511	0.3044	0.2924	0.2977	0.2910
std	0.0114	0.0110	0.0213	0.0130	0.0209	0.0122	0.0173	0.0103

对健康人群脑电信号与癫痫患者不同导联脑电信号的 $rH_{2||1}$ 进行基本研究后, 根据表1、表2的平均值与方差绘制成图, 如图2所示.

由图2可知, 在16个导联信号中, 第一导联 Fp_1 信号区分度较好, 所以取 Fp_1 导联数据做分析对象.

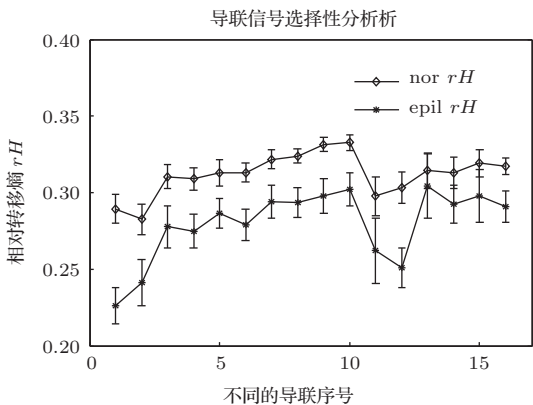


图2 健康人群与癫痫患者不同导联脑电信号的相对转移熵在 $N = 4000$ 时结果比较 (横轴坐标表示的含义是: 1 为 Fp_1 , 2 为 Fp_2 , $\cdots$, 16 为 T_6 , nor rH 为 normal rH , epil rH 为 epilepsy rH)

不同脑电信号的相对转移熵分析. 从样本“normal”及样本“epilepsy”中对每个个体的脑电信号序列分别取数据长度为 $N = 1000, 4000$ 的脑

电信号原始序列. 与上一实验相同, 针对每一个体, 取不同的时间起点, 分别计算十组序列长度为1000, 序列长度为4000的相对转移熵值, 并把每一个体中相同序列长度(比如 $N = 1000$)的十组转移熵值取平均作为这个个体的最终的 rH . 最后分别对正常脑电信号与癫痫脑电信号的相对转移熵进行平均, 得到表3、表4的结果, 以分析是否可以区分这两类.

对健康人群脑电信号与癫痫患者脑电信号的 $rH_{2||1}$ 进行基本研究后, 根据表3、表4的平均值与方差绘制成图, 如图3所示.

由表3可知, 正常脑电的相对转移熵分布在 $[0.2736, 0.2930]$ 范围内, 癫痫脑电的相对转移熵分布在 $[0.2216, 0.2486]$ 范围内, 显然正常脑电与癫痫脑电的 rH 是不一样的. 进一步对正常脑电信号和癫痫脑电信号的 rH 进行研究, 根据表3中的均值和方差画出10组健康人群与癫痫患者脑电信号的相对转移熵的分布范围, 如图3(a)所示. 可以直观的看到正常脑电信号和癫痫脑电信号的 rH 分布明显不同, 正常脑电信号的 rH 大于癫痫脑电信号的 rH . (a)与(b)类似, 且可以发现随着数据量的增加, 两类脑电的 rH 区分更加明显.

表3 基于健康人群脑电信号与癫痫患者脑电信号的相对转移熵 $rH_{2||1}(1000)$

个体	1	2	3	4	5	6	7	8	9	10	mean	std
正常	0.31	0.31	0.27	0.29	0.24	0.25	0.29	0.27	0.31	0.24	0.28	0.01
癫痫	0.31	0.20	0.21	0.21	0.23	0.30	0.25	0.18	0.20	0.22	0.23	0.01

表4 基于健康人群脑电信号与癫痫患者脑电信号的相对转移熵 $rH_{2||1}(4000)$

个体	1	2	3	4	5	6	7	8	9	10	mean	std
正常	0.32	0.32	0.30	0.28	0.23	0.28	0.30	0.27	0.30	0.24	0.28	0.01
癫痫	0.31	0.19	0.20	0.19	0.22	0.26	0.23	0.18	0.22	0.21	0.22	0.01

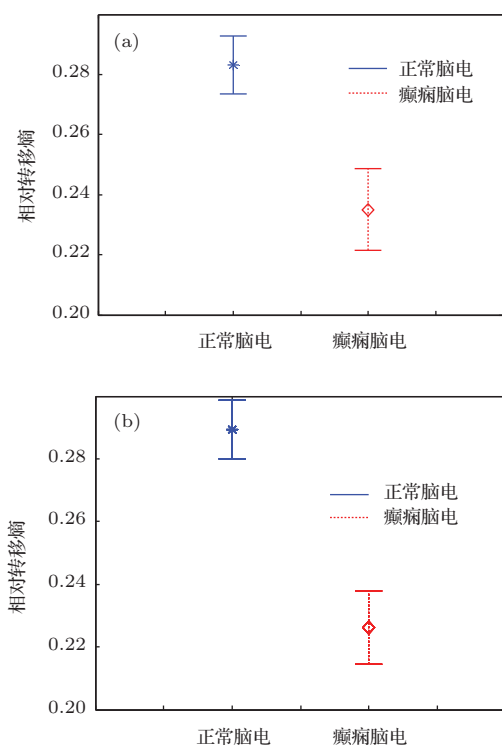


图3 (网刊彩色) 健康人群与癫痫患者脑电信号的相对转移熵在 $N = 1000$ 和 4000 时结果比较 (a) $N = 1000$ 对应表3数据; (b) $N = 4000$ 对应表4数据

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 T 检验. T 值 = 2.905, $P = 0.01 < 0.05$, 达到了1%的极显著水平, 这说明健康人群与癫痫患者脑电信号的相对转移熵的差异性非常显著, 该算法可以有效的对健康人群与癫痫患者脑电信号进行区分.

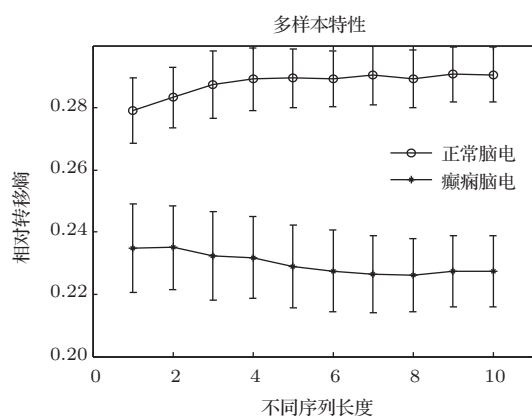


图4 不同数据长度时健康人群与癫痫患者脑电信号的相对转移熵 (横轴坐标表示的含义是: 1为500, 2为1000, ..., 10为5000, 即对500点到5000点的序列长度求其相对转移熵)

实验中, 仅取了两组 ($N = 1000$, $N = 4000$) 数据, 为验证多样本特性. 针对相同的样本对不同数

据长度计算其相对转移熵, 结果如图4所示.

从图4可以看出: 正常脑电信号的相对转移熵较高, 分布在区间 $[0.279, 0.291]$ 上; 癫痫脑电信号的相对转移熵较低, 分布在区间 $[0.226, 0.236]$ 上, 明显具有可分辨性的差异. 另外, 在序列长度很小时, 正常脑电信号与癫痫脑电信号的相对转移熵就有比较大的区别, 且随着序列长度的增加, 区别越来越明显, 当数据长度达到4500时趋于稳定.

由上述的分析中可知, 健康人群的脑电信号与癫痫患者的脑电信号的相对转移熵在平均值差值上存在显著差异, 这说明相对转移熵可以作为区分正常脑电信号与癫痫脑电信号的参数, 这对于癫痫研究判定具有重要的意义.

最后, 验证算法的健壮性. 依据前期实验基础, 从样本“normal”与“epilepsy”中分别取数据长度为 $N = 500i$ ($i = 1, 2, 3, \dots, 9, 10$) 的两类脑电信号原始序列并进行叠加高斯白噪声的计算. 结果如图5所示.

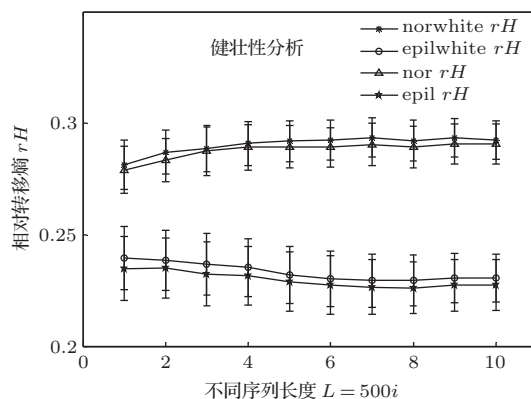


图5 相对转移熵算法健壮性分析 (norwhite rH 为 normal white noise rH , epilwhite rH 为 epilepsy white noise rH , nor rH 为 normal rH , epil rH 为 epilepsy rH)

从图5可以看出, 相对于原始信号的相对转移熵, 加上高斯白噪声后并没有太大变化, 且趋势完全一致, 由此可见本论文所提出的算法健壮性较高, 算法可靠有效.

4 结 论

在原有的转移熵研究上, 针对过程的不可逆性, 本文提出了一种新的相对转移熵计算方法. 使用正向序列与逆向序列的转移概率来计算相对熵, 从而获得了过程的不可逆程度信息. 自转移熵概

念被提出后, 一系列的计算方法被用于转移熵的计算中, 然而大多数的计算方法都对观测数据有较高的要求, 比如要求数据的长度要足够长, 对噪声敏感等等, 这些都限制了转移熵的应用和推广, 而本文所提出的方法仅需少量数据即可达到理想效果. 运用此方法于癫痫患者与健康人群的脑电信号分析, 只需较小的序列长度即可有效的提取出时间不可逆性特征, 区分正常脑电信号与癫痫脑电信号, 并且此方法运用了符号化方法, 受噪声的影响较小, 而加入噪声后影响微乎其微, 可见这一算法的高健壮性.

参考文献

- [1] Daw C S, Finney C E A, Kennel M B 2000 *Phys. Rev. E* **62** 1912
- [2] Costa M, Goldberger A L, Peng C K 2005 *Phys. Rev. Lett.* **95** 198102
- [3] Wang J, Ma Q L 2008 *Chin. Phys. B* **17** 4424
- [4] Wang J, Yu Z F 2012 *Chin. Phys. B* **21** 018702
- [5] Thatcher R W 2012 *Dev. Neuropsychol.* **37** 476
- [6] Petrantonas P C, Hadjileontiadis L J 2012 *IEEE T. on Signal Proc.* **60** 2604
- [7] Musselman M, Djurdjanovic D 2012 *Expert Syst. Appl.* **39** 11413
- [8] Orhan U, Hekim M, Ozer M 2012 *J. Med. Syst.* **36** 2219
- [9] Staniek M, Lehnertz K 2008 *Phys. Rev. Lett.* **100** 158101
- [10] Schreiber T 2000 *Phys. Rev. Lett.* **85** 461
- [11] Porporato A, Rigby J R, Daly E 2007 *Phys. Rev. Lett.* **98** 094101
- [12] Zhang H, Yang M J 1997 *Acta. Biop. Sin.* **13** 340346 (in Chinese) [张辉, 杨明静 1997 生物物理学报 **13** 340346]
- [13] Shen W, Wang J 2011 *Acta Phys. Sin.* **60** 118702 (in Chinese) [沈韡, 王俊 2011 物理学报 **60** 118702]
- [14] Zhang M, Cui C, Ma Q L, Gan Z L, Wang J 2013 *Acta Phys. Sin.* **62** 068704 (in Chinese) [张梅, 崔超, 马千里, 干宗良, 王俊 2013 物理学报 **62** 068704]
- [15] Yao W P, Liu T B, Dai J F, Wang J 2014 *Acta Phys. Sin.* **63** 78704 (in Chinese) [姚文坡, 刘铁兵, 戴加飞, 王俊 2014 物理学报 **63** 78704]
- [16] Wessel N, Ziehmann C, Kurths J, Meyerfeldt U, Schirdewan A, Voss A 2000 *Phys. Rev. E* **61** 733

Analysis on relative transfer of entropy based on improved epileptic EEG*

Wang Ying¹⁾ Hou Feng-Zhen²⁾ Dai Jia-Fei³⁾ Liu Xin-Feng³⁾† Li Jin⁴⁾ Wang Jun¹⁾‡

1 (Image Processing and Image Communications Key Lab., Nanjing Univ. of Posts & Telecomm., Nanjing 210003, China)

2 (School of Science, China Pharmaceutical University, Nanjing 210009, China)

3 (Nanjing General Hospital of Nanjing Military Command, Nanjing, 210002, China)

4 (College of Physics and Information Technology, Shaanxi Normal University - Xi'an 710062, China)

(Received 29 May 2014; revised manuscript received 21 June 2014)

Abstract

EEG (electroencephalogram) is generated by the brain activity and is present in the central nervous system of spontaneous electrical activity, which is an important biological signal. EEG is a very weak and nonlinear as well as irreversible signal. This paper presents a new method to describe it based on the relative entropy of transition probability for the forward and reverse sequences. Besides, we may apply this method to study the normal EEG and epileptic EEG irreversibility, and the experimental results show that the EEG irreversibility of patients who suffer from epilepsy is significantly less than that of normal people. This shows that the relative transfer entropy can be used as a parameter to detect the irreversible degree of EEG for recognizing whether a patient is suffering from epilepsy or not, which may be a positive index for clinical diagnosis.

Keywords: relative transfer entropy, EEG, symbolic, time irreversibility

PACS: 87.85-d, 05.45.-a

DOI: 10.7498/aps.63.218701

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

† Corresponding author. E-mail: xfliu2@gmail.com

‡ Corresponding author. E-mail: wangj@njupt.edu.cn