

This article was downloaded by:[National Taiwan University]
On: 1 August 2008
Access Details: [subscription number 788856085]
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Production Research

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713696255>

Decision tree based control chart pattern recognition

Chih-Hsuan Wang^a, Ruey-Shan Guo^b, Ming-Huang Chiang^b, Jehn-Yih Wong^a

^a Department of Business Administration, Ming Chuan University, Taipei, Taiwan

^b Department of Business Administration, National Taiwan University, Taipei, Taiwan, Republic of China

Online Publication Date: 01 September 2008

To cite this Article: Wang, Chih-Hsuan, Guo, Ruey-Shan, Chiang, Ming-Huang and Wong, Jehn-Yih (2008) 'Decision tree based control chart pattern recognition', International Journal of Production Research, 46:17, 4889 — 4901

To link to this article: DOI: 10.1080/00207540701294619

URL: <http://dx.doi.org/10.1080/00207540701294619>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Decision tree based control chart pattern recognition

CHIH-HSUAN WANG*†, RUEY-SHAN GUO‡,
MING-HUANG CHIANG‡ and JEHN-YIH WONG†

†Department of Business Administration, Ming Chuan University,
Ban-chiao, Taipei, Taiwan, Republic of China

‡Department of Business Administration, National Taiwan University,
Taipei, Taiwan, Republic of China

(Received ■■ February 2007; in final form ■■ ■■ ■■)

This paper presents a new approach to classify six anomaly types of control chart patterns (CCP), of systematic pattern, cyclic pattern, upward shift, downward shift, upward trend, and downward trend. Current CCP recognition methods use either unprocessed raw data or complex transformed features (via principal component analysis or discrete wavelet transform) as the input representation for the classifier. The objective of using selected features is not only for dimension reduction of input representation, but also implies the process of data compression. In contrast, using raw data is often computationally inefficient while using transformed features is very tedious in most cases. Therefore, owing to its computational advantage, using appropriate features of CCP to achieve good classification accuracy becomes more promising in real process implementation. In this study, using three features of CCP shows quite a competitive performance in terms of classification accuracy and computational loading. More importantly, the proposed method presented here has potential to be generalized to medical, financial, and other application of temporal data.

Keywords: Control chart patterns; Feature selection; Decision tree; Anomaly detection; Anomaly classification

1. Introduction

In the current manufacturing industry, SPC/EPC (statistical process control/engineering process control) techniques play an important role to improve product quality and to monitor the process situation. Generally speaking, observed variation of quality characteristics results from either natural variation (common cause) or unnatural variation (assignable cause). Natural variation always exists in the manufacturing process regardless of the fact that how well the product is designed and how adequately the process is maintained. On the contrary, unnatural patterns resulting from unnatural variation are often associated with a specific set of

*Corresponding author. Email: chihwang@mcu.edu.tw

assignable causes. These unnatural patterns often contain valuable information relevant to process parameters and process changes. The two quantities that are commonly monitored in practice are the mean and the range of the sample. The earliest techniques developed by Shewhart involve \bar{X} chart (used for monitoring process mean) and R chart (used for monitoring process variance). Once the sources of assignable causes are correctly identified, quality practitioners can remove them and bring the abnormal process back to the normal condition (natural variation). Unfortunately, control charts are easily used inappropriately without owning prior knowledge or sufficient historical data. Moreover, control charts do not provide any pattern-related information due to lack of the capability to recognize different kinds of unnatural patterns (Guh and Tannock 1999, Yang and Yang 2002).

Similar to other researchers (Pham and Oztemel 1994, Cheng 1997, Pham and Wani 1997, Guh and Hsieh 1999, Guh and Tannock 1999, Yang and Yang 2002, Hassan *et al.* 2003, Yousef 2004, Guh and Shiue 2005, Yang and Yang 2005), anomaly control chart patterns (CCP) illustrated in this research consist of four types listed below and numerous quality practitioners ascribed their corresponding assignable causes to the following (Cheng 1997):

- (1) Trend patterns: A trend can be defined as a continuous movement in either positive or negative direction. Possible causes are tool wear, operator fatigue, equipment deterioration, and so on.
- (2) Shift patterns: A shift can be defined as a sudden change above or below the average of the process. This change may be caused by an alternation in process setting, replacement of raw materials, minor failure of machine parts, or introduction of new workers, and so forth.
- (3) Cyclic patterns: Cyclic behaviours can be observed by a series of peaks and troughs occurring in the process. Typical causes are the periodic rotation of operators, systematic environmental changes or fluctuation in the production equipment.
- (4) Systematic patterns: The characteristic of systematic patterns is that a point-to-point fluctuation has systematically occurred. It means a low point is always followed by a high point and vice versa. Possible causes include difference between test sets and difference between production lines where product is sampled in rotation.

Most of the existing control chart pattern recognition schemes use the unprocessed raw data as the input representation to the classifier. Thus, the large size of input dimension may lead to an inefficient computational time especially when a complex artificial neural network (ANN) based classifier is adopted. In brief, the approach taking the extracted or transformed features of control chart patterns as the classifier input presents the following advantages:

- It is more robust to the amount of noise embedded in time series data.
- It saves huge computation time because of the reduced input dimension.
- Human heuristics can be easily incorporated into the classification decision tree.

Obviously, if a small size of input vector could convey sufficient information to represent the original CCPs, an efficient computation with satisfied classification accuracy could be easily achieved. However, if the process of feature extraction

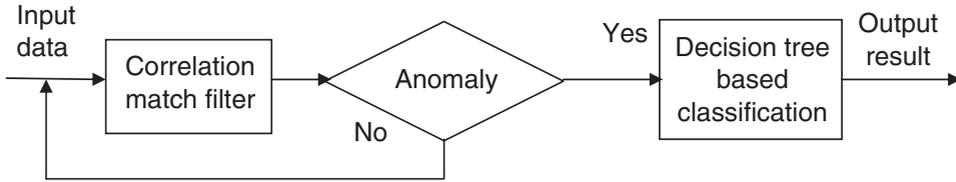


Figure 1. The proposed approach.

(such as principal component analysis) or feature transformation (such as discrete wavelet analysis) is too complex or very tedious, its on-line responsiveness capability with respect to fault detection or anomaly diagnosis may be seriously questioned. By contrast, the advantage of using selected features without changing its original characteristics could generate a rule-based expert system that contains information more explicitly and the rules stored inside the system may be modified or updated more flexibly (Pham and Wani 1997). The proposed framework to achieve this goal is listed in figure 1. A correlation analysis will be first conducted between each input data and the normal reference (natural pattern). Once an anomaly pattern is recognized, feature selection and DT based classification will be followed to provide decision support for output result.

The remainder of this paper is organized as follows. Section 2 briefly reviews the existing approaches and section 3 presents the proposed approach. Computer simulated results, including the training set and the testing set, are conducted in section 4. Finally, conclusions are drawn in section 5.

2. Previous work on control chart pattern recognition

Fault detection and diagnosis systems have played an important role in modern SPC/EPC. To sum up, there are two basic approaches in fault detection system development: a model-based approach and a feature-based approach (Jin and Shi 2001). In the former approach, observations are assumed to follow a time-ordered stochastic process and hence fault models or data distribution characteristics need to be known in advance. The critical concern of using a model-based approach is to have an appropriate model which is sensitive to process faults but robust to process noises. Besides, sufficient historical information of fault models is normally unavailable at the beginning of the manufacturing process in practice. On the other hand, a feature-based approach based on linear/nonlinear components or Fourier/wavelet coefficients is more suitable to deal with a complex process problem, especially when no explicit prior information is available. Features could be extracted through various kinds of component analysis (Lee *et al.* 2004a, b), or could be obtained through Fourier or wavelet transformation (Ganesan *et al.* 2004, Yousef 2004). Extracted or transformed features not only reduce the size of the input dimension but can also decompose the complex functional relationship between time-series data and the associated manufacturing process. However, there is no denying that extracted or transformed features are not intuitive to be understood and may need a long computation time.

Rapid developments in artificial intelligence have motivated researchers to explore the artificial neural network (ANN) based control chart pattern recognizer. Most researchers (Cook and Chiu 1998, Chang and Ho 1999, Ho and Chang 1999) use supervised neural networks, including multi-layer perceptron (MLP) and radial basis function (RBF), to classify different kinds of process signal or control chart patterns. Other researchers (Pham and Oztemel 1994, Hwang and Chong 1995, Yang and Yang 2002, Pacella *et al.* 2004) use unsupervised neural networks, involving learning vector quantization (LVQ) and adaptive resonance theory (ART). However, one disadvantage of using neural networks is that the topology or the structure of neural networks cannot be systematically determined due to its black-box property. In addition, the training of network parameters is usually time-consuming and empirically infeasible for quality practitioners because of the requirement of many and good training samples.

Most of the previously mentioned studies use unprocessed raw data as the input vector to artificial neural network; it normally requires more input neurons and hence produces a large ANN structure. Thus, these data representations are inefficient for pattern recognition since the training of the network is extremely time-consuming. One way to improve the performance of the network is by using less input vectors to reduce the size and the complexity of the network. Hassan *et al.* (2003) use statistical features, such as skewness, kurtosis, and autocorrelation, to improve the performance of an ANN based control chart pattern recognizer. Based on nine chosen shape features (i.e. slope, number of mean crossings, number of least-square crossings, cyclic membership, average slope of line segment, slope difference, area between the pattern and the mean line, area between the pattern and the least-square line, area between the least-square line and the line segments), Pham and Wani (1997) proposed a shape feature based approach to classify six types of control chart patterns. Although the shape features of control chart patterns present more intuitive meaning than the statistical features, Pham and Wani (1997) did not indicate how to generate the heuristics rules, how to select the most discriminative attribute at each node, how to determine the threshold to split the decision tree and when to stop the classification process. Besides, some shape features like 'the area between the original pattern and the regression line' are quite hard to be derived in practice.

To our knowledge, very limited work has been reported on the benefit of using selected or extracted/transformed features so far, such as shape features (Pham and Wani 1997), statistical features (Hassan *et al.* 2003), component analysis (Lee *et al.* 2004a, b) or wavelet features (Jin and Shi 2001, Ganesan *et al.* 2004, Wang *et al.* 2007). The feature based approach has two main steps: (1) extraction of features, and (2) recognition of CCPs using those extracted features (Gauri and Chakraborty 2006). In this research, a new approach based on feature selection and rule induction is proposed to recognize six anomaly types of control chart patterns. Five candidate features of control chart patterns are determined by domain experts. In particular, the selection of the best attribute (most discriminative) at each node of the decision tree and also their priority along the decision tree are systematically determined via comparing the information entropy. Therefore, a rule-based expert system could be easily incorporated and constructed to facilitate an efficient and satisfied unnatural CCP classification.

3. The proposed approach

The proposed approach consists of three main steps: (1) recognition of anomaly control chart patterns, (2) selection of meaningful features from control chart patterns, and (3) classification of control chart patterns using a decision tree (see figure 1). Al-Ghanim and Ludeman (1997) evaluated the correlation between reference vectors and the input vector using an inner product. The principle of their approach is the so-called 'match filter' approach which measures the similarity between each input pattern and a prototype reference vector. Similar to Yang and Yang (2005), the statistical correlation coefficient (see equation (1)) is used to determine whether the input pattern is an anomaly or not. Besides, five candidate features are selected based on domain expertise but the specific features on recognition of control chart patterns will be systematically determined via constructing the classification decision tree.

$$corr = \frac{\sum_{t=1}^n (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^n (x_t - \bar{x})^2} \sqrt{\sum_{t=1}^n (y_t - \bar{y})^2}}, \quad (1)$$

where x_t , \bar{x} (and y_t , \bar{y}) respectively denote the input (reference) vector and its mean, and n is the total length of the observing window.

3.1 Control chart features

Some statistical features such as mean, standard deviation, skewness, kurtosis, and autocorrelation, are adopted in Hassan *et al.* (2003) to improve the performance of an ANN based classifier. Among them, skewness provides the information with regard to the degree of asymmetry and kurtosis measures the relative peak or flatness of its distribution. Their mathematical forms are respectively shown below:

$$mean = \frac{\sum_{t=1}^n x_t}{n} \quad (2)$$

$$std = \sqrt{\frac{\sum_{t=1}^n (x_t - mean)^2}{n}} \quad (3)$$

$$skew = \frac{\sum_{t=1}^n (x_t - mean)^3}{n(std)^3} \quad (4)$$

$$kurt = \frac{\sum_{t=1}^n (x_t - mean)^4}{n(std)^4} \quad (5)$$

$$auto = \frac{1}{n+1-k} [x_0 x_k + x_1 x_{k-1} + \dots + x_{n-k} x_n]. \quad (6)$$

However, due to a lack of intuitive meaning for constructing the decision tree, using statistical features may not be adequate to discriminate the various types of control chart patterns (Pham and Wani 1997). By observing their shape

characteristics of CCPs (see figure 2 and 3), five candidate features are selected as our discriminators in this paper, which are *LS-SLP* (the slope of least-square regression), *LS-ERR* (the sum of least-square error), *MN-ERR* (the sum of mean error), *LS-CROS* (the number of crossings between the original pattern and its least-square regression),

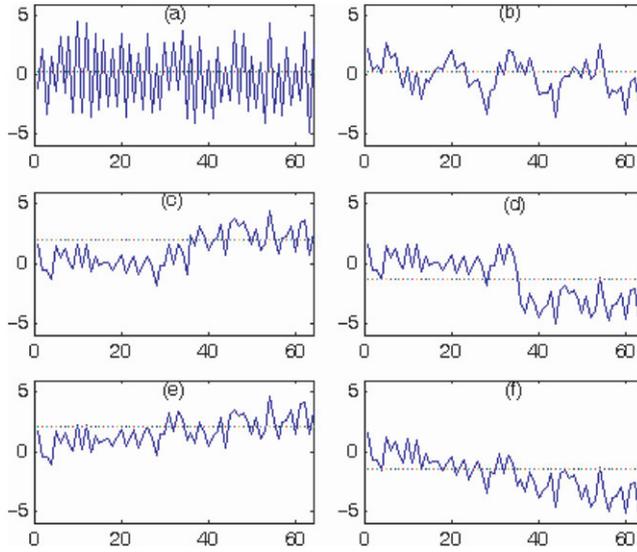


Figure 2. Least-square regression of control chart patterns.

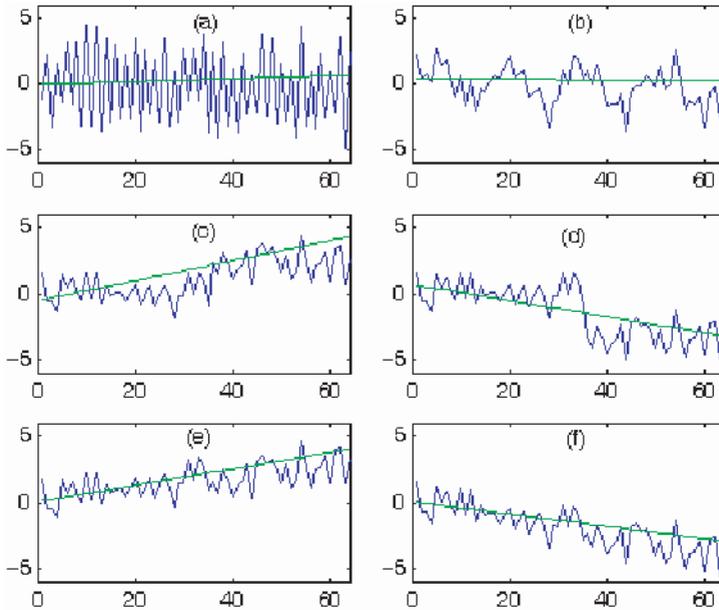


Figure 3. Mean regression of control chart patterns.

and *MN-CROS* (the number of crossings between the original pattern and its mean regression). Their mathematical forms are shown below:

$$LS-SLP = \frac{\sum_{t=1}^n (t - \bar{t})(x_t - \bar{x})}{\sum_{t=1}^n (t - \bar{t})^2} \quad (7)$$

$$LS-ERR = \sum_{t=1}^n (x_t - \hat{x}_{LS}) \quad (8)$$

$$MN-ERR = \sum_{t=1}^n (x_t - \hat{x}_{MN}), \quad (9)$$

where \hat{x}_{LS} and \hat{x}_{MN} respectively denote the estimated signal based on least-square and mean regressions. Besides, typical values of the five candidate features conducted from six types of CCPs are shown in table 1 and their detailed explanations are listed below:

- (1) *LS-SLP*: slope of the least-square line representing the pattern. The magnitude of the slope for both natural and cyclic patterns is approximately zero, while that for trend or shift patterns is obviously greater than zero. Therefore, *LS-SLP* may be a good candidate to differentiate natural and cyclic patterns from trend and shift patterns.
- (2) *LS-CROS*: number of least-square crossings. Not surprisingly, *LS-CROS* is highest for natural and trend patterns, intermediate for shift patterns and lowest for cyclic patterns. Similarly, *LS-CROS* may be suitable to separate natural and trend patterns from other patterns.
- (3) *MN-CROS*: number of mean crossings. *MN-CROS* is lowest for shift patterns, intermediate for cyclic and trend patterns and highest for natural patterns.
- (4) *LS-ERR*: sum of least-square regression error. Natural and trend patterns have lowest least-square error, shift patterns have intermediate least-square error while cyclic patterns have highest least-square error.
- (5) *MN-ERR*: sum of mean regression error. Natural patterns have lowest mean error; cyclic and trend patterns have intermediate mean error while shift patterns have highest mean error.

Obviously, the *LS-SLP* of both systematic pattern and cyclic pattern is approaching zero and much smaller than significantly positive (negative) of

Table 1. Typical value of five selected candidate features.

	<i>LS-SLP</i>	<i>LS-CROS</i>	<i>MN-CROS</i>	<i>LS-ERR</i>	<i>MN-ERR</i>
Cyclic	-0.0111	18.82	19.32	116.94	120.24
Systematic	0.0004	62.59	63.25	583.65	584.78
Up shift	0.0708	23.63	9.38	90.627	181.94
Down shift	-0.0711	24.03	9.095	88.404	180.43
Up trend	0.055	30.15	18.755	58.931	114.68
Down trend	-0.055	30.15	19.185	58.931	115.26

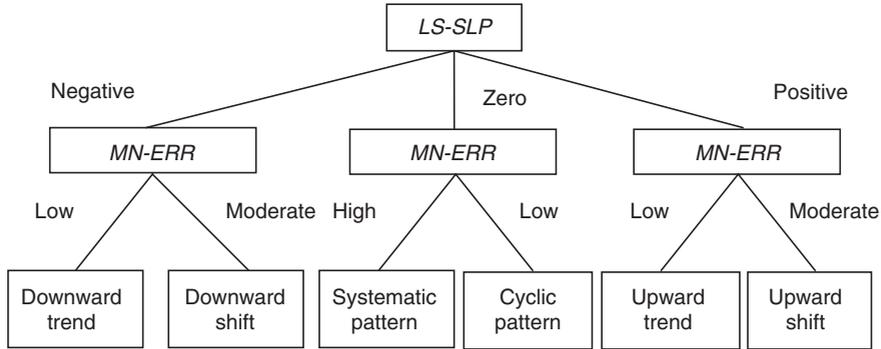


Figure 4. The rule induction for recognition of control chart patterns.

upward (downward) shift pattern or trend pattern. Besides, the $LS-CROS$ is largest in systematic pattern but smallest in cyclic pattern whereas the $MN-CROS$ is smallest in shift pattern. On the other hand, both $LS-ERR$ and $MN-ERR$ are smallest in trend pattern. Therefore, based an appropriate priority arrangement of different features along the decision tree, CCP classification could be easily achieved (see figure 4). Furthermore, figure 2 shows the generation of six types of control chart patterns and their least-square regression (see the solid lines) and figure 3 shows their mean regression (see the dashed lines). For simplification, six anomaly types of control chart patterns in two figures are denoted as: (a) systematic pattern, (b) cyclic pattern, (c) upward shift, (d) downward shift, (e) upward trend, and (f) downward trend.

3.2 The principle of a decision tree

A decision tree is widely used for supervised classification in practice, ranging from medical diagnosis to credit evaluation. Starting from the root node, the algorithm constructs a decision tree in a top-down and divide-and-conquer manner and also employs a greedy search through the decision tree. Moreover, splitting at each internal node represents a test on an attribute and this process will be terminated when all samples in the leaf node belong to the same class. According to Mitchell (1997), decision tree learning is best suited and robust to the problems with the following characteristics:

- Instances can be represented by attribute-value pairs.
- The target function has discrete output values.
- The training set may contain noise or missing attribute values.

The Iterative Dichotomizer 3 (ID3) algorithm (Quinlan 1983, 1986) and its successor, C4.5 (Quinlan 1993), are the main algorithms of research in the field of decision tree learning. In this paper, C4.5 used to induce classification rules for the control chart pattern recognition task. More specifically, C4.5 is an extended form of ID3 with additional characteristics such as the ability to handle continuous attribute, noisy data, and alternative measures for selecting attributes and pruning decision trees (Quinlan 1986). In general, rule induction in C4.5 has three phases (Guh and Shiue 2005). First, an initial, large tree is created from the set of examples.

Second, this tree is pruned by removing the branches with little statistical validity, and, third, the pruned tree is further processed to increase its interpretability. To introduce the process of decision tree induction, assume that the data set S consists of s data samples. Let the class label C_i have m distinct values denoting m distinct classes and let s_i denote the number of samples in class C_i . Using the information theory, the expected information (so called entropy) needed to classify a given sample can be defined by:

$$I(S) = - \sum_{i=1}^m p_i \log_2 p_i, \quad (10)$$

where $p_i = s_i/s$ is the probability that an sample belongs to class C_i . Suppose attribute A has n distinct values, $\{a_1, \dots, a_n\}$, and it can be used to partition the data set S into n subsets, $\{S_1, \dots, S_n\}$, where S_j represents those samples in S that have the attribute value a_j of A . Let s_{ij} denote the number of samples from class C_i in subset S_j . If attribute A is selected as the best attribute for splitting the current node, this attribute should have the largest information gain or the greatest entropy reduction. The expected information based on the partitioning into attribute A 's subsets is given by

$$I(A) = \sum_{j=1}^n \frac{s_j}{s} I(S_j), \quad (11)$$

where the term s_j/s , calculated by the number of samples in S_j divided by the total number of samples in S , acts as the weight of the j th subset. For any given subset S_j ,

$$I(S_j) = - \sum_{i=1}^m p_{ij} \log_2 p_{ij}, \quad (12)$$

where $p_{ij} = s_{ij}/s_j$ is the probability that a sample in S_j belongs to the class C_i . Therefore, the information gain due to branching on attribute A can be described as: $G(A) = I(S) - I(A)$. In other words, $G(A)$ is the expected entropy reduction caused by selecting attribute A . Basically, the attribute with the highest information gain will be selected as the best discriminator for the current node and the recursive process will be continued until all samples in each leaf node belong to the same class.

Unfortunately, when a decision tree is initially built, many branches inside the tree will reflect a lot of anomalies from the training set owing to noise or outliers. This is so-called over-fitting and hence tree pruning approaches are needed to remove unnecessary branches. There are two methods in tree pruning: 'pre-pruning' is achieved by halting the tree construction early and 'post-pruning' removes branches from a fully grown tree. In this paper, pre-pruning is adopted to save the computation time and its details will be described in section 4.2.

4. Simulated results

4.1 Pattern generation

A total of 1200 samples (200 samples in each category) were artificially generated for the data set: one half of each type is used for training and the other half is used

for testing. Without loss of generality, change points of both upward shift and downward shift, are randomly selected around half of the time window. Moreover, the amplitude of cyclic patterns, the magnitude of shift patterns and the slope of trend patterns are also randomly selected within a specific range. The random setting in both change point and shift quantity of unnatural patterns is for the purpose of increasing the adaptive capability of pattern classifiers. All details to generate six types of anomaly control chart patterns are shown from formulae (13)–(18) (also, see figure 2 and 3):

(a) Natural pattern:

$$x(t) = n(t) \sim N(0, 1), \quad (13)$$

where $x(t)$ is a sample at time t (from a standard Gaussian distribution).

(b) Cyclic pattern:

$$x(t) = n(t) + a \sin\left(\frac{2\pi t}{T}\right), \quad (14)$$

where a ($1.5 \leq a \leq 3$) and T ($=15$) respectively denote the amplitude and the period of cyclic patterns.

(c) Systematic pattern:

$$x(t) = n(t) + (-1)^t s, \quad (15)$$

where s ($1.5 \leq s \leq 3$) denotes the magnitude of shift.

(d) Upward shift/downward shift:

$$x(t) = n(t) \pm su(t - t_h), \quad (16)$$

where $u(t)$ stands for a unit step function shown below:

$$u(t - t_h) = \begin{cases} 0, & t < t_h \\ 1, & t \geq t_h. \end{cases} \quad (17)$$

(e) Upward trend/downward trend:

$$x(t) = n(t) \pm dt, \quad (18)$$

where d ($0.05 \leq d \leq 0.1$) stands for the slope of trend patterns.

4.2 Feature selection and pattern classification

Based on domain knowledge, five candidate features suitable for recognizing various types of control chart patterns are selected for the task of pattern classification, namely *LS-SLP*, *LS-CROS*, *MN-CROS*, *LS-ERR* and *MN-ERR*. In brief, each feature has more or less discriminative power to separate various types of control chart patterns. Determining the optimal attribute priority along the decision tree could be achieved by comparing their corresponding information entropy. At each node of the decision tree, the best attribute for branch splitting is selected by searching for a maximal information gain (see the underline marked in each entry of table 2). Note that the number in the first

Table 2. Information gain of various attributes at each node.

	<i>LS-SLP</i>	<i>LS-CROS</i>	<i>MN-CROS</i>	<i>LS-ERR</i>	<i>MN-ERR</i>
Level 0: Root node	1.59	0.98	0.99	1.16	1.12
Level 1-1: Zero-slope	N/A	1	1	1	1
Level 1-2: Positive-slope	N/A	0.23	0.45	0.48	<u>0.61</u>
Level 1-3: Negative-slope	N/A	0.38	0.53	0.42	<u>0.7</u>

column of table 2 represents the level of decision tree (the root node is denoted by level 0).

Obviously, *LS-SLP* is selected as the discriminative attribute at the root node (level 0) and then the data set is accordingly separated into three subgroups: zero-slope (comprising systematic and cyclic patterns), positive-slope (comprising positive trend and shift patterns), and negative-slope (comprising negative trend and shift patterns). Note, since *LS-SLP* is selected at level 0, it could not be selected at the next level (N/A denotes ‘not acceptable’). Similarly, *MN-ERR* is selected to discriminate shift pattern from trend pattern at both level 1-2 and level 1-3. Surprisingly, except *LS-SLP*, all features have the same discriminative power to distinguish systematic pattern from cyclic pattern at level 1-1 since all features of systematic pattern are much different from cyclic pattern. For convenience, *MN-ERR* is used within the zero-slope group to distinguish systematic pattern from cyclic pattern.

The attribute threshold to split each node is approximated by using the mean of their corresponding features. Hence, the threshold to distinguish the ‘zero-slope’ from the ‘positive-slope’ could be set by using the *LS-SLP* mean of those two corresponding subgroups. Moreover, in order to avoid over-fitting, the early termination of the tree-growing process is triggered when the majority category at each leaf node dominates by more than 90%. At last, the construction of the decision tree for CCP recognition can be conducted as shown in figure 4. In fact, a couple of ‘If-Then’ rules embedded in the decision tree could be easily shown. For instance, if the negative *LS-SLP* and the low *MN-ERR* hold simultaneously, the sample will be classified into the category of downward shift.

The classification results via a feature-based decision tree (DT) are presented in table 3. For simplification, the upper (lower) row in each entry respectively denotes the classification results of training (testing) samples. Specifically, most classification error results from misrecognition of cyclic pattern into systematic pattern and from the confusion between shift and trend pattern. In this research, only two features (*LS-SLP* and *MN-ERR*) are selected to separate various types of anomaly control chart patterns. Compared to other schemes, the proposed method not only effectively reduces the size of input dimension but can also achieve competitive classification accuracy (96.5%).

5. Conclusions

In this research, a decision-tree based approach has been proposed to recognize six anomaly types of control chart patterns. In particular, domain knowledge is

Table 3. Classification results of training and testing data set.

	Systematic	Cyclic	Up shift	Down shift	Up trend	Down trend
Systematic	100	0	0	0	0	0
	100	0	0	0	0	0
Cyclic	6	94	0	0	0	0
	7	93	0	0	0	0
Up shift	0	0	94	0	6	0
	0	6	93	0	7	0
Down shift	0	0	0	95	0	5
	0	0	0	95	0	5
Up trend	0	0	2	0	98	0
	0	0	4	0	96	0
Down trend	0	0	0	3	0	97
	0	0	0	4	0	96

incorporated to select five candidate features and the priority of each feature along the decision tree was systematically determined. Based on two specific features (*LS-SLP* and *MN-ERR*) of control chart patterns, the generation of induction rules of the decision tree and the development of an expert system to recognize anomaly CCPs are more explicit and feasible. Besides, experimental results indicate that the proposed method can achieve more than 96% classification accuracy in both training and testing data sets while requiring less computational effort compared to other proposed schemes. Hence, the proposed approach is quite promising for the on-line recognition of control chart patterns. More importantly, the proposed method presented in this study has potential to be generalized to medical, financial, and other application of temporal data.

Acknowledgements

The authors are grateful to many helpful comments from two anonymous referees. The research is supported by National Science Council of Taiwan under Grant NSC-95-2416-H-130-019.

References

- Al-Ghanim, A.M. and Ludeman, L.C., Automated unnatural pattern recognition on control charts using correlation analysis techniques. *Comput. Ind. Eng.*, 1997, **32**, 679–690.
- Chang, S.I. and Ho, E.S., A two-stage neural network for process variance change detection and classification. *Int. J. Prod. Res.*, 1999, **37**, 1581–1599.
- Cheng, C.S., A neural network approach for the analysis of control chart patterns. *Int. J. Prod. Res.*, 1997, **35**, 667–697.
- Cook, D.F. and Chiu, C.C., Using radial basis function neural networks to recognize shifts in correlated manufacturing process parameters. *IIE trans.*, 1998, **30**, 227–234.
- Ganesan, R., Das, T.K. and Venkataraman, V., Wavelet-based multiscale statistical process monitoring: A literature review. *IIE trans.*, 2004, **36**, 787–806.

- Gauri, S.K. and Chakraborty, S., A study on the various features for effective control chart pattern recognition. *Int. J. Adv. Manuf. Techn.*, 2006 [In press].
- Guh, R.S. and Hsieh, Y.C., A neural network based model for abnormal pattern recognition of control charts. *Comput. Ind. Eng.*, 1999, **36**, 97–108.
- Guh, R.S. and Tannock, J.D.T., A neural network approach to characterize pattern parameters in process control charts. *J. Intell. Manuf.*, 1999, **10**, 449–462.
- Guh, R.S. and Shiue, Y.R., On-line identification of control chart patterns using self-organized approaches. *Int. J. Prod. Res.*, 2005, **43**, 1225–1254.
- Hassan, A., Shariff, M.N.B., Shaharoun, A.M. and Jamaludin, H., Improved SPC chart pattern recognition using statistical features. *Int. J. Prod. Res.*, 2003, **41**, 1587–1603.
- Ho, E.S. and Chang, S.I., An integrated neural network approach for simultaneous monitoring of process mean and variance shifts: A comparative study. *Int. J. Prod. Res.*, 1999, **37**, 1881–1901.
- Hwang, H.B. and Chong, C.W., Detecting process non-randomness through a fast and cumulative ART-based pattern recognizer. *Int. J. Prod. Res.*, 1995, **33**, 1817–1833.
- Jin, J. and Shi, J., Automatic feature extraction of waveform signals for in-process diagnostic performance improvement. *J. Intell. Manuf.*, 2001, **12**, 257–268.
- Lee, J.M., Yoo, C.K., Choi, S.W., Vanrolleghem, P.A. and Lee, I.B., Nonlinear process monitoring using kernel principal component analysis. *Chem. Eng. Sci.*, 2004a, **59**, 223–234.
- Lee, J.M., Yoo, C.K., Choi, S.W. and Lee, I.B., Statistical process monitoring with independent component analysis. *J. Process Contr.*, 2004b, **14**, 467–485.
- Mitchell, T.M., *Machine Learning*, pp. 52–80, 1997 (McGraw-Hill: New York).
- Pacella, M., Semeraro, Q. and Anglani, A., Adaptive resonance theory-based neural algorithms for manufacturing process quality control. *Int. J. Prod. Res.*, 2004, **42**, 4581–4607.
- Pham, D.T. and Oztemel, E., Control chart pattern recognition using learning vector quantization networks. *Int. J. Prod. Res.*, 1994, **32**, 721–729.
- Pham, D.T. and Wani, M.A., Feature-based control chart recognition. *Int. J. Prod. Res.*, 1997, **35**, 1875–1890.
- Quinlan, J.R., Learning efficient classification procedures and their application to chess endgames. In *Machine Learning: An Artificial Intelligence Approach*, edited by R.S. Michalski, J.G. Carbonell and T.M. Mitchell, Vol. 1, pp. 463–482, 1983 (Morgan Kaufmann: San Mateo, CA).
- Quinlan, J.R., Induction of decision trees. *Machine Learning*, 1986, **1**, 81–106.
- Quinlan, J.R., *C4.5: Programs for Machine Learning*, 1993 (Morgan Kaufmann: Los Altos, CA).
- Wang, C.H., Kuo, W. and Qi, H., An integrated approach for process monitoring using wavelet analysis and competitive neural network. *Int. J. Prod. Res.*, 2007, **45**, 227–244.
- Yang, M.S. and Yang, J.H., A fuzzy-soft learning vector quantization for control chart pattern recognition. *Int. J. Prod. Res.*, 2002, **40**, 2721–2731.
- Yang, J.H. and Yang, M.S., A control chart pattern recognition scheme using a statistical correlation coefficient method. *Int. J. Prod. Res.*, 2005, **48**, 205–221.
- Yousef, A.A., Recognition of control chart patterns using multi-resolution wavelets analysis and neural networks. *Comput. Ind. Eng.*, 2004, **47**, 17–29.