Neural Network Based Modeling of Hysteresis in Smart Material Based Sensors

Yonghong Tan[1], Ruili Dong[2(✉)], and Hong He[1]

[1] College of Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, China
tany@shnu.edu.cn
[2] College of Information Science and Technology, Donghua University, Shanghai, China
ruilidong@dhu.edu.cn

Abstract. Hysteresis is a nonlinear phenomenon which is involved with dynamics, non-smoothness and multi-valued mapping. It usually exists in elastic materials, smart materials, and energy-storage materials. For describing the characteristic of hysteresis, a basis function based neural network model is proposed in this paper. In this method, the multi-valued mapping of hysteresis is transferred into a one-to-one mapping with an expanded input space involving the input variable and a constructed hysteretic auxiliary function. Thus, the neural network can be employed to approximate the characteristic of hysteresis. Finally, the method is used to the modeling of hysteresis in a smart material based sensor.

Keywords: Hysteresis · Expanded input space · Neural network · Modeling

1 Introduction

It is known that hysteresis usually exists in many smart materials such as piezoceramic, ionic polymer-metal composite (IPMC) and electromagnetic materials etc. [1–3]. Hysteresis is a nonlinear phenomenon with dynamics, non-smoothness and multi-valued mapping. The existence of hysteresis usually affects the performance of actuators or sensors made by these smart materials, e.g. dynamic performance and positioning accuracy etc. Accurate modeling of hysteresis in smart material based actuators or smart material based sensors is important for the design of a model based compensator to reduce the effect of hysteresis.

It has been found out that modeling of hysteresis is a challenge due to its features of multi-valued mapping and non-smoothness. Up till now, there have been some models proposed to describe the hysteresis phenomenon such as Preisach [5], Prandtl-Ishlinskii (PI) [6] and Bouc-Wen models [4]. However, both Preisach model and PI model are having the structure with a sum of weighted hysteretic operators and used for description of rate-independent hysteresis. Moreover, they usually employ lots of hysteretic operators even more than hundred for modeling of hysteresis. Additionally, Bouc-Wen model is usually used to describe the hysteresis between restoring force and displacement. It can also be used to describe the rate-dependent hysteretic behavior.

H. Lu et al. (Eds.): ISNN 2019, LNCS 11555, pp. 162–172, 2019.
https://doi.org/10.1007/978-3-030-22808-8_17

However, the identification of a Bouc-Wen model needs to use a method of nonlinear optimization [4] which may encounter the obstacle that the optimization may trap into local minima. Moreover, the parameters needed to be estimated in the Bouc-Wen model are involved in a group of non-smooth components, the gradients of the hysteresis output with respect to the estimated parameters may not exist at non-smooth points of the system. Furthermore, it is noted that the conventional identification methods cannot be utilized to model hysteresis behavior, directly, due to its characteristic of multi-valued mapping.

For modeling of hysteresis, neural networks are potential measures due to their well-known capability of universal approximation. However, the neural networks are applicable only for modeling the systems with one-to-one mapping between the input and output [7, 8]. Those neural networks will be unable to model the hysteresis which has the features of non-smoothness and multi-valued mapping between the input and output. Hence, using neural networks for modeling hysteresis directly becomes a challenge.

For modeling the hysteresis with non-smoothness and multi-valued mapping, Refs. [7, 8] proposed the so-called expanded input space based neural network hysteresis models. In these models, the sigmoidal functions are used as the active functions of neurons. For the simplification of modeling procedure, in this paper, a simple basis function based neural network is proposed for modeling of hysteresis. In this method, an expanded input space is built to transform the multi-valued mapping between input and output of hysteresis to a one-to-one mapping. Then, a basis function based neural network is applied to modeling of hysteresis on the expanded input space. The proposed basis function has a simple structure and can be trained conveniently just by least-square-type algorithm.

For test the performance of the proposed modeling method, the experimental results of modeling for an IPMC sensor are presented.

2 Hysteresis in Smart Material Based Sensors

Ionic polymer-metal composite (IPMC) is a kind of electroactive polymer (EAP) material, which is also called as artificial muscle. IPMC can generate electric signal correlated with its mechanical deformation or displacement. Based on this property, it can be used as a sensor to measure displacement and deformation of flexible mechanism. It is known that IPMC is usually composed of three layers, namely, an ion-exchange polymer membrane sandwiched in between noble metal electrodes. The negatively charged anions, in the polymer membrane, covalently fixed to polymer chains are balanced by positively charged moving cations. When an externally mechanical force to make IPMC deformed, the moving cations will be redistributed. In this case, IPMC will produce a detectable electric signal (e.g. voltage or current) which is associated with the externally mechanical excitation [2, 9]. This phenomenon enables IPMC to be used as a sensor to measure the displacement, vibration, or deformation of a load. Conversely, based on its inverse electroactive feature, IPMC can also be used as an actuator when voltage is implemented on it.

Moreover, IPMC can be used as a sensor to measure mechanical displacement in humid or flexible case for IPMC has the property that generates electrical signal corresponding to its mechanical deformation.

Just like most smart materials, hysteresis is also inherent in IPMC materials. Figure 1 has shown the experimental result having shown that the hysteresis in IPMC sensor has dynamic drift and rate-dependent characteristic. As hysteresis is a non-smooth function with multi-valued mapping, the conventional neural networks may fail to model it properly due to the conventional neural networks can only be applied to modeling of smooth systems with one-to-one mapping [7, 8].

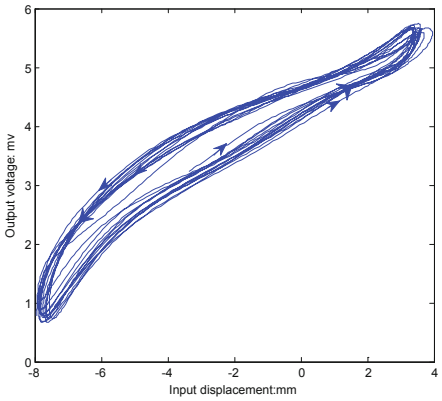

Fig. 1. Rate-dependent behavior of hysteresis with drift in IPMC sensor.

Therefore, in this paper, an important task is to transform the multi-valued mapping of hysteresis to a one-to-one mapping. In order to realize the task to transform the multi-valued mapping between the input and output of hysteresis to a single-valued mapping, an expanded input space will be constructed using the system input and a hysteresis auxiliary function [10]. The hysteretic auxiliary function introduced into the input space is served as an additional coordinate to construct an expanded input space. Actually, the function of the hysteretic auxiliary function is used as an "imagine" of hysteresis that can extract the movement tendency of hysteresis, such as ascending, descending and turning point etc.

3 Expanded Input Space

From the well known Preisach formula to describe hysteresis [10], we have

$$\begin{aligned} H[u](k) &= \iint_{S} \mu(\alpha,\beta)\gamma_{\alpha,\beta}[u](k)d\alpha d\beta \\ &= \iint_{S_{+1}} \mu(\alpha,\beta)d\alpha d\beta - \iint_{S_{-1}} \mu(\alpha,\beta)d\alpha d\beta \end{aligned} \tag{1}$$

where $H[u](k)$ is the output of hysteresis, $\mu(\alpha,\beta)$ is the weighting function and $\gamma_{\alpha,\beta}[u](k)$ is the hysteretic operator. Then, it can be decomposed as

$$\begin{aligned} H_+[u](k) &= H_-[u_p] + \Delta H_+[u - \mathrm{u_p}], \Delta u > 0 \\ H_-[u](k) &= H_+[u_p] - \Delta H_-[u_p - \mathrm{u}], \Delta u < 0 \end{aligned} \tag{2}$$

where u_p is the local extreme of the input, $H_+[u_p]$ and $H_-[u_p]$ represent the dominant maximum and minimum of the output of the hysteretic model, while $\Delta H_+[u - \mathrm{u}_p]$ and $\Delta H_-[u_p - u]$ are the incremental output of the model as $u(k)$ monotonically increases or decreases from the extreme. It is known that the hysteresis described in (1) has the following feature:

For two different time instants k_1 and k_2, if $u(k_1) = u(k_2)$, $H[u](k_1) \neq H[u](k_2)$ due to the different input extremes. Hence, suppose that an auxiliary function satisfies $f[u](k_1) \neq f[u](k_2)$ can be found. Then, we introduce it to the expanded input space to uniquely determine the output of hysteresis on $(u(k), f[u](k))$.

Definition 1: Define the hysteresis auxiliary function as

$$f(k) = f(u_{ex}) + |u(k) - u_{ex}| f(|u(k) - u_{ex}|), \tag{3}$$

where u_{ex} is the local extreme of the input, and $f(x) \geq 0$ is a smooth and monotonic function not dependent on time k, and $f(u_{ex})$ is the current local extreme of the auxiliary function.

Remark: Based on Definition 1, it is known the procedure to construct an auxiliary function, i.e. (i) selecting a piecewise function with the structure shown as in Eq. (3), which includes local extreme of input and the output of function corresponding to the local extreme of input. (ii) each segment of piecewise function is a smooth and monotonic function not dependent on time k. (iii) the switch of the piecewise function is triggered at the extreme of input.

Then, we have

Definition 2: Define

$$U = \{u(k)\} \tag{4}$$

as the input set. Then, the expanded input space is defined as

$$E = \{u(k), f(k)\} \tag{5}$$

Afterward, we have

Theorem: Suppose $u(k) \in U$ and $f(k)$ is the hysteresis auxiliary function defined by (3) and

$$\Phi = \{H(k)\} \tag{6}$$

is the output set of hysteresis. If there exists a continuous mapping Γ, such that

$$E| \xrightarrow{\Gamma} \Phi \tag{7}$$

Then, Γ is a one-to-one mapping.

Proof: Please refer to Appendix.

4 Basis Function Based Neural Network

On the proposed expanded input space, the multi-valued mapping of hysteresis can be transformed to a one-to-one mapping.

Subsequently, a basis function based neural network (BFBNN) on the constructed expanded input space is used to describe the hysteresis characteristic of IPMC sensor. The advantage of using BFBNN is that it has a simpler model structure determined based on the parsimony principle, which does not require a large number of neural basis functions, hysteretic operators, or not rely on empirical skills by comparing with the modeling methods provided by Refs. [10, 11] and [12]. Thus, the corresponding model on expanded input space is defined by

$$y(k) = g[\mathbf{u}(k-1), \mathbf{f}(k-1), \mathbf{y}(k-1)] \tag{8}$$

where $g(\cdot)$ is the mapping between the input space and the output of system, $\mathbf{u}(k-1) = [u(k-1), \cdots, u(k-n_u)]^T$ and $\mathbf{y}(k-1) = [y(k-1), \cdots, y(k-n_y)]^T$ are the input vector and output vector, respectively; $\mathbf{f}(k-1) = [f(k-1), \cdots, f(k-n_f)]^T$ is the output vector of hysteretic auxiliary function. n_u, n_f and n_y are the lags for sequences $\{u\}$, $\{f\}$ and $\{y\}$, respectively. Then, (8) can also be described by:

$$y(k) = \sum_{i=1}^{p} \theta_i \, z_i \, (k), \tag{9}$$

where θ_i is the i*th* weighting factor, $z_i(k)$ is the i*th* basis function of the neural network while p is the number of basis function. Figure 2 illustrates the corresponding structure of the basis function based neural network.

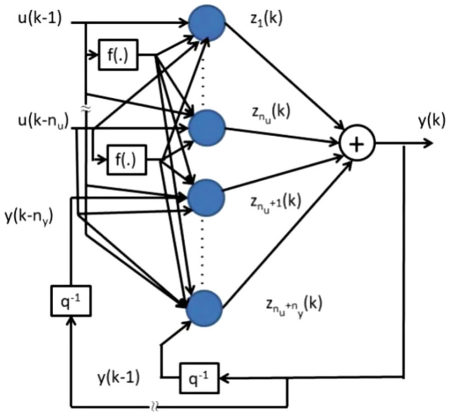

Fig. 2. Structure of basis function based neural network.

The candidates of basis function can be $sin(u)$, $cos(u)$; $1/(1 + exp(-x))$; $exp(-(x - m)^2/\rho^2)$; or $u^i y^j$. The selection of the proper basis function depends upon the specific requirement of application. In this paper, we use $u^i y^j$ as the basis function just for the simple structure of the neural model. Thus, Eq. (8) can be represented by

$$\begin{aligned} y(k) = &\sum_{i_1=1}^{p} \hat{\theta}_{i_1} z_{i_1}(k) + \sum_{i_1=1}^{p} \sum_{i_2=i_1}^{p} \hat{\theta}_{i_1 i_2} z_{i_1}(k) z_{i_2}(k) + \cdots \\ &+ \sum_{i_1=1}^{p} \cdots \sum_{i_q=i_{q-1}}^{p} \hat{\theta}_{i_1 \cdots i_q} z_{i_1}(k) \cdots z_{i_q}(k) + \hat{\theta}_{n_1+1}, \end{aligned} \tag{10}$$

where $\hat{\theta}_{n_1+1}$ is the estimated bias, $\widehat{\boldsymbol{\theta}} = [\hat{\theta}_1, \cdots, \hat{\theta}_{n_1}, \hat{\theta}_{n_1+1}]^T$ is the estimated coefficient vector, $n_1 = (p+q)!/(p!q!) - 1$, and q is the order of model. Define the variable vector as

$$\begin{aligned} \mathbf{h}(k) = &[z_{i_1}(k), \cdots, z_p(k), z_1(k) \cdot z_1(k), \cdots, z_p(k) \cdot z_p(k), \\ &\cdots, \underbrace{z_1(k) \cdots z_1(k)}_{q}, \ldots, \underbrace{z_p(k) \cdots z_p(k)}_{q}, 1]^T. \end{aligned} \tag{11}$$

Then, the BFBNN model can be denoted as:

$$y(k) = \mathbf{h}^T(k) \widehat{\boldsymbol{\theta}}(k-1) + e(k), \tag{12}$$

where $e(k)$ is the modeling residual. Hence, the problem to model the dynamic drift and hysteresis becomes the training of the BFBNN model.

5 Training of Basis Function Based Neural Network

As the expanded input space based BFBNN has a very simple structure shown as in Eq. (12), the least square algorithm can be used for network training. Then, the corresponding recursive training algorithm is shown as follows:

$$e(k) = y(k) - \mathbf{h}^T(k)\widehat{\boldsymbol{\theta}}(k-1), \tag{13}$$

$$\widehat{\boldsymbol{\theta}}(k) = \widehat{\boldsymbol{\theta}}(k-1) + \mathbf{K}(k)e(k), \tag{14}$$

$$S_e(k) = \lambda(k)\mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k) + \mu(k)\hat{\sum}(k), \tag{15}$$

$$\mathbf{K}(k) = \lambda(k)\mathbf{P}(k-1)\mathbf{h}(k)S_e^{-1}(k), \tag{16}$$

$$\mathbf{P}(k) = (1/\mu(k))[\mathbf{P}(k-1) - \mathbf{K}(k)\mathbf{h}^T(k)\mathbf{P}(k-1)], \tag{17}$$

and

$$\hat{\sum}(k) = \hat{\sum}(k-1) + \rho(k)[\hat{\sigma}^2(k) - \hat{\sum}(k-1)], \tag{18}$$

where $\rho(k) \leq 1$ is the convergence factor, $\mu(k) = \frac{\rho(k-1)}{\rho(k)}[1-\rho(k)]$ is the forgetting factor, $\hat{\sigma}^2(k) = \hat{\sigma}^2(k-1) + e^2(k)/k$, $\mathbf{P}(0) = \eta\mathbf{I}$, $\mathbf{I}$ is an identity matrix, $0 < \eta < \infty$, and $\lambda(k)$ is a switch coefficient and defined as

$$\lambda(k) = \begin{cases} 0, & |\varepsilon(k)| \leq |r| \\ \frac{\beta(\hat{\sigma}^2(k) - r^2)}{\mathbf{h}^T(k)P(k-1)\mathbf{h}(k)r^2} & otherwise \end{cases}, \tag{19}$$

where $0 < \beta \leq \mu(k)\hat{\sum}(k)$.

6 Experimental Results

In this experiment, the sampling frequency is chosen as 1000 Hz. Also, the proposed BFBNN model on expanded input space is applied to the modeling of hysteresis behavior of IPMC sensor. To determine the architecture of the model, the criterion shown in the following is defined to evaluate what structure can properly describe the behavior of the IPMC sensor.

$$C(n) = \frac{1}{\mathrm{n}}\sum_{k=1}^{n}(y(k) - \hat{y}(k))^2. \tag{20}$$

Then, $IC(n)$ and $TC(n)$ are used to denote the values of $C(n)$ in training and model validation procedures, respectively. Afterwards, the proposed training algorithm is employed to estimate the coefficients of the model with the corresponding model

architecture. Table 1 presents the corresponding values $IC(n)$ and $TC(n)$ of different model structures, so the structure of the proposed model can be selected as $n_y = 0$, $n_u = 9$, $n_f = 9$, when $q = 2$, and in this case, *TC(n),* the criterion of model validation reaches the smallest value. For comparison, a classical nonlinear auto-regressive and moving-average with exogenous input (NARMAX) is also used to describe the characteristics of the IPMC sensor. The structure parameters of the NARMAX model are chosen as $n_y = 0$, $q = 2$ and $n_u = 9$, respectively.

Table 1. Determination of model structure of IPMC sensor.

$n_y = 0$	$n_u = 7$ $n_f = 7$	$n_u = 8$ $n_f = 8$	$n_u = 9$ $n_f = 9$	$n_u = 10$ $n_f = 10$
$IC(n)$	1.41×10^{-4}	1.01×10^{-5}	1.58×10^{-6}	1.23×10^{-7}
$TC(n)$	1.45×10^{-4}	1.20×10^{-5}	4.68×10^{-6}	1.29×10^{-5}

Figure 3 shows the comparison of model validation between the proposed model and the BFBNN model not on the expanded input space. To show the detail of the model validation results, the data from 0 s to 0.098 s are removed to avoid illustrating the influence of initial values. In Fig. 3(a), the dotted line denotes the output of the proposed model while the dot and dash line represents the output of classical NARMAX model. In Fig. 3(b), the solid line and dotted line denote modeling errors of the proposed model and classical NARMAX model, respectively. Note that the modeling error of the classical NARMAX model illustrates larger fluctuation and the maximum absolute model error of the classical NARMAX model is 0.14 mV, while that of the proposed modeling method is 0.014 mV, respectively. Moreover, Fig. 4 shows the comparison of output versus input curves between the proposed modeling strategy and the classical NARMAX modeling result. Obviously, the proposed model is more suitable to describe the performance of IPMC sensor.

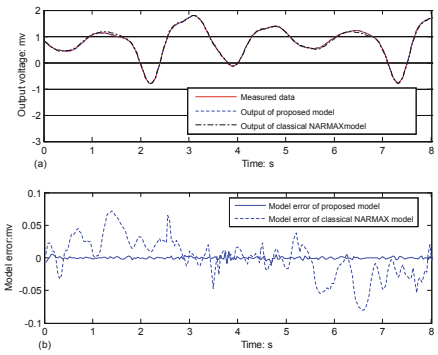

Fig. 3. Comparison of model validation results between the proposed model and the classical NARMAX model.

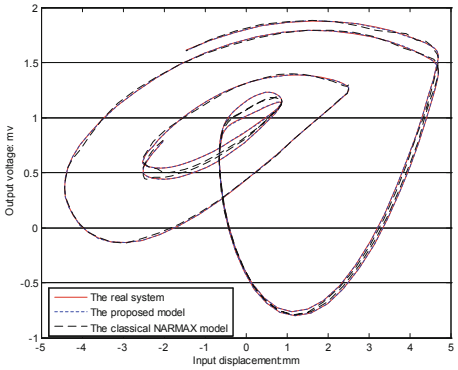

Fig. 4. Comparison of the input-output curves between the proposed model and the classical NARMAX model (the proposed model: dotted line; the classical NARMAX model: dashed line).

7 Conclusions

In this paper, a BFBNN model on the expanded input space to describe the hysteretic and dynamic drifting behavior of IPMC sensor is developed. A method for constructing a proper hysteresis auxiliary function for expanded input space is also presented. Then, a training algorithm is presented to train the BFBNN model of IPMC sensor.

In the modeling experiment, it has shown that the proposed method can achieve better modeling result of hysteresis and dynamic drift in IPMC sensor than a classical NARMAX model.

Acknowledgment. The work presented in this paper has been funded by the National Science Foundation of China under Grants 61671303 and 61571302, the Open Fund of the Key Laboratory of Nano-Devices and Applications, Chinese Academy of Sciences under Grant 18ZS06, the Shanghai Pujiang Program under Grant 18PJ1400100, the Natural Science Foundation of Shanghai under Grant 16ZR1446700, and the project of the Science and Technology Commission of Shanghai under Grant 18070503000.

Appendix

The proof of theorem is as follows:

For any $H(k) \in \Phi$, there, at least, exists a $u(k) \in U$. Thus, $E| \xrightarrow{\Gamma} \Phi$ is a surjective mapping. Figure 5 illustrates the case of surjective mapping of the expanded input space. Now, we prove $E| \xrightarrow{\Gamma} \Phi$ is also a one-to-one mapping.

Case 1: For two different time instants k_1 and k_2, assume $u(k_1) = u(k_2)$ are not extremes, but the output of hysteresis $H(k_1) \neq H(k_2)$ for the effect of the different extremes related to the output of hysteresis. In this case, we can choose a proper $f(k)$

satisfies $f(k_1) \neq f(k_2)$. Then, $E| \xrightarrow{\Gamma} \Phi$ is an injective mapping. To demonstrate the mapping of Case 1, Fig. 6 is presented.

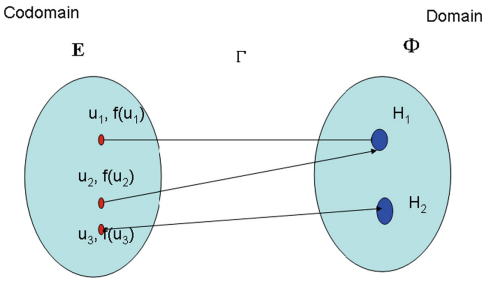

Fig. 5. The surjection of expanded input space.

Case 2: If $u(k_1) = u(k_2)$ and $H(k_1) = H(k_2)$ where $u(k_1)$ and $u(k_2)$ are extremes, based on the characteristic of hysteresis [13], it leads to $H(k_1) = H(k_2)$. In this case, it also results in $f(k_1) = f(k_2)$. In this situation, $E| \xrightarrow{\Gamma} \Phi$ is also an injective mapping. The corresponding description of Case 2 is illustrated in Fig. 7.

In terms of what we discussed, $E| \xrightarrow{\Gamma} \Phi$ is not only a surjective but also an injective mapping. Therefore, it is a bijection, i.e. Γ is a one-to-one mapping.

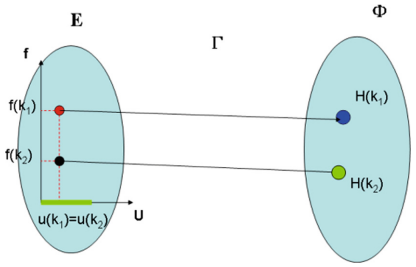

Fig. 6. The injective mapping of expanded input space (Case 1).

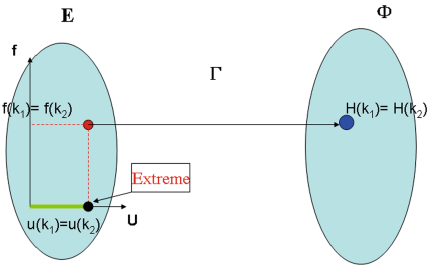

Fig. 7. The injective mapping of expanded input space (Case 2).

References

1. Hu, H.: Compensation of hysteresis in piezoceramic actuators and control of nanopositioning system. Ph.D. thesis of University of Toronto, Canada (2003)
2. Shahinpoor, M., Kim, K.: Ionic polymer-metal composites. Part I. Fundamentals. Smart Mater. Struct. **10**, 819–833 (2001)
3. Mayergoyz, D.: Dynamic Preisach models of hysteresis. IEEE Trans. Magnetics **24**(6), 2925–2927 (1988)
4. Awrejcewicz, J., Dzyubak, L., Lamarque, C.: Modelling of hysteresis using Masing–Bouc-Wen's framework and search of conditions for the chaotic responses. Commun. Nonlinear Sci. Numer. Simul. **13**, 939–958 (2008)
5. Yu, Y., Naganathan, N., Dukkipati, R.: Preisach modeling of hysteresis for piezoceramic actuator system. Mech. Mach. Theory **37**(1), 49–59 (2002)
6. Su, C., Wang, Q., Chen, X.: Adaptive variable structure control of a class of nonlinear systems with unknown Prandtl-Ishlinskii hysteresis. IEEE Trans. Autom. Control **50**(12), 2069–2074 (2005)
7. Dong, R., Tan, Y., Chen, H., Xie, Y.: A neural networks based model for rate-dependent hysteresis for piezoceramic actuators. Sens. Actuators A **143**(2), 370–376 (2008)
8. Deng, L., Tan, Y.: Diagonal recurrent neural network with modified backlash operators for modeling of rate-dependent hysteresis in piezoelectric actuators. Sens. Actuators A **148**(1), 259–270 (2008)
9. Chen, X., Zhu, G., Yang, X., Hung, D., Tan, X.: Model-based estimation of flow characteristics using an ionic polymer-metal composite beam. IEEE/ASME Trans. Mechatron. **18**(3), 932–943 (2013)
10. Zhao, X., Tan, Y.: Neural network based identification of Preisach-type hysteresis in piezoelectric actuator using hysteretic operator. Sens. Actuators A **126**, 306–311 (2006)
11. Nam, D., Ahn, K.: Identification of an ionic polymer metal composite actuator employing Preisach type fuzzy NARX model and Particle Swarm Optimization. Sens. Actuators A **183**, 105–114 (2012)
12. Li, Z., Hao, L.: The identification of discrete Preisach model based on IPMC. In: Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics, Bangkok, Thailand, vol. 21–26, pp. 751–755 (2009)
13. Mayergoyz, I.D.: Mathematical Models of Hysteresis. Springer, New York (1991). https://doi.org/10.1007/978-1-4612-3028-1