

DOI:10.19651/j.cnki.emt.2108444

时间序列复杂性自动度量:快速的弹簧测^{*}

程杰 李文石

(苏州大学电子信息学院 苏州 215006)

摘要:提出了时间序列复杂性自动度量的弹簧测试的一种提速方法。弹簧测试包括5阈值压缩空间中的弹簧图,以及该弹簧图的自相似性刻画数值CC。针对其速度较慢的问题,应用峰谷值映射简化弹簧图,提速自相似性度量。测试方程使用小波映射混沌、统一方程、Rossler超混沌方程。以分岔图为基础的对比判别几何特征,以基于FFT的谱熵复杂度为自动度量金标准,对比改进前后弹簧自相似性度量CC值的结果,给出各测试例的稳定性分析及计算时间比较。结果表明,改进后的弹簧测试能明显区别周期与混沌,增加了超混沌与混沌的辨识度,时间上提升2~20倍。本工作为周期、混沌、随机3分类识别的自动选参的信号复杂性度量贡献了改良的弹簧测试方法,本质是快速提取三维相图(弹簧图)的非对称非线性统计不变量。

关键词:时间序列;多阈值压缩;自动度量;弹簧测试;峰谷值

中图分类号: TP301 **文献标识码:** A **国家标准学科分类代码:** 510.4010

Automatic measure of time series complexity: Rapid spring test

Cheng Jie Li Wenshi

(School of Electronics and Information Engineering, Soochow University, Suzhou 215006, China)

Abstract: A speed-up method of spring test on automatic measure for time series complexity is proposed, including the spring graph in 5-threshold compressed space and the construction creep (CC) rate. For its slow speed, the specific method uses peak and valley values fitting to simplify the spring diagram and to speed up the self-similarity measuring. We select wavelet mapping chaos, unified equation and Rossler hyperchaos equation as test equations. To take the bifurcation diagram as the basic comparison rule of geometric characteristics and to adopt the spectral entropy complexity (SEC) based on FFT as the gold standard. The results of spring self-similarity measure CC before and after the improvement are compared. The calculation time and stability of each test method is given. The results show that the improved method can distinguish between period and chaos and elevate the resolution ratio of hyperchaos and chaos significantly with 2 to 20 times hoisting in terms of time. This work contributes an improved spring test method for the signal complexity measure featured with automatic parameter selection for distinguishing periodic, chaotic and random states. The essence is to extract the asymmetric nonlinear statistical invariants of three-dimensional phase diagram (the spring diagram).

Keywords: time series; multi-threshold compression; automatic measure; spring test; peak and valley values

0 引言

时间序列是一组按时间排序的变量,其可以反映系统不断变化的趋势以及线性或非线性的物理特性。这种时间序列广泛存在于各个研究领域,包括但不限于气象、金融、医疗和军事等^[1]。为了对其进行定性或定量分析,专家们定义了一系列关于时间序列复杂性度量的方法^[2]:庞加莱截面法(poincare section, PS)、Lempel-Ziv 复杂度 LZC

(complexity)、0~1混沌测试、排列熵(permutation entropy, PE)、谱熵复杂度(spectral entropy complexity, SEC)和弹簧测试(spring test)等,研究审美是满足自动模式识别的无参判据设计。

PS、LZC和0~1混沌测试从时域角度分析一维或多维时间序列。其中,PS通过降维展现时间序列状态的几何特征,自动PS生成方法有待更好解决^[3];LZC比照颗粒度将时间序列重构为符号序列,计算其新状态出现速率,而颗

收稿日期:2021-11-25

* 基金项目:江苏省自然科学基金(BK20141196)、江苏省研究生创新基金(KYCX18_2509)项目资助

粒化过程可能丢失系统有用信息^[4];0~1混沌测试应用正弦和余弦进行数据群扩展(构造随机过程)得到 pq 图,定义 p 或 q 变量的渐进增长率 K 值,二值判断混沌或秩序状态,但需要人为选择内部单参数^[5]。

从频域角度分析,SEC应用FFT得到频谱成分分布概率,求熵值进而归一化,不使用内部参数^[6-7],PE则先完成概率化后利用熵值再结合统计学方法,复杂度都为 $O(n \log_2 n)$,但PE需要选取嵌入维度和延迟时间^[8],而SE需要序列非平稳。

遵循SEC的无参计算思想,接受0~1混沌测试的群扩展启发,应用强于分类的非线性映射的弹簧测试,通过符号函数提取足够的相图非线性信息,构建三维特征弹簧图,几何度量弹簧的自相似性,得到复杂数CC值(construction creep),该三分类算法判据能够准确区分周期窗口、混沌或随机状态^[9-11],但是获得CC值将消耗较多计算时间。为加速计算CC值,本工作采用峰谷值采样法,缩减了实际计算数据的长度,在保证准确率不降低条件下,节省了2~20倍的计算时间,选用3个典型测试例进行对比验证。

1 弹簧测试原理

1.1 弹簧测试算法

区别于谱熵复杂度,弹簧测试本质是一种时域上的压缩熵。该法以三维时间序列为输入,若时间序列 $Q(n)$ 为一维时间序列(长度 N),则通过式(1)转化为三维序列,升维抽取策略是每三点抽一点。

$$\begin{aligned} x(i) &= Q(3i-2) \\ y(i) &= Q(3i-1) \\ z(i) &= Q(3i) \end{aligned} \quad (1)$$

定义弹簧测试任一维新变量形如式(2)所示。

$$s_\varphi(n) = \sum_{i=1}^n \{\phi(i) \operatorname{sgn}(\theta_\varphi(i))\} \omega_\varphi(i) \quad (2)$$

其中,被表征项 $\varphi(i)$ 为式(1)所得 $\{x(i), y(i), z(i)\}$

中的一个子集,sgn为符号函数,积分 $\theta_\varphi(n) = \sum_{i=1}^n \varphi(i)$, $n=1,2,\dots,\operatorname{INT}(N/3)$,5阈值压缩函数 $\omega_\varphi(i)$ 的具体构形为式(3)。

$$w_\varphi(n) = \begin{cases} 1, & \rho < \varphi(n) \leq \varphi_{\max} \\ 0.5, & \varphi_{\text{mean}} < \varphi(n) \leq \rho \\ 0, & \varphi(n) = \varphi_{\max} \\ -0.5, & \sigma < \varphi(n) \leq \varphi_{\text{mean}} \\ -1, & \varphi_{\min} < \varphi(n) \leq \sigma \end{cases} \quad (3)$$

黄金分割策略阈值关系为式(4):

$$\begin{aligned} \rho &= \lambda(\varphi_{\max} - \varphi_{\text{mean}}) + \varphi_{\text{mean}} \\ \sigma &= (1-\lambda)(\varphi_{\text{mean}} - \varphi_{\min}) + \varphi_{\min} \end{aligned} \quad (4)$$

其中, $\lambda=0.618$ 。

给定一个时间序列数据集,将其代入式(2)计算得到三维特征空间 $\{s_x(n), s_y(n), s_z(n)\}$,即成功定义了弹簧图。

试算观察传统蔡氏方程状态的弹簧图表征,启发我们采用切片方法,度量弹簧的自相似性,量化了弹簧图的完整程度。相应的CC值关键算法步骤如下。

1)将弹簧最近邻点的欧氏距离 $d(n)$ 归一化为数组 $D(n)$ 。

2)切片 $D(n)$ 得到分层矩阵 $\mathbf{R}^{m \times p}$:

$$\mathbf{R}^{m \times p} = \begin{bmatrix} D_1 & D_2 & \cdots & D_p \\ D_{p-1} & D_{p+2} & \cdots & D_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ D_{(m-1)p+1} & D_{(m-1)p+2} & \cdots & D_{mp} \end{bmatrix} \quad (5)$$

其中, $p = \operatorname{INT}(N/m)$, $m=2,3,\dots,\operatorname{INT}(N/2)$ 。

3)定义切片积累误差 e_i^m :

$$e_i^m = \begin{cases} e_i^m, & R_{i,k}^{m \times p} \in [R_{\delta,k}^{m \times p}(1-\beta), R_{\delta,k}^{m \times p}(1+\beta)] \\ e_i^m + 1, & \text{其他} \end{cases} \quad (6)$$

其中, $i=1,2,\dots,m-1$; $k=1,2,\dots,p$; $\delta=(i+1), (i+2), \dots, m$; 误差精度 β 固定为0.02。

4)最小平均积累误差定义为CC值:

$$CC = \min \left\{ \sum_{i=1}^p e_i^m / (m \cdot (p-1)) \right\} \times 100\% \quad (7)$$

前述分别定义的弹簧图和复杂数(CC值)是为弹簧测试算法。

CC值计算耗时原因:进一步切片比较的前提是矩阵 $\mathbf{R}^{m \times p}$ 大于1维,所以 m 至少从2遍历到 $\operatorname{INT}(N/2)$,而且 m 每递增1次,积累误差 e 的计算次数理论上会增加 $(m-1)$ 次,此过程耗费了时间较多,因此需要加速弹簧测试。

1.2 弹簧测试加速

数值计算建模的一般加速策略:减少遍历次数、变步长、以加代乘^[12-13]。

实验发现,计算点数每增加1000点,计算时间就会增加1倍。为快速判断时间序列的复杂度,基于理解弹簧切片原理,针对弹簧在某一轴的投影,按照式(8)搜索峰谷值,这样既保留了弹簧原有的非对称规律,又减少了 $\mathbf{R}^{m \times p}$ 的切片次数。

$$s'(n') = \min \max \{s_\varphi(n)\} \quad (8)$$

其中, $s'(n')$ 为弹簧图峰谷值组成的新三维序列(n' 是峰谷值个数),实际选取三轴峰谷值最多的一组新序列为式(9):

$$s''(n') = \begin{cases} s'(n'), & n' > 3 \\ s_\varphi(\operatorname{INT}(n'/f)i), i=1,2,\dots,f, & n' \leq 3 \end{cases} \quad (9)$$

当峰谷值个数小于3,安排等距采样,采样率设置为 $f=500$ 。

前述弹簧测试加速原理的算法思想图解如图1所示。

考察图1中的两根弹簧及其子图可知,弹簧在3个特征面的映射均能反映其几何特征,混沌状态为破损弹簧,结果其映射出现类似损坏,而周期状态拥有规则弹簧图斑,其

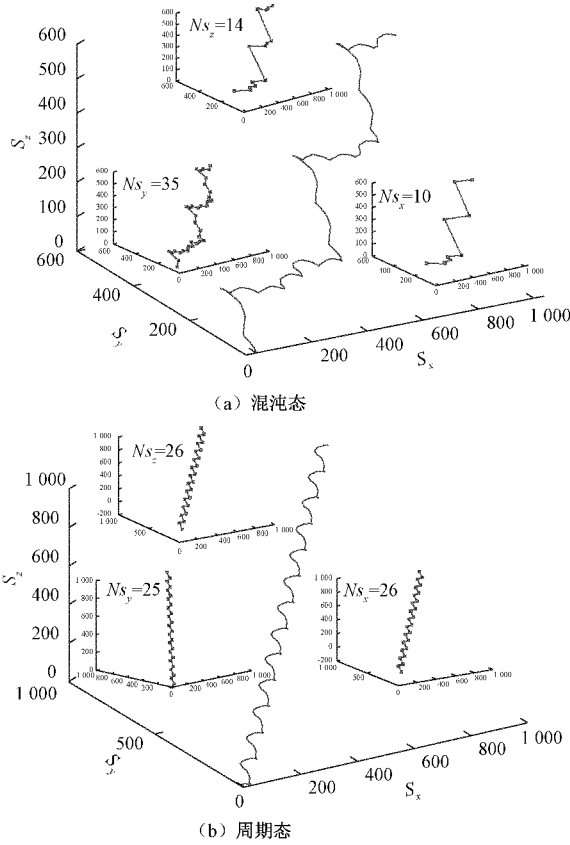


图 1 蔡氏方程的弹簧图及其 x, y, z 方向上映射图

映射结果也是规则的。因此我们确定研究其映射图而不是弹簧本身。可以选某一映射进行切片,但为了得到更加准确的 CC 值,实际记录每个方向上取到峰谷值的总数 N_s ,并且选取拥有最大 N_s 方向上的映射切片,此时的最小平均积累误差记为 rCC(rapid CC),定义为加速计算后的新复杂数。

1.3 误差精度分析

式(6)中的 β 是整个弹簧切片时的误差精度,在经过峰谷值映射之后,欧氏距离度量由最近邻点变为最近峰谷点之间,所以其误差精度需要重新确定。

图 2 所示为在不同 β 下,改进后弹簧图的蔡氏方程的 rCC 与改进之前的 CC 对比。

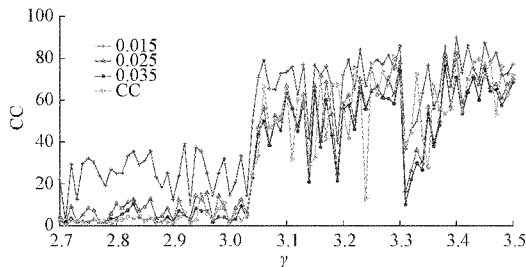


图 2 蔡氏方程在不同 β 下扫描的 CC 图

由图 2 可知,随着 β 减小,最小平均积累误差越大,过小的 β 会引起周期状态与混沌状态的混叠,过大的 β 则又

使计算精度下降。结合实验综合考虑, β 值实际选取 0.035。

在参数 $\gamma=3.24$ 时,CC 出现了负峰,此时的时间序列为混沌状态,改进之后的 CC 值抬升到 50 以上;当参数 $\gamma=3.31$ 时,因为其相图的相对对称性特征,CC 和 rCC 出现整体下降趋势。

2 测试结果与分析

选择 3 个典型测试例小波映射函数、统一方程和 Rossler 方程,以 SEC 为金标准,比较峰谷值映射前后复杂数的波动以及时间消耗对比。实验均在 CPU 型号为 Intel(R) Core(TM) i5-3230M,主频 2.6 GHz 的计算机上,运行 MATLAB2018b;高斯分布伪随机数的 SEC 值是 0.94。

2.1 小波映射函数

小波映射函数^[14]从周期态开始倍分岔进入混沌态再退出混沌态回到周期态,将正分岔与逆分岔结合到一起,其迭代式为:

$$x(n+1) = 1.3(1-x(n))^2 e^{0.5(x(n)-\mu)^2} \quad (10)$$

其中,参数 $\mu \in [-1.5, 1.5]$,分岔图和观测数据经由式(10)迭代 1 000 次,取后 200 点,初值 $x(1)=0.1$ 。

图 3 底部为小波映射分岔图,上方 3 条曲线分别对应复杂特征 SEC、CC 和 rCC。由图 3 可知,在分岔时刻,SEC 出现尖峰,CC 判为周期态,rCC 有轻微变化;当迭代进入混沌态时,3 种特征一致表征复杂变化。在系数 $\mu=-0.5$ 左右时,SEC 值有一块区域明显降低,对比观察置于底部的分岔图,其状态相较于区域外更加复杂,而 CC 与 rCC 值均对应增加,弹簧测试在此处表现出对复杂度更好的刻画能力。

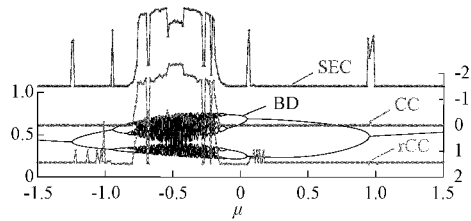


图 3 小波映射的分岔图和不同算法的复杂度值

2.2 统一方程系统

统一方程系统^[15]连接了 Lorenz 系统、Chen 系统和 Lu 系统,其包含了混沌系统簇的所有动力学行为。其微分方程组如式(11)所示。

$$\begin{aligned} dx/dt &= (25a+10)(y-x) \\ dy/dt &= (28-35a)x-xz+(29a-1)y \\ dz/dt &= xy-(8+a)z/3 \end{aligned} \quad (11)$$

其中,参数 $\alpha \in [0, 1]$,观测数据采用 ode45 求解 7 000 点(初值为[0.1,0.1,0.1],步长 0.01,去头 3 000 点)。

图 4 所示为统一方程系统随着参数 α 变化,SEC、CC

与 rCC 值的对比。根据对照算法 SEC 和 CC 值可知,在 α 值为 0.37、0.47、0.58 时系统处于周期态,改进后的弹簧计算出的 rCC 值也能准确找出这 3 个周期谷值窗口,并且在混沌态也有较高的 rCC 值,与对照算法一致。

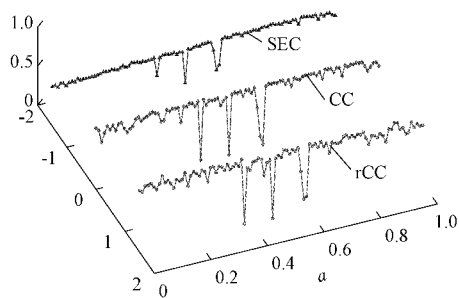


图4 统一方程的不同算法的复杂度值

2.3 Rossler 超混沌系统

Rosler 超混沌系统^[16]可能具有更加复杂的动力学行为,该系统在不同方向产生双曲超混沌吸引子,其微分方程如式(12)所示。

$$\begin{aligned} dx/dt &= -y - z \\ dy/dt &= x + 0.25y + w \\ dz/dt &= 3 + xz \\ dw/dt &= -cz + 0.5w \end{aligned} \quad (12)$$

其中,参数 $c \in [0.2, 0.6]$,观测数据采用 ode45 求解 7 000 点(初值为[0.1, 0.1, 0.1, 0.1],步长 0.01 去头 3 000 点)。

图 5 所示为 Rosler 超混沌系统随参数 c 变化,SEC、CC 与 rCC 值的对比。

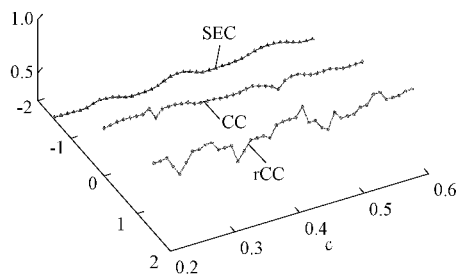


图5 Rosler 方程不同算法的复杂度值

从图 5 可知,对于 Rosler 超混沌方程,从频域角度计算 x 、 y 、 w 维的 SEC 值总体较低,但 z 维的 SEC 出现了 0.9 以上的高值区域;从时域方面计算得到 x - y - z 相图非对称性的 CC 和 rCC 值总体较高,且其变化趋势基本相同,相对于统一方程混沌态的 rCC 值,超混沌态的 rCC 值提高了 20%。

2.4 稳定性分析

此处的稳定性将考察,误差精度 β 值为 0.02 时的 CC 值,与 β 值为 0.035 时的 rCC 值的差值。

图 6 所示为上述 3 个测试例在改进前后的 CC 值变化。因为每个测试例的参数不同,本文等距采样了非周期态下的 14 个参数。可知在混沌态下,CC 值波动大致在

$[-10, 10]$ 之间;对于 Rosler 超混沌态,其 CC 值整体上飘且变化均较大。说明经过改进后的 CC 在混沌态下不会有较大波动,但更易于区分出超混沌态。

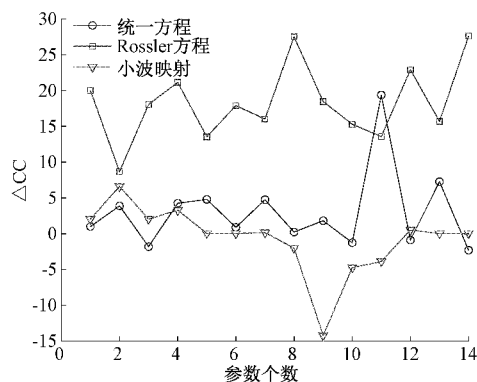


图6 3个测试例的扫参数与误差图

2.5 计算时间对比

表 1 均选取 4 000 点数据长度,记录最耗时情形。

表 1 3 个测试例的不同方法计算时间对比 s

测试例	SEC	CC	rCC
小波映射函数	0.791	7.300	3.100
统一方程系统	0.025	7.100	0.332
Rosler 超混沌系统	0.023	6.900	0.306

结合表 1 可知 CC 的计算复杂度较高,而 SEC 和 rCC 的计算复杂度较低。其中,小波函数的 rCC 比 CC 提升了至少 2 倍速度,而统一方程和 Rosler 方程则提升了最低 20 倍的速度。因此 rCC 能够从时域角度快速判断系统的复杂状态。

3 结 论

弹簧测试给出弹簧图表征时域信号的主要类别,例如周期(完好弹簧)、准周期(较好弹簧)、混沌(受损弹簧)、高斯伪随机数(细长棍),相应地自恰给出复杂数(CC 值分别小于 7、介于 7 和 50 之间、介于 50 和 84 之间、大于 84)。

本文尝试加速了弹簧测试的复杂数 CC 值计算。首先采样弹簧图的峰谷值,再做自相似性度量,跟进微调了误差精度。以谱熵复杂度 SEC 为对照判据,选取 3 个混沌方程实验验证。结果表明,改进后的 CC 值计算时间缩短了 20 倍(ode45 解算混沌方程条件下),并且增加了混沌与超混沌的复杂性辨识度。

进一步的算法验证工作,将涉及混沌检测、染噪测试、过采样研究、稳定性进一步分析等。

参考文献

- [1] PARISI G. Complex systems: A physicist's viewpoint[J]. Physica A, 1999, 263(1-4): 557-564.
- [2] TANG L, LV H L, YANG F M, et al. Complexity

- testing techniques for time series data: A comprehensive literature review[J]. *Chaos, Solitons and Fractals*, 2015, 81: 117-135.
- [3] 张识, 王攀, 张瑞浩, 等. 选取任意庞加截面的新方法[J]. *物理学报*, 2020, 69(4): 64-73.
- [4] 张栋, 陈东伟, 游雅, 等. 基于自适应 Lempel-Ziv 复杂度情感脑电信号特征分析[J]. *计算机应用与软件*, 2014, 31(9): 162-165.
- [5] ARMAND J S, BODO B. A modified 0-1 test for chaos detection in oversampled time series observations [J]. *International Journal of Bifurcation and Chaos*, 2014, 24(5): 1450063-1-13.
- [6] POWELL G, PERCIVAL I. A spectral entropy method for distinguishing regular and irregular motion of Hamiltonian system [J]. *Journal of Physics A: Mathematical and General*, 1979, 12(11): 2053-2071.
- [7] 孙克辉, 贺少波, 何毅, 等. 混沌伪随机序列的谱熵复杂性分析[J]. *物理学报*, 2013, 62(1): 35-42.
- [8] AUDUN M, KHASAWNEH F A. On the automatic parameter selection for permutation entropy [J]. *Chaos*, 2020, 30(3): 033130.
- [9] WU S L, LI Y T, LI W S, et al. Chaos criteria design based on three-threshold sign function [J]. *Chinese Journal of Electronics*, 2019, 28(2): 364-369.
- [10] CAI J W, LI Y T, LI W S, et al. Two entropy-based criteria design for signal complexity measures [J]. *Chinese Journal of Electronics*, 2019, 28(6): 1139-1143.
- [11] 应杰攀, 宋贺伦, 罗晓明. 基于 GaAs/AlGaAs 超晶格受激混沌振荡的随机数发生器设计与实现 [J]. *电子测量技术*, 2021, 44(9): 1-5.
- [12] DUBOIS M A. Modelling nature: A physicist's viewpoint [J]. *Science Asia*, 2009, 35: 1-7.
- [13] WADE J, HEYDARI B. Complexity: Definition and reduction techniques: Some simple thoughts on complex systems [J]. *Complex Systems Design and Management*, 2014, 1234(18): 213-226.
- [14] 于万波, 魏小鹏. 一个小波函数指数参数变化的分岔现象 [J]. *物理学报*, 2006, 55(8): 3969-3973.
- [15] LV J H, CHEN G R, CHENG D Z, et al. Bridge the gap between the Lorenz system and the Chen system [J]. *International of Bifurcation and Chaos*, 2002, 12(12): 2917-2926.
- [16] ROSSLER O E. An equation for hyperchaos [J]. *Physics Letter A*, 1979, 71: 155-157.

作者简介

程杰, 硕士生, 主要研究方向为混沌自动识别。

E-mail: 20204228009@stu.suda.edu.cn

李文石(通信作者), 博士, 教授, 主要研究方向为复杂性自动度量 and 芯片工程。

E-mail: lwshi@suda.edu.cn