第26卷 第1期 自 动 化 学 报 Vol. 26, No. 1
2000年1月 ACTA AUTOMATICA SINICA Jan., 2000

短 文

涨落复杂性在EEG时间序列分析中的应用

陈仲永 钱鸣奇 伍文凯
（浙江大学生物医学工程学系 杭州 310027）
童勤业
（浙江大学非线性应用研究中心 杭州 310027）

摘 要 给出了涨落复杂性定义，并应用在精神分裂症患者的EEG时间序列分析中．通过实验分析表明，涨落复杂性能够用来区分精神患者和正常人之间的EEG，从而有可能为临床脑电分析提出新的量化指标，为精神分裂症等患者的临床诊断提供新的科学方法．

关键词 复杂性，时间序列，EEG．

THE APPLICATION OF FLUCTUATION COMPLEXITY IN ANALYZING EEG TIME SERIES

CHEN Zhongyong QIAN Mingqi WU Wenkai
(*Dept. of Biomedical Engineering, Zhejiang University, Hangzhou* 310027)
TONG Qinye
(*Research Center for Nonlinear Theory and Application, Hangzhou* 310027)

Abstract This paper describes the definition of fluctuation complexity. It is used to analyze the time series for schizophrenics and normal persons EEG. Through a number of experiments, we have found that this complexity can reflect the character of the EEG, and can be used as a feature classifying the schizophrenic patient and normal person.

Key words Complexity, time series, EEG.

1 序言

现有的时间序列处理方法如相关分析、回归分析、频谱分析和统计平均叠加等都是属

1) 国家自然科学基金资助项目(69675020)。
收稿日期 1998-06-08 收修改稿日期 1999-06-02

于线性的处理方法. 线形方法的特点是稳定、平衡、有序和一致性. 然而,随着科学技术的发展,人们又发现了另一类时间序列信号,它表现出来的特性是看起来十分“乱”而类似于噪声,但是其内部却包含着一定的规律. 这一类时间序列表现出非线性的特征. 非线性是以不稳定、不平衡、无序和非一致性为其特征的. 如何刻划和分析这种看来很“乱”而实际又有内在规律的时间序列,这一问题越来越被人们关注. 一门新的非线性时间序列分析学科正在蓬勃地发展起来.

本文以脑电 EEG 信号为处理对象,讨论涨落复杂性分析方法.

2 对脑电 EEG 性质的讨论

脑电 EEG 信号表面看来是十分混乱的,但是其中包含了大量有规律的信息. 脑活动的本身是一非平稳过程. 人从生长发育到衰老死亡也是一个非平稳过程. 按文献报道:人的注意力集中最多也不会超过四秒钟. 人脑可以说时时刻刻在改变着自身的运动状态. 不同的脑功能状态就会有不同的 EEG 特性. 因此可以说 EEG 是一个典型的非平稳过程.

正因为有以上特点,决定了 EEG 信号不能用线性方法处理. 长期以来 EEG 研究的进展远落后于心电图 ECG 的研究,这也说明了线性分析方法的局限性. 目前对非线性时间序列分析和处理有很多方法,如李雅普诺夫方法,分数维分析,熵分析和复杂性分析. 对以上方法我们都作了一定的分析,经比较后本文决定采用复杂性分析方法.

3 涨落复杂性

复杂性分析方法有很多[1~3]. 本文采用涨落复杂性方法来分析 EEG 信号.

一维离散时间动力系统,从闭区间 A 到自身的映射 F

$$F: A \to A, \quad X \to F_r(x). \tag{1}$$

$X \in A$ 是系统的状态,$r \in R$ 是系统的参数空间. 动力系统在时间上离散而在状态空间上连续. 可通过状态 A 的离散化生成一个符号动力学系统. 即将状态空间 A 离散为 N 个勒贝格测度上非空的元素 $A_i(i=1\sim N)$. A_i 和 A_j 两两不相交,且 $\bigcup_{i=1}^{N} A_i = A$,$A$ 上的每个元素用符号 a_i 表示,则动力系统的时间演变可用符号序列 $S=S_0S_1\cdots S_n$ 表示. 在时间间隔 i 后,系统的状态 x_i 的符号 S_i 由在时刻 i 与轨迹相交的元素 A_i 决定,符号 $S_i \in \{a_1, a_1, \cdots a_N\}$ 符号集. 符号动力学系统定义为

$$\sum\nolimits_F \to \sum\nolimits_F, \qquad S_1 \longmapsto \hat{\sigma}_F(S) = S'. \tag{2}$$

序列中每个符号满足条件 $S_{i+1}=S'=\hat{\sigma}_F(S_i)$,$\sum_F$ 是所有允许的符号序列空间. $\hat{\sigma}_F$ 为 $\sum_F$ 上的映射.

在前一状态 A_i 已知时,确定下一个状态 A_j 所需的信息量为

$$G_{ij} = -\log P_{i\to j}, \tag{3}$$

$P_{i\to j}$ 是从状态 A_i 到下一个状态 A_j 的条件转移概率,$P_{i\to j}=\dfrac{P_{ij}}{P_i}$.

平均信息增益 $\langle G \rangle$ 是所有 $i \to j$ 信息的平均值.

$$\langle G\rangle = \sum_{i,j=1}^{N} P_{ij}G_{ij} = -\sum_{i,j=1}^{N} P_{ij}\log P_{i\to j}. \tag{4}$$

利用转移概率的定义,$\langle G\rangle$可简化为

$$\langle G\rangle = -\sum_{i,j=1}^{N} P_{ij}\log P_{ij} + \sum_{i=1}^{N} P_{i}\log P_{i}, \tag{5}$$

其中 P_i 表示系统轨迹落在 A_i 中的概率,P_{ij}是系统在相邻两时间间隔落入状态 A_i 和 A_j 中的联合概率. 如果第二项$\sum_{i=1}^{N} P_i\log P_i$ 是时刻 n 的信息量$-I_n$,即 $I_n=-\sum_{i=1}^{N} P_i\log P_i$,则第一项描述了时刻 $n+1$的信息量 I_{n+1}.

另外,在所有可能的转移 $i\to j$ 中信息损失量 L_{ij}的平均值确定了系统在已经进入下一个状态 A_j 后得到的关于前一个状态 A_i 丢失的信息

$$\langle L\rangle = \sum_{i,j=1}^{N} P_{ij}L_{ij} = -\sum_{i,j=1}^{N} P_{ij}\log P_{i\leftarrow j}, \quad L_{ij} = -\log P_{i\leftarrow j}. \tag{6}$$

其中 $P_{i\leftarrow j}$是从状态 A_j 到前一个状态 A_i 的条件转移概率,则净信息量

$$\Gamma_{ij} = G_{ij} - L_{ij} = \log\frac{P_i}{P_j}. \tag{7}$$

归一化$\sum_{j} P_{i\to j}=1$,平均信息量$\langle\Gamma\rangle=\sum_{i,j} P_{ij}\Gamma_{ij}=0$. 在系统演变过程中,$\Gamma_{ij}$在其平均值上下波动,因而有非零的均方差 σ_Γ^2. 这可理解为净信息量的涨落,由 Bates 和 Shepard 引入复杂性测度[2],我们用 C_F 表示

$$C_F \triangleq \sigma_\Gamma^2 = \langle\Gamma^2\rangle - \langle\Gamma\rangle^2 = \sum_{i,j=1}^{N} P_{ij}\left(\log\frac{P_i}{P_j}\right)^2. \tag{8}$$

当系统是周期行为(周期为 φ)时,状态概率 $P_i=P_j=1/\varphi$, 故 $C_F=\sigma_\Gamma^2=0$. 表明有规律的周期过程是不复杂的.

当系统是完全随机行为时,因为状态空间 A 是等概率的过程分布($P_i=P_j$),故 $C_F=\sigma_\Gamma^2=0$. 说明完全随机的过程也是不复杂的. 而真正复杂的是介于这二者之间,这与文[1]中的 C_2有些相近.

4　实验结果与分析

用上面介绍的涨落复杂性来分析精神分裂症病人的 EEG,并运用统计理论对这些结果进行分析. 所用的脑电数据是在杭州市第七医院采集的,受试者分为二类:精神分裂症患者(98人)、正常人(34人). 取每位受试者的二种状态:睁眼、闭眼.

1) 应用上面介绍的方法计算每个通道复杂性值 C_F,将每个人的8个通道的 C_F 值组合成一8维向量,它代表了一个人的信息. 我们把状态空间 A 离散为 A_i,$i=1,2,\cdots,512$和 $i=1,2,\cdots128$两种,分别计算 C_F. 状态空间 A 的划分如下:根据 EEG 信号的幅值大小,将 EEG 幅值划分为512(128)等份. 这样 P_i 等于落入 A_i 的点数除以总点数. P_{ij}等于当前时刻落入 A_i 且在下一时刻落入 A_j 的点数除以总点数.

根据文献[5]所提出的方法,对所得样本集进行样本均值检验. 检验结果表明:两样本的均值是有差异的,可以据此作为判别准则.

2) 判别分析. 本文采用的是贝叶斯分类方法[4],假设样本 Ω_i 服从 $N(\mu_i,\Sigma_i)$,由文[4]

可得判别函数

$$g_i(x)=\frac{P(\omega_i)}{(2\pi)^{d/2}\|\sum_i\|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu_i)^{\mathrm{T}}\sum{}^{-1}(x-\mu_i)\right),\tag{9}$$

$P(\omega_i)$为类型Ω_i的先验概率,μ_i为类型Ω_i的平均向量,Σ_i为类型Ω_i的协方差矩阵.

对于二类样本,由贝叶斯分类准则,可得到二类样本Ω_i,Ω_j的决策面方程

$$G(x)=\ln g_i(x)-\ln g_j(x)\begin{cases}<\\>\end{cases}0,\quad 则\ x\in\begin{cases}\Omega_i,\\ \Omega_j.\end{cases}\tag{10}$$

采用留一法(Leave-one-out method)所得结果如下表所示.

表1 (状态空间划分为512个元素)

	睁眼状态(δ=1.08)			闭眼状态(δ=1.39)		
	病人	正常人	判对率	病人	正常人	判对率
病人(98)	76	22	77.55%	72	26	73.47%
正常人(34)	4	30	88.24%	4	30	88.24%

表2 (状态空间划分为128个元素)

	睁眼状态(δ=1.36)			闭眼状态(δ=1.07)		
	病人	正常人	判对率	病人	正常人	判对率
病人(98)	61	37	62.24%	63	35	64.28%
正常人(34)	4	30	88.24%	3	31	91.18%

注. δ为两样本集之间的马氏距离$=(\mu_1-\mu_2)^{\mathrm{T}}\Sigma(\mu_1-\mu_2)$,$\Sigma$为两类样本的协方差矩阵. μ_1为精神病患者样本集的均值向量,μ_2为正常人样本集的均值向量.

从以上结果可以得出:

1)涨落复杂性C_F能够反映正常人和精神分裂症的特征,能够区分正常人和精神分裂症的特征.

2)由表1和表2可以看出,对状态空间越细分,复杂性C_F更能反映出EEG序列的细微结构,这样就克服了以往各种复杂性(如Kolmogorov复杂性,C1C2等)过分粗粒化,丢失细节太多的问题.

3)用涨落复杂性在临床EEG中已取得一定成果,这不但为EEG分析提供了一个有力工具,同时此法也可以用于其它非线性时间序列处理.

参 考 文 献

1 Tong Qinye ,Kong Jun. A note on analyzing schizophrenic EEG with complexity measure. *Chaos Soliton & Fractals*, 1996,**7**(3):371~375

2 Bates J E,Shepard H K. Information fluctuation as a measure of complexity. Durham: University of New Hampshire,1991

3 徐京华,吴祥宝. 以复杂度测度刻划人脑皮层上的信息传输. 中国科学(B辑),1994,**24**(1): 57~62

4 李金宗. 模式识别导论,北京:高等教育出版社,1994

5 张尧庭,方开泰著. 多元统计分析引论. 北京:科学出版社,1982

6 李从珠编译. 判别分析,北京:群众出版社,1988

7 Waltor J Freeman. Tutorial on nurobiology from single neuron to brain chaos, *Int. J. of Bifuration and Chaos*, 1992,**2**(3):451~482

陈仲永　男，1972年生．毕业于浙江大学研究生院生物医学工程专业，硕士学位．现在浙江省技术监督局条码中心从事信息处理工作．

钱鸣奇　男，1961年生．浙江大学生命科学与医学工程学系讲师，从事非线性科学的应用研究和教学工作．

伍文凯　男，1970年生．浙江大学研究生院生物医学工程专业在读硕士研究生．专攻非线性技术在生物医学信息中的应用．

童勤业　男，1939年生．浙江大学非线性应用研究中心副主任、教授、博士生导师．主要从事非线性科学在信息处理中的应用研究和教学工作．

第三届全球智能控制与自动化大会关于征文的补充通知

经多方努力，第三届全球智能控制与自动化大会（WCICA’2000）（原名：第三届全球华人智能控制与智能自动化大会，CWC ICIA’2000）已获得 IEEE 的出版书号，因此大会论文将被国际论文检索系统（包括 EI）收录．希望各位作者认真准备论文，踊跃投稿．

WCICA’2000国际程序委员会将采取国内外稿件统一评审方案，审稿将于2000年1月中旬进行．请投稿者提交3份用中文或英文写成的论文全文复印件，并另附上一页，在其上注明：文章标题；作者姓名和单位；联系人姓名及其详细通讯地址（包括电话、传真和电子邮件地址）；论文摘要；3—5个关键词以及所属征文范围．中文稿件请在文末附文章标题、作者姓名和单位、论文摘要以及关键词的英文译文．请在1999年12月31日前将论文寄到以下地址：

230027
安徽合肥　中国科学技术大学自动化系
WCICA’2000组委会

特此通知．

第三届全球智能控制与自动化大会组委会
一九九九年十月

附：征文范围

·智能系统与专家系统·智能控制·神经网络及其应用·模糊系统与模糊控制·机器人学与机器人控制·大系统·调度、规划、管理与决策系统·自治、容错和故障诊断系统·制造系统和 DEDS·计算机辅助设计·信息处理与信息系统·系统理论和控制理论·建模、辨识和估计·自适应控制·变结构控制·非线性系统及其控制·遗传算法·混合动力学系统·H_∞ 控制和鲁棒控制·最优化与最优控制·分布式计算机控制系统·仪器仪表·其它