ARTICLE IN PRESS

ISA Transactions xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

ISA Transactions

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

Practice article

Fault detection and isolation in the challenging Tennessee Eastman process by using image processing techniques

Payman Hajihosseini[a,*], Mohammad Mousavi Anzehaee[b], Behzad Behnam[a]

[a] Department of Electrical Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran
[b] Department of Electrical Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Keywords:
Fault detection and isolation
Image processing
Image texture
2D Wavelet packet transform
Classifier

ABSTRACT

The early fault detection and isolation in industrial systems is a critical factor in preventing equipment damage. In the proposed method, instead of using the time signals of sensors, the 2D image obtained by placing these signals next to each other in a matrix has been used; and then a novel fault detection and isolation procedure has been carried out based on image processing techniques. Different features including texture, wavelet transform, mean and standard deviation of the image accompanied with MLP and RBF neural networks based classifiers have been used for this purpose. Obtained results indicate the notable efficacy and success of the proposed method in detecting and isolating faults of the Tennessee Eastman benchmark process and its superiority over previous techniques.

1. Introduction

Due to the growing demands for higher production quality, the performances of different systems have greatly improved, and the degree of automation and the complexity of various industrial processes such as metallurgical and chemical processes have increased substantially. As an important tool for fulfilling these demands, the condition monitoring of various industrial equipment has received a great deal of attention [1]. Stated differently, with the progresses made in modern engineering processes, and in order to improve the reliability, safety, accessibility and the environmental friendliness of such processes, the detection of process faults has become ever more important. The main goal here is to improve the quality of industrial production in a vast range processes. To this end, many attempts have been made to achieve various fault detection algorithms by using different assumptions [2]. A common opinion among the researchers is that the quick detection of abrupt or incipient faults prevents the occurrence of larger damages and failures in systems [3].

Although large-scale industrial disasters like the ones at the Bhopal Union Carbide and the B.P. Deep Water Horizon oil spill are quite rare, smaller industrial mishaps are very common; and the mismanagement of these events impose an estimated annual cost of $20B on USA alone [4].

The necessity of dealing with these problems has prompted the researchers to focus various research efforts on the subject of Fault Detection and Isolation (FDI). In practice, with the detection of faults and their root causes by the FDI algorithms, the process operators and the repair and maintenance personnel are alerted and given the opportunity to improve the abnormal operating conditions [5].

The diversity of FDI techniques in the past decades has been due to the growing complexity of industrial processes. Because most industrial processes need a shorter fault detection time and online verification, the model-based fault diagnosis methods have become very popular [2]. Nevertheless, because of the high complexity and the extensiveness of industrial processes, lack of accurate process knowledge, and huge modeling expenses, it is difficult to provide a complete and robust mathematical model [6]. Thus, these approaches are limited to processes with few variables [7]. Data-based fault detection methods can be used to solve the abovementioned problems. Due to the high complexity and large dimensions of industrial processes, the execution of FDI algorithms in most of them is accomplished by employing the quantitative data-based techniques. Moreover, most of the quantitative data-based methods are based on classifiers [7].

In Ref. [8] a distributed monitoring scheme based on combination of Kernel Principle Component Analysis (KPCA) and Bayesian diagnosis system is presented to monitor large-scale nonlinear systems. Fault isolation is evaluated in terms of Missed Classification Rate (MCR) and only 5 faults (numbers 1, 2, 4, 5 and 10 which is not included complex faults) among 20 faults of TE process are investigated. Clustering of qualitative event sequences is presented in Ref. [9] for process fault isolation. First 15 faults among 20 faults of the TE process are isolated by using Fault Detection Rate (FDR) index whilst only stepwise faults

* Corresponding author.
E-mail addresses: p_hajihosseini@kiau.ac.ir (P. Hajihosseini), m_mousavi@azad.ac.ir (M.M. Anzehaee), b_behnam@kiau.ac.ir (B. Behnam).

https://doi.org/10.1016/j.isatra.2018.05.002
Received 2 April 2017; Received in revised form 4 January 2018; Accepted 4 May 2018
0019-0578/

Please cite this article as: Hajihosseini, P., ISA Transactions (2018), https://doi.org/10.1016/j.isatra.2018.05.002

are properly diagnosed and diagnosis percent of other faults were not satisfactory. In Ref. [10], deep belief network idea is used for monitoring complex chemical processes. The results of FDI for TE process were based on FDR index whilst faults 15 and 16 were totally unrecognized and faults 5, 9 and 12 were not properly isolated. A multivariable statistical method including Support Vector Machine (SVM) and Principle Component Analysis (PCA) is applied for FDI in Ref. [11]. It is shown that PCA had better classification accuracy% than SVM for TE process, but in order to reach to better results more complex faults of the process (faults 3, 9 and 15) should be excluded in both of them. Exponential Discriminant Analysis is an idea which is introduced in Ref. [12] as an alternative for Fisher Discriminant Analysis (FDA) if a small size of data is reachable. However, only normal state and fault numbers 2, 6 and 8 are investigated. An algorithm named cascade feature selection is used in Ref. [13] for combining feature selection methods. Three different combinations are presented in this regard, but F1 score values were not appropriate for faults 3, 9, 10 and 15 of the TE process. In Ref. [14], an improved case-based reasoning method is derived to evaluate the performance of TE process. Among process faults, first 15 faults are studied while classification accuracy% was reasonable except that fault 2. An FDI method based on Relevance Vector Machine (RVM) suitable for small sample size of process data is presented in Ref. [15]. It is shown that fault numbers 5, 9, 15, 18 and 20 were isolated with less than 80% value of F1 score in TE process. In Ref. [16], a wavelet-based multi-scale data representation method is used to improve PCA ability in FDI operation. Missed detection rate values were not proper for faults 3, 4, 5, 7, 9, 11, 15, 16 and 19 of TE process. Also false alarm rates of faults 9 and 16 were big values. A fuzzy-Bayesian method for change point detection along with an immune-neural network is presented in Ref. [17] for FDI purpose. It is shown that F1 score was between 45.5% and 74.8% for faults 3, 4, 9, 10 and 14 of the TE process. Also, in Ref. [18], the author has discussed about the detection of only 4 process faults (from a total of 20 cases); and the diagnosed faults do not even include the complex faults of the process. In Ref. [19], process fault detection and identification has been done based on a combination of Generative Topographic Mapping (GTM) and Graph Theory (GT) algorithms. This work was performed on only 9 faults among 20 predefined faults of the TE process which was not also included complex faults (faults number 3, 9 and 15). Meanwhile, value of false positive detection rate in each fault has not been computed. Image processing of infrared thermal images in order to fault diagnosis in rotating machinery and cooling radiator have been recently performed in Refs. [20] and [21], respectively. Extracted features in Ref. [20] were standard deviation of pixel values and Gini coefficient while fault isolation results were reported by F1 score. In Ref. [21], statistical features including average, variance, energy, entropy, smoothness, and skewness were extracted relying on 2D wavelet transform of thermal image and its histogram. The results were reported by some indices like precision and recall. Plant-wide assessment of industrial systems' operating state relying on digital color images has been done in Ref. [22]. In this research, an abnormal-degree-function was introduced as a classifier to grade process operation state in range 0–1. It was shown in TE process that the faults 6 and 18 are the most dangerous faults among all.

In Ref. [23], a PCA-based multidimensional visualization technique has been employed for fault detection and isolation. The results of fault diagnosis in that paper indicate deficient accuracies for detecting faults 3, 5, 9, 10, 11, 15, 16, 19 and 20 in the Tennessee Eastman (TE) benchmark process. In Ref. [7], a data-based method called the 'transfer entropy algorithm' has been used for diagnosing process faults. The technique presented in that paper is somewhat complicated, and the model of information transfer between process variables and the cause and effect relationships between them have been used to detect and classify the faults; and of course, good results have been obtained in diagnosing all types of process faults (including complex faults). Also in Ref. [24], a decentralized method has been employed for fault isolation. The advantages of decentralized systems over centralized systems include shorter detection time and higher reliability. By using the mentioned approach, all the process faults have been grouped into appropriate categories, and the differentiation of faults has taken place in independent units. Also, the diagnosis of more complicated faults in the mentioned groups was not possible, and so these faults have been isolated by means of the transfer entropy algorithm. Despite the complexity of this approach, satisfactory results have been obtained in the diagnosis of all process faults.

In summary, common problems in previous works (relying on 1D representation of sensory data) are as follows:

1. Most of them are not focused on all 20 predefined faults of the challenging TE process, especially on more complex faults 3, 9 and 15.
2. In most of them, fault diagnosis evaluation is performed by using one incomplete index like FDR which is not included miss alarms and false alarms.
3. Some of process faults in each paper are isolated with low and improper accuracy.

Also, some references relying on 2D representation of sensory data have focused on other plants except that TE process by using thermal images with different features than this research. Also, there are some differences with other references worked on TE process. Some of them have not covered all process faults and one has focused on investigating danger grade of each process fault [22].

The method proposed for the detection and isolation of process faults in this research is based on the analysis of 2D images obtained from sensor data. Image processing algorithms have had extensively varied applications in engineering sciences, including the robotics, medical, meteorological and military fields. The idea presented in this research is to put together all the sensor data for each of the normal or faulty conditions of a process and thus to form a data matrix for each situation. In fact, instead of analyzing the sensor data in one dimension, the image corresponding to the matrix of all the data obtained for any condition becomes the basis for fault detection in that condition. Then, all simple and complex faults of the TE process are investigated relying on image texture features obtained by Gray Level Co-occurrence Matrix (GLCM) and Gray Level Gradient Co-occurrence Matrix (GLGCM) matrices and also texture features extracted from 2D wavelet packet decomposition of each image along with two global statistical features. At last, classification results are reported in terms of F1 score which is an average between missed alarm and false alarm rates. Another important achievement of this research is that both procedures of fault detection and fault isolation are carried out by using this same idea.

The rest of this paper has been organized as follows: in Section 2, the details of the proposed method have been described. The TE process, as a benchmark process, has been introduced in Section 3. A set of simulation results is presented in Section 4 to comparatively evaluate performance of the proposed method with those of the conventional approaches, and the concluding remarks have been summarized in Section 5.

2. Proposed method

In the proposed method in this research, the time samples of signals obtained from the sensors of a process are placed next to each other to form a numerical matrix. If the number of time samples obtained during the whole sampling interval is n and the number of sensors is m, an $n \times m$ matrix will be obtained for each operating condition (normal or faulty) of the process. Then by considering the equivalent gray level of each of the normalized matrix elements, a corresponding 2D image can be obtained for each of the matrices. Thus, in the normal operating condition, and in each faulty process condition, a 2D image is obtained from the sensor data of that condition. In subsequent steps, these images become the basis for the detection and isolation of process faults.