

# 基于多尺度的人体 ECG 信号质量指数谱分析\*

杨小冬 宁新宝† 何爱军 都思丹

(南京大学电子科学与工程系, 生物医学电子工程研究所, 南京 210093)

(2007 年 6 月 20 日收到, 2007 年 7 月 8 日收到修改稿)

提出一种多重分形质量指数谱曲率和面积的概念, Cantor 集验证结果证明这两个参数在探测混沌时间序列复杂性方面是完全有效的. 引入多尺度的分析方法, 详细研究了不同采样频率和数据点下健康人、心肌缺血患者和心梗患者心电图(ECG)信号的质量指数谱, 并与其他非线性参数进行了比较, 期望从不同个体间的这种差异, 获得区分健康人及有病人的非线性特征值. 分类指标和单因素方差分析(ANOVA)检验结果表明, 该方法达到了最好的综合区分效果. 这一结论在早期诊断和临床应用方面具有一定的价值.

关键词: ECG, 多尺度, 多重分形, 质量指数谱

PACC: 0547, 0555, 8770

## 1. 引言

大量证据表明, 由复杂自调节系统产生的生理信号具有分形结构, 显示非线性动力学系统理论可以帮助理解心脏的机能. 传统的非线性动力学参数, 如关联维、李雅普诺夫指数和近似熵等已成功地应用于心电信号的研究, 证明其具有混沌特性<sup>[1-4]</sup>. 而最近广泛使用的去趋势波动分析方法<sup>[5]</sup>(DFA)则证明了 ECG 信号的长程相关性和幂指数规律, 表明其具有标度不变性即分形特性. 许多健康的生理时间序列, 是极端不平衡和非稳定的, 并且以某种不规则的复杂方式波动. 这表明, 用单一的分形指数难以准确地描述这类信号. 自 1986 年 Halsey 等人<sup>[6]</sup>提出多重分形概念以来, 多重分形便在各个领域取得了迅猛发展. 在生物学中, 对 DNA 序列的研究<sup>[7,8]</sup>、对心跳信息<sup>[9-11]</sup>和脑电信号的分析<sup>[12-13]</sup>都得出了很多重要结论.

以往的研究多集中在对多重分形奇异谱宽度和峰值位置的分析上<sup>[10,14-16]</sup>, 这种方法描述的是最大和最小子集概率测度的差异, 它的大小反映了整个分形结构概率测度分布的不均匀程度. 文献 [17,18] 使用奇异谱面积作为多重分形参数, 认为该谱曲线所形成的面积一定包含了数据序列的所有信息. 不

同于以上方法, 本文详细研究了多重分形质量指数谱  $\tau(q)$ , 提出了分别代表该谱曲线曲率和面积的新参数  $K_{\tau(q)}$  和  $\ln S_{\tau(q)}$ , 并将其应用在对人体心电信号复杂性的分析上. 实验结果表明, 这两个参数的值越大, 代表信号的非线性复杂程度越高. 我们选取不同的生成元测度, 用分形理论模型 Cantor 集证实了这一结论. 本文采集和使用的 ECG 信号为高频心电图(HFECG)信号, 其包含有 100 Hz 以上的频率分量. 该 HFECG 信号的采样频率为 1 kHz, 因此根据采样定理, 信号中 1—500 Hz 范围内的频率成分均可被分析. 一些早期的心脏疾病如心肌缺血等首先反映在心电波形的高频分量上, 往往与时域上的切迹(notches)和扭结(slurs)等相联系. 另外, 引入多尺度的分析方法对生理信号(ECG)进行时空分析, 以不同角度充分揭示该信号的非线性动力学本质, 从而找到最能表达其复杂分形结构和敏感生理信息的有效参数和频率范围. 我们认为, 多尺度多重分形(MSMF)分析的物理意义在于通过改变尺度因子  $\gamma$ , 对一个时间序列进行多尺度的粗粒化实际上就是改变该时间序列的采样频率(粗粒化采样频率), 对其在时域上进行分析, 而对粗粒化后的时间序列再进行多重分形计算则是从空间上对该序列的不同局域进行分析, 研究它的非线性参数. 通过选取适当的尺度因子, 对不同人群大量样本 ECG 信号质量指数谱

\* 国家自然科学基金(批准号: 60501003)资助的课题.

† E-mail: xbning@nju.edu.cn

和奇异谱的比较分析结果表明,参数  $K_{(q)}$  在探测生理时间序列复杂性方面具有明显的优势,在对健康人及有疾病人的区分上达到了良好的效果。

## 2. 多重分形理论

### 2.1. 质量指数和奇异谱

用尺寸为  $L$  的盒子来覆盖实验的测度,并且计算测度在每个盒中的概率,多重分形公式被发展成用来描述一些测度的统计特性<sup>[11]</sup>。定义测度值为  $T_i$  段的奇异测度(概率)为

$$P_i(L) = \frac{T_i}{\sum_{i=1}^N T_i}, \quad (1)$$

其中  $T_i$  是某测度被分成  $N$  等份的时间序列中的第  $i$  份(盒)本文以每段中电压的平均值作为该段的测度值。

由于许多实际的物理场是非均匀的、有强有弱,奇怪吸引子的点集有疏有密。在非线形深化过程中,伴随着不同量级的子串级过程。因此,对于不同区域的混沌或分形,必须用不同的标度来研究,即必须考虑概率测度的  $q$  阶矩。为此,Renyi 提出了关于表征非线性结构普遍性的 Renyi 信息:

$$I_q = \frac{1}{1-q} \log \sum_{i=1}^N P_i^q(L), \quad (2)$$

式中,  $N$  为系统可能的微观状态数目,  $q$  表示不同的标度。当  $q \rightarrow 1$  时, Renyi 信息量就简化为 Shannon 信息了。

相对于概率测度  $q$  阶矩的伸缩指数,一般维  $D_q$  定义为

$$D_q = \lim_{L \rightarrow 0} \frac{-I_q}{\log L} = \frac{1}{q-1} \lim_{L \rightarrow 0} \frac{\log \sum_{i=1}^N P_i^q(L)}{\log L}. \quad (3)$$

多分维  $D_q$  包含了分形理论所涉及的全部维数,并扩展了分形理论的内涵。实际上,  $D_q$  是通过空间各子集奇异测度的  $q$  次方及求和运算,来从整体上反映各子集奇异性程度的。因此,具有不同标度指数的子集通过  $q$  值的改变得以区分,也就把一个分形分成许多具有不同奇异程度的区域来研究,以便于分层次了解分形体的内部结构。

我们引入矩的标度指数即质量指数

$$\tau(q) = \begin{cases} (q-1)D_q, & q \neq 1, \\ 0, & q = 1. \end{cases} \quad (4)$$

(3)式的结论可以用来估计奇异谱  $f(\alpha)$ 。首先构造一个单参数的标准化测度  $\mu_i(q)$ ,其在尺寸为  $L$  盒子中的概率为

$$\mu_i(q, L) = \frac{[P_i(L)]^q}{\sum_{i=1}^N [P_i(L)]^q}. \quad (5)$$

在一般维数的定义中,参量  $q$  提供了对奇异测度不同区域的微观探测,又称权重因子。对  $q > 1$ ,  $\mu_i(q)$  放大测度的强奇异区域,对  $q < 1$ ,  $\mu_i(q)$  强调测度的弱奇异区域,而对  $q = 1$ ,  $\mu_i(1)$  测复制原来的测度。那么,  $\mu_i(q)$  理论测度支架的 Hausdorff 维数  $f(q)$  由下式给出:

$$\begin{aligned} f(q) &= - \lim_{N \rightarrow \infty} \frac{1}{\ln N} \sum_{i=1}^N \mu_i(q, L) \ln[\mu_i(q, L)] \\ &= \lim_{L \rightarrow 0} \frac{\sum_{i=1}^N \mu_i(q, L) \ln[\mu_i(q, L)]}{\ln L}. \end{aligned} \quad (6)$$

另外,关于  $\mu_i(q)$  奇异强度  $\alpha_i = \ln(P_i) / \ln L$  的平均值可由下式计算:

$$\begin{aligned} \alpha(q) &= - \lim_{N \rightarrow \infty} \frac{1}{\ln N} \sum_{i=1}^N \mu_i(q, L) \ln[P_i(L)] \\ &= \lim_{L \rightarrow 0} \frac{\sum_{i=1}^N \mu_i(q, L) \ln[P_i(L)]}{\ln L}. \end{aligned} \quad (7)$$

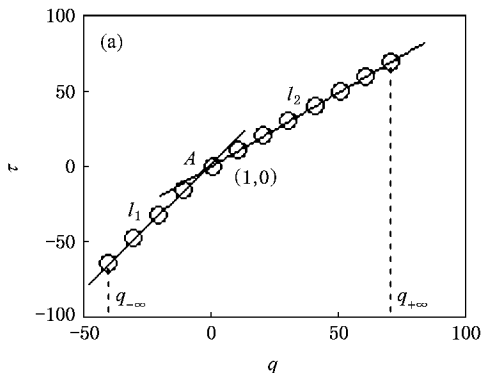
方程(6),(7)给出了作为参量  $q$  函数的 Hausdorff 维数  $f$  和平均奇异强度  $\alpha$  之间的关系。奇异谱曲线  $f(\alpha)$ - $\alpha$  在  $\alpha$  轴分布区域的大小表示多重分形的强弱,大的  $\Delta\alpha$  值与强的多重分形相联系,小的  $\Delta\alpha$  值代表弱的多重分形并趋于单重分形。

多重分形奇异谱  $f(\alpha)$  与一般维  $D_q$  通过质量指数  $\tau(q) = (q-1)D_q$  相联系,用勒让德变换表示为  $f = q\alpha - \tau$  和  $\alpha = d\tau/dq$ 。可以证明  $\lim_{q \rightarrow -\infty} D_q = \alpha_{\max}$ ,  $\lim_{q \rightarrow +\infty} D_q = \alpha_{\min}$ ,  $\alpha_{\min}$  分别是最大、最小奇异性指数。

### 2.2. 质量指数谱的曲率和面积

质量指数曲线揭示了分形程度的强弱。线性的  $\tau(q)$  曲线对应于单重分形,所有点均属于同一个分形维数,而曲线中变化的曲率则表明多重分形,或称分形测度,反映了分形结构复杂性的增加。曲线的曲率越大,代表其信号的非线性程度越强。质量指数谱的曲率和面积可以用来作为原始序列非线性复杂程度的度量。

图 1(a) 是一名 22 岁健康人  $V_4$  导联 ECG 信号的质量指数曲线谱. 从图中可看出, 该曲线由离散的数据点组成, 其疏密程度由  $q$  的取值决定, 两边分别具有不同的线性特性, 并且在点  $(1, 0)$  处形成一个较明显的拐角. 为了定性的分析该曲线, 我们以  $q = 1$  把曲线分为左右两部分, 分别对这两部分实验数据进行最小方差直线拟合, 结果相交于  $A$ , 夹角为  $\varphi$



( $\varphi \in (\pi/2, \pi]$ ) 形成的线段长度分别为  $l_1, l_2$ . 定义质量指数谱的曲率和面积

$$K_{\tau(q)} = \tan \varphi, \tag{8}$$

$$S_{\tau(q)} = \frac{1}{2} l_1 \times l_2 \times \sin \varphi. \tag{9}$$

图 1(b) 所示为该导联奇异谱  $f(\alpha)$  与奇异强度  $\alpha$  的关系曲线.

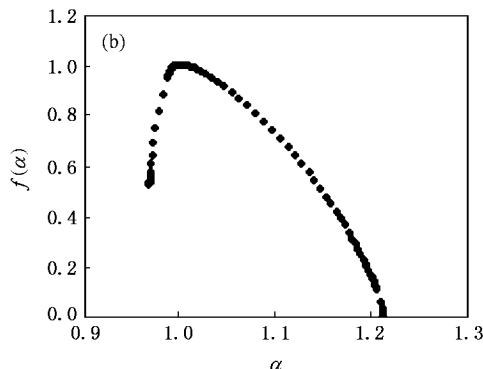


图 1 一名健康人的多重分形曲线 (a) 质量指数谱  $\tau(q)$ - $q$  在  $q = 1$  处形成明显的拐点 (b) 奇异谱  $f(\alpha)$ - $\alpha$

我们用二尺度 Cantor 集进行验证. 该集合产生如下: 首先把单位区间分成两等份, 每一份赋予不同的测度值(如  $p_1$  和  $p_2$ ), 然后把余下的每个部分再用此映射方法继续下去. 到第  $n$  次映射, 共包含  $N = a^n$  个等长小区间( $a = 2$ ), 每段长度  $L = a^{-n}$ , 概率值  $P_i(L) = p_1^{n-k} p_2^k$  ( $k = 0, \dots, n$ ). 经过无限次映射(本文取  $n = 16$ )所生成的点集就叫 Cantor 集. 显然 Cantor 集具有自相似结构, 因此它是一个分形系统. 分别选择不同的生成元测度, 验证结果如表 1 所示( $\gamma = 5$ , 见(10)式). 表 2 是三尺度 Cantor 集( $a = 3, n = 10$ )的验证结果. 结果表明, 随着生成元概率测度差异的增大、系统复杂性的增加,  $K_{\tau(q)}, \ln S_{\tau(q)}$  值也随之升高. 从而证明了这两个参数在探测该混沌时间序列复杂性方面是完全有效的.

表 1 2-Cantor set 验证结果

$p_1/p_2$	$K_{\tau(q)}$	$\ln S_{\tau(q)}$
0.5/0.5	$-1.72 \times 10^{-7}$	-6.77
0.4/0.6	-0.23	7.37
0.3/0.7	-0.53	8.18
0.2/0.8	-0.87	8.45
0.1/0.9	-1.29	8.57

表 2 3-Cantor set 验证结果

$p_1/p_2/p_3$	$K_{\tau(q)}$	$\ln S_{\tau(q)}$
0.35/0.3/0.35	-0.035	5.48
0.3/0.4/0.3	-0.092	6.50
0.4/0.2/0.4	-0.24	7.35
0.2/0.6/0.2	-0.55	8.28
0.1/0.8/0.1	-1.28	8.60

本文用  $K_{\tau(q)}, \ln S_{\tau(q)}$  表征人体心脏电活动的非线性动力学特征参数, 期望从不同个体间的这种差异, 获得区分健康人和有疾病人的非线性特征值.

### 3. 实 验

#### 3.1. 实验对象

本实验的 ECG 数据全部来源于第二军医大学长征医院南京分院心脏内科. 实验对象共分 A, B, C 三组: A 组为健康人组, 共 40 例(平均年龄  $\pm$  标准差  $28.68 \pm 16.31$  岁, 其中老年人 6 例, 血压正常, 无心脏病史; 年轻人 34 例, 窦性心率.); B 组为心肌缺血(ischemia)患者组, 共 24 例(平均年龄  $\pm$  标准差,  $51.67 \pm 18.86$  岁); C 组为心梗(MI)患者组, 共 18 例(平均年龄  $\pm$  标准差  $62.22 \pm 12.37$  岁).

#### 3.2. 实验装置和数据处理

采用南京大学电子科学与工程系生物医学电子

工程研究所自主研制的 NHE-2000 型多功能心电工作站,该检测系统频率响应宽、灵敏度高,可检测显示 100 Hz 以上的高频成分(100 Hz—2 kHz)。A/D 转换位数 12 bit,一个数据用 2 byte 来表示(高 4 位置 0)。受试者保持平躺休息状态,在安静条件下几分钟后开始做同步十二导联常规心电图信号采集。体表的 ECG 信号经放大处理后送入机内专用 A/D 采集卡,每个通道采样频率为 1 kHz,时间长度约为 1 min,所采数据共分 12 组(导联,每个导联采 59797 个点)。采集后的心电数据在实际计算前先经过小波滤波,用的是 bior6.8 小波基函数,滤去呼吸波(低频)、50 Hz 工频干扰噪声(中频)和肌电干扰(高频)。另外,为了消除数据刚开始采集时的不稳定状态,去掉每组数据前面 2 000 个点,并把剩下的点按尺寸  $L = 2$  划分成  $N$  个区间( $1/N \rightarrow 0$ )。设置参数  $q$  为从  $-\infty$  到  $+\infty$  取值,其步长根据点的疏密而变。

## 4. 结果与分析

### 4.1. 多尺度分析

为了寻找最适当的采样频率,我们用多个尺度因子来研究原始序列的不同时间段。对原始时间序列  $x = \{x_1, \dots, x_i, \dots, x_N\}$  构造它的粗粒化时间序列,即  $\{y^{(\gamma)}\}$ ,

$$y_j^{(\gamma)} = \frac{1}{\gamma} \sum_{i=(j-1)\gamma+1}^{j\gamma} x_i, \quad 1 \leq j \leq N/\gamma, \quad (10)$$

其中  $\gamma$  是尺度因子。当  $\gamma = 1$  时,时间序列  $y^{(1)}$  即为

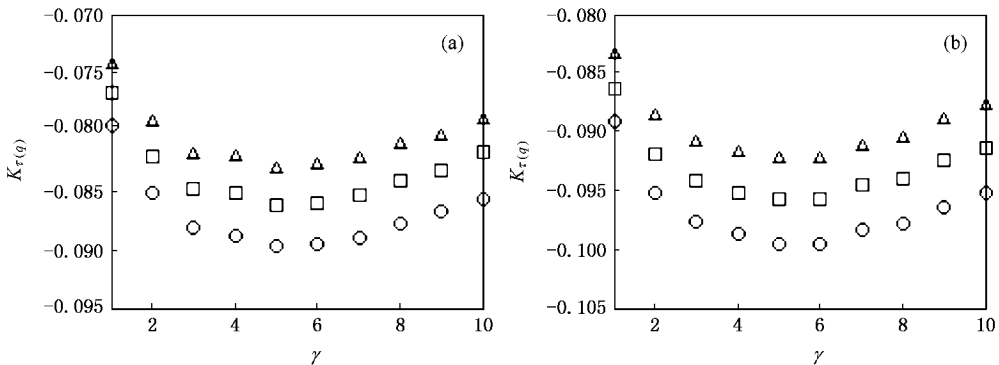


图2 质量指数谱曲率  $K_{\tau(q)}$  (均值)随尺度因子  $\gamma$  变化曲线(该曲线在  $\gamma = 5$  处形成拐点;不同的符号代表数据长度,  $\circ$  为 30240 点,  $\square$  为 40320 点,  $\triangle$  为 55440 点) (a) 心肌梗塞患者 (b) 健康人

### 4.2. 实验结果

我们总共计算了 82 个样本 984 个导联 ECG 信

原始时间序列  $x$ 。

我们分别计算了三种不同长度的时间序列,即 30240、40320 和 55440 数据点(采样频率均为 1 kHz),并把它们粗粒化到最高尺度 10,然后对每一个粗粒化的时间序列计算其质量指数曲线谱,结果如图 2 所示(以心肌缺血患者和健康人为例)。图中显示  $K_{\tau(q)}$  最大值(绝对值)出现的位置与序列长度无关,而仅与尺度因子有关。对每一个尺度因子而言,其绝对值随序列长度的增加而减小。另外,我们还发现  $K_{\tau(q)}$  对每个尺度因子并不相同,而是有一个空间分布。当尺度因子  $\gamma$  集中在 4—6 之间时,该值达到最大,并且一般在  $\gamma = 5$  处形成拐点。经实验发现,当采样频率增加  $n$  倍即  $n$  kHz 时,则尺度因子集中在  $4n-6n$  之间时  $K_{\tau(q)}$  (绝对值)达到最大,相应的拐点位置出现在  $\gamma = 5n$  处。分析表明,对 ECG 信号最敏感的生理信息和复杂的分形结构可能集中在粗粒化频率范围 250—166 Hz(信号频率 125—83 Hz)之间,其中在 200 Hz(信号频率 100 Hz)处或许包含了某些特殊的生命活动规律和实质。心室左右束支传导阻塞、心肌缺血、心肌损伤等早期的心脏疾病首先反映在心电波形的高频分量上,往往与时域上的切迹和扭结等高频小波形式相联系,并且大于 100 Hz。我们认为,多尺度分析的意义就在于,通过构造原始生理时间序列的粗粒化时间序列,寻找能够表达其复杂动力学特征的最有效参数,从而发现一些深层次规律。因此,在下面的分析中,我们采用尺度因子  $\gamma$  (1 kHz 采样频率)来计算多重分形质量指数谱。

号质量指数谱的曲率和面积((8)(9)式,  $\gamma = 5$ ),发现每个个体、不同导联都存在差异,说明具有不同的非线性动力学特性(一般情况下,胸导联计算结果要

大于肢体导联).表 3—表 6 显示不同人群各导联 ECG 信号  $K_{\zeta(q)}$ ,  $\ln S_{\zeta(q)}$  的统计均值和标准差以及每个导联的  $t$  检验分析结果(表中所示分别为健康人/

心肌缺血患者、心肌缺血/心梗患者).结果表明,健康人组各导联多重分形质量指数谱的曲率和面积均明显大于其他两组,而心梗患者组各导联值均最小.

表 3 同步十二导联心电图信号多重分形质量指数谱曲率的对数( $\times 10^{-1}$ ) (均值  $\pm$  标准差(肢体导联))

导联	I	II	III	AVR	AVL	AVF
健康人	$-0.83 \pm 0.15$	$-1.03 \pm 0.13$	$-0.88 \pm 0.19$	$-0.77 \pm 0.08$	$-0.72 \pm 0.14$	$-0.93 \pm 0.16$
心肌缺血	$-0.73 \pm 0.18$	$-0.83 \pm 0.19$	$-0.65 \pm 0.15$	$-0.67 \pm 0.10$	$-0.56 \pm 0.19$	$-0.70 \pm 0.20$
心梗	$-0.57 \pm 0.18$	$-0.58 \pm 0.22$	$-0.55 \pm 0.14$	$-0.48 \pm 0.09$	$-0.49 \pm 0.16$	$-0.57 \pm 0.15$
$P$	0.01518	0.00000	0.00000	0.00006	0.00045	0.00000
	0.00865	0.00038	0.03444	0.00000	0.21190	0.02939

表 4 同步十二导联心电图信号多重分形质量指数谱曲率的对数( $\times 10^{-1}$ ) (均值  $\pm$  标准差(胸导联))

导联	$V_1$	$V_2$	$V_3$	$V_4$	$V_5$	$V_6$
健康人	$-0.91 \pm 0.16$	$-1.04 \pm 0.13$	$-1.16 \pm 0.10$	$-1.08 \pm 0.11$	$-1.18 \pm 0.14$	$-1.08 \pm 0.10$
心肌缺血	$-0.78 \pm 0.25$	$-0.98 \pm 0.15$	$-1.02 \pm 0.18$	$-0.96 \pm 0.17$	$-1.04 \pm 0.15$	$-0.99 \pm 0.16$
心梗	$-0.76 \pm 0.19$	$-0.91 \pm 0.26$	$-0.74 \pm 0.19$	$-0.97 \pm 0.13$	$-0.82 \pm 0.17$	$-0.72 \pm 0.19$
$P$	0.01744	0.08818	0.00035	0.00113	0.00072	0.01008
	0.74677	0.25153	0.00002	0.80264	0.00007	0.00000

表 5 同步十二导联心电图信号多重分形质量指数谱面积的对数(均值  $\pm$  标准差(肢体导联))

导联	I	II	III	AVR	AVL	AVF
健康人	$6.32 \pm 0.17$	$6.56 \pm 0.14$	$6.37 \pm 0.24$	$6.23 \pm 0.11$	$6.17 \pm 0.20$	$6.45 \pm 0.20$
心肌缺血	$6.20 \pm 0.26$	$6.34 \pm 0.26$	$6.08 \pm 0.26$	$6.09 \pm 0.17$	$5.90 \pm 0.39$	$6.15 \pm 0.34$
心梗	$5.91 \pm 0.33$	$5.81 \pm 0.85$	$5.87 \pm 0.31$	$5.75 \pm 0.19$	$5.75 \pm 0.42$	$5.91 \pm 0.30$
$P$	0.02695	0.00007	0.00003	0.00023	0.00052	0.00003
	0.00281	0.00616	0.01983	0.00000	0.23710	0.02685

表 6 同步十二导联心电图信号多重分形质量指数谱面积的对数(均值  $\pm$  标准差(胸导联))

导联	$V_1$	$V_2$	$V_3$	$V_4$	$V_5$	$V_6$
健康人	$6.37 \pm 0.17$	$6.54 \pm 0.13$	$6.68 \pm 0.10$	$6.55 \pm 0.10$	$6.69 \pm 0.13$	$6.62 \pm 0.11$
心肌缺血	$6.20 \pm 0.34$	$6.48 \pm 0.16$	$6.55 \pm 0.22$	$6.44 \pm 0.17$	$6.57 \pm 0.18$	$6.53 \pm 0.19$
心梗	$6.16 \pm 0.31$	$6.11 \pm 1.35$	$6.17 \pm 0.29$	$6.42 \pm 0.16$	$6.27 \pm 0.23$	$6.15 \pm 0.31$
$P$	0.01263	0.11147	0.00234	0.00124	0.00318	0.02630
	0.68013	0.18164	0.00002	0.76110	0.00002	0.00002

进一步对这两个参数进行统计分析,我们先计算 A, B, C 三组中每个个体 ECG 信号十二个导联的

算术均值和标准差(图 3),然后分别再在相同人群(组)内对该十二导联均值统计平均,其结果如表 7

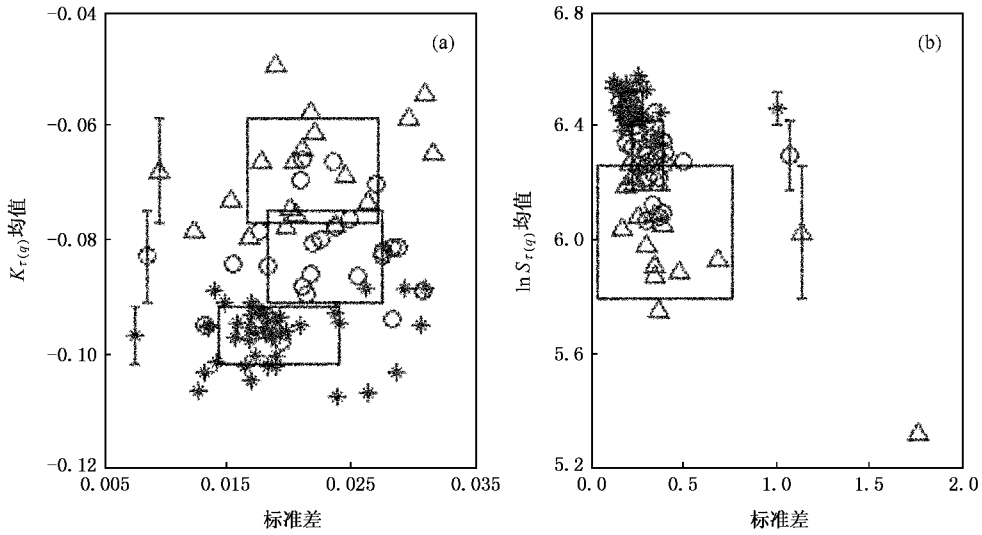


图 3 三组人群中不同个体 ECG 信号质量指数谱曲率和面积的十二导联算术均值和标准差(符号代表不同的人群,\*为健康人,○为心肌缺血患者,△为心梗患者.矩形定义为各自的组内均值和标准差) (a) $K_{\tau(q)}$  (b) $\ln S_{\tau(q)}$

表 7 ECG 质量指数谱曲率和面积(均值±标准差)

参数	健康人(A组)	心肌缺血(B组)	心梗(C组)
$K_{\tau(q)}$	$-0.097 \pm 0.005$	$-0.083 \pm 0.008$	$-0.068 \pm 0.009$
$\ln S_{\tau(q)}$	$6.461 \pm 0.056$	$6.294 \pm 0.119$	$6.023 \pm 0.233$
$\Delta\alpha$	$0.222 \pm 0.012$	$0.189 \pm 0.019$	$0.160 \pm 0.019$
$\ln S_{\chi(\alpha)}$	$-1.663 \pm 0.105$	$-1.898 \pm 0.163$	$-2.193 \pm 0.383$

从上表可以看出,健康人 ECG 信号质量指数谱的  $K_{\tau(q)}$ (绝对值)、 $\ln S_{\tau(q)}$ 均大于其他两类人群,代表其曲线的弯曲程度最大.心肌缺血患者次之,而心梗患者  $\tau(q)$ 曲线的曲率明显小于另两类人群,并接近直线.这表明,健康人心脏处于最佳生理状态,其 ECG 信号吸引子的非线性动力学复杂程度最高.心梗(MI)患者由于冠状动脉血管狭窄或堵塞,严重的心肌缺血可能导致相应组织的类分形结构受损或者发生结构变化<sup>[11]</sup>,这可能是系统的复杂性降低、质量指数曲线趋于线性的原因.图 3(a)显示,我们的方法完全可以把健康人与另外两组区分开,彼此没有任何重叠.在对心肌缺血与心梗患者组的区分上,比较其他几种参数, $K_{\tau(q)}$ 也达到了最好的效果(对它们的有效区分率分别为 87.50%、88.89%,而  $\Delta\alpha$  则分别为 78.04%、77.81%).

### 4.3. 几种方法的比较

我们用两个参数对表 7 中几种方法的统计结果

所示.表中同时列出了对应的奇异谱宽度  $\Delta\alpha$  和面积  $\ln S_{\chi(\alpha)}$ 的计算结果( $\gamma = 5$ )进行比较.

量化分析.首先考虑两组数据间的比率因子  $\eta$ <sup>[9]</sup>,

$$\eta = \frac{(\mu_H - \mu_{Isc})^2}{\sigma_H^2 + \sigma_{Isc}^2}, \quad (11)$$

其中  $\mu_H, \mu_{Isc}$  分别是健康人和心肌缺血患者组各参数的统计均值,  $\sigma_H, \sigma_{Isc}$  为其各自的标准差.对一个良好的分类,组间数据均值的差异应远大于它们的标准差,即  $\eta \gg 1$ .其次,我们计算达到两组数据最小不正确分类状态下边界均值的统计距离  $d^2$ ,

$$d^2 = \left( \frac{\mu_H - \frac{\sigma_H \times \mu_{Isc} + \sigma_{Isc} \times \mu_H}{\sigma_H + \sigma_{Isc}}}{\sigma_H} \right)^2. \quad (12)$$

表 8 列出了几种方法对组间数据不同的  $\eta$  和  $d^2$  值.这两个值越大,表明该方法的分类效果越明显.表中同时列出了不同方法对三组数据的 ANOVA 检验分析结果,其  $P$  值接近零表示至少有一组样本均值与其他两组的样本均值有显著不同.从该结果可以看出,参数  $K_{\tau(q)}$  达到了最好的综合区分效果.

表 8 几种方法分类结果的比较

参数	$\eta_{H,Isc}$	$d_{H,Isc}^2$	$\eta_{Isc,MI}$	$d_{Isc,MI}^2$	$F(ANOVA)$
$K_{\tau(q)}$	2.21	1.16	1.55	0.78	0
$K_{\tau(q)}(\gamma = 1)$	1.91	1.00	1.52	0.76	0
$\ln S_{\tau(q)}$	1.60	0.90	1.08	0.59	0
$\Delta\alpha$	2.23	1.17	1.21	0.61	0
$\ln S_{\chi(\alpha)}$	1.48	0.77	0.50	0.29	$1.04 \times 10^{-12}$

## 5. 结 论

本文以非线性分析手段,研究了不同人群大量样本 ECG 信号的多重分形质量指数谱,得出以下结论:

1. 人体心电图(ECG)信号具有稳定的多重分形特性,同一个体的不同导联具有不同的谱曲率和面积,呈现出一种离散分布的特性。 $K_{\alpha(q)}$ ,  $\ln S_{\alpha(q)}$  这两个参数值越大,代表信号的非线性复杂程度越高。

2.  $K_{\alpha(q)}$  最大值总是出现在粗粒化频率 200 Hz (信号频率 100 Hz) 处,表明在该频率点上 ECG 信号的非线性特性最敏感. 该位置与数据序列长度无关,

而仅与尺度因子有关. 对每一个尺度因子而言,其绝对值随序列长度的增加而减小. 另外,  $K_{\alpha(q)}$  对每个尺度因子并不相同,而是有一个空间分布. 对于采样频率为 1 kHz 的 ECG 时间序列,当尺度因子集中在 4—6 之间时,该值达到最大并且一般在  $\gamma = 5$  处形成拐点. 如果采样频率增加到  $n$  kHz 时,则相应的拐点位置出现在  $\gamma = 5n$  处.

3. 健康人质量指数谱的曲率(绝对值)和面积明显大于其他两类人群,表明健康人 ECG 信号具有最强的非线性动力学特性. 而心梗患者的这两个值最小. 分类指标  $\eta$ ,  $d^2$  和 ANOVA 检验分析结果表明,参数  $K_{\alpha(q)}$  ( $\gamma = 5$ ) 在区分健康人、心肌缺血患者和心梗患者方面达到了最好的效果.

- [ 1 ] Wang Z Z , Ning X B , Zhang Y , Du G H 2000 *Chin . Sci . Bull .* **45** 1628
- [ 2 ] Wang Z Z , Li Z , Wei Y X , Ning X B , Lin Y Z 2002 *Chin . Sci . Bull .* **47** 1845
- [ 3 ] Bian C H , Ning X B 2004 *Chin . Phys .* **13** 633
- [ 4 ] Richman J S , Moorman J R 2000 *Am . J . Physiol .* **278** H2039
- [ 5 ] Peng C K , Buldyrev S V , Goldberger A L , Havlin S , Sciortino F , Simons M , Stanley H E 1992 *Nature* **356** 168
- [ 6 ] Halsey T C , Jensen M H , Kadanoff L P , Procaccia I , Shraiman B I 1986 *Phys . Rev . A* **33** 1141
- [ 7 ] Buldyrev S V , Goldberger A L , Havlin S , Mantegna R N , Matsa M E , Peng C K , Simons M , Stanley H E 1995 *Phys . Rev . E* **51** 5084
- [ 8 ] Grosse I , Herzel H , Buldyrev S V , Stanley H E 2000 *Phys . Rev . E* **61** 5624
- [ 9 ] Amaral L A N , Goldberger A L , Ivanov P C , Stanley H E 1998 *Phys . Rev . Lett .* **81** 2388
- [ 10 ] Amaral L A N , Ivanov P Ch , Aoyagi N , Hidaka I , Tomono S , Goldberger A L , Stanley H E , Yamamoto Y 2001 *Phys . Rev . Lett .* **86** 6026
- [ 11 ] Chhabra A , Jensen R V 1989 *Phys . Rev . Lett .* **62** 1327
- [ 12 ] Popivanov D , Jivkova S , Stomonyakov V , Nicolova G 2005 *Signal Process .* **85** 2112
- [ 13 ] Ma Q L , Ning X B , Wang J , Bian C H 2006 *Chin . Sci . Bull .* **51** 3059
- [ 14 ] Sun X , Xiong G , Fu Z X , Wu Z Q 2000 *Acta Phys . Sin .* **49** 854 ( in Chinese ) [ 孙 霞、熊 刚、傅竹西、吴自勤 2000 物理学报 **49** 854 ]
- [ 15 ] Sun X , Wu Z Q 2001 *Acta Phys . Sin .* **50** 2126 ( in Chinese ) [ 孙 霞、吴自勤 2001 物理学报 **50** 2126 ]
- [ 16 ] Shao Y Z , Zhong W R , Ren S , Cai Zh S , Gong L 2005 *Acta Phys . Sin .* **54** 3290 ( in Chinese ) [ 邵元智、钟伟荣、任 山、蔡志芬、龚 雷 2005 物理学报 **54** 3290 ]
- [ 17 ] Wang J , Ning X B , Chen Y 2003 *Physica A* **323** 561
- [ 18 ] Wang J , Ning X B , Chen Y 2003 *Physica A* **325** 485

# Mass exponent spectrum analysis of human ECG signals based on multiple scale factors<sup>\*</sup>

Yang Xiao-Dong Ning Xin-Bao<sup>†</sup> He Ai-Jun Du Si-Dan

( *Department of Electronic Science & Engineering, Institute for Biomedical Electronic Engineering, Nanjing University, Nanjing 210093, China* )

( Received 20 June 2007 ; revised manuscript received 8 July 2007 )

## Abstract

Life is one of the most complex nonlinear systems and heart is the core of this lifecycle system. The complexity of electrocardiogram ( ECG ) signals may reflect the physiologic function and health status of the heart. In this paper, we introduced two novel parameters of the multifractal mass exponent spectrum curvature and area. The evaluation of Cantor set validated that the two indicators are entirely effective in exploring the complexity of chaotic time series. Using the multiscale analysis method, we studied the mass exponent spectra of ECG signals taken from the cohorts of healthy, ischemia and myocardial infarction ( MI ) sufferer under different sampling frequencies and data lengths. Then we compared these new indicators with other nonlinear parameters and also expected to acquire some valuable nonlinear eigenvalues to distinguish the healthy from the heart diseased through those individual discrepancies. The classification indexes and ANOVA testing results both indicated that our method could achieve better results. These conclusions may be of much value in early diagnoses and clinical applications.

**Keywords** : ECG , multiple scale factors , multifractal , mass exponent spectrum

**PACC** : 0547 , 0555 , 8770

---

<sup>\*</sup> Project supported by the National Natural Science Foundation of China ( Grant No. 60501003 ).

<sup>†</sup> E-mail : xbning@nju.edu.cn