SCIENCE CHINA
Information Sciences

• LETTER •

January 2023, Vol. 66 119203:1–119203:2
https://doi.org/10.1007/s11432-020-3157-5

A novel frequency-dependent hysteresis model based on improved neural Turing machine

Yinan WU[1,2], Yongchun FANG[1,2*], Zhi FAN[1,2] & Cunhuan LIU[1,2]

[1]*Institute of Robotics and Automatic Information System, College of Artificial Intelligence, Nankai University, Tianjin* 300350*, China;*
[2]*Tianjin Key Laboratory of Intelligent Robotics, Nankai University, Tianjin* 300350*, China*

Received 16 May 2020/Revised 1 December 2020/Accepted 6 January 2021/Published online 1 November 2022

Citation Wu Y N, Fang Y C, Fan Z, et al. A novel frequency-dependent hysteresis model based on improved neural Turing machine. Sci China Inf Sci, 2023, 66(1): 119203, https://doi.org/10.1007/s11432-020-3157-5

Dear editor,

The inherent hysteresis of a piezoelectric actuator (PEA) results in intricate nonlinearity between the output displacement and input voltage, which restricts positioning accuracy of the actuator [1, 2]. Hysteresis behavior appears as a coupling of nonlinearity, frequency-dependence and memory characteristic, which makes it difficult to comprehensively characterize hysteresis [3]. To eliminate the effect of hysteresis on the positioning accuracy of PEAs, researchers have proposed various modeling methods to simulate hysteresis [4–6].

To accurately simulate the hysteresis effect while improving the model generalization ability, this study attempts to develop a novel model based on an improved neural Turing machine, which has the ability to simultaneously simulate nonlinearity, frequency-dependence and memory characteristic of hysteresis. The architecture of the proposed modeling method is shown in Figure 1(a). Although it is also feasible to set frequency as one of the nonlinear module inputs to simulate hysteresis, the advantages of the model with the frequency module are mainly reflected in the ability to reduce model training costs and improve the model applicability. More specifically, since the relationship between frequency and a hysteresis area is also nonlinear, it is equivalent to extracting frequency-dependent nonlinearity from hysteresis by designing a separate frequency module, thereby improving the interpretability of the constructed model. More importantly, this design can decrease the nonlinearity of the function fitted by the nonlinear module, so as to reduce the complexity of the network structure in the nonlinear module. Meanwhile, this structure enables distributed model training, which shortens training time while ensuring modeling accuracy.

Modeling and identification of frequency module. As is well known that hysteresis becomes more pronounced as the frequency of input voltage increases within a certain range, and the hysteresis area is positively correlated with frequency due to the frequency-dependence. Therefore, it is applicable to precisely characterize the frequency-dependence by elaborately constructing a back propagation neural network (BPNN) [7]. The structure of the designed BPNN is presented in Figure 1(b), where the input node corresponds to the frequency of the voltage, the output node produces the hysteresis area at the frequency of f, and the hidden layer consists of M nodes denoted as h_m, $m \in \{1, 2, \ldots, M\}$.

Therefore, a hysteresis area is calculated as follows:

$$\hat{a} = \tanh\left(\sum_{m=1}^{M}(\theta_m h_m) + \lambda\right), \tag{1}$$

where θ_m and λ respectively denote the weight and bias between the hidden layer and the output layer, which are further determined by training with the error back propagation algorithm. The hidden node h_m is calculated as follows:

$$h_m = \tanh(\hat{\theta}_m f + \hat{\lambda}_m), \tag{2}$$

where $\hat{\theta}_m$ and $\hat{\lambda}_m$ denote the weight and bias between the input layer and the hidden layer, respectively. $\tanh(\cdot)$ is an activation function with the following form:

$$\tanh(s) = \frac{\mathrm{e}^s - \mathrm{e}^{-s}}{\mathrm{e}^s + \mathrm{e}^{-s}}, \ \tanh(s) \in (-1, 1). \tag{3}$$

The training process of the weight and bias in the BPNN is shown in Appendix A.

Modeling and identification of nonlinear module. The structure of the nonlinear module is shown in Figure 1(c), which is a series-wound structure of an extreme learning machine (ELM) and a BPNN with a single hidden layer. The performance comparison between the proposed structure and a solitary ELM/BPNN structure is presented in Appendix B.

To be specific, for the constructed ELM, the input layer contains three nodes, which are set as follows:

$$\hat{i}_1 = \hat{a}, \ \hat{i}_2 = v_t, \ \hat{i}_3 = r_t, \tag{4}$$

where $\hat{a}$ denotes the hysteresis area produced by the frequency module at the frequency of f, v_t is the input voltage

* Corresponding author (email: fangyc@nankai.edu.cn)

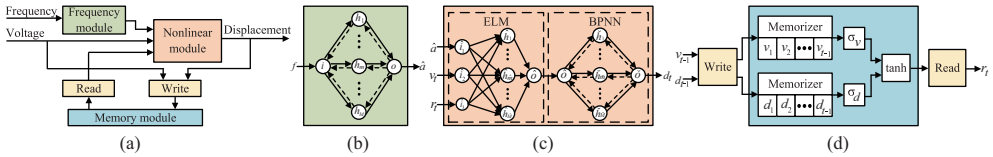

Figure 1 (Color online) The structures of the proposed modeling method and the different modules. (a) The architecture of the proposed model; (b) the frequency module; (c) the nonlinear module; (d) the memory module.

at the moment of t, and r_t represents the data read from the memory module at the moment of t. The calculation process of r_t is presented in detail in the next subsection.

Afterwards, the nodes in the hidden layer are calculated as follows:

$$\hat{h}_{\hat{m}} = \sin\left(\sum_{p=1}^{3}(\delta_{p\hat{m}}\hat{i}_p) + \gamma_{\hat{m}}\right),\ \hat{m}\in\{1,2,\ldots,\hat{M}\}, \quad (5)$$

where $\delta_{p\hat{m}}$ and $\gamma_{\hat{m}}$ represent the weight and bias between the input layer and the hidden layer, which are given randomly and do not need to be updated iteratively, thus effectively reducing the training time.

Hence, the output of the ELM is calculated as follows:

$$\hat{o} = \sum_{\hat{m}=1}^{\hat{M}}(\hat{\delta}_{\hat{m}}\hat{h}_{\hat{m}}), \quad (6)$$

where $\hat{\delta}_{\hat{m}}$ denotes the weights between the hidden layer and the output layer, which needs to be determined by model training.

Further, the output of the ELM is put into the BPNN to produce the final output of the nonlinear module as follows:

$$d_t = \check{o} = \tanh\left(\sum_{\check{m}=1}^{\check{M}}(\check{\varphi}_{\check{m}}\check{h}_{\check{m}}) + \check{\omega}\right), \quad (7)$$

where $\check{\varphi}_{\check{m}}$ and $\check{\omega}$ denote the weight and bias between the hidden layer and the output layer. $\check{h}_{\check{m}}$ is calculated by

$$\check{h}_{\check{m}} = \tanh(\varphi_{\check{m}}\hat{o} + \omega_{\check{m}}), \quad (8)$$

where $\varphi_{\check{m}}$ and $\omega_{\check{m}}$ stand for the weight and bias between the input layer and the hidden layer.

The parameter identification for the nonlinear module is shown in Appendix A.

Mechanism design of memory module. The detailed structure of the memory module is presented in Figure 1(d). To be specific, at the moment of t, the input voltage and the model output displacement at previous moments, written into the memory module, are stored by two memorizers, thus constructing a voltage sequence container $\{v_1, v_2, v_3, \ldots, v_{t-1}\}$ and a displacement sequence container $\{d_1, d_2, d_3, \ldots, d_{t-1}\}$.

Since the voltage closer to the current moment has a stronger correlation with the output displacement at the current moment, the weights applied to the voltage at different moments are set to gradually increase as approaching the current moment. Hence, the weighted average of the voltage sequence container is calculated as follows:

$$\bar{v}_t = \sigma_v(v_1, v_2, \ldots, v_{t-1}) = \sum_{j=1}^{t-1}\zeta_j v_j, \quad (9)$$

where ζ_j increases with the time j, which is determined by

$$\zeta_j = \frac{j}{\sum_{j=1}^{t-1} j} = \frac{2j}{t(t-1)}. \quad (10)$$

Similarly, the displacement sequence container is weighted averaged as follows:

$$\bar{d}_t = \sigma_d(d_1, d_2, \ldots, d_{t-1}) = \sum_{j=1}^{t-1}\tau_j d_j, \quad (11)$$

where τ_j is calculated by

$$\tau_j = \frac{\sum_{j=1}^{t-1} v_j}{\sum_{j=1}^{t-1} d_j}\cdot\frac{j}{\sum_{j=1}^{t-1} j} = \frac{2j\sum_{j=1}^{t-1} v_j}{t(t-1)\sum_{j=1}^{t-1} d_j}. \quad (12)$$

Therefore, the output of the memory module is calculated as follows:

$$r_t = \tanh(\varpi_v\bar{v}_t + \varpi_d\bar{d}_t + \xi), \quad (13)$$

where ϖ_v and ϖ_d are the given weights and ξ denotes a given threshold.

Experiment and analysis. Experimental results are shown in Appendix C.

Conclusion. This study proposes a neural Turing machine-based hysteresis modeling method to simulate hysteresis from the aspects of frequency-dependence, nonlinearity and memory characteristic. Experiments demonstrate the high accuracy and strong generalization ability of the proposed model. Future work will focus on the application of the proposed model for hysteresis compensation.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61633012, 62003172, 21933006).

Supporting information Appendixes A–C. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Li Z W, Song Q K, Cheng L, et al. Snoring detection based on a stretchable strain sensor. Sci China Inf Sci, 2021, 64: 174201
2 Wang J R, Zou Q Z. Rapid probe engagement and withdrawal with force minimization in atomic force microscopy: a learning-based online-searching approach. IEEE/ASME Trans Mechatron, 2020, 25: 581–593
3 Li M, Xi N, Wang Y C, et al. Atomic force microscopy in probing tumor physics for nanomedicine. IEEE Trans Nanotechnol, 2019, 18: 83–113
4 Tao Y D, Li H X, Zhu L M. Rate-dependent hysteresis modeling and compensation of piezoelectric actuators using Gaussian process. Sens Actuat A-Phys, 2019, 295: 357–365
5 Ren Y, Chen M, Liu J Y. Bilateral coordinate boundary adaptive control for a helicopter lifting system with backlash-like hysteresis. Sci China Inf Sci, 2020, 63: 119203
6 Zhou Z P, Tan Y H, Liu X F. State estimation of dynamic systems with sandwich structure and hysteresis. Mech Syst Signal Process, 2019, 126: 82–97
7 An Y, Meng H, Gao Y, et al. Application of machine learning method in optical molecular imaging: a review. Sci China Inf Sci, 2020, 63: 111101