Control Engineering Practice 16 (2008) 1427–1437

Contents lists available at ScienceDirect

Control Engineering Practice

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

Online monitoring of nonlinear multiple mode processes based on adaptive local model approach

Zhiqiang Ge, Zhihuan Song *

State Key Laboratory of Industrial Control Technology, Institute of Industrial Process Control, Zhejiang University, Hangzhou 310027, Zhejiang, China

ARTICLE INFO

Article history:
Received 24 September 2007
Accepted 3 April 2008
Available online 6 June 2008

Keywords:
Online process monitoring
Multiple mode
Least square support vector regression
Local model approach
Two-step information extraction strategy

ABSTRACT

A new adaptive local model based monitoring approach is proposed for online monitoring of nonlinear multiple mode processes with non-Gaussian information. To solve the multiple mode problem, just-in-time-learning (JITL) strategy is introduced. The local least squares support vector regression (LSSVR) model is built on the relevant dataset for prediction. To satisfy the online modeling demand, the real-time problem is considered. Then a two-step independent component analysis–principal component analysis (ICA–PCA) information extraction strategy is introduced to analyze residuals between the real output and the predicted one. Two case studies show that the new proposed method gives better performance compared to conventional methods.

1. Introduction

Multivariate statistical process control (MSPC) schemes such as principal component analysis (PCA) and partial least squares (PLS) are widely used. There were representative and recent research efforts on continuous, batch, dynamic, multiscale and plant-wide processes (AlGhazzawi & Lennox, 2008; Bakshi, 1998; Kresta, MacGregor, & Marlin, 1991; Ku, Storer, & Georgakis, 1995; Narasimhan & Shah, 2008; Nomikos & MacGregor, 1994, 1995a, 1995b; Thornhill & Horch, 2007). However, traditional MSPC schemes are formulated based on the assumption that the process variables are sampling independent, Gaussian distributed and linearly correlated. However, real industrial processes are always nonlinear and run under multiple operating modes, because of product changes and set-point changes, among others. Furthermore, some process variables may not follow Gaussian distribution. Therefore, it is necessary to develop new methods to address these problems.

To overcome the shortcomings of MSPC schemes for nonlinear processes, several nonlinear extensions of PCA have been reported. Kramer (1992) developed a nonlinear PCA method based on the autoassociative neural network. Dong and McAvoy (1996) proposed a nonlinear PCA by combining the principal curve and the neural network. Hiden, Willis, Tham, and Montague (1999) used genetic programming to address the same problem. However, neural network suffers from drawbacks such as complex training strategy. Recently, a nonlinear process monitoring method based on kernel PCA (KPCA) was proposed (Cho, Lee, Choi, Lee, & Lee, 2005; Choi, Lee, Lee, Park, & Lee, 2005; Lee, Yoo, Choi, Vanrolleghem, & Lee, 2004). The main advantage of the KPCA method over other nonlinear PCA approaches is that no nonlinear optimization needs to be involved. Kruger, Antory, Hahn, Irwin, and McCullough (2005) proposed a new nonlinearity measure for principal component and discussed the criteria on the selection of linear or nonlinear PCA for a specific process. Alternatively, Cheng and Chiu (2005) used just-in-time-learning (JITL) model with finite impulse response (FIR) structure to address the same problem. Although the adaptive FIR model can capture the dynamic of a nonlinear process, process nonlinearity might not be adequately modeled with a small number of training samples. Support vector regression (SVR), which only needs a small number of training samples, has been widely used for nonlinear function regression and system identification since the last decade (Vapnik, 1995). Recently, least squares support vector regression (LSSVR) was developed (Suykens, Van Gestel, De Brabanter, De Moor, & Vandewalle, 2002). Because of its improved computational efficiency compared to SVM, the LSSVR model can be updated much faster, which enables online process modeling.

Motivated by the fact that MSPC is not efficient in multiple mode processes, adaptive PCA and PLS methods were developed (Dayal & MacGregor, 1997; Li, Yue, Valle-Cervantes, & Qin, 2000; Qin, 1998; Wang, Kruger, & Irwin, 2005; Wang, Kruger, & Lennox, 2003). Li et al. (2000) presented a monitoring strategy, which built a recursive PCA (RPCA) model to update the monitoring model. To avoid fault accommodation, Wang et al. (2005) proposed an N-step-ahead horizon strategy for fault monitoring. As an alternative, model library based methods have also been introduced in which predefined models match their corresponding

* Corresponding author. Tel.: +86 571 87951442 808; fax: +86 571 87951921.
E-mail address: zhsong@iipc.zju.edu.cn (Z. Song).

0967-0661/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.conengprac.2008.04.004

modes (Chen & Liu, 2000; Hwang & Han, 1999; Zhao, Zhang, & Xu, 2004). However, the transition between two operating modes was always falsely alarmed. To handle this issue, several research papers have been published, including Bhagwat, Srinivasan, and Krishnaswamy (2003a, 2003b) and Srinivasan, Wang, Ho, and Lim (2004). However, they are either model-based or too complicated to be implemented in practice. An overview of the transition process monitoring was given by Kourti (2005). Recently, process knowledge was incorporated into MSPC for time-varying process monitoring (Jin, Lee, Lee, & Han, 2006; Lee, Jin, & Han, 2006). Unfortunately, process knowledge is difficult to obtain from modern complex processes.

In this paper, a new local model approach for monitoring multiple mode and time-varying processes is proposed. The problem can be successfully solved through online updating of the local model. In the new method, a history database which contains normal process data is needed. The LSSVR model is used to extract the nonlinear information from the raw process data. Hence, the resulting residual between the actual process output and the predicted output from the local model is no longer sensitive to process nonlinearity. In the next step, the process output and the predicted one are used in the MSPC analysis to draw monitoring charts.

Independent component analysis (ICA) is an emerging technique for finding several independent variables as linear combinations of measured variables. It can reveal more meaningful information in non-Gaussian data than PCA. Several applications of ICA have been reported (Hyvarinen & Oja, 2000). Lee, Yoo, and Lee (2004a, 2004b) used ICA for process monitoring. Kano, Tanaka, Hasebe, Hashimoto, and Ohno (2004) developed a unified framework for MSPC, which combined PCA-based SPC and ICA-based SPC. A two-step information extraction strategy based on ICA–PCA was proposed by Ge and Song (2007).

The study reported in this paper focuses on developing a new adaptive monitoring scheme for nonlinear and multiple mode processes with non-Gaussian information. This paper focuses on the situation that multiple operating modes are driven by process inputs. First, the LSSVR model is built for online modeling of the nonlinear process. Second, a local modeling strategy is proposed to handle the multiple mode behavior. Finally, a two-step information extraction and monitoring strategy based on ICA–PCA is used for online monitoring.

The remainder of this paper is structured as follows. In Section 2, the local model strategy is described. The real-time modeling problem is considered in Section 3. Section 4 presents the two-step ICA–PCA information extraction and process monitoring strategy. Two case studies are given in Section 5. Finally, some conclusions are made.

2. Local modeling strategy based on JITL and LSSVR

Conventional data-based modeling methods focused on global approaches, such as neural networks, fuzzy set methods, and other kinds of parametric models. However, when dealing with large data sets from industrial processes, traditional approaches become difficult in terms of efficient model structure determination and optimization problem formulation. Furthermore, such models are difficult to be updated on-line when the process-operating mode changes. On the other hand, the idea of local modeling is an approach that represents a nonlinear system with a set of simple local models valid in certain operating regions. There are several well-known examples of the local modeling approach, such as neural-fuzzy network and the T–S fuzzy model. However, most local modeling approaches suffer from the drawback of requiring *a priori* knowledge to determine the partition of operating space. When this information is not available, a complicated training strategy is needed to determine both optimal model structure and parameters of the local model.

To alleviate the aforementioned problem, JITL strategy was recently developed (Bontempi, Bersini, & Birattari, 2001; Cheng & Chiu, 2004, 2005) for modeling nonlinear processes. The approach is based on the ideas of local modeling and database technology. JITL assumes that all available normal observations are stored in the database, and local models are built dynamically on query. Compared to traditional modeling methods, JITL exhibits three main characteristics. First, the model development is based on the current data sample. Secondly, the predicted output of the current data sample is computed by the local model. Finally, the constructed local model is discarded after the predicted output is obtained. Fig. 1 illustrates the difference between traditional methods and JITL (Cheng & Chiu, 2005). Traditional methods typically train the model offline, while JITL can be carried out online. When the process changes frequently, the traditional method not only is time-consuming but also will interrupt the plant operation. On the other hand, JITL shows more flexibility. Therefore, it can be used for online modeling.

There are three main steps in JITL to predict the model output corresponding to the current data sample (Cheng & Chiu, 2005): (1) relevant data samples that match the current data sample are searched in the database by using some nearest neighborhood criteria, (2) a local model is built based on the relevant dataset, and (3) the model output is calculated based on the local model and the current data sample. The local model is then discarded after the predicted output of the current data sample is obtained. When the next data sample comes, a new local model will be built based on the aforementioned procedures. In the previous work on JITL modeling (Cheng & Chiu, 2005), an ARX/FIR model is commonly used as the local model structure. However, LSSVR is more appropriate because of the on-line modeling demand and its low computational cost. LSSVR is a novel machine-learning tool and is especially useful for classification and prediction with small sample cases. The formulation of LSSVR is omitted in the present paper. Details of the algorithm can be found in Suykens et al. (2002).

Having discussed the local model structure, we proceed to describe the JITL algorithm. Suppose a dataset $\{\mathbf{u}_i, \mathbf{y}_i\}_{i=1,2,\ldots,N}$ is collected as the database of JITL, with input data $\mathbf{u}_i \in \mathbf{R}^n$ and output data $\mathbf{y}_i \in \mathbf{R}^n$. For a current sample data $\mathbf{u}_c$, the objective of JITL is to predict the model output $\hat{\mathbf{y}}_c = f(\mathbf{u}_c)$ according to the known database $\{\mathbf{u}_i, \mathbf{y}_i\}_{i=1,2,\ldots,N}$. For the development of the local LSSVR model, the relevant data should first be selected from the database. In the previous work, the distance measure, Euclidean norm $d(\mathbf{u}_c, \mathbf{u}_i) = ||\mathbf{u}_c, \mathbf{u}_i||_2$, is employed to evaluate the similarity between $\mathbf{u}_c$ and $\mathbf{u}_i$. Recently, a new similarity measure (Cheng & Chiu, 2004), which integrates both distance measure and angle measure, was proposed to improve the

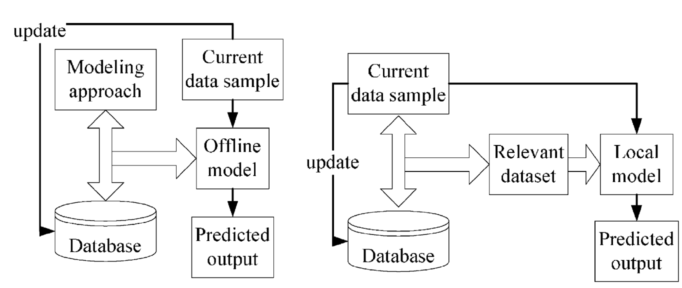

Fig. 1. Comparison between traditional modeling methods and JITL.