

语音信号相空间重构中时间延迟选择的改进的平均位移法*

林嘉宇* 黄芝平** 王跃科** 沈振康*

(国防科技大学电子技术系* 机械电子工程与精密仪器系** 长沙 410073)

摘要 本文在分析时间序列相空间重构中的时间延迟选取的平均位移法基础上,对该法的原有度量进行改进,得到较好的求时间延迟的准则。改进的平均位移法具有更强的理论依据;应用于语音信号相空间重构的仿真实验表明,其度量一般情况下可得到合适的时间延迟。

关键词 语音信号,混沌,相空间重构,时间延迟,平均位移法

分类号 TN912.3

Modification of Average Displacement Method for Selection of Time-delay in Phase Space Reconstruction of Speech Signals

Lin Jiayu* Huang Zhiping** Wang Yueke** Shen Zhenken*

(* Department of Electronic Technology ** Department of Mechanotronics Engineering and Instrumentation, NUDT, Changsha, 410073)

Abstract Based on the analysis of average displacement (AD) method for selection of time delay in phase space reconstruction of speech signals, this paper modifies the original measurement of AD method to get a better criterion for proper time delay selection, which has a sound theoretic basis and can produce better effect.

Key words speech signals, chaos, phase space reconstruction, time delay, average displacement method

近十几年来,混沌信号处理^[1]为人们提供了分析自然现象的全新方法。混沌吸引子的相空间重构一般是分析混沌动力学系统的第一步,根据著名的 Takens 嵌入定理^[2],可使用延迟法。延迟法应用延迟坐标重构相空间矢量,即,

$$x_m(n) = [x(n), x(n + \tau), \dots, x(n + (m - 1)\tau)], \quad (1)$$

其中 $x(n)$ 是已离散化的待研究系统在 n 时刻的观测值, τ 称时间延迟, m 称重构维。如何选择适当的重构维和时间延迟是相空间重构的主要研究内容。

现有的时间延迟的选择方法一般都基于如下的两个准则之一:(1) 序列相关法。让 x_m 内元素之间的相关性减弱,同时 x_m 包含的原动力学系统的信息不丢失;(2) 相空间扩展法。重构相空间轨迹应从相空间的主对角线(τ 很小时)尽可能地扩展,但又不出现折叠。例如,自相关法^[3]、互信息量法^[4]、高阶相关法^[5]等属于第一类准则;填充因子法^[6]、摆动量法^[7]、平均位移法^[8]、SVF 法^[9]等属于第二类准则。

1 AD 方法的度量

求时间延迟的各种方法中,AD 方法具有如下的突出优点:(1) 概念明晰,在相空间几何学中有明确的意义;(2) 数学式简洁,易于计算;(3) 计算复杂度中等;(4) 对数据量要求不大,较小数据量求解结果的可靠性仍较高;(5) 对含噪数据集计算结果的鲁棒性较高;(6) 易于扩展到高维等。因此,AD 方法受到了人们的重视。

定义 m 维相空间重构下的平均位移 $\langle S_m^-(\tau) \rangle$ 如下:

$$\langle S_m^-(\tau) \rangle = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=1}^{m-1} (x_{i+j\tau} - x_i)^2 \quad (2)$$

* 湖南省自然科学基金重点资助项目
1998年11月10日收稿
第一作者:林嘉宇,男,1973年生,博士生

其中, N 为观测序列 $\{x_i\}$ 的点数。如果 $\langle S_m^{-1}(\tau) \rangle$ 为零, 重构相空间轨迹将完全聚集于相空间主对角线, 而随着它的增大, 轨迹将逐渐扩展, 因此 $\langle S_m^{-1}(\tau) \rangle$ 可以作为相空间轨迹扩展度的度量。

2 AD 法原有度量求时间延迟的准则的不足

文献[8]认为, 可选择 AD 法度量值的波形斜率第一次降为其波形初始斜率 40% 以下对应的时间延迟 τ 为所求的时间延迟。我们作出语音信号 f-yu(女声发 'yu' 音, 实验室环境采音, 音长 0.59s, 16kHz 采样, 16bits 线性 PCM 量化, 记为 f-yu) 的二维 AD 值, 如图 1(a) 所示, 则点划线处 ($\tau = 13$) 为其所求。但文献[8]承认, 这种准则具有一定的随意性, 仅仅根据实验结果得到, 其理论根据不强。

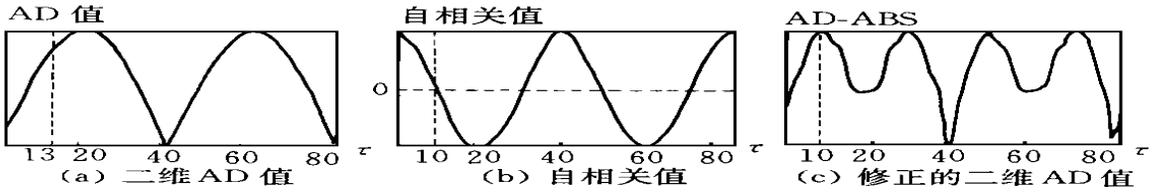


图 1 AD、自相关函数、AD-ABS 间相互关系(语音信号 f-yu)

Fig. 1 Relationship between AD, auto-correlation and AD-ABS (Speech signal 'f-yu')

称和相空间的主对角线垂直的对角线为辅对角线。设二维相空间的两轴为 x 轴和 y 轴, 若其主对角线为 $x = y$, 则其辅对角线为 $x = -y$ 。高维相空间的情形可以类推。那么, 从二维 AD 值和自相关函数值之间的对应关系中可知, 二维 AD 值由波形的谷点到波形的峰点的演化过程, 一般而言对应重构相空间轨迹由相空间的主对角线, 经最佳扩展后, 向其辅对角线聚集的过程。这是文献[8]没有指出的。一般地, 高维 AD 法的情形也是如此。

作出时间延迟为 1、10 和 21 情形下语音信号 f-yu 的二维相空间图, 如图 2 所示。我们仅作出其轨迹的中间 250 个样点。其中, 时间延迟 1 对应二维 AD 的峰点, 10 对应中间值, 21 对应谷点。重构相空间轨迹图证实了我们的想法, 图 2(a) 的轨迹图集中于相空间的主对角线, 图 2(b) 非常扩展, 而图 2(c) 则聚集于相空间的辅对角线。

事实上, AD 法度量的核心是 $\Delta x_i = x_{i+j\tau} - x_i$ 。当 x_i 和 $x_{i+j\tau}$ 的值较接近时, Δx_i 值较小, AD 度量的值总体较小, 表现为聚集于主对角线, 图 2(a) 就是这样; 当 x_i 和 $x_{i+j\tau}$ 的值差别较大, 例如一正一负时, Δx_i 值较大, AD 度量的值总体较大, 表现为聚集于负对角线, 如图 2(c)。这两种情况都是我们所要避免的。

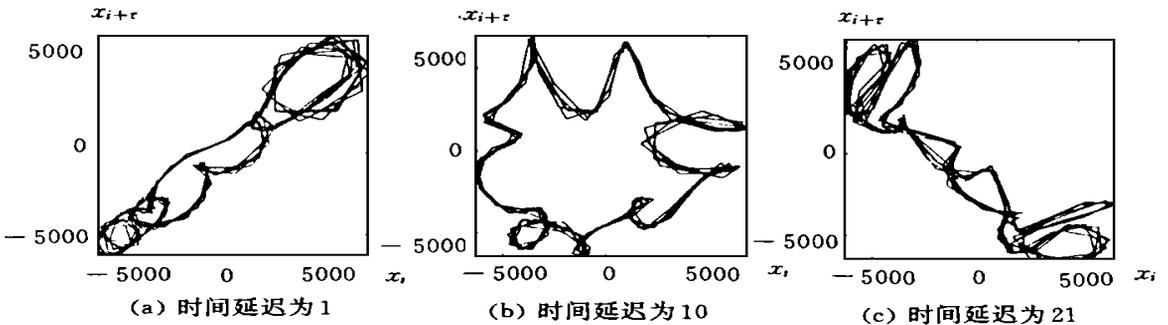


图 2 时间延迟为不同典型值的语音信号 f-yu 的二维重构相空间轨迹(250 个样点)

Fig. 2 Two-dimensional phase space portraits of speech signal 'f-yu' for typical time delays (250 samples)

根据文献[8]的准则求出的时间延迟并不是最佳的。如图 1 所示。对二维重构, 自相关法能提取序列样点间线性相关性最小的时间延迟, 可作为参照; 根据文献[8]的准则所求出的时间延迟 ($\tau = 13$) 显然已使序列样点间具有一定的相关性(自相关值为负, 相空间轨迹略向辅对角线聚集)。因此, 有必要根据

我们在上文所讨论的 AD 法原有度量的意义对其进行修正, 修正后的度量除保留原度量的优点外, 还应能提供易于求解时间延迟的准则。

3 对 AD 法的修正

定义 AD-ABS 如下:

$$\langle S_m^{-1}(\tau) \rangle = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=1}^{m-1} (x_{i+j\tau} - x_i)^2 \quad (3)$$

和(2)式相比, AD-ABS 仅仅将参加计算的值改为其绝对值(ABS 代表绝对值)。但是, 如图 1. c 所示, 二维 AD-ABS 的波形具有如下性质: 自相关函数的峰点对应二维 AD-ABS 的谷点(二维 AD 也是谷点), 重构相空间轨迹聚集于主对角线; 自相关函数的谷点对应二维 AD-ABS 的小谷点(二维 AD 却是峰点), 重构相空间轨迹聚集于辅对角线; 自相关函数的零点对应二维 AD-ABS 的峰点(二维 AD 无特征点对应), 二维重构相空间轨迹扩展。这样, 可得到二维重构时求解时间延迟的准则: 取二维 AD-ABS 度量值的第一个峰点所对应的时间延迟。

改进的主要目的在于将重构轨迹向相空间辅对角线聚集的情况避开, 同时明显地标识出重构轨迹的扩展。将 x_i 和 $x_{i+j\tau}$ 取绝对值后参加原来的二维 AD 计算, 其实是把相空间的其他像限对称地映射到第一象限, 即 $(x_i > 0, x_{i+j\tau} > 0)$ 像限, 原来在第一象限要避免的重构轨迹向主对角线聚集(二维 AD 值最小), 在 $(x_i > 0, x_{i+j\tau} < 0)$ 和 $(x_i < 0, x_{i+j\tau} > 0)$ 像限就演变为避免重构轨迹向辅对角线聚集。

作出 AD-ABS 在语音信号 $f-yu$ 高维重构中的波形图, 见图 3。高维 AD-ABS 法将 AD 法度量的峰点变为小谷点。重构轨迹聚集于相空间主对角线或辅对角线都对应于高维 AD-ABS 度量的谷点, AD-ABS 度量的峰点对应于高维重构时的轨迹扩展。这样得到高维重构中求解时间延迟的准则: 取 AD-ABS 度量值的第一个峰点所对应的时间延迟。

图 4 显示了时间延迟取不同典型值的语音信号 $f-yu$ 三维重构相空间轨迹的一部分。其中, 时间延迟取 7 对应(3)式相应情况的第一个峰点, 13 则对应(2)式的第一个峰点。我们注意到, 图 4(c) 的轨迹基本上和图 4(a) 垂直, 聚集于相空间辅对角线附近。但是, 此例的图 4(c) 仍较为扩展, 这从图 3 的对应度量值可看出。

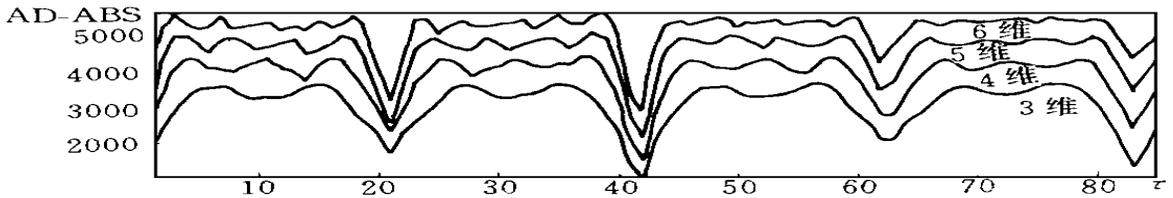


图 3 修正的平均位移法的高维度量对时间延迟的波形(语音信号 $f-yu$)
(波形从低往高排列, 依次为 3 维、4 维、5 维和 6 维)

Fig. 3 Waveform of high-dimensional measurements of modified AD method (speech signal $f-yu$)

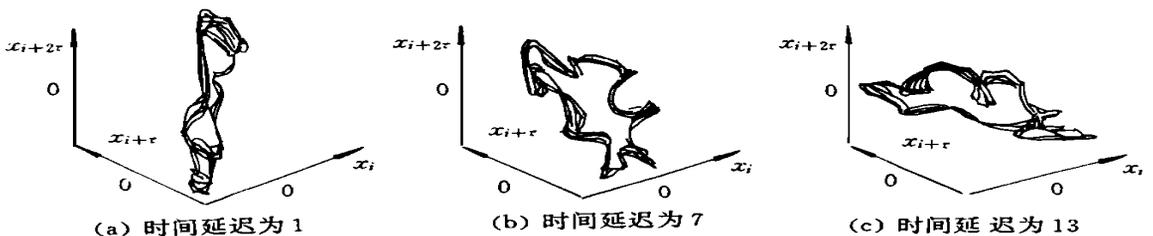


图 4 时间延迟为不同典型值的语音信号 $f-yu$ 的三维重构相空间轨迹(250 个样点)

Fig. 4 Three-dimensional phase space portraits of speech signal $f-yu$ for typical time delays (250 samples)

实验了男女声各一人发汉语基本音素各 56 个的样本测试。语音样本在普通实验室环境下采集, 发音音长为 0.6s 左右, 用普通微机经 16 位声卡采音, 采样率为 16kHz, 数据经 16bits PCM 线性量化。我们比较二维重构的 AD-ABS 法和自相关法。为了排除随机抖动的影响, 我们采取一些方法(如低通滤波)来求 AD-ABS 度量的第一个主峰点。在时间延迟误差为 ± 1 的允许条件下, 两者求解结果 108 个一致, 仅 4 个不一致, 不一致率约为 3.6%。而根据文献[8]的准则求得的结果和自相关法比较的不一致率为 100%。

出现二维 AD-ABS 法和自相关法求解结果不一致的原因大致为: 重构相空间轨迹不向辅对角线聚集, 或在向辅对角线偏转时轨迹分布仍较广、较为凌乱, 所求度量抖动厉害, 波形不光滑。其峰点不和自相关函数的零点对应。如图 5 的二维 AD-ABS 和自相关函数值波形、重构相空间图所示, 其中, 图 5(a) 的第一峰点对应时间延迟为 2, 图 5(b) 的第一零点为 11, 图 5(c) 的重构轨迹仅取 250 个样点。必须指出, 用二维 AD-ABS 法求出的时间延迟, 即使和自相关法不一致, 总体上仍满足重构相空间轨迹扩展的要求。

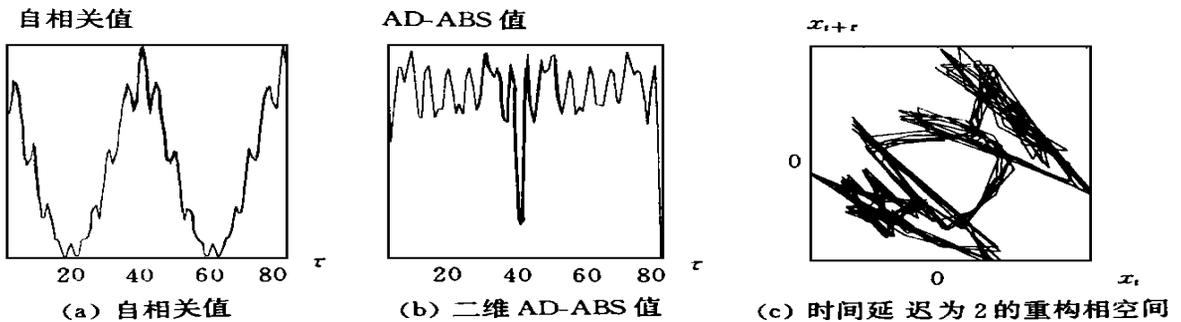


图5 语音信号 $f-i$ 的自相关值、二维 AD-ABS 和小时间延迟下的重构相空间

Fig. 5 Auto-correlation, two-dimensional AD-ABS and phase space portrait of speech signal $f-i$

4 结论

本文在分析求时间延迟的平均位移法基础上, 对该法的原有度量进行修正, 得到较好的求时间延迟的准则。修正的平均位移法具有更强的理论依据, 除了能标识出重构相空间轨迹聚集于相空间主对角线, 还能标识出轨迹聚集于辅对角线的情形。应用修正的平均位移法求解时间延迟, 只需找出其度量的第一主峰点, 准则简单。语音信号相空间重构的仿真实验表明, 其度量一般情况下可得到合适的时间延迟。

参考文献

- 1 Parker T S, Chua L D. Chaos: a tutorial for engineers. Proc of IEEE, 1987, 75 (8): 982 ~ 1008
- 2 Takens F. Detecting strange attractor in turbulence. in Dynamical systems and turbulence, Warwick, 1980, Lecture Notes in Mathematics, Rand and Young eds, 1981, 898: 366 ~ 381
- 3 Albano A M, et al. SVD and Grassberger-Proccaccia algorithm. Phy Rev A, 1988, 38: 3017 ~ 3026
- 4 Fraser A M. Information and entropy in strange attractors. IEEE Tr on IT, 1989, 35 (2): 245 ~ 262
- 5 Albano A M, et al. Using higher-order correlations to define an embedding window. Physica D, 1991, 54: 85 ~ 97
- 6 Buzug T, Pfister. Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors. Phy Rev A, 1992, 45: 7073 ~ 7084
- 7 Buzug T, Pfister G. Comparison of algorithms calculating optimal embedding parameters for delay time coordinates. Physica D, 1992, 58: 127 ~ 137
- 8 Rosenstein M T, Collins J J, Luca C. Reconstruction expansion as a geometry-based framework for choosing proper delay time. Physica D, 1994, 73: 82 ~ 98
- 9 Kember G, Fowler A C. A correlation function for choosing time delays in phase portrait reconstructions. Phy Lett A, 1993, 179: 72 ~ 80