Journal of Process Control 37 (2016) 46–57

Contents lists available at ScienceDirect

Journal of Process Control

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

A novel process monitoring and fault detection approach based on statistics locality preserving projections

He Fei*, Xu Jinwu

Collaborative Innovation Center of Steel Technology, National Engineering Research Center of Flat Rolling Equipment, University of Science and Technology Beijing, Beijing 100083, China

ARTICLE INFO

Article history:
Received 15 October 2014
Received in revised form 15 October 2015
Accepted 7 November 2015
Available online 29 November 2015

Keywords:
Statistics pattern analysis
Locality preserving projections
Process monitoring
Parallel analysis
Fault detection

ABSTRACT

Data-driven fault detection technique has exhibited its wide applications in industrial process monitoring. However, how to extract the local and non-Gaussian features effectively is still an open problem. In this paper, statistics locality preserving projections (SLPP) is proposed to extract the local and non-Gaussian features. Firstly, statistics pattern analysis (SPA) is applied to construct process statistics and grasp the non-Gaussian statistical property using high order statistics. Then, locality preserving projections (LPP) method is used to discover local manifold structure of the statistics. In essence, LPP tries to map the close points in the original space to close in the low-dimensional space. Lastly, T^2 and squared prediction error (SPE) charts of SLPP model are used to detect process faults. One simple simulated system and the Tennessee Eastman process show that the proposed SLPP method is more effective than principal component analysis, LPP and statistics principal component analysis in fault detection performance.

1. Introduction

With the increased demands on plant safety and quality consistency, fault detection technique has been one of the most fascinating topics in industrial process control field. In the past years several fault detection methods have been developed which can be classified into three major categories: analytical methods, knowledge-based methods and data-driven methods. Large amounts of plant data are collected and stored by computer control systems in modern industrial processes. Traditionally, statistical process control (SPC) charts such as Shewhart, cumulative sum (CUSUM) and exponentially weighted moving average (EWMA) charts have been used to monitor processes and improve product quality. However, there are two or more related quality-process characteristics in need of simultaneous monitoring and control in most industry systems. Monitoring these process characteristics independently can be very misleading [1]. This motivates the development of the multivariate statistical process monitoring (MSPM) technique. Some well-known data driven fault diagnosis methods have been developed including principal component analysis (PCA) [2], partial least squares (PLS) [3], canonical variate analysis (CVA) [4] and independent component analysis (ICA) [5]. Many extended methods have been developed including dynamic PCA, neural PCA [6], multiblock and multiway PLS [7], kernel PLS [8], kernel PCA [9], multiscale KPCA and KPLS [10]. Although PCA, PLS and their extended methods have been widely used in the fault detection field, these methods are based on the global structure analysis of the process data.

Recently, some manifold learning methods such as locality preserving projection (LPP) [11], locally linear embedding (LLE) [12], Laplacian eigenmaps (LE) and local tangent space alignment (LTSA) [13] have been presented to analyze the local manifold structure of data set. Among these methods, locality preserving projection (LPP) is a well-known local structure analysis method because it can be simply applied to new data. LPP aims at preserving the neighborhood structure of the data set, while PCA only retains the most variance of the original data. LPP shares many of the properties of nonlinear techniques such as LLE and LE. Hence LPP can reveal the intrinsic geometrical structure of the observed data and find more meaningful low dimensional information hidden in the high-dimensional observations than PCA. LPP has been introduced for process monitoring and fault detection by some researchers. Hu et al. [14] firstly proposed multi-way LPP (MLPP) method for batch process fault detection. Yu [15] utilized LPP for bearing performance degradation assessment. Zhang et al. [16] combined LPP and PCA for process monitoring. In these studies, LPP has been thought as a promising fault detection method. The kernel LPP process monitoring algorithms [17,18]

* Corresponding author. Tel.: +86 10 62334255; fax: +86 10 62334255.
E-mail address: hefei@ustb.edu.cn (F. He).

http://dx.doi.org/10.1016/j.jprocont.2015.11.004
0959-1524/

are also introduced to complete nonlinear industrial process monitoring.

According to the non-Gaussian distribution of the variables besides the local characteristics of the process, in this paper we propose a novel LPP method called statistics LPP (SLPP) for effective process monitoring. Firstly statistics pattern analysis based on a moving window is used to construct process statistics: first order, second order and higher order statistics. The higher order statistics can be used to describe the non-Gaussian distribution. Then we run the LPP on process statistics to reveal the intrinsic geometrical structure of the observed data. The T^2 and SPE charts that are calculated from the normal process are used to fault detection.

The major difference between the traditional PCA/LPP and the SPA fault detection methods is that PCA/LPP monitors the process variables while SPA monitors various statistics of the process variables. The different statistics that capture the different characteristics of the process can be selected to build the model for normal process operation, and various higher-order statistics can be utilized explicitly [19]. In fact, the process statistics under abnormal conditions would show some deviation from the distribution of the process statistics under normal operation.

The concepts of statistics pattern analysis (SPA), locality preserving projection and inference procedure of the SLPP are introduced in Section 2. In Section 3, the performance of SLPP in process monitoring and fault detection is compared with PCA, LPP, and SPCA using simulated experiments, and the Tennessee Eastman (TE) process. Finally, conclusions are drawn in Section 4.

2. Theory

The process monitoring methods in common use, such as PCA, PLS, CVA, etc., are directly built based on original measured variables and do not make use of higher-order statistics for non-Gaussian distribution dataset. The statistics pattern analysis framework was proposed by Wang and He [19]. For a continuous process, an SPA is a collection of various statistics calculated from a window (or a segment) of the process measurements. These statistics capture the characteristics of each individual variable (such as mean, variance, and skewness), the interactions among different variables (such as correlation), as well as process dynamics (such as autocorrelation and cross-correlation). Note that, for different processes, different statistics can be selected to capture the dominant process characteristics such as dynamics and nonlinearity. Then principal component analysis is used to process monitoring, which is called statistics principal component analysis (SPCA). In this work, we apply locality preserving projections to quantify the dissimilarities among the training statistics, which is called statistics locality preserving projections (SLPP) and define two detection indices similar to Hotelling's T^2 (D) and SPE (Q).

2.1. Statistics pattern analysis

If $\mathbf{X}$ denotes the original measured variables, then $\mathbf{X}_k$ is given to denote a window of process measurements, as shown below:

$$\mathbf{X}_k = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m] = \begin{bmatrix} x_1(k-w+1) & x_2(k-w+1) & \cdots & x_m(k-w+1) \\ x_1(k-w+2) & x_2(k-w+2) & \cdots & x_m(k-w+2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(k) & x_2(k) & \cdots & x_m(k) \end{bmatrix} \quad (1)$$

where w is the window width and k is the time index. Generally, three groups of process statistics: first order, second order and high order statistics are defined, which are expressed:

$$\mathbf{S} = [\boldsymbol{\mu} \,|\mathbf{V}|\, \Xi] \quad (2)$$

$\boldsymbol{\mu}$ in Eq. (2) denotes the first order statistics, i.e., variable means (μ_i), whose elements are calculated from the data contained in the window:

$$\mu_i = \frac{1}{w}\sum_{l=0}^{w-1} x_i(k-l) \quad (3)$$

$\mathbf{V}$ in Eq. (2) denotes the second order statistics, which include variance (v_i), correlation ($r_{i,j}$), autocorrelation (r_i^d), and cross-correlation ($r_{i,j}^d$) of different variables.

$$v_i = \frac{1}{w}\sum_{l=0}^{w-1} [x_i(k-l)-\mu_i]^2 \quad (4)$$

$$r_{i,j} = \frac{1}{w}\frac{\sum_{l=0}^{w-1}[x_i(k-l)-\mu_i][x_j(k-l)-\mu_j]}{\sqrt{v_i v_j}} \quad (5)$$

$$r_i^d = \frac{1}{w-d}\frac{\sum_{l=d}^{w-1}[x_i(k-l)-\mu_i][x_i(k+d-l)-\mu_i]}{v_i} \quad (6)$$

$$r_{i,j}^d = \frac{1}{w-d}\frac{\sum_{l=d}^{w-1}[x_i(k-l)-\mu_i][x_j(k+d-l)-\mu_j]}{\sqrt{v_i v_j}} \quad (7)$$

where d denotes the time lag between the variables.

Ξ in Eq. (2) denotes the high order statistics, including skewness (γ_i), kurtosis (κ_i), and other high-order cumulants.

$$\gamma_i = \frac{\frac{1}{w}\sum_{l=0}^{w-1}[x_i(k-l)-\mu_i]^3}{\left(\frac{1}{w}\sum_{l=0}^{w-1}[x_i(k-l)-\mu_i]^2\right)^{3/2}} \quad (8)$$

$$\kappa_i = \frac{\frac{1}{w}\sum_{l=0}^{w-1}[x_i(k-l)-\mu_i]^4}{\left(\frac{1}{w}\sum_{l=0}^{w-1}[x_i(k-l)-\mu_i]^2\right)^{2}} - 3 \quad (9)$$

Higher-order statistics quantify the non-Gaussian probability distribution of the process variables, Specifically, skewness measures the asymmetry and kurtosis measures the "peakedness" of the process variable distribution. Note that, for a normal distribution, its skewness and kurtosis are both zero [20].

With different statistics computed, the final statistics for a window of original data is obtained by stacking all selected statistics into a row vector, and the statistics obtained from different windows are augmented together to build the training matrix, as shown in Fig. 1.

2.2. Locality preserving projections

While PCA handles high dimensional, noisy and correlated data by projecting the data onto a lower dimensional subspace which contains the most variance, LPP is a linear dimensionally reduction technique that optimally preserves the neighborhood structure of the data set. LPP can be seen as an alternative to PCA.

The LPP algorithm seeks a transformation matrix $\mathbf{A}$ to project high-dimensional input data $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n] \in R^{m\times n}$ into a low-dimensional subspace $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n] \in R^{l\times n}$ ($l < m$), such that $\mathbf{y}_i$ represents $\mathbf{s}_i$, where $\mathbf{y}_i = \mathbf{A}^T\mathbf{s}_i$.