Journal of Process Control 52 (2017) 45–56

Contents lists available at ScienceDirect

Journal of Process Control

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

Process control of time-varying systems using parameter-less self-organizing maps

Young Jae Choung [a], Jihoon Kang [b], Seoung Bum Kim [a,*]

[a] School of Industrial Management Engineering, Korea University, 145 Anam-Ro, Seongbuk-Gu, Seoul 02841, Republic of Korea
[b] Algorithm Research Team, Samsung SDS, 689 Umyeon-Dong, Seocho-Gu, Seoul 06765, Republic of Korea

ARTICLE INFO

Article history:
Received 11 July 2016
Received in revised form
19 December 2016
Accepted 7 February 2017

Keywords:
Multivariate process control
Control chart
Self-organizing map
Time-varying process
Data mining
Machine learning

ABSTRACT

Traditional control charts, such as Hotelling's T^2, are effective in detecting abnormal patterns. However, most control charts do not take into account a time-varying property in a process. In the present study, we propose a parameter-less self-organizing map-based control chart that can handle a situation in which changes occur in the distribution or parameter of the target observations. The control limits of the proposed chart are determined by estimating the empirical level of significance on the percentile using the bootstrap method. Experimental results obtained by using simulated data and actual process data from the manufacturing process for a thin-film transistor-liquid crystal display demonstrate the effectiveness and usefulness of the proposed algorithm.

1. Introduction

Statistical process control (SPC) is a popular technique for process monitoring. A control chart is a widely used tool in SPC to detect the abnormal status of a process and maintain control of the process [1]. The Shewart control chart is the most popular of these charts. It is used to efficiently monitor the quality of a single process variable when the process data are normally distributed. However, univariate control charts like the Shewart may not be appropriate for modern manufacturing systems in which a process may simultaneously affect a number of correlated quality variables. Various multivariate control charts have been proposed to achieve this simultaneous monitoring of correlated variables. One of the most widely used is Hotelling's T^2 control charts that assume that multivariate normal distribution governs the underlying distribution of the process observations [2,3]. However, the actual distribution is usually unknown and difficult to estimate accurately, especially when the number of samples is not sufficiently large enough to approximate the asymptotic distribution [4]. Further, many modern manufacturing systems exhibit nonlinearity because of their complicated processes. In such cases, the unregulated use of a linear modeling approach such as a T^2 chart may not be effective [5].

Numerous monitoring methods have been developed to address the limitations of traditional control charts. Nonlinear principal component analysis (PCA) methods were proposed to address the nonlinearity of modern manufacturing systems [6–8]. Kramer [6] proposed a nonlinear PCA technique that used an auto-associative neural network. Dong and McAvoy [7] developed a nonlinear PCA method based on the principal curve and the neural network. Hiden et al. [8] addressed the same problem by combining genetic programming and nonlinear PCA. Recently, a monitoring method based on kernel PCA was proposed to address a nonlinear process [9–13]. Kernel PCA has an advantage over other nonlinear PCA methods because it does not require nonlinear optimization. However, kernel PCA contains an assumption that the observations in the extended space conform to a Gaussian distribution. Thus, selection of an optimal number of PCs in the kernel space is important [11,13,14].

The Gaussian mixture model (GMM) is another method developed to address the nonlinear property of data [5]. The GMM uses a combination of Gaussian components to describe the data set from complex industrial processes. The number of Gaussian components can be determined automatically by the F-J algorithm, which is a variant of an expectation-maximization algorithm [15]. Hence, the combination of the Gaussian components approximates the nonlinear process [5]. However, when the observations described by each of the Gaussian components do not follow a Gaussian distribution, the results from a GMM-based control chart may be unsatisfactory [16,17]. It is possible that monitoring performance may soon

* Corresponding author.
E-mail addresses: youngjae8734@gmail.com (Y.J. Choung), jihoon.kang82@gmail.com (J. Kang), sbkim1@korea.ac.kr (S.B. Kim).

http://dx.doi.org/10.1016/j.jprocont.2017.02.005

be improved with the emergence of different non-Gaussian data-modeling techniques [5,18].

Another alternative is a self-organizing map-based control chart. Yu and Xi [21] proposed a minimum quantization error (MQE) chart based on a self-organizing map (SOM). Because a neural network-based method is useful to describe the nonlinearity properties of process data, an MQE chart also can deal with nonlinearity. Nevertheless, all of these nonlinear monitoring methods were formulated based on an assumption that the process data is time-invariant, which means the distribution, mean, and covariance of the data do not change after the offline training procedure.

In reality, many industrial processes contain both nonlinear and time-varying properties because of fluctuations in process raw materials, slow shifts in the set points, aging of the main process components, seasoning effects, and catalyst deactivation [5,15,19]. For these reasons, it is difficult to apply a traditional monitoring approach to time-varying data [20]. The use of traditional SPC techniques with time-varying data will degrade the performance of the monitoring scheme. Therefore, several techniques have been proposed to improve monitoring performance with these processes. Wold [22] proposed use of exponentially weighted moving average (EWMA) filters in conjunction with PCA and partial least squares (PLS). Rigopoulos et al. [23] discussed the use of a similar moving window scheme. Rannar et al. [24] used a hierarchical PCA to address adaptive batch monitoring, which is similar to EWMA and PCA. Li et al. [25] developed two adaptive PCA algorithms. However, all of these methods accumulate the data used, resulting in adaptation speed degradation as the data size increases [15]. Wang et al. [26] proposed a more computationally efficient moving-window algorithm similar to recursive PCA. Its adaptation is achieved by calculating the current correlation matrix from the previous one instead of using the old process data. Liu [27] extended Wang's method into the nonlinear field by using a moving-window kernel PCA approach. The kernel PCA model is updated as the moving-window slides along the data. Jeng [28] discussed combining both recursive PCA and moving window PCA to exploit both methods. More recently, Xie and Shi [15] used moving window-based update logic with the GMM. Their approach could address the multimode and time-varying problems and could approximate monitoring of the nonnormal and nonlinearity problems. However, as noted earlier, results from a GMM-based control chart may be unsatisfactory when the observations described by each of the Gaussian components do not follow a Gaussian distribution [16,17].

In the present paper, we propose an alternative nonparametric control chart that can monitor both time-varying and nonlinear processes. We call the proposed method a parameter-less self-organizing map (PLSOM)-based control chart because it uses a PLSOM algorithm[29] that addresses both the nonlinear and time-varying problems.

The remainder of this paper is organized as follows. In Section 2, related works are introduced. Section 3 describes the proposed PLSOM-based control chart for time-varying and nonlinear processes. Section 4 presents a simulation study to examine the performance of the proposed PLSOM-based control chart under various scenarios. Section 5 presents a case study using a thin-film transistor-liquid crystal display process. Section 6 consists of our concluding remarks.

2. Related work

2.1. Self-organizing map

A SOM, introduced by Kohonen in 1982, is an unsupervised learning method that consists of a set of neurons that gradually adapt to input data by competitive learning. These neurons also represent data points and describe the data similar to the way k-means center points function [30]. It creates ordered neurons that preserve the topology of the mapped data. The adaptation of neurons is based on a similarity measure, which is usually Euclidean distance. The SOM has been successfully applied in various engineering applications, including pattern recognition, image analysis, process monitoring and control, and fault diagnosis [31]. A brief description of a SOM algorithm is as follows:

Step 1. At the start of a SOM algorithm, all weights w_j of the neurons are initialized with random values.

Step 2. An input vector $X = \left\{x_1, x_2, \ldots, x_p\right\}$ is presented at epoch (or iteration) t.

Step 3. The best matching unit (or winning neuron) is selected. The best matching unit (BMU) is determined by minimizing the Euclidean distance between the input vector X and the weight vectors w_j:

$$\|X - w_{bmu}\| = \min_j \|X - w_j\|, j = 1, 2, \ldots, N, \tag{1}$$

where w_{bmu} is the BMU for input vector X, operator $\| \ \|$ denotes the Euclidean distance, and N is the number of neurons in the SOM.

Step 4. The weight of the neurons is updated using Eqs. (2) and (3).

$$w_i(t+1) = w_i(t) + \Delta w_i(t), \tag{2}$$

$$\Delta w_i(t) = d(t) h_{c,i}(t) [x(t) - w_i(t)], \tag{3}$$

where $w_i(t+1)$ is the weight of neuron i at iteration $t+1$, $d(t)$ is the learning rate, and $h_{c,i}(t)$ is the neighborhood function. The learning rate $d(t)$ and the neighborhood function $h_{c,i}(t)$ are decreased in accordance with the annealing scheme.

Step 5. If iteration t reaches a predefined number T, the procedure ends. Otherwise, steps 2–4 are repeated for the next input vector.

2.2. MQE chart

A SOM has outstanding noise tolerance techniques that do not require any assumptions on the statistical distribution of the monitored processes. These features make a SOM an effective technique for quality control [21]. The SOM-based monitoring method with the moving window approach called an MQE chart was proposed by Yu and Xi [21].

The monitoring statistic of an MQE chart is calculated by the following equation:

$$MQE = \|X - w_{bmu}\|, \tag{4}$$

where X is the input vector, and w_{bmu} is the weight vector of the BMU.

The MQE statistic indicates the deviation of input vector X from the weight vector of the BMU. If the MQE value exceeds a predetermined threshold, the process is considered as in an abnormal state. An MQE chart is an effective tool to monitor nonlinear processes. However, the performance of the monitoring model can be degraded if normal changes occur, such as catalyst deactivation or equipment aging [15].

2.3. Time-adaptive GMM chart for time-varying processes

The GMM describes a complex process with several Gaussian components. Hence, the GMM model can be effective for non-Gaussian, nonlinear, and multimode process monitoring [5]. To address normal change problems, Xie et al. [15] proposed a time adaptive monitoring scheme based on the GMM.