Journal of Process Control 54 (2017) 81–89

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Dual least squares support vector machines based spatiotemporal modeling for nonlinear distributed thermal processes

Kang-Kang Xu [a,b], Han-Xiong Li [b,*], Hai-Dong Yang [c,d]

[a] College of Mechanical and Electrical Engineering, Central South University, Changsha, China
[b] Department of Systems Engineering and Engineering Management City University of Hong Kong, Hong Kong, China
[c] School of Mechanical Science & Engineering, Huazhong University of Science and Technology, China
[d] Guangdong Engineering Research Center for Green Manufacturing & Energy Efficiency Optimization, Guangdong University of Technology, Guangzhou 510006, China

ARTICLE INFO

Article history:
Received 3 May 2016
Received in revised form 4 November 2016
Accepted 10 March 2017

Keywords:
System modeling
Karhunen-Loève decomposition
Least squares support vector machines
Distributed parameter system
Rademacher complexity

ABSTRACT

In this paper, a dual least squares support vector machines (LS-SVM) is proposed to model the thermal process. The infinite-dimensional system is first transformed into a finite-dimensional system through space-time separation. Then, the dual LS-SVM model is to approximate the two nonlinearities embedded in the system. Through space-time synthesis, the dual LS-SVM based spatiotemporal model is able to approximate the complex DPS with inherent coupled nonlinearities. The generalization performance of the proposed model is discussed using Rademacher complexity. Finally, simulations on a curing process demonstrate the effectiveness of the proposed modeling method.

1. Introduction

Many industrial processes, such as, semiconductor manufacturing, lithium ion battery and curing process, belong to nonlinear distributed parameter systems (DPSs) [1,2], which are described in partial differential equations (PDEs) with boundary and initial conditions. These kinds of systems have strong spatiotemporal dynamics that are extremely difficult to model and control [3]. A finite-dimensional model that could approximate the spatiotemporal dynamics of DPS is desirable in engineering applications [4]. In recent years, quite lots of research are reported on modeling and control of DPSs [5–8].

The space-time separation based methods, using spatial basis function (BF) expansion, have been widely applied for model reduction of the parabolic type DPSs because of the slow/fast separation property [1,2]. After the spatial BFs are selected, the finite-dimensional model can be derived for the temporal dynamics by using the Galerkin method [9] or traditional data-based identification methods [10,11].

There are many discretization method used to obtain the ordinary differential equation (ODE) model. Such as: finite difference method (FDM) [12] and finite element method (FEM) [13], etc. These methods using local spatial BFs can be easily applied to transform the infinite dimensional model into the finite dimensional temporal model. However, these methods will often lead to a high-order temporal model for a good model approximation, which may not be suitable for controller design. Though the spectral method [14,15] using global spatial BFs is able to obtain the analytical model, it requires the system to have homogeneous boundary conditions. In summary, all these methods require the system to be known.

When the DPS of the thermal processes is unknown in most of real-world applications, the modeling has to rely on the data-based approach. The Karhunen–Loève (KL) method [16–19], which is also called proper orthogonal decomposition (POD) or principal component analysis (PCA), is widely used to obtain principal empirical eigenfunctions (EEFs) from the experimental data. KL method is not a discretization method. It is an algorithm enabling a stochastic field to be represented with a minimum number of degree of freedom. Compared with the above discretization methods, KL method can obtain low-order ODE model efficiently without sacrifice the model precision. When the proper spatial BFs are learned from data, the low-order temporal model of DPS can be easily constructed using the existing methods, such as the least squares support vector machines (LS-SVM) [20], the neural network [21], Volterra [22], fuzzy model [23] and block-oriented models [24,25].

* Corresponding author.
E-mail addresses: mehxli@cityu.edu.hk, 632461502@qq.com (H.-X. Li).

http://dx.doi.org/10.1016/j.jprocont.2017.03.006
0959-1524/

It is well-known that LS-SVM can approximate any nonlinear function, and SVM combination with KL method has ever been extended to DPS modeling by Qi [26]. However, this method is not designed for the process with two coupled nonlinear dynamics, and it is only suit for the general DPS process. Many of thermal processes, like the example discussed in this paper, have complex nonlinearities that consist of two coupled nonlinear functions with respect to the inputs and process output. It is well known that the modeling performance highly depends on how much the model structure matches the process. So our paper is to develop a dual model structure for the widely existing process that contains two inherently coupled nonlinearities. Under this model structure, different methods like SVM, LS-SVM, NN, etc can all be applied. This paper will focus on the LS-SVM method. As the model structure matches well with the process, it can achieve better performance. Once the spatiotemporal model is estimated, Rademacher complexity is developed to find the generalization error bounds [27,28]. Rademacher complexity has been proved to be an efficient method in the analysis of many learning problems like pattern classification and regression. However, this method has not been used in DPS.

In this study, a Dual LS-SVM modeling approach is proposed for the nonlinear distributed thermal processes with inherent coupled nonlinear dynamics. Under the space-time separation with KL decomposition, the temporal dynamics will be estimated using two LS-SVMs. After the spatiotemporal model been acquired, the generalization performance of the model is discussed using Rademacher complexity. The simulation experiment on the selected cure oven [15,22] will be carried out for further validation and the comparison study.

2. Problem description

Many of industrial thermal processes can be expressed in the general form as follows:

$$\rho c\frac{\partial T}{\partial t}=\frac{\partial}{\partial x}(\underline{k}\frac{\partial T}{\partial x})+\frac{\partial}{\partial y}(\underline{k}\frac{\partial T}{\partial y})+\frac{\partial}{\partial z}(\underline{k}\frac{\partial T}{\partial z})+f_c(T)+f_r(T)+\rho Q \quad (1)$$

Subject to the Neumann boundary conditions:

$$\begin{aligned}&\frac{\partial T}{\partial x}|_{x=0}=0 \quad \frac{\partial T}{\partial x}|_{x=x_0}=0\\&\frac{\partial T}{\partial y}|_{y=0}=0 \quad \frac{\partial T}{\partial y}|_{y=y_0}=0\\&\frac{\partial T}{\partial z}|_{z=0}=0 \quad \frac{\partial T}{\partial z}|_{z=z_0}=0\end{aligned} \quad (2)$$

where $T=T(x, y, z, t)$ denotes the temperature (°C) at time t and location (x, y, z), $x\in[0, x_0]$, $y\in[0, y_0]$ and $z\in[0, z_0]$ are spatial coordinates, c is the specific heat coefficient (J/kg°C), $f_c(T)$ and $f_r(T)$are unknown nonlinear effects of convection and radiation, respectively. $Q=Q(x, y, z, t)$ is the heating source, ρ and $\underline{k}$ are the density (kg/m^3) and the thermal conductivity (W/m°C) respectively.

The thermal conductivity $\underline{k}$ and the density ρ of Eq. (1) are dependent on the temperature, and can be expressed as

$$\underline{k}=k_0+\bar{k}(T),\ \rho=\frac{\rho_0}{1+\bar{\rho}(T)},$$

where: k_0 and ρ_0 are nominal values around the working point, $\bar{k}(T)$ and $\bar{\rho}(T)$ are functions of $T(x, y, z, t)$.

Thus, Eq. (1) can be rewritten to the follow form:

$$\frac{\partial T}{\partial t}=k_1\nabla^2 T+F(T)+\frac{1}{c}Q \quad (3)$$

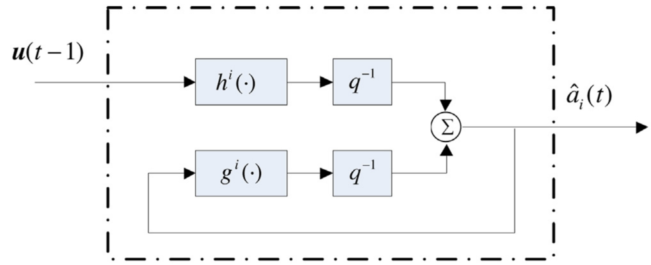

Fig. 1. Structure of the ODE model $a_i(t)$.

where: $\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$ is the Laplacian operator, $k_1=k_0/\rho_0 c$ is a constant,

$$F(T)=\frac{k_0\bar{\rho}(T)}{\rho_0 c}\nabla^2 T+\frac{1+\bar{\rho}(T)}{\rho_0 c}(\bar{k}(T)\nabla^2 T+\frac{\partial\bar{k}(T)}{\partial x}\frac{\partial T}{\partial x}+\frac{\partial\bar{k}(T)}{\partial y}\frac{\partial T}{\partial y}+\frac{\partial\bar{k}(T)}{\partial z}\frac{\partial T}{\partial z}+f_c(T)+f_r(T))$$

is an unknown nonlinear function with respect to T. It is obvious that the right hand of Eq. (3) has two nonlinear function $F(\cdot)$ and $Q(\cdot)$, where $Q(\cdot)$ is a nonlinear function about $\boldsymbol{u}(t)$.

The model Eq. (3) described in PDE cannot be used directly for online estimation and control due to its infinite-dimensional property. A finite-dimensional ODE model is usually needed for practical applications. With the help of KL method (see in Appendix A), the spatiotemporal measurements $T(x, y, z, t)$ can be decoupled into the following equations using a set of spatial BFs $\left\{\phi_i(x, y, z)\right\}_{i=1}^{\infty}$. Due to the slow/fast properties of the parabolic system (1), the slow modes of the dynamics will be remained in the following expression.

$$T_n(x, y, z, t)=\sum_{i=1}^{n}\phi_i(x, y, z)\cdot a_i(t) \quad (4)$$

$$Q_n(x, y, z, t)=\sum_{i=1}^{n}\phi_i(x, y, z)\cdot b_i(t) \quad (5)$$

where $a_i(t)$ is the ODE model of the spatiotemporal model (1), n is the order of the ODE model that can be estimated using Eq. (A.12), $b_i(t)$ is a nonlinear function of the manipulated inputs $\boldsymbol{u}(t)=[u_1(t), u_2(t)\cdots u_{n_u}(t)]^T$ with $n_u\in R$. As the derivation in Appendix B, $a_i(t)$ can be expressed as follows:

$$a_i(t)=g^i(a_i(t-1))+h^i(\boldsymbol{u}(t-1)) \quad (6)$$

The structure of $a_i(t)$ can be described in Fig. 1, where q is the forward operator. It is obvious that the ODE model $a_i(t)$ is composed of two nonlinear block $g^i(\cdot)$ and $h^i(\cdot)$ with respect to the manipulated inputs and the process outputs. In order to model this complex dynamics efficiently, the model structure should be designed to have the similar configuration to handle these two nonlinear modules.

According to the above analysis, the original system (1) is a nonlinear infinite-dimensional model. To model this system, KL will be first used to transform the infinite-dimensional model into a finite-dimensional model, where the model order n can be estimated using Eq. (A.12). Then a novel dual-model structure will be specially designed for DPS that contains two inherently coupled nonlinearities (Eq. (6)). Under this model structure, different methods like SVM, LS-SVM, NN, etc all can be applied. This paper will use LS-SVM to model the dual-model structure.