Expert Systems with Applications 41 (2014) 7579–7595

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Bivariate quality control using two-stage intelligent monitoring scheme

Ibrahim Masood [a,*], Adnan Hassan [b]

[a] Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia
[b] Faculty of Mechanical Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia

ARTICLE INFO

Article history:
Available online 5 June 2014

Keywords:
Balanced monitoring
Bivariate quality control
Statistical features
Synergistic artificial neural network
Two-stage monitoring

ABSTRACT

In manufacturing industries, it is well known that process variation is a major source of poor quality products. As such, monitoring and diagnosis of variation is essential towards continuous quality improvement. This becomes more challenging when involving two correlated variables (bivariate), whereby selection of statistical process control (SPC) scheme becomes more critical. Nevertheless, the existing traditional SPC schemes for bivariate quality control (BQC) were mainly designed for rapid detection of unnatural variation with limited capability in avoiding false alarm, that is, imbalanced monitoring performance. Another issue is the difficulty in identifying the source of unnatural variation, that is, lack of diagnosis, especially when dealing with small shifts. In this research, a scheme to address balanced monitoring and accurate diagnosis was investigated. Design consideration involved extensive simulation experiments to select input representation based on raw data and statistical features, artificial neural network recognizer design based on synergistic model, and monitoring–diagnosis approach based on two-stage technique. The study focused on bivariate process for cross correlation function, $\rho = 0.1–0.9$ and mean shifts, $\mu = \pm 0.75–3.00$ standard deviations. The proposed two-stage intelligent monitoring scheme (2S-IMS) gave superior performance, namely, average run length, $ARL_1 = 3.18–16.75$ (for out-of-control process), $ARL_0 = 335.01–543.93$ (for in-control process) and recognition accuracy, RA = 89.5–98.5%. This scheme was validated in manufacturing of audio video device component. This research has provided a new perspective in realizing balanced monitoring and accurate diagnosis in BQC.

1. Introduction

In manufacturing industries, when quality feature of a product involves two correlated variables (bivariate), an appropriate SPC charting scheme is necessary to monitor and diagnose these variables jointly. Specifically, process monitoring refers to the identification of process condition either in a statistically in-control or out-of-control, whereas process diagnosis refers to the identification of the source variable(s) for out-of-control condition. In addressing this issue, the traditional SPC charting schemes for BQC such as χ^2 (Hotelling, 1947), multivariate cumulative sum (MCUSUM) (Crosier, 1988), and multivariate exponentially weighted moving average (MEWMA) (Lowry, Woodall, Champ, & Rigdon, 1992; Prabhu & Runger, 1997) are known to be effective in monitoring aspect. Unfortunately, they are merely unable to provide diagnosis information, which is greatly useful for a quality practitioner in finding the root cause error and solution for corrective action. Since then, major researches have been focused on diagnosis aspect. Shewhart-based control charts with Bonferroni-type control limits (Alt, 1985), principle component analysis (PCA) (Jackson, 1991), multivariate profile charts (Fuchs & Benjamini, 1994), T^2 decomposition (Mason, Tracy, & Young, 1995) and Minimax control chart (Sepulveda & Nachlas, 1997), among other, have been investigated for such purpose. Further discussions on this issue can be found in Lowry and Montgomery (1995), Kourti and MacGregor (1996), Mason, Tracy, and Young (1997) and Bersimis, Psarakis, and Panaretos (2007).

In the related study, development in soft computing technology has motivated researchers to explore the use of machine learning (ML) technology for automatically recognizing SPC chart patterns towards improving capability in monitoring and diagnosis. Identification of these patterns coupled with engineering knowledge of the process would lead to more specific diagnosis information. Expert systems (ES) (Chih & Rollier, 1994; Chih & Rollier, 1995), Fuzzy inference system (FIS) (Wang & Chen, 2001), artificial neural network (ANN), decision tree learning (DT) (Guh & Shiue, 2005), and support vector machine (SVM) (Cheng & Cheng, 2008) methods, among others, have been studied in designing the advanced SPC pattern recognition schemes. Extensive literature review revealed that most of the proposed schemes were developed based on

* Corresponding author. Tel.: +607 4537700.
E-mail addresses: ibrahim@uthm.edu.my (I. Masood), adnan@fkm.utm.my (A. Hassan).
URLs: http://www.uthm.edu.my (I. Masood), http://www.utm.my (A. Hassan).

http://dx.doi.org/10.1016/j.eswa.2014.05.042
0957-4174/

research in ANN models such as an integrated bivariate SPC-ANN (Chen & Wang, 2004; Niaki & Abbasi, 2005; Yu, Xi, & Zhou, 2009), novelty detector ANN (Zorriassatine, Tannock, & O'Brien, 2003), modular ANN (Guh, 2007), ensemble ANN (Yu & Xi, 2009), multi-module-structure ANN (El-Midany, El-Baz, & Abd-Elwahed, 2010), hybrid learning ANN (Salehi, Bahreininejad, & Nakhai, 2011), an integrated ANN-SVM (Salehi, Kazemzadeh, & Salmasnia, 2012), and feature-based ANN (Masood & Hassan, 2013).

The integrated bivariate SPC-ANN schemes combined the traditional SPC chart(s) with an ANN model. The traditional SPC chart(s) role for monitoring the existence of unnatural variation in bivariate process, whereas an ANN model roles for diagnosing the sources of variation. In that case, an ANN model is utilized only when necessary, that is, when an out-of-control signal is triggered. Inversely, the other schemes such as novelty detector ANN consist of fully ANN or fully ML-based model for monitoring and diagnosing simultaneously. In that case, an ANN model is continuously utilized, that is, for triggering out-of-control signal and then, for identifying the sources of variation. Further discussion on these schemes can be found in (Masood & Hassan, 2010; Hachicha & Ghorbel, 2012).

1.1. Problem situation and solution concept

When dealing with monitoring and diagnosis of bivariate process variation in mean shifts, based on process monitoring viewpoint, an effective bivariate SPC scheme should be able to identify out-of-control condition as quickly as possible at the shortest ARL_1 (average run length for out-of-control process, $ARL_1 \Rightarrow 1$). Concurrently, it should be able to maintain small false alarm at the longest ARL_0 (average run length for in-control process, $ARL_0 \gg 200$). Nevertheless, the existing traditional SPC schemes were mainly designed by focusing on rapid detection of out-of-control condition ($ARL_1 \Rightarrow 1$) but it has limited capability in avoiding false alarm ($ARL_0 \leqslant 200$). Fig. 1 illustrates the concepts of imbalanced monitoring vs. balanced monitoring as the central theme for this investigation.

Based on diagnosis viewpoint, an effective bivariate SPC scheme should be able to identify the source variable(s) of out-of-control condition as accurate as possible. Nevertheless, it is difficult to correctly recognize when dealing with small shifts ($\leqslant$1.0 standard deviation). Chih and Rollier (1994), Chih and Rollier (1995), Zorriassatine, Tannock, and O'Brien (2003), Chen and Wang (2004) and Yu and Xi (2009), for examples, have reported less than 80% accuracy for diagnosing mean shifts at 1.0 standard deviation. Among others, only Guh (2007) and Yu et al. (2009) reported the satisfied results ($\geqslant$90% accuracy).

The imbalanced monitoring and lack of diagnosis capability as mentioned above need further investigation. In order to minimize erroneous decision making in BQC, it is essential to enhance the overall performance towards achieving balanced monitoring (rapidly detect process variation/mean shifts with small false alarm as shown in Fig. 1) and accurate diagnosis (accurately identify the sources of variation/mean shifts). Additionally, the BQC applications are still relevant in today's manufacturing industries. In solving this issue, a two-stage intelligent monitoring scheme (2S-IMS) was designed to deal with dynamic correlated data streams of bivariate process. This paper is organized as follows. Section 2 describes a modeling of bivariate process data streams and patterns. Section 3 presents the framework and procedures of the 2S-IMS. Section 4 discusses the performance of the proposed scheme in comparison to the traditional SPC. Section 5 finally outlines some conclusions.

2. Modeling of bivariate process data streams and patterns

A large amount of bivariate samples is required for evaluating the performance of the 2S-IMS. Ideally, such samples should be tapped from real world. Unfortunately, they are not economically available or too limited. As such, there is a need for modeling of synthetic samples based on Lehman (1977) mathematical model. Further discussion on data generator can be found in Masood and Hassan (2013).

In bivariate process, two variables are being monitored jointly. Let $X_{1\text{-}i} = (X_{1\text{-}1},\ldots,X_{1\text{-}24})$ and $X_{2\text{-}i} = (X_{2\text{-}1},\ldots,X_{2\text{-}24})$ represent 24 observation samples for process variable 1 and process variable 2 respectively. Observation window for both variables start with samples $i = (1,\ldots,24)$. It is dynamically followed by $(i + 1)$, $(i + 2)$ and so on. When a process is in the state of statistically in-control, samples from both variables can be assumed as identically and independently distributed (*i.i.d.*) with zero mean ($\mu_0 = 0$) and unity standard deviation ($\sigma_0 = 1$). Depending on process situation, the bivariate samples can be in low correlation ($\rho = 0.1$–0.3), moderate correlation ($\rho = 0.4$–0.6) or high correlation ($\rho = 0.7$–0.9). Data correlation (ρ) shows a measure of degree of linear relationship between the two variables. Generally, this data relationship is difficult to be identified using Shewhart control chart as shown in Fig. 2. On the other hand, it can be clearly indicated using scatter diagram. Low correlated samples yield a circular pattern (circular distributed scatter plot), moderate correlated samples yield a perfect ellipse pattern, whereas high correlated samples yield a slim ellipse pattern.

Disturbance from assignable causes on the component variables (variable-1 only, variable-2 only, or both variables) is a major source of process variation. This occurrence could be identified by various causable patterns such as mean shifts (sudden shifts), trends, cyclic, systematic or mixture. In this research, investigation was focused on sudden shifts patterns (upward and downshift shifts) with positive correlation ($\rho > 0$). Seven possible categories of bivariate patterns were considered in representing the bivariate process variation in mean shifts as follows:

- N (0,0): both variables $X_{1\text{-}i}$ and $X_{2\text{-}i}$ remain in-control.
- US (1,0): $X_{1\text{-}i}$ shifted upwards, while $X_{2\text{-}i}$ remains in-control.
- US (0,1): $X_{2\text{-}i}$ shifted upwards, while $X_{1\text{-}i}$ remains in-control.
- US (1,1): both variables $X_{1\text{-}i}$ and $X_{2\text{-}i}$ shifted upwards.
- DS (1,0): $X_{1\text{-}i}$ shifted downwards, while $X_{2\text{-}i}$ remains in-control.
- DS (0,1): $X_{2\text{-}i}$ shifted downwards, while $X_{1\text{-}i}$ remains in-control.
- DS (1,1): both variables $X_{1\text{-}i}$ and $X_{2\text{-}i}$ shifted downwards.

Reference bivariate shift patterns based on mean shifts ±3.00 standard deviations are summarized in Fig. 3. Their structures are unique to indicate the changes in process mean shifts and data correlation. The degree of mean shifts can be identified when the center position shifted away from zero point (0,0).

3. Two-stage intelligent monitoring scheme

As noted in Section 1, an integrated MSPC-ANN was combined in a single-stage monitoring scheme (direct monitoring–diagnosis) as proposed in Chen and Wang (2004), Niaki and Abbasi (2005), and Yu et al. (2009). The other schemes based on fully ANN-based models as proposed in Zorriassatine, Tannock, and O'Brien (2003), Guh (2007), Yu and Xi (2009) and El-Midany et al. (2010) also can be classified as a single-stage monitoring scheme. In this research, two-stage monitoring scheme was investigated by integrating the powerful of MEWMA control chart and Synergistic-ANN model for improving the monitoring–diagnosis performance. Framework and pseudo-code (algorithm) for the proposed scheme are summarized in Figs. 4 and 5 respectively. It should be noted that an initial setting as follows needs to be performed before it can be put into application:

- Load the trained raw data-ANN recognizer into the system.