Neurocomputing 149 (2015) 652–666

Contents lists available at ScienceDirect

Neurocomputing

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

Noise-resistant joint diagonalization independent component analysis based process fault detection ☆

Xuemin Tian [a], Lianfang Cai [a], Sheng Chen [b,c,*]

[a] College of Information and Control Engineering, China University of Petroleum, Qingdao 266580, China
[b] Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
[c] Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

ARTICLE INFO

Article history:
Received 11 February 2014
Received in revised form
27 May 2014
Accepted 4 August 2014
Communicated by J. Zhang
Available online 15 August 2014

Keywords:
Fault detection
Independent component analysis
Mixing matrix
Measurement noise
Time-delayed covariance matrices
Kurtosis

ABSTRACT

Fast independent component analysis (FastICA) is an efficient feature extraction tool widely used for process fault detection. However, the conventional FastICA-based fault detection method does not consider the ubiquitous measurement noise and may exhibit unsatisfactory performance under the adverse effects of the measurement noise. To solve this problem, we propose a new process fault detection method based on noise-resistant joint diagonalization independent component analysis (NRJDICA), which explicitly takes the measurement noise into consideration. Specifically, the NRJDICA algorithm is developed to estimate the mixing matrix and the independent components (ICs) by whitening the measured variables and performing the joint diagonalization of the whitened variables' time-delayed covariance matrices. The relationships between the kurtosis statistics of the ICs and the fourth-order cross cumulant statistics of the measured variables are then derived based on the estimated mixing matrix to help sorting the estimated ICs and selecting the dominant ICs. The serial correlation information of each dominant IC is next estimated by using a moving window technique, based on which a monitoring statistic is constructed to conduct fault detection. The simulation studies using a three-variable system and a continuous stirred tank reactor show that the proposed method has superior fault detection performance over the traditional FastICA-based fault detection.

1. Introduction

With the advancement in measurement technology and distributed control systems, modern industrial processes have become increasingly more complex. To ensure process safety and stability as well as to maintain high quality of final products, reliable and timely fault detection has emerged as an essential task. Due to the convenient availability of substantial measured data in industrial plants, multivariate statistical analysis methods, which can extract meaningful feature information from large amounts of the measured data for detecting various faults or abnormal situations of industrial processes, have attracted much attention from both process engineers and academic researchers [1–18]. Principal component analysis (PCA), as one of the classical multivariate statistical analysis approaches, has found wide-ranging applications in the fault detection field [11–16,19–22]. PCA projects the high-dimensional and correlated measured variables onto a smaller set of the uncorrelated latent variables called the principal components (PCs) that retain most of the original variance. However, PCA can only utilize the second-order zero-delayed covariance information and it cannot take the meaningful time-delayed covariance statistical information or the higher-order statistical information of the measured data into consideration [23–25], which may lead to insufficient feature extraction and unsatisfactory fault detection performance. Moreover, fault detection based on PCA assumes that the measured data follow a multivariate Gaussian distribution, in order to derive the control limits of Hotelling's T-squared (T^2) and the companion squared prediction error (SPE) monitoring statistics. In practice, industrial process data usually obey non-Gaussian distribution due to process nonlinearity, operating condition shifts or other reasons [26,27], and the control limits derived based on the Gaussian distribution assumption may be ill-suited for fault detection purpose as the resulting fault indication may be biased.

More recently, fault detection based on independent component analysis (ICA) has become a hot topic [2–6,8,10,23–40]. ICA was originally derived for solving the blind source separation

☆This work is supported by the National Natural Science Foundation (NNSF) of China (61273160), the NNSF of Shandong Province of China (ZR2011FM014), the Doctoral Fund of Shandong Province (BS2012ZZ011), the Fundamental Research Funds for the Central Universities (14CX02174A), and the Postgraduate Innovation Funds of China University of Petroleum (CX2013060).

* Corresponding author at: Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK.
E-mail addresses: tianxm@upc.edu.cn (X. Tian), cailianfang@163.com (L. Cai), sqc@ecs.soton.ac.uk (S. Chen).

http://dx.doi.org/10.1016/j.neucom.2014.08.009
0925-2312/

problem [41–48] and was introduced to fault detection by Kano et al. [29]. As a multivariate statistical analysis method, ICA can exploit the higher-order statistical information [41–45] or the second-order time-delayed covariance statistical information [32,47] to extract mutually independent latent variables called independent components (ICs) from the measured non-Gaussian variables, and it can be regarded as a useful extension of PCA. ICA is especially suitable for process fault detection as real-world processes are usually non-Gaussian [23–27]. Different criteria have been considered to develop various ICA algorithms [41], including the maximization of the non-Gaussian measures, such as negentropy, the minimization of the mutual information and the maximum likelihood estimation. Among the existing ICA algorithms, the fast ICA algorithm (FastICA) [41] based on the maximum negentropy criterion is widely used in the ICA-based fault detection methods because of its fast convergence rate and good non-Gaussian feature extraction ability [34,35]. Lee et al. [25] developed a FastICA-based fault detection method by utilizing the FastICA to extract the non-Gaussian ICs from the measured data and constructing three elliptical-type monitoring statistics, known as I^2, I_e^2 and SPE. Hsu et al. [36] argued that I^2, I_e^2 and SPE may not always appropriately capture the characteristics of the extracted ICs by the FastICA because of the ICs' skewed distributions, and developed a rectangular-type monitoring statistic named the adjusted outlyingness to monitor non-Gaussian processes. After concluding that both the elliptical-type and rectangular-type monitoring statistics may not always accurately estimate the nonlinear feature space boundary of normal operating condition (NOC), Lee et al. [37] constructed a monitoring statistic by using a local outlier factor method on the ICs extracted by the FastICA, which can effectively determine the nonlinear decision boundary of NOC. Zhao et al. [6] combined the ideas of partial least squares (PLS) and FastICA under the same mathematical umbrella by constructing a dual-objective optimization criterion that can consider higher-order statistical independence and quality-related requirements simultaneously. The modified independent component regression method proposed in [6] can extract the latent variables, which have close correlation with the quality properties and are more comprehensible for regression modeling.

Furthermore, by considering different process characteristics, such as nonlinearity, dynamic or multi-modality, researchers have proposed various improved fault detection methods based on the FastICA. In particular, Stefatos and Hamza [38] developed a dynamic ICA method for dynamic process fault detection by augmenting the original measured data with the previous observations and applying the FastICA to extract the ICs from the augmented data. Odiowei and Cao [39] integrated the canonical variate analysis with the FastICA and developed a state-space ICA based fault detection method for dynamic processes. Tian et al. [34] proposed a multiway kernel FastICA method based on feature samples for monitoring nonlinear batch processes. Cai et al. [33] presented a nonlinear process fault detection method by integrating the kernel FastICA with a newly emerging manifold learning method known as locality preserving projection. Basically, kernel FastICA integrates kernel PCA (KPCA) with FastICA [2,3,24,34], and thus it combines the advantages of both KPCA [22] and FastICA [41]. In other words, kernel FastICA possesses the unique capability of data processing which KPCA alone does not have. More specifically, kernel FastICA can indirectly excavate the second-order time structure information of time-series data by utilizing the higher-order information to extract mutually independent kernel ICs. This is important because the second-order time structure information, such as the time-delayed covariance information, can be utilized as a viable alternative for the higher-order information [45]. By contrast, KPCA is blind to the second-order time structure information, and can only utilize the second-order zero-delayed covariance information to extract kernel PCs which are only uncorrelated but not independent. Therefore, kernel FastICA is more appropriate for time-series data than KPCA. Zhang [2] also pointed out that ICs can reveal more dynamic information from the measured data than PCs To account for the process multimodal characteristics, Rashid and Yu [28] proposed a hidden Markov model based adaptive FastICA approach for monitoring non-Gaussian processes with multimodality. Zhang et al. [3] developed a multimodal process fault detection method based on the Kronecker product and modified kernel FastICA. In the above studies, the FastICA has been employed to extract non-Gaussian ICs for fault detection. However, this widely used FastICA is basically a “noise-free” algorithm, which adopts the noise-free ICA model and does not explicitly take the influence of the measurement noise into account. In reality, the measurement noise corruption always exists in industrial processes, as pointed out by Ge and Song [12] and Wang [49]. Under the adverse effects of measurement noise, the FastICA may not conduct effective and reliable feature extraction from the measured data. Moreover, the traditionally used monitoring statistics are also vulnerable to the measurement noise. These limitations may degrade the performance of the FastICA-based fault detection methods drastically.

Recently, researchers have begun to focus on the challenging problem of how to extract more accurately the features from the measured noisy data for fault detection. Kim and Lee [16] extended the conventional PCA to the probabilistic PCA (PPCA) for process monitoring, which considers the noise information of the measured data. Zhu et al. [14] further proposed the robust mixture PPCA for process monitoring. Compared to the PPCA, the robust mixture PPCA can reduce the negative effect of outliers and deal with the missing data problem more effectively. However, PPCA based techniques are only suitable for linear processes. By adopting a kernel technique, Ge and Song [12] proposed the kernel generalization of PPCA to extract nonlinear features for monitoring nonlinear processes. To meet the critical demand of identifying fault variables that contribute to process faults, Chen and Sun [20] proposed a probabilistic contribution analysis method in conjunction with the PPCA model based on the concept of missing variable. Noting that the PPCA requires the same noise level in all the measured variables, Kim et al. [17] utilized the factor analysis (FA), which can be regarded as an extension of PPCA and is capable of dealing with practical situations where the noise levels are different in different measured variables. Considering that fault information may not have a definite mapping relationship to a single factor and the useful information captured by some factors may be submerged in the insignificant information of the other factors, Jiang and Yan [18] further proposed the weighted FA (WFA) for process monitoring, which is capable of highlighting the useful information that is submerged in the insignificant information by suppressing the latter. These existing works have led to some successes in practical applications, but there still exist many critical and important issues which require further analysis and research.

Specifically, most existing works assume that both latent variables and noise variables obey Gaussian distributions. This assumption is usually violated in real-life industrial processes where the related variables often have non-Gaussian characteristics, owing to various reasons, such as shifting operating conditions, feedstock changes, production strategy changes, non-Gaussian disturbances, and process nonlinearity [18,27]. The assumption that all the noise variables have the identical power level, as required by the PPCA-related methods [12,14,16,20], is also too rigorous and unrealistic. Thus, an appropriate multivariate statistical analysis approach is urgently demanded, which can overcome these above-mentioned limitations and is capable of meeting the practical requirements of fault detection in real-life