

基于高阶复杂性测度的心率变异信号分析

朱家富^{1,2}, 杨浩¹, 何为¹

(1. 重庆大学电气工程学院电工技术研究所, 重庆 400044;

2. 渝西学院物理与电子信息工程系, 重庆 402168)

摘要: 心率变异性反映了交感神经和迷走神经对心血管系统的综合调节作用, 是评价心血管系统功能的重要指标。复杂度是描述时间信号序列信息量的一个重要参数, 传统算法中的过分粗略化会丢失大量的有用信息, 而高阶复杂度的引入可较大幅度地避免这一问题。在利用 Lorenz 模型数据仿真的基础上, 分别对 25 例正常人样本和 25 例充血性心力衰竭病人样本的心率变异信号的 1~15 阶 Kolmogorov 复杂度进行了计算, 通过对比分析后确定, 5 阶 Kolmogorov 复杂度在临床医学上可以为分析心率变异信号获得更为理想的效果。

关键词: 心率变异信号; Kolmogorov 高阶复杂度; 充血性心力衰竭; Lorenz 模型

中图分类号: R318.04

1 引言

心率变异性(heart rate variability, HRV)是连续心搏间瞬时心率的微小涨落, 是窦性心率在一定时间内周期性改变的现象, 其体现了心率或心动周期的波动性。近代医学证明, 心率的这种波动性并非是对外在环境变化的响应, 而是由心血管系统内在机制作用的结果, 即使是正常人在安静状态下的心率也是在一定范围内波动的, 也存在心率变异性。影响心率波动的因素包括脑的高级神经活动、中枢神经系统的自发性节律活动、呼吸活动以及由压力和化学感受器传入的心血管反向活动等, 这些因素最终通过对心脏的交感神经和迷走神经的综合调节作用而形成心率变异性。心率变异性可以反映自主神经系统的活性, 是反映交感-副交感神经张力及其平衡的重要指标, 是心力衰竭的一项强有力的预测因子。

心率波动呈现显著的非线性动力学特性^[1], 这点已经被许多学者证实并接受, 可以用非线性系统的各种特征量来描述 HRV 信号。由于 HRV 信号具有动态变化大、稳态时间短、数据长度短等特点, 在多个非线性特征量中, 复杂度最适合用来分析 HRV 信号。而一阶复杂度指标在区别正常人和病人时区分度较小, 我们通过引入高阶复杂度指标, 对 50 个不同样本的心率变异信号的 1~15 阶 Kolmogorov 复杂度进行了计算, 通过对比分析发现, 5 阶 Kolmogorov 复杂度可以为分析心率变异

信号获得更为理想的效果。

2 复杂性测度

复杂度的最初定义是由 Kolmogorov 于 1965 年提出的^[2], 表征为能够产生某一 (0,1) 序列所需的最短程序的比特数, 后来由 Lempel 和 Ziv 等人给出了实现这种定义复杂度的具体算法^[3,4], 并称为 Kolmogorov 复杂度或算法复杂度, 本文用 C_k 来表示, 其大小体现了时间序列的信息含量^[5]。

Kolmogorov 复杂度的基本算法描述为: 以时间序列的均值为参考, 大于均值取“1”, 小于均值取“0”, 重构出一新的相同数据长度的 (0,1) 序列; 计算 (0,1) 序列中禁止字或禁止字符串的个数, 被称为绝对复杂度 $C(n)$; 由于任何序列的绝对复杂度 $C(n)$ 都趋近于一个定值 $b(n)$,

$$b(n) = \frac{n}{\log_2 n} \quad (1)$$

所以定义相对复杂度, 也称归一化复杂度 C_k , 为

收稿日期: 2003-11-17

基金项目: 重庆大学高电压与电工新技术教育部重点实验室基金 (2002040)

通讯作者: 朱家富, 电话: (023)65106977-8101,

E-mail: szzjf@163.com

$$C_k = \frac{C(n)}{b(n)} = C(n) \frac{\log_2 n}{n} \quad (2)$$

C_k 是一个与序列长度 N 有关的物理量, 可以看出完全随机序列的 C_k 趋近于 1, 而规律的或周期的序列的 C_k 趋近于 0。

我国著名学者徐京华等^[6]于 1994 年提出了新的复杂度 C_1 、 C_2 的定义, 它是描述非线性时间序列的又一个复杂性测度。本研究中以 Kolmogorov 复杂度 C_k 为复杂性测度研究指标。

3 高阶复杂度

在计算时间序列的复杂度时, 首先将原序列以序列的均值为参考字符化, 得到一个与原序列相同长度的(0, 1)字符序列, 这一过程的实质是将原序列颗粒化, 而这一过度粗略化行为丢失了原序列中所包含的幅度信息, 使得新序列中的高幅波起主要作用, 而淹没了原时间序列中的低幅波, 忽略了其中的细微变化信息。为弥补这一缺陷, 我们引入高阶复杂度的概念, 以减小粗略化对原序列信息的丢失^[7]。其方法是对原序列与序列均值的差值的绝对值重构新的序列, 如原序列为 $x_i (i=1, \dots, N)$, N 为数据长度, 则重构的新时间序列为:

$$y_i = \left| x_i - \frac{1}{N} \sum_{j=1}^N x_j \right| \quad (i, j=1, \dots, N) \quad (3)$$

然后将新序列 y_i 字符化, 计算所得的复杂度称为原序列 x_i 的二阶复杂度。多次重复序列的重构过程, 可以计算得到原序列的任意高阶复杂度。以如(4)式所示的 Lorenz 模型为例, 来仿真不同阶次的复杂度描述非线性时间序列信号中信息的不同特征:

$$\begin{cases} \dot{x} = \sigma(y-x) \\ \dot{y} = rx - y - xz \\ \dot{z} = -bz + xy \end{cases} \quad (4)$$

取系数 $\sigma=10$ 、 $r=28$ 、 $b=8/3$, 初始值取(1,2,3), 步长取 0.01, 利用四阶 Runge-Kutta 公式通过 5 000 次迭代后取 x 的值为实验用非线性时间序列, 以此为对象分别计算其不同数据长度下的 1~15 阶复杂度 C_k , 计算结果如图 1 所示。

由图 1 可以看出, 时间序列的复杂度随序列的长度 N 的增大而相对减小, 而随着阶次的增大而增大。高阶复杂度存在一定的限度, 超过一定的阶次限度后, 复杂度的数值将会出现振荡, 失去了原有的稳定性。合理地选择高阶复杂度的阶次, 以增加

时间序列复杂度的区分度, 同时保持原有的稳定性, 可以在对时间序列的复杂度进行分析时获得更理想的效果。

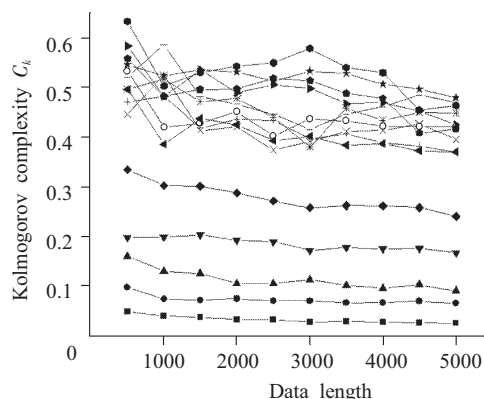


Fig.1 Curves of complexity measure C_k in different orders and data length

—■—: OR1; —●—: OR2; —▲—: OR3; —▼—: OR4;
—◆—: OR5; —◀—: OR6; —▶—: OR7; —●—: OR8;
—★—: OR9; —◆—: OR10; —○—: OR11; —+—: OR12;
—×—: OR13; —*—: OR14; — — —: OR15

4 分析方法与结果

4.1 HRV 信号的提取

心率变异信号可以从完整的心电信号(electrocardiogram signal, ECG)中提取, 常用心动周期内的信号采样点数或者心动周期时间来表示。本文采用的原始心电信号是利用 Holter 以 128 Hz 采样信号、12 位 A/D 转换电路采样所得, 为保证能获得足够长度的 HRV 信号, 样本心电信号必须连续采集 20 min 以上。HRV 信号从心电信号中分离得到, 其数值的大小采用连续心动周期内(相邻的 R-R 间期)信号采样点的数目 n 来表示, 它是一个与时间相关联的量。实际测得的 R-R 间期在时间轴上是不等间隔的, 采用 Kobayashi 的线性插值法将不等间隔的 R-R 间期序列变成等间隔的采样数据序列^[8], 每个样本的计算用 HRV 信号数据的长度均取 1 000 点。图 2 显示的是心电信号和心率变异信号片段。

实验数据分两组样本, 分别是正常人样本组和充血性心力衰竭病人样本组, 每组均为 25 个样本个体。其中正常人样本组中男性 17 人, 女性 8 人, 最小年龄 22, 最大年龄 46, 平均年龄 26.8; 充血性心力衰竭病人样本组中男性 11 人, 女性 14 人, 最小年龄 34, 最大年龄 72, 平均年龄 58.8。

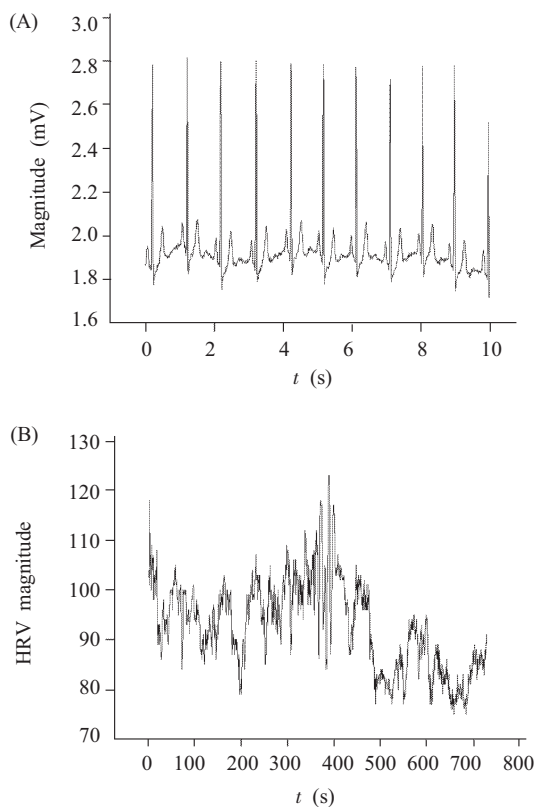


Fig.2 Examples of signal waves in time. (A) ECG signal waves; (B) HRV signal waves

4.2 HRV 信号的复杂度

我们对两个样本组共 50 个样本个体的 HRV 信号的 1~15 阶的复杂度分别作了计算，并按不同的组和不同的阶次进行统计，其统计平均值的结果如图 3 所示。由图 3 可以看出，不管是正常人

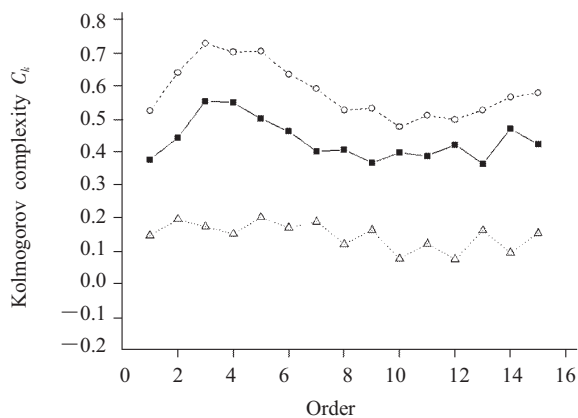


Fig.3 Comparing curves of mean complexity C_k of two groups in different orders
 —■—: Control; —○—: Normal; —△—: Difference

组还是充血性心力衰竭病人组，其复杂度 C_k 的数值在四阶以下随着阶数的增大而增加，在四阶以上随着阶数的增大而减小；而通过不同的样本组之间的对比可以看出，无论在多少阶次下，正常人组的复杂度 C_k 的平均值总比充血性心力衰竭病人组的大，其间的差值约为 0.2 左右。

同时，我们对每个样本组的 25 个样本个体的复杂度 C_k 的分布进行了统计，其统计结果如图 4 所示，图中的 x 轴表示复杂度 C_k 的值处在坐标值的 ± 0.5 范围内， y 轴表示复杂度的阶次， z 轴表示在一定阶次下 C_k 的值处于某一范围内的样本的个数。从图 4 可以看出，不管是正常人样本还是病人样本，在低阶(<4)和高阶(>7)时，样本的复杂度相对较低，其 C_k 数值也比较分散，而样本的 4~7 阶复杂度的数值相对较大，且复杂度数值的分布也相对集中。

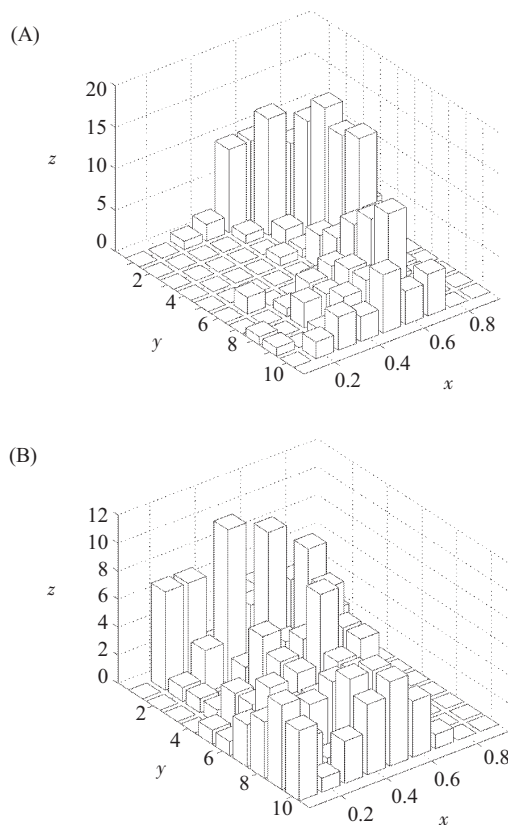


Fig.4 Plots of statistical distribution of two sample groups in different orders scaled in the y -axis, each 25 samples, and the z -axis represents the amount of samples which complexity indexes locate in the range (± 0.5) scaled in the x -axis. (A) Normal group; (B) Patient populations group

4.3 HRV 高阶复杂度阶次的选择

由于样本的复杂度 C_k 存在较大的个体差异,因此,合理地选择复杂度的阶次更有利于区分正常人和病人,最佳阶次的选择以最大可能地区分正常样本和病理样本为宜。我们分别计算了两组样本的一

阶至十五阶 Kolmogorov 复杂度,其计算结果显示,两个样本组的样本个体的复杂度 C_k 在四阶以下和七阶以上都具有较大的动态范围,两个样本组之间的 C_k 图形存在较大的混叠,反过来可以认为,依据样本的 C_k 我们无法准确判断样本的性质。而在五阶复杂度附近,正常样本组和病人样本组的复杂度指数 C_k 出现了明显的分层,两样本组之间的混叠较少,其中以五阶复杂度的区分最为明显,这样更有利于区分样本的性质。图 5 为两种样本个体在三个典型不同阶次下的 C_k 对比图。

从图 4 的 C_k 分布统计结果来看,五阶 Kolmogorov 复杂性 C_k 在同类样本中的分布最集中,而不同类样本之间的差异最大,分层最为明显。因此,利用 HRV 信号的五阶 Kolmogorov 复杂性指数 C_k 可以作为识别充血性心力衰竭的重要参数。

5 讨 论

复杂性测度是描述非线性时间序列信息量的一个重要参数,它有多种定义方式,但以 Kolmogorov 复杂度 C_k 最为常用。高阶复杂度的根本是在分析信号时找到信号中是高幅波还是低幅波起主要作用为切入点,减小粗略化对原序列信息的丢失,突出了原始信号中的细微变化信息特征,增加对信号序列性质识别的区分度。分析表明,在利用复杂度描述非线性时间序列的信息特征时,并非阶次越高越好,超过一定的阶次限度后,复杂度的数值将会出现振荡,失去原有的稳定性。最佳阶次的选择以复杂度的数值区分度最大、随数据长度变化的稳定性最好为宜。

实验所采用的病例为充血性心力衰竭病人,病例的确定以临床诊断为准。本研究的方法与结果是否适用于其它类心血管类疾病病理样本还有待作进一步的研究与临床试验。

研究表明,在分析正常人样本组和充血性心力衰竭病人样本组的 HRV 信号时,五阶 Kolmogorov 复杂度比其它阶次复杂度更能体现样本的性质,五阶复杂度是分析充血性心力衰竭病人 HRV 信号的最佳阶次。依据这一结论可以作为临床诊断的辅助手段,提高诊断的准确率。复杂度阶次的确定依分析时间序列样本的性质而异,并且只有在大量样本实验的基础上,根据统计结果所得到的结论才具有更高的科学性和准确性。

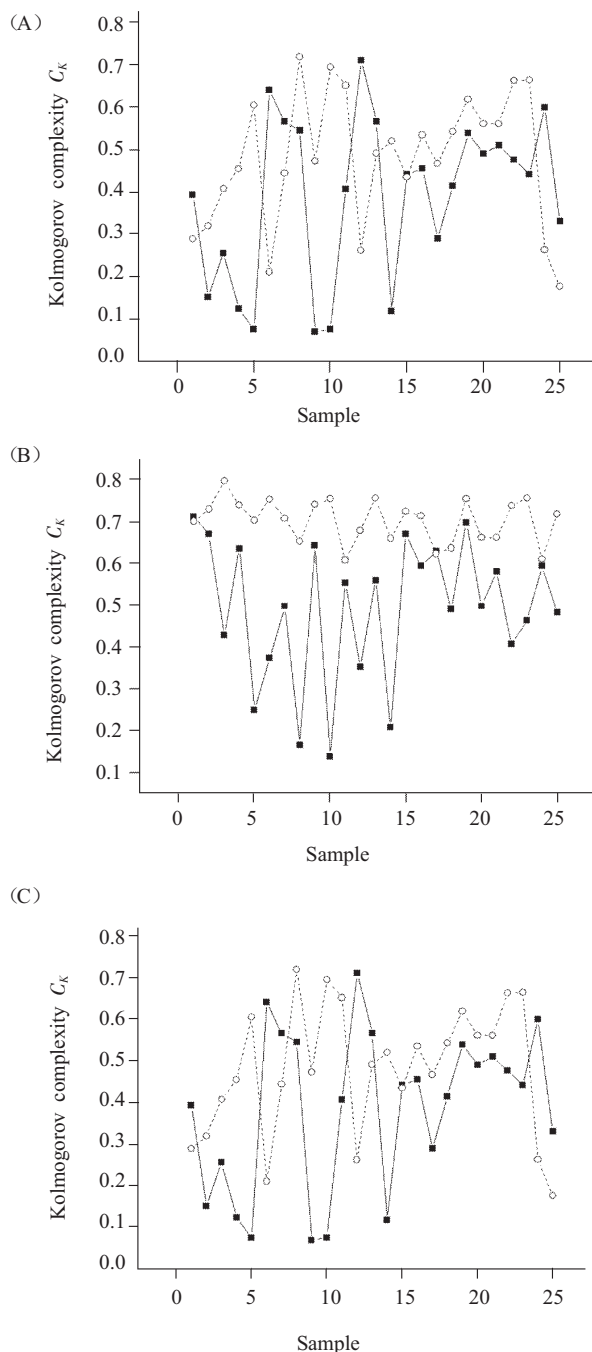


Fig.5 Comparing curves of C_k in different orders of two groups. (A) 1st order complexity measure; (B) 5th order complexity measure; (C) 10th order complexity measure

—■—: Control; —○—: Normal

参考文献:

- [1] Zwiener U, Hoyer D, Bayer R. Deterministic chaotic and periodic properties of heart rate and arterial pressure fluctuations and their mediation in piglets. *Cardiovascular Research*, 1996,31:455~462
- [2] Kolmogorov AN. Three approaches to the quantitative indefination of information. *Inform Tuams*, 1965,1:3~11
- [3] Lempel A, Ziv J. On complexity of finite sequences. *IEEE Trans Inform Theor*, 1976,22:75~88
- [4] Kasper K, Schuster HG. Easily calculable measure for complexity of sparial-temporal, pattens. *Physics Review*, 1987,A36: 843~848
- [5] 陈仲永, 伍文凯, 童勤业, 严筱刚. 基于复杂性测度的帕金森病人 EEG 分析. *生物医学工程学杂志*, 1999,16(2):218~221
- [6] Xu JH, Liu ZR, Liu R. The measures of sequence complexity for EEG studies. *Chaos, Soliton and Fractals*, 1994,4: 2111~2119
- [7] 孟欣, 邹睿, 邵颀, 欧阳楷. 简化 EEG 为符号序列及对其复杂度的研究. *山东生物医学工程*, 1995,14(2):9~13
- [8] Kobayashi M, Musha T. 1/f Fluctuation of heartbeat period. *IEEE Trans Biomed*, 1988,58:203~211

ANALYSIS OF HRV SIGNALS BASED ON THE HIGH-ORDER COMPLEXITY MEASUREMENT

ZHU Jia-fu^{1,2}, YANG Hao¹, HE Wei¹

(1. College of Electrical Engineering, Chongqing University, Chongqing 400044, China;

2. Department of Physics & Electronic Information Engineering, Western Chongqing University, Chongqing 402168, China)

Abstract: The heart rate variability(HRV) reflects the synthetical function of the parasympathetic-sympathetic nerve to adjust and control the heart and blood system, and it is an important index to evaluate the function of the heart and blood system. The complexity is a valuable parameter representing the information contained in a time series, but the traditional algorithm of measure complexity is so rough that a lot of information contained in the time series will be lost, and this can be avoided by higher order complexity. Based on a simulation of Lorenz model, the all Kolmogorov complexity indexes of two groups of samples of 25 normal populations and 25 congestive heart failure patients were calculated and analyzed, and an important conclusion was given that the 5th order complexity was the best equal complexity index to represent the HRV.

Key Words: HRV signals; High-order Kolmogorov complexity; Congestive heart failure; Lorenz model