Control Engineering Practice 50 (2016) 69–83

Contents lists available at ScienceDirect

Control Engineering Practice

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

An improved weighted recursive PCA algorithm for adaptive fault detection

Ivan Portnoy*, Kevin Melendez, Horacio Pinzon, Marco Sanjuan

Universidad del Norte, Barranquilla, Colombia

ARTICLE INFO

Article history:
Received 30 June 2015
Received in revised form
9 January 2016
Accepted 24 February 2016
Available online 9 March 2016

Keywords:
Fault detection
False alarm
Recursivity
Time-drifting
Eigenvalue
Eigenvector
PCA

ABSTRACT

A novel weighted adaptive recursive fault detection technique based on Principal Component Analysis (PCA) is proposed to address the issue of the increment in false alarm rate in process monitoring schemes due to the natural, slow and normal process changes (aging), which often occurs in real processes. It has been named as weighted adaptive recursive PCA (WARP).

The aforementioned problem is addressed recursively by updating the eigenstructure (eigenvalues and eigenvectors) of the statistical detection model when the false alarm rate increases given the awareness of non-faulty condition. The update is carried out by incorporating the new available information within a specific online process dataset, instead of keeping a fixed statistical model such as conventional PCA does. To achieve this recursive updating, equations for means, standard deviations, covariance matrix, eigenvalues and eigenvectors are developed. The statistical thresholds and the number of principal components are updated as well.

A comparison between the proposed algorithm and other recursive PCA-based algorithms is carried out in terms of false alarm rate, misdetection rate, detection delay and its computational complexity. WARP features a significant reduction of the computational complexity while maintaining a similar performance on false alarm rate, misdetection rate and detection delay compared to that of the other existing PCA-based recursive algorithms. The computational complexity is assessed in terms of the Floating Operation Points (FLOPs) needed to carry out the update.

1. Introduction

PCA is a dimensionality reduction technique used for fault detection that optimally captures the maximum variance of the data in a low-dimensional space, and has been widespread used in process monitoring (Amanian, Salahshoor, Jafari & Mosallaei., 2007; Choi & Lee, 2004; Iwashita, 1997; Jeng, 2010; McGregor & Kourti, 1995; Russell, Chiang, & Braatz 2000; Qin, 2003). The large amount of observations gathered from sensors and actuators is turned into a couple of meaningful measures such as the T^2 and the Q statistics (Chen, Kruger, Meronk & Leung, 2004; Chiu & Ling, 2009; Iwashita, 1997). Its ultimate advantage is to perform the monitoring procedures like univariate charts, comparing a calculated measure with a statistical threshold, both arranged in a single plot (Chow, Tan, Tabe, Zhang, & Thornhill, 1999).

Among the drawbacks ascribed to conventional PCA, perhaps the major one is that, once the fault detection models have been structured in the training step, their monitoring schemes remain invariant. This feature becomes a significant disadvantage considering that real industrial processes usually demonstrate slow time-varying behaviors, such as catalyst deactivation, heat exchanger fouling, equipment and sensor aging and process time-drifting (Chen and Liao, 2002; Gallagher, Wise, Butler, White & Barna, 1997; Wold, 1994; Yingwei, Shuai, & Yongdong, 2012). As a consequence, false alarms will eventually occur unless the underlying statistical structure is updated.

When processes exhibit slow-varying changes, adaptive/recursive approaches are more suitable to address the false alarm issue. On the other hand, when processes exhibit several different operation conditions, the multimode monitoring approaches should be implemented (Zhiqiang, Zhihuan & Furong, 2013). Another type of monitoring methods, which tackles both time-varying and multimode processes, is found in literature, such as the *just-in-time-learning model* proposed by Cheng and Chiu (2005), the *adaptive local model* proposed by Ge and Song (2008), and the external analysis combined with ICA (Independent Component Analysis) proposed by Kano, Hasebe, Hashimoto and Ohno (2004). An overview of these methods is summarized in Fig. 1.

Many adaptive approaches have been developed based on PCA and Partial Least Squares (PLS) algorithms as these techniques

* Corresponding author.
E-mail addresses: iportnoy@uninote.edu.co (I. Portnoy), vkevin@uninorte.edu.co, kmelendez09@hotmail.com (K. Melendez), horacio.pinzon@gmail.com, hcoronado@uninorte.edu.co (H. Pinzon), msanjuan@uninorte.edu.co (M. Sanjuan).

http://dx.doi.org/10.1016/j.conengprac.2016.02.010
0967-0661/

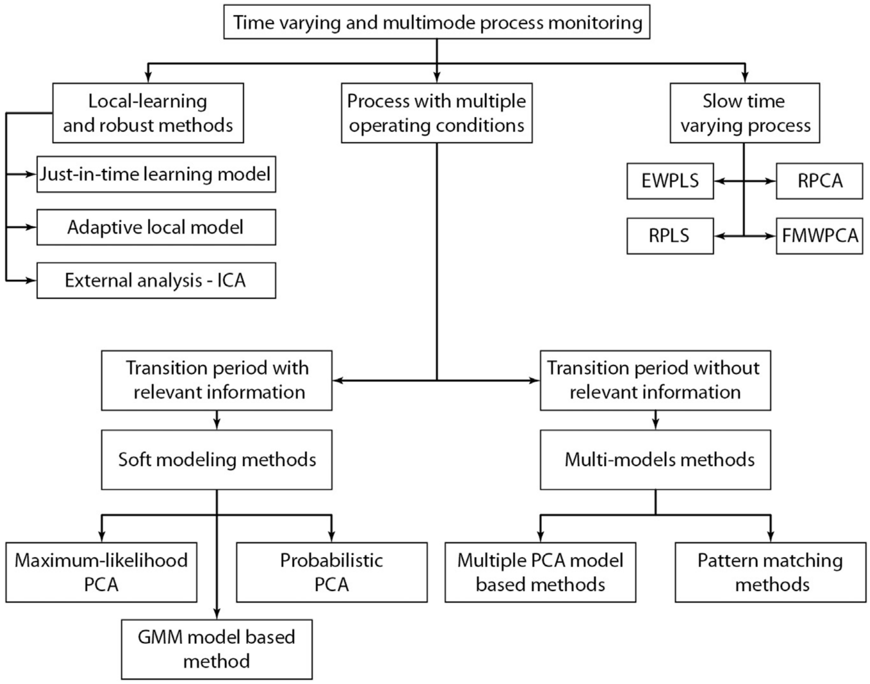

Fig. 1. Time-varying and multimode process monitoring overview.

register the more successful industrial implementations. Alkaya and Eker (2011) proposed a PCA-based variance-sensible fault detection algorithm combined with a dynamic threshold, mitigating false alarms caused by time-drifting behaviors by following the T^2 statistic trend and adjusting the detection threshold. Including the time-varying behavior and variable autocorrelation is a key feature for a robust fault detection scheme, being one the most important development branches on the fault detection field. Some approaches combine univariate techniques like Exponentially Weighted Moving Average (EWMA) and Cumulative Sum (CUSUM) charts with PCA, e.g. Wold (1994) discussed the use of EWMA filters in conjunction with PCA and PLS. Ku, Storer, and Georgakis (1995) proposed a modification to PCA to include time-lagged information to mitigate the temporal correlation among the process variables; this method is considered a dynamic version of the conventional PCA, hence dynamic PCA (DPCA) (Maravelakis & Castagliola 2009; Rato & Reis 2013; Weihua & Qin 2001).

Wang, Kruger, and Irwin (2005) presented a fast moving window PCA approach to improve monitoring efficiency of time-varying processes monitoring and Liu, Kruger, and Littler (2009) developed a moving window kernel PCA for non-linear time-varying process. Rigopoulos, Arkun, Kayihan and Hanezyc (1996) used a similar window scheme to identify significant modes in a simulated paper machine profile. Rannar, MacGregor and Wold (1997) used a hierarchical PCA for adaptive batch monitoring in a similar way to EWMA-based PCA. For industrial processes with multiple modes, different multimode approaches for process monitoring have been developed, such as the real-time monitoring approach proposed by Hwang and Han (1999). Moreover, when the monitoring of transition period between two different operation modes is a requirement, soft modeling algorithms offer a alternative to perform the fault detection procedures (Choi, Martin & Morris, 2005; Ge & Song, 2010; Yu & Qin, 2008).

Besides adaptive approaches, there are also solutions relying on the periodic incorporation of new process data, thus recursively updating the statistical fault detection model. Dayal and MacGregor (1997) developed a recursive exponentially weighted PLS method for adaptive control in industrial processes, Wang, Kruger and Lennox (2003) built a recursive PLS (RPLS) model for adaptive monitoring in complex industrial processes, Naik, Yin, Ding and Zhang (2010) propose algorithms to deal with recursive identification of parity-based fault detection systems, updating their eigenstructure after every new measurement, which improves fault detection performance against frequent shifts in operation point or parameter variations. Qin (1998a,1998b) proposed several RPLS algorithms for both offline and on-line process modeling allowing the adaptation to process changes and dealing with a large number of data samples. These algorithms include a block-wise RPLS with a moving window and a forgetting factor adaptation scheme and a block-wise RPLS off-line used to reduce computation time and computer memory usage in PLS regression and cross-validation.

Like adaptive approaches, recursive techniques are developed based on PCA to take advantage of its widespread implementation. Jeng (2010) proposed a recursive PCA (RPCA) algorithm based on a rank-one matrix update. This algorithm pre-treats data to be mean-centered however, it does not perform an auto-scaling operation neglecting the effects of such changes on the standard deviations of process variables. Besides, this update is made after every new measurement (sample by sample), making it inconvenient due to the large amount of FLOPS (Floating Operation Points per Second) required. Weihua, Yue, Valle-Cervantes, and Qin (2000) proposed two PCA-based algorithms using a rank-one modification and a Lanczos tridiagonalization, respectively. After a computational complexity assessment, Weihua et al. concluded that the algorithm based on rank-one modification is less demanding. The rank-one algorithm carries out an auto-scaling operation to consider the changes on standard deviations of process variables; nevertheless it requires two spectral decompositions to update the eigenstructure. In addition, the formulas used to update the covariance matrix and standard deviations may be improved in order to lower their complexity. This algorithm also features a forgetting factor μ to weight current and new datasets.

In this paper, a new weighted adaptive recursive PCA-based algorithm is developed. A comparison between the proposed algorithm and other recursive PCA-based algorithms (Weihua et al., 2000; Jeng, 2010) is carried out in terms of false alarm rate, misdetection rate, detection delay and computational complexity. The paper is organized as follows: in Section 2, the background about