SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

March 2021, Vol. 64 132201:1–132201:18
https://doi.org/10.1007/s11432-019-2878-y

Rapid dynamical pattern recognition for sampling sequences

Weiming WU[1], Qian WANG[2], Chengzhi YUAN[3] & Cong WANG[2*]

[1]*School of Automation Science and Engineering, South China University of Technology, Guangzhou* 510641*, China;*
[2]*School of Control Science and Engineering, Shandong University, Jinan* 250061*, China;*
[3]*Department of Mechanical, Industrial and Systems Engineering, University of Rhode Island, Kingston* RI 02881*, USA*

Received 23 October 2019/Revised 17 January 2020/Accepted 29 February 2020/Published online 1 February 2021

Abstract In this paper, based on deterministic learning, we propose a method for rapid recognition of dynamical patterns consisting of sampling sequences. First, for the sequences yielded by sampling a periodic or recurrent trajectory (a dynamical pattern) generated from a nonlinear dynamical system, a sampled-data deterministic learning algorithm is employed for modeling/identification of inherent system dynamics. Second, a definition is formulated to characterize similarities between sampling sequences (dynamical patterns) based on differences in the system dynamics. Third, by constructing a set of discrete-time dynamical estimators based on the learned knowledge, similarities between the test and training patterns are measured by using the average L_1 norms of synchronization errors, and general conditions for accurate and rapid recognition of dynamical patterns are given in a sampled-data framework. Finally, numerical examples are discussed to illustrate the effectiveness of the proposed method. We demonstrate that not only a test pattern can be rapidly recognized corresponding to a similar training pattern, but also the proposed recognition conditions can be verified step by step based on historical sampling data. This makes a distinction compared with the previous work on rapid dynamical pattern recognition for continuous-time nonlinear systems, in which the recognition conditions are difficult to be verified by using continuous-time signals.

Keywords deterministic learning, dynamical pattern recognition, sampling sequence, synchronization

Citation Wu W M, Wang Q, Yuan C Z, et al. Rapid dynamical pattern recognition for sampling sequences. Sci China Inf Sci, 2021, 64(3): 132201, https://doi.org/10.1007/s11432-019-2878-y

1 Introduction

Temporal/dynamical pattern recognition is an important and challenging issue that has attracted great attention in various areas, such as aerospace, biology, biomedicine, and control engineering [1–6]. Considering that the information of temporal patterns is embedded in time, recognition of temporal patterns differs from that of static patterns. In research and practical applications, temporal patterns are often measured by using sensors and are represented in a form of sampling sequences (a category of temporal sequences). Therefore, it is of great importance to investigate the problem of modeling and recognition of temporal/dynamical patterns represented by sampling sequences (temporal sequences) [7, 8].

There are many methods proposed for modeling and recognition of temporal sequences, for example, hidden Markov models (HMMs) and recurrent neural networks (RNNs) [9–18]. HMMs can be used to model temporal sequences as a series of hidden states with a probabilistic dependency [9]. This approach has been applied to speech and video recognition tasks [10–12]; however, the applicability of HMMs is limited to relatively simple and stationary temporal sequences due to the time-invariant nature of these models [12]. In contrast, RNNs can be considered as an alternative powerful approach for processing temporal sequences [13–22]. They have been shown capable of emulating the evolution of temporal sequences by simulating discrete-time dynamical systems [14–16]. So far, RNNs and their variants have been successfully applied to various sequence recognition tasks such as speech recognition and natural language processing [16–22]. This has been possible not only owing to the application of advanced unit

* Corresponding author (email: wangcong@sdu.edu.cn)

architectures, such as long short-term memory (LSTM) [23], and gated recurrent unit (GRU) [17], to address the problems of vanishing and exploding gradient arising in early RNNs [24], but also to the construction of deep network architectures, such as stacked RNNs [18], deep transition RNNs [14], and convolutional RNNs [25], to enable hierarchical processing of temporal sequences. However, despite the growing use of RNNs, theoretical development of such networks is not fully investigated [26], especially concerning a theoretical problem of whether the true system dynamics within temporal sequences can be accurately modeled/identified.

Recently, a deterministic learning approach has been proposed for accurate identification and rapid recognition of dynamical patterns [27–30], in which they are defined as periodic or recurrent trajectories generated from continuous-time nonlinear dynamical systems and thus belong to a specific class of temporal patterns. Through deterministic learning, unknown system dynamics within dynamical patterns can be accurately modeled in local regions along periodic or recurrent trajectories [27]. A similarity definition for dynamical patterns was given based on differences in their system dynamics, and a mechanism for rapid recognition of a test dynamical pattern from a set of training dynamical patterns has been proposed [30]. Specifically, according to this method, training patterns are firstly trained through deterministic learning, and then, the learned knowledge is reused to construct a set of estimators to represent these training patterns. When a test dynamical pattern is presented to estimators, a set of state synchronization errors (recognition errors) can be generated. Rapid recognition of test dynamical patterns can be achieved according to the principle of the smallest synchronization error. The deterministic learning approach for identification and recognition of dynamical patterns has been applied to enable modeling and rapid detection of small oscillation faults yielded from several classes of nonlinear dynamical systems [31–36].

It should be noted that the aforementioned results on dynamical pattern recognition have been achieved based on the frameworks of continuous-time nonlinear systems. In this paper, based on deterministic learning, we further propose a method for rapid recognition of dynamical patterns consisting of sampling sequences. First, for the sampling sequences yielded by sampling a periodic or recurrent trajectory (a dynamical pattern) generated from a nonlinear dynamical system, a sampled-data deterministic learning algorithm [37, 38] can be employed to perform accurate modeling/identification of the inherent system dynamics. Second, a definition to characterize similarities between sampling sequences (dynamical patterns) is formulated based on differences in system dynamics. The proposed definition implies that the sampling period needs to be small enough. Third, by constructing a set of discrete-time dynamical estimators based on the learned knowledge for a set of training patterns, similarities between the test and training patterns can be evaluated through considering average L_1 norms of synchronization errors. Particularly, by analyzing recognition error systems, general conditions for accurate and rapid recognition of dynamical patterns in sampling sequences are proposed in a sampled-data framework.

In contrast with the previous studies on rapid dynamical pattern recognition for continuous-time nonlinear systems [30, 31], the present study makes a distinction in that the proposed recognition conditions can be verified step by step based on the historical sampling data in the sampled-data framework. However, concerning the continuous-time dynamical pattern recognition [30,31], general recognition conditions are difficult to be verified by using continuous-time signals. Even in the case when misrecognition occurs, the situation is difficult to analyze due to the lack of condition verification. Furthermore, compared with the previously obtained results on modeling/recognition with regard to temporal sequences using HMMs and RNNs [9–18], the proposed approach has the following features: (1) true inherent system dynamics in dynamical patterns can be modeled/identified accurately; (2) rapid recognition can be achieved based on analyzing differences in system dynamics of corresponding dynamical patterns.

The remainder of the paper is organized as follows. The research problem is formulated in Section 2. In Section 3, the modeling phase and the similarity of dynamical patterns consisting of sampling sequences are presented. The proposed rapid dynamical pattern recognition method is described in Section 4. The results of numerical simulations conducted on the nonlinear oscillation systems are discussed in Section 5. Finally, a conclusion is drawn in Section 6.

Notation. $\mathbb{R}$ and $\mathbb{N}$ denotes the set of real numbers and the set of natural numbers, respectively; for any $x \in \mathbb{R}$, $|x|$ denotes its absolute value; $\mathbb{R}^n$ denotes the set of $n \times 1$ real column vectors; for any $x \in \mathbb{R}^n$, x^{T}, $\|x\|$ and $\|x\|_{A1}$ denote its transpose, Euclidean norm, and average L_1 norm, respectively; for any $t_1, t_2 \in \mathbb{N}$, $[t_1, t_2]$ denotes an integer set $\{t_1, t_1 + 1, \ldots, t_2 - 1, t_2\}$.

2 Problem formulation

Consider a general nonlinear dynamical system:

$$\dot{X} = F(X; p^s), \quad X(0) = X_0, \tag{1}$$

where $X = [x_1, \ldots, x_n]^{\mathrm{T}} \in \mathbb{R}^n$ denotes the state vector, $F(X; p^s) = [f_1(X; p^s), \ldots, f_n(X; p^s)]^{\mathrm{T}}$ represents the system dynamics, in which each $f_i(X; p^s)$ is a smooth and unknown continuous nonlinear function, p^s is a vector of system parameters, $\Phi(X_0; p^s)$ denotes the state trajectory started with the initial condition $X(0) = X_0$. Different recurrent trajectories[1)] generated from the above dynamical system (1) with different parameters are referred to as different dynamical patterns [30].

The sampling sequences yielded by sampling the trajectories (dynamical patterns) are represented as $\phi^s_{T,N}(X_0; p^s) := \{X(0), X(T), \ldots, X((N-1)T)\}$ (or ϕ^s for conciseness of presentation) with T being the sampling period and N being the total number of steps.

The Euler sampled-data model of system (1) can be expressed as

$$X[k+1] = X[k] + TF(X[k]; p^s) + \varepsilon(k; T), \quad X[0] = X_0, \tag{2}$$

where $X[k] := [x_1[k], \ldots, x_n[k]]^{\mathrm{T}} \in \mathbb{R}^n$ denotes the state vector of the Euler sampled-data model with $X[k] = X(kT)$, and $\varepsilon(k; T)$ denotes the modeling error of the Euler sampled-data model. The modeling error $\varepsilon(k; T)$ can be very small by choosing a small enough T. Thus, we have the following assumption regarding the sampling period.

Assumption 1. The sampling period T is a small enough positive constant.

Under Assumption 1, the sampling sequences ϕ^s can be approximately described by the Euler approximation model:

$$X[k+1] = X[k] + TF(X[k]; p^s), \quad X[0] = X_0, \tag{3}$$

where $X[k] := [x_1[k], \ldots, x_n[k]]^{\mathrm{T}} \in \mathbb{R}^n$ denotes the state vector of the Euler approximation model with $X[k] = X(kT)$.

The proposed rapid dynamical pattern recognition method contains two phases, the modeling phase and the recognition phase.

• Modeling phase. Consider the training set $\mathcal{P} := \{\phi^1, \phi^2, \ldots, \phi^M\}$, which contains M different training patterns, where each training pattern ϕ^s $(s = 1, \ldots, M)$ has different system parameters p^s[2)]. The aim of the modeling phase is to model/learn the inherent system dynamics $F(X; p^s)$ of training patterns ϕ^s and store the learned knowledge for further use.

• Recognition phase. Consider the test pattern yielded by sampling the recurrent trajectory generated from the following dynamical system:

$$\dot{X} = F(X; p^r), \quad X(0) = X_0, \tag{4}$$

where the vector of system parameters p^r is different from all p^s in training patterns ϕ^s, $s \in \{1, \ldots, M\}$. The test pattern is represented by sampling sequences $\phi^r_{T,N}(X_0; p^r) := \{X(0), X(T), \ldots, X((N-1)T)\}$ (or ϕ^r for conciseness of presentation). It can also be described by the Euler approximation model (3) with different system dynamics $F(X[k]; p^r)$.

The aim of the recognition phase is to search rapidly from the training patterns ϕ^s $(s = 1, \ldots, M)$ for those similar to the test pattern ϕ^r based on the inherent system dynamics of the corresponding patterns in the sampled-data framework.

3 Modeling and similarity of dynamical patterns consisting of sampling sequences

In this section, a sampled-data deterministic learning algorithm [38] is firstly presented for modeling of dynamical patterns consisting of sampling sequences. Some important steps about the sampled-data deterministic learning algorithm will be introduced for the completeness of presentation. Next, a similarity definition for sampling sequences (dynamical patterns) will be formulated.

1) Recurrent trajectories are the most important types (though not all types) of trajectories generated from nonlinear dynamical systems, which include periodic, quasi-periodic and even chaotic trajectories (see [28, 39] for a rigorous definition of the recurrent trajectory).

2) As shown in literature on nonlinear systems, different p^s, and sometimes different X_0, can generate different state trajectories.

3.1 Modeling dynamical patterns consisting of sampling sequences

Consider all training patterns ϕ^s $(s = 1, \ldots, M)$ in the training set $\mathcal{P}$. Each training pattern ϕ^s can be described by the Euler approximation model (3) with a different dynamical function vector $F(X[k]; p^s)$. For each training pattern ϕ^s, the following RBFN (radial basis function network)-based identifiers are used to model the unknown dynamical function $f_i(\cdot; p^s)$:

$$\hat{x}_i^s[k+1] = x_i[k] + \alpha_i(\hat{x}_i^s[k] - x_i[k]) + T\hat{W}_i^{s\mathrm{T}}[k]S[X[k]], \quad i \in \{1, \ldots, n\}, \tag{5}$$

where $\hat{x}_i^s$ is the state of identifiers, $X[k] = [x_1[k], \ldots, x_n[k]]$ denotes the state of the sampling sequence, $0 < \alpha_i < 1$ is the gain of identifiers, $\hat{W}_i^s$ denotes the weights of the RBFNs, and $S(X[k]) = [s_1(\|X[k] - \xi_1\|), \ldots, s_n(\|X[k] - \xi_m\|)]^{\mathrm{T}}$ is the regressor vector of RBFNs, in which $s_j(\|X[k] - \xi_j\|) = \mathrm{e}^{\frac{-(X[k]-\xi_j)^{\mathrm{T}}(X[k]-\xi_j)}{\eta^2}}$, $j \in \{1, \ldots, m\}$ is a radial basis function with $\xi_j \in \Xi$ being neurons constructed on a regular lattice in the input space, m being the number of neurons and η being the width of the receptive field.

The NN (neural network) weight update law is given by

$$\hat{W}_i^s[k+1] = \hat{W}_i^s(k) - T\gamma S(X[k])e_i(k+1), \quad i \in \{1, \ldots, n\}, \tag{6}$$

where γ is the learning gain of the update law, and $e_i[k] := \hat{x}_i[k] - x_i[k]$ denotes the tracking errors.

By using the above identifiers (5) and weight update law (6), it is rigorously proven in [38] that the system dynamics of sampling sequences can be locally-accurately modeled and stored in constant RBFNs.

Lemma 1 ([38])**.** Consider the sampling sequence ϕ^s sampled from the dynamical system (1) with constant sampling period T, dynamical RBFN identifiers (5), and NN weights update laws (6), where $\hat{W}_i^s(0) = 0$, $\forall i \in \{1, \ldots, n\}$. By choosing appropriate parameters α_i and γ, locally-accurate approximation for the unknown dynamics $f_i(\cdot; p^s)$ of system (1) along the sampling sequence ϕ^s is obtained by the time-invariant RBFNs $\bar{W}_i^{s\mathrm{T}} S(X[k])$, where

$$\bar{W}_i^s := \frac{1}{k_b - k_a + 1} \sum_{k=k_a}^{k_b} \hat{W}_i^s[k], \quad i \in \{1, \ldots, n\} \tag{7}$$

with $[k_a, k_b]$ denoting a time interval after the transient process.

Remark 1. The input trajectories of RBFNs are assumed to be period or recurrent trajectories in this paper. The neuron centers of RBFNs should be placed on a regular lattice to cover the compact set of recurrent input trajectories. It is shown in [27, 28, 38] that the corresponding regressor vector is persistently exciting, provided that the input variables to the RBFNs belong to certain neighborhoods of the neuron centers. This PE (persistence of excitation) condition leads to exponential stability of the identification error system along the recurrent system trajectory. Consequently, the estimated parameters (i.e., the weights of RBFNs) converge to their true values exponentially. Further, the inherent dynamics of sampling sequences can be locally-accurately modeled by the RBFNs. More detail analytical results of the sampled-data deterministic learning algorithm are reported in [38].

Based on the above deterministic learning algorithm, the inherent dynamics $F(X; p^s)$ within training patterns can be modeled/identified from the sampling sequences ϕ^s. The learned knowledge is stored in constant RBFNs and will be effectively reused in the following recognition phase.

3.2 Similarity of dynamical patterns consisting of sampling sequences

The following definition is formulated to characterize the similarity between a test pattern ϕ^r and a training pattern ϕ^s.

Definition 1. Consider a test dynamical pattern ϕ^r and a training dynamical pattern ϕ^s yielded by sampling the state trajectories generated from the systems (4) and (1), respectively. The test pattern ϕ^r is said to be similar to the training pattern ϕ^s, when the trajectory of ϕ^r stays within a neighborhood region of ϕ^s, and the differences in dynamics are small along the trajectory of ϕ^r, i.e.,

$$\max_{X[k] \in \phi^r} |f_i(X[k]; p^r) - f_i(X[k]; p^s)| < \epsilon_i^*, \quad \forall k \geqslant k_0,\ i \in \{1, \ldots, n\}, \tag{8}$$

where $\epsilon_i^* > 0$ denotes the similarity measure between the two dynamical patterns consisting of sampling sequences.

Different from the similarity for the continuous-time system [30], the definition proposed herein requires that the sampling period be small; otherwise, the differences in dynamics along the entire state trajectory cannot be effectively represented because of the large gap between neighboring sampling points.

In the modeling phase (Subsection 3.1), the inherent system dynamics within training patterns ($f_i(X[k]; p^s)$) can be modeled and approximated by time-invariant RBFNs $\bar{W}_i^{s\mathrm{T}} S(X[k])$ in a local region along the state trajectory of sampling sequences. The local region Ω_{ϕ^s} (or called the approximation region) can be expressed as follows [30]:

$$\Omega_{\phi^s} = \{X \,|\, \mathrm{dist}(X, \phi^s) < d_s \Rightarrow |\bar{W}_i^{s\mathrm{T}} S(X) - f_i(X; p^s)| < \varsigma_i^*,\ i \in \{1, \ldots, n\}\}, \tag{9}$$

where $\mathrm{dist}(X, \phi^s)$ denotes the smallest Euclidean distance from a fixed point X to a given trajectory of ϕ^s (i.e., $\mathrm{dist}(X, \phi^s) = \min_{Z \in \phi^s} \|X - Z\|$), $d_s > 0$ and $\varsigma_i^* > 0$ are constants, and ς_i^* denotes the approximation error within Ω_{ϕ^s}.

It implies that the system dynamics $f_i(X; p^s)$ of the training pattern ϕ^s along the trajectory of the test pattern ϕ^r can be represented by the well-trained RBFNs $\bar{W}_i^{s\mathrm{T}} S(X)$ with a small RBFN approximation error less than ς_i^*, if the trajectory of ϕ^r stays within the approximation region of the trajectory of ϕ^s. Based on this, Definition 1 can be rewritten as follows.

Definition 2. Consider the test dynamical pattern ϕ^r and the training dynamical pattern ϕ^s yielded by sampling the state trajectories generated from the system (4) and (1), respectively. The test pattern ϕ^r is said to be similar to the training pattern ϕ^s, when the trajectory of ϕ^r stays within the approximation region Ω_{ϕ^s} (given in (9)) of ϕ^s, and the differences in dynamics are small along the trajectory of ϕ^r, i.e.,

$$\max_{X[k] \in \phi^r} |f_i(X[k]; p^r) - \bar{W}_i^{s\mathrm{T}} S(X[k])| < \epsilon_i^* + \varsigma_i^*, \quad \forall k \geqslant k_0,\ i \in \{1, \ldots, n\}, \tag{10}$$

where ϵ_i^* and ς_i^* are defined by (8) and (9), respectively.

In Section 4, a discrete-time state synchronization technique is introduced to further deal with the comparison of dynamics between test and training patterns.

4 Dynamical pattern recognition and performance analysis

Consider all well-trained weights of RBFNs $\bar{W}_i^s$, $i \in \{1, \ldots, n\}$, $s \in \{1, \ldots, M\}$ for all training patterns. The following discrete-time dynamical estimators are constructed for all training pattern φ^s, $s \in \{1, \ldots, M\}$ in the training set $\mathcal{P}$:

$$\bar{x}_i^s[k+1] = u_i[k] + b_i(\bar{x}_i^s[k] - u_i[k]) + T\bar{W}_i^{s\mathrm{T}} S(U[k]), \quad i \in \{1, \ldots, n\}, \tag{11}$$

where $\bar{x}_i^s$ denotes the state of the dynamical estimator for training pattern φ^s, $U[k] = [u_1[k], \ldots, u_n[k]]^{\mathrm{T}}$ denotes the input of the estimator, and $0 < b_i < 1$ is the gain of the estimator.

Consider the test pattern ϕ^r sampled from (4) with a constant sampling period T. As stated in Section 2, the Euler approximation model can be used to represent the sampling sequence ϕ^r with a small sampling period

$$x_i[k+1] = x_i[k] + T f_i(X[k], p^r), \quad i \in \{1, \ldots, n\}, \tag{12}$$

where $X[k] = [x_1[k], \ldots, x_n[k]]$ denotes the state of the Euler approximation model.

The test pattern sequence ϕ^r is fed to all dynamical estimators (11) in parallel. Then the synchronization error (or called recognition error) can be calculated by $\tilde{x}_i^s[k] := \bar{x}_i^s[k] - x_i[k]$, $\forall i = \{1, \ldots, n\}$. Moreover, from (11) and (12), the synchronization error can be regarded as a state of the following discrete-time dynamical system (called the recognition error system):

$$\tilde{x}_i^s[k+1] = b_i \tilde{x}_i^s[k] + T(\bar{W}_i^{s\mathrm{T}} S(X[k]) - f_i(X[k]; p^r)), \quad \forall s \in \{1, \ldots, M\}, \tag{13}$$

where $\tilde{x}_i^s[k]$, $\forall i \in \{1, \ldots, n\}$ denotes the synchronization errors for the sth training pattern.

Note that the differences in dynamics $\bar{W}_i^{s\mathrm{T}} S(X[k]) - f_i(X[k]; p^r)$ appear in the right-hand side of (13). It indicates that the similarity measure in Definition 2 will reflect in the synchronization error. As a result, small differences in dynamics will lead to small synchronization errors.

For improving the reliability of recognition, the following average L_1 norm is used to facilitate the recognition decision making:

$$\|\tilde{x}_i^s[k]\|_{A1} := \frac{1}{T_e}\sum_{j=k-T_e+1}^{k}|\tilde{x}_i^s[j]|, \quad \forall k \geqslant T_e, \tag{14}$$

where $T_e \in N$ is a pre-set positive number and denotes the range of the average L_1 norm.

In the following, the theoretical analysis of the recognition process is presented.

Theorem 1. Consider the training set $\mathcal{P} = \{\phi^1, \phi^2, \ldots, \phi^M\}$, the test pattern ϕ^r, and the synchronization errors (13) generated from the discrete-time estimators (11). If the following conditions hold, then the similar training pattern ϕ^{s^*} can be rapidly recognized according to the smallest synchronization error principle (i.e., $\|\tilde{x}_i^{s^*}[k]\|_{A_1} < \|\tilde{x}_i^s[k]\|_{A_1},\ \forall i \in \{1, \ldots, n\},\ s \in \{1, \ldots, M\} - \{s^*\},\ k \geqslant T_e$).

(1) There exists a training pattern ϕ^{s^*}, $s^* \in \{1, \ldots, M\}$ similar to the given test pattern ϕ^r; i.e., the differences in the inherent system dynamics between the corresponding patterns satisfy

$$|\bar{W}_i^{s^*\mathrm{T}}S(X[k]) - f_i(X[k]; p^r)| < \epsilon_i^* + \varsigma_i^*, \quad \forall k \geqslant k_0,\ i \in \{1, \ldots, n\}. \tag{15}$$

(2) For the other patterns ϕ^s, $s \in \{1, \ldots, M\} - \{s^*\}$, there exists a time interval $\mathcal{I}_k = [T_{ak}, T_{bk}] \subset \mathcal{L}_k$, s.t.

$$|\bar{W}_i^{s\mathrm{T}}S(X[j]) - f_i(X[j]; p^r)| > \epsilon_i^* + \varsigma_i^* + \mu_i, \quad \forall j \in \mathcal{I}_k,\ \forall i \in \{1, \ldots, n\}, \tag{16}$$

where $\mu_i > 0$ denotes the similarity distinction, $l_k := T_{bk} - T_{ak} + 1$ denotes the length of the interval $\mathcal{I}_k$ and $\mathcal{L}_k = [k - T_e + 1, k]$ denotes the interval of average L_1 norm.

(3) The length of the interval $\mathcal{I}_k$ satisfies

$$l_k \geqslant \log_{b_i}\frac{1}{\frac{4(\epsilon_i^*+\varsigma_i^*)}{\mu_i}+3} + \frac{T_e}{1+\frac{\mu_i}{2(\epsilon_i^*+\varsigma_i^*)}} + 1 := \underline{l}. \tag{17}$$

Proof. The complete proof can be divided into three steps. First, it is proven that when condition (1) is satisfied, there exists an upper bound of the synchronization error's average L_1 norm of the similar training pattern ϕ^{s^*} (i.e., $\|\tilde{x}_i^{s^*}[k]\|_{A_1} < \frac{T(\epsilon_i^*+\varsigma_i^*)}{1-b_i},\ \forall k \geqslant T_e$). Second, it is proven that under condition (2), there exists a lower bound of the synchronization error's average L_1 norm of the other training patterns $\phi^s,\ s \in \{1, \ldots, M\} - \{s^*\}$ (i.e., $\|\tilde{x}_i^s[k]\|_{A_1} > \frac{T(l_k-l')(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{T_e(1-b_i)},\ \forall k \geqslant T_e$). Finally, with the satisfaction of condition (3), $\frac{T(l_k-l')(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{T_e(1-b_i)} \geqslant \frac{T(\epsilon_i^*+\varsigma_i^*)}{1-b_i}$ holds, which leads to $\|\tilde{x}_i^{s^*}[k]\|_{A_1} < \|\tilde{x}_i^s[k]\|_{A_1},\ \forall k \geqslant T_e$.

Step 1. Consider the synchronization error of the similar training pattern s^*:

$$\tilde{x}_i^{s^*}[k+1] = b_i\tilde{x}_i^{s^*}[k] + T(\bar{W}_i^{s^*\mathrm{T}}S(X[k]) - f_i(X[k]; p^r)). \tag{18}$$

The initial condition of the synchronization error is set to zero, which can be easily implemented by choosing $\bar{x}_i[k_0] = x_i[k_0]$, since $x_i[k_0]$ is available. Thus, the synchronization error (18) satisfies

$$\tilde{x}_i^{s^*}[k] = \sum_{j=k_0}^{k-1} Tb_i^{k-1-j}(\bar{W}_i^{s^*\mathrm{T}}S(X[j]) - f_i(X[j], p^r)), \quad \forall k \geqslant k_0. \tag{19}$$

From condition (1), we have $|\bar{W}_i^{s^*\mathrm{T}}S(X[k]) - f_i(X[k]; p^r)| < \epsilon_i^* + \varsigma_i^*$. The absolute value of the synchronization error (19) satisfies

$$\begin{aligned}
|\tilde{x}_i^{s^*}[k]| &= \left|\sum_{j=k_0}^{k-1} Tb_i^{k-1-j}\big(\bar{W}_i^{s^*\mathrm{T}}S(X[j]) - f_i(X[j], p^r)\big)\right| \\
&\leqslant \sum_{j=k_0}^{k-1} Tb_i^{k-1-j}\big|\big(\bar{W}_i^{s^*\mathrm{T}}S(X[j]) - f_i(X[j], p^r)\big)\big| \\
&< T\sum_{j=k_0}^{k-1} b_i^{k-1-j}(\epsilon_i^* + \varsigma_i^*) < \frac{T(\epsilon_i^* + \varsigma_i^*)}{1-b_i}.
\end{aligned} \tag{20}$$

Furthermore, by introducing the average L_1 norm (14) in the synchronization error (19), we have

$$\|\tilde{x}_i^{s^*}[k]\|_{A_1} = \frac{1}{T_e}\sum_{j=k-T_e+1}^{k}|\tilde{x}_i^{s^*}[j]| < \frac{1}{T_e}\sum_{j=k-T_e+1}^{k}\frac{T(\epsilon_i^*+\varsigma_i^*)}{1-b_i} = \frac{T(\epsilon_i^*+\varsigma_i^*)}{1-b_i}, \quad \forall k \geqslant T_e. \tag{21}$$

Step 2. Consider the synchronization error of the other patterns ϕ^s, $s \in \{1,\ldots,M\}-\{s^*\}$:

$$\tilde{x}_i^s[k+1] = b_i\tilde{x}_i^s[k] + T(\bar{W}_i^{s\mathrm{T}}S(X[k]) - f_i(X[k];p^r)). \tag{22}$$

From condition (2), we have a time interval $\mathcal{I}_k = [T_{ak}, T_{bk}] \subset \mathcal{L}_k$, s.t.

$$|\bar{W}_i^{s\mathrm{T}}S(X[j]) - f_i(X[j];p^r)| > \epsilon_i^* + \varsigma_i^* + \mu_i, \quad \forall j \in \mathcal{I}_k,\ \forall i \in \{1,\ldots,n\}. \tag{23}$$

Thus, the average L_1 norm of the synchronization error in (22) satisfies

$$\|\tilde{x}_i^s[k]\|_{A_1} = \frac{1}{T_e}\sum_{j=k-T_e+1}^{k}|\tilde{x}_i^s[j]| > \frac{1}{T_e}\sum_{j\in\mathcal{I}_k}|\tilde{x}_i^s[j]|, \tag{24}$$

where the synchronization error in the interval $\mathcal{I}_k$ satisfies

$$\tilde{x}_i^s[\tau] = b_i^{\tau-T_{ak}}\tilde{x}_i^s[T_{ak}] + \sum_{j=T_{ak}}^{\tau-1}Tb_i^{\tau-1-j}(\bar{W}_i^{s\mathrm{T}}S(X[j]) - f_i(X[j];p^r)), \quad \forall\tau \in \mathcal{I}_k. \tag{25}$$

Define an interval $\mathcal{I}_k' = \{k\,|\,|\tilde{x}_i^s[k]| < \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}\} \subset \mathcal{I}_k$. Based on the interval $\mathcal{I}_k'$, synchronization error's analysis of (25) with three cases of different initial conditions $\tilde{x}_i^s[T_{ak}]$ are listed as follows (See Appendix A for the complete proof).

Case (i). If $\tilde{x}_i^s[T_{ak}] \in \mathcal{I}_k'$, then

$$|\tilde{x}_i^s[\tau]|\begin{cases} < \dfrac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}, & \forall\tau \in [T_{ak}, T_{ak}'], \\ \geqslant \dfrac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}, & \forall\tau \in [T_{ak}'+1, T_{bk}], \end{cases} \tag{26}$$

where $T_{ak}' - T_{ak} + 1 \leqslant l' = \log_{b_i}\frac{1}{\frac{4(\epsilon_i^*+\varsigma_i^*)}{\mu_i}+3} + 1$.

Case (ii). If $\tilde{x}_i^s[T_{ak}] \notin \mathcal{I}_k'$, and $\tilde{x}_i^s[T_{ak}]$ has the same sign with $\bar{W}_i^{s\mathrm{T}}S(X[T_{ak}]) - f_i(X[T_{ak}];p^r)$, then $|\tilde{x}_i^s[\tau]| \geqslant \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$, $\forall\tau \in \mathcal{I}_k$.

Case (iii). If $\tilde{x}_i^s[T_{ak}] \notin \mathcal{I}_k'$, and $\tilde{x}_i^s[T_{ak}]$ has the different sign with $\bar{W}_i^{s\mathrm{T}}S(X[T_{ak}]) - f_i(X[T_{ak}];p^r)$, then

$$|\tilde{x}_i^s[\tau]|\begin{cases} \geqslant \dfrac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}, & \forall\tau \in [T_{ak}, T_{ak}'-1], \\ < \dfrac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}, & \forall\tau \in [T_{ak}', T_{bk}'], \\ \geqslant \dfrac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}, & \forall\tau \in [T_{bk}'+1, T_{bk}], \end{cases} \tag{27}$$

where $T_{bk}' - T_{ak}' + 1 = l' = \log_{b_i}\frac{1}{\frac{4(\epsilon_i^*+\varsigma_i^*)}{\mu_i}+3} + 1$.

Based on the above analysis, we can conclude that the largest length of the interval $\mathcal{I}_k'$ is $l' = \log_{b_i}\frac{1}{\frac{4(\epsilon_i^*+\varsigma_i^*)}{\mu_i}+3} + 1$. Thus, from (24), we have

$$\begin{aligned}\|\tilde{x}_i^s[k]\|_{A_1} &= \frac{1}{T_e}\sum_{j=k-T_e+1}^{k}|\tilde{x}_i^s[j]| > \frac{1}{T_e}\sum_{j\in\mathcal{I}}|\tilde{x}_i^s[j]| \\ &= \frac{1}{T_e}\left(\sum_{j\in\mathcal{I}_k-\mathcal{I}_k'}|\tilde{x}_i^s[j]| + \sum_{j\in\mathcal{I}_k'}|\tilde{x}_i^s[j]|\right)\end{aligned}$$

$$> \frac{1}{T_e}\sum_{j\in\mathcal{I}_k-\mathcal{I}'_k}|\tilde{x}_i^s[j]| > \frac{1}{T_e}\sum_{j\in\mathcal{I}_k-\mathcal{I}'_k}\frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$$
$$= \frac{T(l_k-l')(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{T_e(1-b_i)}. \tag{28}$$

Step 3. Finally, if (17) holds with $l' = \log_{b_i}\frac{1}{\frac{4(\epsilon_i^*+\varsigma_i^*)}{\mu_i}+3}+1$, we have

$$l_k - l' > \frac{T_e}{1+\frac{\mu_i}{2(\epsilon_i^*+\varsigma_i^*)}} = \frac{T_e(\epsilon_i^*+\varsigma_i^*)}{\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2}}. \tag{29}$$

By combining (28) with (29), we can get that

$$\|\tilde{x}_i^s[k]\|_{A_1} > \frac{T(l_k-l')(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{T_e(1-b_i)} = \frac{T(\epsilon_i^*+\varsigma_i^*)}{1-b_i}. \tag{30}$$

Further, from (21) and (30), it can be guaranteed that $\|\tilde{x}_i^s[k]\|_{A_1} > \|\tilde{x}_i^{s^*}[k]\|_{A_1},\ \forall k \geqslant T_e$.

Remark 2. The recognition conditions indicate that the test pattern can be rapidly recognized according to the smallest synchronization error principle, if (a) the differences in dynamics between the test and the similar training pattern is small, (b) and the others (the differences in dynamics between the test and the other training patterns) are large. In the sampled-data framework, these recognition conditions can be verified step by step based on historical sampling data. This makes a distinction compared with the previous results for continuous-time nonlinear systems [30, 31]. It is worth noting that, the recognition conditions are very difficult to be verified by using continuous-time signals. Even when misrecognition occurs, the situation is difficult to be analyzed due to the lack of condition verification.

Remark 3. The above result implies that the recognition time is less than or equal to T_e (the range of the average L_1 norm). In previous study [31], T_e is set to be the period of the test pattern. By contrast, T_e defined herein can be less than the period of the test pattern. Smaller T_e will lead to less time for recognition. Moreover, for the recognition conditions given in [31], the selection of gains of dynamical estimators relies on the prior knowledge of dynamical patterns. This restriction is removed in this paper, which makes the conditions more practical and meaningful.

5 Simulation

In this section, numerical simulation examples will be given to demonstrate the effectiveness of our proposed method. On one hand, we will verify the result of the rapid recognition (i.e., the recognition time is less than T_e). On the other hand, we will further verify the recognition conditions step by step based on the historical sampling data. This makes a distinction compared with the previous results for continuous-time nonlinear systems [30, 31], in which the recognition conditions are difficult to be verified by using continuous-time signals.

5.1 Data description and identification of system dynamics

Two types of nonlinear oscillation systems [40] are used to generate sampling sequences. One is the Duffing system:

$$\begin{aligned}\dot{x}_1 &= x_2,\\ \dot{x}_2 &= -p_2x_1 - p_3x_1^3 - p_1x_2 + q\cos(\omega t),\end{aligned} \tag{31}$$

where $f_d(X) = -p_2x_1 - p_3x_1^3 - p_1x_2$ is the dynamical function of the Duffing system with p_1, p_2, p_3, q, ω being the system parameters.

The other is the Duffing-VanderPol system:

$$\begin{aligned}\dot{x}_1 &= x_2,\\ \dot{x}_2 &= p_1(1-x_1^2)x_2 - p_2x_1 - p_3x_1^3 + q\cos(\omega t),\end{aligned} \tag{32}$$

Table 1 System parameters of training patterns

Pattern	p_1	p_2	p_3	q	ω	Initial state X_0
Duf1	1.2	−1.5	1	0.9	1.8	[0.438; 0.07713]
Duf2	0.4	−1.5	1	0.9	1.8	[0.438; 0.07713]
Duf3	0.55	−1.1	1	1.498	1.8	[0.438; 0.07713]
Duf4	0.2	−1.1	1	1.498	1.8	[0.438; 0.07713]
DVan1	0.6	1	0.8	1	1.498	[1.3; 2.2]
DVan2	0.6	1	1.3	1	1.498	[1.3; 2.2]

Table 2 Transformation parameters of different systems

System	Shifting of x_1 (S_{h1})	Scaling of x_1 (S_{c1})	Shifting of x_2 (S_{h2})	Scaling of x_2 (S_{c2})
Duf(1,2)	0	1/1.2	−0.8	1/1.2
Duf(3,4)	0	1/3.5	0	1/3.5
DVan	0	1/5	0	1/5

where $f_v(X) = p_1(1 - x_1^2)x_2 - p_2x_1 - p_3x_1^3$ is the dynamical function of the Duffing-VanderPol system with p_1, p_2, p_3, q, ω being the system parameters.

For each type of oscillation system, we use specific parameters to generate different dynamical patterns within sampling sequences with sampling period 0.01 s and total time 150 s. Six groups of sampling sequences (four for the Duffing system and two for the Duffing-Vanderpol system) are taken as the data of training patterns (or TRP for short). Specific parameters of different TRPs are shown in Table 1, where Duf2, Duf4, and DVan2 belong to the category of chaotic trajectories.

Since the trajectories of different patterns stay within areas of different sizes in phase space, a preprocessing step is needed to normalize all the sampling sequences, such that the trajectories stay within the same region $[-1, 1] \times [-1, 1]$. Specifically, the formulas $x_{1\text{tran}} = (x_1 + S_{h1})S_{c1}$, $x_{2\text{tran}} = (x_2 + S_{h2})S_{c2}$ are taken into account. And the specific transformation parameters of different systems are shown in Table 2. Similarly, the preprocessing step corresponding to the training pattern is also used to normalize the data of the test pattern (or TEP for short). Note that the system dynamics will also be changed in the normalization (i.e., $f_{d\text{tran}}(x_{1\text{tran}}, x_{2\text{tran}}) = f_d(\frac{x_{1\text{tran}}}{S_{c1}} - S_{h1}, \frac{x_{2\text{tran}}}{S_{c2}} - S_{h2})$ and $f_{v\text{tran}}(x_{1\text{tran}}, x_{2\text{tran}}) = f_v(\frac{x_{1\text{tran}}}{S_{c1}} - S_{h1}, \frac{x_{2\text{tran}}}{S_{c2}} - S_{h2})$). Nevertheless, the morphological character of the original dynamics is not changed.

Before the recognition process, we need to identify the system dynamics in all training patterns. To this end, we construct RBFNs containing 21×21 neurons covering the region $[-1, 1] \times [-1, 1]$ regularly. The receptive field width is set as $\eta = 0.1$. The deterministic learning algorithm presented in Subsection 2.2 is used to identify the dynamic information in these training patterns[3)]. Figure 1 shows the knowledge representation of all TRPs. Figure 2 shows the NN approximation of system dynamics along the TEP's trajectory (f_d for the dynamical function of Duffing systems, and f_v for the dynamical function of Duffing-VanderPol systems). It is shown that good NN approximation of system dynamics can be achieved along the TEP's trajectory via the deterministic learning algorithm.

Next, two scenarios will be considered. In the first scenario, we select Duf1, Duf2, DVan1, and DVan2 as the four TRPs in order. The pattern, whose parameters are close to Duf1, is selected as the TEP. The objective of this scenario is to verify the effectiveness of the main results in this paper (verify the proposed recognition conditions step by step based on the historical sampling sequences). In the second scenario, Duf3, Duf4, DVan1, and DVan2 are selected as the four TRPs in order. We choose another pattern in Duffing systems as the TEP. The second scenario mainly explores what is the most representative training patterns in our method. The specific system parameters of TEPs in two scenarios are shown in Table 3. Figure 3 shows the state trajectories of test and training patterns in two scenarios.

5.2 Scenario 1

In this scenario, Duf1, Duf2, DVan1, and Dvan2 are selected as the TRPs. A group of discrete-time dynamical estimators are constructed with the well-training RBFNs[4)].

3) To identify the dynamics f_d in Duffing patterns, the identifier is slightly modified as $\hat{x}[k+1] = x_2[k] + \alpha(\hat{x}[k] - x_2[k]) + T\bar{W}^{\mathrm{T}}[k]S[X[k]] + Tq\cos(\omega kT)$ to remove the influence of the time-varying term $q\cos(\omega t)$.

4) Similarly, the estimator is slightly modified as $\bar{x}^s[k+1] = x_2[k] + b(\bar{x}^s[k] - x_2[k]) + T\bar{W}^{s\mathrm{T}}[k]S[X[k]] + Tq\cos(\omega kT)$ to remove the influence of the time-varying term $q\cos(\omega t)$.

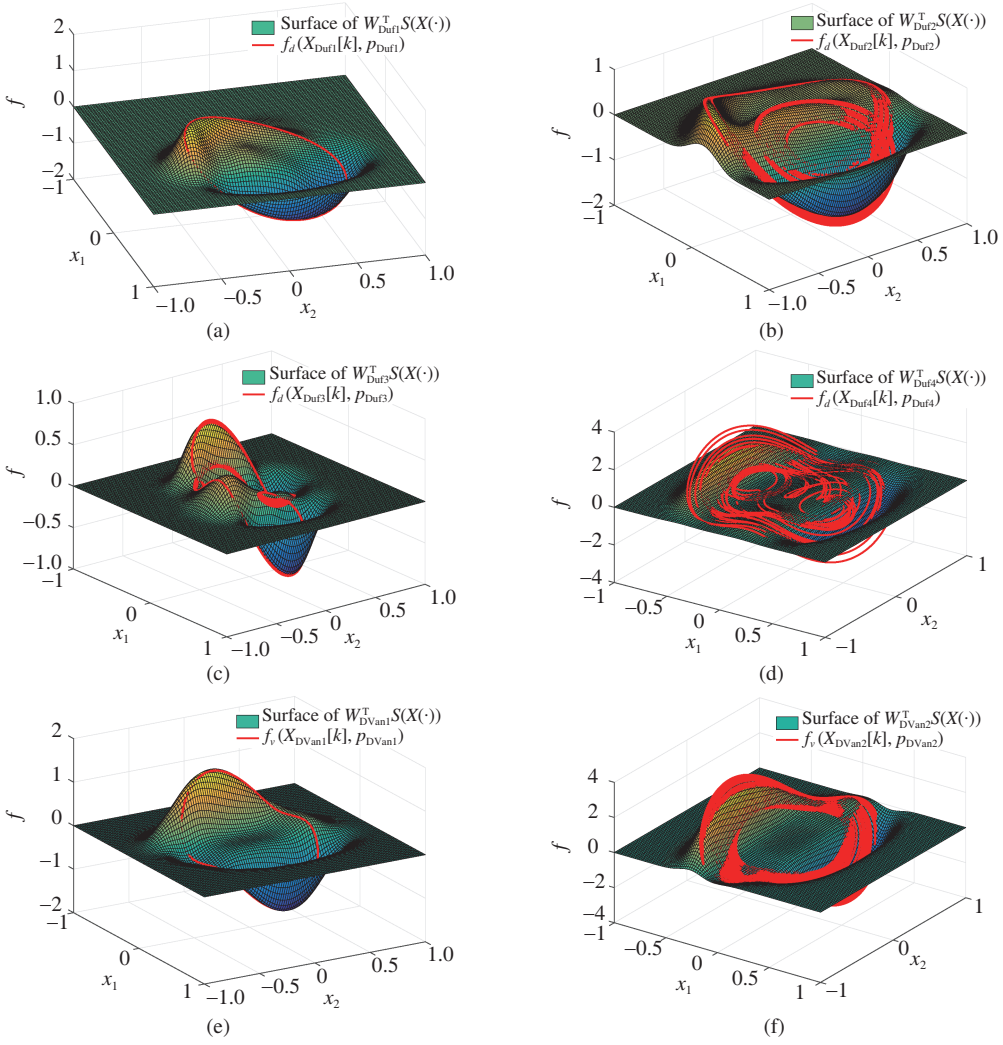

Figure 1 (Color online) Function approximation of f_d of (a) Duf1, (b) Duf2, (c) Duf3, and (d) Duf4 in space. Function approximation of f_v of (e) DVan1 and (f) DVan2 in space.

In the recognition phase, we input the test pattern sequences into the discrete-time dynamical estimators and further calculate the synchronization error and its average L_1 norm as shown in Figures 4(a) and (b), where the gain of estimators is set as $b = 0.1$ and the range of the average L_1 norm T_e is set as the period of TEP (i.e., 350 steps). It is shown in Figure 4(b) that the most similar pattern (TRP1) is recognized in 51 steps (less than T_e steps) successfully. Further, based on the historical sampling data, the information of dynamic differences in the sense of Definition 2 can be calculated and shown in Figure 4(c). It can be found out that the figure of the synchronization error (i.e., Figure 4(a)) is quite similar to the figure of the dynamic differences (i.e., Figure 4(c)).

Next, the effectiveness of the proposed recognition conditions will be verified step by step. Figure 5 shows the information about the dynamic differences between the TEP and TRPs in the steady-state process. Specifically, in this scenario, TRP1 is the most similar training pattern to the TEP. As shown in the black dashed-line box in Figure 5, the dynamic differences between TEP and TRP1 are less than 0.03, i.e., $\epsilon^* + \varsigma^* = 0.03$ in condition (1). For the second similar pattern TRP2, the average L_1 norm interval $\mathcal{L}_k$ is shown in the green dashed-line box. Moreover, the corresponding subinterval $\mathcal{I}_k$ as described in condition (2) can be shown in the blue dashed-line box, where the similarity distinction is $\mu = 0.17$ (i.e., the dynamic differences are larger than $\epsilon^* + \varsigma^* + \mu = 0.2$ as shown in the red dashed-line box). Note that the length l_k of the subinterval $\mathcal{I}_k$ is changed over time. Based on the historical sampling data, we can calculate the relations of l_k for different TRPs, as shown in Figure 6(a). Note that since the trajectory of TEP is periodic, the relation shown in Figure 6(a) will repeat periodically.

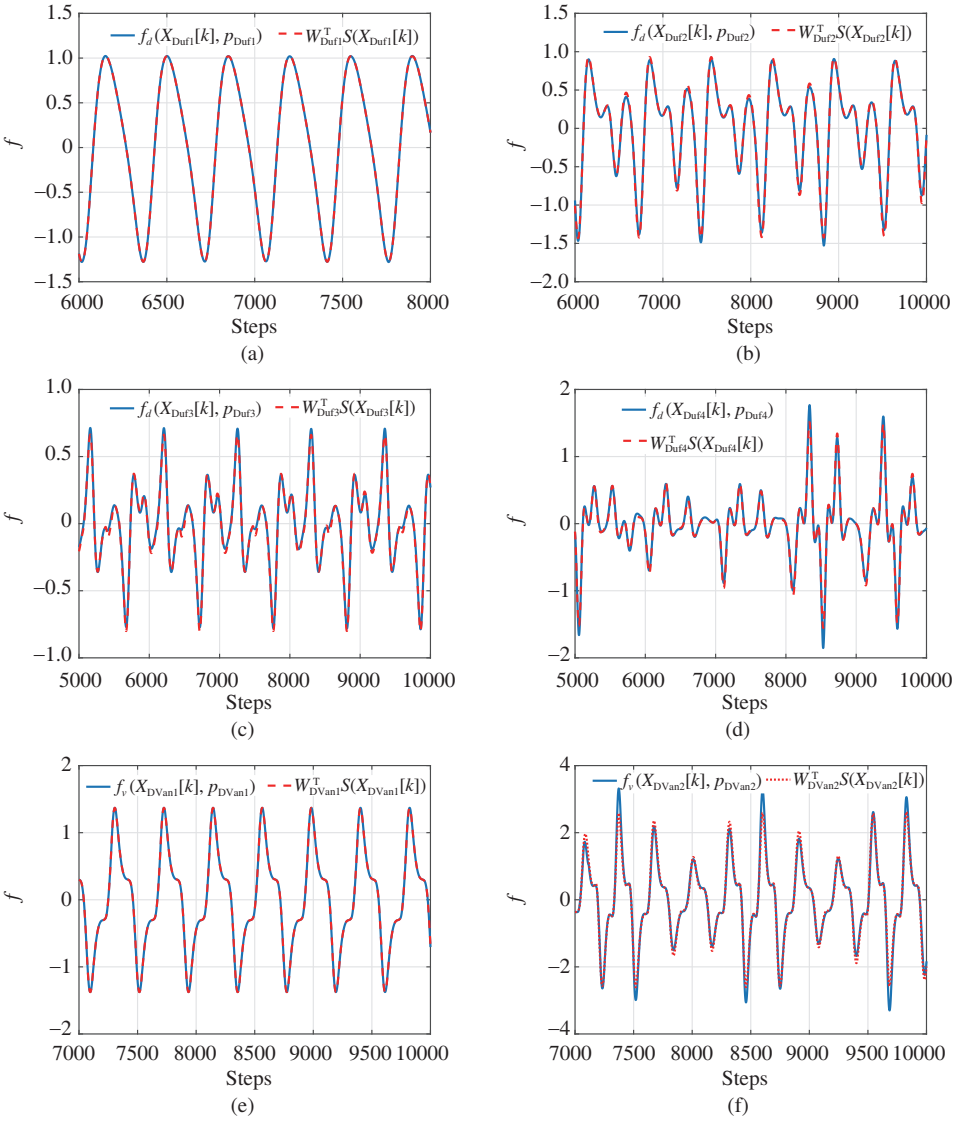

Figure 2 (Color online) Function approximation of f_d of (a) Duf1, (b) Duf2, (c) Duf3, and (d) Duf4 along the time axis. Function approximation of f_v of (e) DVan1 and (f) DVan2 along the time axis.

Table 3 System parameters of test patterns

Pattern	p_1	p_2	p_3	q	ω	Initial state X_0
TEP in Scenario 1 (Duf)	1.22	−1.5	1	0.9	1.8	[0.438; 0.07713]
TEP in Scenario 2 (Duf)	2	−1.3	1	1.498	1.8	[0.438; 0.07713]

Based on the results from above analysis, we know that $\epsilon^* + \varsigma^* = 0.03$, $\mu = 0.17$, $T_e = 350$, $b = 0.1$. Thus, we can calculate $\underline{l} = 54.0689$ from condition (3). It is seen in Figure 6(a) that for TRP2 $\min_{i\in[i_b+T_e-1,i_e]} l_k = 95$, for TRP3 $\min_{i\in[i_b+T_e-1,i_e]} l_k = 118$, for TRP4 $\min_{i\in[i_b+T_e-1,i_e]} l_k = 92$, where i_b and i_e denote the beginning point and the ending point of the steady-state process of TEP, respectively. It is verified that $\min_{i\in[i_b+T_e-1,i_e]} l_k$ of TRP2–TRP4 are larger than the theoretical value $\underline{l}$. Thus, all three conditions of the theorem can be guaranteed simultaneously.

As a result, the most similar pattern will be recognized in T_e steps based on the average L_1 norm of the synchronization error. Note that when $k < T_e$, the average L_1 norms of the synchronization errors are unavailable. To this end, a modified average L_1 norm $|||\tilde{x}[k]|||_{A1} = \frac{\sum_{j=1}^{k} \tilde{x}[j]}{k}, k < T_e$ is used instead of the original one in this simulation. Nevertheless, this modification does not affect the recognition result that the most similar pattern will be recognized in T_e steps.

At the end of the identification process, we have to verify whether the trajectory of the TEP stays

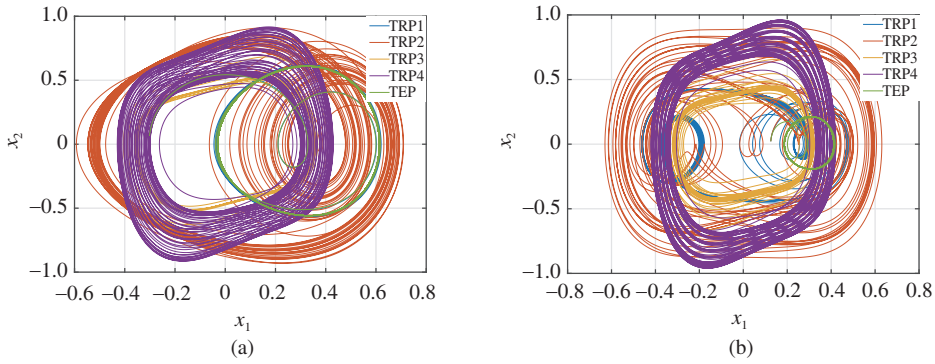

Figure 3 (Color online) State trajectories in (a) the first and (b) the second scenarios.

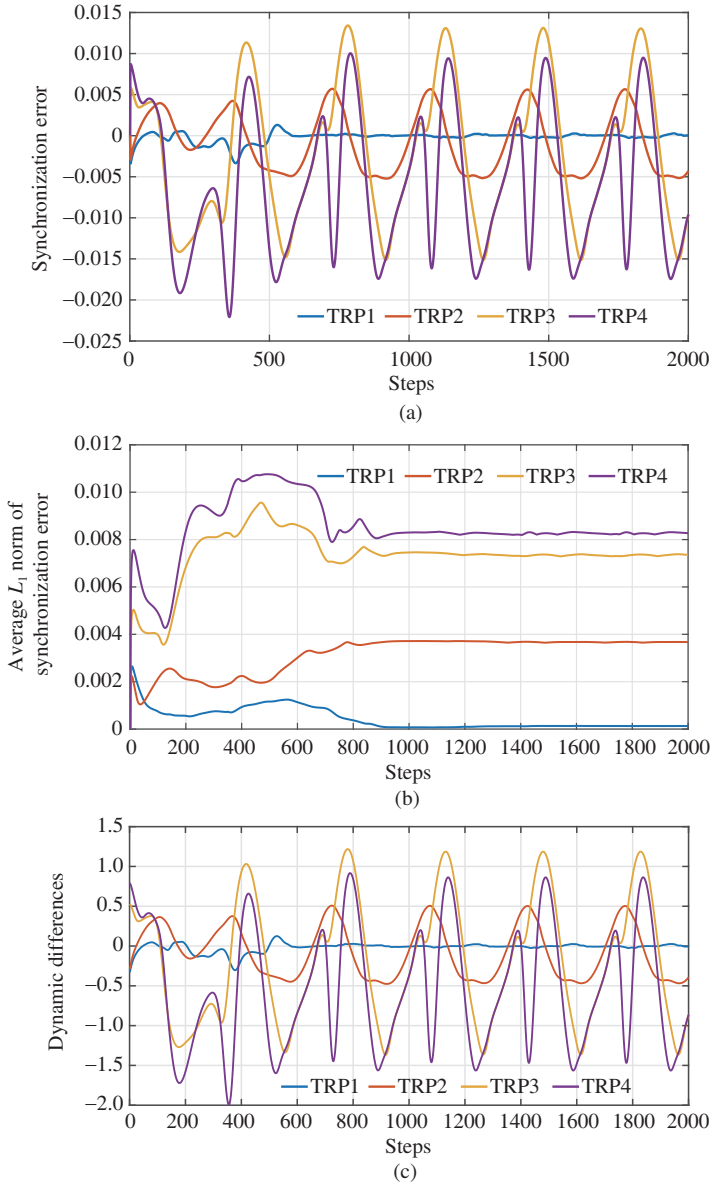

Figure 4 (Color online) (a) Synchronization error of four training patterns in the first scenario; (b) average L_1 norm of synchronization error of four training patterns; (c) information of dynamic differences in the sense of Definition 2.

within the approximation region of the similar TRP. We use $d(k) = \min_{i\in[i_{b1},i_{e1}]} \|X_{\mathrm{TEP}}(k) - X_{\mathrm{TRP1}}(i)\|_2$ to describe the distance from each point of the TEP trajectory to the TRP1, where i_{b1}, i_{e1} denote the

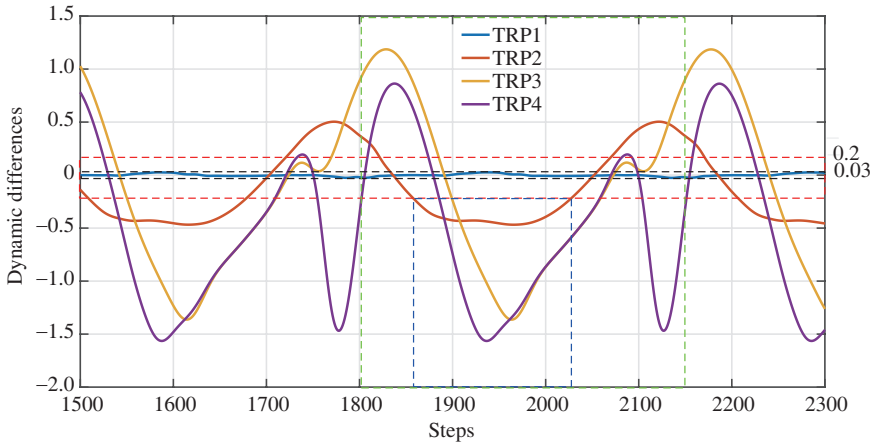

Figure 5 (Color online) Information of dynamic differences in the sense of Definition 2 in the steady-state process.

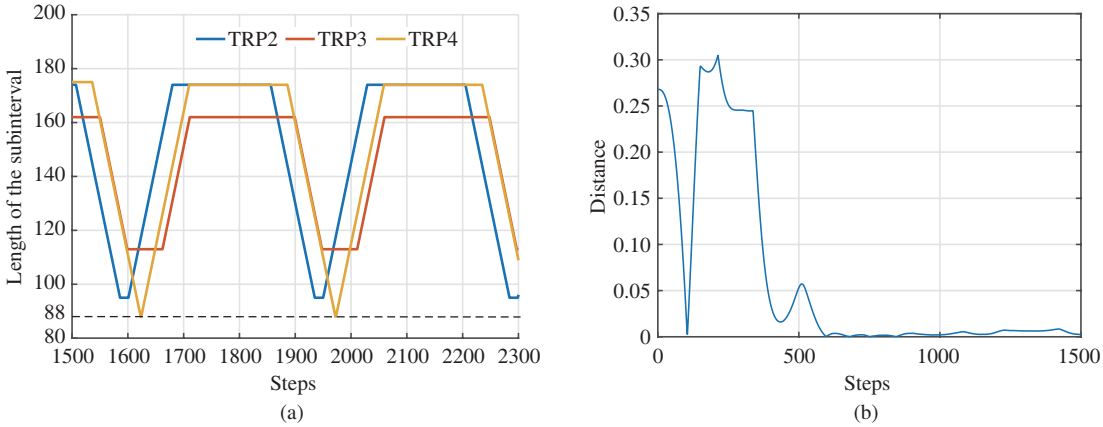

Figure 6 (Color online) (a) The length of subinterval of the information of dynamic differences of different TRPs; (b) distance from TEP to TRP1.

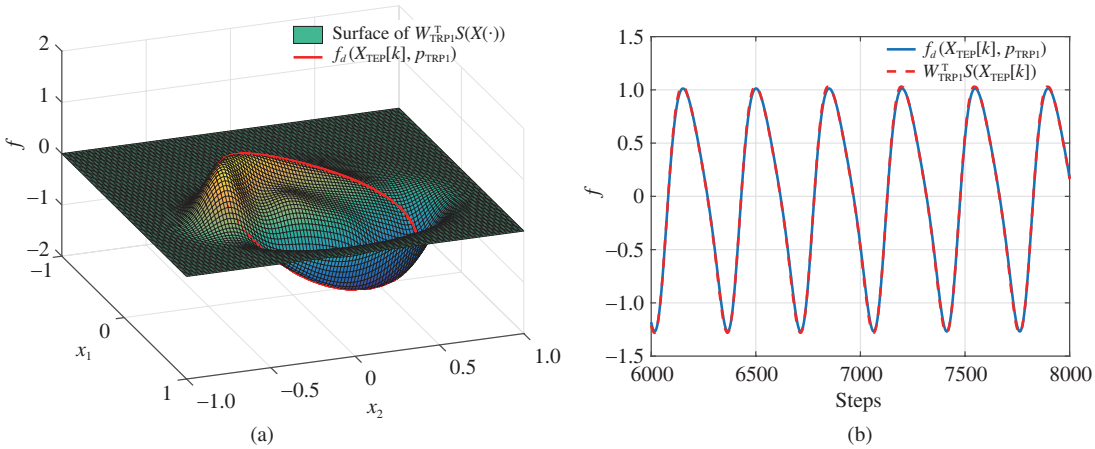

Figure 7 (Color online) (a) RBF representation of f_d of TRP1 along the test pattern trajectories in space; (b) RBF representation of f_d of TRP1 along the time axis.

beginning point and the ending point of the steady-state process of TRP1. It is shown in Figure 6(b) that when the TEP trajectory enters the steady-state process, the distance to the TRP1 is much small (smaller than 0.01). It is revealed that TEP stays within the approximation region of the similar TRP (i.e., the TRP1). Furthermore, it is shown in Figures 7(a) and (b) that good NN approximation of system dynamics of TRP1 can be achieved along the TEP's trajectory when the TEP stays within the approximation region of the TRP1.

To ensure the effective recognition, condition (1) indicates that there must exist a training pattern similar to the test pattern in the training set $\mathcal{P}$ (i.e., the dynamic differences $\epsilon^* + \varsigma^*$ are small enough). This condition can be usually satisfied in real applications (by obtaining enough patterns). Besides, another key factor of condition (1) is that the trajectory of the test pattern should stay within the approximation region of the similar training pattern. Fortunately, this problem can be verified by using a simple numerical calculation at the end of the recognition process. Condition (2) shows that there must exist a similarity distinction $\mu > 0$ between the similar training pattern and the other training patterns. It means that the test pattern can be accurately recognized based on the average L_1 norm of the synchronization errors, if there exists a subinterval $\mathcal{I}_k$ with sufficient similarity distinction μ in the average L_1 norm interval $\mathcal{L}_k$. And the dynamic differences in the complementary interval are not limited (i.e., the dynamic differences between the test pattern and the training pattern can be zero in the complementary interval of $\mathcal{I}_k$). It should be pointed out that after choosing T_e, in each time-interval $\mathcal{L}_k$, the length l_k of the subinterval $\mathcal{I}_k$ is changed over time. Based on the historical sampling data, the relations of l_k for different TRPs can be calculated, as shown in Figure 6(a). Condition (3) reveals that there exists a relationship between the distinction μ and the required length $\underline{l}$ of the interval $\mathcal{I}_k$ in condition (2). All these conditions can be verified step by step based on historical sampling data.

Remark 4. This makes a distinction compared with the previous results [30, 31] for continuous-time nonlinear system. The proposed recognition conditions herein can be verified step by step based on historical sampling data in the sampled-data framework. It is worth noting that the precise description of dynamical functions is often unavailable in practical engineering. Nevertheless, sampling data of the test dynamical pattern can be recorded. Based on the historical sampling data, the sampled-data deterministic learning algorithm is employed for accurately modeling dynamics of the test pattern along its trajectory. Based on this, the proposed recognition conditions can be verified by using the modeling result.

Although the expression of the dynamical function is unavailable in real-world applications, we can use the sampled-data deterministic learning algorithm to model the dynamics of the test pattern again based on the historical sampling data. Therefore, the proposed recognition conditions can also be verified in real-world applications.

5.3 Scenario 2

In the following scenario, we further investigate the selection of the most representative training pattern. Duf3, Duf4, DVan1, and Dvan2 are selected as the TRPs. A group of discrete-time dynamical estimators are constructed with the well-training RBFNs[5)].

Similarly, we input the test pattern sequences into the discrete-time dynamical estimators and further calculate the synchronization error and its average L_1 norm as shown in Figures 8(a) and (b), where the gain of estimators is set as $b = 0.1$. As shown in Figure 8(b), TRP2 is recognized as the most similar pattern to TEP. As shown in Figure 8(c), TRP2 is recognized as the most similar pattern to TEP. Figure 8(c) shows the information of dynamic differences in the sense of Definition 2, which is calculated by using the historical sampling data.

In this scenario, both TRP1 and TRP2 belong to the dynamical system (the Duffing system), which is the same as the TEP. Their trajectories are shown in Figure 9(a). It can be seen that the trajectory of the TRP2 is a chaotic trajectory, which explores much larger areas in phase space than the periodic trajectory of the TRP1. In this case, most of the trajectory of TEP stays in the approximation region of TRP2 rather than in the approximation region of TRP1. The distances from the TEP trajectory to the TRP1 and TRP2 are calculated by the formula introduced in Subsection 5.2, $d_1(k) = \min_{i\in[i_{b1},i_{e1}]} \|X_{\mathrm{TEP}}(k) - X_{\mathrm{TRP1}}(i)\|_2$ and $d_2(k) = \min_{i\in[i_{b2},i_{e2}]} \|X_{\mathrm{TEP}}(k) - X_{\mathrm{TRP2}}(i)\|_2$, respectively. As shown in Figure 10, compared with the distance to the TRP1, the distance from the TEP to the TRP2 is much smaller. It is clearly verified in Figure 11 that the RBF approximation error of TRP1 is much larger than the error of TRP2. Consequently, even though the actual dynamic differences between TEP and TRP1 are smaller than those between TEP and TRP2 slightly, it is recognized that the TRP2 is the most similar pattern rather than the TRP1.

The result indicates that compared with chaotic trajectories (like the TRP2), periodic trajectories (like the TRP1) explore limited areas in phase space. It leads to a limited approximation region of the corre-

5) The estimator is slightly modified as $\bar{x}^s[k+1] = x_2[k] + b(\bar{x}^s[k] - x_2[k]) + T\bar{W}^{s\mathrm{T}}[k]S[X[k]] + Tq\cos(\omega kT)$ to remove the influence of the time-varying term $q\cos(\omega t)$.

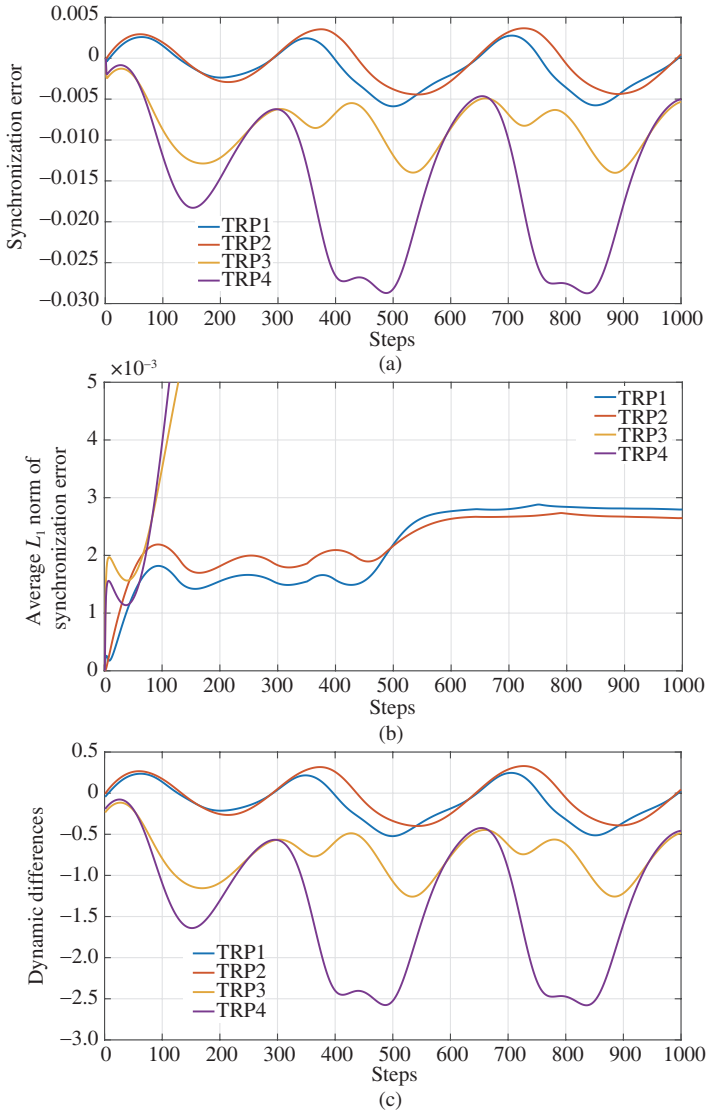

Figure 8 (Color online) (a) Synchronization error of four training patterns in the second scenario; (b) average L_1 norm of synchronization error of four training patterns; (c) information of dynamic differences in the sense of Definition 2.

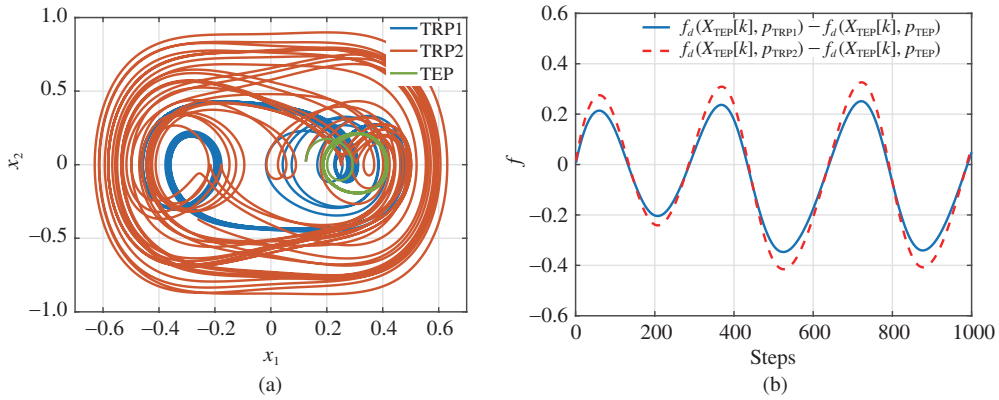

Figure 9 (Color online) (a) Comparison in trajectories; (b) dynamic differences between TEP and TRPs.

sponding TRP. When the parameters of TEP are much different from the periodic TRP (like the TRP1), most of the TEP trajectory may not stay within limited approximation region of the corresponding TRP. By contrast, the chaotic patterns are more spatially expanded, which implies that chaotic patterns are

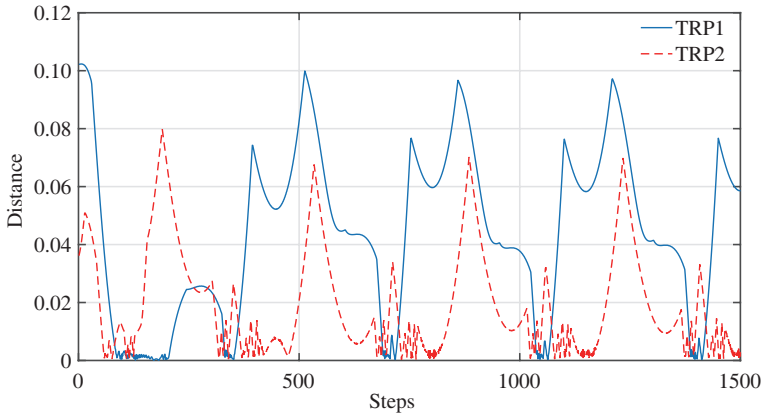

Figure 10 (Color online) Distance from TEP to TRPs.

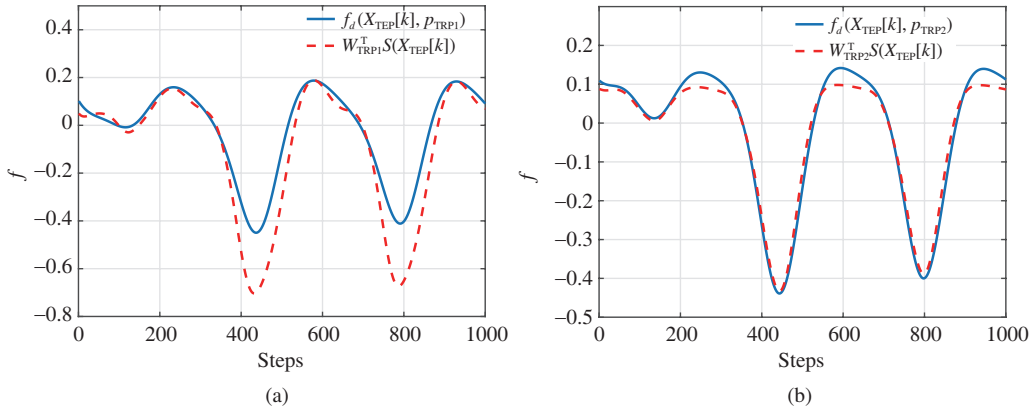

Figure 11 (Color online) RBF representation of f_d of (a) TRP1 and (b) TRP2 along the time axis.

more suitable for use as the dynamical patterns stored in the training set.

6 Conclusion

In the present paper, we have proposed a rapid dynamical pattern recognition method for the sampled-data framework. In the modeling phase, the sampling sequences of the training patterns are modeled by using sampled-data deterministic learning. The obtained knowledge is then employed to construct a set of discrete-time dynamical estimators to represent all training patterns. In the recognition phase, average L_1 norms of synchronization errors are considered to reflect similarities between the test and training patterns. General conditions for accurate and rapid recognition of dynamical patterns are established. We would like to emphasize that the proposed recognition conditions can be verified step by step based on historical sampling data in the sampled-data framework. This makes a distinction compared with the previous results for continuous-time nonlinear systems, in which the recognition conditions are difficult to be verified using continuous-time signals. As a follow-up to the present paper, real applications on fault detection and fault isolation will be implemented.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant No. 61890922) and in part by National Major Scientific Instruments Development Project (Grant No. 61527811).

References

1 Kadous M W. Temporal classification: extending the classification paradigm to multivariate time series. Dissertation for Ph.D. Degree. Kensington: University of New South Wales, 2002
2 Dietterich T G. Machine learning for sequential data: a review. In: Proceedings of Joint IAPR International Workshops SSPR and SPR, 2002. 15–30
3 Keogh E, Ratanamahatana C A. Exact indexing of dynamic time warping. Knowl Inf Syst, 2005, 7: 358–386
4 Xi X P, Keogh E, Shelton C, et al. Fast time series classification using numerosity reduction. In: Proceedings of International Conference on Machine Learning, 2006

5 Wang M, Wang Z D, Chen Y, et al. Adaptive neural event-triggered control for discrete-time strict-feedback nonlinear systems. IEEE Trans Cybern, 2020, 50: 2946–2958
6 Huang Z K, Cao J D, Raffoul Y N. Hilger-type impulsive differential inequality and its application to impulsive synchronization of delayed complex networks on time scales. Sci China Inf Sci, 2018, 61: 078201
7 Ljung L. Perspectives on system identification. IFAC Proc Volume, 2008, 41: 7172–7184
8 Yang Q, Wu X D. 10 challenging problems in data mining research. Int J Info Tech Dec Mak, 2006, 5: 597–604
9 Gales M, Young S. The application of hidden Markov models in speech recognition. FNT Signal Process, 2007, 1: 195–304
10 Rabiner L R. A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE, 1989, 77: 257–286
11 Yamato J, Ohya J, Ishii K. Recognizing human action in time-sequential images using hidden Markov model. In: Proceedings of Computer Vision and Pattern Recognition, 1992. 379–385
12 Turaga P, Chellappa R, Subrahmanian V S, et al. Machine recognition of human activities: a survey. IEEE Trans Circ Syst Video Technol, 2008, 18: 1473–1488
13 LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521: 436–444
14 Pascanu R, Gulcehre C, Cho K, et al. How to construct deep recurrent neural networks. In: Proceedings of International Conference on Learning Representations, 2014
15 Sutskever I, Martens J, Hinton G E. Generating text with recurrent neural networks. In: Proceedings of International Conference on Machine Learning, 2011. 1017–1024
16 Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks. In: Proceedings of Advances in Neural Information Processing Systems, 2014. 3104–3112
17 Cho K, van Merrienboer B, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: Proceedings of Conference on Empirical Methods in Natural Language Processing, 2014
18 Graves A, Mohamed A R, Hinton G. Speech recognition with deep recurrent neural networks. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2013
19 Huang Z Y, Tang J, Xue S F, et al. Speaker adaptation of RNN-BLSTM for speech recognition based on speaker code. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2016. 5305–5309
20 Li X G, Wu X H. Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2015. 4520–4524
21 Li J, Xu H, He X W, et al. Tweet modeling with LSTM recurrent neural networks for hashtag recommendation. In: Proceedings of IEEE International Joint Conference on Neural Networks, 2016. 1570–1577
22 Wang M S, Song L, Yang X K, et al. A parallel-fusion RNN-LSTM architecture for image caption generation. In: Proceedings of IEEE International Conference on Image Processing, 2016. 4448–4452
23 Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput, 1997, 9: 1735–1780
24 Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Netw, 1994, 5: 157–166
25 Xu K, Ba J, Kiros R, et al. Show, attend and tell: neural image caption generation with visual attention. In: Proceedings of International Conference on Machine Learning, 2015. 2048–2057
26 Charles A, Yin D, Rozell C. Distributed sequence memory of multidimensional inputs in recurrent networks. J Mach Learn Res, 2017, 18: 181–217
27 Wang C, Hill D J. Learning from neural control. IEEE Trans Neural Netw, 2006, 17: 130–146
28 Wang C, Hill D J. Deterministic Learning Theory for Identification, Recognition, and Control. Boca Raton: CRC Press, 2009
29 Wang M, Zhang Y W, Wang C. Learning from neural control for non-affine systems with full state constraints using command filtering. Int J Control, 2018, 47: 1–15
30 Wang C, Hill D J. Deterministic learning and rapid dynamical pattern recognition. IEEE Trans Neural Netw, 2007, 18: 617–630
31 Wang C, Chen T R. Rapid detection of small oscillation faults via deterministic learning. IEEE Trans Neural Netw, 2011, 22: 1284–1296
32 Lin P, Wang C, Chen T R. A stall warning scheme for aircraft engines with inlet distortion via deterministic learning. IEEE Trans Control Syst Technol, 2018, 26: 1468–1474
33 Lin P, Wang M, Wang C, et al. Abrupt stall detection for axial compressors with non-uniform inflow via deterministic learning. Neurocomputing, 2019, 338: 163–171
34 Chen T R, Wang C, Hill D J. Rapid oscillation fault detection and isolation for distributed systems via deterministic learning. IEEE Trans Neural Netw Learn Syst, 2014, 25: 1187–1199
35 Chen T R, Wang C, Hill D J. Small oscillation fault detection for a class of nonlinear systems with output measurements using deterministic learning. Syst Control Lett, 2015, 79: 39–46
36 Chen T R, Wang C, Chen G, et al. Small fault detection for a class of closed-loop systems via deterministic learning. IEEE Trans Cybern, 2019, 49: 897–906
37 Yuan C Z, Wang C. Design and performance analysis of deterministic learning of sampled-data nonlinear systems. Sci China Inf Sci, 2014, 57: 032201
38 Wu W M, Wang C, Yuan C Z. Deterministic learning from sampling data. Neurocomputing, 2019, 358: 456–466
39 Fradkov A L, Evans R J. Control of chaos: methods and applications in engineering. Annu Rev Control, 2005, 29: 33–56
40 Chen G R, Dong X N. From Chaos to Order: Methodologies, Perspectives and Applications. Singapore: World Scientific, 1998

Appendix A Analysis in the interval $\mathcal{I}'_k$

Let $H_i^s[k] := \bar{W}_i^{s\mathrm{T}} S(X[k]) - f_i(X[k], p^r)$, $\forall k \in \mathcal{I}_k$. The solution of the synchronization error (13) can be expressed as follows:

$$\tilde{x}_i^s[k] = b_i^{k-T_{ak}} \tilde{x}_i^s[T_{ak}] + \sum_{j=T_{ak}}^{k-1} T b_i^{k-1-j} H_i^s[j]. \tag{A1}$$

Since we do not know any information about the synchronization error in the beginning of $\mathcal{I}_k$ (i.e., in T_{ak} steps), we have to discuss three cases: (i) the magnitude of $\tilde{x}_i^s[T_{ak}]$ is small; (ii) the magnitude of $\tilde{x}_i^s[T_{ak}]$ is large and $\tilde{x}_i^s[T_{ak}]$ has the same sign with $H_i^s[k]$; (iii) the magnitude of $\tilde{x}_i^s[T_{ak}]$ is large and $\tilde{x}_i^s[T_{ak}]$ has a different sign with $H_i^s[k]$.

In order to facilitate analysis, we assume the sign of $H_i^s[k]$ is positive in follows. (The opposite situation can directly carried out by following the same proof with the negative sign.)

Case (i). If $|\tilde{x}_i^s[T_{ak}]| < \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$, the synchronization error satisfies:

$$\begin{aligned}\tilde{x}_i^s[k] &> -b_i^{k-T_{ak}}\frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}+\sum_{j=T_{ak}}^{k-1}Tb_i^{k-1-j}H_i^s[j]\\ &> -\frac{Tb_i^{k-T_{ak}}(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}+\frac{T(1-b_i^{k-T_{ak}})(\epsilon_i^*+\varsigma_i^*+\mu_i)}{1-b_i}\\ &= -\frac{Tb_i^{k-T_{ak}}(2\epsilon_i^*+2\varsigma_i^*+\frac{3\mu_i}{2})}{1-b_i}+\frac{T(\epsilon_i^*+\varsigma_i^*+\mu_i)}{1-b_i}.\end{aligned} \tag{A2}$$

Next, we have to estimate the maximum time for passing the area $|\tilde{x}_i^s| < \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$. With the property (A2), we can estimate the maximum passing time by using the following condition:

$$\begin{aligned}&-\frac{Tb_i^{k-T_{ak}}(2\epsilon_i^*+2\varsigma_i^*+\frac{3\mu_i}{2})}{1-b_i}+\frac{T(\epsilon_i^*+\varsigma_i^*+\mu_i)}{1-b_i}\geqslant\frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i},\\ &\Leftrightarrow\frac{Tb_i^{k-T_{ak}}(2\epsilon_i^*+2\varsigma_i^*+\frac{3\mu_i}{2})}{1-b_i}\leqslant\frac{T\mu_i}{2(1-b_i)}.\end{aligned} \tag{A3}$$

Based on the analysis in (A3), we can conclude that $|\tilde{x}_i^s[k]| > \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$ satisfies, if

$$k > T_{ak}+\log_{b_i}\frac{\mu_i}{4\epsilon_i^*+4\varsigma_i^*+3\mu_i}. \tag{A4}$$

In case (i), a time interval $\mathcal{I}_k' = \{k\,|\,|\tilde{x}_i^s[k]| < \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}\} = [T_a, T_{ak}']$ appears in the beginning of the time interval $\mathcal{I}_k$. The maximum time length of the interval $\mathcal{I}_k'$ is $T_{ak}' - T_{ak} + 1 \leqslant l' := \log_{b_i}\frac{\mu_i}{4\epsilon_i^*+4\varsigma_i^*+3\mu_i} + 1$. In the time interval $k \in \mathcal{I}_k - \mathcal{I}_k'$, $|\tilde{x}_i^s[k]| > \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$.

Case (ii). If $|\tilde{x}_i^s[T_{ak}]| \geqslant \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$ and $\tilde{x}_i^s[T_{ak}]$ has the same sign with $H_i^s[k]$, the synchronization error satisfies

$$\begin{aligned}\tilde{x}_i^s[k] &> b_i^{k-T_{ak}}\frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}+\sum_{j=T_{ak}}^{k-1}Tb_i^{k-1-j}H_i^s[j]\\ &> \frac{Tb_i^{k-T_{ak}}(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}+\frac{T(1-b_i^{k-T_{ak}})(\epsilon_i^*+\varsigma_i^*+\mu_i)}{1-b_i}\\ &= \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}+\frac{T(1-b_i^{k-T_{ak}})(\frac{\mu_i}{2})}{1-b_i}.\end{aligned} \tag{A5}$$

We know that $\frac{T(1-b_i^{k-T_{ak}})(\frac{\mu_i}{2})}{1-b_i}$ is always larger than 0. Thus, we have

$$\tilde{x}_i^s[k] > \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}. \tag{A6}$$

In case (ii), the time interval $\mathcal{I}_k' = \{k\,|\,|\tilde{x}_i^s[k]| < \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}\}$ will be an empty set $\mathcal{I}_k' = \emptyset$. In the time interval $k \in \mathcal{I}_k - \mathcal{I}_k'$, $|\tilde{x}_i^s[k]| > \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$.

Case (iii). If $|\tilde{x}_i^s[T_{ak}]| \geqslant \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$ and $\tilde{x}_i^s[T_{ak}]$ has a different sign with $H_i^s[k]$, there exists a time interval $\mathcal{I}_k' = \{k\,|\,|\tilde{x}_i^s[k]| < \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}\} = [T_{ak}', T_{bk}']$ in $\mathcal{I}_k$ such that

$$|\tilde{x}_i^s[k]| \geqslant \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}, \quad \forall k \in [T_{ak}, T_{ak}'-1], \tag{A7}$$

$$|\tilde{x}_i^s[k]| \leqslant \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}, \quad \forall k \in [T_{ak}', T_{bk}'], \tag{A8}$$

$$|\tilde{x}_i^s[k]| \geqslant \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}, \quad \forall k \in [T_{bk}'+1, T_{bk}]. \tag{A9}$$

In case (iii), we can conclude that the time length of $\mathcal{I}_k'$ is $T_{bk}' - T_{ak}' + 1 = l' = \log_{b_i}\frac{\mu_i}{4\epsilon_i^*+4\varsigma_i^*+3\mu_i} + 1$ according to the analysis of case (i). In the time interval $k \in \mathcal{I}_k - \mathcal{I}_k'$, $|\tilde{x}_i^s[k]| > \frac{T(\epsilon_i^*+\varsigma_i^*+\frac{\mu_i}{2})}{1-b_i}$ holds.