基于两参数熵的广义相位排列熵算法

康欢, 张小凤*, 张光斌

陕西师范大学物理学与信息技术学院, 西安 710119

* 通信作者. E-mail: xiaofengzhang71@snnu.edu.cn

收稿日期: 2018-08-06; 修回日期: 2018-09-25; 接受日期: 2018-12-13; 网络出版日期: 2019-08-30

国家自然科学基金 (批准号: 11574191, 11674208) 资助项目

摘要 广义排列熵 ($PE_{q,\delta}$) 取恰当参数时能放大系统的微小变化, 而信号的相位相比于幅度包含更多关键信息. 本文将相位信息引入到广义排列熵中, 提出了广义相位排列熵 ($PPE_{q,\delta}$) 算法, 验证了广义相位排列熵在检测系统动态改变方面的优势, 分析了参数 q, δ 选取对于系统动态改变的影响, 探究了数据长度和噪声对于广义相位排列熵的影响. 最后, 将广义相位排列熵算法运用到异常心电信号检测分析中. 仿真结果表明, 当广义排列熵和广义相位排列熵的 q, δ 取值相同时, 广义相位排列熵对于同一动态变化的检测效果更为显著. 无论是针对 logistic 映射还是异常心电信号的动态变化检测, 当 $q > 0$, $\delta > 0$ 时, 减小 q 值或者增大 δ 值, 或者同时减小 q 值增大 δ 值, 都能提高广义排列熵和广义相位排列熵对于动态改变的检测效果. 此外, 数据长度改变对于广义排列熵和广义相位排列熵基本无影响且两者对噪声均具有良好的鲁棒性.

关键词 广义相位排列熵, 动态改变检测, logistic 映射, 心电信号, 非线性分析

1 引言

随着非线性动力学理论的发展, 众多非线性分析方法相继被提出. 熵分析作为分析系统动力学特征的非线性方法, 被广泛运用于描述系统复杂性及评估系统的动力学特征.

1948 年, 香农 (Shannon) 首次将熵这一概念引入到信息论中, 使通信技术取得跨越式发展, 同时也为后来研究者将熵这一概念引入到信号处理中做了铺垫. 为解决系统复杂性的测度, Pincus^[1] 于 1991 年提出近似熵的概念以表征系统复杂性, 近似熵值越大, 表示系统生成新模式的几率越大, 则系统复杂性与混沌程度越高. 2000 年, Richman 等^[2] 学者针对近似熵的不足, 提出了通过剔除自匹配来减小近似熵中的偏差从而使结果更加准确的样本熵. 2002 年, Costa 等^[3,4] 提出了多尺度熵, 用于研究 HRV (heart rate variability) 信号的复杂性. 该方法通过对原始序列进行时间尺度下的“粗粒化”处理, 然后计算其样本熵, 以期待获得信号在不同尺度下的复杂性, 熵值越大意味着时间序列在该尺度

引用格式: 康欢, 张小凤, 张光斌. 基于两参数熵的广义相位排列熵算法. 中国科学: 信息科学, 2019, doi: 10.1360/N112018-00197
Kang H, Zhang X F, Zhang G B. Generalized phase permutation entropy algorithm based on the two-index entropy (in Chinese). Sci Sin Inform, 2019, doi: 10.1360/N112018-00197

下的信号不规则性越强. 同年, Bandt 等^[5] 基于比较时间序列相邻值的大小提出了排列熵. 由于排列熵是对时间序列按照升序排列后, 取升序排列序列在原相空间重构序列中的位置索引进行概率求解, 因此排列熵算法具有很好的抗噪能力, 同时算法简单、容易理解, 已被广泛运用于生理电信号处理^[6]、股票数据分析^[7]、工程异常信号分析等^[8]. 为了进一步优化排列熵算法在不同领域中的应用, 各种改进算法被提出, 如 Renyi 排列熵^[9,10]、加权排列熵^[11,12]、多尺度排列熵^[13,14]、广义排列熵^[15] 等. 其中广义排列熵 ($PE_{q,\delta}$) 是将两参数熵公式 $S_{q,\delta}$ ^[16] 引入到排列熵中, 在 q, δ 取恰当值时, 广义排列熵能放大系统中的微小变化^[15]. 然而, 相位同步和幅度关联是功能独立的现象^[17], 且相位相比于幅度包含更多关键信息^[18,19]. 考虑到广义排列熵算法的优点和信号的相位包含更多关键信息的特点, 本文在广义排列熵中引入相位信息提出广义相位排列熵 ($PPE_{q,\delta}$), 分析广义相位排列熵在系统动态检测方面的性能、熵参数 q, δ 改变对广义相位排列熵检测系统动态变化的影响和不同数据长度与信噪比 (signal-to-noise ration, SNR) 时的广义相位排列熵熵值变化情况. 最后, 将广义相位排列熵运用于异常心电信号分析.

2 广义相位排列熵

设一时间序列 $\{X(i), i = 1, 2, \dots, N\}$, 利用 Hilbert 变换求其解析序列 $Z(i)$, 即

$$\begin{aligned} Z(i) &= X(i) + jH[X(i)] = A(i)e^{j\Theta(i)}, \\ H[X(i)] &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{X(i)}{i-t} dt, \end{aligned} \quad (1)$$

则 $\{X(i), i = 1, 2, \dots, N\}$ 的相位序列为

$$\left\{ \Theta(i) = \arctan \frac{H[X(i)]}{X(i)}, i = 1, 2, \dots, N \right\}. \quad (2)$$

接下来, 依据广义排列熵的基本原理^[15], 利用信号的相位信息, 对相位序列进行相空间重构得重构矩阵如下:

$$\begin{bmatrix} \Theta(1) & \Theta(1+\tau) & \cdots & \Theta(1+(m-1)\tau) \\ \Theta(2) & \Theta(2+\tau) & \cdots & \Theta(2+(m-1)\tau) \\ & & \vdots & \\ \Theta(j) & \Theta(j+\tau) & \cdots & \Theta(j+(m-1)\tau) \\ & & \vdots & \\ \Theta(K) & \Theta(K+\tau) & \cdots & \Theta(K+(m-1)\tau) \end{bmatrix}, \quad j = 1, 2, \dots, K, \quad (3)$$

式中, m 和 τ 代表相空间重构的嵌入维数和时间延迟, 矩阵中每一行可以看作一个重构分量记为 $S(j) = \{\Theta(j), \Theta(j+\tau), \dots, \Theta(j+(m-1)\tau)\}$, $j = 1, 2, \dots, K$, 其中, $K = N - (m-1)\tau$ 为重构分量的总个数.

对每个重构分量 $S(j)$ 按照数值大小进行升序排列, 提取升序排列后每个元素在原重构序列 $S(j)$ 中的位置索引并将该位置索引记为 l , 从而得到相应的位置序列, 记为

$$M(j) = \{l_1, l_2, \dots, l_m\}, \quad j = 1, 2, \dots, N - (m-1)\tau. \quad (4)$$

计算相同位置序列出现的概率, 即

$$P_v = \frac{\text{Num}_v}{N - (m-1)\tau}, \quad 1 \leq v \leq N - (m-1)\tau, \quad (5)$$

式中, Num_v 表示相同位置序列的个数, $N - (m-1)\tau$ 表示位置序列总个数. 利用 Ribeiro 等^[16] 提出的两参数熵 $S_{q,\delta}$ 计算熵值并定义该熵值为广义相位排列熵 ($\text{PPE}_{q,\delta}$), 公式为

$$S_{q,\delta} = \sum_{v=1} P_v \left[\ln_q \left(\frac{1}{P_v} \right) \right]^\delta, \quad q \in \mathbb{R}, \quad \delta > 0, \quad (6)$$

式中, 当 $q \neq 1$ 时, $\ln_q(\frac{1}{P_v}) = \frac{(\frac{1}{P_v})^{1-q} - 1}{1-q}$; 当 $q = 1$ 时, $\ln_q(\frac{1}{P_v}) = \ln \frac{1}{P_v}$. 当 $S_{q,\delta}$ 中 $\delta = 1$ 时, $S_{q,1}$ 为 Tsallis 熵 (Tsallis entropy, S_q) 的表达式; 当 $S_{q,\delta}$ 中 $q = 1$ 时, $S_{1,\delta}$ 被称为广义黑洞熵 (generalized black-hole entropy, S_δ); 当 $S_{q,\delta}$ 中 $q = 1$ 且 $\delta = 1$ 时, $S_{1,1}$ 则转化为香农熵.

对式 (6) 中 $S_{q,\delta}$ 做微小改变, 即

$$S_{q,\delta} = \sum_{v=1} P_v \left\{ \text{sign} \left[\ln_q \left(\frac{1}{P_v} \right) \right] \left| \ln_q \left(\frac{1}{P_v} \right) \right|^\delta \right\}, \quad q \in \mathbb{R}, \quad \delta \in \mathbb{R}, \quad (7)$$

这样, q, δ 属于实数集, 有利于全面分析改变参数 q, δ 对系统动态变化检测的影响^[15].

3 广义相位排列熵性能分析

为了分析广义相位排列熵在信号动态变化检测方面的性能, 运用非线性系统分析中常用的 logistic 映射, 对广义相位排列熵算法的性能进行仿真研究. logistic 映射的表达式为

$$x_{i+1}(n) = r(n)x_i(n)[1 - x_i(n)], \quad (8)$$

其中, $r(n)$ 表示控制参数, 一般的取值范围为 $0 < r(n) \leq 4$, 本文中 $r(n)$ 的变化区间为 (3.5, 4.0). logistic 序列初值为 $x_0(n) = 0.65$, $r(0) = 3.5$. 为了与广义排列熵对 logistic 映射动态变化检测作对比, 此处各参数取值与文献 [15] 完全一致, 即取连续 $r(n)$ 的递增步长为 0.0005, 对于每个 $r(n)$, 序列 $x(n)$ 的长度为 5000, 取后 3000 个数据进行熵计算, 序列相空间重构的嵌入维数取 $m = 3$, 时间延迟取 $\tau = 1$, 仿真结果如图 1 所示. 图 1(a) 和 (b) 分别为 $m = 3, \tau = 1, q, \delta$ 取不同值时计算的广义排列熵 ($\text{PE}_{q,\delta}$) 和广义相位排列熵 ($\text{PPE}_{q,\delta}$). 对于图 1(a), 在 $q \neq 1, \delta \neq 1$ 时, $\text{PE}_{q,\delta}$ 在 logistic 映射空白窗处的熵值下降比 $q = 1, \delta = 1$ 时显著, 由此得出与文献 [15] 相一致的结论, 即广义排列熵在 q, δ 取恰当值时能放大系统的微小变化. 对于图 1(b), 在 $q \neq 1, \delta \neq 1$ 时, $\text{PPE}_{q,\delta}$ 在 logistic 映射空白窗处的熵值下降比 $q = 1, \delta = 1$ 时更为显著, 同样得出广义相位排列熵在 q, δ 取恰当值时也能放大系统的微小变化. 对比图 1(a) 和 (b) 可以发现, 在 q, δ 取值相同时, $\text{PPE}_{q,\delta}$ 比 $\text{PE}_{q,\delta}$ 在同一空白窗处的熵值下降更多, 即在同等条件下, 广义相位排列熵相比于广义排列熵, 能更为有效地检测出系统的动态变化.

为了进一步研究 q, δ 的影响, 取 $-2 < q \leq 2, \delta = 1$ 及 $q = 1, -2 < \delta \leq 2$, 分别分析 q, δ 在取不同值时, 广义相位排列熵对于 logistic 映射动态变化的检测结果. 考虑到研究 q, δ 改变对广义相位排列熵检测系统动态变化的计算量相对较大, 仿真中 $r(n)$ 的递增步长为 0.001, 对于每个 $r(n)$, 序列 $x(n)$ 的长度为 2500, 取 $x(n)$ 的后 1000 个数据进行熵计算. 一般地, 在仿真中嵌入维数可取 $m = 3, 4, \dots, 7$ ^[5]. 但是, 若取 $m = 3$, 得到的位置序列可能的种类有 $3! = 6$, 则位置序列种类数目相对较少不利于对信号

图 1 (网络版彩图) 广义排列熵与广义相位排列熵对 logistic 映射动态变化检测的对比

Figure 1 (Color online) Comparison of two kinds of entropy for the logistic map dynamic changes detection when $m = 3$, $\tau = 1$. (a) Generalized permutation entropy ($PE_{q,\delta}$); (b) generalized phase permutation entropy ($PPE_{q,\delta}$)

 图 2 (网络版彩图) $m = 5$, $\tau = 1$, $0 < q \leq 2$, $\delta = 1$ 时, $PE_{q,\delta}$ 与 $PPE_{q,\delta}$ 对于 logistic 映射的动态变化检测

Figure 2 (Color online) Detecting dynamic changes in logistic map using $PE_{q,\delta}$ and $PPE_{q,\delta}$ when $m = 5$, $\tau = 1$, $0 < q \leq 2$, $\delta = 1$. (a) $PE_{q,1}$ when $0 < q \leq 2$; (b) $PPE_{q,1}$ when $0 < q \leq 2$

进行全面分析; 相反, 若取 $m = 7$, 虽然能使熵值变化更精确, 但是会大大增加计算量. 因此在本文之后的仿真中, 均取相空间重构嵌入维数 $m = 5$. 同时, 相空间重构中时间延迟 τ 的取值越大, 位置序列总个数 $N - (m - 1)\tau$ 越小, 则求解概率时包含原始序列信息越少, 所以文中所有处均取相空间重构时间延迟 $\tau = 1$.

图 2 为当 $m = 5$, $\tau = 1$, $0 < q \leq 2$, $\delta = 1$ 时, 以步长 $\Delta q = 0.1$, 改变熵参数 q 值, 广义排列熵 ($PE_{q,\delta}$) 和广义相位排列熵 ($PPE_{q,\delta}$) 对于 logistic 映射动态变化的检测结果. 由图 2(a) 和 (b) 的仿真

图3 (网络版彩图) $m=5$, $\tau=1$, $-2 < q \leq 0$, $\delta=1$ 时, $PE_{q,\delta}$ 与 $PPE_{q,\delta}$ 对于 logistic 映射的动态变化检测
Figure 3 (Color online) Detecting dynamic changes in logistic map using $PE_{q,\delta}$ and $PPE_{q,\delta}$ when $m=5$, $\tau=1$, $-2 < q \leq 0$, $\delta=1$. (a) The $PE_{q,1}$ when $-2 < q \leq 0$; (b) the $PPE_{q,1}$ when $-2 < q \leq 0$

结果可见, 当 $0 < q \leq 2$, $\delta=1$ 时, 随着 q 减小, $PE_{q,1}$ 和 $PPE_{q,1}$ 整体呈上升趋势, 且当 $0 < q \leq 1$ 时, $PE_{q,1}$ 和 $PPE_{q,1}$ 的上升趋势比较明显. 此时, $PE_{q,1}$ 和 $PPE_{q,1}$ 对于 logistic 映射动态变化检测效果明显增强. 在 q, δ 取值相等的条件下, 对比 $PE_{q,1}$ 和 $PPE_{q,1}$ 发现, $PPE_{q,1}$ 对于 logistic 系统的动态变化放大程度远比 $PE_{q,1}$ 显著. 由此得出, 当 $q > 0$, $\delta=1$ 时, 减小 q 值, 可以提高 $PE_{q,1}$ 和 $PPE_{q,1}$ 对 logistic 系统动态变化的检测效果, 且在 q 取值相同的条件下, $PPE_{q,1}$ 比 $PE_{q,1}$ 的检测效果更加显著.

图3为当 $m=5$, $\tau=1$, $-2 < q \leq 0$, $\delta=1$ 时, 以步长 $\Delta q=0.1$, 改变熵参数 q 值, 广义排列熵 ($PE_{q,\delta}$) 和广义相位排列熵 ($PPE_{q,\delta}$) 对于 logistic 映射动态变化的检测结果. 由图3(a)可知, 当 $-2 < q \leq 0$, $\delta=1$ 时, 随着 q 减小, 每次 logistic 系统出现空白窗后, $PE_{q,1}$ 都会出现明显的上升毛刺, 直至 $q=-2$ 时, $r(n)=3.86$ 处的毛刺直接影响对 logistic 系统动态变化的整体判断. 对于图3(b), 当 $-2 < q \leq 0$, $\delta=1$ 时, 随着 q 减小, 在 logistic 映射的非空白窗处, $PPE_{q,1}$ 的起伏波动也明显被放大, 表现为毛刺加剧. 由此得出, 当 $q < 0$, $\delta=1$ 时, $PE_{q,1}$ 和 $PPE_{q,1}$ 都会产生毛刺, 但是 $PPE_{q,1}$ 产生的毛刺不影响对系统动态变化的判断.

图4为当 $m=5$, $\tau=1$, $q=1$, $0 < \delta \leq 2$ 时, 以步长 $\Delta \delta=0.1$, 改变熵参数 δ 值, 广义排列熵 ($PE_{q,\delta}$) 和广义相位排列熵 ($PPE_{q,\delta}$) 对于 logistic 映射动态变化的检测结果. 由图4(a)和(b)的仿真结果可见, 当 $q=1$, $0 < \delta \leq 2$ 时, 随着 δ 值增大, $PE_{1,\delta}$ 和 $PPE_{1,\delta}$ 整体呈上升趋势, 且对 logistic 映射动态变化的检测效果越来越显著; 但是, 相比于 $PE_{1,\delta}$ 的上升幅度, 随着 δ 增大, $PPE_{1,\delta}$ 上升的更快, 且对于 logistic 动态变化检测效果更为明显. 由此说明, 当 $q=1$, $\delta > 0$ 时, 增大 δ 值, 可以提高 $PE_{1,\delta}$ 和 $PPE_{1,\delta}$ 对于 logistic 系统动态变化的检测效果, 且 δ 取值相同的条件下, $PPE_{1,\delta}$ 比 $PE_{1,\delta}$ 的检测效果更为显著.

图5为 $m=5$, $\tau=1$, $q=1$, $-2 < \delta \leq 0$ 时, 以步长 $\Delta \delta=0.1$, 改变熵参数 δ 值, 广义排列熵 ($PE_{q,\delta}$) 和广义相位排列熵 ($PPE_{q,\delta}$) 对于 logistic 映射动态变化的检测结果. 由图5(a)和(b)可见, 当 $q=1$, $-2 < \delta \leq 0$ 时, 随着 δ 值减小, $PE_{1,\delta}$ 和 $PPE_{1,\delta}$ 整体呈下降趋势, 但是两者在 logistic 系统出现混沌空白窗时熵值增大, 在非空白窗处熵值减小, 这与实际恰好相反. 这是由于为了全面分析 δ 值的影响, 人为扩充了 δ 的取值范围, 使其属于实数域. 而且, 当 $q=1$, $-2 < \delta \leq 0$ 时, 随着 δ 值减小,

图 4 (网络版彩图) $m = 5$, $\tau = 1$, $q = 1$, $0 < \delta \leq 2$ 时, $PE_{q,\delta}$ 与 $PPE_{q,\delta}$ 对于 logistic 映射的动态变化检测
 Figure 4 (Color online) Detecting dynamic changes in logistic map using $PE_{q,\delta}$ and $PPE_{q,\delta}$ when $m = 5$, $\tau = 1$, $q = 1$, $0 < \delta \leq 2$. (a) The $PE_{1,\delta}$ when $0 < \delta \leq 2$; (b) the $PPE_{1,\delta}$ when $0 < \delta \leq 2$

图 5 (网络版彩图) $m = 5$, $\tau = 1$, $q = 1$, $-2 < \delta \leq 0$ 时, $PE_{q,\delta}$ 与 $PPE_{q,\delta}$ 对于 logistic 映射的动态变化检测
 Figure 5 (Color online) Detecting dynamic changes in logistic map using $PE_{q,\delta}$ and $PPE_{q,\delta}$ when $m = 5$, $\tau = 1$, $q = 1$, $-2 < \delta \leq 0$. (a) The $PE_{1,\delta}$ when $-2 < \delta \leq 0$; (b) the $PPE_{1,\delta}$ when $-2 < \delta \leq 0$

$PE_{1,\delta}$ 和 $PPE_{1,\delta}$ 对于 logistic 系统动态变化检测并无明显放大, 表现为熵值属于区间 $(0,1)$. 由此说明, 当 $q = 1$, $\delta < 0$ 时, $PE_{1,\delta}$ 和 $PPE_{1,\delta}$ 对于 logistic 系统动态变化检测结果与实际相反, 且减小 δ 值, $PE_{1,\delta}$ 和 $PPE_{1,\delta}$ 对于 logistic 系统动态变化检测效果无明显增强.

接着对于 q 减小且 δ 增大时, 广义排列熵 ($PE_{q,\delta}$) 与广义相位排列熵 ($PPE_{q,\delta}$) 在检测 logistic 系统动态变化时的规律进行分析. 分别取 $q = 0.3$, $\delta = 1.7$, $q = 0.4$, $\delta = 1.6$, $q = 0.5$, $\delta = 1.5$, $q = 1$, $\delta = 1$, 计算 logistic 映射的 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$, 结果如图 6(a) 和 (b) 所示. 图 6 的仿真结果显示, 当 $q > 0$ 且 $\delta > 0$ 时, q 值选取越小且 δ 取值越大, $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 对于 logistic 系统动态变化检测效果越显著; 同时, 在 q, δ 取值相同的条件下, $PPE_{q,\delta}$ 对于动态变化的检测效果比 $PE_{q,\delta}$ 显著的多. 然而, q 越小且 δ 取值越大, $PE_{q,\delta}$ 在 logistic 系统出现空白窗后产生的上升毛刺越明显, 而 $PPE_{q,\delta}$ 并未出现毛刺. 因

图 6 (网络版彩图) $m = 5$, $\tau = 1$ 时, q, δ 取不同值, $PE_{q,\delta}$ 与 $PPE_{q,\delta}$ 对于 logistic 映射的动态变化检测

Figure 6 (Color online) Detecting dynamic changes in logistic map with different value of q, δ using $PE_{q,\delta}$ and $PPE_{q,\delta}$ when $m = 5$, $\tau = 1$. (a) The generalized permutation entropy ($PE_{q,\delta}$); (b) the generalized phase permutation entropy ($PPE_{q,\delta}$)

图 7 (网络版彩图) $m = 5$, $\tau = 1$ 时, 不同数据长度和信噪比 (SNR) 时的 $PE_{q,\delta}$ 与 $PPE_{q,\delta}$

Figure 7 (Color online) The $PE_{q,\delta}$ and $PPE_{q,\delta}$ in different length of data and signal-to-noise ratio (SNR) when $m = 5$, $\tau = 1$. (a) The sequence of Gaussian wave packets; (a1) enlarged view of each wave packet in a Gaussian wave packet sequence; (b) the $PE_{q,\delta}$ and $PPE_{q,\delta}$ in different length of data when $m = 5$, $\tau = 1$; (c) the $PE_{q,\delta}$ and $PPE_{q,\delta}$ in different SNR when $m = 5$, $\tau = 1$

此, 在之后的分析时可以同时使 q, δ 取较优值, 但是要尽可能的选用 $PPE_{q,\delta}$ 去测量, 从而使检测结果更为显著.

为了探究数据长度和噪声对于广义相位排列熵的影响, 产生一数据长度为 10000 的 Gauss 波包

序列, 如图 7(a) 所示. 图 7(b) 为 $m = 5, \tau = 1$ 时, 以 1000 为递增步长改变数据长度所对应的广义排列熵 ($PE_{q,\delta}$) 和广义相位排列熵 ($PPE_{q,\delta}$). 由图 7(b) 的仿真结果可见, 当 $m = 5, \tau = 1$ 且 q, δ 取不同值时, 不同数据长度下, $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 的值基本保持稳定, 并不会因数据长度改变而出现明显熵值波动. 由此说明, 数据长度对 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 影响较小. 图 7(c) 为 $m = 5, \tau = 1$ 时, 对图 7(a) 所示数据长度为 10000 的 Gauss 波包序列加入信噪比 (SNR) 初始值为 -15 dB, 递增步长为 0.5 dB 的 Gauss 噪声, 重复计算 30 次得到的各信噪比时广义排列熵 ($PE_{q,\delta}$) 和广义相位排列熵 ($PPE_{q,\delta}$) 的统计平均熵值. 由图 7(c) 的仿真结果可见, 对于数据长度为 10000 的 Gauss 波包序列, 不同 SNR 时的 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 值都相对稳定, 若观察具体熵值, 不同 q, δ 时的 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 都随着 SNR 的增加有轻微的减小. 由此说明 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 对于噪声具有良好的鲁棒性.

4 广义相位排列熵在异常心电信号检测中的应用

心脏疾病是当代社会致死的主要疾病之一, 对心脏疾病的诊断依赖于对心电信号特征的分析 and 提取. 因此, 准确抓取心电信号中的变化特征对于疾病诊断尤为重要. 心室纤颤 (ventricular fibrillation, VF) 是心脏病病人突然性死亡最为常见的致死疾病, 而心室扑动 (ventricular flutter, VFL) 被认为是心室纤颤发病的前兆. 同时, 心室扑动是由严重的室性心动过速 (ventricular tachycardia, VT) 逐渐演变而成. 基于此, 本节分析广义相位排列熵对这 3 种实际信号动态变化的检测效果, 以期望为研究心脏活动和进一步早期临床疾病诊断提供一些新的思路 and 手段.

借助国际标准数据库 PhysioNet^[20], 选取编号为 cu02 的室性心动过速信号, 编号为 cu01 的心室纤颤信号和编号为 418 的心室扑动信号, 各取这 3 种信号病发状态左右共计 2500 个数据点. 为了使熵值变化更贴适于实际数据信息变化且考虑到 C-C 方法计算得到的窗口长度 $w = 200$ 会损失过多起始处的熵值变化信息^[21], 此处取窗口长度为 $w = 100$, 时间延迟 $t_0 = 1$ 生成各异常心电信号的子序列矩阵, 即此时各子序列为最大重叠序列, 再计算子序列矩阵中每一子序列的 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$, 从而得到各异常心电信号的熵值变化. 其中, 计算各熵值时的相空间重构嵌入维数为 $m = 5$, 时间延迟为 $\tau = 1$, 仿真结果如图 8~10.

图 8(a) 所示为 cu02 室性心动过速信号, 其中数据点 0~1500 为室性心动过速病发阶段, 数据点 1500~2500 为正常阶段, 图 8(b) 和 (c) 分别为 $q = 1, \delta = 1, q = 0.5, \delta = 1.5, q = 0.4, \delta = 1.6$ 时室性心动过速信号的 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 熵值图. 由图 8(b) 和 (c) 的仿真结果可以看出, 随着 q 减小 δ 增大, $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 的熵值都相应的增加, 且室性心动过速病发阶段与正常阶段各自熵值的差别越来越大. 但是, 在 q, δ 取值相等的条件下, 相比于 $PE_{q,\delta}$, $PPE_{q,\delta}$ 能更容易地区分出室性心动过速病发阶段和正常阶段.

图 9(a) 所示为 418 心室扑动信号, 数据点 600~1500 为心室扑动病发阶段, 其余处为正常阶段. 图 9(b) 和 (c) 分别为 $q = 1, \delta = 1, q = 0.5, \delta = 1.5, q = 0.4, \delta = 1.6$ 时心室扑动信号的 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 熵值图. 与图 8(b) 和 (c) 一致, 图 9(b) 和 (c) 显示随着 q 减小 δ 增大, $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 的熵值都相应增加且对于心室扑动信号状态变化检测越来越清晰, 特别是 $PPE_{q,\delta}$ 能很清楚地显示出心室扑动病发阶段与正常阶段相互转换发生的具体数据点, 即正常阶段转变为病发阶段发生在数据点 600 左右, 病发阶段转变为正常阶段发生在数据点 1500 左右.

如图 10(a) 所示为 cu01 心室纤颤信号, 数据点 1100~2500 为心室纤颤病发阶段, 其余处为正常心电信号. 图 10(b) 和 (c) 分别为 $q = 1, \delta = 1, q = 0.5, \delta = 1.5, q = 0.4, \delta = 1.6$ 时心室纤颤信号的 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 熵值图. 相比于图 8 和 9, 无论 q, δ 取何值, $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 都能明显地检测出心室纤颤病

图 8 (网络版彩图) $PE_{q,\delta}$ 与 $PPE_{q,\delta}$ 对比分析 $m=5, \tau=1$ 时的室性心动过速信号

Figure 8 (Color online) $PE_{q,\delta}$ vs. $PPE_{q,\delta}$ in the case of the VT signal when $m=5, \tau=1$. (a) The VT signal; (b) comparison between $PE_{q,\delta}$ with different parameters q, δ for the VT signal; (c) comparison between $PPE_{q,\delta}$ with different parameters q, δ for the VT signal

图 9 (网络版彩图) $PE_{q,\delta}$ 与 $PPE_{q,\delta}$ 对比分析 $m=5, \tau=1$ 时的心室扑动信号

Figure 9 (Color online) $PE_{q,\delta}$ vs. $PPE_{q,\delta}$ in the case of the VFL signal when $m=5, \tau=1$. (a) The VFL signal; (b) comparison between $PE_{q,\delta}$ with different parameters q, δ for the VFL signal; (c) comparison between $PPE_{q,\delta}$ with different parameters q, δ for the VFL signal

发阶段. 对比相同 q, δ 时的 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 发现, 对于心室纤颤检测, $PPE_{q,\delta}$ 比 $PE_{q,\delta}$ 更能清晰刻画数据的起伏及状态变化. 同时, 随着 q 减小 δ 增大, 病发阶段与正常阶段之间的 $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 的熵值差增加.

上述分析显示, 室性心动过速、心室扑动、心室纤颤, 这 3 种疾病在病发阶段的系统复杂性依次

图 10 (网络版彩图) $PE_{q,\delta}$ 与 $PPE_{q,\delta}$ 对比分析 $m = 5, \tau = 1$ 时的心室纤颤信号

Figure 10 (Color online) $PE_{q,\delta}$ vs. $PPE_{q,\delta}$ in the case of the VF signal when $m = 5, \tau = 1$. (a) The VF signal; (b) comparison between $PE_{q,\delta}$ with different parameters q, δ for the VF signal; (c) comparison between $PPE_{q,\delta}$ with different parameters q, δ for the VF signal

降低, 且正常阶段的系统复杂性高于病发阶段. 与 logistic 映射得出的结论一致, 在 q, δ 取相同值时, $PPE_{q,\delta}$ 对于这 3 种疾病的动态变化检测相比于 $PE_{q,\delta}$ 更为显著. 此外, 随着 q 减小且 δ 增大, $PE_{q,\delta}$ 和 $PPE_{q,\delta}$ 对于这 3 种疾病的病发状态检测效果明显增强. 因此, 合理选择 q, δ 值, $PPE_{q,\delta}$ 能更好、更准确地检测出病理系统的动态变化.

5 总结

本文将相位信息引入到广义排列熵中, 提出广义相位排列熵算法. 借助标准模型 logistic 映射, 对比分析了广义相位排列熵在检测系统动态变化方面的性能, 探究了两参数熵中的 q, δ 改变对于 logistic 系统动态变化检测的影响, 并给出了如何恰当地选取 q, δ 值以提高广义排列熵和广义相位排列熵对于微小变化的检测效果. 通过 Gauss 波包序列, 研究了信号的数据长度和信噪比变化对广义相位排列熵的影响. 最后, 我们将广义相位排列熵运用到异常心电信号的检测分析中. 仿真表明: 当广义排列熵和广义相位排列熵的 q, δ 取值相同时, 对于系统动态变化检测, 广义相位排列熵比广义排列熵更为明显. 当 $q > 0, \delta > 0$ 时, q 取值越小, δ 取值越大, 无论是针对 logistic 系统还是异常心电信号的动态变化检测, 广义排列熵和广义相位排列熵都能放大系统的动态改变, 使检测结果更为显著. 此外, 数据长度变化对于广义排列熵和广义相位排列熵基本无影响且两者对噪声均具有良好的鲁棒性. 广义相位排列熵在系统动态变化检测领域具有一定优势, 能够为之医生临床诊断疾病提供一定参考.

参考文献

- 1 Pincus S M. Approximate entropy as a measure of system complexity. Proc Natl Acad Sci USA, 1991, 88: 2297–2301
- 2 Richman J S, Moorman J R. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol-Heart Circ Physiol, 2000, 278: 2039–2049

- 3 Costa M, Goldberger A L, Peng C K. Multiscale entropy analysis of complex physiologic time series. *Phys Rev Lett*, 2002, 89: 068102
- 4 Costa M, Goldberger A L, Peng C K. Multiscale entropy analysis of biological signals. *Phys Rev E*, 2005, 71: 021906
- 5 Bandt C, Pompe B. Permutation entropy: a natural complexity measure for time series. *Phys Rev Lett*, 2002, 88: 174102
- 6 Cao Y H, Tung W W, Gao J B, et al. Detecting dynamical changes in time series using the permutation entropy. *Phys Rev E*, 2004, 70: 046217
- 7 Jiang J, Shang P J, Zhang Z Q, et al. Permutation entropy analysis based on Gini-Simpson index for financial time series. *Phys A Stat Mech Appl*, 2017, 486: 273–283
- 8 Feng F Z, Rao G Q, Jiang P C, et al. Analysis on comparison of permutation entropy with correlation dimension for abnormality detection of a rolling bearing. In: *Proceedings of International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, 2012. 767–771
- 9 Zhao X J, Shang P J, Huang J J. Permutation complexity and dependence measures of time series. *Europhys Lett*, 2013, 102: 40005
- 10 Mammone N, Duun-Henriksen J, Kjaer T W, et al. Differentiating interictal and ictal states in childhood absence epilepsy through permutation Rényi entropy. *Entropy*, 2015, 17: 4627–4643
- 11 Fadlallah B, Chen B D, Keil A, et al. Weighted-permutation entropy: a complexity measure for time series incorporating amplitude information. *Phys Rev E*, 2013, 87: 022911
- 12 Gao J F, Chen K Y, Wu P, et al. Leakage detection of pipeline based on weighted-permutation entropy. In: *Proceedings of the 29th Chinese Control and Decision Conference*, 2017. 2820–2823
- 13 Liu T B, Yao W P, Wu M, et al. Multiscale permutation entropy analysis of electrocardiogram. *Phys A Stat Mech Appl*, 2017, 471: 492–498
- 14 Xia J N, Shang P J, Wang J, et al. Permutation and weighted-permutation entropy analysis for the complexity of nonlinear time series. *Commun Nonlinear Sci Numer Simul*, 2016, 31: 60–68
- 15 Xu M J, Shang P J. Generalized permutation entropy analysis based on the two-index entropic form $S_{q,\delta}$. *Chaos*, 2015, 25: 053114
- 16 Ribeiro M S, Nobre F D, Tsallis C. Probability distributions and associated nonlinear Fokker-Planck equation for the two-index entropic form $S_{q,\delta}$. *Phys Rev E*, 2014, 89: 052135
- 17 Bruns A, Eckhorn R, Jokeit H, et al. Amplitude envelope correlation detects coupling among incoherent brain signals. *Neuroreport*, 2000, 11: 1509–1514
- 18 Schyns P G, Thut G, Gross J. Cracking the code of oscillatory activity. *PLoS Biol*, 2011, 9: 1001064
- 19 Ng B S W, Logothetis N K, Kayser C. EEG phase patterns reflect the selectivity of neural firing. *Cerebral Cortex*, 2013, 23: 389–398
- 20 Goldberger A L, Amaral L A N, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 2000, 101: 215–220
- 21 Zhao X L, Ren M R, Zhao Y T, et al. Nonlinear analysis of ECG signal based on permutation entropy. *Signal Process*, 2010, 19: 90–93 [赵小磊, 任明荣, 张亚庭, 等. 基于排列熵的心电信号非线性分析. *信号处理*, 2010, 19: 90–93]

Generalized phase permutation entropy algorithm based on the two-index entropy

Huan KANG, Xiaofeng ZHANG* & Guangbin ZHANG

College of Physics and Information Technology, Shaanxi Normal University, Xi'an 710119, China

* Corresponding author. E-mail: xiaofengzhang71@snnu.edu.cn

Abstract The generalized permutation entropy ($PE_{q,\delta}$) with appropriate parameters can amplify the minor changes of the system, however, the phase of the signal contains more critical information than the amplitude. This paper introduces the phase information into the generalized permutation entropy and proposes the generalized phase permutation entropy ($PPE_{q,\delta}$) algorithm. At the same time, we verify the advantages of $PPE_{q,\delta}$ in detecting the dynamic changes of the system, analyze the influence of q , δ selection on the dynamic change of the system, and explore the effect of data length and noise for the $PPE_{q,\delta}$. Finally, the $PPE_{q,\delta}$ is applied to the analysis of abnormal ECG signals. The results show that when the value of q , δ are the same, the $PPE_{q,\delta}$ has a more significant effect on the detection of the same dynamic change. Whether for the logistic map or the abnormal ECG signal dynamic changes detection, when $q > 0$ and $\delta > 0$, the effect of $PE_{q,\delta}$ and $PPE_{q,\delta}$ can be improved by decreasing q value, or increasing δ value, or both changing simultaneously. Furthermore, the change of data length has no effect for the $PE_{q,\delta}$ and $PPE_{q,\delta}$ mainly, and both of them are robust to noise.

Keywords generalized phase permutation entropy, dynamic changes detection, logistic map, ECG signal, non-linear analysis

Huan KANG was born in 1994. She is pursuing a master's degree at Shaanxi Normal University. Her research direction is intelligent information processing and detection technology.

Xiaofeng ZHANG was born in 1971. She received her Ph.D. degree in underwater acoustic engineering from Northwestern Polytechnical University, Shaanxi, in 2003. Currently, she is a professor at Shaanxi Normal University. Her research interests include sound controlling, intelligent information processing and detection technology.

Guangbin ZHANG was born in 1971. He received his Ph.D. degree from National Laboratory of Radar Signal Processing at Xi'dian University, Shaanxi, in 2006. He is a professor in School of Physics and Information Technology of Shaanxi Normal University. He focuses on acoustic sensor and intelligent information processing and detection technology.