

DOI:10.11918/202504054

融合 CNN 和 BiLSTM 的挠性太阳帆时变惯性张量辨识

夏子阅,倪智宇,董凯凯,沈宇豪,王士新

(沈阳航空航天大学 航空宇航学院,沈阳 110136)

摘要: 为准确辨识挠性太阳帆航天器在轨展开过程中的惯性张量参数,并提升复杂干扰环境下参数辨识模型的精度和鲁棒性,结合卷积神经网络(CNN)的局部特征提取能力和双向长短期记忆网络(BiLSTM)的时间序列建模优势,提出了一种基于CNN-BiLSTM混合网络的数据驱动型参数辨识方法。首先,建立挠性太阳帆航天器的姿态-振动耦合动力学模型,并利用领域随机法生成大量训练数据。其次,构建CNN-BiLSTM网络参数辨识模型,并采用预先设计的训练策略对其进行训练。最后,使用训练完成的CNN-BiLSTM模型在不同的测量噪声条件下辨识太阳帆的惯性张量,并比较分析了该模型和单一模型的辨识结果。仿真结果表明,所构建的网络模型在无干扰的条件下,辨识出的系统惯性张量参数的平均相对误差低于1%;而在具有干扰力矩和较低信噪比的复杂测量噪声条件下,辨识结果的平均相对误差仍不超过1.5%,性能显著优于CNN和BiLSTM模型。验证了本文方法能有效克服单一网络模型的不足,不仅能准确辨识出挠性太阳帆的时变惯性张量,还能增强在复杂干扰环境中的鲁棒性。

关键词: 太阳帆;参数辨识;惯性张量;卷积神经网络;双向长短期记忆网络

中图分类号: V448

文献标志码: A

文章编号: 0367-6234(2026)04-0023-09

Identification of time-varying inertia tensor of flexible solar sail based on CNN and BiLSTM

XIA Ziyue, NI Zhiyu, DONG Kaikai, SHEN Yuhao, WANG Shixin

(College of Aerospace Engineering, Shenyang Aerospace University, Shenyang 110136, China)

Abstract: To accurately identify the inertia tensor parameters of flexible solar sail spacecraft during on-orbit deployment and improve the accuracy and robustness of the parameter identification model in complex disturbance environments, this paper proposed a data-driven parameter identification method based on a convolutional neural network (CNN) and bidirectional long short-term memory (BiLSTM) hybrid network. This method combined the local feature extraction capability of CNN with the time series data modeling advantage of BiLSTM. First, the paper established the attitude-vibration coupling dynamics model of the flexible solar sail spacecraft and generated massive training data using domain randomization. Second, a CNN-BiLSTM network parameter identification model was constructed and trained using a pre-designed training strategy. Finally, the trained model was applied to identify the inertia tensor of the solar sail under different measurement noise conditions, and the identification results of this model and single models were compared and analyzed. Simulation results demonstrate that the constructed network model has a mean relative error of less than 1% in identifying the system inertia tensor parameters under disturbance-free conditions. In addition, under complex measurement noise conditions with disturbance torques and lower signal-to-noise ratios, the mean relative error of the identification results still does not exceed 1.5%, which is significantly better than that of the CNN and BiLSTM models. The proposed method effectively addresses the shortcomings of single network models, not only accurately identifying the time-varying inertia tensor of the flexible solar sail but also enhancing robustness in complex disturbance environments.

Keywords: solar sail; parameter identification; inertia tensor; convolutional neural network; bidirectional long short-term memory network

收稿日期: 2025-04-19; 录用日期: 2025-08-05; 网络首发日期: 2026-02-03

网络首发地址: <https://link.cnki.net/urlid/23.1235.t.20260202.1457.004>

基金项目: 国家自然科学基金(62388101); 辽宁省自然科学基金(2024-BS-153); 辽宁省教育厅基本科研项目(JYTMS20230254)

作者简介: 夏子阅(2001—),男,硕士研究生

通信作者: 倪智宇, nizhiyu@sau.edu.cn

太阳帆作为一种以光压为动力的新型航天器,因其无需消耗燃料、长期运行等优势,在诸多领域中备受青睐^[1-3]。挠性太阳帆在轨展开过程中,其自身构型的改变导致整体的惯性张量参数随时间发生变化。若无法实时准确获取惯性张量的变化信息,那么当展开过程中发生非预期情况时,将难以实施有效的姿态调整进而导致其失稳。因此有必要采用在轨辨识等手段对其时变惯性张量进行有效准确的辨识,以修正航天器的控制系统参数,从而确保挠性太阳帆展开过程中的安全性。

随着以深度学习为代表的驱动型方法的迅猛发展,神经网络算法已经开始广泛应用于航天领域^[4]。在系统辨识方向,由于其具有从掺杂复杂噪声的数据中提取出鲁棒特征的能力^[5],以及不需要系统模型的先验信息、依靠系统的输入输出数据即可实现辨识的优势^[6],因此逐渐被研究人员所关注。目前针对基于深度学习的航天器惯性参数辨识问题已有相关的研究成果。孙俊等^[7]针对传统辨识方法的运算量较大的问题,提出了一种基于归一化最小均方准则的 Adaline 神经网络方法用于辨识航天器的惯性参数。金晨迪等^[8]提出了一种基于卷积神经网络(convolutional neural network, CNN)的组体航天器的惯性参数辨识方法。赵光玖^[9]利用 CNN 对航天器的转动惯量进行辨识。Chu 等^[10]基于深度神经网络(deep neural network, DNN)实现了对具有复杂角速度测量噪声的组体航天器的惯性矩阵辨识。Wu 等^[11]在文献[10]的基础上开展了进一步的研究,为提高 DNN 的计算效率,建立了一种优化 DNN 结构的惯性矩阵辨识模型。初未萌等^[12]利用长短期记忆网络(long short-term memory, LSTM)实现了对空间机器人抓捕过程中的系统惯性张量辨识。通过分析上述文献可知,现有研究普遍采用单一网络模型进行辨识,这些模型在面临低信噪比、复杂噪声类型和外部干扰力矩的环境时,其辨识精度和鲁棒性能有限。此外当前成果多集中于刚体航天器模型,关于利用深度学习方法辨识具有时变和姿态-振动耦合特性的挠性航天器的惯性参数的研究则相对匮乏。

针对上述问题,本文提出了一种融合 CNN 和双向 LSTM(bidirectional long short-term memory, BiLSTM)网络的混合网络模型——CNN-BiLSTM,以提升在复杂环境干扰下的挠性太阳帆的辨识精度与鲁棒性。本文方法充分利用 CNN 在提取数据重要特征方面的优势,以及 BiLSTM 在捕捉时间序列依赖关系方面的强大能力,构建了具有双重抗噪机制

和强大的特征提取能力的混合神经网络辨识模型。该模型不仅能准确辨识挠性太阳帆的时变惯性张量,更能增强辨识模型在复杂干扰环境下的抗干扰能力。仿真结果表明,在复杂噪声环境中所提出的 CNN-BiLSTM 参数辨识模型对太阳帆时变惯性张量参数的辨识精度要显著优于单一的 CNN 和 BiLSTM 网络模型,有效提升了在低信噪比的复杂测量条件下对挠性太阳帆惯性参数辨识的精度与鲁棒性。同时仿真结果验证了此方法的可行性,为后续的实际工程应用提供了理论依据。

1 太阳帆航天器的动力学模型描述

1.1 模型简化与坐标系

太阳帆航天器的构型和相关坐标系如图 1 所示,采用方形构型,由展开机构和 12 U 立方星平台组成。展开机构主要包括 4 根支撑杆、帆膜和相关部件,并收纳于立方星平台内。本文为减小建模复杂度,对航天器作出以下假设:1)将航天器模型简化为具有挠性附件的中心刚体模型,其中中心刚体结构视为长方体,支撑杆等效为固定端约束的悬臂梁模型,忽略其轴向变形;2)忽略帆膜与支撑杆之间的相互作用,并将帆膜的质量均匀等效分配到 4 根支撑杆上;3)仅分析挠性附件振动对系统姿态转动的影响,忽略其对系统平动的作用。

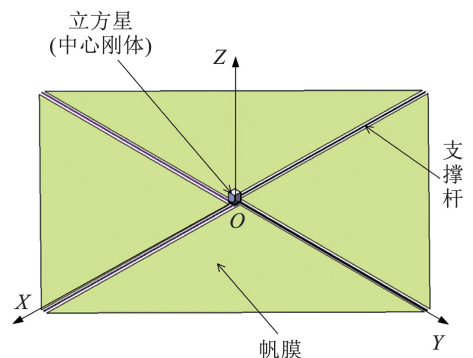


图 1 太阳帆航天器的构型

Fig. 1 Configuration of solar sail spacecraft

定义太阳帆航天器的本体坐标系 $OXYZ$ 的原点 O 位于帆膜的几何中心, OZ 垂直于帆膜并指向立方星一侧, OX 与支撑杆平行并指向支撑杆的远端, OY 则与前两者组成右手正交系;惯性系 $O_e X_e Y_e Z_e$ 的原点 O_e 位于地球质心, $O_e X_e$ 在赤道平面内指向春分点, $O_e Z_e$ 与地球的自转轴重合, $O_e Y_e$ 则与前两者组成右手正交系。

1.2 太阳帆姿态-振动耦合动力学方程

太阳帆在轨展开过程中,当姿态角的变化较小

时,航天器的姿态-振动耦合动力学模型可表示为^[13]:

$$J_s(t)\ddot{\boldsymbol{\varphi}} + \sum_{i=1}^4 \mathbf{P}(t)\ddot{\boldsymbol{\eta}}_i = \boldsymbol{\tau}(t) \quad (1)$$

$$\ddot{\boldsymbol{\eta}}_i + 2\xi_i\boldsymbol{\Omega}_i(t)\dot{\boldsymbol{\eta}}_i + \boldsymbol{\Omega}_i^2(t)\boldsymbol{\eta}_i + \mathbf{P}^T(t)\ddot{\boldsymbol{\varphi}} = 0 \quad (2)$$

式中: $J_s(t)$ 为航天器的时变惯性张量, $\boldsymbol{\varphi}$ 为航天器的姿态角, $\mathbf{P}(t)$ 为支撑杆相对于航天器转动的刚柔耦合系数矩阵, $\boldsymbol{\eta}_i$ 为第*i*个支撑杆的模态坐标, $\boldsymbol{\tau}(t)$ 为输入力矩, ξ_i 、 $\boldsymbol{\Omega}_i(t)$ 分别为第*i*个支撑杆的模态阻尼比矩阵和模态频率对角矩阵。

定义向量 $\mathbf{x} = [\boldsymbol{\varphi}^T, \dot{\boldsymbol{\varphi}}^T, \boldsymbol{\eta}^T, \dot{\boldsymbol{\eta}}^T]^T$, 矩阵 $\mathbf{K}(t) = \boldsymbol{\Omega}_i^2(t)$ 、 $\mathbf{C}(t) = 2\xi_i\boldsymbol{\Omega}_i(t)$, 则式(1)、(2)可以转化为如下的状态方程形式:

$$\dot{\mathbf{x}} = (\mathbf{D}^{-1}\mathbf{A})\mathbf{x} + (\mathbf{D}^{-1}\mathbf{B})\mathbf{u} \quad (3)$$

其中:

$$\mathbf{D} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & J_s(t) & \mathbf{0} & \mathbf{P}(t) \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^T(t) & \mathbf{0} & \mathbf{I} \end{bmatrix}, \mathbf{u} = \boldsymbol{\tau}(t),$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & -\mathbf{K}(t) & -\mathbf{C}(t) \end{bmatrix}, \mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

通过式(3)可以计算得到太阳帆的姿态角速度 $\dot{\boldsymbol{\varphi}}$ 和角加速度 $\ddot{\boldsymbol{\varphi}}$ 响应信号,以用于后续网络模型的训练和测试。在实际的在轨航天器系统中,角速度通常可以由陀螺仪测量获得,而角加速度可以通过对测量出的角速度进行一次微分并结合滤波算法来获取。

2 基于 CNN-BiLSTM 网络的参数辨识方法

2.1 数据集的生成

本文提出的 CNN-BiLSTM 参数辨识模型是一种数据驱动方法,因此需要大量的数据来训练网络模型。

领域随机法(domain randomization)是深度学习一种常用的数据生成方法,其目的是通过在源域中训练出的模型泛化到目标域中,即将设定的参数空间视为目标域,在该空间内随机选取足够多的参数视为源域,通过使用源域内的参数所训练的模型,即可在目标域中取得较好的泛化能力^[14]。数据集的具体生成方法如下。

1) 通过查阅立方星的设计规范文件,确定太阳

帆航天器的质量变化范围,并从该范围内随机生成 N 组不同的惯性张量的初始值 $J_s^k(0)$ ($k = 1, 2, \dots, N$),再根据太阳帆的展开过程计算出 N 组航天器的时变惯性张量 $J_s^k(t)$ 。

2) 根据式(3),对具有不同 $J_s^k(t)$ 的太阳帆航天器施加相同的力矩 $\boldsymbol{\tau}(t)$ 激励,以获取每组时变惯性张量 $J_s^k(t)$ 所对应的状态-控制序列 $s^k(t)$ ($k = 1, 2, \dots, N$),其中每个时刻 $s^k(i) = [\dot{\boldsymbol{\varphi}}^k(i), \ddot{\boldsymbol{\varphi}}^k(i), \boldsymbol{\tau}^k(i)]$ 。

3) 将序列 $s^k(t)$ 和惯性张量 $J_s^k(t)$ 组合成数据集的数据样本。

为了使网络模型能够捕捉时序数据中的时间依赖信息,需要利用滑动窗口技术来构建数据样本的输入部分和标签。即宽度为 L 的滑动窗口沿状态-控制序列 $s^k(t)$ 向前移动到下一个时刻,直到序列数据的末尾,在此过程中,窗口每次移动时截取的维度为 L 的子序列 $\mathbf{x}^k(t) = [s^k(t-L), s^k(t-L+1), \dots, s^k(t)]$ 作为数据样本的输入部分,而 t 时刻的航天器惯性张量 J_s^k 则作为数据样本的标签。

由于数据集中包含不同量纲的数据,直接传入神经网络可能导致网络难以训练,因此在构建数据样本的输入部分和标签之前,需要对每个状态-控制序列和对应的时变惯性张量进行无量纲化处理。本文选择 Z-score 标准化进行处理:

$$\hat{z}_i(t) = \frac{z_i(t) - \mu_i}{\sigma_i} \quad (i = 1, 2, \dots, N) \quad (4)$$

式中: $z_i(t)$ 为第*i*个需要处理的原始数据, μ_i 、 σ_i 分别为 $z_i(t)$ 的均值和标准差。经过式(4)处理后,利用滑动窗口技术构建的 $\mathbf{x}^k(t)$ 和对应时刻的惯性张量 J_s^k 将重新作为数据样本的输入部分和标签。

2.2 CNN-BiLSTM 模型的搭建

图2为本文建立的 CNN-BiLSTM 参数辨识模型的具体结构,使用深度学习框架 Tensorflow 搭建而成。主要由 CNN 和 BiLSTM 网络串联组成,其中 CNN 网络由两个一维卷积层和两个最大池化层构成,且卷积层与池化层交替连接;而 BiLSTM 网络则由一个双向 LSTM 层、两个全连接层和输出层组成。

输入数据 $\mathbf{x}^k(t)$ 传入神经网络后,卷积核将对其进行卷积运算,接着通过池化层对提取出的高阶特征进行池化操作,以获得数据的关键特征。在经过多次卷积和池化操作后,CNN 将筛选出输入数据的重要特征信息。同时降低数据维度,从而提高网络模型的训练速度^[15]。

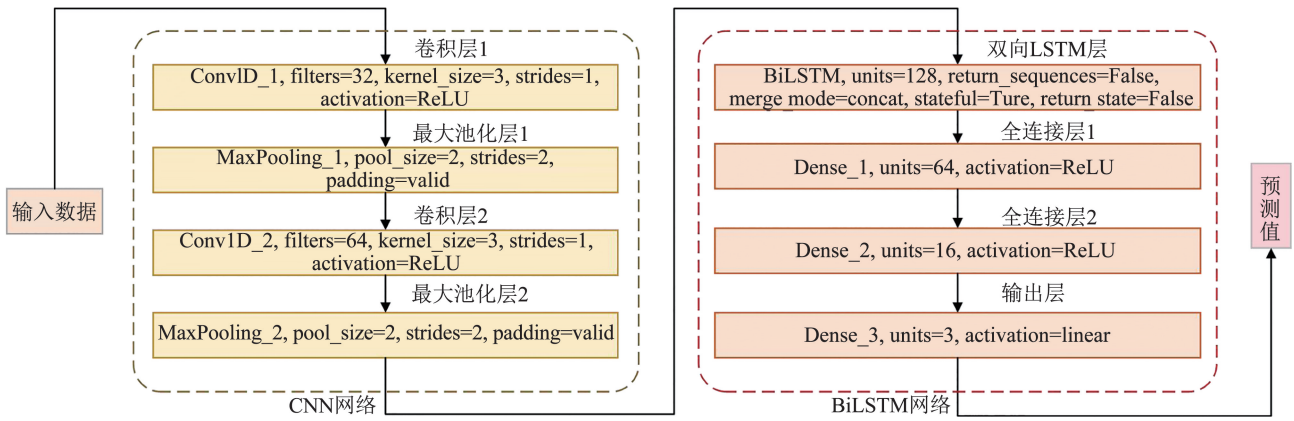


图 2 CNN-BiLSTM 参数辨识模型的具体结构

Fig. 2 Detailed structure of CNN-BiLSTM parameter identification model

提取出的重要特征信息将作为双向 LSTM 层的输入,该网络层的具体结构如图 3 所示,由一个前向传播和一个反向传播的 LSTM 层组成,两者之间既相互独立又同时传播。任意单独的 LSTM 层都是由多个 cell 单元组成,内部含有输入门、遗忘门和输出门结构,这 3 个门结构控制特征信息的传递和储存,其计算过程如下^[16]:

$$\begin{cases} i_t = \text{sigmoid}(W_i h_{t-1} + U_i \hat{s}_t + b_i) \\ f_t = \text{sigmoid}(W_f h_{t-1} + U_f \hat{s}_t + b_f) \\ o_t = \text{sigmoid}(W_o h_{t-1} + U_o \hat{s}_t + b_o) \\ g_t = \tanh(W_g h_{t-1} + U_g \hat{s}_t + b_g) \\ c_t = f_t c_{t-1} + i_t g_t \\ h_t = o_t \tanh(c_t) \end{cases} \quad (5)$$

式中: \hat{s}_t 为当前时刻的输入向量,即 CNN 提取出的特征信息; i_t, f_t, o_t 分别为 cell 的输入门值、遗忘门值

和输出门值(t 表示时间步); g_t, c_t, h_t 分别为 cell 产生的候选记忆、新记忆和输出的隐藏状态; W_i, W_f, W_o, W_g 分别为输入门、遗忘门、输出门和产生候选记忆时所需的循环权重矩阵; U_i, U_f, U_o, U_g 分别为输入门、遗忘门、输出门和产生候选记忆时所需的输入权重矩阵; b_i, b_f, b_o, b_g 分别为输入门、遗忘门、输出门和产生候选记忆时所需的偏置向量。

与传统的 LSTM 层相比,双向 LSTM 层输出的隐藏状态 h_t^{bi} 都由前向传播 LSTM 的隐藏状态 h_t^f 和反向传播 LSTM 的隐藏状态 h_t^b 拼接而成^[17]。这一特点可以使其同时捕捉数据中的正向和反向的时间依赖信息,以便深入挖掘当前时刻数据与过去及未来时刻数据的内在联系,从而全面利用数据的特征来提高网络模型的预测性能^[18]。

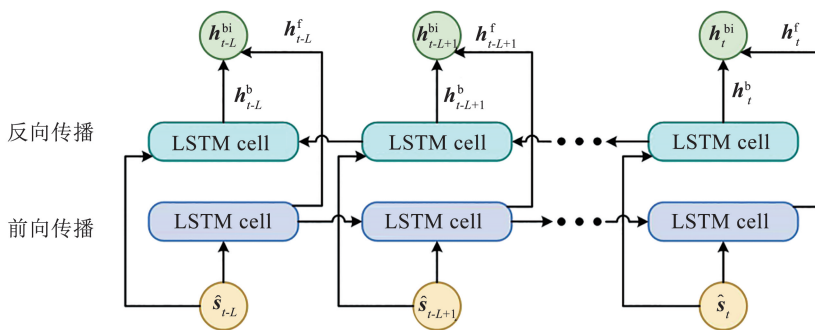


图 3 双向 LSTM 层的典型结构

Fig. 3 Typical structure of BiLSTM layer

网络模型的最后部分则由多个全连接层和一个输出层组成,该部分将双向 LSTM 层的输出隐藏状态 h_t^{bi} 映射到与航天器惯性张量维度相同的空间,从而得到惯性张量预测值。每层的计算过程如下:

$$y_i = f(y_{i-1} W_i + b_i) \quad (6)$$

式中: y_{i-1}, y_i 分别为上一层的输出向量和当前层的

输出向量, W_i, b_i 分别为当前层的权重矩阵和偏置向量, $f(\cdot)$ 为激活函数。

2.3 模型的训练与辨识过程

图 4 为 CNN-BiLSTM 模型的训练与辨识过程,网络模型需要经过训练才能实现参数辨识,具体的训练过程步骤如下。

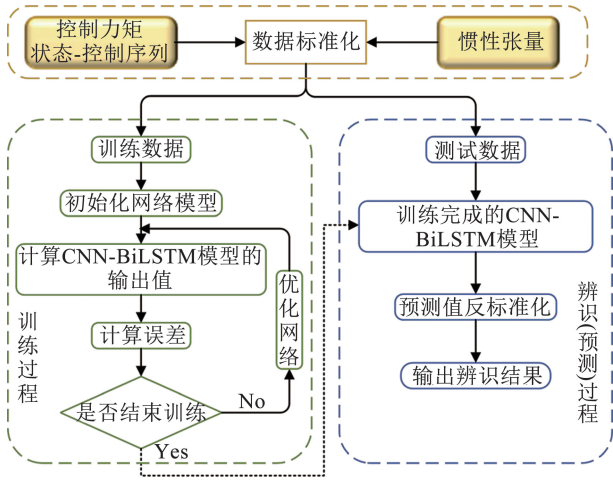


图4 CNN-BiLSTM 模型的训练与辨识过程

Fig. 4 Training and identification process of CNN-BiLSTM model

Step 1 (采集训练数据) 基于所述的数据集生成方法获得 Z-score 标准化处理后的训练数据集。

Step 2 (网络初始化) 网络开始训练时,首先对其每一层的权重和偏置进行初始化。对于卷积层和全连接层采用 He 初始化方法^[19];而对于双向 LSTM 层则采用 Xavier 初始化方法。

Step 3 (计算网络模型的输出值) 输入数据依次经过各网络层的处理后,得出网络模型的输出值。

Step 4 (计算误差) 通过设定的损失函数,计算网络的输出值与标签值之间的误差。由于参数辨识属于回归问题,因此选择均方误差 (mean-square error, MSE) 作为网络模型的损失函数^[12]。

Step 5 (判断是否结束训练) 如果网络达到预定的训练次数,则训练结束;否则,继续训练直到条件满足。

Step 6 (优化网络) 网络如果继续训练,将通过 Adam 优化方法,更新其权重和偏置,然后返回 Step3 直到训练结束。

训练完成后,网络模型学习到的知识将用于其预测,而网络模型的辨识过程本质上就是其预测过程。具体的辨识过程如下:首先依据数据集生成的方法获得标准化处理后的测试数据集;其次将其传送到训练完成的网络模型中进行计算;最后利用下式将计算出的预测值进行还原,所得结果即为航天器惯性张量的辨识值。

$$y = \sigma_L y_p + \mu_L \quad (7)$$

式中: y 为还原后的数据,也是惯性张量的辨识值; y_p 为网络的预测值; μ_L 、 σ_L 分别为标签的均值和标准差。

3 仿真实验验证

3.1 算例模型与数据生成

生成数据所使用的太阳帆航天器模型参数见

表 1,通过查阅立方星设计规范,可知 12 U 立方星的质量在 12 ~ 24 kg 之间。在仿真过程中,由于受支撑杆材料和尺寸等影响,太阳帆在展开初期仍可将其视为刚体模型,因此本文仅考虑支撑杆从伸展 20% 状态到完全展开状态的情况。此外,由于本体坐标系建立在航天器的惯性主轴上,因此只对转动惯量 J_x 、 J_y 和 J_z 进行辨识。

表 1 太阳帆航天器模型的参数

Tab. 1 Parameters of solar sail spacecraft model

参数	数值
立方星尺寸/m	0.23 × 0.23 × 0.34
整星质量范围/kg	[12, 24]
支撑杆全展开尺寸/m	7.00 × 5.03 × 10 ⁻² × 3.30 × 10 ⁻⁴
单根支撑杆质量/kg	0.152
支撑杆阻尼比	1 × 10 ⁻²
支撑杆展开速度/(m·s ⁻¹)	5.6 × 10 ⁻²
帆膜总质量/kg	0.231

由于式(1)、(2)中的刚柔耦合系数 $P(t)$ 和模态频率矩阵 $\Omega(t)$ 在展开过程中是时变且难以确定的,在算例中为了降低计算的复杂度,并保持其时变特性,本文对动力学模型进行了必要的简化:只选取支撑杆的前 3 阶振动模态频率进行研究,并采用离散化的思想获得展开过程中的 $\Omega(t)$ 和 $P(t)$,这是工程中常用的一种处理时变参数的近似策略。具体而言,分别计算当支撑杆展开到 20%、40%、60%、80% 和完全展开时的 5 种状态下对应的刚柔耦合系数矩阵和模态频率值,利用这 5 种状态下每一段时间内的定常耦合系数来代替该时间段内各个时间节点对应的时变耦合系数;而系统的模态频率值采用分段 3 次 Hermite 插值方法,利用上述 5 种情况下的频率值,插值得到该系统的时变频率变化结果。各时间段内的刚柔耦合系数矩阵见附录,而前 3 阶模态频率值如图 5 所示。

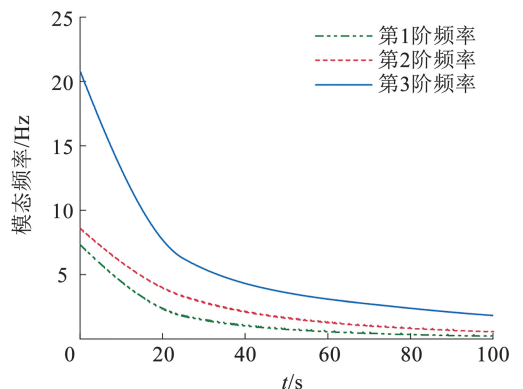


图 5 支撑杆的前 3 阶模态频率

Fig. 5 First three modal frequencies of support rod

仿真时间设为 100 s, 采样频率为 10 Hz。为激发太阳帆的状态响应, 将输入力矩设为如图 6 所示的正弦信号。太阳帆的姿态角和角速度的初始值均为 0, 支撑杆的初始模态坐标参数为 $\boldsymbol{\eta}(0) = \mathbf{0}, \dot{\boldsymbol{\eta}}(0) = \mathbf{0}$ 。设定滑动窗口的宽度 L 为 50, 结合给定的输入力矩和表 1 中的参数, 基于数据集生成所述的方法, 生成了 200 组数据用于训练网络, 50 组数据用于测试网络。

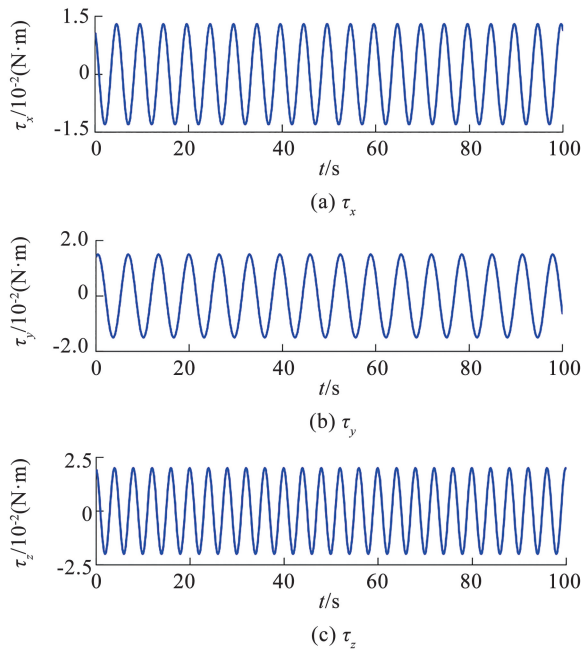


图 6 输入力矩

Fig. 6 Input torque

3.2 算法有效性验证

基于预先设计的模型训练流程对网络模型进行训练, 设定 Adam 优化方法的学习率为 1×10^{-4} , 训练过程中每次迭代的批次大小为 50, 对整个网络进行 50 次训练。训练所采用的计算平台为 Nvidia RTX 4070 Ti, 训练时间约为 712 s。图 7 为网络训练时的均方误差变化曲线, 由图 7 结果可知, 随着训练次数的增加, 均方误差逐渐收敛并维持在较低水平。训练完成后, 网络模型在训练集与测试集上的均方误差非常接近, 分别为 2.239×10^{-5} 和 2.308×10^{-5} 。这表明所采用的优化方法适用于所构建的网络模型, 且模型具有较好的泛化性。

本文选取测试集中的一组整星质量为 22.5 kg 时的数据对 CNN-BiLSTM 参数辨识模型的有效性进行验证。为了定量评价辨识性能, 采用平均相对误差 (mean relative error, MRE) 作为评价准则, 其公式为

$$R_{ME} = \frac{1}{m} \sum_{i=1}^m \left| \frac{J_r - \tilde{J}}{J_r} \right| \times 100\% \quad (8)$$

式中: R_{ME} 为平均相对误差, m 为采样次数, J_r, \tilde{J} 分别为转动惯量的理论值和辨识值。图 8 为太阳帆在轨展开过程中的时变惯性张量辨识结果, 计算结果表明 CNN-BiLSTM 网络模型准确辨识出了航天器的时变惯性张量, 其所计算出的辨识值曲线与理论值曲线几乎重合, 辨识结果的平均相对误差低于 1%。

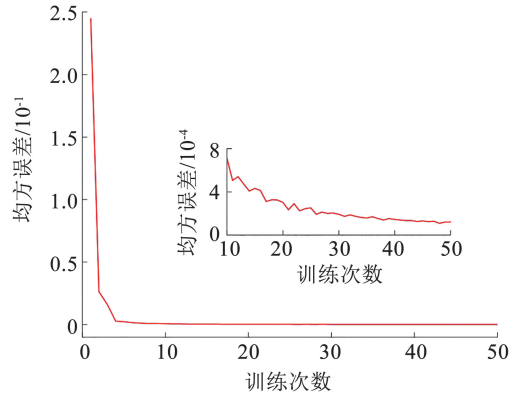


图 7 网络训练过程中的均方误差变化

Fig. 7 Mean square error variation during network training

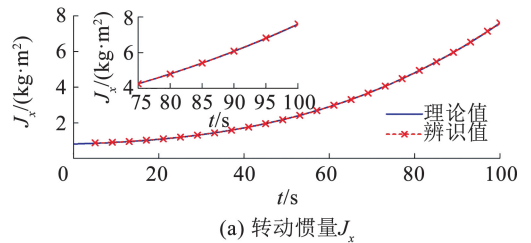
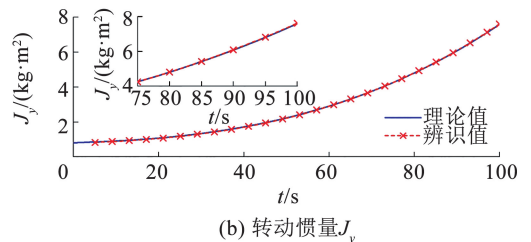
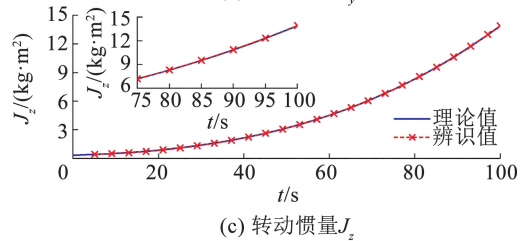
(a) 转动惯量 J_x (b) 转动惯量 J_y (c) 转动惯量 J_z

图 8 CNN-BiLSTM 模型在无干扰条件下的辨识结果

Fig. 8 Identification results of CNN-BiLSTM model under disturbance-free conditions

上述的仿真算例中, 并没有考虑外界干扰对算法的影响, 而在实际情况中, 航天器会不可避免地受到外界干扰, 接下来将对 CNN-BiLSTM 参数辨识模型的抗干扰能力进行讨论。此外, 为了体现该模型的优越性, 还将其与单一 CNN 和 BiLSTM 网络模型的辨识性能进行对比。其中 CNN 模型不包含双

向 LSTM 结构, 而 BiLSTM 模型则未采用卷积与池化结构。除此之外, 3 种模型在超参数设置以及训练策略等方面均保持一致。为了模拟空间环境干扰, 在输入力矩 $\tau(t)$ 中加入如图 9 所示的干扰力矩 τ_d , 同时对姿态角速度和角加速度加入混合测量噪声 δ_n , 其表达式分别为:

$$\delta_n = \delta_w + \delta_c \quad (9)$$

$$\delta_w = \mu_w(t) \quad (10)$$

$$\delta_c = \mu_w(t) + 0.05\mu_w(t-1) \quad (11)$$

式中: δ_w 、 δ_c 分别为高斯白噪声和高斯有色噪声, μ_w 则是均值为 0、方差为信号强度百分比的高斯随机信号。

在干扰条件下航天器时变惯性张量的辨识结果如图 10 所示, 表 2 给出了 3 种网络模型在不同条件下的平均相对误差。结果表明, 相较于单一网络模型, 本文所提方法在受到干扰力矩及方差为 0.3%

信号强度的测量噪声影响时, 其参数辨识精度仍能保持较高水平, 平均相对误差与理想无干扰条件下的结果十分接近, 且两种情况下的平均相对误差均控制在 1.0% 以内。

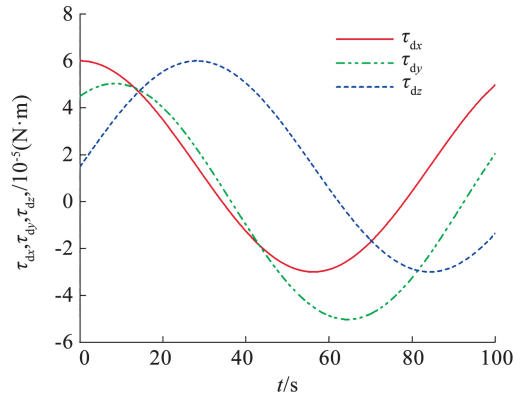
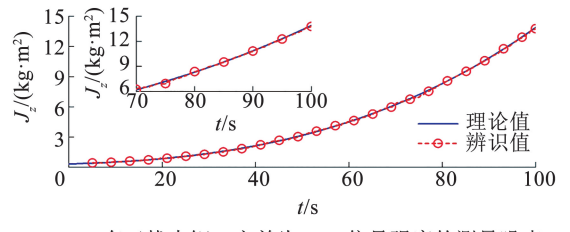
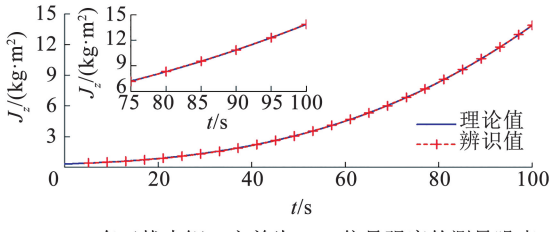
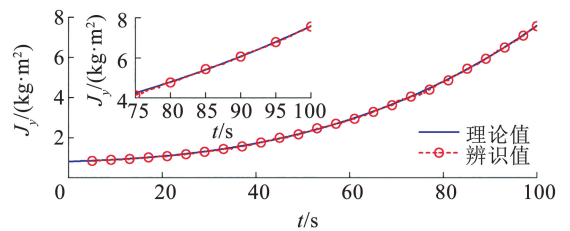
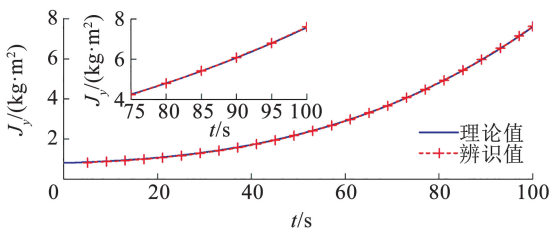
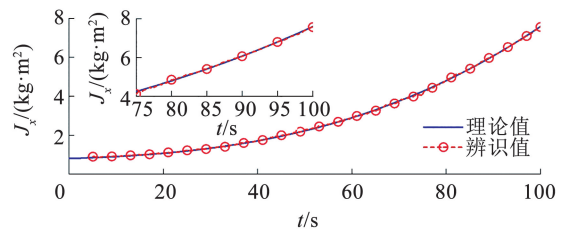
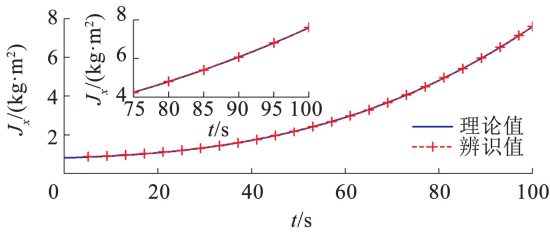


图 9 干扰力矩 τ_d

Fig.9 Disturbance torque τ_d



(a) 有干扰力矩、方差为0.3%信号强度的测量噪声

(b) 有干扰力矩、方差为3.0%信号强度的测量噪声

图 10 CNN-BiLSTM 模型在外界干扰条件下的辨识结果

Fig.10 Identification results of CNN-BiLSTM model under external disturbance conditions

随着测量噪声量级的增大(方差为 3.0% 的信号强度), 3 种网络模型的辨识精度都有所降低, 但本文提出的 CNN-BiLSTM 网络模型仍然保持了较高的辨识精度, 平均相对误差在 1.5% 以下。这都归因于混合网络架构设计: 通过卷积层和池化层实现对噪声的初步过滤, 再结合双向 LSTM 层从含噪声的数据中进行时序信息的深度挖掘, 从而形成双重

抗噪机制; 相比之下, BiLSTM 网络模型虽然可以利用门控机制和双向信息处理能力来过滤部分噪声, 但由于局部特征提取能力的不足, 其噪声抑制能力减弱; 而 CNN 网络模型则因噪声干扰造成数据的局部特征失真, 同时缺乏对时序数据的时间依赖信息捕捉能力, 导致其性能急剧下降。

表 2 3 种网络模型在不同条件下惯性张量辨识结果的平均相对误差

Tab. 2 Mean relative errors of inertia tensor identification results for three network models under different conditions

辨识参数	无干扰力矩和测量噪声/%			有干扰力矩, 方差为 0.3% 信号强度的测量噪声/%			有干扰力矩, 方差为 3.0% 信号强度的测量噪声/%		
	CNN-BiLSTM	CNN	BiLSTM	CNN-BiLSTM	CNN	BiLSTM	CNN-BiLSTM	CNN	BiLSTM
J_x	0.280	0.367	0.848	0.362	0.660	0.965	0.920	4.410	1.914
J_y	0.739	0.366	1.719	0.762	0.649	1.820	1.000	4.391	2.607
J_z	0.553	0.828	2.458	0.632	1.212	2.612	1.363	6.197	4.675

综上所述,本文构建的参数辨识模型性能表现优异,能够有效弥补单一 CNN 和 BiLSTM 网络模型存在的缺陷,不仅能准确辨识出太阳帆系统的时变惯性张量参数,还具有较好的鲁棒性。

4 结 论

1) 本文将 CNN 的数据特征提取能力和 BiLSTM 强大的双向时间序列建模能力相融合,通过领域随机法生成的训练数据和设计的训练策略优化网络参数,构建了一种基于 CNN-BiLSTM 混合网络的数据驱动型参数辨识模型,实现了对挠性太阳帆航天器在轨展开过程中的时变惯性张量参数的准确辨识。

2) 相较于传统 CNN 和 BiLSTM 单一网络模型,本文提出的 CNN-BiLSTM 参数辨识模型由于结合两者优势,在辨识性能上取得了显著提升。结果表明,该模型在无干扰的理想工况下可实现平均相对误差低于 1.0% 的高精度辨识;在存在低信噪比测量噪声(方差为 3.0% 的信号强度)及干扰力矩的复杂工况下,仍能保持稳定的辨识精度,平均相对误差控制在 1.5% 以内,说明所构建参数辨识模型具有较强的准确性和鲁棒性。

3) 本文方法完全基于数据驱动,不需要预先已知系统精确的数学模型,只依赖航天器的测量设备和执行机构即可完成对其时变惯性张量的辨识任务,工程实施性较强。

4) 仿真结果初步证明了所提方法对辨识挠性太阳帆时变惯性张量的可行性和抵抗外部干扰的鲁棒性,为未来的工程应用提供了理论基础。但是本文尚未考虑模型简化而造成的动力学模型的不确定性对辨识的影响,后续需要对此类问题开展分析研究。此外在未来的工作中还将对实际工程问题开展相关的地面实验测试,利用实验数据训练网络模型,以验证本文方法在工程中的有效性。

参 考 文 献

[1] 吴利平, 郭毓, 吴益飞, 等. 采用帆板-滑块执行结构的太阳帆姿态控制[J]. 哈尔滨工业大学学报, 2018, 50(9): 141
WU Liping, GUO Yu, WU Yifei, et al. Attitude control for solar-

sail using rotating panel-sliding mass actuator[J]. Journal of Harbin Institute of Technology, 2018, 50(9): 141. DOI: 10.11918/j.issn.0367-6234.201707072

- [2] ZHAO Pengyuan, WU Chenchen, LI Yangmin. Design and application of solar sailing: a review on key technologies [J]. Chinese Journal of Aeronautics, 2023, 36(5): 125. DOI: 10.1016/j.cja.2022.11.002
- [3] BERTHET M, SCHALKWYK J, ÇELİK O, et al. Space sails for achieving major space exploration goals: historical review and future outlook[J]. Progress in Aerospace Sciences, 2024, 150: 101047. DOI: 10.1016/j.paerosci.2024.101047
- [4] SILVESTRINI S, LAVAGNA M. Deep learning and artificial neural networks for spacecraft dynamics, navigation and control [J]. Drones, 2022, 6(10): 270. DOI: 10.3390/drones6100270
- [5] 初未萌. 在轨捕获过程空间组合体惯量矩阵智能辨识[D]. 大连: 大连理工大学, 2021
CHU Weimeng. Intelligent identification of inertia matrix of combined spacecraft in on-orbit capturing [D]. Dalian: Dalian University of Technology, 2021. DOI: 10.26991/d.cnki.gdllu.2021.004895
- [6] 曹凯, 朱勇, 高强, 等. 深度强化学习在自动控制领域研究现状与展望[J]. 排灌机械工程学报, 2023, 41(6): 638
CAO Kai, ZHU Yong, GAO Qiang, et al. Research status and prospect of deep reinforcement learning in automatic control [J]. Journal of Drainage and Irrigation Machinery Engineering, 2023, 41(6): 638. DOI: 10.3969/j.issn.1674-8530.22.0108
- [7] 孙俊, 张世杰, 马也, 等. 空间非合作目标惯性参数的 Adaline 网络辨识方法[J]. 航空学报, 2016, 37(9): 2799
SUN Jun, ZHANG Shijie, MA Ye, et al. Adaline network-based identification method of inertial parameters for space uncooperative targets[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(9): 2799. DOI: 10.7527/S1000-6893.2015.0349
- [8] 金晨迪, 康国华, 郭玉洁, 等. 基于深度学习的航天器组合体惯性参数在轨智能辨识[J]. 中国空间科学技术, 2019, 39(2): 1
JIN Chendi, KANG Guohua, GUO Yujie, et al. On-orbit intelligent identification of combined spacecraft's inertia parameter based on deep learning[J]. Chinese Space Science and Technology, 2019, 39(2): 1. DOI: 10.16708/j.cnki.1000-758X.2019.0003
- [9] 赵光玖. 基于深度学习的航天器控制系统参数辨识[D]. 长沙: 国防科技大学, 2021
ZHAO Guangjiu. Parameter identification of spacecraft control system based on deep learning [D]. Changsha: National University of Defense Technology, 2021. DOI: 10.27052/d.cnki.gzjgu.2021.001333
- [10] CHU Weimeng, WU Shunan, HE Xiao, et al. Deep learning-based inertia tensor identification of the combined spacecraft [J]. Proceedings of the Institution of Mechanical Engineers, Part G;

Journal of Aerospace Engineering, 2020, 234 (7): 1356. DOI: 10.1177/0954410020904555

[11] WU Shunan, CHU Weimeng, WU Zhigang, et al. Inertia matrix identification of combined spacecraft using a deep neural network with optimized network structure[J]. Advances in Space Research, 2024, 73 (3): 1979. DOI: 10.1016/j.asr.2023.11.024

[12] 初未萌, 杨今朝, 邬树楠, 等. 基于 LSTM 的空间机器人系统惯性张量在轨辨识[J]. 航空学报, 2021, 42(11): 524615
CHU Weimeng, YANG Jinzhao, WU Shunan, et al. LSTM-based on-orbit identification of inertia tensor for space robot system[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(11): 524615. DOI: 10.7527/S1000-6893.2020.24615

[13] MENG Bin, ZHAO Yunbo. The dynamics characteristics of flexible spacecraft and its closed-loop stability with passive control [J]. Journal of Systems Science and Complexity, 2021, 34 (3): 860. DOI: 10.1007/s11424-020-9268-8

[14] 刘金鹏. 基于 LSTM 的小型共轴双旋翼无人机不确定性建模与控制[D]. 长沙: 国防科技大学, 2022
LIU Jinpeng. LSTM based modeling and control of an uncertain coaxial dual-rotor UAV [D]. Changsha: National University of Defense Technology, 2022. DOI: 10.27052/d.cnki.gzjgu.2022.000175

[15] 薛佳帆, 张航维, 何广华, 等. 基于深度学习的海工平台运动预测[J]. 哈尔滨工业大学学报, 2024, 56(8): 163
XUE Jiafan, ZHANG Hangwei, HE Guanghua, et al. Motion prediction of offshore platforms based on deep learning[J]. Journal of Harbin Institute of Technology, 2024, 56(8): 163. DOI: 10.11918/202307085

[16] 张氢, 江伟哲, 李恒. MCKD-Teager 能量算子结合 LSTM 的滚动轴承故障诊断[J]. 哈尔滨工业大学学报, 2021, 53(7): 68
ZHANG Qing, JIANG Weizhe, LI Heng. Combined MCKD-Teager energy operator with LSTM for rolling bearing fault diagnosis[J]. Journal of Harbin Institute of Technology, 2021, 53(7): 68. DOI: 10.11918/202006114

[17] TANG Xueming, WU Nan, PAN Ying. Prediction of particulate matter 2.5 concentration using a deep learning model with time-frequency domain information [J]. Applied Sciences, 2023, 13(23): 12794. DOI: 10.3390/app132312794

[18] ZHANG Chu, MA Huixin, HUA Lei, et al. An evolutionary deep learning model based on TVFEMD, improved sine cosine algorithm, CNN and BiLSTM for wind speed prediction [J]. Energy, 2022, 254: 124250. DOI: 10.1016/j.energy.2022.124250

[19] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Delving deep into rectifiers: surpassing human-level performance on imagenet classification [C]// 2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2016: 1026. DOI: 10.1109/ICCV.2015.123

附录:

1) 时间段 0 ~ 25 s, 刚柔耦合系数矩阵为

$$\begin{bmatrix} -2.100 \times 10^{-3} & 1.100 \times 10^{-2} & 1.150 \times 10^{-2} & -3.065 \times 10^{-4} & -0.107 & -0.103 \\ 3.065 \times 10^{-4} & 0.107 & 0.103 & -2.100 \times 10^{-3} & 1.100 \times 10^{-2} & 1.150 \times 10^{-2} \\ 0.151 & -3.834 \times 10^{-4} & -5.391 \times 10^{-5} & 0.151 & -3.834 \times 10^{-4} & -5.391 \times 10^{-5} \\ 2.100 \times 10^{-3} & -1.100 \times 10^{-2} & -1.150 \times 10^{-2} & 3.065 \times 10^{-4} & 0.107 & 0.103 \\ -3.065 \times 10^{-4} & -0.107 & -0.103 & 2.100 \times 10^{-3} & -1.100 \times 10^{-2} & -1.150 \times 10^{-2} \\ 0.151 & -3.834 \times 10^{-4} & -5.391 \times 10^{-5} & 0.151 & -3.834 \times 10^{-4} & -5.391 \times 10^{-5} \end{bmatrix}$$

2) 时间段 25 ~ 50 s, 刚柔耦合系数矩阵为

$$\begin{bmatrix} -3.000 \times 10^{-3} & 2.080 \times 10^{-2} & -1.020 \times 10^{-2} & -2.940 \times 10^{-4} & -0.368 & 0.167 \\ 2.940 \times 10^{-4} & 0.368 & -0.167 & -3.000 \times 10^{-3} & 2.080 \times 10^{-2} & -1.020 \times 10^{-2} \\ 0.405 & -2.867 \times 10^{-4} & 8.870 \times 10^{-5} & 0.405 & -2.867 \times 10^{-4} & 8.870 \times 10^{-5} \\ 3.000 \times 10^{-3} & -2.080 \times 10^{-2} & 1.020 \times 10^{-2} & -2.940 \times 10^{-4} & 0.368 & -0.167 \\ -2.940 \times 10^{-4} & -0.368 & 0.167 & 3.000 \times 10^{-3} & -2.080 \times 10^{-2} & 1.020 \times 10^{-2} \\ 0.405 & -2.867 \times 10^{-4} & 8.870 \times 10^{-5} & 0.405 & -2.867 \times 10^{-4} & 8.870 \times 10^{-5} \end{bmatrix}$$

3) 时间段 50 ~ 75 s, 刚柔耦合系数矩阵为

$$\begin{bmatrix} 3.700 \times 10^{-3} & 2.790 \times 10^{-2} & 5.900 \times 10^{-3} & 4.621 \times 10^{-4} & -0.717 & -0.134 \\ -4.621 \times 10^{-4} & 0.717 & 0.134 & 3.700 \times 10^{-3} & 2.790 \times 10^{-2} & 5.900 \times 10^{-3} \\ -0.731 & -4.544 \times 10^{-4} & -1.161 \times 10^{-4} & -0.731 & -4.544 \times 10^{-4} & -1.161 \times 10^{-4} \\ -3.700 \times 10^{-3} & -2.790 \times 10^{-2} & -5.900 \times 10^{-3} & -4.621 \times 10^{-4} & 0.717 & 0.134 \\ 4.621 \times 10^{-4} & -0.717 & -0.134 & -3.700 \times 10^{-3} & -2.790 \times 10^{-2} & -5.900 \times 10^{-3} \\ -0.731 & -4.544 \times 10^{-4} & -1.161 \times 10^{-4} & -0.731 & -4.544 \times 10^{-4} & -1.161 \times 10^{-4} \end{bmatrix}$$

4) 时间段 75 ~ 100 s, 刚柔耦合系数矩阵为

$$\begin{bmatrix} 4.200 \times 10^{-3} & -3.280 \times 10^{-2} & 3.100 \times 10^{-3} & 6.724 \times 10^{-4} & 1.110 & -9.220 \times 10^{-2} \\ -6.724 \times 10^{-4} & -1.110 & 9.220 \times 10^{-2} & 4.200 \times 10^{-3} & -3.280 \times 10^{-2} & 3.100 \times 10^{-3} \\ -1.115 & 6.659 \times 10^{-4} & -3.151 \times 10^{-4} & -1.115 & 6.659 \times 10^{-4} & -3.151 \times 10^{-4} \\ -4.200 \times 10^{-3} & 3.280 \times 10^{-2} & -3.100 \times 10^{-3} & -6.724 \times 10^{-4} & -1.110 & 9.220 \times 10^{-2} \\ 6.724 \times 10^{-4} & 1.110 & -9.220 \times 10^{-2} & -4.200 \times 10^{-3} & 3.280 \times 10^{-2} & -3.100 \times 10^{-3} \\ -1.115 & 6.659 \times 10^{-4} & -3.151 \times 10^{-4} & -1.115 & 6.659 \times 10^{-4} & -3.151 \times 10^{-4} \end{bmatrix}$$