

文章编号: 1001-0920(2011)05-0789-05

## 基于小波变换的城市市区典型路段交通流量预测

马林才<sup>1,2</sup>, 许玮珑<sup>2</sup>, 刘大学<sup>2</sup>

(1. 浙江大学 动力机械及车辆工程研究所, 杭州 310027; 2. 浙江交通职业技术学院 汽车学院, 杭州 311112)

**摘要:** 提出一种基于“先分-独立预测-再合”思想的新型交通流量预测模型. 其应用小波变换技术, 对交通流量数据进行分解和重构, 将得到的低频概貌信号和高频细节信号, 分别用3次平滑指数和 ARIMA 进行拟合, 并应用逐步回归方法建立基于原始交通流量数据, 以及高、低频拟合信号的回归方程. 对某城市市区典型路段交通流量数据进行验证分析, 结果表明, 该方程具有较高的拟合精度, 优于 ARIMA 单独使用的效果.

**关键词:** 小波变换; 交通流量; 自回归求和滑动平均; 回归分析

中图分类号: U491

文献标识码: A

### Prediction model of traffic flow along typical roads in city urban district based on wavelet transform

MA Lin-cai<sup>1,2</sup>, XU Wei-long<sup>2</sup>, LIU Da-xue<sup>2</sup>

(1. Institute of Power Machinery and Vehicular Engineering, Zhejiang University, Hangzhou 310027, China; 2. Automobile School, Zhejiang Institute of Communications, Hangzhou 311112, China. Correspondent: MA Lin-cai, E-mail: lincaima@sohu.com)

**Abstract:** A new kind prediction model of traffic flow based on the thought “decomposition-independently prediction-combination” is described in this paper. First, the traffic flow data are decomposed and reconstructed by using the wavelet transform technique. And the low frequency outline signal and high frequency detail signal are fitted respectively by using the three time exponential smoothing method and ARIMA. Then stepwise regression technology is used to establish the regression equation, in which the primitive traffic flow data are the dependent variable, and high, low-frequency signal's fitting data are the independent variables. The confirmation analysis is carried on with traffic flow data from typical roads in some city urban districts. The results show that the regression equation has a high fitting precision, and is better than the ARIMA function.

**Key words:** wavelet transform; traffic flow; ARIMA; regression analysis

## 1 引言

城市交通流量预测是智能运输系统中的路径诱导和交通流控制的必要条件, 交通流量信息的实时性和可靠性直接关系到交通管理与控制的效果, 因此短时交通流量预测是一个研究热点. 目前交通流量的预测方法很多, 如简单移动平均、自回归求和滑动平均 (ARIMA) 以及 Kalman 滤波等<sup>[1]</sup>, 但都有优缺点. 为了发挥各种预测方法的优点, 对时间序列预测的研究主要集中在各种方法的组合预测模型上, 如: ARIMA 模型与神经网络组合预测, 灰色预测与支持向量机的组合预测, 神经网络与支持向量机的组合预测等<sup>[2]</sup>. 同时, 基于小波理论的组合预测模型也被陆续提出,

如小波理论与时间序列的结合, 小波包与最小二乘支持向量机结合, 小波理论与卡尔曼滤波的结合等<sup>[3]</sup>. 显然, 组合预测模型利用了各类模型的优点, 预测结果更为理想, 但在使用时显得复杂.

本文提出一种基于“先分-独立预测-再合”思想的新型实用预测模型: 先进行小波变换; 再进行3次平滑指数和 ARIMA 独立预测; 最后进行逐步回归建立回归方程, 从而对城市交通流量进行有效预测. 结果表明, 该模型简单实用, 易于操作, 预测精度较为明显.

## 2 基本理论

### 2.1 小波变换简介

小波思想可追溯到1910年 Harr 提出的小波规范

收稿日期: 2010-03-05; 修回日期: 2010-07-23.

基金项目: 浙江省交通运输厅科技基金项目(2008T37).

作者简介: 马林才(1975—), 男, 副教授, 博士生, 从事汽车运用工程的研究; 许玮珑(1973—), 男, 讲师, 硕士, 从事交通工程、场站规划等研究.

正交基;但直到 1980 年, Morlet 引入了小波的概念,对地震信号进行分析和处理,并与 Graussman 共同提出了连续小波变换的体系. 1986 年, Meyer 第 1 次构造了具有一定衰减性的光滑小波,其二进伸缩平移可构成  $L^2(R)$  空间的规范正交系. 1987 年, Mallat 利用多分辨率分析的概念统一了以前各种小波的构造方法,并提出了著名的 Mallat 金字塔型快速算法<sup>[4-5]</sup>,其地位不亚于快速傅立叶变换在经典信号分析中的地位.

小波变换的含义是:把某一被称为基本小波(也叫母小波)的函数  $\psi(t)$  作位移  $\tau$  后,再在不同尺度  $a$  下与待分析信号  $x(t)$  作内积,即

$$WT_x(a, \tau) = \langle x(t), \psi_{a,\tau}(t) \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \cdot \psi \left[ \frac{t-\tau}{a} \right] dt. \quad (1)$$

其中:  $a > 0$  是尺度因子;  $\tau$  代表位移,其值可正可负;  $\psi_{a,\tau}(t) = \frac{1}{\sqrt{a}} \cdot \psi \left[ \frac{t-\tau}{a} \right]$  是基本小波的位移与尺度伸缩. 式(1)中不但  $t$  是连续变量,而且  $a$  和  $\tau$  也是连续变量,因此,称其为连续的小波变换(CWT).

小波变换具有多分辨(也叫多尺度)的特点,可以由粗及精地逐步观察信号.适当地选择基本小波(如: Morlet 小波, Haar 小波, Daubechines 小波, 样条小波等),使之在时域上为有限支撑,在频域上也比较集中,这样便可使小波变换在时、频两域都具有表征信号局部特性的能力,而且小波变换经适当离散化后能构成标准正交系,有利于检测信号的奇异点或瞬态.由于上述特性,有人把小波变换誉为:分析信号的数学显微镜.

将尺度与位移离散化之后得到离散栅格的小波变换.通常按 2 的幂级数进行离散化,即取  $a = 2^j$ ,  $j \in Z$ ; 位移取  $\tau = k \cdot 2^j T_s$ . 为简化起见,设  $T_s = 1$ , 便有

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi \left[ \frac{t - 2^j k T_s}{2^j} \right] = 2^{-(j/2)} \psi(2^{-j}t - k),$$

将在这些点上计算得到的小波变换记作

$$WT_x(j, k) = \int x(t) \psi_{j,k}^*(t) dt. \quad (2)$$

在二分的情况下, Mallat 从函数的多分辨率空间分解概念出发,在小波变换和多分辨率之间建立起联系.即:把平方可积函数空间  $x(t) \in L^2R$  作逐级二分分解,产生一组逐级包含的子空间.这些子空间之间有逐级包含、逐级替换等特性,且有

$$P_{j-1}x(t) = \sum_k x_k^{(j-1)} \phi_{j-1,k}(t),$$

$$x_k^{(j-1)} = \langle x(t), \phi_{j-1,k}(t) \rangle; \quad (3)$$

$$P_j x(t) = \sum_k x_k^j \phi_{j,k}(t),$$

$$x_k^{(j)} = \langle x(t), \phi_{j,k}(t) \rangle. \quad (4)$$

其中:  $P_j x(t)$  是  $x(t)$  在  $V_j$  中的平滑逼近,即  $x(t)$  在分

辨率  $j$  下的概貌;  $x_k^{(j)}$  是其在分辨率  $j$  下的离散逼近. 又有

$$D_j x(t) = \sum_k d_k^j \psi_{j,k}(t),$$

$$d_k^{(j)} = \langle x(t), \psi_{j,k}(t) \rangle; \quad (5)$$

$$P_{j-1}x(t) = P_j x(t) + D_j x(t). \quad (6)$$

其中:  $D_j x(t)$  是  $x(t)$  在  $W_j$  中的投影,反映  $P_{j-1}x(t)$  和  $P_j x(t)$  两级平滑逼近之间的细节差异.因此,称  $D_j x(t)$  为分辨率  $j$  下的离散细节,而其离散值  $d_k^j$  便是小波变换  $WT_x(j, k)$ .

## 2.2 ARIMA 模型

对于平稳时间序列,自回归移动平均模型 ARMA 可提供精确度较高的短期预测.对于如交通流量数据等非平稳的时间序列,却不能直接使用 ARMA 模型,需通过差分后将原序列转换为平稳时间序列才可使用,这便涉及到另外一个模型——ARIMA 模型,即:如果  $w_t = \Delta^d y_t$  是平稳序列,则称  $y_t$  是  $d$  阶齐次非平稳序列,这里  $\Delta$  表示差分.如果有序列  $w_t$ ,则可通过对  $w_t$  进行  $d$  次求和得到  $y_t$ ,即  $y_t = \sum_{i=0}^{d-1} w_{t-i}$ . 求和算子和差分算子是互逆运算.如果  $w_t = \Delta^d y_t$  且  $w_t$  是一个 ARMA( $p, q$ ) 过程,则称  $y_t$  是 ( $p, d, q$ ) 阶自回归求和和滑动平均过程,简称为 ARIMA( $p, d, q$ )<sup>[1]</sup>.

在确定 ARIMA( $p, d, q$ ) 的阶时,可依据自相关函数和偏相关函数的截尾性和拖尾性进行识别.对参数进行估计之后,应通过检验模型残量是否为白噪声来判断模型是否合理.

## 2.3 平滑指数法模型

平滑指数模型是时间序列预测法中的一种基本预测模型,在该模型中,平滑指数  $\alpha$  的取值非常关键.本文采用 3 次平滑指数模型:一般数据接近稳定时,取  $\alpha = 0.05 \sim 0.20$ ; 数据序列波动不大,长期趋势变动缓慢时,取  $\alpha = 0.3 \sim 0.5$ ; 数据序列波动较大,长期趋势变动快捷、明显时,取  $\alpha = 0.6 \sim 0.9$ .

## 2.4 逐步回归模型

在回归分析中,逐步回归是其中的一种优秀的变量筛选技术.对于常见的二次型回归方程,并不是所有系数均显著,如何选择贡献显著的项,删除不重要的项,即为变量筛选技术.逐步回归是预先确定入选和淘汰阈值,用于决定变量能否入选或剔除.筛选过程中不仅要选入变量,而且要将已选进方程的变量进行比较,删除对因变量影响不明显的变量.如此反复地选进和剔除,直到回归方程和系数均显著为止<sup>[6]</sup>.

## 3 基于小波变换的交通流量预测步骤

由信号和系统理论可知,信号的高频分量反映

信号的变化, 低频分量则反映信号的趋势. 本文选用 Daubechies 小波基, 将交通流量信号用小波变换分解为低频成分与高频成分来进一步观察和处理, 即

$$\phi(t) = \sqrt{2} \sum_{n=0}^{2N-1} h_n \phi(2t - n), \quad (7)$$

$$\psi(t) = \sqrt{2} \sum_{n=0}^{2N-1} g_n \phi(2t - n). \quad (8)$$

其中  $g_n = (-1)^n h_{2N-n-1}, n = 0, 1, \dots, 2N - 1$ .

这种小波具有正交性和紧支撑性, 可采用快速算法进行计算, 而且可通过有限冲激响应(FIR)滤波器来实现.

通过对交通流量进行小波变换的基础上, 将交通流量时间序列分解为多个分量, 对各信号分量分别进行预测, 再合成最终的预测结果<sup>[7-9]</sup>, 极大地提高了预测准确度. 具体步骤如图1所示.

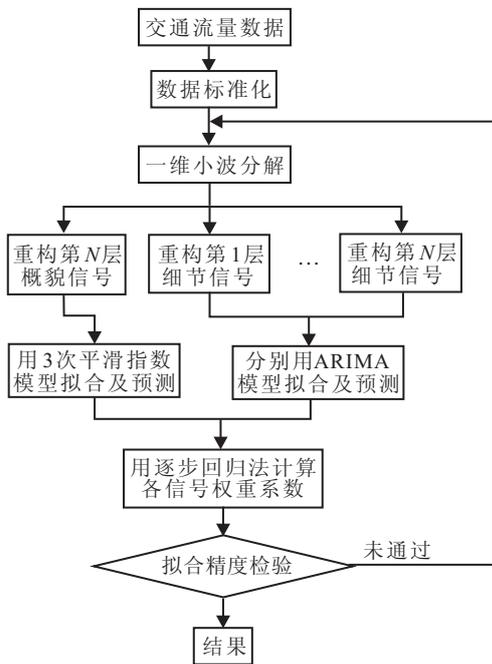


图1 基于小波变换的交通流量组合预测步骤

### 4 城市市区典型路段交通状况调查与分析

#### 4.1 交通状况调查实施办法

本次调查以杭州市为例, 选取的典型路段为: 路段1: 莫干山路(萍水路-申花路路段, 长度为703 m); 路段2: 环城北路(建国路-环城东路路段, 长度为682 m), 路段3: 西溪路(紫荆花路-东方中学路段, 长度为730 m). 调查时间为: 2009-02-20的06:30-19:30. 调查项目为: 为期13 h的双向分车型流量, 以15 min为间隔记录. 本文中的算法以路段2的环城北路西向东的交通流量数据为试验数据, 其他路段数据为检验数据.

#### 4.2 城市市区典型路段交通流量统计

根据《公路工程技术标准》和《城市道路设计标

准》等标准, 将城市交通流中摩托车、小轿车、大中小客车、大中小货车等换算出标准小汽车当量.

### 5 应用过程

采样的交通流量数据样本共52个. 前48个流量数据用于建立模型, 后4个数据用于模型的预测实验. 图2为前48个原始交通流数据图.

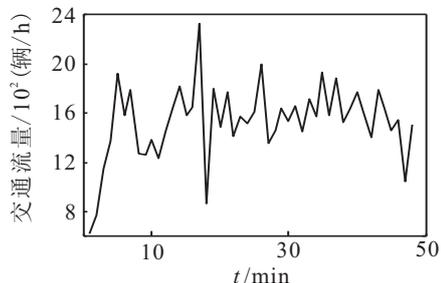


图2 前48个交通流量数据图

首先, 对交通流量数据进行小波分解, 选用的小波为“dB4”, 分解层次为2; 然后对各层次的分量进行重构. 图3为前48个交通流数据的小波重构结果.

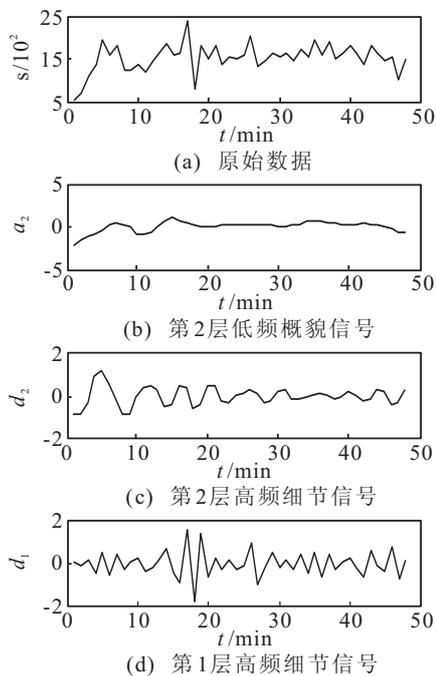


图3 交通流量标准化数据的重构

其次, 第2层的低频概貌信号, 采用3次平滑指数模型, 其中  $\alpha$  取0.9. 对于第1到第2层高频细节信号采用 ARIMA 模型. 表1列出了高频细节信号的自相关函数和偏相关函数值. 根据赤池信息准则(AIC)定阶原则, 第1层和第2层偏相关函数值分别是9, 4后截尾; 自相关函数值分别是1, 3后拖尾. 因此, 在第1层高频细节信号中, 采用  $p = 9, q = 1$  作为模型的阶; 在第2层高频细节信号中, 采用  $p = 4, q = 3$  作为模型的阶. 即: 对于第1层高频细节信号采用 ARIMA(9, 0, 1) 模型, 对于第2层高频细节信号采用 ARIMA(4, 0, 3) 模型, 分别进行拟合和对后4个信号的预测.

表1 高频细节信号的自相关函数和偏相关函数值

项数	第1层高频细节信号		第2层高频细节信号	
	自相关函数值	偏相关函数值	自相关函数值	偏相关函数值
1	-0.737	-0.737	0.399	0.399
2	0.293	-0.549	-0.435	-0.707
3	-0.066	-0.426	-0.554	0.058
4	0.012	-0.343	-0.224	-0.411
5	-0.012	-0.327	0.057	-0.116
6	-0.022	-0.511	0.181	-0.151
7	0.150	-0.254	0.141	-0.209
8	-0.263	-0.317	0.012	-0.127
9	0.247	-0.316	-0.011	0.008
10	-0.108	0.033	0.036	-0.081
11	-0.107	-0.226	-0.039	-0.136
12	0.277	-0.053	-0.107	0.027

再次,令因变量 $Y$ 为交通流量标准化值,自变量 $X_1$ 为第1层高频细节拟合信号值,自变量 $X_2$ 为第2层高频细节拟合信号值,自变量 $X_3$ 为第2层低频概貌拟合信号值.经逐步回归,得到如下方程:

$$Y = -0.045 + 1.002X_1 + 1.038X_2 + 1.283X_3. \quad (9)$$

方程(9)的调整相关系数 $R^2 = 0.915$ ,标准误差 $\delta^* = 0.305$ , $F = 157.083 > F_{0.975}(3, 44) = 3.45$ .经检验,方程和系数均显著可信.

最后,将各信号拟合值和预测值代入方程(9)中,可得到交通流量的拟合及预测结果,即小波变换的拟合和预测值.与单纯采用ARIMA模型直接预测相比,其预测效果显著提升,见图4和图5.表2为采用平均绝对误差和均方误差两个指标进行拟合和预测效果的比较.

由图4,图5和表2的对比可以发现,与单独使用

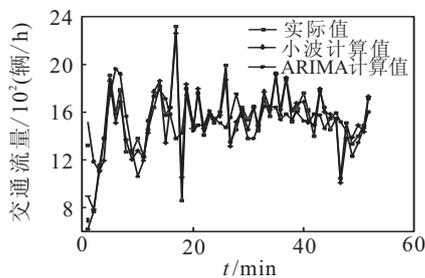


图4 小波变换与传统ARIMA模型的拟合与预测效果对比

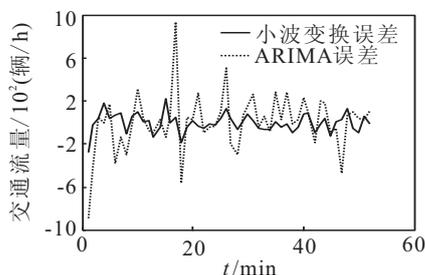


图5 小波变换与传统ARIMA模型的拟合与预测误差对比

表2 小波变换的拟合和预测效果评价

评价 指标	计算 公式	小波预 测误差	传统ARIMA 预测误差
平均绝对误差	$MAE = \frac{1}{n} \sum_{t=1}^n  Y_t - \hat{Y}_t $	61.74	177.19
均方误差	$MSE = \frac{1}{n} \sqrt{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}$	11.83	37.48

ARIMA模型相比,小波变换后分层拟合精度更高.

## 6 结 论

1) 应用小波变换原理,对原始交通数据进行了分解和重构.分别采用3次平滑指数拟合低频概貌信号,自回归求和滑动平均(ARIMA)拟合高频细节信号,逐步回归技术对交通流量、高、低频信号进行回归分析.该方法拟合精度高、可操作性强,更为简单实用.

2) 与单独使用ARIMA模型拟合预测相比,小波变换后,充分运用不同模型的优势,进行分层拟合预测,可获得更好的时间序列数据预测精度.

3) 本文方法的核心思想是“先分-独立预测-再合”.核心算法是“小波变换-常用预测算法-回归分析”.

## 参考文献(References)

- [1] 窦慧丽,刘好德,吴志周,等.基于小波分析和ARIMA模型的交通流预测方法[J].同济大学学报(自然科学版),2009,37(4):486-494.  
(Dou H L, Liu H D, Wu Z Z, et al. Study of traffic flow prediction based on wavelet analysis and autoregressive integrated moving average model[J]. J of Tongji University (Natural Science), 2009, 37(4): 486-494.)
- [2] 张晓利.基于小波分析与神经网络的交通流短时预测方法[J].信息与控制,2007,36(4):467-475.  
(Zhang X L. The forecasting approach for short-term traffic flow based on wavelet analysis and neural network[J]. Information and Control, 2007, 36(4): 467-475.)
- [3] 谭满春,李英俊,徐建闽.基于小波消噪的ARIMA与SVM组合交通流预测[J].公路交通科技,2009,26(2):127-138.  
(Tan M C, Li Y J, Xu J M. A hybrid ARIMA and SVM model for traffic flow prediction based on wavelet denoising[J]. J of Highway and Transportation Research and Development, 2009, 26(2): 127-138.)
- [4] 秦中元,牟轩沁,洪伟.基于小波变换的交通车流密度分析[J].公路交通科技,2001,18(1):56-59.  
(Qin Z Y, Mu X Q, Hong W. Analysis of traffic density based on wavelet transform[J]. J of Highway and Trans on Research and Development, 2001, 18(1): 56-59.)

- [5] 杨福生. 小波变换的工程分析与应用[M]. 北京: 科学出版社, 2006.  
(Yang F S. The engineering analysis and application of wavelet transform[M]. Beijing: Science Press, 2006.)
- [6] 王毅, 汪洋, 蔡华. 基于小波变换的逐步线性回归分析预测大坝位移[J]. 测绘信息与工程, 2008, 33(4): 16-17.  
(Wang Y, Wang Y, Cai H. Dam displacement forecast with stepwise regression analysis based on wavelet transform[J]. J of Geomatics Aug, 2008, 33(4): 16-17.)
- [7] 杨芳明, 朱顺应. 基于小波的短时交通流预测[J]. 重庆交通学院学报, 2006, 25(3): 99-102.  
(Yang F M, Zhu S Y. Short-term traffic flow forecasting based wavelet[J]. J of Chongqing Jiaotong University, 2006, 25(3): 99-102.)
- [8] 李存军, 杨儒贵, 张家树. 基于小波分析的交通流量预测方法[J]. 计算机应用, 2003, 23(12): 7-8.  
(Li C J, Yang R G, Zhang J S. Traffic prediction based on wavelet analysis[J]. Computer Applications, 2003, 23(12): 7-8.)
- [9] 陈淑燕, 王炜, 瞿高峰. 短时交通量时间序列的小波分析模糊马尔柯夫预测方法[J]. 东南大学学报(自然科学版), 2005, 35(4): 637-640.  
(Chen S Y, Wang W, Qu G F. Short-term traffic flow time series forecasting based on wavelet analyses-fuzzy Markov prediction model[J]. J of Southeast University(Natural Science Edition), 2005, 35(4): 637-640.)

(上接第784页)

- [8] Rabbat M, Nowak R. Distributed optimization in sensor networks[C]. Proc of the 3th Int Symposium on Information Processing in Sensor Networks. Berkeley: Association for Computing Machinery, 2004: 20-27.
- [9] Lin X, Johansson M, Boyd S. Simultaneous routing and resource allocation via qual decomposition[J]. IEEE Trans on Communications, 2004, 52(7): 1136-1144.
- [10] Vicente L, Masgrau E. Analysis of the desired-response influence on the convergence of gradient-based adaptive algorithms[J]. IEEE Trans on Circuits and Systems I: Regular Papers, 2008, 55(2): 1257-1266.
- [11] Hou Y T, Shi Y, Pan J, et al. Maximizing the lifetime of wireless sensor networks through optimal single-session flow routing[J]. IEEE Trans on Mobile Computing, 2006, 5(9): 1255-1266.

(上接第788页)

- [13] Han K H, Kim J H. Genetic quantum algorithm and its application to combinatorial optimization problem[C]. Proc of the 2000 IEEE Congress on Evolutionary Computation. La Jolla, 2000, 2: 1354-1360.
- [14] 罗利军, 李银山, 李彤, 等. 李雅普诺夫指数谱的研究与仿真[J]. 计算机仿真, 2005, 22(12): 285-288.  
(Luo L J, Li Y S, Li T, et al. Research and simulation of Lyapunov's exponents[J]. Computer Simulation, 2005, 22(12): 285-288.)
- [15] 陆振波. 混沌时间序列分析与预测工具箱V2.0[CP/OL]. [2009-06-14]. [http://luzhenbo.88uu.com.cn/program/chaosToolbox2p0\\_trial.rar](http://luzhenbo.88uu.com.cn/program/chaosToolbox2p0_trial.rar).  
(Lu Zhen-bo. Chaotic time series analysis and prediction matlab toolbox trial version 2.0 [CP/OL]. (2009-06-14). [http://luzhenbo.88uu.com.cn/program/chaosToolbox2p0\\_trial.rar](http://luzhenbo.88uu.com.cn/program/chaosToolbox2p0_trial.rar).)
- [16] 赵媛媛, 吴云洁. 基于实验数据的控制系统 Lyapunov 指数谱计算[J]. 计算机仿真, 2004, 21(11): 58-59.  
(Zhao Y Y, Wu Y J. Computation of Lyapunov exponents using experimental data of control system[J]. Computer Simulation, 2004, 21(11): 58-59.)
- [17] Djamai L, Coirault P. Estimation of Lyapunov exponents by using the perception[C]. Proc of the American Control Conf Anchorage. Alaska, 2002: 5150-5155.