

## 应用多尺度化的基本尺度熵分析心率变异性\*

严碧歌 赵婷婷<sup>†</sup>

(陕西师范大学物理学与信息技术学院, 西安 710062)

(2010年10月13日收到; 2011年3月10日收到修改稿)

本文采用多尺度化的基本尺度熵方法, 针对心率变异性信号进行了分析, 研究发现多尺度化的基本尺度熵可以区分不同生理病理信号, 包括健康人、充血性心力衰竭患者和房颤心律失常患者的心率变异性信号, 以及健康人白天黑夜的心率变异性信号. 通过对健康人代理数据的分析, 发现房颤心律失常患者与代理数据的熵值趋势相似, 研究表明房颤心律失常患者的心率变异性信号更多的是反映生理信号的线性特征, 而对环境变化不能很好的进行自我调节.

**关键词:** 多尺度化的基本尺度熵, 心率变异性, 充血性心力衰竭, 房颤心律失常

**PACS:** 87.85.Ng

## 1. 引言

心脏是一个复杂的、受到神经、血压、激素等多种因素调节控制的、开放的、非线性的和远离平衡态的动力学系统. 心脏组织结构和控制结构的复杂性是心脏电信号复杂的非线性特性的基础. 近20年来, 许多研究者在这方面作了深入的研究.

心脏电信号中研究最多的是心率变异性信号 (heart rate variability, HRV). 心率变异性是指逐次心跳间期之间的时间差异<sup>[1,2]</sup>. 临床 HRV 信号的研究最早要归功于 1965 年 Hon 和 Lee<sup>[3]</sup> 发现胎儿宫内缺氧时, 最先反映出来的是其 RR 间期变异. 心率变异性信号中隐藏着许多生理系统潜在的机理, 可以反映心脏的生理功能和健康状态. 心率变异性的异常可以作为预测心脏事件或猝死的一个重要手段, 其临床价值越来越受到重视<sup>[4,5]</sup>.

20 世纪 90 年代以来, 非线性的概念和研究方法广泛应用到 HRV 的分析中, 并取得了一定的进展<sup>[6,7]</sup>, 其中比较有代表性的一类方法是熵分析法<sup>[8-11]</sup>. 熵是热力学中一种度量混乱与随机程度的测度, 也被广泛应用到时间序列分析中. 心跳间隔序列作为确定性非线性系统产生的非随机信

号, 仅用概率分布函数描述是不够的, 间隔排列的先后顺序同样也蕴含大量的内在动力学特性, 熵测量则能在一定程度上衡量这种时间先后的关联性. 熵测度以其方法简单、运算快速、抗干扰能力强的优点为探测和捕捉时间序列中的有用信息提供了方便, 例如近似熵 (approximate entropy, AE), 作为时间序列复杂性的测度, 一经提出, 马上在众多的科学领域得到广泛应用<sup>[12-14]</sup>. 此外, 还有样本熵 (sample entropy, SE), 基本尺度熵 (base-scale entropy, BE), 多尺度熵 (multiscale entropy, MSE) 等, 都在心率波动性或其他生理信号分析方面有较成功的应用. 然而近似熵方法需要采用一个表示连续数据点的向量来估计, 然后用序列中的数据点依次积分来计算近似熵值, 明显的噪声干扰和非平稳数据都会影响近似熵得到有意义的结果<sup>[15]</sup>. 样本熵和基本尺度熵都是在一个最小的尺度上估计序列长度为  $m$  和  $m+1$  之间的不同, 而忽略了其他的尺度. 多尺度熵对心跳间隔序列的分析更多的是反映年龄上的差别, 而对充血性心力衰竭患者 (CHF) 无法形成临床上的诊断依据<sup>[16]</sup>.

为了更全面的研究心率变异性信号的复杂性, 本文利用多尺度化的基本尺度熵方法<sup>[17]</sup>, 对心率变异性信号进行了分析.

\* 陕西省自然科学基金 (批准号: 2010JM1007) 和基础研究经费 (批准号: GK200902053) 资助的课题.

<sup>†</sup> 通讯联系人. E-mail: 601111ting@163.com

## 2. 多尺度化的基本尺度熵

对于数据长度为  $N$  的时间序列  $x$ ,

$$\{x(i); 1 \leq i \leq N\}, \quad (1)$$

给定时间尺度因子  $\tau$ , 对原始时间序列  $x$  按以下方式进行粗粒化<sup>[18]</sup>:

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{(j-1)\tau+1}^{j\tau} x_i, \quad (1 \leq j \leq N/\tau). \quad (2)$$

将该时间序列嵌入  $m$  维相空间. 对于每一个  $y(i)$ , 取  $m$  个点组成  $m$  维矢量

$$Y(i) = [y(i), y(i+L), \dots, y(i+(m-1)L)], \quad (3)$$

这里  $m$  是嵌入维数,  $L$  是延迟时间. 取  $L = 1$ , 那么共有  $N - m + 1$  个  $m$  维矢量. 对于每一个  $m$  维矢量, 计算出基本尺度 BS, BS 定义为  $m$  维矢量中所有相邻点数据间隔 (或大小) 的差值方均根值, 即

$$BS(i) = \sqrt{\frac{\sum_{j=1}^{m-1} [(y(i+j) - y(i+j-1))]^2}{m-1}} \quad (4)$$

根据基本尺度我们选择符号划分标准为  $\alpha \times BS$ , 然

后将每一个  $m$  维矢量转换成  $m$  维矢量符号序列  $S_i(Y(i)) = \{s(i), s(i+1), \dots, s(i+m-1)\}$ ,  $s \in A$  ( $A = 0, 1, 2, 3$ ). 具体的转换如下:

$$S_i(Y(i)) = \begin{cases} 0, & \bar{y} < y_{i+k} \leq \bar{y} + a \times BS, \\ 1, & y_{i+k} > \bar{y} + a \times BS, \\ 2, & \bar{y} - a \times BS < y_{i+k} \leq \bar{y}, \\ 3, & y_{i+k} \leq \bar{y} - a \times BS, \end{cases} \quad (5)$$

这里  $i = 1, 2, 3, \dots, N - m + 1, k = 0, 1, 2, \dots, m - 1$ .  $\bar{y}$  代表第  $i$  个  $m$  维矢量的平均值, BS 代表第  $i$  个  $m$  维矢量的基本尺度, 符号 0, 1, 2, 3 仅仅作为每一个划分区域的记号, 具体数值的大小并没有实际的意义.  $a$  是一个特殊参数.  $a$  的值取过大会导致细节信息的丢失, 不能很好的捕捉到信号中的动态信息.  $a$  的值取过小会导致噪声的明显的影响.

为了计算多尺度化的基本尺度熵值, 我们先统计  $m$  维矢量符号序列  $S_i$  的分布概率  $P(S_i)$ . 包含 0, 1, 2, 3 四种符号的  $m$  维矢量符号序列  $S_i$  ( $m$  - words) 共有  $4^m$  种不同的组合形式  $\pi$ . 统计每一种不同的组合形式在整个  $N - m + 1$  个  $m$  维矢量所占的概率为

$$p(\pi) = \frac{\#\{t | 1 \leq t \leq N - m + 1, (y_t, \dots, y_{t+m-1}) \text{ has form } \pi\}}{N - m + 1}, \quad (6)$$

其中, #表示个数.

$m$  维矢量的多尺度化的基本尺度熵定义为

$$H(m) = - \sum P(\pi) \log_2 P(\pi). \quad (7)$$

为了计算方便,  $m$  的取值一般为  $m = 3, 4, 5, 6, 7, N$  的取值一般只要大于  $4^m$  即可.

## 3. 应用 MBE 分析 HRV 信号

我们首先研究健康人, 房颤心律失常 (atrial fibrillation, AF) 患者和充血性心力衰竭 (congestive heart failure, CHF) 患者的多尺度化的基本尺度熵值. 根据经验选取  $M = 4, \alpha = 0.2$  来计算熵值 (参数  $m = 3, 4, 5$  对结果没有太大影响). 分别计算了 30 个健康人, 29 个充血性心力衰竭患者和 20 个房颤心律失常患者的 HRV 信号, 原始数据总长  $N = 90000$ , 时间尺度  $\tau$  从 1 到 20 变化. 所有数据均来自文献 [19]. 计算结果如图 1 所示.

从图 1 可以看出, 在尺度  $\tau = 1$  时, CHF 患者熵

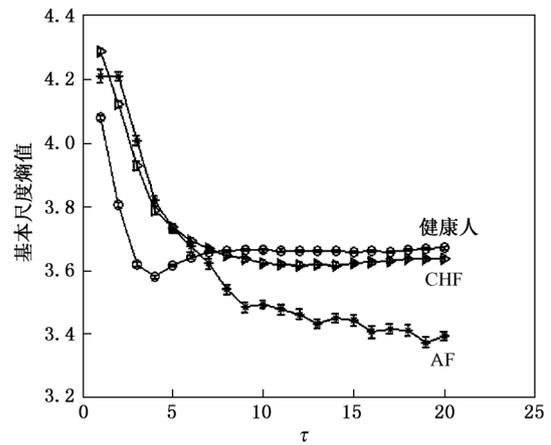


图 1 HRV 的 MBE 分析 (数值用平均值  $\pm$  标准差表示)

值最大, AF 患者熵值居中, 健康人熵值最小, 随着时间尺度的增加, CHF 患者和 AF 患者的熵值开始下降, 并且 AF 患者熵值下降幅度远大于 CHF 患者, 而健康人的熵值先下降后升高, 在  $\tau \geq 8$  时, 三组样本进入平台区, 并且以健康人熵值为最高, CHF 患者

次之, AF 患者为最低, 这与尺度 1 时的情况完全不同. 说明健康人的 HRV 序列在小的时间尺度下, 更集中于特定的波动模式, 但在大的时间尺度下, 却有更丰富的波动模式. 因此, 健康人的心率变异性是小尺度时的规则性加上大尺度的不确定性, 从而可以适应外界环境的不同时间尺度的复杂变化<sup>[20]</sup>. 健康人的熵值代表最佳生理状态, CHF 患者和 AF 患者远离了健康人的熵值, 当心脏本身病变时, 心脏内组织结构的自谐控制受到破坏, 造成心脏活动偏离最佳状态.

为了深入研究 HRV 序列的多尺度性质, 我们对图 1 中健康人的数据首先进行了傅里叶变换, 然后打乱振幅, 在进行反傅里叶变换, 得到一组代理数据, 代理数据保留了原始数据的线性特征, 而消除了非线性特征. 代理数据的熵值如图 2 所示.

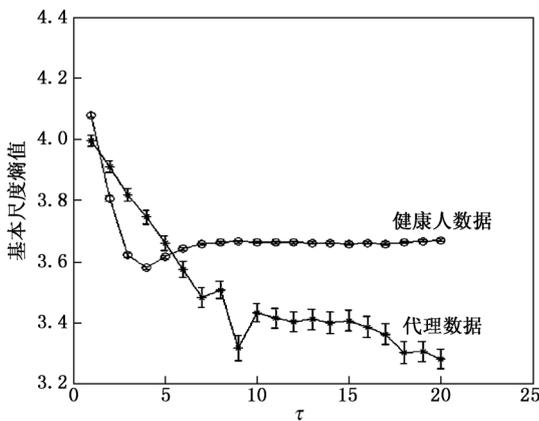


图 2 健康人数据与代理数据的 MBE 分析数值用平均值  $\pm$  标准差表示

由图 2 我们可以看出, 在尺度  $\tau = 1$  时, 代理数据的熵值低于健康人, 在尺度  $2 \leq \tau \leq 5$  时, 代理数据熵值高于健康人, 在  $\tau \geq 6$  时, 代理数据熵值又再次低于健康人的值. 和图 1 相比较, 我们发现 AF 患

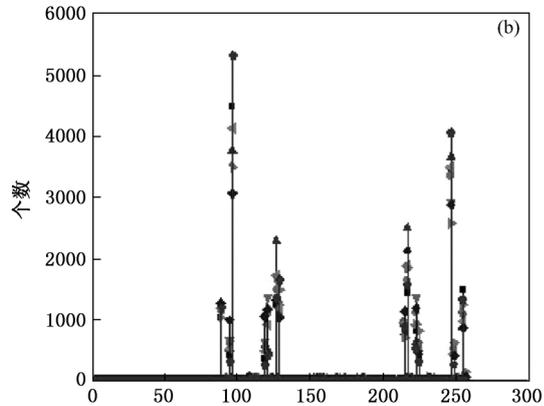
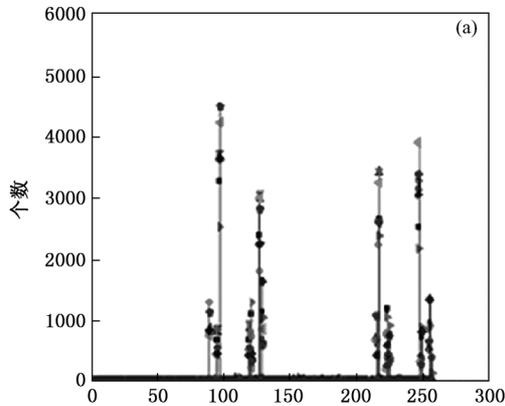


图 4 (a) 白天 (b) 黑夜的  $m$ -words 组合形式分布直方图

者熵值与代理数据熵值的趋势一致, 而代理数据保留了原数据的线性特征, 消除了非线性特性, 说明 AF 患者的 HRV 数据更多的是反应生理信号的线性特性而不是动力学的非线性特性.

为了进一步验证多尺度化的基本尺度熵, 我们利用此方法来分析白天黑夜的 HRV 序列数据. 觉醒和睡眠是生命活动必需的两个相互转化的过程, 睡眠过程的研究, 对我们了解人体生理机理有重要的意义, 还可以发现某些疾病的隐患或致病基因. 因而, 对白天黑夜 HRV 序列的研究日益引起专家学者的重视<sup>[21-23]</sup>. 本文分析了 8 个健康人 (平均年龄  $27 \pm 5$ ) 的白天黑夜数据, 分析结果如图 3 所示.

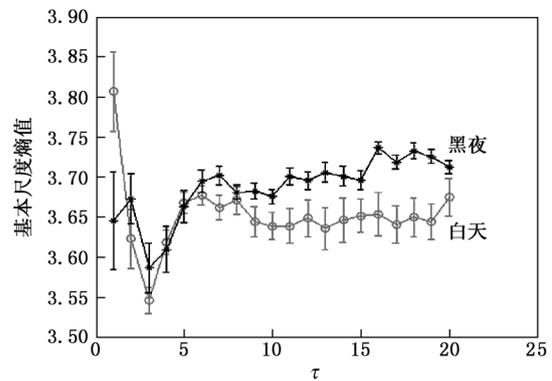


图 3 白天黑夜 HRV 的 MBE 分析数值用平均值  $\pm$  标准差表示

对于白天黑夜的 HRV 数据, 我们又分别计算了它们原始序列的  $m$ -words 组合形式. 横坐标 word index 的范围从 1 到 256 (即  $m$ -words 组合形式从 0000, 0001, ..., 3333), 纵坐标表示的是各种组合状态出现的次数. 如图 4 所示.

由图 3 可以发现, 在尺度  $\tau = 1$  时, 白天数据的熵值比黑夜的熵值要大, 对于大尺度  $\tau \geq 6$  的情况, 白天数据的熵值比黑夜数据的熵值要小, 表明白天

数据比黑夜数据表现出更高的规律性. 从图 4(a) 可以发现, 白天的数据中较高概率分布的  $m$ -words 组合形式分别为 (1200) (1332) (3120) 和 (3312), 相对应的  $x$  轴坐标为 96, 126, 216, 246. 这与李锦等人<sup>[24]</sup>的研究结果是一致的. 从图 4(b) 可以发现, 黑夜的数据中较高概率分布的  $m$ -words 组合形式相对应的  $x$  轴坐标同样也是 96, 126, 216, 246. 但黑夜的  $m$ -words 组合形式图中  $x$  轴坐标为 96 和 246 的两个值有上升的趋势, 而 126 和 216 的两个值有下降的趋势. 与白天相比较, 在夜间, 人体的基本代谢率下, 体内合成加强, 副交感神经占支配地位, 副交感神经会产生心搏变慢, 消化腺分泌增加等反应, 这些都有助于身体能量的储存, 消除疲劳, 也可恢复体力. 这些生理信息, 也许可以解释图 4 的结论, 但还有待于进一步的研究.

#### 4. 结 论

通过对 HRV 信号的研究发现, 对于多尺度化的基本尺度熵值, 当时间尺度发生变化时, 其值也发生相应变化. 在尺度  $\tau = 1$  时, CHF 患者的熵值最高, AF 患者的熵值居中, 健康人的熵值最低, 随着时间尺度  $\tau$  的增加, 最后健康人的熵值为最高, CHF 患者的熵值居中, AF 患者的熵值最低. 表明健康人

的 HRV 是小时间尺度下的特定性与大时间尺度下的丰富性相结合, 而 CHF 和 AF 患者则是在小尺度下显示出更多的不确定性, 大尺度下又缺乏足够的应变性, 显现出来心脏系统不能很好的适应外界环境的变化. 通过对 AF 患者和代理信号熵值的比较, 发现 AF 患者的 HRV 序列更多的反映的是生理信号的线性特性而非动力学信息.

利用 MBE 方法分析白天黑夜数据的结果表明, 在尺度  $\tau = 1$  时, 白天数据的熵值比黑夜的熵值要大, 对于大尺度  $\tau \geq 6$  的情况, 白天数据的熵值比黑夜数据的熵值要小, 说明白天数据比黑夜数据表现出更高的规律性. 而对于原始序列的  $m$ -words 组合形式, 两者数据中较高概率分布的  $m$ -words 组合形式相对应的  $x$  轴坐标都是 96, 126, 216, 246. 但黑夜的  $m$ -words 组合形式图中对应的  $x$  轴坐标为 96 和 246 的两个值有上升的趋势, 而 126 和 216 的两个值有下降的趋势, 对于这个结论的理论解释, 还有待于进一步研究.

本文分析了 CHF 和 AF 疾病, 以及 8 个健康人的白天黑夜数据, 并将 AF 疾病与健康人的代理数据相比较, 结果表明多尺度化的基本尺度熵可以区分不同的生理病理信号, 说明此方法对科学研究有重要的现实意义, 可以广泛的应用到其他病理信号的分析中.

- [1] Malik M, Camm A J 1990 *Clin. Cardiol.* **13** 570
- [2] Kjellgren O, Gomes J A 1993 *Am. Heart J.* **125** 204
- [3] Hon E H, Lee S T 1965 *Am. J. Obstet Gynecol.* **87** 814
- [4] Kleiger R E, Miller J P, Bigger J T, Moss A J 1987 *Am. J. Cardiol.* **59** 256
- [5] Marek M, Thomas F, A. John C 1990 *Am. J. Cardiol.* **66** 1049
- [6] Ning X B, Bian C H, Wang J, Chen Y 2006 *Chin. Sci. Bull.* **51** 385
- [7] Bian C H, Ning X B 2004 *Chin. Phys.* **13** 633
- [8] Goldberger A L, Amaral L A N, Hausdorff J M, Ivanov P C, Peng C K, Eugene Stanley H 2002 *PNAS* **99** 2466
- [9] Costa M, Goldberger A L, Peng C K 2002 *Phys. Rev. Lett.* **89** 068102
- [10] Yang A C C, Hseu S S, Yien H W, Goldberger A L, Peng C K 2003 *Phys. Rev. Lett.* **90** 108103
- [11] Bian C H, Ning X B 2004 *Physica A* **337** 149
- [12] Pincus S M 1991 *Proc Natl Acad Sci.* **88** 2297
- [13] Pincus S M, Goldberger A L 1994 *Am. J. Physiol Heart Circ Physiol* **266** H1643
- [14] Joydeep B 2000 *Acta Neurobiol Exp.* **60** 495
- [15] Pincus S, Singer BH 1996 *Proc Natl Acad Sci. USA* **93** 2083
- [16] Thuraisingham R A, Gottwald G A 2006 *Physica A* **366** 323
- [17] Huang X L, Cui S Z, Ning X B, Bian C H 2009 *Acta Phys. Sin.* **58** 8160 (in Chinese) [黄晓林, 崔胜忠, 宁新宝, 卞春华 2009 物理学报 **58** 8160]
- [18] Yang X D, Ning X B, He A J, Du S D 2008 *Acta Phys. Sin.* **57** 1514 (in Chinese) [杨小冬, 宁新宝, 何爱军, 都思丹 2008 物理学报 **57** 1514]
- [19] <http://physionet.caregroup.harvard.edu/physiobank/database/>
- [20] Huang X L, Cui S Z, Ning X B, Bian C H 2009 *Acta Phys.* **58** 8160
- [21] Ichimaru Y, Katayama S 1994 *Front MedBiol Eng* **6** 117
- [22] Unbehauen A, Mrowka R, Schwarz V, Patzak A, Lewinsohn D, Schubert E 1995 *Wien Med Wochenschr.* **145** 467
- [23] Yeragani V K, Sobolewski E, Jampala V C, Kay J, Yeragani S, Igel G 2000 *Biological Psychiatry* **41** 113
- [24] Li J, Ning X B 2007 *Physica A* **384** 423

# Multiscale base-scale entropy analysis of heart rate variability signal<sup>\*</sup>

Yan Bi-Ge Zhao Ting-Ting<sup>†</sup>

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

(Received 13 October 2010; revised manuscript received 10 March 2011)

## Abstract

Multiscale base – scale entropy is introduced in this paper. We use it to analyze heart rate variability series. The results show that multiscale base-scale entropy can identify patterns generated from healthy and pathologic states, and can distinguish daytime and nighttime heartbeat time series. We also calculate the multiscale base-scale entropy of surrogate signal (phase randomized data), compare it with the entropy of atrial fibrillation signal, and find that the trends of two entropies are similar to each other, which indicates that atrial fibrillation reflects the linear characteristics of physiological signals. Multiscale base-scale entropy method has potential applications to studying a wide variety of other physiologic and physical time series data.

**Keywords:** multiscale base-scale entropy, heart rate variability, congestive heart failure, atrial fibrillation

**PACS:** 87.85.Ng

---

<sup>\*</sup> Project supported by the Natural Science Foundation of Shaanxi Province, China (Grant No. 2010JM1007), the Fundamental Research Funds for the Central Universities, China (Grant No. GK200902053).

<sup>†</sup> Corresponding author. E-mail: 601111ting@163.com