Mechanical Systems and Signal Processing 60-61 (2015) 644–655

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Plant-wide quantitative assessment of a process industry system's operating state based on color-spectrum ☆

Sun Kai [a,b,*], Gao Jianmin [a], Gao Zhiyong [a], Jiang Hongquan [a], Gao Xu [a]

[a] *State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, PR China*
[b] *College of Mechanical & Electrical Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, PR China*

ARTICLE INFO

Article history:
Received 11 October 2014
Received in revised form
8 November 2014
Accepted 20 January 2015
Available online 23 February 2015

Keywords:
Color-spectrum
Operating state assessment
Process industry system
Data visualization
Multivariate time series

ABSTRACT

This paper presents a general theoretical framework to assess the operating state of a process industry system quantitatively based on meshing the theory of scientific data visualization and digital image processing. First, a series of color-spectrum, which represent the operating state of the system, is formed by mapping the monitor data set to a group of digital color images. Second, the common feature of color-spectrum, which is named benchmark-color-spectrum, is extracted as a standard of the normal state. Third, the abnormal degree can be quantified by calculating the difference of the benchmark-color-spectrum with observed color-spectrum. At last, a plant-wide operating state of the system in a period of time can be shown by plotting quantitative abnormal degree. Two case is included to illustrate the proposed method and its appropriateness. One is a general process industry system simulator named Tennessee Eastman Process. Another is an air compressor group which belongs to a real chemical plant.

1. Introduction

A modern process industry system is a typical distributed and complex electromechanical system. It is composited of large-scale of mechanical equipment, various chemical reactors and computerized monitor systems. Its core supervisory system named distributed control system (DCS) has thousands of various sensors which upload monitor data every minute. Therefore, a huge monitor data set, which is a nonlinear multivariate time series essentially, has been accumulated. That is, the data-driven operating state assessment has been transferred to a multivariate time series analysis, which is a challenge for decades.

Many remarkable efforts have been implemented to meet this challenge. There are two categories of data-driven operating state assessment of process industry system: qualitative and quantitative. The conventional qualitative methods are expert system [1–4] and qualitative trend analysis (QTA) [5–9]. Initial attempts at the application of expert system can be found in Henley [1], Chester et al. [2]. Tarifa and Scenna [3] have proposed a hybrid system that used signed directed graph (SDG) and fuzzy logic. An expert system approach for fault diagnosis in batch processes is discussed in Scenna [4]. The limitation of an expert system approach is system-specific. Their representation power is quite limited and they are difficult

☆ This project is supported by National Natural Science Foundation of China (Grant no. 51175402).
* Correspondence to: State Key Laboratory for Manufacturing Systems Engineering, School of Mechanical Engineering, Xi'an Jiaotong University, Xian Ning West Road No. 28, Xi'an 710049, Shannxi, PR China. Tel.: +86 18629404731.
E-mail addresses: clara_sun96@hotmail.com, clara_sun96@163.com (S. Kai).

http://dx.doi.org/10.1016/j.ymssp.2015.01.029
0888-3270/

to update. Qualitative trend analysis(QTA) including trend analysis and prediction is an important component of process monitoring and supervisory. Cheung and Stephanopoulos [5] introduce triangulation to represent trends. Janusz and Venkatasubramanian [6] identify a comprehensive set of primitive-based language which any trend can be represented. Mah [7] use piecewise-linear elements to illustrate trend. Vedam [8] utilize wavelet, neural and B-Splines-based to extract trend. Recent applications of qualitative trends analysis integrate signed directed graphs and principal component analysis [9]. The quantitative methods can be broadly classified as non-statistical or statistical methods, such as Neural networks, principal component analysis (PCA) [10–12], partial least squares (PLS) [13] etc.

These conventional methods usually select some “key” equipment and use their monitor data to analyze the system's operating state. The remainder elements have been ignored because they are “unimportant”. Are they real unimportant? The investigation shows the most severe accidents due to insignificant equipment which has been ignored. Moreover, it is difficult to distinguish “key” sensors because of the complicate of the system. Therefore, all equipment of the system should be analyzed simultaneously in order to evaluate the plant-wide operating state.

In order to meet this challenge, combining scientific data visualization [14] and digital image processing, a new method named color-spectrum assessment was proposed. The monitor data set of process industry system is a multiple variables time series with large scale of variables which is inherently a typical multidimensional multivariate data set. Considering the definition of scientific data visualization which is concerned with exploring data and information in such a way to gain understanding and insight into the data [15]. Visualization offers a method for seeing the unseen [14]. Moreover, it simples the monitor data set by mapping it from a limitless element's value range of float matrix to a limited range of integer matrix. Following the visualization process, the monitor data set can be converted into a displayable image named color-spectrum which is the data base of analyzing plant-wide operating state of a process industry system. The color-spectrum is inherently a digital image. Since the pixels of a image has natural characteristics of massive, nonlinear and high correlation, the impact of these influencing factors to the monitor data set has been eliminated. So the problem has been transferred from a difficult problem of a multivariate data set analysis to a simple requirement of digital image process. In order to assess the abnormal degree of the plant-wide operating state quantitatively, abundant approaches of digital image processing can be used.

The outline of this paper is the following: in Section 2, a rule was defined to transfer the period monitor data set to color-spectrum. In Section 3, a new algorithm has been introduced of quantizing the trend of plant-wide operating state by calculating the abnormal degree of the color-spectrum. In Section 4, two case has been used to verify the algorithm. One is a general chemical process simulator named Tennessee Eastman Process. Another is an air compressor group of a real chemical plant. This paper is closed with a conclusion of the method in Section 5.

2. Forming color-spectrum vector based on monitor data set

Considering a process industry system with n sensors marked with the number j, $1 \le j \le n$, $j \in N$. It's system period is T. All sensors upload monitor data every Δt. During a special system period, it accumulates m monitor data ($m = T/\Delta t$, $m \in N$) for every sensor, which is defined as monitor data matrix X.

Definition 1. Data matrix X is a $m \times n$ matrix of monitor data set of every system period. The row of matrix X is composited with the sensors of system. And the column represents time, as the following equation:

$$X = \left.\begin{array}{c} \textit{sensors} \\ \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}_{m \times n} \end{array}\right\} time \tag{1}$$

Meanwhile, image the time duration of the historical monitor data set is Lh. Obviously, Lh is longer than T, where is $Lh = \kappa T, \kappa \in N$. Therefore, the kth $(0 \le k \le \kappa - 1, k \in N)$ runtime period can be shown as a time interval $[kT, (k+1)T]$. The whole observation of the system can be divided into κ interval such as $[0, T], [T, 2T], \ldots, [(\kappa - 1)T, \kappa T]$. So the kth runtime period's monitor data set can be shown as $U_k = \left\{ x_i^j \middle| kT \le i \le (k+1)T, 1 \le j \le n, i, j \in N \right\}$. Besides, the monitor data set $U = \left\{ U_k \middle| 1 \le k \le \kappa \right\}$ is a multivariable discrete time series that can be transferred to a vector of data matrix X.

Definition 2. Data Matrix vector $\overrightarrow{X}$ is a one dimensional vector which elements is data matrix of each system period. As the following equation:

$$\overrightarrow{X} = [X_1, X_2, \ldots, X_k, \ldots, X_\kappa]_{1 \times \kappa} \tag{2}$$

Then, define a rule χ to color every elements of X_k in order to form a color-spectrum Γ_k.

Definition 3. Data color rule χ colors numerical data to corresponding pixel value. As the following steps.

(1) Form data interval $[x_{\min}, x_{\max}]$, $x \in X$.
(2) Get the number of color as color_num.
(3) Separate the data interval and get the length of each subinterval $L = (X_{max} - X_{min}/color\_num)$.