Chemometrics and Intelligent Laboratory Systems 158 (2016) 31–40

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Nonlinear and robust statistical process monitoring based on variant autoencoders

Weiwu Yan [a,*], Pengju Guo [a], Liang gong [a], Zukui Li [b]

[a] Automation Department, Shanghai Jiao Tong University, Shanghai 200240, China
[b] Chemical and Materials Engineering, University of Alberta, Edmonton, AB T6G2V4, Canada

ARTICLE INFO

Article history:
Received 22 May 2016
Received in revised form 11 August 2016
Accepted 13 August 2016
Available online 15 August 2016

Keywords:
Process monitoring
Variant autoencoders
Control limits
Kernel density estimation

ABSTRACT

Autoencoders (AEs) are an effective means for nonlinear feature extraction and dimension reduction. Variant autoencoders are an improvement over traditional AEs in terms of robustness. This paper proposes a novel nonlinear and robust process-monitoring approach based on variant autoencoders (variant AEs), which include denoising autoencoders (DAE) and contractive autoencoders (CAE). The CAE and DAE are powerful for extracting robust and nonlinear feature representations or manifold structures underlying data from industrial processes. Next, an online monitoring model is built through constructing new test statistic H^2 based on the robust feature representations. The control limits are determined by kernel density estimation. The proposed method was applied to the Tennessee Eastman process (TE process) to evaluate its monitoring performance, and it demonstrated outstanding process-monitoring performance through the experimental results, especially for the barely detectable faults, such as 3, 5, 9, 10, 11, 15, 19, 20 and 21. Variant AEs monitoring provides a simple and very effective process-monitoring method for industrial processes.

1. Introduction

With the rapid development of modern industry, industrial processes have experienced an increase in the scale and complexity. The requirement for the reliability and security of industrial processes is growing significantly. Online process monitoring and fault diagnosis are key factors that ensure product quality and operational safety. Process-monitoring approaches are roughly classified into two categories, that is, model-driven approaches and data-driven approaches. Model-driven approaches require an accurate mathematical model of the system or plant [1]. Because it is often difficult to obtain such an accurate mathematical model for complex industrial processes, data-based approaches, which are based on historic process data [2], rather than model-based approaches have been widely used for process monitoring.

Data-based process monitoring is typically accomplished by comparing the actual behavior of the process with a model representing normal or desirable process behavior. The detection of process faults is based on monitoring of the deviation between the actual process behavior and that predicted by the model, with a fault condition flagged when these deviations exceed certain predetermined limits [24]. The data-driven process monitoring is usually accomplished using conventional multivariate statistical methods, typically unsupervised learning, such as PCA (principal component analysis), PLS (partial least squares), ICA (independent component analysis), FDA (Fisher discriminant analysis), ForeCA (Forecastable Component Analysis), etc. [3–7,24–26].

The core issue of these data-based process monitoring schemes is to extract robust feature representations or manifold structure from industrial process data based on which a statistic is constructed in principal component subspace or established in residual subspace. PCA is the most widely used process monitoring method; it extracts a few principal component variables from a number of highly relevant process variables. Through a linear space transformation, most of the characteristics of the original variables are reserved and the correlation between variables is removed. PLS [4] is fairly similar to PCA. It also requires that original data satisfy a Gaussian distribution. Compared with PCA, the principal components in ICA [5] are statistically independent and contain higher-order statistical information between variables. FDA [6] determines a series of linear transform directions by maximizing discrete degree between classes and minimizing discrete degree within class. However, it is difficult to find an appropriate linear transform direction in which a maximum classification is available when the relationship between process variables is nonlinear. ForeCA is a feature-extraction method for multivariate time series [7]. Unfortunately, these linear monitoring methods often result in high rates of false positives in process monitoring because of the nonlinearity in real industrial processes.

Nonlinear methods have been proposed to address the problem of nonlinearity in process monitoring. Nonlinear principal component analysis (NLPCA) is commonly considered as a nonlinear generalization

* Corresponding author.
E-mail address: yanwwsjtu@sjtu.edu.cn (W. Yan).

http://dx.doi.org/10.1016/j.chemolab.2016.08.007
0169-7439/© 2016 Published by Elsevier B.V.

of standard PCA. Gnanadesikan [8] proposed a general PCA method in which the original vector is extended to a new vector by a nonlinear mapping, after which linear principal component analysis is applied to the new vector. Kramer [9] proposed an autoassociative neural network with five layers. However, it is very difficult to determine an effective middle layer or the number of neurons of the inverse mapping layer. Hastie and Stuetzle [10] proposed a principal curve method, in which a smooth curve is produced by the nonlinear feature extracted from the sample data. In recent years, the kernel methods have seen increasing application to industrial processes to address this issue of nonlinearity. Schölkopf [11] proposed kernel principal component analysis (Kernel PCA), based on which Lee [12] proposed a nonlinear performance-monitoring method. Its main idea is that the data from low-dimensional input space is mapped to high-dimensional feature space via nonlinear kernel function, and then the principal component analysis is applied in the high-dimensional feature space. Zhang et al. [14] proposed a monitoring method based on kernel FDA. Widodo applied KICA to fault diagnosis of induction motors [20]. Nonetheless, kernel-based algorithms require considerable time to build the kernel matrix and train the monitoring model. These methods mainly consider the nonlinearity in process monitoring and lack robustness.

Currently, robustness and nonlinearity are the two key issues in process monitoring because of the complexity of industrial processes. To address the problems mentioned above, this paper proposed novel process-monitoring approaches based on variant autoencoders. Autoencoders (AE) are artificial neural networks used for representing (encoding) a set of data, typically for the purpose of dimensionality reduction. AE is equivalent to PCA when the active function is linear and the number of hidden layers is less than that of the input layers. Thus, AE is regarded as a generalization of PCA. Kramer [9] proposed a process monitoring method based on autoassociative. However, AE sometimes has poor robustness. Recently, variant autoencoders (variant AE) have become more widely used for learning generative models of data. Variant AE, such as denoising autoencoders (DAE) and contractive autoencoders (CAE), are considered an improvement in terms of robustness over traditional AEs and are able to extract robust hidden representations of process data. DAE is an outstanding technique with attractive application to robust feature extraction [21]. CAE explicitly supports the robustness of hidden representations by adding the Jacobean term of hidden representations to the loss function [13]. DAE and CAE also perform well in approximating multivariable nonlinear and complex functions. This paper employs DAE and CAE to extract robust and nonlinear feature representations of process data. Next, two statistics are constructed and the control limits of the statistics are estimated by the kernel density estimation method.

In this paper, we discuss novel nonlinear and robust process monitoring approaches based on variant AEs. Variant AEs monitoring models include two monitoring models which are denoising autoencoders model and contractive autoencoders model respectively. The CAE and DAE are employed to extract robust and nonlinear feature representations or manifold structures underlying data from industrial processes. A new test statistics, that is statistic H^2 , is proposed and constructed based on the robust feature representations to detect fault information. The proposed methods were applied to the Tennessee Eastman process (TE process) to evaluate its monitoring performance.

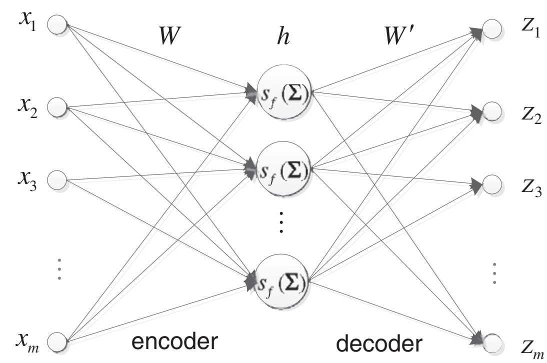

Fig. 2. Architecture of the AEs neural network.

2. Preliminary

2.1. Principal component analysis

PCA [14] is a multivariate statistical method which transforms multiple related variables into a lesser number of unrelated variables. It transforms high-dimensional variable space into low-dimensional variable space while retaining most of the information in the original data. In PCA-based process monitoring, the number of principal components is first determined by the cumulative contribution rate method. Then, the principal component model is set up based on the loading vector and the control limits which are calculated through normal process data. Finally, the value of statistics is calculated from test data, through the principal component model.

Considering a sample matrix X which consists of n sample points and m variables, the principal matrix T is the linear combination of X, that is, $T = XW$ where $W = (w_1, w_2, \ldots, w_m)$ is the loading matrix, w_i the i_{th} loading vector, $T = (t_1, t_2, \ldots, t_m)$, and t_i the i_{th} principal vector. X is decomposed of the sum of the outer product between t_i and w_i, as in

$$X = TW^T = \sum_{i=1}^{m} t_i {w_i}^T \quad (1)$$

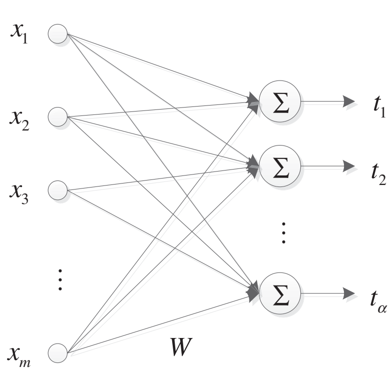

Fig. 1. Architecture of the PCA Neural network.

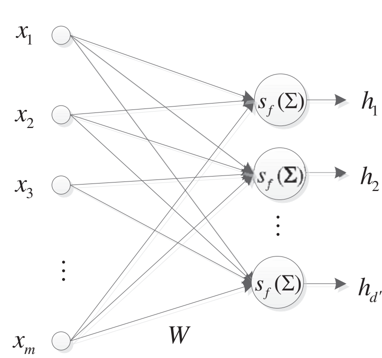

Fig. 3. Architecture of the encoder.