Journal of Process Control 69 (2018) 44–57

Contents lists available at ScienceDirect

Journal of Process Control

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

Multi-phase batch process monitoring based on multiway weighted global neighborhood preserving embedding method

Yongyong Hui [a,b], Xiaoqiang Zhao [a,b,c,*]

[a] College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China
[b] Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou University of Technology, Lanzhou, 730050, China
[c] National Experimental Teaching Center of Electrical and Control Engineering, Lanzhou University of Technology, Lanzhou, 730050, China

ARTICLE INFO

Article history:
Received 1 October 2017
Received in revised form 1 April 2018
Accepted 29 June 2018

Keywords:
Batch
Process monitoring
Multi-phase
Global-local
Probability weighted
Gaussian mixture model

ABSTRACT

A multi-phase batch process monitoring method based on multiway weighted global neighborhood preserving embedding (MWGNPE) is proposed. MWGNPE has three advantages. Firstly, for the multi-phase feature of batch process, gaussian mixture model (GMM) method is used to divide phases by clustering characteristics. Secondly, after the multiple phases have been divided, global and local structures are extracted by using global neighborhood preserving (GNPE) method. Thirdly, probability density estimation characteristic of GMM is introduced to estimate the probability density of the extracted global and local structures. It can amplify useful information and suppress noise. These three advantages make MWGNPE well suit for batch process monitoring. A full MWGNPE model is combined with the cluster and the density estimation characteristic of GMM concurrently to improve the effect of fault detection in batch process monitoring. The effectiveness and advantages of proposed method are verified by a numerical system and the penicillin fermentation process. The results show that the proposed method can effectively capture the fault information hidden in process data and has the superiority compared with other conventional methods.

1. Introduction

Batch process is an important production mode in industrial production and has been widely used in modern process industry, especially in the pharmaceutical, food and semiconductor industries [1]. Compared with continuous industrial production process, batch process has the characteristics of complex mechanism, limit product life cycle, multi-phase and so on. The process is vulnerable to a variety of uncertain factors. In order to improve product quality and the safety of production process, it is urgent to establish a corresponding monitoring system to carry out effective monitoring.

The methods of multivariate statistical process monitoring such as principal component analysis (PCA) [2], partial least squares (PLS) [3] have been used for batch process monitoring. Some methods based on dynamic, nonlinear, multimodal, non-Gaussian and other characteristics [4–9] are proposed to improve batch process monitoring effects. Such as independent component analysis (ICA) [10], canonical variable analysis (CVA) [11], kernel entropy component analysis(KECA) [12], fisher discriminant analysis (FDA) [13] and gaussian mixture model (GMM) [14,15] are used to offset the defects of methods based on PCA and PLS. In many instances, these methods have good performances for dimension reduction and process monitoring, but they only can extract the global structure of process data, the potential local structure cannot be extracted [16,17]. As a result, it is difficult to extract the complete structure of process data.

On the other hand, He et al. [18] proposed neighborhood preserving embedding (NPE) which was a dimensionality reduction algorithm. It is different with global structure preserving algorithms like PCA and PLS. NPE preserves the neighborhood structure by projecting neighboring points of original higher dimensional space into low dimensional space and has been widely used in industrial process monitoring [19,20].

A good dimensionality reduction algorithm should reveal the inherent properties of a data set effectively. In other words, adjacent data points in original space should maintain the neighborhood in dimensionality reduction process, and the data points that are distant each other should also maintain this characteristic. NPE can only find the relationship between the nearest neighbor data points in dimensionality reduction process, the distant data points are not specifically described, which cannot guarantee

* Corresponding author at: College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China.
E-mail address: xqzhao@lut.cn (X. Zhao).

https://doi.org/10.1016/j.jprocont.2018.06.012
0959-1524/

the relationship between alienated data points in low-dimensional structure [21,22]. In order to preserve global and local structures and extract data information, Zhang et al. [17] proposed global-local structure analysis (GLSA) algorithm, which combined the objective function of global structure and local structure. Also, Yu [23] proposed LGPCA-based method which found an optimally preserve of processing local and global information. Zhao et al. [24] proposed a global neighborhood preserving embedding (GNPE) method, which integrated PCA and NPE successfully to preserve the global and local structures and the performance was better than PCA and NPE.

However, for batch process, there are multiple operating phases, each phase has its own unique characteristics, and production process shows a relatively large difference [25]. The above algorithms don't take into account the unique characteristics of different phases and only establish a unified statistical model. This would suppress some useful information of process data and affect process monitoring.

Each phase has distinct process correlation characteristics, a fault also has phase characteristic. Zhao et al. [26] proposed ELM-based fault detection and diagnosis method which conducted fault diagnosis for batch process more precisely. To preserve the time sequence of different phase while maintaining the kernel parameters unchanged for computation, Liu et al. [27] proposed phase partition for nonlinear batch process monitoring. To explicitly address the assumption that process data followed unimodal and Gaussian distribution, Yu [28] proposed a process monitoring method through manifold regularization-based GMM with global/local information.

After dividing phases, the above methods treat the different elements in each phase equally, the significance of different elements is not considered in projection matrix for process monitoring. The existing fault information could not be projected exactly in the extracted projection matrix. Therefore, the fault information would be scattered and could not converge. So how to make the most of the characteristics of each element in projection matrix is vital for efficient and reliable process monitoring.

Considering the above analyses, we propose a multi-phase batch process monitoring method based on MWGNPE to improve the effect of process monitoring. Since GMM can classify complex process data and obtain probability density estimation, batch process is divided into different operation phases by using GMM. After classification, the process data structure of each phase is similar and it is easy to carry out statistical analysis, so GNPE method is used for dimensional reduction and feature extraction in each phase. Because GNPE can extract effective characteristics from process data, it preserves global and local structure information. Based on the features extracted by GNPE, the probability density of each obtained elements in projection matrix is estimated by GMM. Since the probability density between an abnormal data and a normal data is different, it can identify an abnormal state, thereby fault information is enhanced and the noise is suppressed. The efficiency of the proposed method is improved. A numerical system and the penicillin fermentation process are used to verify the effect of the proposed method in batch process monitoring.

2. Neighborhood preserving embedding (NPE)

Neighborhood Preserving Embedding (NPE) is a linear approximation of LLE. First, k nearest neighbors are used to construct neighborhood graph. Before projecting the data into the low-dimension space, weight coefficient matrix W of the neighborhood graph should be assigned first. If node i to j has a connection edge, then edge weight value is w_{ij}. If no connection, the weight value is 0. The weight coefficient matrix can be obtained by solving optimal solution of Eq. (1).

$$\begin{aligned}\Phi(w) = \min \sum_i \left| |x_i - \sum_j w_{ij} x_j | \right|^2 \\ \sum_j w_{ij} = 1, i = 1, 2, \cdots n\end{aligned} \tag{1}$$

If w_{ij} can reconstruct data point x_i in spaceR^m, it can also reconstruct corresponding points $y_i$in space R^m. Therefore, mapping transformation matrix $A\,(a_1, \cdots, a_d)$ can be obtained by solving optimal solution of Eq. (2).

$$\begin{aligned}A_{opt} &= \sum_i \left(y_i - \sum_j w_{ij} y_j \right)^2 \\ &= y^T (I-W)^T (I-W) y \\ &= a^T X (I-W)^T (I-W) X^T a \\ &= a^T X M X^T a\end{aligned} \tag{2}$$

where$M = (I-W)^T (I-W)$, the constraint is$y^T y = a^T X X^T a^T = 1$. This transforms the problem of the transformation matrix into generalized eigenvalue of Eq. (3):

$$XMX^T a = \lambda XX^T a \tag{3}$$

The smallest d eigenvalues $(\lambda_1 \le \lambda_2, \cdots, \le \lambda_d)$ corresponding to the eigenvectors in Eq. (3) form a transformation matrix $A = (a_1, a_2, \cdots, a_d)$. Therefore, the original space X is mapped into feature space by mapping relation $Y = A^T X$. A is the mapping transformation matrix.

3. Gaussian mixture model (GMM)

For training sample set $X = \{x_1, x_2, \ldots, x_m | x_i \in R^{1\times n}\}$, GMM algorithm assumes that X is a mixed distribution model consisting of K Gaussian distributions which mean is $\mu_j$and variance isδ_j. Its typical applications include cluster and probability density estimation [15,29]. $x_i$has a corresponding probability value $g(x_i|C_j)$for each cluster, the conditional probability of $x_i$belonging to the jth cluster is $g(x_i|C_j)$, which is shown as Eq. (4).

$$g(x_i|C_j) = \frac{1}{2\pi \left|\sum_j\right|^{1/2}} \exp\left(\frac{(x_i - \mu_j)\sum_j^{-1}(x_i - \mu_j)^T}{2}\right) \tag{4}$$

where $i = 1, 2, \cdots, m; j = 1, 2, \cdots, K$. $\sum_j$and μ_j denote the covariance matrix and the mean vector of the jth clusterC_j, respectively. Thus the probability density function of x_i can be expressed as K Gaussian model:

$$p(x_i) = \sum_{j=1}^{K} w_k g(x_i|C_j) \tag{5}$$

Parameter$\Theta = \{\{w_1, \mu_1, \sum_1\}, \ldots \{w_K, \mu_K, \sum_K\}\}$can be determined by expectation maximization (EM) algorithm. Given training data X, K Gaussian model and the initial parameter $\Theta^{(0)} = \{\{w_1^{(0)}, \mu_1^{(0)}, \sum_1^{(0)}\}, \ldots, \{w_K^{(0)}, \mu_K^{(0)}, \sum_K^{(0)}\}\}$, EM repeats iteratively to update the parameters to ensure that the likelihood of training data