Jurnal
Teknologi

Full Paper

FAULT DIAGNOSIS BASED ON MULTI-SCALE CLASSIFICATION USING KERNEL FISHER DISCRIMINANT ANALYSIS AND GAUSSIAN MIXTURE MODEL AND K-NEAREST NEIGHBOR METHOD

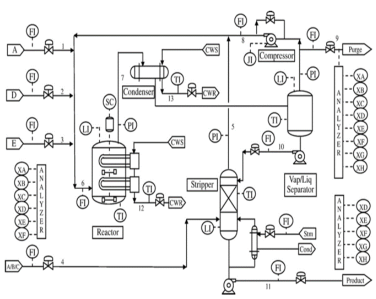
Norazwan Md Nor*, Mohd Azlan Hussain, Che Rosmani Che Hassan

Department of Chemical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

Article history
Received
17 November 2015
Received in revised form
25 February 2016
Accepted
28 March 2016

*Corresponding author
norazwan_mn@siswa.um.edu.my

Graphical abstract



Abstract

Effective fault monitoring, detection and diagnosis of chemical processes is important to ensure the consistency and high product quality, as well as the safety of the processes. Fault diagnosis problems can be considered as classification problems as these techniques have been proposed and greatly improved over the past few years. However, a chemical process is often characterized by large scale and non-linear behavior. When linear discriminant analysis is used for fault diagnosis in the system, a lot of incorrect diagnosis will occur. As solution, this paper presents a novel approach for feature extraction and classification framework in chemical process systems based on wavelet transformation and discriminant analysis. The proposed multi-scale kernel Fisher discriminant analysis (MSKFDA) method used the combination of kernel Fisher discriminant analysis (KFDA) and discrete wavelet transform (DWT) to improve the classification performance as compared to conventional approaches. A DWT is applied to extract the process dynamics at different scales by decomposed the data into multiple scales, analyzed by the KFDA and only dynamical characteristics with important information was reconstructed by inverse discrete wavelet transform (IDWT). Then, Gaussian mixture model (GMM) and K-nearest neighbor (KNN) method were individually applied for the fault classification using the output from the MSKFDA approach. These two classifiers are evaluated and compared based on their performance on the Tennessee Eastman process database. The proposed framework for GMM and KNN classifiers had achieved average classification accuracies of 84.72% and 82.00%, respectively, with the results show significant improvement over existing methods in fault detection and classification.

Keywords: Fault Diagnosis, Discrete Wavelet Transform, Fisher discriminant analysis, Gaussian mixture model, K nearest neighbor

1.0 INTRODUCTION

Effective fault monitoring, detection and diagnosis of chemical processes is important to ensure the consistency and high product quality, as well as the safety of the processes. Any abnormal process operation should be detected and the root causes to be diagnosed early to reduce the risk of catastrophic accident and economic losses before the corrective actions could move the plant back to normal operation condition.

Various methods of process fault monitoring, detection and diagnosis are developed based on signal processing and data classifications. In general, these fault detection and diagnosis methodologies are broadly divided into three types: quantitative model-based methods; qualitative model-based methods; and process history based methods [1].

The process history based or data-driven method concerns with the transformation of large amounts of data into a particular form of knowledge and distinctive representation. Availability of vast amounts of process data has encouraged researchers to develop and improve the data-driven-based and multivariable statistical process monitoring based methods to extract key process information. However, compared to fault detection, the problem within fault classification and diagnosis have not yet been properly solved, and still present important practical limitations that make this area an open for further research [2].

Fault diagnosis problems can be considered as classification problems as these techniques have been proposed and greatly improved over the past few years. For example, Bayesian classifier [3-5], Principal Component Analysis (PCA) [6-7], Fisher Discriminant Analysis (FDA) [8-9], Partial Least Squares (PLS) [10], Artificial Neural Networks (ANN) [11-12], Support Vector Machine (SVM) [13], and other techniques have been applied in fault classification problems. However, a chemical process is often characterized by large scale and non-linear behaviour. When linear FDA is used for fault diagnosis in non-linear system, a lot of incorrect diagnosis results will occur. As solution to deal with the nonlinear system, and to improve the classification ability, kernel-based FDA, called kernel FDA (KFDA), is introduced.

On the other hand, discrete wavelet transform (DWT) also is considered as an effective tool for signal processing and classification for fault detection and diagnosis scheme. Discrete wavelet analysis decomposes the high-frequency part further and adaptively selects relative frequency based on character of signal to obtain a better resolution for analysing process. This wavelet analysis could provide local feature in both time and frequency domains and has the feature of multi-scale, which enables wavelet analysis to distinguish the abrupt components of the signal.

In this paper, two different fault detection and diagnosis system are proposed. Process history data are extracted using DWT before analysed by KFDA. Then, two different classifiers are proposed to be combined with multi-scale KFDA approach: The GMM as a parametric classifier and KNN as a non-parametric one. KNN is chosen because of its simplicity and practicality. In the other hand, GMM is robust to variation and unbalanced number of involved samples for different classes of detection.

Gaussian mixed model (GMM) is one of the general tool for probability distribution function estimation. It has been widely applied in monitoring of various processes. However, GMM method and its literature study is more focused on fault detection while little work has been reported to address the more challenging fault diagnosis issue [14]. GMM assumes the data under modelling is generated via a probability density distribution. Through the use of expectation maximization (EM) method, we can simultaneously identify the optimum set of parameters of GMM in an iterative manner and approximate the data distribution and pattern similarity. The mathematical formula of GMM and the derivation are a bit lengthy. Refer the references for details [15].

K nearest neighbour (KNN) is a non-parametric classifier. This method classifies a sample of data according to the distance between this sample and some pre-labelled training samples. If most of the nearest samples to the unknown sample are from a specified class, the test sample will be assigned to that class. It used when train set embrace every possible faults included in the test data [16].

The paper is organized as follows. Section 2 presents the proposed multi-scale KFDA with different classifiers; GMM and KNN with application to Tennessee Eastman process case study. The results and discussion are included in Section 3 while Section 4 concludes the paper.

2.0 METHODOLOGY

2.1 Case Study

Tennessee Eastman (TE) process as described by Downs and Vogel [17] in Figure 1 was used as a case study. The process includes a total of 52 variables, as listed in Table 1, with 21 different faults were simulated for the process. The data set for the process and the details can be obtained from the Multi-scale Systems Research Laboratory [18]. To investigate the efficiency of multi-scale classification of KFDA with GMM and KNN, three classes of faulty data was simulated from the TEP simulator [8]. This three classes are corresponded to faults 4, 9, and 11 as listed in Table 2.

Fault 4 and 11 are selected because both faults are associated with reactor cooling water inlet temperature but with different type of fault (step change and random variation, respectively) while fault 9 is selected for its random variation fault but in different location (D feed temperature). Thus, misclassification is expected for these overlap dataset. The multi-scale classification of KFDA with GMM and KNN were programmed using MATLAB 7.1.

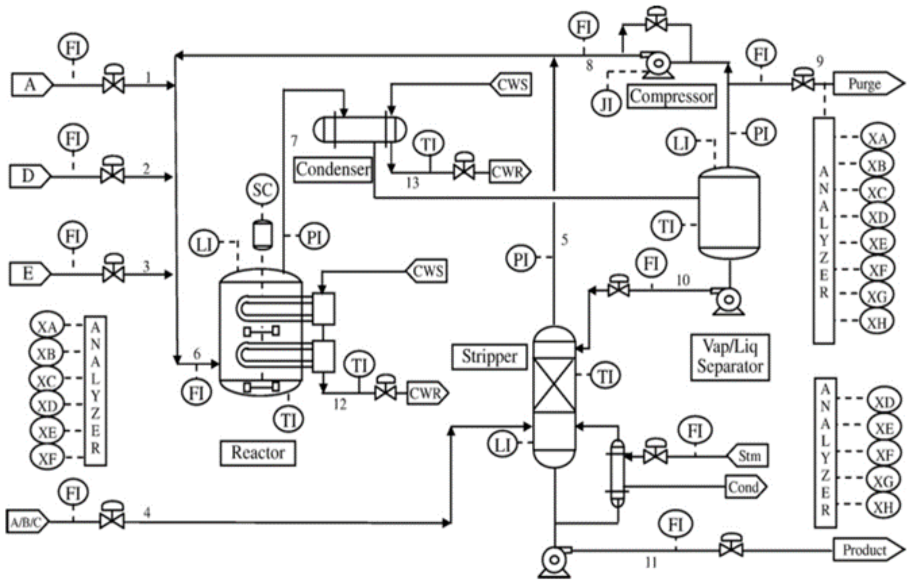


Figure 1 Tennessee Eastman process diagram

Table 1 Measured and Manipulated Variables of the TE Process

Variable	Description	Variable	Description
XMEAS(1)	A feed Stream 1	XMEAS(27)	Reactor feed component E
XMEAS(2)	D feed Stream 2	XMEAS(28)	Reactor feed component F
XMEAS(3)	E feed Stream 3	XMEAS(29)	Purge component A
XMEAS(4)	Total feed Stream 4	XMEAS(30)	Purge component B
XMEAS(5)	Recycle flow	XMEAS(31)	Purge component C
XMEAS(6)	Reactor feed rate	XMEAS(32)	Purge component D
XMEAS(7)	Reactor pressure	XMEAS(33)	Purge component E
XMEAS(8)	Reactor level	XMEAS(34)	Purge component F
XMEAS(9)	Reactor temperature	XMEAS(35)	Purge component G
XMEAS(10)	Purge rate	XMEAS(36)	Purge component H
XMEAS(11)	Separator temperature	XMEAS(37)	Product component D
XMEAS(12)	Separator level	XMEAS(38)	Product component E
XMEAS(13)	Separator pressure	XMEAS(39)	Product component F
XMEAS(14)	Separator underflow	XMEAS(40)	Product component G
XMEAS(15)	Stripper level	XMEAS(41)	Product component H
XMEAS(16)	Stripper pressure	XMV(1)	D feed flow Stream 2
XMEAS(17)	Stripper underflow	XMV(2)	E feed flow Stream 3
XMEAS(18)	Stripper temperature	XMV(3)	A feed flow Stream 1
XMEAS(19)	Stripper steam flow	XMV(4)	Total feed flow Stream 4
XMEAS(20)	Compressor work	XMV(5)	Compressor recycle valve
XMEAS(21)	Reactor cooling water outlet temp.	XMV(6)	Purge valve
XMEAS(22)	Separator cooling water outlet temp.	XMV(7)	Separator product liquid flow
XMEAS(23)	Reactor feed component A	XMV(8)	Stripper product liquid flow
XMEAS(24)	Reactor feed component B	XMV(9)	Stripper steam valve
XMEAS(25)	Reactor feed component C	XMV(10)	Reactor cooling water flow
XMEAS(26)	Reactor feed component D	XMV(11)	Condenser cooling water flow

Table 2 Selected three fault classes

Class	Fault	Fault description	Type
1	Fault 4	Reactor cooling water inlet temperature	Step change
2	Fault 9	D feed temperature	Random variation
3	Fault 11	Reactor cooling water inlet temperature	Random variation

2.2 Multi-Scale Kernel FDA Feature Extraction

The detailed in multi-scale kernel FDA (MSKFDA) application procedure of fault diagnosis were discussed briefly in this section. The normalization of the data was done with each of the variables were linearly scaled to the range of [0, 1]. It is important to scale the data before applying the multivariable methods to avoid the variables with greater numerical range dominating those with smaller numerical range. Then, each of the m variables is first decomposed individually by applying discrete wavelet transformation (DWT).

Then, the kernel FDA is performed on the wavelet coefficients for each selected scale. Appropriate numbers of component loading vectors are retained and the wavelet coefficients are reconstructed at each selected scale. In this work four scales (s=4) are used for discrete wavelet transformation (DWT) of the original signal. After that, the wavelet coefficients larger than a selected threshold corresponding to a significant event are retained. The variables consisting of deterministic components are reconstructed from the retained wavelet coefficients through inverse discrete wavelet transformation (IDWT) and the loadings of the extracted deterministic components are computed.

The new observations are projected into lower dimensional subspace. This subspaces measures the systematic or state variations occurring in the process. KFDA is used to search the optimal one-dimensional discriminant direction between the fault data and the normal data. The outputs of this step were used as an input for GMM and KNN classification step individually.

After reducing the original data dimension using KFDA, which finds a set of orthogonal discriminant vectors, a GMM is constructed to describe the faulty patterns. In the GMM, each local model represents one fault. In systems with complex patterns, it is difficult to characterize each fault pattern using only a single Gaussian model. This problem is overcome by increasing the number of Gaussians used to represent a fault. The mean and covariance obtained from each faulty dataset are taken as the initial parameters of the corresponding local Gaussian model. Then, the model parameters are fine-tuned through the training procedure.

For KNN application, the test sample probability is adaptively estimated without any prior assumption, except k which will be chosen such that the best results obtain on train data. For test sample classification, all the train data points must be saves and the distances between the test sample and all the training samples must be calculated and sorted. Then, the k that is nearest to the neighbours are selected to make the final decision. The efficiency of the multi-scale KFDA-GMM and KFDA-KNN-based fault detection and diagnosis system were validated by comparing it to the traditional FDA and KFDA method. The classification results are shown in Table 3 and Table 4.

Table 3 Comparison among Diagnosis Accuracy

	Diagnosis accuracy (%)	
	Multi-scale KFDA-GMM	Multi-scale KFDA-KNN
Fault 4	89.48	82.75
Fault 9	72.81	76.50
Fault 11	91.88	86.75
Average	84.72	82.00

Table 4 Diagnosis Accuracy Using Different Approaches for Selected Faults

	Multi-scale KFDA-GMM	Multi-scale KFDA-KNN	FDA [8]	KFDA-Bayes [18]	PCA-KNN [18]
Diagnosis accuracy (%)	84.72	82.00	62	48.05	47.00

3.0 RESULTS AND DISCUSSIONS

In this paper the multi-class classification problem was studied with the implementation of multi-scale KFDA with the GMM and KNN classifiers. The proposed method was implemented to detect the faults of the TE process.

Figure 2 shows a normal (no fault) and Fault 4 decomposition coefficients for approximation and details after transformation by DWT approach for level 1, 3 and 5. From the figures, the approximation coefficient for Fault 4 data of the transformed signal clearly shows a significant differences in amplitude of the plot compared to the normal data. This implied that disturbance or faulty event has occurred in this data. The detailed coefficient for level 5 (d5) also shows some distinctive characteristics to differ the normal and faulty condition in the database. After fault data are decomposed by DWT wavelet analysis, KFDA is performed on these multi-scaled fault data, which offers important supplemental classification information to KFDA.

Figure 3 and 4 show the selected fault classification based on FDA and multi-scale KFDA projection for the case study database, respectively. The data consists of different type of faults, which are Fault 4, Fault 9 and Fault 11. From Figure 3, FDA is able to classify the Fault 4 and Fault 9 data, but failed to separate and classify Fault 11. The FDA method was unable to distinguish the Fault 11 since the separation between-classes is not large enough while the distribution within-classes also was quite large. This is because all variables are in a same level without proper variable weighting and the data sets are masked with irrelevant information.

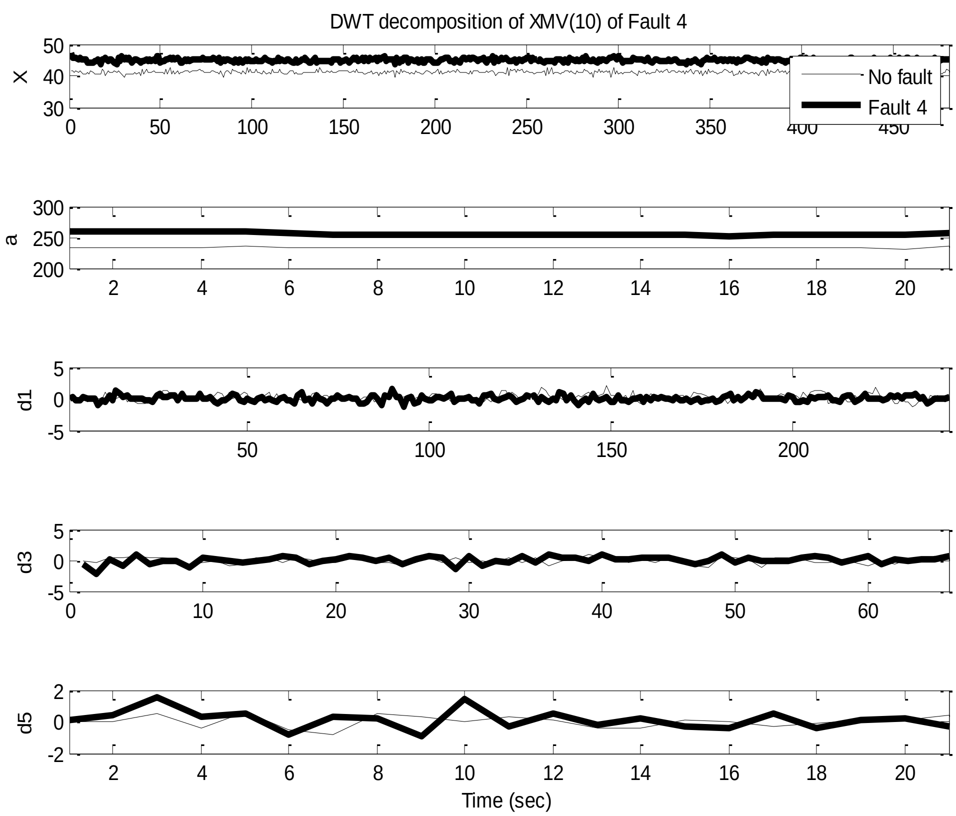


Figure 2 Variable XMV10 of normal and Fault 4 approximate and detail coefficients; approximate level (a), first decomposed level (d1), third decomposed level (d3), and fifth decomposed level (d5)

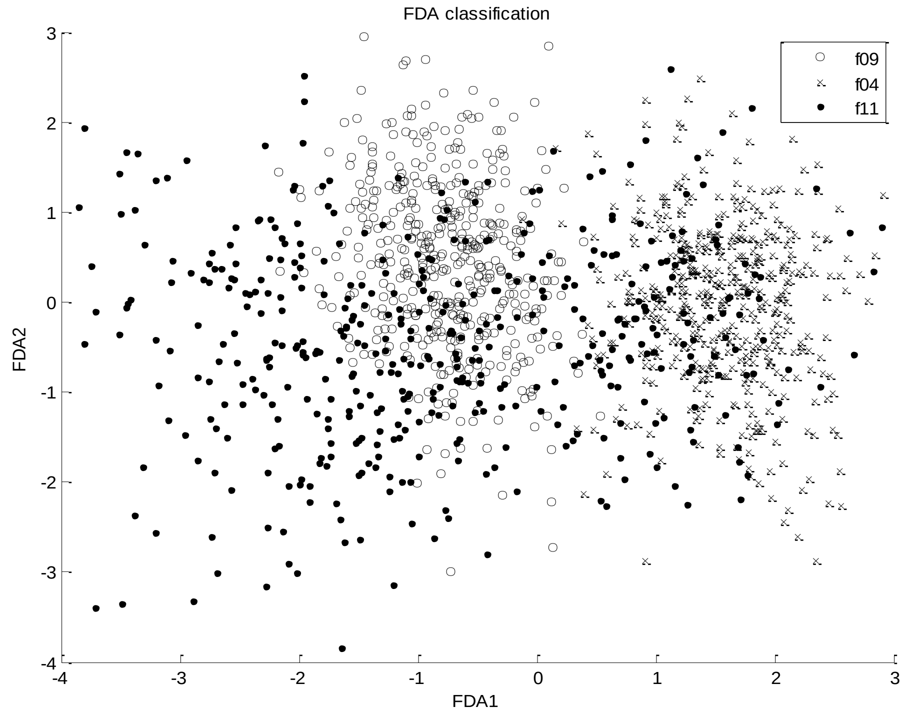


Figure 3 Selected fault classification projection using normal FDA

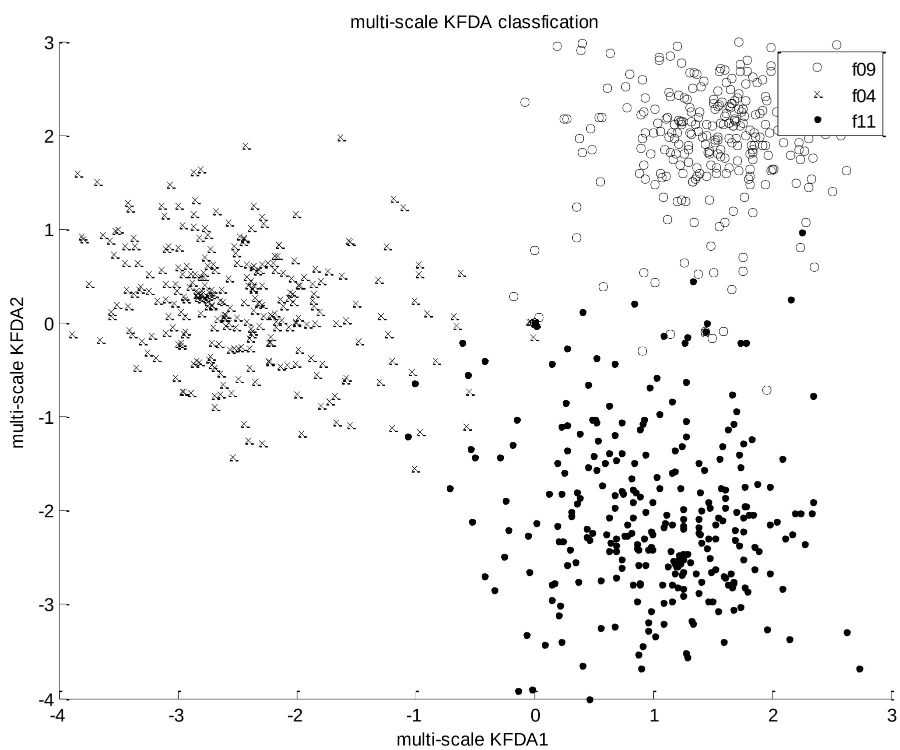


Figure 4 Selected fault classification projection using multi-scale KFDA

The integration of DWT with KFDA improved the features extraction that are relevant to the abnormal operation in both time and frequency domain and lead to better classification. Figure 4 shows the projection data onto the first two multi-scale KFDA vectors. From the figure, there is a large separation in between-classes distribution while the scattering of within-classes distance is shorten compared to FDA, which shows that the discriminant power of multi-scale KFDA is better than FDA.

The classification accuracy is evaluated by two types of classifiers: Gaussian mixture method (GMM) and K-Nearest Neighbor (KNN). Table 3 presents the validation results obtained using the combination of DWT and KFDA for feature extraction and classification. From the table, classification and diagnosis performance of GMM is found to be the best averagely with 84.72% accuracy compared to KNN classifier with 82%. However, KNN classifier also shows decent performance, especially in Fault 9 classification. This is due to the method's advantage in determining the boundaries for Fault 9 via a k coefficient selection method.

A summary of the classification results for FDA, KFDA-Bayesian, PCA-KNN, MSKFDA-GMM and KFDA-KNN is tabulated in Table 4. The table lists the average diagnosis accuracy rate by utilizing the selected faults of TE process dataset. Comparing the diagnosis accuracy percentage presented in the table, it can be observed that the diagnostic accuracy percentage for MSKFDA-GMM and MSKFDA-KNN are significantly higher than the traditional FDA method. On average, the multi-scale KFDA classifications (MSKFDA-GMM and MSKFDA-KNN) also produced higher diagnosis accuracy than KFDA-Bayesian and PCA-KNN approaches. Multi-scale KFDA-GMM has the highest average of classification accuracy (84.72%) compare with MSKFDA-KNN (82.00%), FDA (62%), KFDA-Bayes (48.05%) and PCA-KNN (47.00%).

Through this case study, results show that the proposed multi-scale KFDA with GMM and KNN have superior capability in diagnosing faults when compared to traditional FDA methods, its modified method of KFDA-Bayes and methods of PCA-KNN. The use of the multi-scale KFDA classification method can help operators immediately rectify the process when a fault occurs. This is particularly important for the process industries such as chemical processes which may involve plant safety problems.

4.0 CONCLUSIONS

In this paper, we have presented the application of multi-scale KFDA with GMM and KNN classifier-based for fault diagnosis system. The effectiveness of multi-scale KFDA is demonstrated with the help of fault classification of TE process. From the results, the combination of multi-scale feature extraction using DWT with KFDA and GMM method for classification proved the proposed method suitable for implementation in fault diagnosis system of chemical processes. The data discrimination and fault detection based on multi-scale KFDA methodology enhanced the diagnosis proficiency by taking into consideration the multi-scale information compared to other methods that considered only single scale nature. Moreover, it can provide a better separation of the deterministic features and improve the features extraction that are relevant to a faulty situation from both time and frequency domain aspects. By comparing performance of the classification accuracy of FDA, PCA-KNN, KFDA-Bayes and multi-scale KFDA (MSKFDA-GMM and MSKFDA-KNN) for handling the TE process database, the results showed that the performance of the classifier by multi-scale KFDA-GMM and KFDA-KNN were better than the others. Further research can extend the proposed method to a batch process which is more complex in contrast to a continuous process.

Acknowledgement

The authors are thankful to Ministry of Higher Education Malaysia for their ASTS (SLAB) scholarship and financial support toward the project.

References

[1] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin. 2003. A Review of Process Fault Detection and Diagnosis Part III: Process History Based Methods. *Computers & Chemical Engineering*. 27: 327-346.

[2] I. Yélamos, G. Escudero, M. M. Graells, L. Puigjaner, I. Yelamos, and G. Escurado. 2009. Performance assessment of a Novel Fault Diagnosis System Based on Support Vector Machines. *Computers & Chemical Engineering.* 33(1): 244-255.

[3] S. Verron, T. Tiplica, and A. Kobi. 2006. Fault Diagnosis with Bayesian Networks : Application to the Tennessee Eastman Process. *IEEE International Conference on Industrial Technology*. 98-103.

[4] J. Liu and D.-S. Chen. 2009. Fault Detection and Identification Using Modified Bayesian Classification on PCA Subspace. *Industrial and Engineering Chemistry Research.* 48(6): 3059-3077.

[5] A. D. Sagale and S. G. Kale. 2014. Hybrid Model For Intrusion Detection Using Naive Bayesian And Support Vector Machine. *International Journal of Computing and Technology.* 1(3): 56-59.

[6] C. K. Lau, K. Ghosh, M. A. Hussain, and C. R. C. Hassan. 2013. Fault Diagnosis of Tennessee Eastman Process with Multi-scale PCA and ANFIS. *Chemometrics and Intelligence Laboratory Systems.* 120: 1-14.

[7] K. Feng, Z. Jiang, W. He, and B. Ma. 2011. A Recognition and Novelty Detection Approach Based on Curvelet Transform, Nonlinear PCA and SVM with Application to Indicator Diagram diagnosis. *Expert Systems with Applications.* 38(10): 12721-12729.

[8] L. H. Chiang, M. E. Kotanchek, and A. K. Kordon. 2004. Fault Diagnosis based on Fisher discriminant Analysis and Support Vector Machines. *Computer & Chemical Engineering.* 28: 1389-1401.

[9] J. Li and P. Cui. 2009. Improved Kernel Fisher Discriminant Analysis for Fault Diagnosis. *Expert Systems with Applications.* 36(2): 1423-1432.

[10] H. W. Lee, M. W. Lee, and J. M. Park. 2009. Multi-scale Extension of PLS Algorithm for Advanced On-line Process Monitoring. 2009. *Chemometrics and Intelligence Laboratory Systems.* 98(2): 201-212.

[11] S. H. Fourie and P. De Vaal. 2000. Advanced Process Monitoring Using an On-lne Non-linear Multiscale Principal Component Analysis Methodology. *Computers & Chemical Engineering*. 24: 755-760.

[12] A. Hajnayeb, A. Ghasemloonia, S. E. Khadem, and M. H. Moradi. 2011. Application and Comparison of an ANN-based Feature Selection Method and the Genetic Algorithm in Gearbox Fault Diagnosis. *Expert Systems with Applications.* 38(8): 10205-10209.

[13] H. Keskes, A. Braham, and Z. Lachiri. 2013. Broken Rotor Bar Diagnosis in Induction Machines Through Stationary Wavelet Packet Transform and Multiclass Wavelet SVM. *Electric Power Systems Research*. 97: 151-157.

[14] J. Yu. 2013. A New Fault Diagnosis Method of Multimode Processes Using Bayesian Inference Based Gaussian Mixture Contribution Decomposition. *Engineering Applications of Artificial Intelligence*. 26: 456-466.

[15] T. Chen and J. Zhang. 2010. On-line Multivariate Statistical Monitoring of Batch Processes Using Gaussian Mixture Model. *Computers & Chemical Engineering*. 34: 500-507.

[16] M. H. Gharavian, F. Almas Ganj, A. R. Ohadi, and H. Heidari Bafroui. 2013. Comparison of FDA-based and PCA-based Features in Fault Diagnosis of Automobile Gearboxes. *Neurocomputing*. 121: 150-159.

[17] J. J. Downs and E. F. Vogel. 1993. A Plant-wide Industrial Process Control Problem. *Computers & Chemical Engineering*. 17: 245-255.

[18] Y. Liu, L. Ye, P. Zheng, X. Shi, B. Hu, and J. Liang. 2010. Multiscale Classification and Its Application to Process Monitoring. *Journal of Zhejiang University-SCIENCE C (Computers & Electronics)*. 11(6): 425-434.