Int J Adv Manuf Technol (2012) 62:291–301
DOI 10.1007/s00170-011-3799-z

ORIGINAL ARTICLE

An expert system for control chart pattern recognition

Monark Bag · Susanta Kumar Gauri · Shankar Chakraborty

Received: 1 April 2011 / Accepted: 21 November 2011 / Published online: 2 December 2011

Abstract This paper focuses on the design and development of an expert system for on-line detection of various control chart patterns so as to enable the quality control practitioners to initiate prompt corrective actions for an out-of-control manufacturing process. Using this expert system developed in Visual BASIC 6, all the nine most commonly observed control chart patterns, e.g., normal, stratification, systematic, increasing trend, decreasing trend, upward shift, downward shift, cyclic, and mixture can be recognized well, employing an optimal set of seven shape features. Based on an observation window of 32 data points, it can plot the control chart, compute the control limits, identify the control chart pattern, calculate the process capability index, determine the maximum run length, and identify the starting point of the maximum run length. After pattern recognition, it can also inform the users about various root assignable causes associated with a particular pattern along with the necessary pre-emptive actions. It opens up wide opportunities for quality improvement and real-time applications in diverse manufacturing processes. This developed expert system is built for a vertical drilling process and its recognition performance is tested using simulated process data.

M. Bag
Indian Institute of Information Technology,
Allahabad 211012, India
e-mail: monark@iiita.ac.in

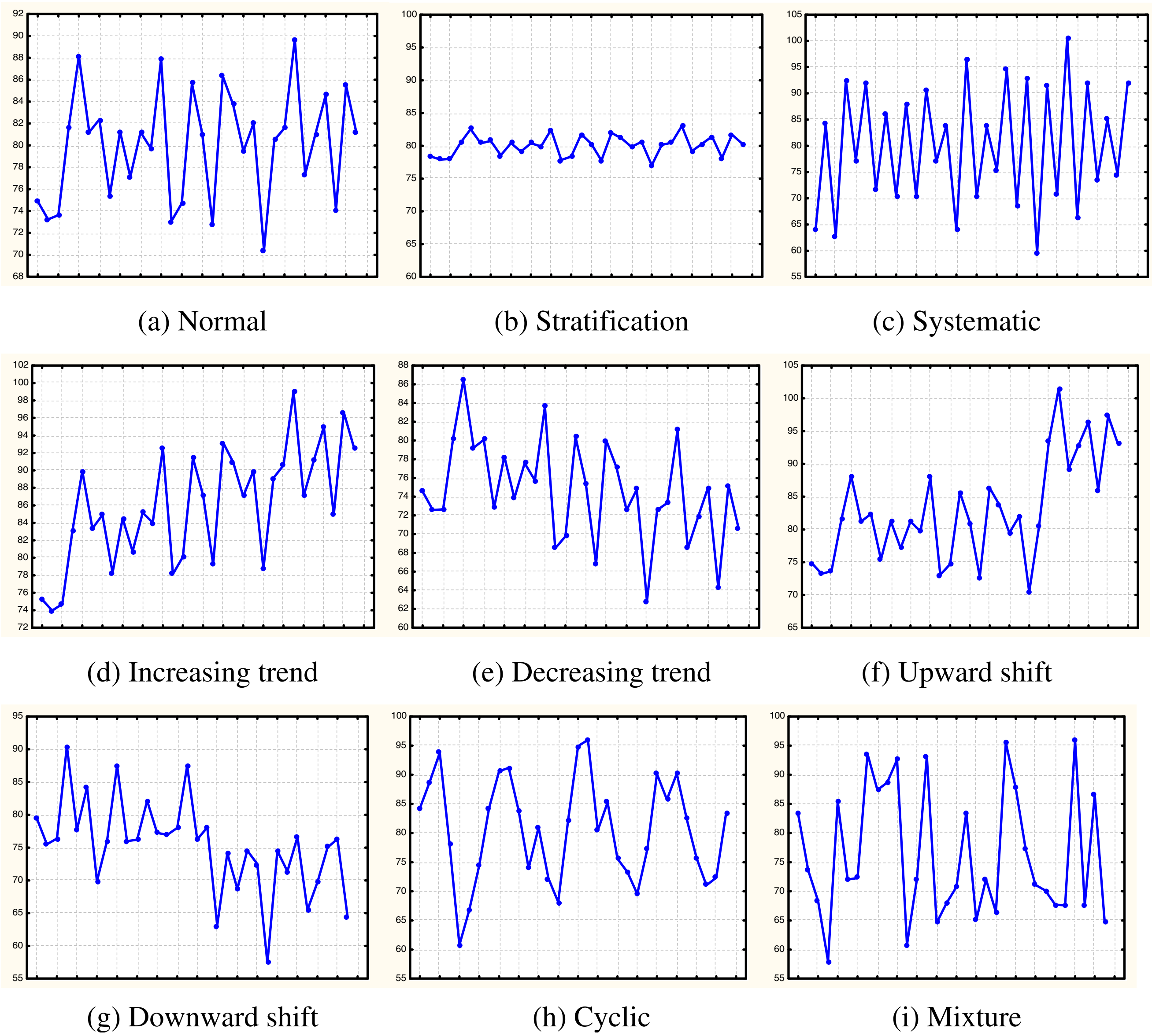
S. K. Gauri
SQC & OR Unit, Indian Statistical Institute,
203, B. T. Road,
Kolkata 700108, India
e-mail: susantagauri@hotmail.com

S. Chakraborty (✉)
Department of Production Engineering, Jadavpur University,
Kolkata 700 032 West Bengal, India
e-mail: s_chakraborty00@yahoo.co.in

Keywords Control chart pattern · Pattern recognition · Shape feature · CART algorithm · Expert system

1 Introduction

In order to achieve global competitive advantage, every organization is trying to improve its product quality at each stage of the manufacturing process. Statistical process control (SPC) is one of the most effective tools of total quality management, which is used to monitor process variations and improve the quality of production. Control charts, mostly in the form of $\overline{X}$ chart, are widely used as aids in maintaining quality and achieving the objective of detecting trends in quality variation before defective parts/products are actually produced. In any continuous manufacturing process, variations from the established standards are mainly of two types. One is assignable cause variation, such as those due to faulty manufacturing equipment or irresponsible personnel or defective material or a broken tool. The other one is normal chance variation, resulting from the inherent non-uniformities that exist in machines or operators or materials or processes. The $\overline{X}$ chart usually exhibits various types of patterns [1, 2], e.g., normal (NOR), stratification (STA), systematic (SYS), increasing trend (UT), decreasing trend (DT), upward shift (US), downward shift (DS), cyclic (CYC), and mixture (MIX), as shown in Fig. 1. Generation of these patterns for a normal manufacturing process can be simulated using the equations, as given in Appendix.



(a) Normal (b) Stratification (c) Systematic

(d) Increasing trend (e) Decreasing trend (f) Upward shift

(g) Downward shift (h) Cyclic (i) Mixture

Fig. 1 Nine control chart patterns

Only the normal pattern is indicative that the process is operating under random chance causes, i.e., in statistical control. The remaining patterns are unnatural and are associated with impending problems requiring pre-emptive actions. The task of control chart pattern (CCP) recognition is basically associated to accurately identify the unnatural CCPs so that prompt corrective actions can be initiated by the operators. Identification and analysis of the unnatural patterns require considerable experience and skill from the part of the quality control practitioners. However, usually, they are lacking the skill and expertise needed for interpretation of the control chart patterns. Therefore, the development of a knowledge-based expert system can help the operators and quality control practitioners to identify the possible sources of variation and take necessary decisive actions.

2 Past researches on expert system for CCP recognition

Evans and Lindsay [3] proposed a framework for developing expert systems for SPC applications. Its knowledge base was partitioned into three sets, i.e., (a) domain-independent, analysis rules for determining whether or not the sample observations indicate a lack-of-control, (b) interpretive rules for analyzing the control chart patterns in terms of process changes, and (c) domain-dependent diagnostic rules for assisting to determine assignable causes and corrective actions. Cheng and Hubele [4] designed an expert system for all the problem-solving aspects of SPC, and addressed all the issues concerning integration of monitoring, interpreting, diagnosing, planning, and statistical consulting for SPC. Kuo and Mital [5] reviewed the existing quality control expert systems and recommended a set of quality engineering techniques that should be used to form the knowledge base. It was pointed out that many organizations had been faced with a shortage of experienced quality controllers and individuals who could train and educate others on SPC techniques. The authors also mentioned that the present trend had been to develop a quality control system and apply it throughout the company (company-wide quality control—CWQC or total quality control—TQC). Pham and Oztemel [6] integrated an expert system and a neural network-based pattern recognizer for analyzing and interpreting control charts. The expert system had an on-line process monitoring package to detect general out-of-control situations and a diagnosis module to suggest corrective actions. The pattern recognizer was an on-line system comprising two neural networks and a heuristics module designed to identify incipient process abnormalities from control chart patterns. Swift and Mize [7] developed an expert system to detect and analyze various patterns of variation (trend, cycle, mixture, shift, stratification, and systematic) that could occur in manufacturing quality control charts. Statistical significance tests as interpretive rules were also used to determine the patterns of variation. Once the pattern was identified, the expert system could provide the user with possible causes for the out-of-control process, magnitude of the out-of-control condition, and starting and stopping points of the recognized pattern. Hooks et al. [8] presented a model of an expert system to enhance dimensional tolerancing and data analysis in quality control. It would also serve a dual role as a technological link to a CIM environment through the use of IGES-CAD data. Tsacle and Aly [9] designed and developed an expert system to advise the hospital personnel how to measure and control their processes effectively using different types of control charts, such as $\overline{X}$ and R-charts, P-charts, IC-charts and Individual-X, and Moving range charts. A complete step-by-step interactive session with the expert system was also shown. Finally, its effectiveness was evaluated and the feasibility of linking it directly to other SPC software packages was explored. Vosniakos and Wang [10] proposed a quality information system framework for mechanical component manufacturing industries to support both the planning and operation activities. Guh et al. [11] developed a hybrid intelligent tool (IntelliSPC) in which a neural network-based control chart pattern recognition system, an expert system-based control chart alarm interpretation system and a quality cost simulation system were integrated for on-line process control. It was designed to provide the quality control practitioners with the status of the process (in-control or out-of-control), plausible causes for the out-of-control situation, and cost-effective actions against the out-of-control situation. Paladini [12] presented the guidelines for structuring a decision supporting expert system to help with those decisions related to determine the need or convenience of carrying out inspection. Once the opportunity to carry it out was defined, the expert system could help the user to select the type of inspection to adopt from amongst (a) automatic or sensorial inspection, (b) inspection by samples or complete, (c) acceptance or rectifying, and (d) inspection by attributes or by variables. Although the past researchers have proposed different frameworks for developing expert systems for quality control and process monitoring, the development of a real time and user interactive expert system for effective control chart pattern recognition has been first accredited to Guh et al. [11]. That neural network-based CCP recognition system has the drawbacks of complicated network architecture, incomprehensible and unchangeable decision rules which compel to the development and augmentation of a simple, easily understandable, and more dynamic expert system for CCP recognition.

This paper focuses on the design and development of an expert system for on-line application which can detect all the nine most commonly observed control chart patterns, i.e., NOR, STA, SYS, UT, DT, US, DS, CYC, and MIX to enable the quality control practitioners to initiate prompt corrective actions for an out-of-control process. All these nine types of CCPs are simulated and the values of an optimal set of seven shape features are extracted which are then analyzed using the classification and regression tree (CART) algorithm to formulate the related decision rules for the expert system for CCP recognition. This expert system is built for a vertical milling process and its recognition performance is validated using simulated process data.

3 Determination of the decision rules for expert system

Decision rules are required for detection of the abnormal process conditions. Most of the past researchers [3, 6, 7] used statistical properties as the decision rules. On the other hand, Guh et al. [11] adopted neural network-based algorithms for detection of the abnormal control chart patterns. The use of rules based on statistical properties in the expert system has the difficulty that similar statistical properties may be derived for some patterns of different classes, which may create problems of incorrect recognition. The advantage with neural network is that it is capable of handling noisy measurements requiring no assumption about the statistical distribution of the monitored data. However, the major disadvantage with neural network is the difficulty in understanding how a particular classification decision has been reached and also in determining the details of how a given pattern resembles with a particular class. Pham and Wani [13] and Gauri and Chakraborty [14, 15] highlighted that each control chart pattern has its own geometric shape and various related features can represent this shape. Different patterns can, therefore, be efficiently discriminated based on these shape features extracted from the control chart plot. The advantage of shape features is that those can be extracted from lesser number of observations without losing order of the data. Gauri and Chakraborty [16] later carried out an extensive research work on the feature-based approach for CCP recognition. They extracted 30 shape features from an observation window of 32 data points in such a way that the feature values are independent of the process mean and standard deviation. Then, they selected an optimal set of seven shape features using a CART algorithm-based [17] systematic approach. Based on those seven shape features, they developed the related heuristic rules by which eight types of CCPs, e.g., NOR, STA, SYS, UT, DT, US, DS, and CYC can be discriminated efficiently.

Since the extracted shape features represent the main characteristics of the original data in a condensed form, the feature-based heuristic rules facilitate efficient pattern recognition. Moreover, the feature-based heuristic approach has a distinct advantage that the quality control practitioners can clearly understand how a particular pattern has been identified while using the relevant shape features. Therefore, it is decided to use the feature-based heuristics as the decision rules for the developed expert system. However, since this expert system aims at recognizing the MIX pattern too, the heuristic rules, as proposed by Gauri and Chakrabory [16], cannot be used here. Therefore, the feature-based heuristic rules are developed afresh.

3.1 Selection of the important shape features

All the 30 shape features, as considered by Gauri and Chakraborty [16], are first critically examined to evaluate if some of those features are at all capable of discriminating the MIX pattern or it is necessary to define new features for recognizing the MIX pattern. It is observed that some of the features, as proposed by Gauri and Chakraborty [16], can well discriminate the MIX pattern from all other patterns. Therefore, it is decided to select the most appropriate set of features from those 30 shape features.

The CART-based systematic approach, as adopted by Gauri and Chakraborty [16], is applied here for the purpose of selecting the most appropriate set of shape features that can recognize all the nine CCPs, including the MIX pattern. Interestingly, it is found that a different set of seven shape features can well discriminate all the nine CCPs. Out of these seven features, four features are extracted without segmentation of the observation window, two features are extracted with pre-defined segmentation of the observation window into four segments of equal size, and one feature is extracted with criterion-based segmentation of the observation window. These shape features are described as below:

(a) Sign of slope of the least square (LS) line representing the overall pattern (SB):

The slope (B) of the LS line fitted to the data points in an observation window is given by the following equation:

$$B = \sum_{i=1}^{N} y_i(t_i - \bar{t}) / \sum_{i=1}^{N} (t_i - \bar{t})^2 \quad (1)$$

where $t_i = ic$ ($i = 1,2,3,\ldots,N$) is the distance of ith time point of observation from the origin, c is a constant linear distance used to represent a given sampling interval on the control chart plot, y_i is the observed value of a quality characteristics at ith time point, N is the size of the observation window and $\bar{t} = \sum_{i=1}^{N} t_i/N$. Then, the feature SB can be defined as follows: SB=1 if B≥0, and SB=0 if B<0. This feature can better discriminate UT versus DT and US versus DS patterns.

(b) Ratio between variance of the data points in the observation window (SD^2) and mean sum of squares of errors (MSE) of the LS line representing the overall pattern (RVE):

$$RVE = \left[\frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2\right] \Bigg/ \left[\frac{1}{N-2}\left\{\sum_{i=1}^{N}(y_i - \bar{y})^2 - \frac{\left(\sum_{i=1}^{N} y_i(t_i - \bar{t})\right)^2}{\sum_{i=1}^{N}(t_i - \bar{t})^2}\right\}\right] \quad (2)$$

The magnitude of RVE for NOR, STA, SYS, CYC, and MIX patterns is approximately 1, while for trend and shift patterns, it is greater than 1.

(c) Area between the overall pattern and the LS line per interval in terms of SD^2 (ALSPI):

$$ALSPI = [ALS/(N-1)]/SD^2;\ SD^2 = \sum_{i=1}^{N} (y_i - \bar{y})^2/(N-1) \tag{3}$$

where ALS is the area between the pattern and fitted LS line. The value of ALS can be easily computed by summing the areas of the triangles and trapeziums that are formed by the LS line and overall pattern. The magnitude of ALSPI is the highest for STA pattern, lowest for SYS pattern, and intermediate for all other patterns.

(d) Proportion of the sum of number of crossovers to mean line and the LS line (PSMLSC):

$$PSMLSC = \sum_{i=1}^{N-1} \left(O_i + O_{\rm i}'\right)/2N \tag{4}$$

where $O_{\rm i}$=1 if $(y_i - \bar{y})(y_{i+1} - \bar{y}) < 0$, otherwise, $O_{\rm i}$=0, and $\bar{y}$ is the mean value of N data points; $O_{\rm i}' = 1$ if $\left(y_i - y_i'\right) \times \left(y_{i+1} - y_{i+1}'\right) < 0$, otherwise, $O_i' = 0$, and y_i' is the least square estimate of ith data point. The PSMLSC value is maximum for SYS pattern, intermediate for NOR, STA, UT, DT, and MIX patterns, and lesser for CYC, US, and DS patterns.

The following two features are extracted with segmentation of the observation window into four segments of equal size. The behavior of the process within a segment can be represented by the midpoint of the segment, which is given as:

$$\left[\left\{\sum_{i=k}^{k+(N/4)-1} t_{\rm i}/(N/4)\right\}, \left\{\sum_{i=k}^{k+(N/4)-1} y_{\rm i}/(N/4)\right\}\right]$$

where k=1, (N/4+1), (2 N/4+1), (3 N/4+1) for the first, second, third, and fourth segments, respectively. A combination of two midpoints can be obtained in $C_2^4 = 6$ ways implying that six straight lines can be drawn passing through the midpoints of these four segments. Similarly, six subsets of N/2 data points can be formed taking a combination of two segments in six ways. So, six LS lines can also be fitted to six subsets of N/2 data points.

(e) Range of slopes of straight lines passing through six pairwise combinations of midpoints of four equal segments (SRANGE):

$$SRANGE = \text{maximum}\left(s_{jk}\right) - \text{minimum}\left(s_{jk}\right);\ (j = 1, 2, 3;\ k = 2, 3, 4;\ j < k) \tag{5}$$

where $s_{\rm jk}$ represents the slope of the straight line passing through the midpoints of jth and kth segments. The magnitude of SRANGE will be higher for shift patterns than trend patterns. The value of SRANGE will also be higher for CYC pattern than NOR, STA, SYS, and MIX patterns, unless each segment of CYC pattern consists of a complete cycle.

(f) Ratio of mean sum of squares of errors (MSE) of the LS line fitted to overall data and average MSE of the LS lines fitted to six subsets of N/2 data points (REAE):

$$REAE = MSE/\left[\sum_{j,k} MSE_{jk}/6\right];\quad (j = 1, 2, 3;\ k = 2, 3, 4; j < {\rm k}) \tag{6}$$

where $MSE_{\rm jk}$ is the mean sum of squares of errors of the LS line fitted to the observations in jth and kth segments. The magnitude of REAE is greater than 1 for CYC and shift patterns, and about 1 for NOR, STA, SYS, and trend patterns. In case of MIX pattern, the value of REAE is less than 1. Thus, REAE value can differentiate MIX pattern from all other patterns.

The last feature is extracted using a criterion-based segmentation of the observation window into two segments, where the defined criterion is minimization of the pooled MSE (PMSE) of the two LS lines fitted to the two segments. In this segmentation approach, sizes of the two segments may vary in order to satisfy the desired criterion. Assuming that at least 10 data points are required for fitting an LS line, the least square lines are fitted to all possible two segments and the segmentation which leads to the minimum PMSE is selected and then the following feature is extracted.

(g) Sum of absolute slope difference between the LS line representing the overall pattern and the individual line segment (SASDPE):

$$SASDPE = \sum_{j=1}^{2} \left|B - B_j\right| (j = 1, 2) \tag{7}$$

where B is the slope of the LS line representing the overall pattern and $B_{\rm j}$ is the slope of the LS line fitted to jth segment. The magnitude of SASDPE is higher for shift patterns than trend patterns. On the other hand, the value of SASDPE will be higher for MIX, CYC, and SYS patterns than NOR and STA patterns.

For the purpose of assessing the degree of association between the selected features, a set of 9,000 sample patterns consisting of all the nine types of patterns is simulated. From each sample pattern in this set, all the selected seven features are extracted. So 9,000 values for each of the seven features are obtained and then the correlation coefficients between all the pairs of the selected features are estimated. Table 1 shows the values of pairwise correlation coefficients between the selected seven features. The table reveals that

Table 1 Pairwise correlation coefficients between seven shape features

Shape feature	SB	RVE	ALSPI	PSMLSC	SRANGE	REAE	SASDPE
SB	1.00	0.01	−0.01	0.23	−0.16	−0.09	−0.05
RVE		1.00	−0.26	−0.34	0.03	0.18	0.01
ALSPI			1.00	−0.04	−0.34	−0.05	−0.41
PSMLSC				1.00	−0.43	−0.36	−0.19
SRANGE					1.00	0.58	0.36
REAE						1.00	0.13
SASDPE							1.00

the degree of association between these seven shape features is considerably low. Therefore, these seven shape features are considered to be appropriate for developing the decision rules of the expert system for CCP recognition.

3.2 Derivation of the decision rules for CCP recognition

Ideally, sample patterns should be collected from a real manufacturing process. However, a large number of patterns is required for deriving and validating the decision rules for CCP recognition. Since those are not economically available from the manufacturing processes, simulated data are often used. This is a common approach adopted by the other researchers also.

A set of 1,000 series of standard normal variate of length 32 is generated first. These standard normal data are then transformed to 1,000 series of NOR, STA, SYS, UT, DT, US, DS, CYC, and MIX patterns using the equations given in Appendix. Thus, a set of 9,000 (1,000 × 9) sample patterns is generated from 1,000 series of standard normal variate. Similarly, nine more sets of sample patterns of size 9,000 in each are generated for developing the decision rules of the expert system. The only difference between these 10 sets of patterns is in the random generation of standard normal variate and values of different pattern parameters within their respective limits. From each set of the sample patterns, values of the selected shape features are extracted, which are then subjected to CART analysis (available in STATISTICA package). The CART analysis results in a binary decision tree, which is nothing but a set of heuristics based on the values of the selected features for CCP classification. The specifications for the CART analysis are the same as taken by Gauri and Chakraborty [16]. Similarly, the values of the selected seven features are extracted from the other nine sets of sample patterns and subjected to CART analysis. These result in 10 different decision trees, i.e., 10 sets of decision rules for CCP recognition. The percentage of correct classification given by the decision trees is shown in Table 2. It is observed that the set of decision rules represented by the decision tree number 4 (shown in Fig. 2) results in the maximum percentage of correct classification, and so this set of decision rules is selected for deployment in the expert system.

4 The developed expert system

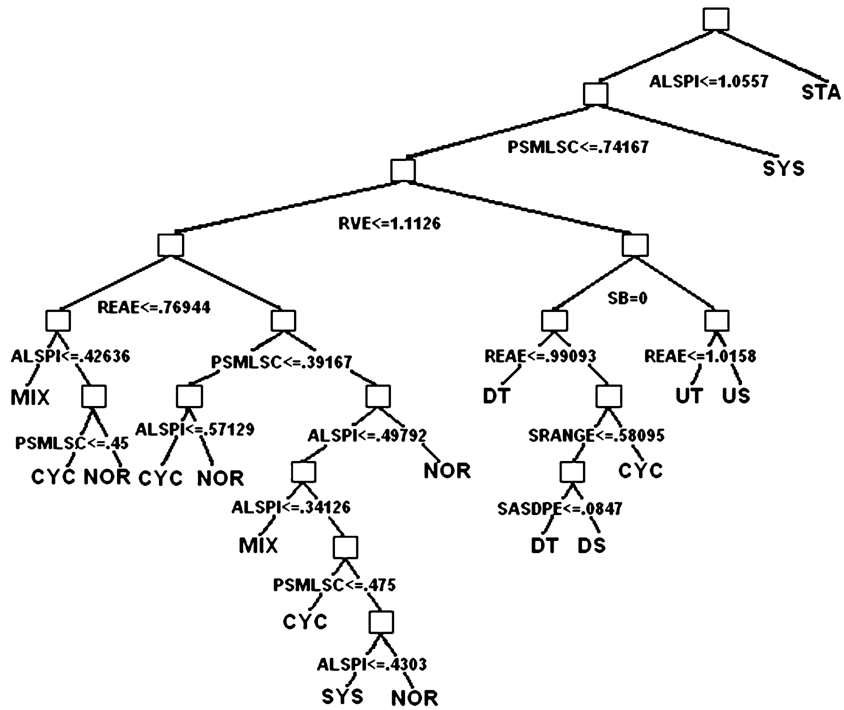
An expert system is a computer program with the capability to emulate the decision making and reasoning ability of a human expert. Unlike conventional programs, expert systems are designed and developed to solve complex problems by reasoning about knowledge. An expert system has a unique structure, different from the traditional programs. It has three main components, i.e., user interface, knowledge base, and inference engine. The user interface acts like a dialog box to interact and communicate with the users. The knowledge base (static or dynamic) stores the valuable knowledge of the human experts in the form of a database and the inference engine reasons about the knowledge base. The inference engine is a computer program designed to produce a reasoning based on simple if…then rules. In simple expert systems, the user provides the data and receives the desired results immediately, and the reasoning is invisible.

The main advantage of expert system is its ability to interact with the user as it asks the user step-by-step about the problem and identify the solution gradually. In expert

Table 2 Performance of 10 different decision trees

Decision tree number	Number of terminal nodes	Percentage of correct classification
1	22	94.69
2	20	93.27
3	23	94.43
4	18	95.27
5	22	94.94
6	19	94.20
7	21	95.10
8	23	94.69
9	21	95.12
10	20	94.74

Fig. 2 Selected decision rules for the expert system in the form of a tree



system, the knowledge base can be written much faster than a conventional program. The reliability of an expert system is the same as the reliability of a knowledge base and is much higher than a classical program. It has also high scalability; as the rules are written in plain language, those can be easily modified, added, or deleted. As it is run by a true logic, it can explain to the user why a question has been asked and how a final deduction has been arrived at. Lastly, valuable knowledge may disappear with the death, resignation, or retirement of an expert. But when stored in the database of an expert system, it becomes eternal.

The expert system has a major drawback. Accumulation of knowledge and converting this into meaningful rules are often quite difficult. The expert knowledge is not well understood; there is a lack of rules, rules are contradictory, and some are poorly written and unusable.

With the aim to effectively recognize all the nine control chart patterns, an expert system is developed in Visual BASIC 6 and it is made platform-independent. The expert system can plot the related control chart, compute the control limits, identify the concerned pattern, calculate the process capability index, estimate the maximum run length value, and identify the starting point of maximum run length. It can also detect an out-of-control situation and identify the probable root causes behind that situation while guiding the quality control practitioners with the possible remedial/corrective actions. This expert system is designed to monitor and diagnose the occurrence of different CCPs for a vertical drilling process. The root assignable causes [18] behind the generation of these CCPs are stored in the knowledge base and the valuable opinions of the experts are used to make a link between the type of the identified pattern and associated assignable causes with the necessary pre-emptive actions. Its data input section consists of 32 boxes where an operator or quality control practitioner can easily enter the $\overline{X}$ values from a particular manufacturing process. In this paper, 32 $\overline{X}$ values (with a subgroup size 4) are taken with a process mean of 80 mm. It means that in each time interval, the operator collects four products, measures the corresponding dimensions, and enters the mean dimensional value for that interval in the data input box. The expert system consists of some functional buttons along with a picture box. After entering all the $\overline{X}$ values in the data input box, pressing the "Plot" functional bottom automatically draws the relevant control chart in the picture box. The "Control Limit" key computes the values of the central, lower control, and upper control limits.

This expert system has also the option where the user can set the lower and upper specification limits as required for the concerned manufacturing process. Based on these specification and control limits, it can compute the process capability index (C_p) when the user presses the "C_p" button. When "Max RL" and "Starting Point of Max RL" functional bottoms are clicked, the maximum run length and the starting pint of that maximum run length in the control chart plot are respectively displayed. These help the user to detect the occurrence of an unnatural pattern with its starting point. After pressing the "Feature Value" key, the expert system calculates the considered seven shape feature values and subsequently displays those values. Based on the computed values of these seven shape features and the in-built decision rules, it can now recognize the exact control chart pattern, when the user clicks the "Identified Pattern" button. According to the decision rules, if the pattern has ALSPI value greater than 1.0557, it would be a stratification pattern. Similarly, if the pattern has ALSPI value less than or equal to 0.42636, REAE value less than or equal to 0.76944, RVE value less than or equal to 1.1126, and PSMLSC value less than or equal to 0.74167, then it would identify the pattern as a mixture. The hard copy of the control chart pattern can be made available after clicking the "Print" button within the picture box. The "RESET" button allows the user to enter a new set of pattern data in the input box and the "EXIT" button closes the expert system. After clicking the 'Pattern Details' bottom, another pop-down window would appear containing four functional keys, i.e., "Parameter Value/s," "Possible Cause/s," "Remedial Action", and "Print." The Parameter Value/s button displays the values of different parameter(s) associated with an identified control chart pattern. Pressing of the Possible Cause/s key identifies various assignable causes as present behind the generation of that abnormal pattern for the vertical drilling process and similarly, the Remedial Action button guides the user to take various pre-emptive actions to avoid manufacturing of defective drilled components. The hard copy of these details can be obtained after clicking the Print button.

Figure 3 shows the output of the expert system where based on the input data and decision rules, the control chart is identified having a decreasing trend pattern. The Pattern Details screen display for that identified decreasing trend pattern is exhibited in Fig. 4. The gradient (g) for that pattern is calculated as 1.8229. Figure 5 gives the screen

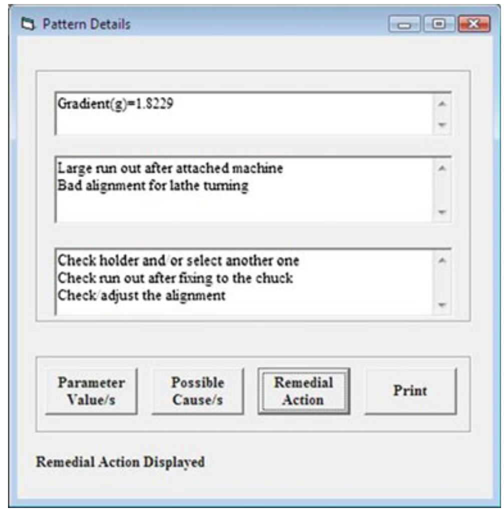
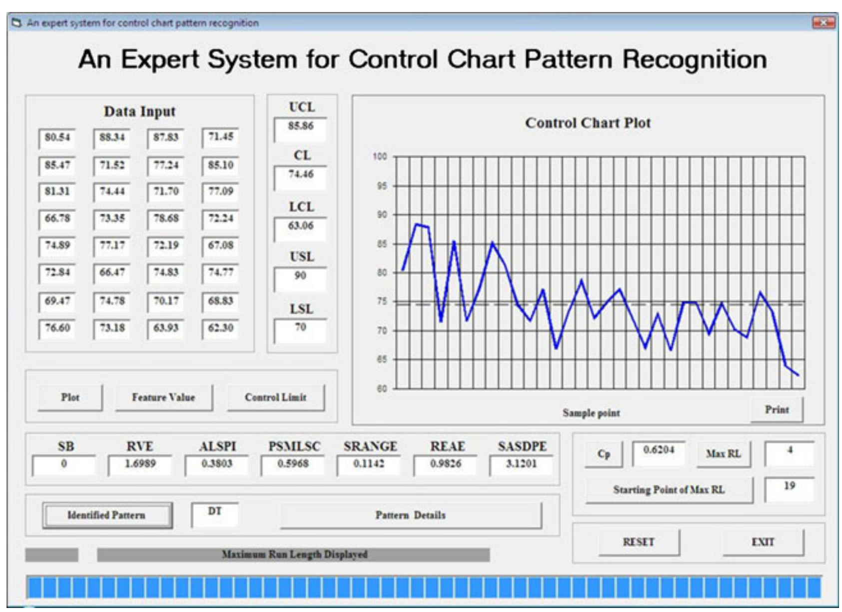


Fig. 4 Pattern details for a decreasing trend pattern

display for an identified mixture pattern. The process mean is determined as 84.1756 mm which has been shifted from the set process mean of 80 mm because of the occurrence of the mixture pattern. In Fig. 5, by double clicking on the displayed control chart, the user can highlight various data points on the identified pattern. Figure 6 displays the expert system output where the pattern is identified as cyclic, having period (T)=11 and amplitude (a)=6.615. Although, this expert system is developed for monitoring and diagnosing various root assignable causes behind generation of different control chart patterns for a vertical drilling process, it can also be applicable to other manufacturing processes only by changing the knowledge base.

Traditionally, control chart patterns are interpreted visually and the subjectivity associated with the visual analysis

Fig. 3 Screen display for a decreasing trend pattern



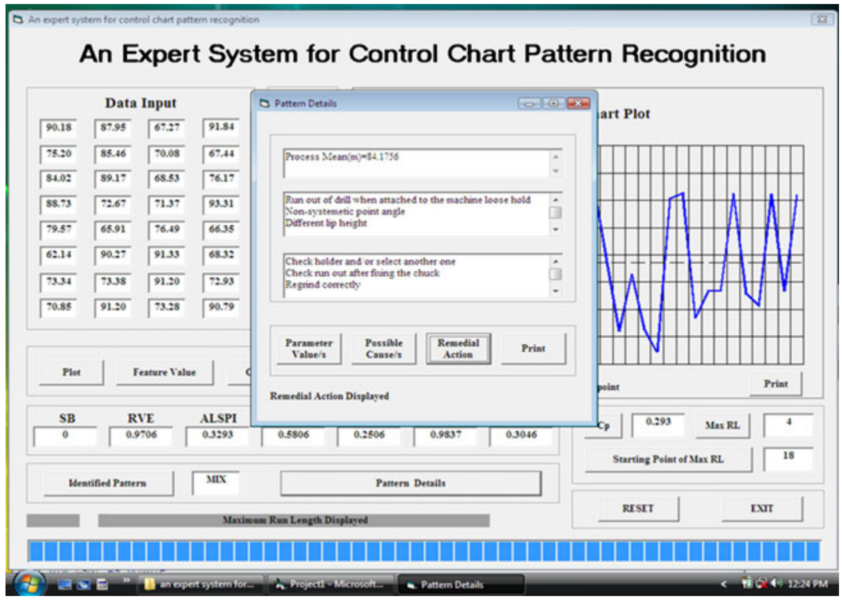


Fig. 5 Screen display for a mixture pattern

of the patterns usually leads to a large amount of error. Wrong classification of patterns often costs heavily. This is because the potential causes for different types of patterns are different. For example, the cyclic pattern is likely to be observed when some process variables are present at first, and then absent on a more or less regular basis. The causes behind cyclic pattern include temperature and humidity changes, operator fatigue, rotation of operators, electrical fluctuations, etc. On the other hand, the upward or downward shift pattern may result from the introduction of new materials, machines or any other new processing variable, a change in the inspection method or standard, or a change in the skill, attentiveness or motivation of the operators, or change in the level of any existing variable.

The basic purpose of the developed expert system is to minimize the error by eliminating the subjectivity in the interpretation of control chart patterns. It is, therefore, important to assess the performance of the developed expert system rigorously before its actual deployment in real-time process monitoring applications. Ideally, the performance of the expert system should be evaluated by testing with a large number of datasets collected from real-time manufacturing processes. Again, these datasets should be collected from normal as well as various abnormal process conditions. But collection of a large number of real-time process data representing various process conditions is extremely difficult. So it is planned to assess the performance of the expert system using simulated process data.

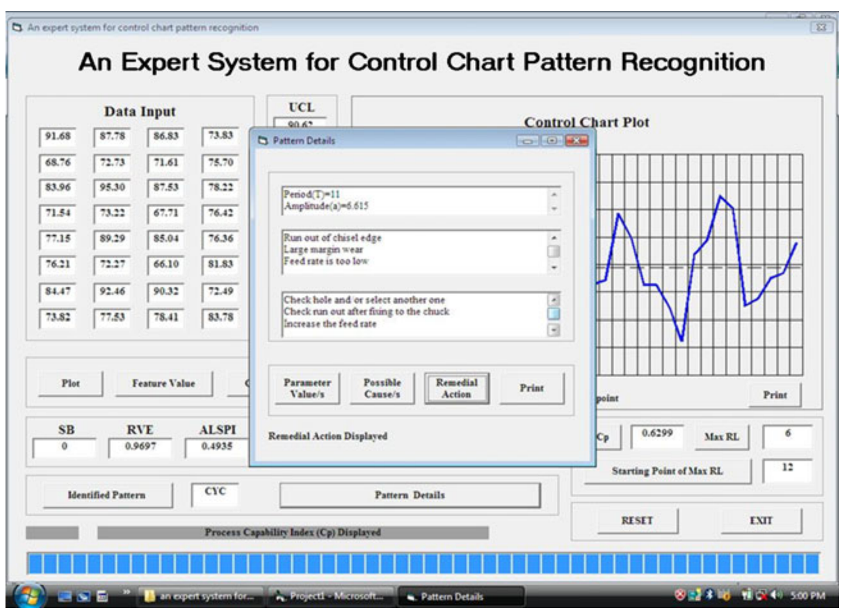


Fig. 6 Expert system output for a cyclic pattern

For the purpose of evaluating the CCP recognition performance of the expert system, a set of 1,000 series of standard normal data of 32 observations in each is simulated first. Using the same 1,000 series of standard normal data, sample data for NOR, STA, SYS, UT, DT, US, DS CYC, and MIX patterns are generated using the equations shown in Appendix. Thus, the set of sample pattern data for evaluating the performance of the expert system contains 9,000 (1,000×9) series of data. Each series of pattern data is separately fed into the expert system and the outcome in terms of recognition of the control chart pattern is then observed. The overall results are summarized in Table 3 as a confusion matrix. The confusion matrix depicts the tendency of the expert system to classify a known pattern into a correct class or into any of the other eight possible (wrong) classes. Each entry in the confusion matrix is expressed in percentage. Examination of the results in Table 3 indicates that the developed expert system is well capable in recognizing NOR, STA, SYS, UT, DT, and MIX patterns. However, substantial amount of confusion occurs while recognizing shift and MIX patterns. The shift patterns are confused with trend patterns, and CYC patterns are confused with MIX patterns. Upward shift patterns (5.6%) are incorrectly recognized as increasing trend patterns and 5.3% of downward shift pattern are incorrectly recognized as decreasing trend patterns. On the other hand, 3.1% of CYC patterns are incorrectly recognized as MIX patterns and 2.9% of MIX patterns are wrongly classified as CYC patterns.

The results in Table 3 suggest that further improvements are required for reduction of confusion between the shift and trend patterns. Attempts should also be made to reduce the confusion between CYC and MIX patterns. One possible way can be the identification of some new features that would be more powerful in discriminating shift patterns from trend patterns and CYC patterns from MIX patterns, and then developing the decision rules. Another possible approach may be the development of a hybrid intelligent system, which can integrate two or more artificial intelligence tools (e.g., ANN and heuristics) together for the purpose of CCP recognition.

Montgomery [1] considered two sets of data on inside diameter of piston rings for an automotive engine produced by a forging process and observed that for the first case, the forging process was under statistical control, and for the second case, the process was out-of-control having a shift in the process mean. These same datasets are tested using the developed expert system and it is observed that for the first dataset, the identified control chart pattern is normal with central, lower, and upper control limits as 74.0014, 73.8662, and 74.0149 mm, respectively. On the other hand, the identified pattern for the second dataset is upward shift with central limit=74.0040 mm, lower control limit=73.9998 mm, upper control limit=74.0081 mm, shift magnitude (s)=0.01371, maximum run length=7, and starting point of maximum run length=33. These exactly corroborate with the findings of Montgomery [1].

5 Conclusions

An expert system is developed in Visual BASIC 6 which is capable to plot the control chart, compute the control limits, identify the control chart pattern, calculate the process capability index, determine the maximum run length, and identify the starting point of the maximum run length. It employs an optimal set of only seven shape features to identify the nine most commonly observed control chart patterns. The major advantage of this expert system is that it cannot only recognize a particular control chart pattern, but also display various assignable causes behind that pattern along with the necessary remedial actions to help the quality control practitioners for effective decision making. It would be widely acceptable to any manufacturing process to prevent production of defective products while improving the overall quality level. It would also act as a poka-yoke

Table 3 Confusion matrix for the expert system

Known pattern class	Identified pattern class								
	NOR	STA	SYS	CYC	UT	US	DT	DS	MIX
NOR	95.5	0.1	0.2	1.6	0.6	0.6	0.4	0.4	0.6
STA	0.3	99.6	0	0.1	0	0	0	0	0
SYS	0.2	0	98.5	0.2	0	0	0	0	1.1
CYC	1.1	0	1.3	93.3	0	1	0	0.2	3.1
UT	1.0	0	0	0.4	96.2	2.4	0	0	0
US	0.6	0	0	1.3	5.6	92.1	0	0	0.4
DT	1.3	0	0	0.7	0	0	94.7	3.3	0
DS	0.4	0	0.6	2.0	0	0	5.3	91.6	0.1
MIX	0.2	0	1.2	2.9	0	0.1	0	0.2	95.4

device for the manufacturing industry to achieve the goal of zero defect and total quality management. Future research may make improvements in several directions. Development of an expert system based on identified statistical features, and making a comparative performance study between the statistical feature-based and shape feature-based expert systems may be the scope of this paper. Future research may address the inclusion of concurrent patterns where two or more unnatural patterns exist simultaneously (e.g., a systematic pattern with a cyclic behavior).

Appendix

Equations for generation of various patterns in a given normal process

Suppose, the value of a standard normal variate at *i*th (i=1,2,…,32) time point is r_i, and the observed value at *i*th time point is y_i. Then, various patterns of length 32 for a normal process with mean μ and standard deviation σ can be generated using the following equations:

(a) Normal pattern $$y_i = \mu + r_i\sigma \quad (8)$$

(b) Stratification pattern $$y_i = \mu + r_i\sigma' \quad (9)$$

(c) Systematic pattern $$y_i = \mu + r_i\sigma + d \times (-1)^i \quad (10)$$

(d) Increasing trend pattern $$y_i = \mu + r_i\sigma + ig \quad (11)$$

(e) Decreasing trend pattern $$y_i = \mu + r_i\sigma - ig \quad (12)$$

(d) Upward shift pattern $$y_i = \mu + r_i\sigma + ks \quad (13)$$

(e) Downward shift pattern $$y_i = \mu + r_i\sigma - ks \quad (14)$$

(f) Cyclic pattern $$y_i = \mu + r_i\sigma + a\ \sin(2\pi i/T) \quad (15)$$

(g) Mixture pattern $$y_i = \mu + r_i\sigma + (-1)^w m \quad (16)$$

where,

- σ' random noise for stratification pattern
- a amplitude of cyclic variation
- g magnitude of gradient for the trend pattern
- d magnitude of the systematic pattern
- k parameter determining the shift position
- s magnitude of the shift
- i discrete time point at which the pattern is sampled
- T period of a cycle
- m magnitude of the mixture pattern
- w a binary integer value dependent on a random number p ($0<p<1$) and a pre-specified probability value $b=mp$, which determines the shifting between distributions. The value of b is fixed as 0.4, and thus, $w=0$ if $p<0.4$ and $w=1$ if $p\geq0.4$.

In this paper, for simulation of various patterns, the values of different process parameters are chosen as follows: $\mu=80$, $\sigma=5$, $0.2\sigma\leq\sigma'\leq0.4\sigma$, $1\sigma\leq d\leq3\sigma$, $0.05\sigma\leq g\leq0.1\sigma$ (for UT), $-0.1\sigma\leq g\leq-0.05\sigma$ (for DT), $1.5\sigma\leq s\leq2.5\sigma$ (for US), $-2.5\sigma\leq s\leq-1.5\sigma$ (for DS), P=9, 17 or 25, $1.5\sigma\leq a\leq2.5\sigma$, T= 8 or 16, $1.5\sigma\leq m\leq2.5\sigma$, and $0\leq p\leq1$.

References

1. Montgomery DC (2001) Introduction to statistical quality control. Wiley, New York
2. Electric W (1958) Statistical quality control handbook. Western Electric Company, Indianapolis
3. Evans JR, Lindsay WM (1988) A framework for expert system development in statistical quality control. Comput Ind Eng 14:335–343
4. Cheng C-S, Hubele NF (1992) Design of a knowledge-based expert system for statistical process control. Comput Ind Eng 22:510–517
5. Kuo T, Mital A (1993) Quality control expert systems: a review of pertinent literature. J Intell Manuf 4:245–257
6. Pham DT, Oztemel E (1995) Integrated neural network and expert system toll for statistical process control. Proc Inst Mech Eng B J Eng Manuf 209:91–97
7. Swift JA, Mize JH (1995) Out-of-control pattern recognition and analysis for quality control charts using lisp-based systems. Comput Ind Eng 28:81–91
8. Hooks K, Rabelo L, Velasco T (1995) An expert system framework for a CIM based quality inspection system. Comput Ind Eng 29:159–163
9. Tsacle EG, Aly NA (1996) An expert system model for implementing statistical process control in the health care industry. Comput Ind Eng 31:447–450
10. Vosniakos G, Wang J (1997) A software system framework for planning and operation of quality control in discrete part manufacturing. Comput Integr Manuf Syst 10:9–25
11. Guh R-S, Tannock JDT, O'Brien C (1999) IntelliSPC: a hybrid intelligent tool for on-line economical statistical process control. Expert Syst Appl 17:195–212
12. Paladini EP (2000) An expert system approach to quality control. Expert Syst Appl 18:133–151
13. Pham DT, Wani MA (1997) Feature-based control chart pattern recognition. Int J Prod Res 35:1875–1890
14. Gauri SK, Chakraborty S (2006) Feature-based recognition of control chart patterns. Comput Ind Eng 51:726–742
15. Gauri SK, Chakraborty S (2007) A study on the various features for effective control chart pattern recognition. Int J Adv Manuf Technol 34:385–398
16. Gauri SK, Chakraborty S (2009) Recognition of control chart patterns using improved selection of features. Comput Ind Eng 56:1577–1588
17. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth and Brooks, Monterey
18. Technical Resource for Manufacturing Professionals (www.carbidedepot.com/formulas-drills-troubelshoot.htm)