Structural Engineering and Mechanics, Vol. 26, No. 4 (2007) 377-392
DOI: http://dx.doi.org/10.12989/sem.2007.26.4.377

Sliding mode control based on neural network for the vibration reduction of flexible structures

Yong-an Huang†

School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an, 710072, China
School of Mechanical Science & Engineering, Huazhong University of Science & Technology, Wuhan, 430074, China

Zi-chen Deng‡

School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an, 710072, China
State Key Laboratory of Structural Analysis of Industrial Equipment, Dalian University of Technology, Dalian, 116024, China

Wen-cheng Li‡†

School of Science, Northwestern Polytechnical University, Xi'an, 710072, China

(Received March 29, 2006, Accepted February 26, 2007)

Abstract. A discrete sliding mode control (SMC) method based on hybrid model of neural network and nominal model is proposed to reduce the vibration of flexible structures, which is a robust active controller developed by using a sliding manifold approach. Since the thick boundary layer will reduce the virtue of SMC, the multilayer feed-forward neural network is adopted to model the uncertainty part. The neural network is trained by Levenberg-Marquardt backpropagation. The design objective of the sliding mode surface is based on the quadratic optimal cost function. In course of running, the input signal of SMC come from the hybrid model of the nominal model and the neural network. The simulation shows that the proposed control scheme is very effective for large uncertainty systems.

Keywords: sliding mode control; neural network; flexible structure; hybrid model.

1. Introduction

Flexible structures provide significant merits over the commonly used rigid structures due to their lightweight properties. However, these kinds of structures are rather difficult to be controlled precisely in that a very high (theoretically infinite) dimensional state space needs employing to

† Postdoctor, E-mail: yahuang@hust.edu.cn
‡ Professor, Corresponding author, E-mail: dweifan@nwpu.edu.cn
‡† Lecturer, E-mail: wenchengli@nwpu.edu.cn

represent the accurate dynamics system. In practical applications, an effective control for the flexible structure is needed in a variety of applications ranging from robot arms to aerospace structures. In the past decades, the active control absorbs much interesting about the vibration reduction of flexible structures (Qiu *et al.* 2002, Jung *et al.* 2004, El-Sinawi 2004). However there isn't a very ideal control method, whose main problem had been listed in Cavallo *et al.* (1999). Controller designs for these control systems can be adaptive control (Suleman and Costa 2004), robust control (Stavroulakis *et al.* 2005), neural network control (Cheng and Patel 2003), and so on, which have their own advantages and disadvantages.

Sliding mode control (SMC) is an ideal candidate method to control the vibration of flexible structures since it is insensitive to parameters' changes or external disturbances (Hung *et al.* 1993, Huang and Deng 2005). However, there are uncertainties and finite time delays in a practical plant and limitation in the physical actuators. All these non-ideal conditions will result in a phenomenon called chattering which is the main obstacle in practical systems. The high frequency component of chattering maybe excites unmodeled high-frequency dynamics parts which may induce failure of the control scheme (Edwards and Spurgeon 1998). In literatures, there are many suggestions to reduce the chattering, the most popular of which is a boundary layer approach (Lee *et al*. 2001). However the thickness depends on the value of the switching gain, and the value of the switch gain depends on the plant's uncertainties part. In other word, the high uncertainties need a high switch gain, and the high gain needs a thick boundary layer. However, the thick boundary layer will lost the virtue of the sliding mode control. Hence, how to obtain a chattering-free SMC becomes an important topic with an acceptable suppression result for the vibration of flexible structures. Recently, the intelligent techniques based on fuzzy logic, neural networks and other techniques have been proposed to solve the chattering problem to improve the control performance of SMC (Lee *et al*. 2001, Ertugrul and Kaynak 2000). Furthermore, there are many literatures about combining intelligent techniques with sliding mode control techniques (Hussain and Ho 2004, Wai *et al.* 2004).

Because of the main reason of chattering due to the poor modeling, one promising method is to use the neural network to model the uncertain parts or the external disturbance. Finally, the control scheme will be given based on the hybrid model of the nominal system and the neural network. If you want to know something about the identification using neural network, there are many papers investigating on the neural network modeling for nonlinear systems (Yazdizadeh and Khorasani 2002). In this paper, a novel sliding mode control scheme assisted by neural network will be presented for depressing the vibration of flexible structures using discrete state-space equation. The neural network is used to model the uncertainties of the practical plant. The sliding mode surface is defined by the optimal method. This paper is organized as follows: How to get the discrete system is simply introduced in Section 2. Section 3 presents sliding mode control theory and discusses how to get the sliding mode surface and design the controller. Section 4 presents how to design the controller using hybrid model of nominal system and neural network. The multilayer feed-forward neural network and the training method is introduced in Section 5. The last section shows the simulation of the flexible using the SMC based on the hybrid model.

2. Discuss on establishing discrete dynamics system

The dynamics model of the multi-degree of freedom system is as follows

$$\mathbf{M}\ddot{\mathbf{X}}(t) + \mathbf{C}\dot{\mathbf{X}}(t) + \mathbf{K}\mathbf{X}(t) = \mathbf{f}(t) + \mathbf{D}\mathbf{u}(t) \tag{1}$$

where, matrixes **M**, **C** and **K** are the mass, the damp and the stiff matrixes of $n \times n$ dimension, respectively; Matrix **D** is the control force position matrix of $n \times m$ dimension; $\mathbf{u}(t) = [u_1(t) \;\; u_2(t) \;\; \dots \;\; u_m(t)]^T$ is the control force vector; $\mathbf{X}(t) = [x_1(t) \;\; x_2(t) \;\; \dots \;\; x_n(t)]^T$, $\dot{\mathbf{X}}(t) = [\dot{x}_1(t) \;\; \dot{x}_2(t) \;\; \dots \;\; \dot{x}_n(t)]^T$ and $\ddot{\mathbf{X}}(t) = [\ddot{x}_1(t) \;\; \ddot{x}_2(t) \;\; \dots \;\; \ddot{x}_n(t)]^T$ are the displacement, the velocity and the acceleration vector respectively; $f(t)$ is the external disturbance.

In order to use the sliding mode control, the dynamics model needs to be transformed into state-space system as follows

$$\dot{\mathbf{Y}}(t) = \mathbf{A}\mathbf{Y}(t) + \mathbf{B}\mathbf{D}\mathbf{u}(t) + \mathbf{B}\mathbf{f}(t) = \mathbf{A}\mathbf{Y}(t) + \mathbf{B}[\mathbf{D}u(t) + \mathbf{f}(t)] = \mathbf{A}\mathbf{Y}(t) + \mathbf{B}\mathbf{U}(t) \tag{2}$$

where, $\mathbf{Y}(t) = \begin{bmatrix} \mathbf{X}(t) \\ \dot{\mathbf{X}}(t) \end{bmatrix}_{2n \times 1}$, $\mathbf{A} = \begin{bmatrix} 0 & \mathbf{I}_n \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix}_{2n \times 2n}$, $\mathbf{B} = \begin{bmatrix} 0 \\ \mathbf{M}^{-1}\mathbf{D} \end{bmatrix}_{2n \times m}$, $\mathbf{I}_n$ is $n \times n$ identity matrix, and $\mathbf{U}(t) = \mathbf{D}\mathbf{u}(t) + \mathbf{f}(t)$.

In practical control systems, discrete control methods are usually adopt to depress the vibration. Suppose the sample time is T_d, in the interval of which the control force keeps constant

$$\mathbf{U}(t) = \mathbf{U}(kT_d), \quad kT_d \le t \le (k+1)T_d \tag{3}$$

We can get the result of continuous state-space system (2) as follows

$$\mathbf{Y}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{Z}(t_0) + \int_{t_0}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{U}(\tau)d\tau \tag{4}$$

If $t_0 = kT_d$ and $t = (k+1)T_d$, the continuous system can be transformed into the discrete system as follows

$$\mathbf{Y}((k+1)T_d) = \mathbf{A}_d(T_d)\mathbf{Y}(kT_d) + \mathbf{B}_d(T_d)\mathbf{U}(kT_d) \tag{5}$$

where, $\mathbf{A}_d(T_d) = e^{\mathbf{A}T_d}$ is a matrix of $2n \times 2n$ dimension, $\mathbf{B}_d(T_d) = \int_0^T e^{\mathbf{A}v}dv\mathbf{B}$ is a matrix of $2n \times m$ dimension, and $v = (k+1)T_d - \tau$. Eq. (5) can be expressed in brief form

$$\mathbf{Y}(k+1) = \mathbf{A}_d\mathbf{Y}(k) + \mathbf{B}_d\mathbf{U}(k) \tag{6}$$

3. Sliding mode control for the flexible structure vibration

3.1 Theory of sliding mode control

One of the most attractive features of SMC is the high robustness when it holds on the sliding surface. The design method of the sliding mode controller will be discussed, and some important characteristics will be given, more about which can be gotten from (Edwards and Spurgeon 1998, DeCarlo *et al.* 1988). SMC is a control method that switches rapidly between two values, and hold the state-space on a special surface. But there are differences between the discrete system and the continuous system. The discrete state-space variables are limited on the time sequence $(t = 0, T_d, 2T_d, \ldots, kT_d, \ldots)$. So it is less possible that the discrete state-space reach the sliding surface.

Without loss of generalities, the general SMC problem is represented by the linear nominal discrete state-space model (6). SMC scheme generally contains two fundamental steps as follows:

Step 1: The sliding mode surface represents a desired system dynamics. An appropriate sliding mode surface needs designing for the flexible dynamics system to achieve the special performance after the system entering the sliding mode surface. The sliding mode surface can be written as $\mathbf{S}(k)$, whose dimension is usually the same as to the dimension of the control force $\mathbf{u}(k)$.

Step 2: Design the control law for the flexible structure so that the reaching and sliding conditions on the sliding surfaces can be satisfied, which means any initial states $\mathbf{Y}(k)$ outside the sliding surface can be driven to reach the surface in finite time and hold on the surface. SMC law as follows

$$u_i(k) = \begin{cases} u_i^+, S_i(k) > 0 \\ u_i^-, S_i(k) < 0 \end{cases}, \quad i = 1, 2, \ldots, m \tag{7}$$

If the above two steps are designed successfully, the states $\mathbf{Y}(k)$ in the both sides of the sliding surface $\mathbf{S}(k) = \mathbf{0}$ will be driven onto the sliding surface, and converged to the origin finally. If the ideal control law is designed, attractive characteristics mentioned above can be realized, and the vibration of flexible structures will be reduced successfully.

3.2 Design sliding surface for the flexible structure

In order to illustrate the surface design techniques used in this paper, let us consider the nonsingular transformation matrix $\mathbf{T} \in R^{2n \times 2n}$, which transforms system (6) into the canonical form. The transformation equation is given as follows

$$\mathbf{Z}(k) = \mathbf{T}\mathbf{Y}(k) \tag{8}$$

where, $\mathbf{T} = \begin{bmatrix} \mathbf{I}_{2n-m} & -\mathbf{B}_1\mathbf{B}_2^{-1} \\ 0 & \mathbf{I}_m \end{bmatrix}$, $\mathbf{B}_1 \in R^{(2n-m) \times m}, \mathbf{B}_2 \in R^{m \times m}$. The matrix $\mathbf{B}_2$ is invertible.

Substitute the transformation Eq. (8) into the system (5) to get the canonical form. The original Eq. (5) can be written as follows

$$\mathbf{Z}(k+1) = \tilde{\mathbf{A}}_d\mathbf{Z}(k) + \tilde{\mathbf{B}}_d\mathbf{U}(k) \tag{9}$$

or

$$\begin{bmatrix} \mathbf{Z}_1(k+1) \\ \mathbf{Z}_2(k+1) \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{A}}_{d11} & \tilde{\mathbf{A}}_{d12} \\ \tilde{\mathbf{A}}_{d21} & \tilde{\mathbf{A}}_{d22} \end{bmatrix} \begin{bmatrix} \mathbf{Z}_1(k) \\ \mathbf{Z}_2(k) \end{bmatrix} + \begin{bmatrix} 0 \\ \tilde{\mathbf{B}}_{d2} \end{bmatrix} \mathbf{U}(k) \tag{10}$$

where, $\tilde{\mathbf{B}}_d = \mathbf{T}\mathbf{B}_d = [0 \;\; \tilde{\mathbf{B}}_{d2}]^T$, $\mathbf{Z}_1(k) \in R^{2n-m}$, $\mathbf{Z}_2(k) \in R^m$, $\tilde{\mathbf{A}}_{d11}$ is $(2n-m) \times (2n-m)$ compatible dimension matrix block of the matrix $\tilde{\mathbf{A}}_d = \mathbf{T}\mathbf{A}_d\mathbf{T}^{-1}$.

The sliding surface $\mathbf{S}(k)$ is supposed to be linear, and has the following form

$$\mathbf{S}(k) = \mathbf{G}\mathbf{Y}(k) \tag{11}$$

Submit Eq. (8) into Eq. (11), then we can get

$$\mathbf{S}(k) = \tilde{\mathbf{G}}\mathbf{Z}(k) = \tilde{\mathbf{G}}_1\mathbf{Z}_1(k) + \tilde{\mathbf{G}}_2\mathbf{Z}_2(k) = \mathbf{0} \tag{12}$$

where, $\tilde{\mathbf{G}} = \mathbf{G}\mathbf{T}^{-1}$. And then we have

$$\tilde{\mathbf{Z}}_2(k) = -\tilde{\mathbf{G}}_2^{-1}\tilde{\mathbf{G}}_1\tilde{\mathbf{Z}}_1(k) = -\Psi\mathbf{Z}_1(k) \tag{13}$$

where, $\Psi = \tilde{\mathbf{G}}_2^{-1}\tilde{\mathbf{G}}_1$.

Substitute Eq. (13) into the first equation of Eq. (10), so that the system representation on the sliding surface becomes

$$\begin{cases} \mathbf{Z}_1(k+1) = (\tilde{\mathbf{A}}_{d11} - \tilde{\mathbf{A}}_{d12}\Psi)\mathbf{Z}_1(k) \\ \mathbf{Z}_2(k) = -\Psi\mathbf{Z}_1(k) \end{cases} \tag{14}$$

Eq. (14), from which we can see the order of the controlled system has been reduced, are made up of $2n-m$ dimensions independent variables, and it's very convenient for designing the controller.

Theorem 1. Supposed the system $(\mathbf{A}_d, \mathbf{B}_d)$ is controllable, the system $(\mathbf{A}_{d11}, \mathbf{A}_{d12})$ is controllable.

Let's discuss how to gain the sliding surface by optimal method. The optimal cost function of the large system to be minimized is

$$\begin{aligned} J &= \sum_{k=1}^{\infty} \mathbf{Y}(k)^T\mathbf{Q}\mathbf{Y}(k) = \sum_{k=1}^{\infty} \mathbf{Z}(k)^T\tilde{\mathbf{Q}}\mathbf{Z}(k) \\ &= \sum_{k=1}^{\infty} [\mathbf{Z}_1(k)^T \;\; \mathbf{Z}_2(k)^T] \begin{bmatrix} \tilde{\mathbf{Q}}_{11} & \tilde{\mathbf{Q}}_{12} \\ \tilde{\mathbf{Q}}_{21} & \tilde{\mathbf{Q}}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Z}_1(k) \\ \mathbf{Z}_2(k) \end{bmatrix} \end{aligned} \tag{15}$$

Substitute Eq. (10) into Eq. (15), we can get the state-space feedback matrix

$$\Psi = (\mathbf{A}_{d12}^T\mathbf{P}\mathbf{A}_{d12} + \tilde{\mathbf{Q}}_{12})^{-1}(\mathbf{A}_{d12}^T\mathbf{P}\mathbf{A}_{d11} + \tilde{\mathbf{Q}}_{12}^T) \tag{16}$$

where, the matrix **P** is the result of the following equation, which can be solved by MATLAB.

$$\mathbf{A}_{d11}^T\mathbf{P}\mathbf{A}_{d11} - \mathbf{S} - (\mathbf{A}_{d11}^T\mathbf{P}\mathbf{A}_{d12} + \tilde{\mathbf{Q}}_{12})(\mathbf{A}_{d12}^T\mathbf{P}\mathbf{A}_{12} + \tilde{\mathbf{Q}}_{22})^{-1}(\mathbf{A}_{d12}^T\mathbf{P}\mathbf{A}_{d11} + \tilde{\mathbf{Q}}_{12}^T) + \tilde{\mathbf{Q}}_{d11} = 0 \tag{17}$$

where, $\tilde{\mathbf{Q}} = [\mathbf{T}^{-1}]^T\mathbf{Q}\mathbf{T}^{-1} = \begin{bmatrix} \tilde{\mathbf{Q}}_{11} & \tilde{\mathbf{Q}}_{12} \\ \tilde{\mathbf{Q}}_{21} & \tilde{\mathbf{Q}}_{22} \end{bmatrix}$.

We can get the sliding matrix as follows

$$\mathbf{G} = \tilde{\mathbf{G}}\mathbf{T} = [\Psi \vdots \mathbf{I}]\mathbf{T} \tag{18}$$

3.3 Discuss how to design controller for the flexible structure

The sliding surface has been determinated, so that we can design the controller for the vibration system to drive the states onto the sliding surface.

A discrete-time sliding surface is defined in the linear combination of the discrete-time state as

$$\mathbf{S}(k) = \mathbf{GY}(k) \tag{19}$$

where, $\mathbf{S}(k)$ is sliding surface. As the system state enters the sliding surface, the first-order difference equation of the sliding variable is equal to 0

$$\mathbf{S}(k+1) - \mathbf{S}(k) = \mathbf{0} \tag{20}$$

From Eqs. (9), (10), (19) and (20), the equivalent control force can be obtained from the following form

$$\mathbf{U}_{eq} = -(\mathbf{GB}_d)^{-1}(\mathbf{G}(\mathbf{A}_d - \mathbf{I})\mathbf{Y}(k)) \tag{21}$$

The controller is designed in virtue of the Lyapunov function, whose candidate can be defined as

$$V(k) = \mathbf{S}(k)^T\mathbf{S}(k) \tag{22}$$

Let $\Delta V(k) = V(k+1) - V(k)$, we can get

$$\Delta V(k) = [\mathbf{S}(k+1) + \mathbf{S}(k)]^T \Delta\mathbf{S}(k) \tag{23}$$

where, $\Delta\mathbf{S}(k) = \mathbf{S}(k+1) - \mathbf{S}(k)$. In order to stable the system, the function $\Delta V(k) < 0$. We define the equation as follows

$$\Delta V(k) + \mathbf{S}(k)^T\mathbf{T}_d\mathbf{WS}(k) + \mathbf{S}(k)^T\mathbf{T}_d V sign(\mathbf{S}(k)) = 0 \tag{24}$$

where, **W** and **V** are all positive definite matrixes. From Eqs. (23) and (24), we can get

$$\mathbf{S}(k)^T[\mathbf{H}(k)^T\Delta\mathbf{S}(k) + \mathbf{T}_d\mathbf{WS}(k)^T + \mathbf{T}_d V sign(\mathbf{S}(k))] = 0 \tag{25}$$

where, $\mathbf{H} = \mathbf{I} + \mathbf{S}(k+1)[\mathbf{S}(k)^T\mathbf{S}(k)]^{-1}\mathbf{S}(k)^T$. If $\mathbf{S}(k) \neq 0$, we can get the following equation

$$\mathbf{H}(k)^T\Delta\mathbf{S}(k) + \mathbf{T}_d\mathbf{WS}(k) + \mathbf{T}_d V sign(\mathbf{S}(k)) = 0 \tag{26}$$

Supposed that $\mathbf{H}(k) = \mathbf{I}$, we can get the control force as follows

$$\begin{aligned}\mathbf{U} &= -(\mathbf{GB}_d)^{-1}[(\mathbf{G}(\mathbf{A}_d - \mathbf{I})\mathbf{Y}(k)) - [\mathbf{H}(k)^T]^{-1}(Ws(k) + Vsign(S(k)))] \\ &= \mathbf{U}_{eq} + (\mathbf{GB}_d)^{-1}[\mathbf{H}(k)^T]^{-1} + (\mathbf{T}_d\mathbf{WS}(k) + \mathbf{T}_d V sign(\mathbf{S}(k))) \\ &= \mathbf{U}_{eq} + (\mathbf{GB}_d)^{-1}(\mathbf{T}_d\mathbf{WS}(k) + \mathbf{T}_d V sign(\mathbf{S}(k)))\end{aligned} \tag{27}$$

where, **U** is the control force used on the flexible structure in the following section.

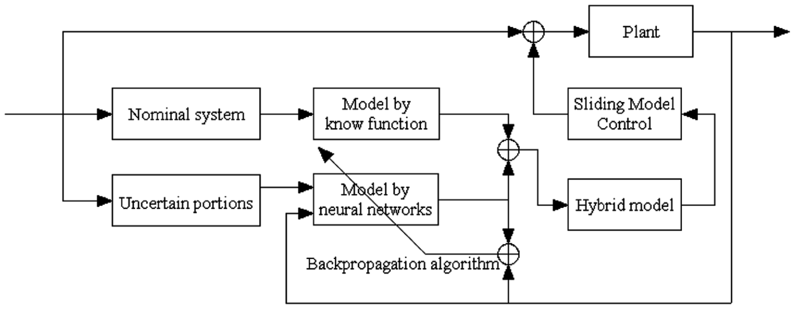


Fig. 1 The control system structure

4. Controller based on the hybrid model of neural network

In this section, the system is established by combining the nominal system and the uncertain portions. Supposed there are uncertainties and external disturbances existing in the system, such as the rand disturbance and uncertainties in the mass, the damp and the stiff.

Define the complex system as follows

$$\mathbf{Y}(k+1) = (\mathbf{A}_d + \Delta\mathbf{A}_d)\mathbf{Y}(k) + (\mathbf{B}_d + \Delta\mathbf{B}_d)\mathbf{U}(k) + \mathbf{d} \tag{28}$$

where, $\mathbf{d}$ is randomly external disturbance. The nominal system is modeled by known dynamics model, and the uncertain portion is modeled by neural networks. The flow chart of the control system is shown in Fig. 1.

Define the neural networks as

$$NET(\mathbf{Y}, \mathbf{U}) = \Delta\mathbf{A}_d\mathbf{Y}(t) + \Delta\mathbf{B}_d\mathbf{U}(t) + \mathbf{d} \tag{29}$$

Introduce Eq. (29) into Eq. (28), the following equation can be gained.

$$\mathbf{Y}(k+1) = \mathbf{A}_d\mathbf{Y}(k) + \mathbf{B}_d\mathbf{U}(k) + NET(\mathbf{Y}, \mathbf{U}) \tag{30}$$

where, $NET(\mathbf{Y}, \mathbf{U})$ is a multilayer feed-forward neural network. If the neural network is trained enough, the neural network can denote the uncertain part.

But the quality of the sliding mode control will decrease with the boundary layer becoming thick. The thickness of the boundary layer depends on the uncertainties of the system. In order to increase the control quality, we need to decrease the thickness of the boundary layer.

The sliding mode control force corresponding to the uncertain system is as follows

$$\begin{aligned}\mathbf{U} = &-(\mathbf{G}\mathbf{B}_d)^{-1}[\mathbf{G}(\Delta\mathbf{A}_d\mathbf{Y}(k) + \Delta\mathbf{B}_d\mathbf{U}(k) + \mathbf{d}) \\ &+ (\mathbf{G}(\mathbf{A}_d - \mathbf{I})\mathbf{Y}(k)) - (\mathbf{T}_d\mathbf{W}\mathbf{S}(k) + \mathbf{T}_d V sign(\mathbf{S}(k)))]\end{aligned} \tag{31}$$

In order to gain the control force $\mathbf{U}$, we need estimate the uncertainty in Eq. (30) as follows

$$\mathbf{G}(\Delta\mathbf{A}_d\mathbf{Y}(k) + \Delta\mathbf{B}_d\mathbf{U}(k) + \mathbf{d}) \tag{32}$$

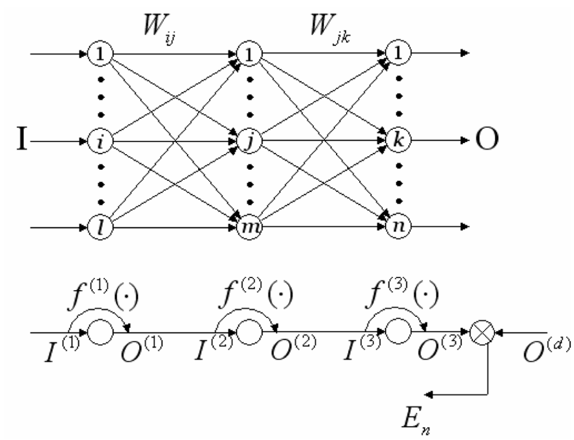


Fig. 2 The structure of the Multilayer Fedforward Neural Network

The neural network method is a promising method. Set the output of the estimator as

$$\mathbf{U} = -(\mathbf{GB}_d)^{-1}\mathbf{G}[NET(\mathbf{Y}, \mathbf{U}) + (\mathbf{G}(\mathbf{A}_d - \mathbf{I})\mathbf{Y}(k)) - \mathbf{T}_d(\mathbf{WS}(k) + Vsign(\mathbf{S}(k)))] \tag{33}$$

5. Training the neural network

5.1 Neural network model

The multilayer feed-forward neural network, which is a typical feed-forward network, is adopted in this paper. This kind of network, which has at least one hidden layer, has been proved that it can approach any nonlinear function in any precision (Hornik 1991, Funahashi 1989, Cybenko 1989). The architecture of a multilayer feed-forward neural network with n inputs and m outputs can be seen in Fig. 2.

The neural network outputs are given as following

(1) The input layer: Input $I_i^{(1)}, \quad i = 1, 2, \ldots, l$

Output $O_i^{(1)} = I_i^{(1)}, \quad i = 1, 2, \ldots, l$

(2) The hidden layer: Input $I_j^{(2)} = \sum_{i=1}^{l}(w_{ji}O_i^{(1)} - \theta_j^{(2)}), \quad j = 1, 2, \ldots, m$

Output $O_j^{(2)} = f^{(2)}(I_j^{(2)}), \quad j = 1, 2, \ldots, m$

(3)The output layer: Input $I_k^{(3)} = \sum_{j=1}^{m}(w_{kj}O_j^{(2)} - \theta_k^{(3)}), \quad k = 1, 2, \ldots, n$

Output $O_k^{(3)} = f^{(3)}(I_k^{(3)}), \quad k = 1, 2, \ldots, n$

where, $\theta_j^{(2)}$ and $\theta_k^{(3)}$ are the bias of the hidden layer and the output layer, respectively. The l, m and n are the number of nodes for the input layer, the hidden layer and the output layer. The $f^{(2)}$ and

$f^{(3)}$ are the active function of the hidden layer and the output layer respectively. The active function used in the paper is hyperbolic tangent function in the hidden layer and the linear function in the output layer.

5.2 Initial training of neural networks

This section describes how to obtain a well-trained network before incorporating into the proposed control strategies. There are four main problems needed to be solved: (1) How to generate the training and validation data sets; (2) how to preprocess of data sets; (3) how to train the neural networks; (4) how to validate the neural network (Hussain and Ho 2004).

The cost function can be defined as follows

$$J = \frac{1}{2}(O^{(3)} - O^{(d)})^2 \tag{34}$$

where, $O^{(3)}$ is the real output of the neural network, and $O^{(d)}$ is the ideal output.

The Levenberg-Marquardt (L-M) method combines the gradient descent method and the Newton method. When the number of the weight is not large, the main virtue is that the velocity of convergence is very fast. The speed of this calculation method is higher than others in one or several order of magnitudes. This can be seen from Table 1 and Table 2. Table 1 displays the comparison of convergence rate between standard gradient descent methods. Table 2 displays the comparison of convergence rate between fast gradient methods.

Now, an example is adopted to show the training velocity. The nonlinear system is

$$y(k) = \frac{y(k-1)}{1 + y^2(k-1)} + u^3(k-1) \tag{35}$$

Table 1 comparison of convergence rate between standard gradient descent methods

Algorithm	Time(s)	Mean square error
Gradient descent with adaptive learning rate backpropagation	215.411	0.0001363
Gradient descent with momentum backpropagation	179.8750	0.0026547
Gradient descent backpropagation	153.7190	0.0026894

Table 2 Comparison of convergence rate between fast gradient methods

Algorithm	Time(s)	Iterations
Conjugate gradient backpropagation with letcher-Reeves updates	2.9342	193
Conjugate gradient backpropagation with Powell-Beale restarts	2.6593	159
Scaled conjugate gradient backpropagation	2.9172	274
BFGS quasi-Newton backpropagation	1.7190	102
One step secant backpropagation	5.7062	408
Levenberg-Marquardt backpropagation	0.6219	7
RPROP backpropagation	2.3124	266

and the input is

$$u(k) = 0.2\sin\frac{2k\pi}{25} + 0.3\sin\frac{k\pi}{15} + 0.3\sin\frac{k\pi}{75} \tag{36}$$

The computer performance: CPU is PIV2.4G, and memory is 512M. The convergence time of different algorithm are displayed in Table 1 and Table 2.

From Table 1, we can see algorithms can not meet square error 0.0001, even if iterations reach 20000 times, and the calculation time is longer than 150 seconds. But data in Table 2 show that not only the accuracy of the fast gradient methods reach the square error 0.0001, but also their calculation time is shorter than 150 seconds. Levenberg-Marquardt backpropagation is the best method of all these algorithms.

Remark: the ending condition of Table 1 is 20000 times iterations, and that in Table 2 is Mean square error 0.0001.

Before training the neural network, the training and validation data sets have to be given. First, we set rand external force onto the practical plant, which in this study is represented by the nominal dynamics model and the uncertain part. Secondly, the data generated by open-loop system are used to train the neural network. To train the neural network, the input signals and states of the original system are adopted as inputs of the neural network, and the value, which is the subtraction of the velocity and the acceleration signal between the practical system and the nominal system, is adopted as the ideal output of the neural network.

In the end, the following equation can be given

$$NET(x_1, x_2, u) = \begin{bmatrix} \Delta\dot{x}_1 \\ \Delta\dot{x}_2 \end{bmatrix} = \begin{bmatrix} \dot{\hat{x}}_1 - \dot{x}_1 \\ \dot{\hat{x}}_2 - \dot{x}_2 \end{bmatrix} \tag{37}$$

where, $\dot{\hat{x}}_1$ and $\dot{\hat{x}}_2$ represent the velocity and the acceleration of the practical system respectively. $\dot{x}_1$ and $\dot{x}_2$ represent the velocity and the acceleration of the nominal respectively.

6. Simulation

6.1 The parameters of the flexible structure and the discrete system

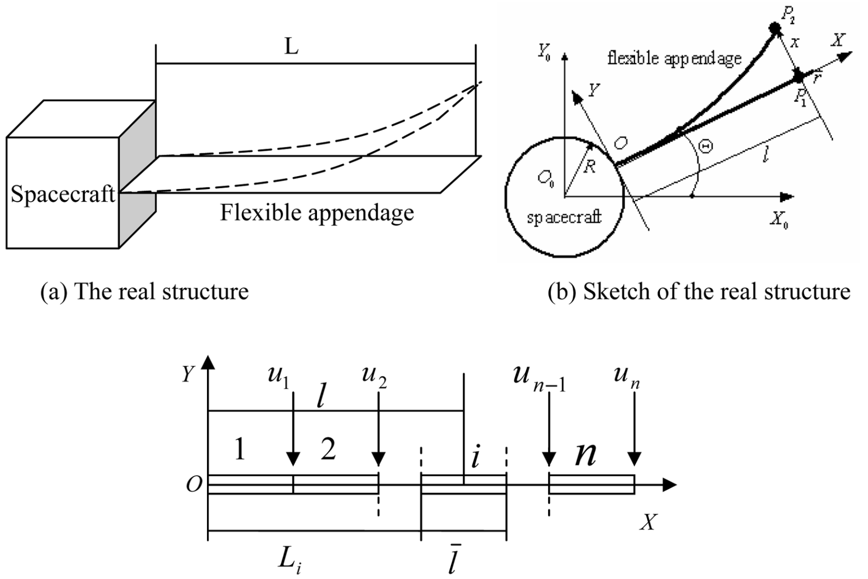
The structure adopted is shown in Fig. 3, and supposed that the control force acts on every node. In this paper, the flexible appendage is divided into three elements. The main parameters coming from Hang *et al.* (1985) are listed as following: $EI = 0.399847 \times 10^7$ N·m^2, $L = 35.05$ m, $\rho = 25.95$ kg/m, $u_{max} = 300\ N$. The displacement vector and the control vector are

$$\mathbf{X} = [x_1 \ \ \theta_1 \ \ x_2 \ \ \theta_2 \ \ x_3 \ \ \theta_3]^T \tag{38}$$

$$\mathbf{U} = [u_1 \ \ u_2 \ \ u_3]^T \tag{39}$$

The mass matrix and the element matrix of rigidity are

$$\mathbf{M} = diag(303, 0, 303, 0, 303, 0) = diag(m_1, 0, m_2, 0, m_3, 0) \tag{40}$$



(a) The real structure

(b) Sketch of the real structure

(c) The finite element model of the flexible appendage

Fig. 3 Flexible appendage of spacecraft

$$[\mathbf{K}^e] = \frac{EI}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix} \tag{41}$$

and the whole matrix of rigidity can be written as following

$$[\mathbf{K}] = \frac{EI}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l & 0 & 0 \\ 6l & 4l^2 & -6l & 2l^2 & 0 & 0 \\ -12 & -6l & 24 & 0 & -12 & 6l \\ 6l & 2l^2 & 0 & 8l^2 & -6l & 2l^2 \\ 0 & 0 & -12 & -6l & 24 & 0 \\ 0 & 0 & 6l & 2l^2 & 0 & 8l^2 \end{bmatrix} = \begin{bmatrix} K_1 \\ K_2 \\ K_3 \\ K_4 \\ K_5 \\ K_6 \end{bmatrix} \tag{42}$$

and the input vector is as following

$$\mathbf{DU} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ 0 \\ u_2 \\ 0 \\ u_3 \\ 0 \end{bmatrix} \tag{43}$$

and the dynamics model of the system is

$$\mathbf{M}\ddot{\mathbf{X}} + \mathbf{K}\mathbf{X} = \mathbf{D}\mathbf{U} \tag{44}$$

or

$$\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{bmatrix} + \begin{bmatrix} K_1 \\ K_3 \\ K_5 \end{bmatrix}\mathbf{X} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \tag{45}$$

$$\begin{bmatrix} \theta_2 \\ \theta_4 \\ \theta_6 \end{bmatrix} = \begin{bmatrix} k_{22} & k_{24} & k_{26} \\ k_{42} & k_{44} & k_{46} \\ k_{62} & k_{64} & k_{66} \end{bmatrix}^{-1}\begin{bmatrix} k_{21} & k_{23} & k_{25} \\ k_{41} & k_{43} & k_{45} \\ k_{61} & k_{63} & k_{65} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{46}$$

After submitting Eq. (46) into Eq. (45), we can get

$$\mathbf{M}\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{bmatrix} + \mathbf{V}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \tag{47}$$

where, $\mathbf{M} = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}$, $\mathbf{V} = \begin{bmatrix} k_{11} & k_{13} & k_{15} \\ k_{31} & k_{33} & k_{35} \\ k_{51} & k_{53} & k_{55} \end{bmatrix} + \begin{bmatrix} k_{22} & k_{24} & k_{26} \\ k_{42} & k_{44} & k_{46} \\ k_{62} & k_{64} & k_{66} \end{bmatrix}^{-1}\begin{bmatrix} k_{21} & k_{23} & k_{25} \\ k_{41} & k_{43} & k_{45} \\ k_{61} & k_{63} & k_{65} \end{bmatrix}$.

Transform Eq. (47) into the state-space model

$$\dot{\mathbf{Y}}(t) = \mathbf{A}\mathbf{Y}(t) + \mathbf{B}\mathbf{u}(t) \tag{48}$$

where, $\mathbf{Y}(t) = [x_1 \;\; x_2 \;\; x_3 \;\; \dot{x}_1 \;\; \dot{x}_2 \;\; \dot{x}_3]^T$

$$\mathbf{B} = \begin{bmatrix} 0 \\ \mathbf{M}^{-1}\mathbf{D} \end{bmatrix} = 10^2 \times \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.33 & 0 & 0 \\ 0 & 0.33 & 0 \\ 0 & 0 & 0.33 \end{bmatrix},\quad \mathbf{A} = \begin{bmatrix} 0 & \mathbf{I} \\ \mathbf{M}^{-1}\mathbf{V} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -7969 & 7969 & 0 & 0 & 0 & 0 \\ 7969 & -15938 & 7969 & 0 & 0 & 0 \\ 0 & 7969 & -15938 & 0 & 0 & 0 \end{bmatrix}.$$

Now transform Eq. (48) into the discrete system

$$\mathbf{Z}(k+1) = \tilde{\mathbf{A}}_d\mathbf{Z}(k) + \tilde{\mathbf{B}}_d\mathbf{U}(k) \tag{49}$$

where,

$$\tilde{\mathbf{A}}_d = \begin{bmatrix} 1 & 0 & 0 & 0.002 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0.002 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0.002 \\ -1.5921 & 1.5913 & 0.0008 & 0.9968 & 0.0032 & 0 \\ 1.5913 & -3.1825 & 1.5904 & 0.0032 & 0.9936 & 0.0032 \\ 0.0008 & 1.5904 & -3.1834 & 0 & 0.0032 & 0.9936 \end{bmatrix}, \quad \tilde{\mathbf{B}}_d = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.6590 & 0.0004 & 0 \\ 0.0004 & 0.6586 & 0.0004 \\ 0 & 0.0004 & 0.6586 \end{bmatrix}.$$

6.2 Simulation and results

From the above section, the nominal state-space equation can be obtained. In order to simulation the practical system, we supposed that there are rand uncertainties in the mass matrix and the stiff matrix. The state (displacement and velocity) of the real system, the nominal system and the error between the real and the nominal system can be given. The displacement and the velocity of the first degree of freedom are compared in Fig. 4 and Fig. 5.

In order to determinate the optimal sliding surface, we define the parameter **Q** of Eq. (15) as

$$\mathbf{Q} = 10^5 \times diag(0.0030, 0.0030, 0.0030, 2.4173, 4.8345, 4.8345) \tag{50}$$

We can get the optimal surface **G** in Eq. (18) as follows

$$\mathbf{G} = \begin{bmatrix} 0.0354 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0.0250 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0.0250 & 0 & 0 & 1 \end{bmatrix} \tag{51}$$

To get the control force **U** given by Eq. (33), the neural network needs to be trained. The input of the neural network is $[Z(k-1)^T \vdots Z(k)^T \vdots f^T]^T$, and the output of the neural network is $\mathbf{Z}(k+1)$. There are 15 neurons in input layer, 18 neurons in hidden layer, 6 neurons in output layer. The

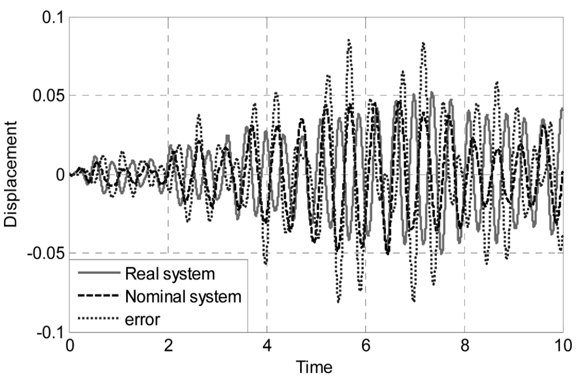


Fig. 4 Displacement of the real system, nominal system and the error between the real and the nominal system

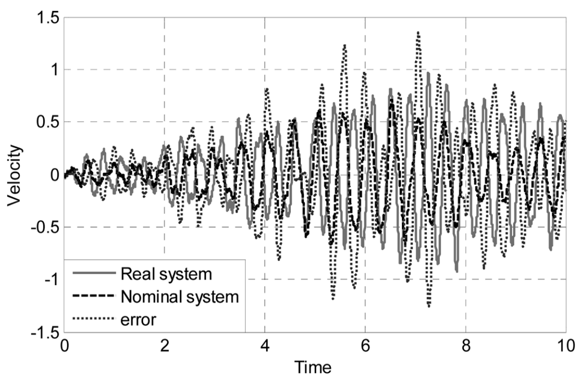


Fig. 5 Velocity of the real system, nominal system and the error between the real and the nominal system

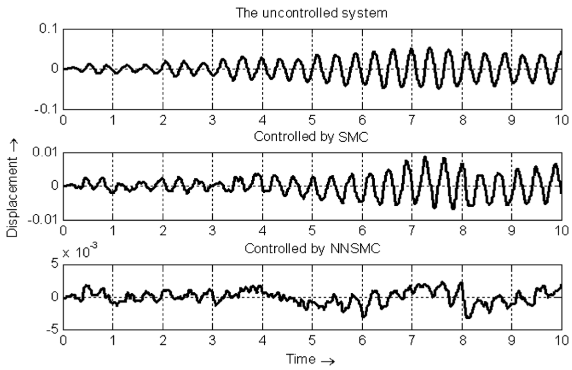


Fig. 6 Displacement of the first DOF of uncontrolled system, system controlled by SMC and NNSMC

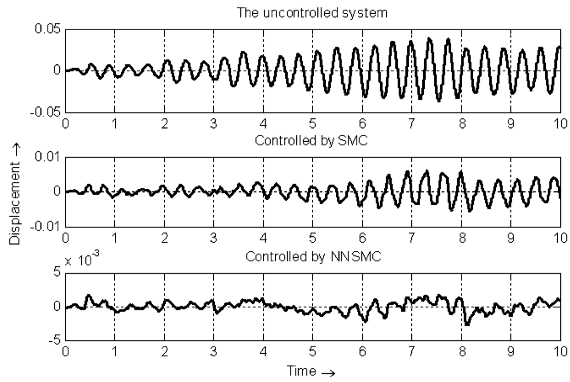


Fig. 7 Displacement of the second DOF of uncontrolled system, system controlled by SMC and NNSMC

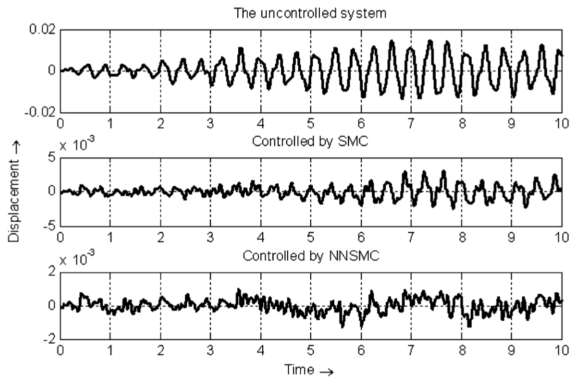


Fig. 8 Displacement of the third DOF of uncontrolled system, system controlled by SMC and NNSMC

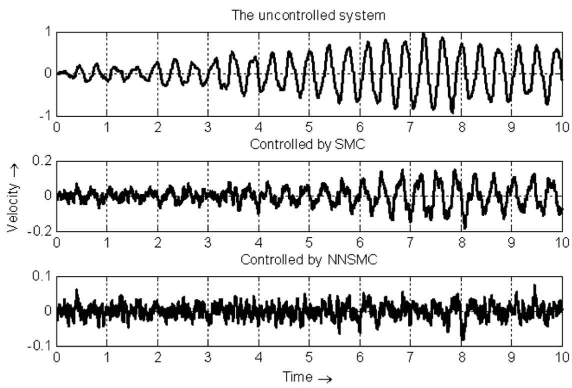


Fig. 9 Velocity of the first DOF of uncontrolled system, system controlled by SMC and NNSMC

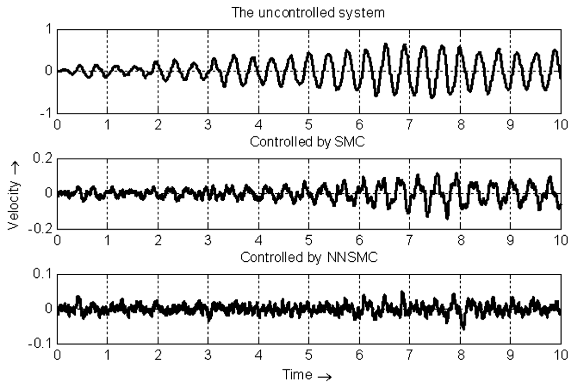


Fig. 10 Velocity of the second DOF of uncontrolled system, system controlled by SMC and NNSMC

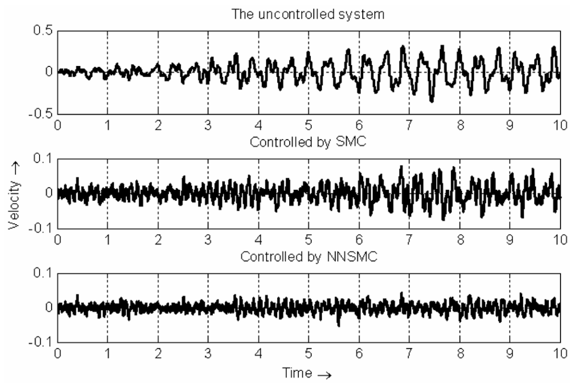


Fig. 11 Velocity of the third DOF of uncontrolled system, system controlled by SMC and NNSMC

Table 3 The max values of the three kinds of system

Max values	Uncontrolled system	Controlled by SMC	Controlled by NNSMC
State 1	0.0519	0.0087	0.0023
State 2	0.0395	0.0059	0.0017
State 3	0.0148	0.0031	0.0010
State 4	0.9677	0.1514	0.0751
State 5	0.6475	0.1126	0.0481
State 6	0.3188	0.0799	0.0446

number of iterations is 200. The active function is hyperbolic tangent sigmoid transfer function. The max allowable control force equals to the 20% of the max value of the external disturbance. Simulation results are given in Figs. 6-11. Even though the system is subjected to parameter uncertainty and external disturbance, the vibration can be depressed by using either SMC scheme or NNSMC(Neural network sliding mode control) scheme. But the control quality of NNSMC is far better than SMC.

The Figs. 6-8 are results of displacement of uncontrolled system, system controlled by SMC and system controlled by NNSMC. In Fig. 6, the max displacement values are 0.0519, 0.0087 and 0.0023, which are corresponding to the first degree of freedom of the uncontrolled system, system controlled by SMC and NNSMC, respectively. The other states are listed in the Table 3. From the Table 1, it can be seen that the NNSMC scheme is more effective than the SMC.

7. Conclusions

In this paper, a new sliding mode control scheme is presented based on the hybrid model of the neural network and the nominal system, which has been proved to be very effective for the uncertain, multi-input and multi-output system. Although SMC can reduce the vibration of flexible structures, the control quality of NNSMC is far better than SMC.

The discrete system model of the flexible structure has been systematically constructed in this paper. Based on the discrete system, SMC has been presented to control the vibration of the flexible structure. To decrease the thickness of the boundary layer or eliminate the chattering coming from the parameters uncertainties, the neural network has been imported into the model.

Finally, the simulation results presented demonstrate that the algorithm discussed is a very promising method for the uncertain system.

Acknowledgements

The authors are grateful for supports from the National Natural Science Foundation of China (10572119, 10632030 and 50625516), the Program for New Century Excellent Talents in University (NCET-04-0958), the National Fundamental Research Program of China (973) (2003CB716207) and the Open Foundation of State Key Laboratory of Structural Analysis of Industrial Equipment of China.

References

Cavallo, A., Maria, G.D. and Setola, R. (1999), “A sliding manifold approach for vibration reduction of flexible systems”, *Automatica*, **35**(10), 1689-1696.

Cheng, X.P. and Patel, R.V. (2003), “Neural network based tracking control of a flexible macro-micro manipulator system”, *Neural Networks*, **16**(2), 271-286.

Cybenko, G. (1989), “Approximation by superpositions of a sigmoidan function”, *Math. Control Signal*, **2**(4), 303-314.

DeCarlo, R.A., Zak, S.H. and Matthews, G.P. (1988), “Variable structure control of nonlinear multivariable systems: A tutorial”, *Proc. of the IEEE*, **76**(3), 212-232.

Edwards, C. and Spurgeon, S.K. (1998), *Sliding Mode Control: Theory and Applications*, Taylor & Francis, London.

El-Sinawi, A.H. (2004), “Active vibration isolation of a flexible structure mounted on a vibrating elastic base”, *J. Sound Vib*., **271**(1-2), 323-337.

Ertugrul, M. and Kaynak, O. (2000), “Neuro sliding mode control of robotic manipulators”, *Mechatronics*, **10**(1-2), 239-263.

Funahashi, K.-I. (1989), “On the approximate realization of continuous mappings by neural networks”, *Neural Networks*, **2**(3), 183-192.

Hang, C., Ih, C. and Wang, S.J. (1985), “Application of adaptive control to space station”, AIAA-85-1970. 707-724.

Hornik, K. (1991), “Approximation capabilities of multilayer feedforward networks”, *Neural Networks*, **4**(2), 251-257.

Huang, Y.A. and Deng, Z.C. (2005), “Decentralized sliding mode control for a spacecraft flexible appendage based on fintie element method”, *Chinese J. Aeronaut.*, **18**(3), 230-236.

Hung, J.Y., Gao, W.B. and Hung, J.C. (1993), “Variable structure control: A survey”, *IEEE T. Ind. Electron.,* **40**(1), 2-22.

Hussain, M.A. and Ho, P.Y. (2004), “Adaptive sliding mode control with neural network based hybrid models”, *J. Process Control*, **14**(2), 157-176.

Jung, W.J., Jeong, W.B., Hong, S.R. and Choi, S.B. (2004), “Vibration control of a flexible beam structure using squeeze-mode ER mount”, *J. Sound Vib*., **273**(1-2), 185-199.

Lee, H., Kim, E. and Kang, H.-J. (2001), “A new sliding-mode control with fuzzy boundary layer”, *Fuzzy Set Syst.*, **120**(1), 135-143.

Qiu, Y.Y., Duan, B.Y., Wei, Q., Nan, R.D. and Peng, B. (2002), “Optimal distribution of the cable tensions and structural vibration control of the cable-cabin flexible structure”, *Struct. Eng. Mech.,* **14**(1), 39-56.

Stavroulakis, G.E., Foutsitzi, G., Hadjigeorgiou, E., Marinova, D. and Baniotopoulos, C.C. (2005), “Design and robust optimal control of smart beams with application on vibrations suppression”, *Adv. Eng. Softw.*, **36**(11-12), 806-813.

Suleman, A. and Costa, A.P. (2004), “Adaptive control of an aeroelastic flight vehicle using piezoelectric actuators”, *Comput. Struct.,* **82**(17-19), 1303-1314.

Wai, R.-J., Lin, C.-M. and Hsu, C.-F. (2004), “Adaptive fuzzy sliding-mode control for electrical servo drive”, *Fuzzy Set Syst.,* **143**(2), 295-310.

Yazdizadeh, A. and Khorasani, K. (2002), “Adaptive time delay neural network structures for nonlinear system identification”, *Neurocomputing*, **47**(1-4), 207-240.