Neurocomputing 285 (2018) 60–73

Contents lists available at ScienceDirect

Neurocomputing

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

Brief papers

# Root cause diagnosis of quality-related faults in industrial multimode processes using robust Gaussian mixture model and transfer entropy

Liang Ma, Jie Dong, Kaixiang Peng*

*Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, PR China*

ARTICLE INFO

*Article history:*
Received 22 October 2017
Revised 13 December 2017
Accepted 12 January 2018

*Keywords:*
Quality-related
Root cause diagnosis
Multimode processes
Robust Gaussian mixture model
Bayesian inference
Transfer entropy

ABSTRACT

Modern complex industrial processes often have multiple operating modes due to various factors, such as different manufacturing strategies, alterations of feedstock and compositions, etc. In this paper, a practical technology or solution of quality-related fault diagnosis is put forward for industrial multimode processes. Different from traditional data-based fault diagnosis methods, the alternative approach is focused more on root cause diagnosis. The new scheme addresses the quality-related fault detection issue with a developed robust Gaussian mixture model and modified Mahalanobis distance. Then, a Bayesian inference-based robust Gaussian mixture contribution index is designed to analyze the potential root-cause variables. Meanwhile, a combination of transfer entropy and direct transfer entropy-based cause and effect extraction methodologies is proposed for root cause diagnosis of quality-related faults. Finally, the whole proposed framework is applied to a real industrial multimode finishing mill process, where the performance and effectiveness are further demonstrated from real industrial data.



## 1. Introduction

As the increasing market demands for multi-species, multi-specifications as well as high-quality products, modern industrial processes depend more on producing small batch and high value-added products. However, in industrial practice, due to various reasons such as production strategy changes, feedstock and operating condition shift, the process covers multiple operating modes which have distinct process correlation characteristics [1–3]. Once a fault occurs during the multimode process operation, it may propagate to the whole or specific control loops by means of information and/or material flow pathways, which will affect the overall process performance and the final products' quality [4–8]. Accordingly, in order to keep high efficiency of operation and ensure stability of product quality for industrial multimode processes, automated detection of quality-related faults and accurate diagnosis of the root causes are of urgent necessity, which have recently attached more and more attention both in academia and engineering domains.

In general, process monitoring and fault diagnosis (PM–FD) methods can be divided into two categories, which are the model-based and the data-driven techniques. Although the model-based methods have been deeply studied and a lot of fruitful results have been achieved [9–15], these techniques present some implementation difficulties in complex industrial systems such as steel-rolling, chemical, paper-making, oil-refining and etc., where the first-principle models of plants are generally difficult to be established. In comparison, data-driven methods, thanks to their simple forms and fewer requirements on the design and engineering efforts, have become more and more popular nowadays [16–22].

In actual production processes, production and product quality, such as the concentration in a chemical process, the thickness of a steel roll between two stands in a steel mill process as well as the concentration of liquor in an alumina evaporation process, are in general difficult to directly measure online or the measurement is implemented after the production process is completed. As a result, identifying covariances or correlations model between process variables and quality variables is particularly important. Among all the available modeling methods, partial least squares (PLS) [23–25], principal component regression (PCR) [26–28] and canonical variable analysis (CVA) [29] are three popular ones, which have been intensively investigated for quality-related fault detection and applied in diverse processes. The Hotelling's $T^2$ and squared prediction error (SPE) indexes along with the corresponding control limits are developed to serve as quality-related fault detection, then the decomposition or reconstruction-based contribution indexes can be derived to identify the faulty variables [2,30–33]. Even though these methods show strong applicabilities in multivariate statistical process monitoring (MSPM) area, the $T^2$

* Corresponding author.
*E-mail address:* kaixiang@ustb.edu.cn (K. Peng).

https://doi.org/10.1016/j.neucom.2018.01.028
0925-2312/

and SPE control limits in above methods are established from the scores and residuals of Gaussian distribution approximately, which may not be fulfilled under different operating conditions leading to multiple steady states.

In order to extend classical MSPM methods to multimode processes, many modified approaches like multiple principle component analysis (PCA)/PLS have been presented, while their major limitation lies in the fact that the number of operating modes existing in the process should be a-priori known and fixed [1,34–38]. More recently, the independent component analysis (ICA)-based monitoring approaches to detect the process faults from non-Gaussian data have been developed to tackle this issue [39–45]. Different from PCA/PLS-based methods, higher-order statistic of negentropy is used to measure the inherent non-Gaussianity of process data instead of the second-order statistics-based variance/covariance. However, in practice, due to the frequent shifting of operating conditions, the negentropy index does not completely reflect the multimodality of process data, which may affect the monitoring performance of ICA. Though some supervised techniques such as support vector machine (SVM) and localized Fisher discriminant analysis (LFDA) have been used for process monitoring with multimodal non-Gaussianity, they also require labeled data from normal and faulty operating conditions [46–48]. As an alternative technique, a density-based Gaussian mixture model (GMM) method has shown strong capability in monitoring non-Gaussian process with multiple operating modes, which has been extensively applied to various continuous or batch processes [2,31,49–51]. The GMM-based process monitoring method decomposes the process data into multiple Gaussian components, in which each Gaussian component corresponds to an individual operating mode. Herein, the underlying multimodality of process data can be completely preserved and the abnormal operation can be detected either by $T^2$/SPE index within a particular mode or Bayesian inference across multiple modes [2,31].

After a fault detected by above methods, next challenges are to identify the candidate dataset of faulty variables and root cause diagnosis of abnormal events from multimode plant operation. Nevertheless, in fact, comparing with the fault detection for multimode processes, the research for root cause diagnosis has just started and is far from prefect. Although several data-based causal analysis techniques, such as cross-correlation function (CCF) [52,53], Granger causality (GC) [6,54], transfer entropy (TE) [29,55–57] and Bayesian network (BN) [58–62] have obtained satisfactory results without requiring deep process knowledge, little work has been reported to address the more challenging root cause diagnosis issue on multimode processes. As a useful supplement, in this work, a new framework for root cause diagnosis of quality-related faults in industrial multimode processes is put forward, including:

- developing a robust GMM (RGMM)-based quality-related fault detection approach;
- presenting a novel Bayesian inference-based robust Gaussian mixture contribution (BIRGMC) index for identification of candidate dataset of faulty variables;
- proposing a combination of TE and direct TE (DTE)-based cause and effect extraction methods for root cause diagnosis of quality-related faults;
- examining the applicability of the new framework to an actual finishing mill process (FMP).

The rest of this paper is organized as follows. In Section 2, the preliminaries and the problems to be addressed are given. Section 3 introduces RGMM-based quality-related fault detection method for multimode processes. After that, Section 4 is dedicated to the root cause diagnosis of quality-related faults. Then, the presented scheme is implemented on a real FMP in Section 5. In the end, concluding remarks and future works are presented in Section 6.

## 2. Preliminaries and problem formulation

Consider an available historical data $\mathcal{D}$ collected from $N$ different samples for a multimode process, each sample contains measurements of process variables $\mathbf{x} \in \mathbb{R}^m$ and quality variable $y \in \mathbb{R}$

$$\mathcal{D} = \left\{ \begin{bmatrix} y(1) \\ \mathbf{x}(1) \end{bmatrix}, \begin{bmatrix} y(1) \\ \mathbf{x}(2) \end{bmatrix}, \dots, \begin{bmatrix} y(1) \\ \mathbf{x}(N) \end{bmatrix} \right\} = \left\{ \mathbf{d}(1), \mathbf{d}(2), \dots, \mathbf{d}(N) \right\} \tag{1}$$

Then, the probability density function (PDF) of a simple point $\mathbf{d}$ can be represented as

$$p(\mathbf{d}|\mathbf{\Theta}) = \sum_{k=1}^{C} \omega_k g(\mathbf{d}|\theta_k) \tag{2}$$

where $C$ is the number of Gaussian components, $\omega_k$ is the prior probability of the $k$th component satisfying $\Sigma_{k=1}^{C} \omega_k = 1$, $\theta_k = \{\mu_k, \Sigma_k\}$ and $\mathbf{\Theta} = \{\theta_1, \theta_2, \dots, \theta_C\}$ represent the sets of local and global Gaussian model parameters, respectively.

For the $k$th component, its Gaussian density function $g(\mathbf{d}|\theta_k)$ is given by

$$g(\mathbf{d}|\theta_k) = \frac{1}{(2\pi)^{m/2}|\Sigma_k|^{1/2}} \exp\left[ -\frac{1}{2} (\mathbf{d} - \mu_k)^T \Sigma_k^{-1} (\mathbf{d} - \mu_k) \right] \tag{3}$$

Let $\mathbf{Z} = \{\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(N)\}$ be the missing data in which $\mathbf{z}_i \in \{1,\ 2,\ \cdots,\ C\}$. If $\mathbf{z}_i = k$, it means that the $i$th simple point belongs to the $k$th component, then $z_{ki} = 1$, otherwise $z_{ki} = 0$. To construct a GMM for quality-related fault detection, the parameters in $\mathbf{\Theta} = \{\{\omega_1, \theta_1\}, \dots, \{\omega_C, \theta_C\}\}$ need to be estimated from the sample dataset $\mathcal{D} = \{\mathbf{d}(1), \mathbf{d}(2), \dots, \mathbf{d}(N)\}$ by the well-known expectation-maximization (EM) algorithm [63] until the following log-likelihood function increases to a local maximum

$$L(\omega, \mathbf{\Theta}; \mathcal{D}, \mathbf{Z}) = \sum_{i=1}^{N} \sum_{k=1}^{C} z_{ki} \ln[\omega_k g(\mathbf{d}|\theta_k)] \tag{4}$$

- E-step: according to [63] and Bayes' theorem, compute the latent variables $\hat{z}_{ki}$

$$\hat{z}_{ki} = \frac{\omega_k g(\mathbf{d}_i|\theta_k)}{\Sigma_{s=1}^{C} \omega_s g(\mathbf{d}_i|\theta_s)} \tag{5}$$

- M-step: update the model parameters

$$\mu_k = \frac{\Sigma_{i=1}^{N} \hat{z}_{ki} \mathbf{d}_i}{\Sigma_{i=1}^{N} \hat{z}_{ki}} \tag{6}$$

$$\Sigma_k = \frac{\Sigma_{i=1}^{N} \hat{z}_{ki} (\mathbf{d}_i - \mu_k)(\mathbf{d}_i - \mu_k)^T}{\Sigma_{i=1}^{N} \hat{z}_{ki}} \tag{7}$$

$$\omega_k = \frac{\Sigma_{i=1}^{N} \hat{z}_{ki}}{N} \tag{8}$$

Although the EM algorithm has obtained ideal results for estimating the distribution parameters of multimode operating data, three issues must be taken into account:

- the number of Gaussian components $C$ needs to be given a priori;
- the local searching approach may fall into local maximal;
- the collinearity among variables may lead the covariance matrix $\Sigma_k$ to be ill-conditioned.

To cope with these problems, in this paper, motivated by Yang's work [64], we present a robust way to obtain the number of Gaussian components automatically by the classical EM algorithm and information entropy. Based on it, we address the quality-related