Journal of the Korea Academia-Industrial cooperation Society
Vol. 17, No. 7 pp. 374-379, 2016

http://dx.doi.org/10.5762/KAIS.2016.17.7.374
ISSN 1975-4701 / eISSN 2288-4688

Empirical Process Monitoring Via On-line Analysis of Complex Process Measurement Data

Hyun-Woo Cho
Department of Industrial and Management Engineering, Daegu University

복잡한 공정 측정 데이터의 실시간 분석을 통한 공정 감시

조현우
대구대학교 산업경영공학과

Abstract On-line process monitoring schemes are designed to give early warnings of process faults. In the artificial intelligence and machine learning fields, reliable approaches have been utilized, such as kernel-based nonlinear techniques. This work presents a kernel-based empirical monitoring scheme with a small sample problem. The measurement data of normal operations are easy to collect, whereas special events or faults data are difficult to collect. In such situations, noise filtering techniques can be helpful in enhancing the process monitoring performance. This can be achieved by the preprocessing of raw process data and eliminating unwanted variations of data. In this work, the performance of several monitoring schemes was demonstrated using three-dimensional batch process data. The results showed that the monitoring performance was improved significantly in terms of the detection success rate.

요 약 실시간 공정 감시 체계는 공정에서 발생된 이상상황을 조기에 감지하도록 설계된다. 공정 감시를 위해 패턴 인식과 머신 러닝 분야에서 비선형 방법론을 비롯한 다양한 방법들이 사용되고 있는 상황이다. 본 연구에서는 데이터의 불균형이 존재하는 공정으로부터 얻은 데이터에 기반한 공정 감시 모델을 제시한다. 정상 조업영역의 과거 데이터는 쉽게 얻을 수 있지만 특정 이상 상황에 대한 이상 데이터는 상대적으로 많지 않다. 이러한 상황에서는 필터링 방법의 활용이 공정 감시 성능 향상에 도움이 될 수 있는데 이는 데이터 모델링에 필요 없는 데이터 산포를 제거하거나 필터링함으로써 달성된다. 본 연구에서는 다양한 선형 및 비선형 방법에 기반한 모니터링 모델들의 감시 성능을 회분식 공정 데이터를 활용하여 비교 검증하였으며 이를 통해 향상된 감시 성능을 얻을 수 있었다.

Keywords : Monitoring, Nonlinear methods, Noise filtering, Process data, Quality improvement

1. Introduction

Process monitoring methods have been extensively studied as one of essential topics of statistical process control. This is largely due to the fact that most of manufacturing or production processes are susceptible to unexpected abnormalities such as process faults, breakdowns and malfunctions[1]. Unfortunately, these special or uncommon events are apt to give a negative impact on final product quality. Thus the main goal of process monitoring is to detect the occurrence of a fault[2]. Among industrial processes, batch process monitoring is quite difficult to implement because these processes have challenging issues like nonlinear process behavior, finite duration of operation time, etc. A batch process operation includes a set of tasks like

This research was supported by the Daegu University Research Grant, 2012.
*Corresponding Author : Hyun-Woo Cho (Daegu University)
Tel: +82-2-850-6540 email: hwcho@daegu.ac.kr
Received April 21, 2016 Revised June 8, 2016
Accepted July 7, 2016 Published July 31, 2016

charging ingredients, processing them under controlled conditions, and discharging final product[3].

Recent reports showed that the development of monitoring schemes depends on the use of multivariate statistical methods due to the availability of real-time and historical process data. Nonlinear kernel techniques based on Fisher discriminant analysis, principal component analysis, and partial least squares have been extensively used in practical monitoring problems[4-6]. Such empirical models are preferred because of widespread sensor and data measurement technology in production processes.

When monitoring a process using process measurement data, data sets for specific classes may be under-sampled or not enough to build empirical monitoring models. Frequently, measurement data of normal operations without faults are easy to gather, but the measurement data of out-of-control states or faults are expensive to collect. Such an unbalanced measurement data can be covered by adopting the appropriate method which depends mainly on normal data. Support vector data description (SVDD) is quite helpful in describing samples of high density areas of normal operating conditions. This is able to adapt to the real shape of normal samples and seeks to find flexible boundary with a minimum volume.

This work presents the utilization of an empirical model-based quality monitoring approach to batch processes. The monitoring scheme is combined with nonlinear representation of raw process data. In addition, preprocessing or filtering schemes are implemented for better monitoring performance. The nonlinear representation of finite and three-way batch data is applied. Prior to building empirical monitoring models, filtering of the data is performed to trim the irrelevant information of the process data. In this work, several monitoring schemes are evaluated, in which two filtering techniques are also tested. Due to the characteristics of batch data, the selection of estimation approaches for future observation is also discussed with the comparison of monitoring performance. The performance of the proposed process monitoring schemes is demonstrated using batch process data. This paper is organized as follows: an introduction of multivariate statistical techniques followed by monitoring performance comparison using batch process data. Finally, concluding remarks are given.

2. Method

A linear version of principal component representation (PCA), is used to decompose correlated original variables into an uncorrelated set of linear principal components. In most cases, only several components are enough to explain the data variability. It seeks to decompose the data matrix **X** into the sum of the outer products of score vectors (**t**) and loading vectors (**p**) plus a residual matrix (**E**):

$$\mathbf{X} = \sum_{r=1}^{R} \mathbf{t}_r \mathbf{p}_r^{\mathrm{T}} + \mathbf{E} \qquad (1)$$

On the other hand, the goal of linear Fisher discriminant analysis (FDA) is to find certain directions in original variables, along which hidden groups are discriminated as clearly as possible [6]. As an extension of linear FDA, nonlinear kernel FDA (KFDA) executes linear FDA in the feature space F. As a result, the discriminant weight vector is determined by maximizing between-class scatter matrix while minimizing total scatter matrix, which are defined in feature space:

$$\mathbf{S}_b^{\Phi} = \frac{1}{M}\sum_{i=1}^{C} c_i (\mathbf{m}_i^{\Phi} - \mathbf{m}^{\Phi})(\mathbf{m}_i^{\Phi} - \mathbf{m}^{\Phi})^{\mathrm{T}} \qquad (2)$$

$$\mathbf{S}_b^{\Phi} = \frac{1}{M}\sum_{i=1}^{C} c_i (\mathbf{m}_i^{\Phi} - \mathbf{m}^{\Phi})(\mathbf{m}_i^{\Phi} - \mathbf{m}^{\Phi})^{\mathrm{T}} \qquad (3)$$

By maximizing following Fisher criterion the optimal discriminant vectors can be obtained.

$$J^{\Phi}(\boldsymbol{\psi}) = \frac{\boldsymbol{\psi}^{\mathrm{T}} \mathbf{S}_b^{\Phi} \boldsymbol{\psi}}{\boldsymbol{\psi}^{\mathrm{T}} \mathbf{S}_t^{\Phi} \boldsymbol{\psi}} \qquad (4)$$

As one of filtering techniques, orthogonal signal correction (OSC) is a PLS-based solution which removes unwanted variation. In this work, the OSC method is used so that a coding is introduced where each column in **Y** matrix contains information about class memberships of samples. The binary **Y** matrix has a structure where each row sums to unity. The first step of an OSC is to calculate the first PC score vector, and actual correction vector is produced:

$$\mathbf{t}^* = \{\mathbf{I} - \mathbf{Y}(\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T\}\mathbf{t}. \quad (5)$$

Then PLS weight vector **w** is computed such that $\mathbf{Xw}=\mathbf{t}^*$, which is followed by the calculation of a new score vector $\mathbf{t}=\mathbf{Xw}$. Finally, a loading vector **p** is computed and the correction term $\mathbf{tp}^T$ is subtracted from **X**, giving the residual. The next components can be calculated in a similar way. An alternative approach is discriminant partial least squares (PLS), which is the classical PLS algorithm applied to classification problems.

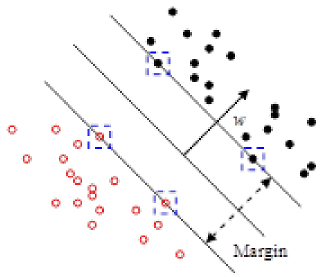

Fig. 1. A simple representation of SVM

One common way to use PLS in classification problems is to introduce a coding in which each column in **Y** contains information about the class memberships of samples [7]. Support vector machine (SVM) is a training algorithm to learn classification and regression rules of patterns from raw data. It is basically a linear method that is nonlinearly mapped from the input data space. In a real computation, input data are first mapped into high dimensional feature space. As shown in Fig. 1, in the feature space optimal decision function is obtained having a maximum margin, in which the decision function satisfies inequality constraints

$$y_i(\mathbf{w}\Phi(\mathbf{x}_i)+b)-1 \geq 0 \ \forall_i \quad (6)$$

Based on the optimal decision function non-separable problems are solved by a dual problem:

$$L_d = \sum \alpha_i - 1/2 \sum \alpha_i \alpha_j y_i y_j \Phi(\mathbf{x}_i)\Phi(\mathbf{x}_j) \quad (7)$$

Training SVM is to find α_i, b, and support vectors with given kernel function parameters and C.

Support vector data description (SVDD) is a one-class classification method to envelop samples or objects within a high-dimensional space with the volume as small as possible[4]. It is necessary to find μ and R that has the minimum volume of hyper-sphere containing all samples. In the end the minimization problem can be denoted as follows:

$$F(R,\boldsymbol{\mu}) = R^2 + C\sum_{i=1}^{I}\xi_i \quad (8)$$

Here, the parameter C represents the trade-off between the volume of the sphere and the number of samples outside it, and thus it should be minimized with the following constraints:

$$\|\mathbf{x}_i - \boldsymbol{\mu}\|^2 \leq R^2 + \xi_i, \xi_i \geq 0 \quad (9)$$

To obtain more flexible boundaries, inner products of samples are replaced by a kernel function.

3. Results

This part demonstrates the monitoring performance of the proposed scheme which utilizes nonlinear kernel method combined with preprocessing techniques for three dimensional process data. The test process is a polyvinyl chloride batch process. Here a straight resin polymerization process is initiated by vinyl chloride monomer. This process contains a polymerization reactor, reflux condenser, agitator, and cooling jacket. Eleven process variables are automatically measured on-line. A total of 170 batches are used in building nonlinear kernel monitoring models.

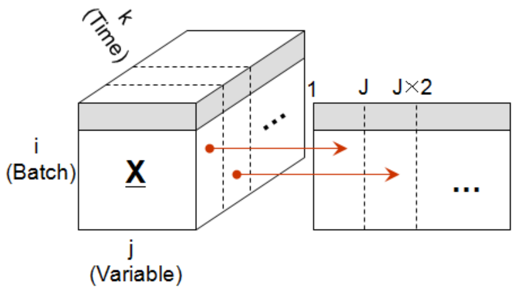

Fig. 2. Batch Data and future values

In order to overcome the limitation of three-way characteristics of batch process data, as stated before, future observations of a new batch, i.e., unmeasured data parts of current batch operation, should be estimated appropriately as shown in Figure 2. It is due to the fact that a new batch operation is not complete until the end of its operation.

A filtering or preprocessing of process data is performed prior to main model building in order to obtain monitoring results using several monitoring schemes, To evaluate difference the monitoring performance based on several multivariate projection methods, the two representation techniques were applied. In addition, two filtering methods for the test process are tested. That is, discriminant partial least squares and orthogonal signal correction are considered.

Table 1. Monitoring rates (%) with library values

	Model1	Model2	Model3	Model4
Ft1	84	89	93	95
Ft2	89	90	92	95
Ft3	84	89	93	94
Ft4	83	86	92	97
Ft5	87	91	92	93
Ft6	85	89	93	96
Ft7	78	83	84	89
Avg.	84	88	91	94

The selection of a kernel function in implementing monitoring models using different representation and preprocessing techniques was evaluated with the test of various kernel functions. In this work, second-order polynomial kernel was chosen to capture nonlinearity of the data. The monitoring results for the seven test batches of Fault1 (denoted as Ft1) through Fault7 (Ft7) are summarized in Table 1. As shown in Table 1 monitoring accuracy values (%), i.e., percentage of detection success rate of faults, are listed to evaluate the monitoring performance of four monitoring schemes denoted as Model1 through Model4. Here monitoring accuracy is the proportion of the observations correctly detected.

As shown in Table 1, Model1 denotes the monitoring scheme of using discriminant partial least squares (DPLS), kernel PCA (KPCA), and SVDD. DPLS is a PLS regression model for the discrimination of different data clusters or groups, and KPCA is a nonlinear kernel version of linear PCA. In addition, Model2 indicates the Model1 with the use of OSC instead of DPLS. Similar to the relationship between Model1 and Model2, on the other hand, Model3 differs from Model1 in that it utilizes KFDA rather than KPCA. The only difference of Model4 is the use of OSC instead of DPLS of Model3.

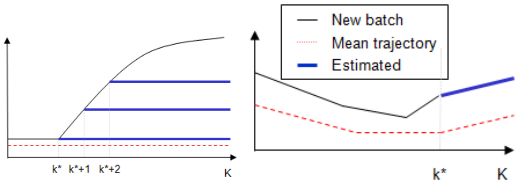

Fig. 3. Current deviation estimation approach

As reported in table 1, the Model4 monitoring scheme showed the best monitoring performance in that it yielded the highest monitoring accuracy for all test batches. The Model4 monitoring method produced the best monitoring performance, i.e., average monitoring success rate of 94.1%. Meanwhile, it is observed that the average monitoring success rate values of Model1, Model2, and Model3 are 84, 88, and 91, respectively.

It should be noted that the overall monitoring performance of using KFDA, i.e., Model3 and Model4, outperforms those of using KPCA, i.e., Model1 and Model2, irrespective of preprocessing techniques used. Consequently, using KFDA monitoring methods has significantly improved the monitoring performance for this test process. It may be meaningful to operating personnel, who have to take remedial actions using the monitoring results.

As executed in Table 1, monitoring results are obtained using the same monitoring models such as Model1 through Model4 monitoring schemes, which are given in Table 2 and Table 3. On the other hand, the difference between Table 1 and these two tables is which estimation approaches to use in estimating future observations of the batch process.

Table 2. Monitoring rates (%) with PCA values

	Model1	Model2	Model3	Model4
Ft1	83	88	92	95
Ft2	87	89	91	93
Ft3	87	88	90	92
Ft4	88	86	91	96
Ft5	86	87	91	93
Ft6	89	90	92	94
Ft7	79	82	83	85
Avg.	86	87	90	93

Table 3. Monitoring rates (%) with current deviation approach

	Model1	Model2	Model3	Model4
Ft1	79	82	87	90
Ft2	86	91	91	92
Ft3	83	85	87	89
Ft4	84	88	89	91
Ft5	82	86	89	91
Ft6	83	83	86	93
Ft7	79	81	84	83
Avg.	82	85	88	90

Specifically, results of Table 1 were obtained using the future value estimation of fault library approach [4]. On the other hand, PCA projection-based estimation method was used for Table 2 whilst current deviation method was applied to produce the monitoring results in Table 3. For more information about the current deviation approach this is illustrated as shown in Figure 3. As shown in Table 2 and Table 3, overall observations of monitoring results are quite similar to Table 1 in that the Model4 produced the best monitoring accuracy in all the faults tested.

As shown in Table 2, for example, the Model4 yielded maximum success rates. As analyzed in Table 1, in case of comparing the effect of different estimation methods, fault library method of Table 1 outperformed performance of Table 2 and Table 3. In particular, the difference between PCA-based estimation method and current deviation method can be seen by comparing these tables. On the other hand, the use of KFDA improved monitoring performance significantly when compared to using KPCA.

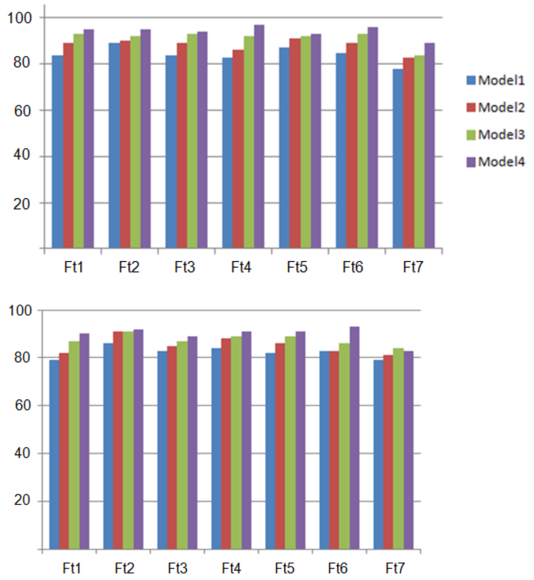

Fig. 4. Comparison of monitoring results

As shown in Figure 4, relative performance of the overall results can be easily distinguished graphically. Here upper one shows the results of using library values while lower of current deviation approach.

4. Conclusion

This work proposed the use of nonlinear kernel representation of complex process measurement data. It also utilized the filtering techniques in order to provide reliable quality monitoring performance for batch processes. In this work a case study on the batch process has been executed with different monitoring schemes. It has shown that the use of appropriate nonlinear and filtering techniques produced reliable monitoring results on the test process. It turned out that KFDA outperformed KPCA in monitoring nonlinear complex processes. In terms of future values, the three estimation methods were tested, and the fault library approach showed the best performance in Model4. Taking into accounts of frequent use of batch processes, kernel-based nonlinear technique is quite helpful to make monitoring decision in an on-line basis.

References

[1] S. J. Qin, "Statistical process monitoring: basics and beyond", Journal of Chemometrics, 17, pp. 480‐502, 2003.
DOI: http://dx.doi.org/10.1002/cem.800

[2] S. Bersimis, S. Psarakis, and J. Panaretos, "Multivariate statistical process control charts: an overview", Quality and Reliability Engineering International, 23, pp. 517‐543, 2007.
DOI: http://dx.doi.org/10.1002/qre.829

[3] X. Meng, A. J. Morris, and E. B. Martin, "On-line monitoring of batch processes using PARAFAC representation", Journal of Chemometrics, 17, pp. 65-81, 2003.
DOI: http://dx.doi.org/10.1002/cem.776

[4] V. A. Sotiris, P. W. Tse, and M. G. Pecht, "Anomaly detection through a bayesian support vector machine", IEEE Transactions on Reliability, 59, pp. 277-286, 2010.
DOI: http://dx.doi.org/10.1109/TR.2010.2048740

[5] R. Lombardo, J.-F. Durand, and A. P. Leone, "Multivariate additive PLS spline boosting in agro-chemistry studies", Current Analytical Chemistry, 8, pp. 236-253, 2012.
DOI: http://dx.doi.org/10.2174/157341112800392661

[6] L. H. Chiang, E. L. Russell, and R. D. Braatz, "Fault monitoring in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis", Chemometrics and Intelligent Laboratory Systems, 50, pp. 243-252, 2000.
DOI: http://dx.doi.org/10.1016/S0169-7439(99)00061-1

[7] J. A. Westerhuis, S. de Jong, and A. K. Smilde, "Direct orthogonal signal correction", Chemometrics and Intelligent Laboratory Systems, 56, pp. 13-25, 2001.
DOI: http://dx.doi.org/10.1016/S0169-7439(01)00102-2

Hyun-Woo Cho [Regular member]

- Aug. 2003 : POSTECH., Industrial Eng., PhD
- Aug. 2003 ~ Aug. 2007 : GIT/UT, Research Associate
- Sep. 2007 ~ Feb. 2011 : SEC, Senior Engineer
- Mar. 2011 ~ Current : Daegu Univ., Dept. of Industrial. & Management Eng., Professor

<Research Interests>
Intelligent Process Monitoring, Data Mining