

水下目标辐射噪声相空间重构参数选取方法

陆 扬, 孙良义, 廖述常

(广东湛江 91388 部队, 广东湛江 524022)

摘要: 利用自相关函数法和平均互信息法求取相空间重构的延迟参数, 对实测数据计算结果进行了分析比较, 确定了相空间重构时间延迟参数的选择方法。在用伪最近邻法进行相空间重构嵌入维数的选取时, 使用时空曲线法计算 Theiler Window 的窗口值 W , 并以此为依据剔除由于时间上的邻近而造成的空间上的邻近的最近邻点, 改进了伪最近邻计算嵌入维数的方法。实测数据计算结果表明, 该方法能够获取合理、有效的嵌入维参数。

关键词: 辐射噪声; 相空间重构; 嵌入维数; 时空曲线

中图分类号: TB561

文献标识码: A

文章编号: 1000-3630(2008)-05-0754-04

A method of phase space reconstruction parameters selection of underwater targets radiated noises

LU Yang, SUN Liang-yi, LIAO Shu-chang

(Unit 91388, PLA., Zhanjiang 524022, Guangdong, China)

Abstract: Reconstructing phase space via time series of one dimension is an important subject in nonlinear theory. Based on the autocorrelation function and mutual information methods, the phase space reconstruction delay parameters of three types of underwater targets radiated noises are calculated and determined. In the embedding dimension calculating process, the Theiler Window value- W is estimated by time-space separate curve. Thus the spatial nearest neighbors, resulted from the temporal neighborhood, can be estimated and eliminated. The calculation method of embedding dimension based on False Nearest Neighbors method is improved and the results indicate that the method can be used to obtain the embedding dimension reasonably and efficiently.

Key words: radiated noise; phase space reconstruction; embedding dimension; time-space separate curve

1 引 言

水下目标自动识别技术, 是水声对抗技术研究领域的一个重要研究内容。目标自动识别技术的关键在于目标固有特征的提取和具有良好推广性能的分类器的设计。其中, 目标固有特征提取技术的优劣直接影响到目标分类识别的成功率。

近年来, 水声信号的非线性特征提取研究十分引人注目^[1]。在非线性信号处理中^[2], 相空间重构是揭示时间序列数据内在本质特性的一个非常关键的

环节。本文基于非线性动力学中的相空间重构理论^[3], 计算确定了实测噪声数据相空间重构延迟参数。为了解决伪最近邻点法选取嵌入维数过程中对非几何特性的最近邻点的误统计问题, 对伪最近邻点法作了改进并应用于实测噪声信号, 为辐射噪声时间序列的相空间重构提供了一种适用的方法。

2 延迟法相空间重构原理

若 d 维动力系统 $dz/dt=f(z)$ 的吸引子轨道是 $z(t)$, 且 $z(t)=z, g:R^d \rightarrow R$ 是一物理可观测过程。以 $t=t_0$ 为采样的初始时刻, T 为采样间隔对 g 采样, 得到时间序列 $\{x_n\}$, 其中 $x_n=x(t_0+nT)=g(z(t_0+nT))$ 。则动力系统的原始吸引子轨道可由映射 $H:R^d \rightarrow R^m$ 得到。

收稿日期: 2007-11-30; 修回日期: 2008-03-10

作者简介: 陆扬 (1981-), 男, 江苏扬州人, 高级工程师, 工学博士, 从事水声信号处理研究。

通讯作者: 陆扬, E-mail: CG300@sina.com.cn

m 维空间轨道上的点表示为： $x_n=H(z_n)=(x_n, x_{n+1}, \dots, x_{n+(m-1)\tau})$ 。其中， m 为正数，称为嵌入维数， τ 称为时间延迟。对连续的动力学系统来说，时间延迟是 τT 。正确地选择重构参数 m 和 τ ，能准确重构反映原动力学系统特征的相空间。

3 相空间重构时延参数选取

延迟时间的选择应该遵循这样的标准：吸引子要完全展开，即吸引子在相空间沿各个方向上的扩张应该大致相同。常用的有自相关函数法，互信息法 (Mutual Information) 等。

3.1 自相关函数法

在高维情况下，为了通过超球面逼近一个点集，原则上在所有大于或等于 τ 的时延处，自相关函数应该等于零。因此，可以选择时间序列自相关函数过零点作为相空间重构的时延参数。

若 $x(n)$ 为一由测量系统得到的时间序列 ($n=1, 2, \dots, N$)，则归一化的时间序列的自相关函数定义如下：

$$C(\tau) = \frac{\sum_{m=1}^M [x(m+\tau) - \bar{x}][x(m) - \bar{x}]}{\sum_{m=1}^M [x(m) - \bar{x}]^2} \quad (1)$$

式中

$$\bar{x} = \frac{1}{M} \sum_{m=1}^M x(m) \quad (2)$$

M 为分析的样本的长度， \bar{x} 为样本的均值， τ 为时延值。

3.2 平均互信息法

Fraser 应用 Shannon 的信息理论^[4]，引入了平均互信息的概念 (Average Mutual Information)，并以平均互信息的第一局部最小值作为最佳延时 τ 的选取。对于观测时间序列进行计算时，取 $s(n)$ 为 S 系统的元素， $s(n)$ 的延时 $s(n+T)$ 为 Q 系统的元素，则 S, Q 两系统的平均互信息为：

$$I(\tau) = \sum_{n=1}^N P(s(n), s(n+\tau)) \cdot \lg \left[\frac{P(s(n), s(n+\tau))}{P(s(n)) \cdot P(s(n+\tau))} \right] \quad (3)$$

且 $I(\tau) \geq 0$ 。

特别是当 S, Q 两系统完全独立时，有：

$P(s(n), s(n+\tau)) = P(s(n)) \cdot P(s(n+\tau)) \Rightarrow I(\tau) = 0$ (4)
计算 $I(\tau)$ 的关键在于二维概率密度的求取，当用直

方图进行求取时，盒子尺寸大小对计算结果影响很大。对于互信息的计算，Fraser 等给出了一种等概率划分空间格子^[4]，这种方法原理上不易掌握，计算繁琐，不便于应用。采用另一种常用的方法，即等间距划分空间格子^[5]的方法进行二维概率密度的求取^[5]。

3.3 实测噪声数据时间延迟参数计算

分别用自相关法和互信息法对 3 种目标噪声数据进行计算，每种目标取 3 个样本，数据长度为 5000，采样率为 25kHz。

图 1 为辐射噪声原始数据，其中图 1(a)、1(b)、1(c) 分别对应目标 A、B、C。

图 2~4 给出了 A、B、C 3 种目标辐射噪声平均互信息法和自相关法的计算结果。图中时间延迟的单位为样本的采样间隔。

三种目标平均互信息值第一次最小点比较集中，而用自相关法计算的结果比较分散。

由于自相关法计算结果仅表示两变量间的线性独立性，而平均互信息法计算值能够较好地反映两变量间的非线性相关性。因此采用平均互信息第一次最小值作为相空间重构的时间延迟。

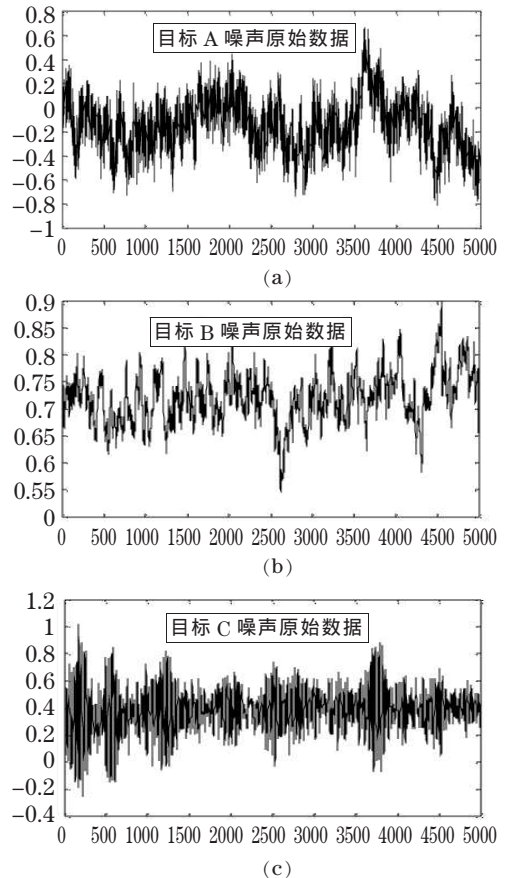


图 1 目标 A、B、C 辐射噪声原始数据
Fig.1 Original data of target A, B and C

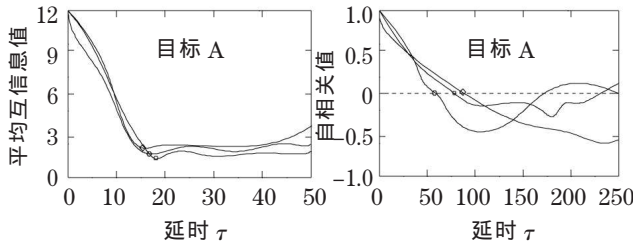


图2 目标 A 平均互信息法及自相关法计算结果

Fig.2 Calculated results of average mutual information and autocorrelation for target A

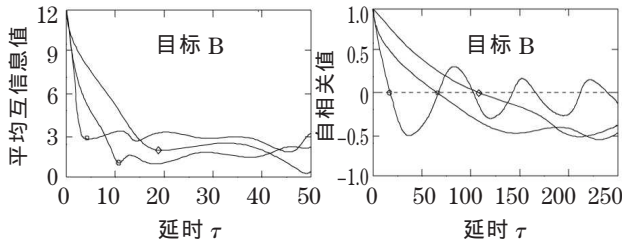


图3 目标 B 平均互信息法及自相关法计算结果

Fig.3 Calculated results of average mutual information and autocorrelation for target B

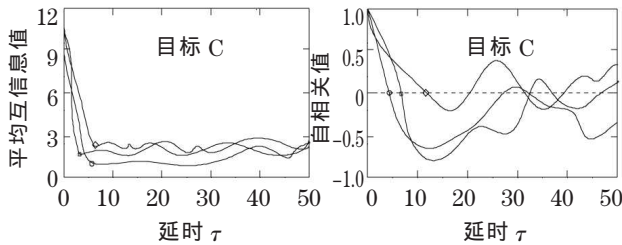


图4 目标 C 平均互信息法及自相关法计算结果

Fig.4 Calculated results of average mutual information and autocorrelation for target C

目标 A 重构时间延迟取 17 个样本点, 目标 B 取 11, 目标 C 取 6。

4 相空间重构嵌入维数选取

4.1 伪最近邻法嵌入维数选取

计算吸引子的分形维数, 取分形维数随嵌入维数的增加开始收敛的值为嵌入维数, 这种方法不客观, 并且计算量非常巨大。而奇异值分解法对计算精度要求很高, 计算精度对计算结果的影响太大。功率谱上限截止频率法物理意义不明显。采用伪最近邻法选取相空间重构嵌入维, 具体运算过程可用数学公式表述如下:

令 $x(n)$ 为观测的时间序列, τ 为延时。根据 Takens 定理^[3], 在 d 维相空间中形成的 d 维矢量为

$$Y(n) = \{x(n), x(n+\tau), \dots, x(n+(d-1)\tau)\} \quad n=1, 2, \dots, N \quad (5)$$

取相空间中一点 $Y(k) = \{x(k), x(k+\tau), \dots, x(k+(d-1)\tau)\}$, 求出其最邻近点 $Y^{NN}(k)$ (取欧氏距离作为为距离尺度), 令 $R_d(k)$ 为 $Y(k)$ 与 $Y^{NN}(k)$ 的距离, 则:

$$R_d^2(k) = [x(k) - x^{NN}(k)]^2 + [x(k+\tau) - x^{NN}(k+\tau)]^2 + \dots + [x(k+(d-1)\tau) - x^{NN}(k+(d-1)\tau)]^2 \quad (6)$$

维数由 d 增加到 $d+1$ 时, $d+1$ 维中 $Y(k)$ 与 $Y^{NN}(k)$ 的距离增加到 $R_{d+1}(k)$:

$$R_{d+1}^2(k) = R_d^2(k) + [x(k+d\tau) - x^{NN}(k+d\tau)]^2 \quad (7)$$

当以下两个准则满足其一时, 则称该两近邻点为假最近邻点。

准则(1):

$$\frac{x(k+d\tau) - x^{NN}(k+d\tau)}{R_d(k)} > R_T \quad (8)$$

$$\text{准则(2): } \frac{R_{d+1}(k)}{R_d(k)} > A_{tol} \quad (9)$$

$$R_A^2(K) = \frac{1}{N} \sum_{k=1}^N [x(k) - \frac{1}{N} \sum_{k=1}^N x(k)]^2 \quad (10)$$

通常, $10 \leq R_T \leq 50$, $A_{tol} = 2$ 。准则(2)为两近邻点距离与吸引子本身尺寸大小相比较得出的。

4.2 改进的伪最近邻法

每个相点最近邻点的确定, 是基于以下假设: 点对之间的分离距离是由重构相空间的几何特性产生的, 而不是由于点对的时间上的邻近造成的。然而, 现实情况是: 许多相点的最近邻点是时间上的邻近点 (在同一条轨道上), 而不是相空间中的最近邻点 (在相邻轨道上)。因此, 必须将这些由于时间上的邻近引起的最近邻点剔除。只有这样, 用最近邻点法进行嵌入维数的选取计算才更为合理, 得到的结果也更为可信。

时间上的邻近引起的关联需用 Theiler Window 来加以区分。在点对距离的计算中, 对于下标为 i, j 的点对, 需满足:

$$|i-j| \geq W \quad (11)$$

其中 W 为 Theiler Window 的数值。通过 Theiler Window 就可以剔除由于时间上邻近而导致的空间距离邻近的点对。

对于一维时间序列 $x_n = x_1, x_2, \dots, x_N, x(n)$ 为重构相空间中的相点, 设每一对点对在时间上的间隔为 Δn , 空间上的距离为 ε 。在时间上间隔为 Δn 的点对中, 定义相空间中距离小于 ε 的点对所占的百分数为 $P(|x(n+\Delta n) - x(n)| \leq \varepsilon)$ 。计算中, 固定 P 为一系列的值, 从 0 到 1 以 0.2 的间隔增长, 对每个 P 值, 绘出空间距离 ε 对于时间间隔 Δn 的关系曲线, 则所有 P 值对应的曲线构成了以空间上的距离 ε

为纵轴、时间上的间隔 Δn 为横轴的时空分离曲线图。当时间间隔较小时,点对总是呈现出空间上的邻近。随着时间间隔的增加,相空间中的点对开始扩散,点对的分布趋于水平。以此时的 Δn 值作为 Theiler Window 的参量 W 值。需要剔除的就是时间间隔小于 Theiler Window 的参量 W 时的点对。

4.3 嵌入维数计算实例

选择目标噪声 5000 点数据,对数据进行归一化处理。选择重构参数嵌入维数 $m=6, m=6$,时间延迟 $\tau=7$ 进行相空间重构,固定 P 为 0.2、0.4、0.6、0.8。绘制时空分离曲线,横轴为时间间隔 Δn ,纵轴为 $\lg \varepsilon$, ε 为相空间中点对的距离。为计算方便,不采用欧氏距离,而是将点对之间的距离定义为两个相点的最大矢量的差。

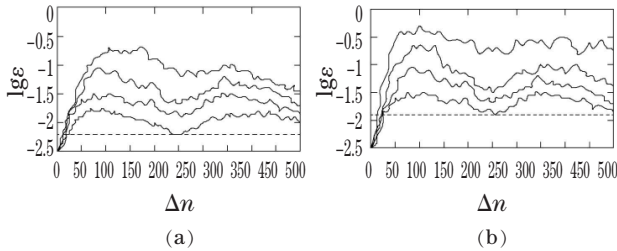


图 5 嵌入维数为 6,8 时的时空分离曲线

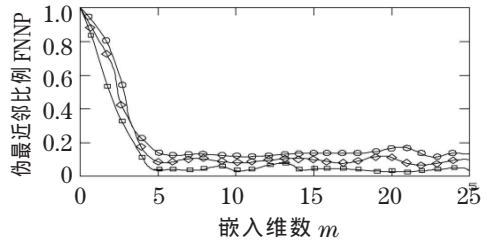
Fig.5 Time-space separate curve for $m=6$ and $m=8$

图 5(a)和 5(b)分别为重构嵌入维数为 6,8 时的时空分离曲线,曲线在 Δn 为 250 点处出现极小值,可以认为这个极小值就是排除了时间上的相邻的时空分离情况。以这个极小值绘出水平线,两图中与时空分离曲线相交在时空曲线第一次振荡前的饱和值 MST (minimum saturate time scale) 都在 25 左右,当 $\Delta t \leq MST$ 时,时空曲线增加很快;当 $\Delta t \geq MST$ 时,稳定的周期振荡占据了整个系统。按照 Theiler 理论,至少应该放弃 MST 个时间步长以内的点。考虑一定的冗余,本文选择 $W=2MST$,即 50。

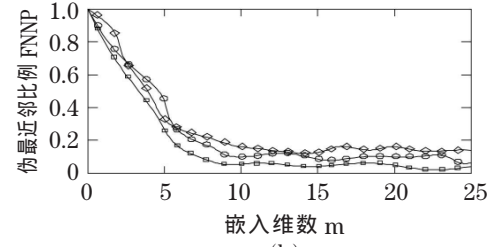
剔除时间步长 50 以内的点对,再依据 4.1 节的计算方法对相空间吸引子的每个点求其最近邻点,最后统计伪最近邻点在总的吸引子矢量点中的比例,判定嵌入维数(结果见图 6)。

图 6(a)、6(b)、6(c)分别给出了 A、B、C 3 种目标的伪最近邻比例计算结果,其中每种目标 3 个样本。

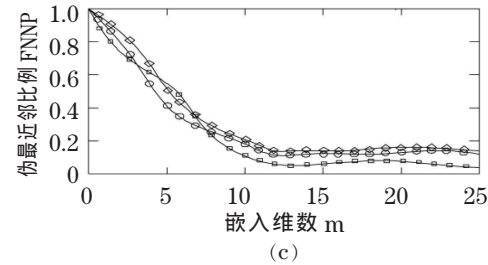
根据伪最近邻比例曲线第一次最小值确定嵌入维数,可以得到嵌入维数计算结果:目标 A 嵌入维数为 5;目标 B 嵌入维数为 10;目标 C 嵌入维数为 13。



(a)



(b)



(c)

图 6 目标 A、B、C 的伪最近邻计算结果

Fig.6 FNNP calculated Results of target A,B and C

5 结束语

本文通过自相关函数法和平均互信息法求取相空间重构延迟参数,分析比较了实测数据计算结果,确定了 3 类水中目标辐射噪声的重构延迟参数。对伪最近邻点法计算嵌入维数的算法作了改进,并将改进后的算法应用于实测噪声数据。从原理及实测数据计算结果来看,改进的伪最近邻方法能够确定较为合理、可信的相空间重构嵌入维数。因用于计算的目标噪声数据的数量和类型有限,该方法及所得到的一些结论尚需使用大量其他类型的目标数据作进一步的验证研究。

参 考 文 献

- [1] 章新华. 基于智能信息处理理论的水下目标识别研究[D]. 浙江大学博士论文, 1996.
- [2] 吕金虎, 陆君安, 陈士华. 混沌时间序列分析及其应用[M]. 武汉: 武汉大学出版社, 2002.
- [3] Takens F. Determing strang attractors in turbulence[J]. Lecture notes in Math, 1981, (898): 361-381.
- [4] Fraser A M, Swinney H L. Independent coordinates for strange attractors from information[J]. Phys. Rev. 1986 A, **33**: 1134.
- [5] Nichols J M, Nichols J D. Attractor reconstruction for non-linearsystem[J]. Mathematical Biosciences 2001, (171): 21.