

原子力显微镜中微悬臂梁分布参数系统的Hammerstein模型

徐运扬[†], 徐康康

(中南大学 高性能复杂制造国家重点实验室, 湖南 长沙 410012)

摘要: 为提高原子力显微镜(atomic force microscope, AFM)中微悬臂梁分布参数模型的精度, 本文提出了包含非线性时空特性的改进模型, 在此基础上简化控制器的结构. 首先加入非线性补偿项修正传统分布参数模型; 然后采用Karhunen-Loève(K-L)方法提取系统主导空间基函数, 实现系统输出的时空变量分离. 利用求解得到的时间系数和系统激励, 建立系统时域Hammerstein模型, 使系统无限维偏微分方程模型转化为时域有限维常微分方程形式, 控制器的设计无需考虑空间耦合的影响; 最后, 利用最小二乘支持向量机结合奇异值分解法辨识模型中的参数. 与传统分布参数模型进行仿真和实验结果比较, 验证了方法的有效性.

关键词: 原子力显微镜; 最小二乘向量机; Hammerstein模型; 分布参数; 系统辨识

中图分类号: TP18 **文献标识码:** A

Hammerstein model for distributed parameter system of micro-cantilever in atomic-force microscope

XU Yun-yang[†], XU Kang-kang

(State Key Laboratory of High Performance Complex Manufacturing, Central South University, Changsha Hunan 410012, China)

Abstract: To improve the accuracy of the distributed-parameter model of the micro-cantilever in an atomic-force microscope (AFM), we propose a modified model which contains the nonlinear spatial-temporal properties. On this basis, the structure of the controller is simplified and a nonlinear compensation term is added to correct the traditional distributed parameter model. Next, the Karhunen-Loève (K-L) decomposition method is applied to extract the dominant spatial basis function of the system, achieving the space-time decomposition. Then, a temporal Hammerstein model is identified by the temporal coefficients obtained from the decomposition and the input signal, which transforms the infinite dimensional partial differential equation model into the finite-dimensional ordinary differential equation model, making it possible to design the controller without considering the space coupling. Finally, we use the least-squares-support-vector-machine algorithm and the singular-value decomposition method to identify the parameters of the model. Simulation and experimental comparisons with the traditional distributed parameter model are given to validate the effectiveness of the proposed method.

Key words: atomic force microscope; least-squares vector machine; Hammerstein model; distributed parameter; system identification

1 引言(Introduction)

过去20年间, 纳米技术的飞速发展使包括医学、生物学和材料科学在内的诸多领域的研究进入纳米层面. 其中被广泛应用于纳米成像和操作的工具是原子力显微镜(atomic force microscope, AFM)^[1], 其发明对科技的发展产生了深远的影响.

AFM中的柔性微悬臂梁是扫描和成像系统中的关键部件^[2], 其一端固定在压电陶瓷驱动器上, 另一端带有针尖并悬空. 微悬臂梁属于分布参数系统(distributed parameter system, DPS), 此类对象的模型为偏微分方程形式, 其参数时空分布且有无穷维特性.

欧拉-伯努利梁分布参数方程是较常用的描述微悬臂梁动力学特性的模型^[3], 它的精度虽高于传统线性集中参数模型, 但其推导过程包含了一些假设与近似处理, 且仅能反应悬臂梁的线性动态特性. 实际的柔性微悬臂梁存在复杂的非线性特性和参数不确定性^[4-5], 这会影响微悬臂梁探针系统的定位和控制效果. 因此, 建立更精确的微悬臂梁可控模型对探针的准确定位和力检测十分重要, 这能进一步提高AFM扫描成像的精度.

对无穷维时空耦合系统的建模与控制, H. S. Tzou和J. H. Ding在论文[6]中采用空间分布的传感器、驱

收稿日期: 2014-06-23; 录用日期: 2014-11-06.

[†]通信作者. E-mail: xuyunyang2008@163.com; Tel.: +86 18670041623.

国家重点基础研究发展规划项目(“973”计划)(2011CB013104)资助.

Supported by National Program on Key Basic Research Project (“973” Program) (2011CB013104).

动器和控制器结合最优线性二次状态反馈控制算法实现了对弹性抛物面壳体的振动主动控制, 借助多个采样点来设计控制量. 这种在无穷维上直接设计控制器的方法精度高, 效果显著. 若利用时空变量分离和低维近似则可将控制量设定在时域内, 无需考虑空间信息, 简化控制器的结构, 便于应用. 实现这一目标的方法大体上有两类: 一类是基于机理的方法, 如谱方法^[7]; 另一类是基于数据的方法, 如正则化正交分解^[8-9](Karhunen-Loève, K-L), 以及奇异值分解等. 谱方法将时空变量表示为一系列的空间基函数与时间项系数的组合, 将原模型方程进行时间空间特征的分, 然后利用Galerkin截断^[10]将无穷维时空耦合系统降阶为有限维时域系统, 并基于低维的时域模型设计控制器. 但这种方法的局限性在于, 必须首先获得原分布参数系统的完整偏微分方程, 而实际情况下对于复杂工业过程, 精确机理模型难以建立. 针对谱方法的不足, 文献[11-12]提出一种完全基于数据的黑箱建模法, 利用K-L分解, 从大量的系统输出采样数据(snapshots)中提取出能反应系统特性的一组空间基函数, 进行输出变量时空分离, 再利用得到的时间项系数和系统输入建立时域模型, 从而在无需系统完整先验知识的前提下达到了模型降维和在时域设计控制器的目的.

柔性微悬臂梁的受迫振动中存在机理尚难以准确描述的非线性特性, 而传统的线性模型对于非线性的描述能力非常有限. 针对非线性分布参数系统的建模, 一类块结构的非线性动态模型成为新的研究热点^[13], 这类模型的结构包含了线性动态模块和无记忆的非线性静态模块, 模型简单, 参数辨识容易, 相比传统的线性动态模型, 其更能反应系统的实际特性, 模型精度更高. 按照块的连接形式, 可分为Hammerstein模型、Wiener模型, 以及组合形式的Hammerstein-Wiener, Wiener-Hammerstein模型. 由于Hammerstein模型非线性部分的计算无需过去的输入、输出信息, 其计算量少于Wiener模型, 因而Hammerstein模型的应用最为广泛, 本文采用的就是Hammerstein建模法.

对Hammerstein模型参数辨识的研究, 向微等^[14]将线性动态部分和非线性静态部分分别用Laguerre级数和非线性基表示, 辨识得到两部分的参数. Chen^[15]提出递推最小二乘法辨识对模型参数进行辨识. Krzyzak等^[16]使用递推核回归法和协相关法辨识Hammerstein模型的参数, 对模型的线性部分使用协相关系数求解, 模型的非线性部分由递推核回归法得到. Pelckmans和Goethal等^[17]使用最小二乘支持向量机对Hammerstein模型参数加以辨识, 采用冗余参数法以及增加约束, 辨识得到了线性部分和非线性部分的系数.

本文对AFM微悬臂梁非线性分布参数系统的建模

进行了研究, 将基于模块的非线性Hammerstein模型的应用扩展到分布参数系统中, 利用K-L分解提取出反应系统本身特性的空间基函数, 实现对系统输出的时空分离, 将时空分布参数系统建模转化为传统的时域建模问题, 使得后续的控制设计无需考虑空间信息, 简化了控制器的结构. 由于最小二乘支持向量机能很好地表示非线性过程^[18], 且其参数具有冗余性, 因此, 采用最小二乘支持向量机算法, 通过增加约束条件, 结合奇异值分解, 辨识得到了Hammerstein模型中的参数.

2 改进的非线性欧拉-伯努利梁模型(Improved nonlinear Euler-Bernoulli beam model)

2.1 传统欧拉-伯努利梁模型(Traditional Euler-Bernoulli beam model)

微悬臂梁是原子力显微镜扫描成像系统中的关键部件, 其一端固定在压电陶瓷上, 另一端和探针相连并悬空. 在扫描的过程中, 探针和样品表面间的时变原子力会使梁受迫振动, 产生动态弯曲变形. 微悬臂梁-探针系统结构示意图见图1所示.

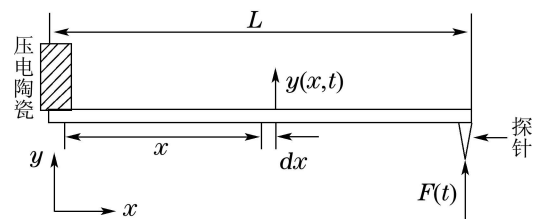


图1 AFM柔性微悬臂梁示意图

Fig. 1 Structure of flexible AFM micro-cantilever

柔性微悬臂梁是一个复杂的非线性分布参数系统, 其参数空间分布和柔性特性使其在外力激励下的动态特性存在非线性时空耦合. 完全基于理论得到的描述其动态特性的分布参数欧拉-伯努利梁方程^[19]

$$m \frac{\partial^2 y(x, t)}{\partial t^2} + c \frac{\partial y(x, t)}{\partial t} + EI \frac{\partial^4 y(x, t)}{\partial x^4} = -F(t)\delta(x - L) \quad (1)$$

满足如下边界条件:

$$y(0, t) = 0, \quad y_x(0, t) = 0, \quad (2)$$

$$y_{xx}(L, t) = 0, \quad EI y_{xxx}(L, t) = -F(t)\delta(x - L), \quad (3)$$

其中: m 是线密度, c 是阻尼常数, E 是弹性模量, I 是截面惯性矩, L 是悬臂梁的长度, $\delta(x - L)$ 是狄拉克函数, $F(t)$ 是针尖处所受原子力, $y(x, t)$ 是微悬臂上的质点相对于原平衡位置的形变量, 它是空间位置 and 时间的函数.

2.2 改进的欧拉-伯努利梁模型(Improved Euler-Bernoulli beam model)

以上模型的精度虽较传统线性集中参数模型高,

但其仅能反应微悬臂梁线性动态特性,实际的柔性微悬臂梁中还存在复杂的非线性特性和参数不确定性^[4]. 这会影响微悬臂梁探针系统的定位和控制精度. 因此,为提高微悬臂梁模型的精度,在此对原机理模型加以改进,使其能体现系统的非线性不确定性,模型形式如下:

$$m \frac{\partial^2 y(x, t)}{\partial t^2} + c \frac{\partial y(x, t)}{\partial t} + EI \frac{\partial^4 y(x, t)}{\partial x^4} + f(y(x, t)) = -F(t)\delta(x - L), \quad (4)$$

式中非线性函数 $f(\cdot)$ 是柔性微悬臂梁模型的非线性补偿项,它是位移 $y(x, t)$ 、空间位置 x 以及时间 t 的函数.

3 微悬臂梁分布参数系统Hammerstein模型 (Hammerstein model for micro-cantilever distributed parameter system)

3.1 微悬臂梁时空分布参数模型(Time-space distribution parameter model of micro-cantilever)

在传统Hammerstein模型中引入空间变量,可将Hammerstein模型的应用扩展到分布参数系统^[17],见图2所示. 其中: ζ, x 为定义在空间域 Ω 内的空间变量, ζ 是输入信号的空间位置, x 是输出位移的空间位置, $u(\zeta, t)$ 代表 t 时刻的输入, $y(x, t)$ 表示 t 时刻的输出, $e(x, t)$ 代表未建模的动态特性和随机扰动. $f(\cdot)$ 是静态非线性函数, $G(\cdot)$ 是线性动态部分.

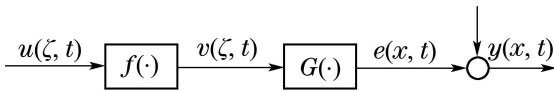


图2 时空分布参数系统Hammerstein模型

Fig. 2 Structure of spatial-temporal distributed parameter Hammerstein model

此时系统的输入输出关系可表示为如下形式:

$$y(x, t) = \sum_{\tau=0}^t \int_{\Omega} g(x, \zeta, \tau) f(u(\zeta, t - \tau)) d\zeta + e(x, t). \quad (5)$$

3.2 微悬臂梁时域Hammerstein模型 (Time domain Hammerstein model of micro-cantilever)

对分布参数系统而言,时空变量能用一个有限数目的空间基函数和时间项系数的乘积之和表示^[20]:

$$y(x, t) = \sum_{i=1}^n \phi_i(x) y_i(t). \quad (6)$$

用K-L方法对系统输出进行时空变量分离,进而可以利用时域输入输出信息建立微悬臂梁时域Hammerstein模型,建模思路如图3所示.

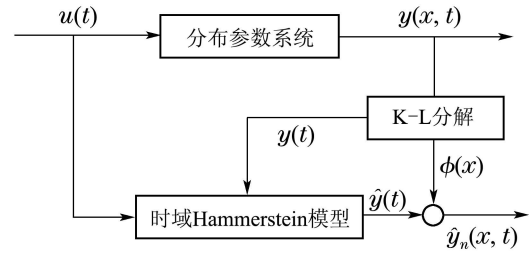


图3 微悬臂梁分布参数系统时域Hammerstein建模思路

Fig. 3 Temporal Hammerstein model methodology for micro-cantilever distributed parameter system

4 Karhunen-Loève分解法 (Karhunen-Loève expansion method)

K-L(Karhunen-Loève)分解法,又称正交分解技术或主成分分析,能以最小的自由度来表示一个随机域^[8]. 利用K-L技术对测得的数据进行时空正交分解,能从大量系统实测输出采样数据(snapshots)中提取出能反应系统特征的空间基函数. 基函数的形式有很多,如傅里叶级数、样条函数、Chebyshev多项式、Jacobi多项式、拉格朗日多项式等. 数学上已经证明,在给定模型精度的前提下,基于K-L分解获得的空间基函数数组能最优地描述实测数据集^[21].

K-L时空分离后的时域Hammerstein模型结构如图4所示.

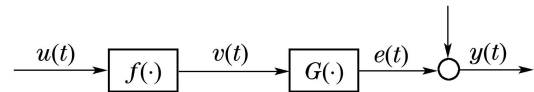


图4 非线性分布参数系统的时域Hammerstein模型

Fig. 4 Temporal Hammerstein model of nonlinear distributed parameter system

假设所得输出数据 $\{y(x_i, t)\}_{i=1, t=1}^{N, T}$ 在时间和空间域上均匀采样,定义全体平均值和内积的概念如下:

全体平均值:

$$\langle y(x, t) \rangle = \frac{1}{T} \sum_{t=1}^T y(x, t). \quad (7)$$

内积:

$$(y(x), g(x)) = \int_{\Omega} y(x) g(x) dx. \quad (8)$$

寻找最能体现采样信息特征的空间基函数 $\{\phi_i(x)\}_{i=1}^n$ 等同于求解如下目标函数的极小值问题:

$$\begin{aligned} \min & \langle [y(x, t) - \sum_{i=1}^n \phi_i(x) (\phi_i(x), y(x, t))]^2 \rangle \\ & (\phi_i(x), \phi_i(x)) = 1, \phi_i(x) \in L^2(\Omega), i = 1, \dots, n, \end{aligned} \quad (9)$$

式中: $\phi_i(x)$ 是系统空间基函数,内积 $(\phi_i(x), y(x, t))$ 的值就是第 i 阶基函数对应的时间项系数,约束条件 $(\phi_i(x), \phi_i(x)) = 1$ 保证了空间基函数的唯一性. 上式

的含义是时空重组数据和实测采样数据间误差平方的均值最小。

此约束优化问题可用拉格朗日条件极值法表示如下:

$$J = \langle [y(x, t) - \sum_{i=1}^n \phi_i(x)(\phi_i(x), y(x, t))]^2 \rangle + \sum_{i=1}^n \lambda_i ((\phi_i(x), \phi_i(x)) - 1). \quad (10)$$

方程有解的必要条件是

$$\frac{dJ(\phi + \delta\phi)}{d\delta}(\delta = 0) = 0, (\phi, \phi) = 0. \quad (11)$$

将式(7)–(8)(10)代入式(11)中, 并令 $\phi(x)$ 为任意函数, 则有解的必要条件可简化成

$$\int_{\Omega} R(x, \zeta)\phi_i(\zeta)d\zeta = \lambda_i\phi_i(x), (\phi, \phi) = 1, \quad (12)$$

其中 $R(x, \zeta) = \langle (y(x, t), y(\zeta, t)) \rangle$ 是空间两点间的相关方程, 它是对称正定的. **K–L**分解的核心问题就是计算特征值和对应的特征向量. λ_i 是第*i*阶空间基函数对应的特征值.

特征方程 $\phi_i(x)$ 可以写成采样值的线性组合

$$\phi_i(x) = \sum_{t=1}^T \beta_{it}y(x, t). \quad (13)$$

把方程(13)代入式(12)中, 得如下特征值问题:

$$\int_{\Omega} \frac{1}{T} \sum_{t=1}^T y(x, t)y(\zeta, t) \sum_{k=1}^T \beta_k y(x, t)d\zeta = \lambda_i \sum_{t=1}^T \beta_t y(x, t). \quad (14)$$

定义两点间的时间相关方程

$$C_{tk} = \frac{1}{T} \int_{\Omega} y(\zeta, t)y(\zeta, k)d\zeta. \quad (15)$$

方程(14)中的 $N \times N$ 维特征值问题可以简化成如下的 $T \times T$ 维特征值问题:

$$C\beta_i = \lambda_i\beta_i. \quad (16)$$

由上式可以解得特征向量 β_1, \dots, β_T , 代入方程(13)中可解得特征方程 $\phi_1(x), \dots, \phi_T(x)$.

由于时间相关方程矩阵对称半正定, 因此计算得到的特征方程满足正交性. 定义系统的总能量等于所有特征值的和, 对每个特征空间基函数, 其所占系统总能量的比例如下:

$$E_i = \frac{\lambda_i}{\sum_{i=1}^n \lambda_i}. \quad (17)$$

通常认为, 能涵盖系统总能量 99% 的特征空间基函数的数目决定方程的阶数^[9]. 对函数集 $\{\phi_i(x)\}_{i=1}^n$, 有关系式

$$\sum_{i=1}^n \langle (y(x, t), \phi_i)^2 \rangle = \sum_{i=1}^n \lambda_i^2 \geq \sum_{i=1}^n \langle (y(x, t), \phi_i)^2 \rangle. \quad (18)$$

从上式可以看出, 在一类线性组合表示方法中,

K–L分解最优, 即在满足一定逼近误差的条件下, 所需**K–L**基的个数是最少的.

5 Hammerstein模型最小二乘支持向量机参数辨识(Hammerstein model identified by LS-SVM)

最小二乘支持向量机(least squares support vector machine, **LS–SVM**)和传统的二次规划问题不同, 它使用最小二乘系数作为损失函数, 给定一个用于训练的数据集合 $\{(x_i, y_i)\}_{i=1}^N$, 由**LS–SVM**理论, 输入空间 R^n 通过一个非线性函数 $\phi(x_i)$ 被映射到一个高维特征空间 Z , 在这个特征空间中, 采用如下的表达式来近似未知函数^[22]:

$$y(x) = \omega^T \varphi(x) + b, \omega \in Z, b \in \mathbb{R}. \quad (19)$$

这就使得原样本空间中非线性函数估计变成了高维特征空间中的线性函数估计, 式中的参数 ω, b 待求. 这一回归问题, 根据结构风险最小原理^[23], 可以表示为一个等式约束问题, 因而**LS–SVM**的优化问题

$$\min_{\omega, e} J(\omega, e) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \rho \sum_{i=1}^N e_i^2 \quad (20)$$

满足等式约束条件

$$y_i = \omega^T \varphi(x_i) + b + e_i, i = 1, 2, \dots, N. \quad (21)$$

目标函数(20)的第1项对应模型的泛化能力, 第2项表示模型的精确性, 正的常数 ρ 是模型泛化能力和精确性的一个折中参数, 可以根据需要进行调整, e_i 是第*i*组实验数据和预测输出间的误差.

时域**Hammerstein**模型输入 $u(t)$ 和输出 $y(t)$ 能用以下关系式表达:

$$y_i(t) = \sum_{j=1}^n \beta_j y_i(t-j) + \sum_{k=1}^m C_k f(u(t-k)) + e(t) \sum_{j=1}^n \beta_j y_i(t-j) + \sum_{k=1}^m \sum_{l=1}^{n_u} c_k f(u_l(t-k)) + e(t), \quad (22)$$

其中: $u(t) = [u_1(t) \dots u_{n_u}(t)]$ 表示时域输入信号, $y_i(t)$ 表示时域输出值, n, m 表示系统输入输出的阶数, $f(\cdot)$ 是系统非线性部分的方程, n_u 表示时域输入号 $u(t)$ 的个数, $\beta = [\beta_1 \dots \beta_n]^T$ 和 $C_k = [c_1 \dots c_1 \dots c_{n_u}]^T$ 以及 $c_1 = [c_1(1) \dots c_1(m)]$ 是系统回归参数.

根据**LS–SVM**理论, 时域**Hammerstein**模型中的非线性项可以表示为

$$c_1(k)f(u_1(t-k)) = c_1(k)(\omega^T \varphi(u_1(t-k)) + d). \quad (23)$$

于是总的**Hammerstein**模型可以写成如下形式:

$$y_i(t) = \sum_{j=1}^n \beta_j y_i(t-j) + \sum_{k=1}^m (\omega_k^T \varphi(u(t-k)) +$$

$$\mathbf{C}_k d) + e(t), \quad (24)$$

式中: $\boldsymbol{\omega} = [\mathbf{b}_1 \cdots \mathbf{b}_{n_u} \boldsymbol{\omega}]^T$, $\sum_{k=1}^m \mathbf{C}_k d = \mathbf{c}$. 因此, 原问题的最优化问题可以写成如下形式:

$$\min_{\boldsymbol{\omega}_k, \mathbf{c}, \beta_j, e} J(\boldsymbol{\omega}_k, \beta_j, e) = \frac{1}{2} \sum_{k=1}^m \boldsymbol{\omega}_k^T \boldsymbol{\omega}_k + \frac{1}{2} \sum_{j=1}^n \beta_j^T \beta_j + \frac{\rho}{2} \sum_{t=r}^T e^2(t). \quad (25)$$

约束条件为

$$\begin{cases} \sum_{j=1}^n \beta_j y_j(t-j) + \sum_{k=1}^m (\boldsymbol{\omega}_k^T \varphi(\mathbf{u}(t-k))) + \\ \mathbf{c} + e(t) - y_i(t) = 0, \\ \sum_{t=1}^T \boldsymbol{\omega}_k^T \varphi(\mathbf{u}(t)) = 0, \end{cases} \quad (26)$$

式中: $t = r, \dots, T$, $r = \max(m, n) + 1$.

为求解方程(25)在约束条件(26)下的优化问题, 使用拉格朗日乘子法构造拉格朗日函数如下:

$$\begin{aligned} L(\boldsymbol{\omega}_k, \beta_j, e, \alpha_t, \delta_k, \mathbf{c}) = & J(\boldsymbol{\omega}_k, \beta_j, e) - \\ & \sum_{t=r}^T [\sum_{j=1}^n \beta_j y_i(t-j) + \sum_{k=1}^m (\boldsymbol{\omega}_k^T \varphi(\mathbf{u}(t-k))) + \\ & \mathbf{c} + e(t) - y_i(t)] - \sum_{t=1}^T \delta_k \boldsymbol{\omega}_k^T \varphi(\mathbf{u}(t)), \end{aligned} \quad (27)$$

式中 α_t, δ_k 是拉格朗日乘子. 由拉格朗日函数的最优值条件

$$\begin{cases} \frac{\partial L}{\partial \boldsymbol{\omega}_k} = 0, & \frac{\partial L}{\partial \beta_j} = 0, \\ \frac{\partial L}{\partial e} = 0, & \frac{\partial L}{\partial \alpha_t} = 0, \\ \frac{\partial L}{\partial \mathbf{c}} = 0, & \frac{\partial L}{\partial \delta_k} = 0 \end{cases} \quad (28)$$

得到如下线性方程组:

$$\begin{bmatrix} 0 & 0 & \mathbf{1}^T & 0 \\ 0 & 0 & \mathbf{y}_p & 0 \\ 1 & \mathbf{y}_p^T & \mathbf{D} + \rho^{-1} \mathbf{I} & \mathbf{E} \\ 0 & 0 & \mathbf{E}^T & \mathbf{T} \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\delta} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mathbf{y}_i \\ 0 \end{bmatrix}, \quad (29)$$

式中:

$$\mathbf{D}_{(T-r+1) \times (T-r+1)} = \sum_{k=1}^m \varphi^T(\mathbf{u}(\tau-k)) \varphi(\mathbf{u}(t-k)), \quad \tau, t = r, \dots, T,$$

$$\mathbf{E}_{(T-r+1) \times m} = \sum_{\tau=1}^T \varphi^T(\mathbf{u}(\tau)) \varphi(\mathbf{u}(t-k)), \quad k = 1, \dots, m, \quad t = r, \dots, T,$$

$$\mathbf{Q}_{u(T \times T)}(\tau, t) = \varphi^T(\mathbf{u}(\tau)) \varphi(\mathbf{u}(t)), \quad t, \tau = 1, \dots, T,$$

$$\mathbf{T} = \mathbf{1}_T^T \mathbf{Q}_u \mathbf{I}_m,$$

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_T \end{bmatrix}, \quad \boldsymbol{\delta} = [\delta_1 \cdots \delta_m]^T.$$

待求的参数 $\alpha_t, \delta_k, \beta_i, \mathbf{c}$ 可以通过求解方程组(29)得到, 下面通过奇异值分解法求解回归参数 $\mathbf{C}_k (k = 1, \dots, m)$, 构造等式(30):

$$\begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \vdots \\ \mathbf{C}_m \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}^T(\mathbf{u}(1)) \\ \hat{\mathbf{f}}^T(\mathbf{u}(2)) \\ \vdots \\ \hat{\mathbf{f}}^T(\mathbf{u}(T)) \end{bmatrix}^T = \begin{bmatrix} \beta_T \cdots \beta_r & 0 \\ & \beta_T & \beta_r \\ & & \ddots & \ddots \\ 0 & & & \beta_T \cdots \beta_r \end{bmatrix} \times \begin{bmatrix} \mathbf{D}_{T,1} & \mathbf{D}_{T,2} & \cdots & \mathbf{D}_{T,T} \\ \mathbf{D}_{T-1,1} & \mathbf{D}_{T-1,2} & \cdots & \mathbf{D}_{T-1,T} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{D}_{r-(m-1),1} & \mathbf{D}_{r-(m-1),2} & \cdots & \mathbf{D}_{r-(m-1),T} \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_m \end{bmatrix} \sum_{t=1}^T \begin{bmatrix} \mathbf{Q}_{u(t,1)} \\ \mathbf{Q}_{u(t,2)} \\ \vdots \\ \mathbf{Q}_{u(t,T)} \end{bmatrix}, \quad (30)$$

其中: $\hat{\mathbf{f}}(\mathbf{u}(k))$ 是非线性部分的估计值, $\underline{\mathbf{f}}(\mathbf{u}(k)) = \mathbf{f}(\mathbf{u}(k)) - \mathbf{c}$. 因此非线性部分 \mathbf{f} 为

$$\mathbf{f}(\mathbf{u}) = \hat{\mathbf{f}}(\mathbf{u}) + \frac{\mathbf{c}}{\sum_{k=1}^m \mathbf{C}_k}.$$

方程(30)等式右边为已知项, 因此对其进行奇异值分解即可求得系数 \mathbf{C}_k .

6 实验与仿真研究(Experimental and simulation study)

为验证模型的效果, 对方程(31)描述的微悬臂梁系统进行实验结合仿真研究:

$$m \frac{\partial^2 y(x, t)}{\partial t^2} + c \frac{\partial y(x, t)}{\partial t} + EI \frac{\partial^4 y(x, t)}{\partial x^4} + f(y(x, t)) = -F(t) \delta(x - L). \quad (31)$$

以上偏微分方程满足如下边界条件与初始条件:

$$y(0, t) = 0, \quad y_x(0, t) = 0, \quad (32)$$

$$y_{xx}(L, t) = 0, \quad EI y_{xxx}(L, t) = -F(t). \quad (33)$$

将方程(1)描述的模型定义为微悬臂梁机理模型, 方程(31)描述的模型为微悬臂梁改进的非线性模型.

对AFM而言, 受限于激光位置检测传感器的数量

和空间布置, 目前只能测得AFM微悬臂梁针尖处的弯曲变形值, 梁上除针尖外其余点的形变量不能同时准确测得. 由于本文的目的是在机理模型的基础上辨识出等效可控时域模型, 用于针尖的定位和控制, 因此辨识等效可控模型所需非针尖数据可由改进后的机理模型数值仿真得到. 当改进后的机理模型输出与实验结果足够接近时, 辨识所得等效可控模型就可以满足实际需求. 因此, 先设定改进非线性模型的近似形式, 非针尖点的数据由改进后的近似模型通过有限差分法产生. 对比针尖处实验数据和改进模型数值仿真产生的结果, 不断修正改进模型中的参数值和非线性补偿项的形式, 以使二者的值尽量接近, 最终使方程(31)描述的模型尽可能的接近实际系统, 修正后的模型将用于数值仿真以获得足量的实验数据.

AFM实验仪器如图5所示.



图5 AFM实验仪器

Fig. 5 AFM experimental instruments

微悬臂梁参数如下:

$$m = 4.08 \times 10^{-7} \text{ kg/m}, E = 169 \text{ Gpa},$$

$$L = 232 \mu\text{m}, w = 40 \mu\text{m},$$

$$d = 4 \mu\text{m}, c = 0.01 \text{ kg/ms},$$

其中 w 和 d 分别是梁的宽度和厚度, 非线性补偿项估计值可由针尖处实验数据采用多项式拟合或者神经网络近似替代.

系统的响应数据由实验结合数值仿真有限差分得到, 一共获得700组输入输出数据, 采样间隔0.01 s, 仿真时长7 s. 前400组数据用于训练模型, 后300组数据用于检验模型的精度.

为评估所建模型的效果, 定义如下性能参数:

模型预测误差绝对值:

$$e(x_i, t) = |y(x_i, t) - \hat{y}(x_i, t)|. \quad (34)$$

均方根误差:

$$\text{RMSE}(t) = \left(\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (y(x_i, t) - \hat{y}(x_i, t))^2 \right)^{\frac{1}{2}}. \quad (35)$$

相对均方根误差:

$$\text{RRMSE}(t) = \left(\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left(\frac{y(x_i, t) - \hat{y}(x_i, t)}{y(x_i, t)} \right)^2 \right)^{\frac{1}{2}}. \quad (36)$$

微悬臂末端处针尖所受原子力 $F(t)$ 是输入信号, 如图6所示; 仿真产生的悬臂梁受迫振动形变如图7所示. 本文取前5阶空间基函数, 使用K-L分解得到的满足正交性前5阶空间基函数如图8所示.

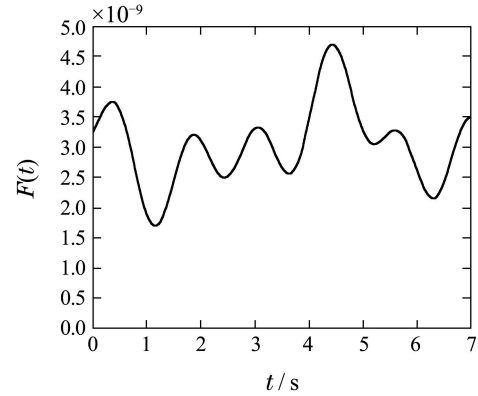


图6 输入信号曲线

Fig. 6 Input signal curve

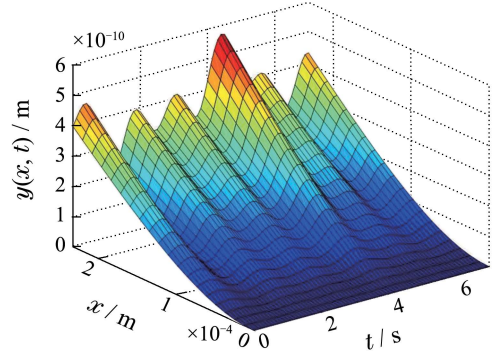


图7 系统响应输出数据

Fig. 7 System response output data

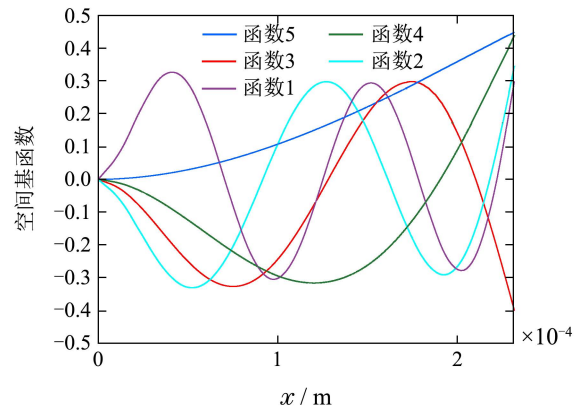


图8 K-L分解得到的空间基函数

Fig. 8 Spatial basis function acquired by K-L decomposition

Hammerstein模型参数调整方法如下: 先将 m, n 设为某固定值, 然后调整 ρ 和 δ , 使用10折交叉验证

法^[24],使得模型在满足最小训练误差的前提下复杂度最小.得到最优 ρ 和 δ 值后,再在此基础上调节 m 和 n ,得到 m 和 n 的最优值.经验证,最终取Hammerstein模型的阶数 m 和 n 分别为3和6,模型正则化参数 $\rho=100$, $\delta=0.001$.

模型第1阶和第5阶预测与仿真时域响应对比分别如图9-10.改进的非线性模型预测输出如图11,机理模型预测输出如图12,改进非线性模型预测误差如图13,机理模型预测误差如图14.

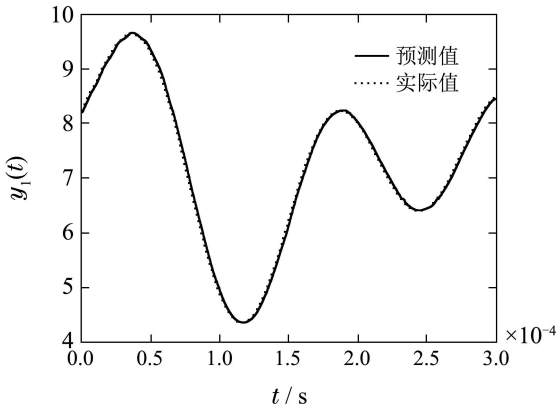


图9 时间系数 $y_1(t)$

Fig. 9 Temporal coefficient $y_1(t)$

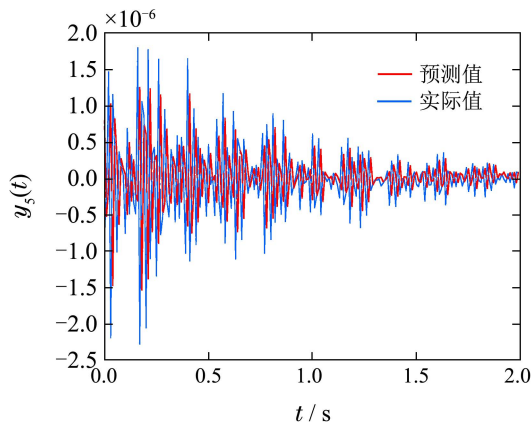


图10 时间系数 $y_5(t)$

Fig. 10 Temporal coefficient $y_5(t)$

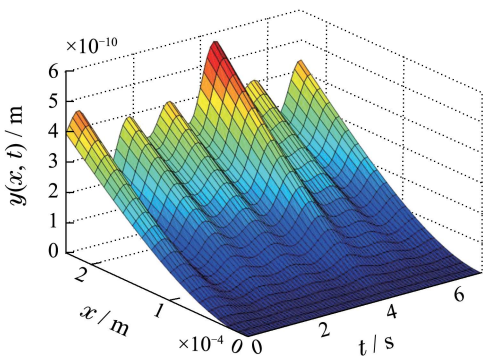


图11 改进非线性模型输出预测值

Fig. 11 Predicted output of the modified nonlinear model

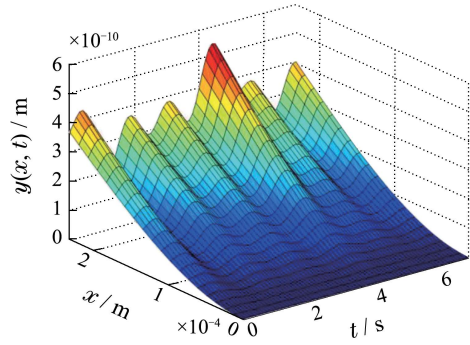


图12 机理模型输出预测值

Fig. 12 Predicted output of the theoretical model

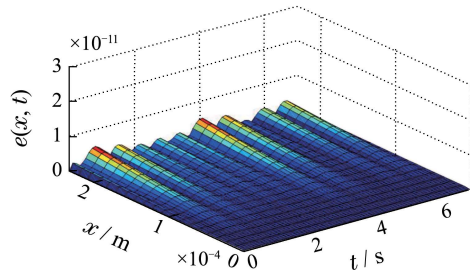


图13 改进的非线性模型预测误差绝对值

Fig. 13 Predicted absolute error of modified nonlinear model

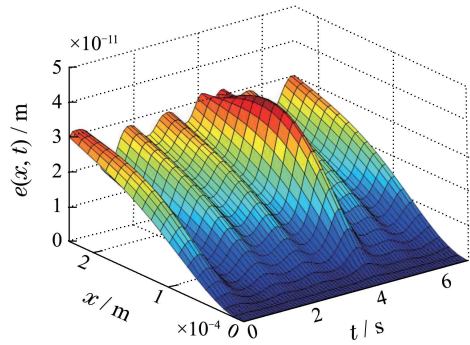
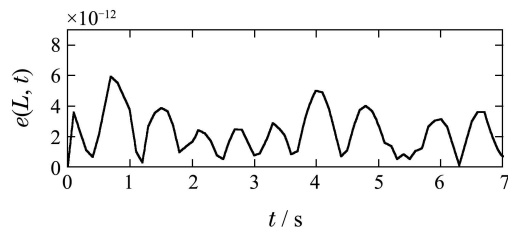


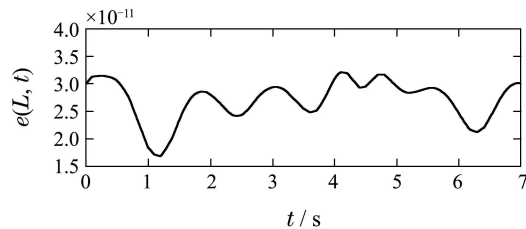
图14 机理模型预测误差绝对值

Fig. 14 Predicted absolute error of theoretical model

微悬臂末端针尖处的预测误差对比如图15.从对比图中可以看出,改进的非线性模型预测误差较机理模型明显减小,机理模型的均方根误差达到了 1.7321×10^{-11} m,而改进的非线性模型均方根误差仅为 1.2452×10^{-12} m.机理模型的相对均方根误差为11.27%,而改进非线性模型的相对均方根误差仅为0.73%.这说明论文提出的非线性模型较原机理模型更能反应系统的真实情况.



(a) 改进非线性模型误差绝对值



(b) 机理模型误差绝对值

图15 微悬臂梁针尖处模型误差绝对值对比

Fig. 15 Absolute error comparison in tip position

7 结论(Conclusions)

论文针对AFM柔性微悬臂梁非线性分布参数系统,对系统输出进行Karhunen-Loève分解,提取出反应系统特征的主导空间基函数,实现了系统输出的时空分离.利用分离得到的时间项系数和系统输入激励信号,建立了时域Hammerstein模型,使得后续的控制设计无需考虑空间信息,简化了控制器的结构.采用最小二乘支持向量机和奇异值分解法对模型中的参数进行辨识,避免了过去迭代法辨识的繁琐步骤.本文的方法无需知道对象精确的偏微分方程,减少了对系统先验知识的依赖,计算量少,便于实施.仿真结果表明,所建模型预测准确,精度较原机理模型高,更能反应系统的真实情况.建模方法对微悬臂梁以及此类非线性分布参数系统皆具有较高的实际应用价值.

参考文献(References):

- [1] 刘延辉. 原子力显微镜及其在各个领域的应用 [J]. 科技导报, 2003, 19(3): 22 – 24.
(LIU Yanhui. Atomic force microscope and its application [J]. *Science & Technology Review*, 2003, 19(3): 22 – 24.)
- [2] 白春礼. 扫描力显微镜 [M]. 北京: 科学出版社, 2000: 10 – 12.
(BAI Chunli. *Scanning Force Microscopy* [M]. Beijing: Science Press, 2000: 10 – 12.)
- [3] STARK M, GUCKENBERGER R, STEMMER A, et al. Estimating the transfer function of the cantilever in atomic force microscopy: a system identification approach [J]. *Journal of Applied Physics*, 2005, 98(11): 114904.
- [4] LEE S I, HOWELL S W, RAMAN A, et al. Nonlinear dynamics of micro-cantilevers in tapping mode atomic force microscopy: a comparison between theory and experiment [J]. *Physical Review B*, 2002, 66(11): 115409.
- [5] 鲍海飞, 李昕欣, 王跃林. 微悬臂梁法向弹性系数的标定方法与分析 [J]. 测试技术学报, 2006, 20(1): 21 – 23.
(BAO Haifei, LI Xinxin, WANG Yuelin. Normal spring constant calibration methods and analysis of micro-cantilever [J]. *Journal of Test and Measurement Technology*, 2006, 20(1): 21 – 23.)
- [6] TZOU H S, DING J H. Optimal control of precision paraboloidal shell structronic systems [J]. *Journal of Sound and Vibration*, 2004, 276(1/2): 273 – 291.
- [7] GOTTLIEB D, ORSZAG S A. *Numerical Analysis of Spectral Methods: Theory and Applications* [M]. Philadelphia: SIAM, 1993: 7 – 21.
- [8] LI H X, QI C K, ZHANG H T. A multi-channel spatiotemporal Hammerstein modeling approach for nonlinear distributed parameter processes [J]. *Journal of Process Control*, 2008, 19(5): 85 – 99.
- [9] CHO D H, PARK H M. The use of Karhunen-Loève decomposition for the modeling of distributed parameter systems [J]. *Chemical Engineering Science*, 1996, 51(1): 81 – 98.
- [10] DELFOUR M, TROCHU F, HAGAR W. Discontinuous Galerkin methods for ordinary differential equations [J]. *Mathematics of Computation*, 1981, 36(154): 455 – 473.
- [11] QI C K, LI H X. A time/space separation-based hammerstein modeling approach for nonlinear distributed parameter process [J]. *Computers & Chemical Engineering*, 2009, 33(7): 1247 – 1260.
- [12] QI C K, YU Y G, LI H X. A spatio-temporal volterra modeling approach for a class of distributed industrial processes [J]. *Journal of Process Control*, 2009, 19(2): 1126 – 1142.
- [13] BAI E W. A blind approach to Hammerstein model identification [J]. *IEEE Transactions on Signal Process*, 2002, 50(7): 1610 – 1619.
- [14] 向微, 陈宗海. 基于Hammerstein模型描述的非线性系统的辨识新方法 [J]. 控制理论与应用, 2007, 24(1): 143 – 147.
(XIANG Wei, CHEN Zonghai. New identification method of nonlinear system based on Hammerstein models [J]. *Control Theory & Applications*, 2007, 24(1): 143 – 147.)
- [15] CHEN H F. Pathwise convergence of recursive identification algorithms for Hammerstein systems [J]. *IEEE Transactions on Automatic Control*, 2004, 49(10): 1641 – 1649.
- [16] KRZYŻAK A, PARTYKA M A. Global identification of nonlinear Hammerstein systems by recursive kernel approach [J]. *Nonlinear Analysis, Theory, Methods and Application*, 2005, 63(6): 1263 – 1272.
- [17] PELCKMANS K, GOETHALS I, SUYKENS J A K, et al. On model complexity control in identification of Hammerstein systems [C] // *Proceedings of The 44th IEEE Conference on Decision and Control, and The European Control Conference*. Sevilla, Spain: IEEE, 2005: 1203 – 1208.
- [18] 贾立, 李训龙. Hammerstein模型辨识的回顾与展望 [J]. 控制理论与应用, 2014, 31(1): 1 – 3.
(JIA Li, LI Xunlong. Identification of Hammerstein model: review and prospect [J]. *Control Theory & Applications*, 2014, 31(1): 1 – 3.)
- [19] RUBIO F J, VAZQUEZ R, STARK R W. Transfer function analysis of the micro cantilever used in atomic force microscopy [J]. *IEEE Transactions on Nanotechnology*, 2006, 5(6): 692 – 700.
- [20] LI H X, QI C K, ZHANG H T, et al. Greatly enhancing the modeling accuracy for distributed parameter systems by nonlinear time/space separation [J]. *Physica A*, 2007, 376(2): 215 – 222.
- [21] BERKOOZ G, HOLMES P, LUMLEY J L. The proper orthogonal decomposition in the analysis of turbulent flows [J]. *Annual Review of Fluid Mechanics*, 1993, 25(1): 539 – 575.
- [22] SONG H Y, GUI W H, YANG C H. Identification of Hammerstein-Wiener model with least squares support vector machine [C] // *Proceedings of the 26th Chinese Control Conference*. Zhang Jiajie, China: Beihang University Press, 2007: 260 – 263.
- [23] 桂卫华, 宋海鹰, 阳春华. Hammerstein-Wiener模型最小二乘支持向量机辨识及其应用 [J]. 控制理论与应用, 2008, 25(3): 394 – 396.
(GUI Weihua, SONG Haiying, YANG Chunhua. Hammerstein-Wiener model identified by least-squares-support-vector machine and its application [J]. *Control Theory & Applications*, 2008, 25(3): 394 – 396.)
- [24] JIANG S H, GUI W H, YANG C H. Vertical quench furnace Hammerstein fault predicting model based on least squares support vector machine and its application [C] // *Chinese Control and Decision Conference*. Guilin, China: Northeastern University Press, 2009: 203 – 206.

作者简介:

徐运扬 (1989–), 男, 硕士研究生, 目前研究方向为系统建模与微机电系统控制, E-mail: xuyunyang2008@163.com;

徐康康 (1990–), 男, 博士研究生, 目前研究方向为系统建模与温度场管理, E-mail: 779771244@qq.com.