Issues in Development of Artificial Neural Network-Based Control Chart Pattern Recognition Schemes

Article *in* European Journal of Scientific Research · January 2010

2 authors:

Ibrahim Masood
Universiti Tun Hussein Onn Malaysia

Adnan Hassan
Universiti Teknologi Malaysia

European Journal of Scientific Research
ISSN 1450-216X Vol.39 No.3 (2010), pp.336-355

http://www.eurojournals.com/ejsr.htm

Issues in Development of Artificial Neural Network-Based Control Chart Pattern Recognition Schemes

Ibrahim Masood
Faculty of Mechanical and Manufacturing Engineering
Universiti Tun Hussein Onn Malaysia
Batu Pahat, Johor, Malaysia
E-mail: ibrahim@uthm.edu.my
Tel: +607-4537700; Fax: +607-4536080

Adnan Hassan
Faculty of Mechanical Engineering
Universiti Teknologi Malaysia, Skudai, Johor, Malaysia
E-mail: adnan@fkm.utm.my

Abstract

Control chart pattern recognition has become an active area of research since late 1980s. Much progress has been made, in which there are trends to heighten the performance of artificial neural network (ANN)-based control chart pattern recognition schemes through feature-based and wavelet-denoise input representation techniques, and through modular and integrated recognizer designs. There is also a trend to enhance it's capability for monitoring and diagnosing multivariate process shifts. However, there is a lack of literature providing a critical review on the issues associated to such advances. The purpose of this paper is to highlight research direction, as well as to present a summary of some updated issues in the development of ANN-based control chart pattern recognition schemes as being addressed by the frontiers in this area. The issues highlighted in this paper are highly related to input data and process patterns, input representation, recognizer design and training, and multivariate process monitoring and diagnosis. Such issues could be useful for new researchers as a starting point to facilitate further improvement in this area.

Keywords: Artificial Neural Network, Control Chart Pattern Recognition, Feature-Based, Wavelet-Denoise, Modular and Integrated Recognizers, Multivariate Process Monitoring and Diagnosis

1. Introduction

Development in manufacturing technology and system has enabled automation in product processing and quality control. Advances in manufacturing technology included processing methods and precision machines, whereas advances in manufacturing system included flexible manufacturing system and quality control system such as automation and robotics, inspection method, automatic gauging and sensing, on-line data acquisition, and automated monitoring and diagnosis system. However, the modernization still yields variation that has affected quality, production capacity and delivery time.

Wear and tear, vibration, machine breakdown, inconsistent material and lack of human operators are the common sources of variation in manufacturing processes.

Traditionally, statistical process control (SPC) was used only for monitoring and identifying process variation. Advances in SPC charting have moved from merely statistical and economic control to diagnosis purposes through control chart pattern identification. The development in soft computing technology such as artificial intelligence (AI) has encouraged investigation on the application of expert systems, artificial neural network (ANN) and fuzzy sets theory for automated recognition of control chart patterns (CCPs). Application of ANN-based models, among others, has realized the computerized decision making in SPC towards replacing human interpretation. The modernization of the SPC schemes is ultimately aims to diagnose the source of variation with minimum human intervention.

Since late 1980s, control chart pattern recognition (CCPR) has become an active area of research. A useful review on the application of ANN for CCPR was provided by Zorriassatine and Tannock (1998). Since then much progress has been made in which the performance of ANN-based CCPR schemes have been enhanced through feature-based and wavelet-denoise input representation techniques, modular and integrated recognizer designs, and multivariate process monitoring and diagnosis. However, there is a lack of updated critical review on such issues. Therefore, this paper aims to update on the research issues and provide research direction in the development of ANN-based CCPR schemes.

This paper begins with an overview of ANN for pattern recognition. It is followed by the advances in ANN-based CCPR schemes. Then, details discussion is focused on issues in development of ANN-based CCPR schemes with respect to input data and process patterns, input representation, recognizer design and recognizer training. In conclusion, the important issues and direction of research in CCPR are highlighted.

2. Artificial Neural Network in Pattern Recognition

ANN is a massively parallel-distributed processor that has the ability to learn, recall and generalize knowledge (Haykin, 1999). It is recognized as an important and emerging methodology in the area of classification.

ANN is flexible, adaptive and can better handle noise and changes in the patterns. The advantage with an ANN-based pattern recognizer is that it does not require the provision of explicit rules or templates. Rather, it learns to recognize patterns from examples during the training phase. It has the ability to classify an arbitrary pattern not previously encountered. ANN offers useful properties and capabilities such as non-linearity, input and output mapping, adaptability and fault tolerance, among others. These attributes are needed for recognizing and classifying data which are often contaminated with noise, unknown distribution and incomplete as found in CCPs (Schalkoff, 1997; Haykin, 1999).

ANN acquires knowledge through a learning process and inter-neuron connection strengths (synapse weights) are used to store the knowledge. A learning algorithm is used to modify the synapse weights so as to achieve the target. ANN can tailor itself to the training data. A well-trained ANN is able to generalize knowledge. It will produce a reasonable output for input that has never been encountered during training/learning. Although ANN training requires considerable computation, the recall process is very fast. ANN is also suitable for implementation using very-large-scale-integrated (VLSI) technology such as in the form of chip that can replace the need for continuously monitoring by personal computer (Zurada, 1992; Patterson, 1996; Scalkoff, 1997; Haykin, 1999).

Specific focus on pattern recognition aspects using ANN can be found in Pao (1989), Ripley (1994), Bishop (1995), and Padya and Macy (1995). ANN has been widely implemented in pattern recognition application such as for hand-written characters (Zeki and Zakaria, 2000), printed characters (Amin, 2000), grain grading (Utku, 2000), bio-signals (Christodoulou and Pattichis, 1999) and speech signals (Loh *et al.*, 2000). ANN is also being researched for application in CCPR.

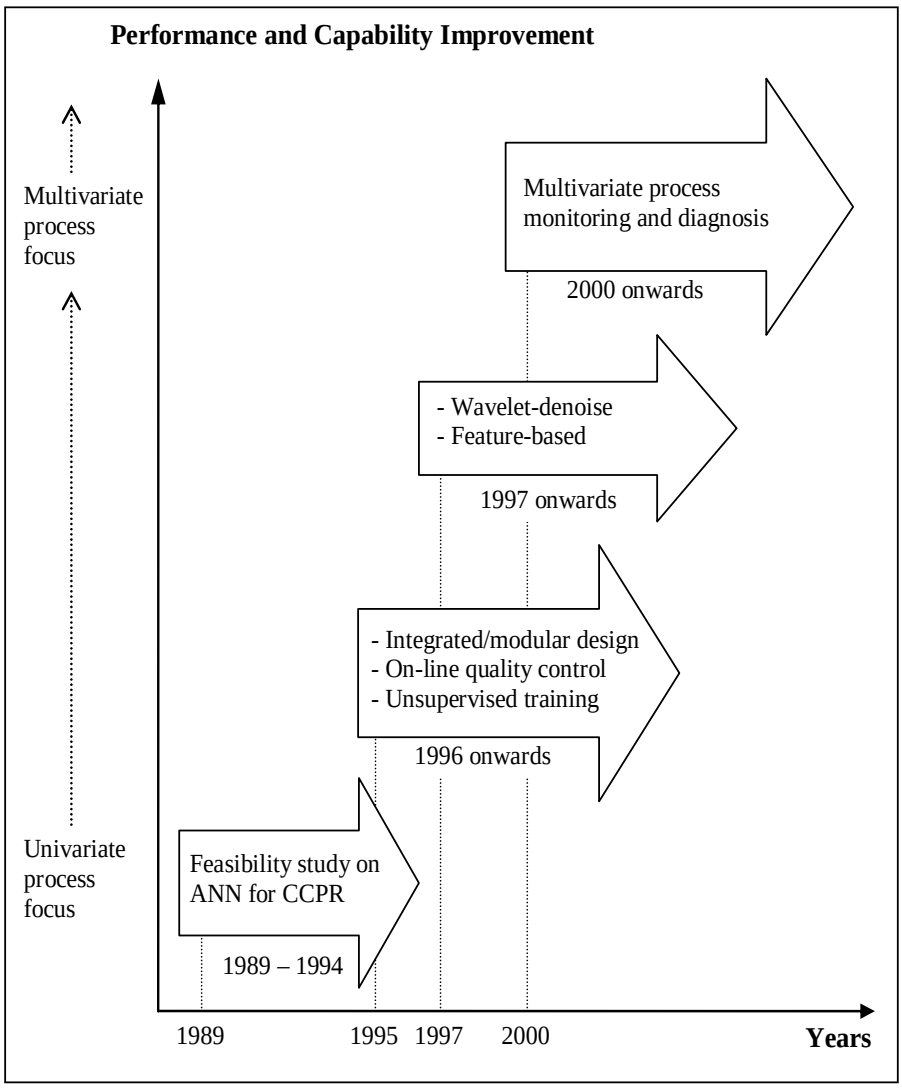
3. Advances in ANN-Based Control Chart Pattern Recognition Schemes

Advances in ANN-based CCPR schemes are shown in Figure 1. Around late 1980s and early 1990s, the application of ANN began to replace the rule based expert system in recognition and interpretation of univariate CCPs. Encouraging results from a comparative study between ANN and conventional control charts as reported by Pugh (1989) attracted further investigation on ANN application to SPC.

The earliest reported works have focused on feasibility study of ANN for implementation in CCPR schemes (see Hwarng and Hubele, 1991; 1993; Velasco, 1993; Pham and Oztemel, 1993; 1994). Then, Hwarng (1995a; 1995b; 1997), Hwarng and Cheng (1995; 1997), Al-Ghanim (1997), Tontini (1996; 1998), Anagun (1998), Guh *et al.* (1999a; 1999b), Guh and Hsieh (1999), Guh and Tannock (1999), and Pham and Chan (1998; 1999; 2001) addressed other issues beyond the levels of feasibility study. The reported works included unsupervised training ANN recognizer, modular and integrated recognizer designs and on-line quality control.

Since late 1990s until recent years, feature-based and wavelet-denoise input representation techniques were investigated for improving the recognition performance of ANN. The most significant works included wavelet-ANN (Al-assaf, 2004; Assaleh and Al-assaf, 2005; Cheng *et al.*, 2007; Wang *et al.*, 2007), shape features-ANN (Pham and Wani, 1997; Gauri and Chakraborty, 2006; 2008) and statistical features-ANN (Hassan *et al.*, 2003).

Figure 1: Advances in ANN-Based CCPR schemes

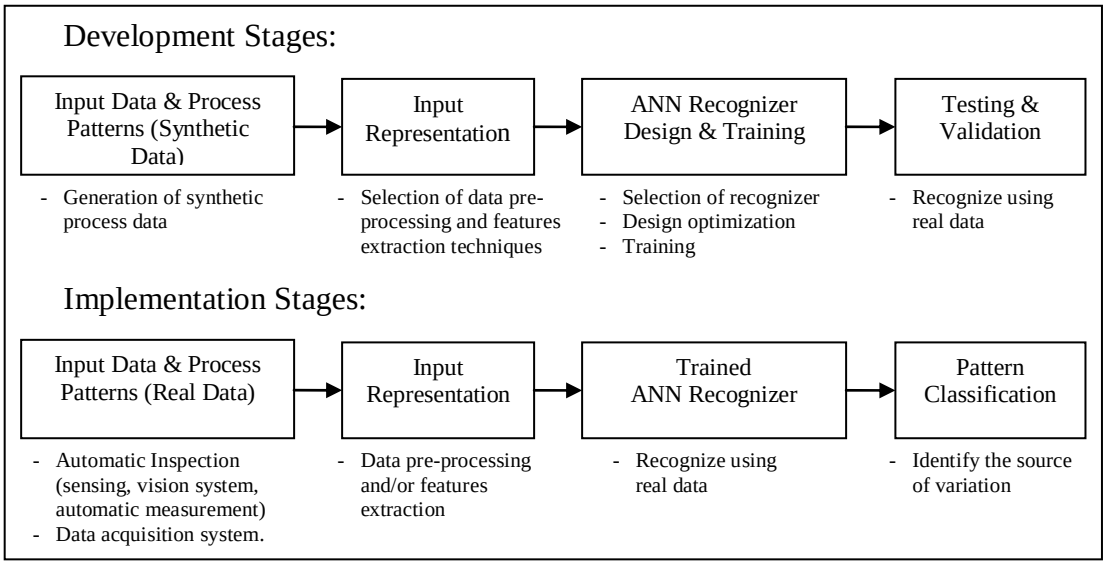


The other recent works focused on monitoring and diagnosis of multivariate/bivariate process mean shifts and variance shifts (see Zorriassatine *et al.*, 2003; Cheng and Wang, 2004; Niaki and Abbasi, 2005; Guh, 2007; Cheng and Cheng, 2008; Yu and Xi, 2009). Guh (2007) proposed a modular-ANN design, whereas Yu and Xi (2009) proposed an ensemble-ANN design and applied feature-based input representation technique.

4. Issues in Development of ANN-Based CCPR Schemes

Development stages of ANN-based CCPR schemes comprise input data and process patterns, input representation, recognizer design and training, and testing and validation. Research issues addressed in the literatures are dispersed throughout the development stages as shown in Figure 2.

Figure 2: Development and Implementation Stages of ANN-Based CCPR schemes



4.1. Input Data and Process Patterns

Observation samples (raw data) of process patterns should represent real manufacturing process situations such as independent or dependent processes, on-line (in-process) quality control, and sufficient or limited amount of data. Sufficient data may represent the situation in long-run production while limited data may represent short-run production or small lot sizes.

Ideally, samples should be tapped from real process environment. However, since a large amount of data are required for training ANN recognizer, synthetic (artificial) samples commonly generated using Monte-Carlo simulation approach, which has been widely adopted in most researches. Cheng (1997) noted that for the case where the patterns cannot be expressed mathematically, data must be collected from real process.

4.1.1. Patterns for Univariate (Independent) Processes

The univariate process patterns are well-defined by *Shewhart X-bar* CCPs as shown in Figure 3. A stable process can be indicated by normal pattern, whereas unstable process can be indicated by abnormal patterns (i.e., upward and downward shifts, upward and downward trends, cyclic, systematic, stratification, and mixture).

Concurrent pattern is another type of univariate process pattern. Figure 4 shows an example of concurrent pattern, which comprises a mixer of trend and cyclic patterns. Among the earliest works on

the ANN application for concurrent pattern recognition can be found in Guh and Tannock (1999) and Guh and Hsieh (1999). Guh and Hsieh (1999) reported excellent recognition performance for trend patterns and low performance for cyclic patterns. Later, Chen *et al.* (2007) reported better results through a wavelet-ANN scheme. It should be noted that the recognition of concurrent patterns was not addressed in multivariate process cases.

In practice, shift patterns indicate there are changes in material, operator or machine. Trend patterns indicate tool wear. Cyclic patterns indicate voltage fluctuation in power supply (Chen *et al.*, 2007). Nelson (1985) noted that stratification pattern represent the stratification of two subgroups data with different averages, whereas mixture pattern occur when two different populations of data are mixed from either one (not both populations) and made up the data average. One may confuse between 'mixture pattern' and 'mixed abnormal patterns'. Guh and Hsieh (1999) referred concurrent patterns as 'mixed abnormal patterns'.

Figure 3: Common CCPs for Univariate Processes

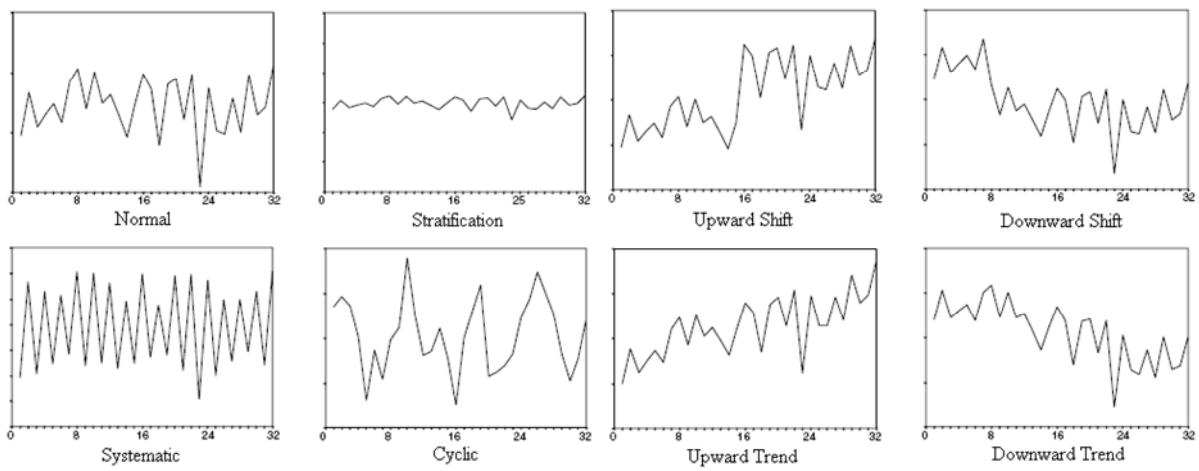
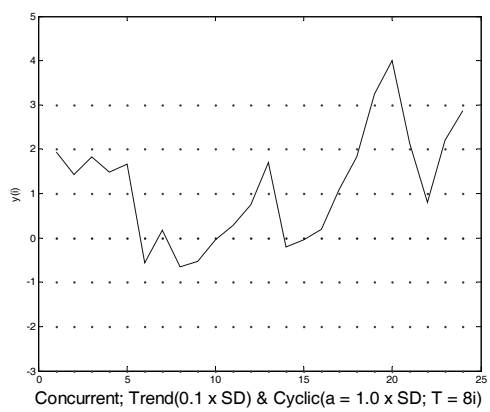


Figure 4: Concurrent Pattern - a Mixer of Trend and Cyclic Patterns



Selection of CCPs parameters is important for training and testing the ANN recognizers. Among the important parameters included window size, random noise, mean shift (for shift patterns), trend slope (for trend patterns), cycle amplitude and cycle period (for cyclic pattern), and systematic departure (for systematic pattern). Table 1 attempts to summarize the common values of CCPs parameters as used in several researches. Based on the magnitude of mean shifts, most of recent works obviously concern on moderate and large shifts (1.0 to 3.0 standard deviations). Less attention is given to smaller shift (less than 0.7 standard deviations).

Table 1: Common Values of Parameters of Abnormal Control Chart Patterns

		Abnormal control chart pattern (Range of parameter in standard deviation)				
Researches	**WS**	**US* / DS**** (*s*)	**UT*/ DT**** (*g*)	**CYC** (*a*) / (*T*)	**SYS** (*d*)	**STRA** (*σ'*)
Guh (2005), Guh and Shieu (2005)	24	(1.0, 3.0)* / (-3.0, -1.0)**	(0.12, 0.28)* / (-0.12, -0.28)**	(1.0, 3.0) / -8	(1.0, 3.0)	
Hassan and Nabi Baksh (2005), Hassan *et. al.* (2006)	20	(0.7, 2.5)* / (-2.5, -0.7)**	(0.015, 0.025)* / (-0.025,-0.015)**	(0.5, 2.5) / (10)		
Gauri and Chakraborty (2006; 2007)	32	(1.5, 2.5)* / (-2.5, -1.5)**	(0.05, 0.1)* / (-0.1, -0.05)**	(1.5, 2.5) / (8,16)	(1.0, 3.0)	(0.2, 0.4)
Chen *et. al.* (2007)	24	(1.0, 3.0)* / (-3.0, -1.0)**	(0.1, 0.26)* / (-0.26, -0.1)**	(1.0, 3.0) / (4, 8)	(1.0, 3.0)	

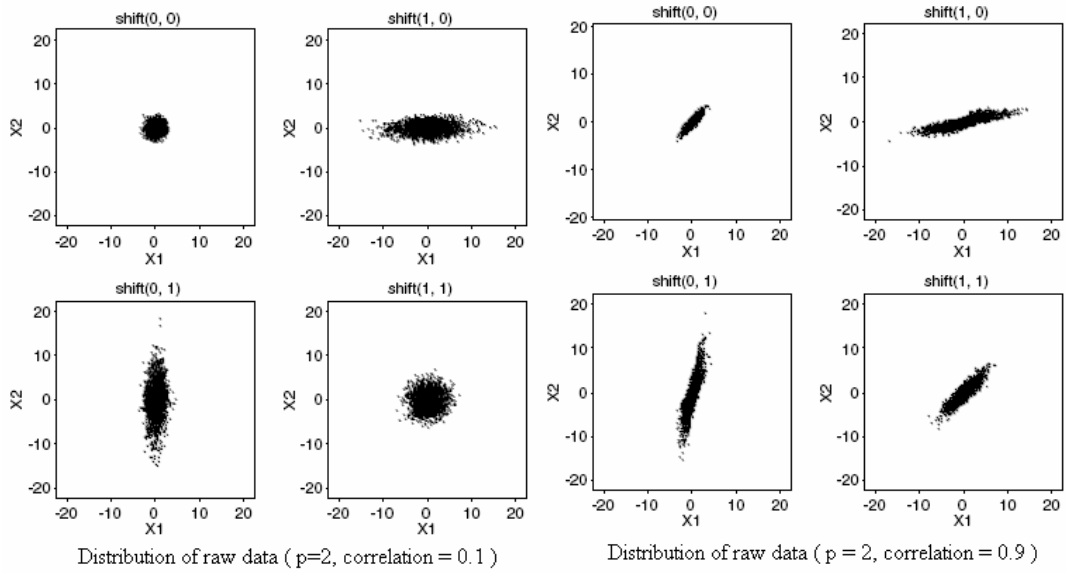
where:

σ' - Random noise for stratification pattern
a - Amplitude of cyclic pattern
g - Magnitude of gradient for trend pattern
d - Magnitude of systematic departure
s - Magnitude of mean shift
T- Period of a cycle for cyclic pattern

4.1.2. Patterns for Multivariate (Dependent) Processes

In joint monitoring of multivariate processes, process patterns should be able to indicate the concurrent effects from process variation and data correlation. The unique structures of *Shewhart X-bar* CCPs can provide useful meaning about variation in mean for independent process variables. Unfortunately, they are unable to indicate the linear correlation between two dependent process variables. On the other hand, T^2 statistic patterns can show the variation in multivariate shifts and data correlation but unable to diagnose the source variables that responsible for the shift. Lack of standard reference patterns for multivariate processes seems can be a reason for limited progress in multivariate CCPR schemes. Only Cheng and Cheng (2008) provided the distribution of raw data for bivariate processes based on scatter diagram as shown in Figure 5. It was limited to variation in mean shifts and variance shifts. Motivation to investigate reference multivariate patterns and extracted features based on scatter diagram could be useful for further improvement in multivariate CCPR schemes performances.

Figure 5: Distribution of Raw Data for Bivariate Processes

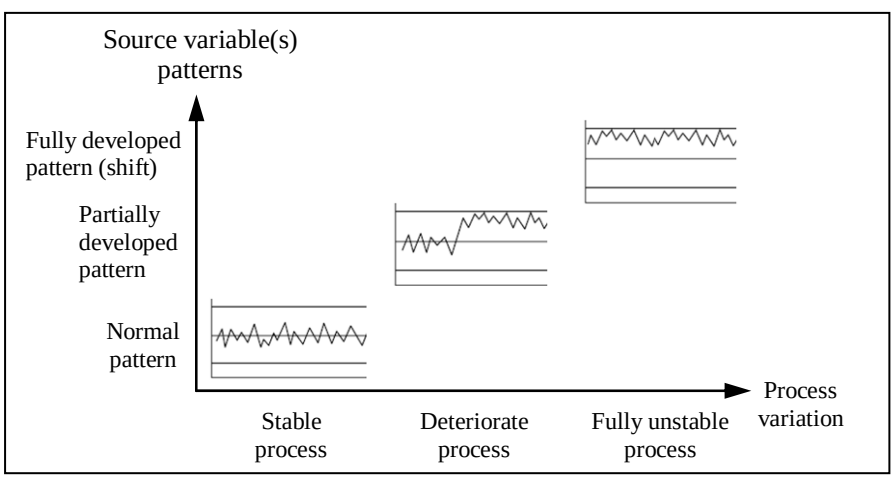


(a) Variation in Mean Shifts (b) Variation in Variance Shifts

4.1.3. Dynamic Patterns for On-Line Quality Control

On-line quality control refers to process control during actual production (Taguchi *et al.*, 1989; Kapur, 1993; Hassan *et al.*, 2000). On-line monitoring involves identification of process status, that is, either in statistically stable state or in statistically unstable state. In practice, an established process variable or quality characteristic is monitored from a stable condition, which is represented by normal pattern on *Shewhart X-bar* control chart. Disturbance from assignable cause may deteriorate the process variable suddenly or gradually into an unstable condition. Process deterioration initially results in partially developed patterns. Then, it is slowly developing into fully developed pattern. The stages towards an unstable condition can be illustrated in Figure 6.

Figure 6: Changes in a Process Variable



The starting point when a stable process actually starts to deteriorate and reaches recognizable patterns is commonly unpredictable. Modeling of training patterns for such a situation is an important issue in realizing a truly automated and an intelligent CCPR scheme for on-line quality control implementation. Therefore, partially developed patterns and dynamic patterns approaches have been addressed for training and testing the ANN recognizers respectively. Guh and Shieu (2005) and Guh (2007), among others, have used different percentage of partially developed patterns for different patterns magnitudes in training the ANN. For example, based on recognition window size of 12, Guh (2007) has set the starting shift at point 11 to obtain a fast detection for large mean shifts (±2.50, ±2.75, ±3.00 standard deviations). For moderate shifts (±1.75, ±2.00, ±2.25 standard deviations), the starting shift were set at point 9. Then, the starting shift were set at the middle of the recognition window for smaller shifts (±1.00, ±1.25, ±1.50 standard deviations). On the other hand, Hassan and Nabi Baksh (2008) trained the ANN recognizer using 75 to 90 per cent partially developed patterns, where the patterns magnitudes were generated randomly. Guh *et al.* (1999a; 1999b) and Guh and Shieu (2005) noted that training and testing ANN recognizers using fully developed patterns was ineffective for on-line application.

4.2. Input Representation

Input representation is an approach to represent input signal of patterns into an ANN recognizer. Numerous techniques have been proposed to represent input signals, which is aims to have effective training and improved recognition performance. In this paper, they are categorized into raw data-based, feature-based and wavelet-denoise input representation techniques.

4.2.1. Raw Data-Based

In most researches, the observation samples (raw-data) are transformed using common pre-processing techniques, namely standardization and normalization. In several works, the standardized samples are further pre-processed using zoning and binary encoding procedures (see Hwarng and Hubele, 1993; Hwarng and Chong, 1995).

Standardization is a procedure to linearly transform the samples (X_t) into standard normal variates (Z_t). It can be performed using the following equation (Nelson, 1989):

$$Z_t = (X_t - \mu) / \sigma \tag{1}$$

where μ and σ are the mean and standard deviation for a statistically stable process. At stable process condition, Z_t satisfies normal distribution (identically and independently distributed) within a range between [–3, 3] with zero mean and unity standard deviation.

Demuth and Beale (1998) noted that training ANN recognizers will be more efficient when input representation fall within a certain range. Normalizing the standardized samples (Z_t) into a compact range, normally between [0, 1] or [–1, 1] could minimize the effect from random noise. Standardization and normalization procedures are also called as re-scaling and they often useful particularly when the values of samples differ significantly (Bishop, 1995). Normalization procedure has also been used to re-scale the extracted features between [–1, 1] for representing to ANN recognizers (Hassan *et al.*, 2003; Gauri and Chakraborthy, 2006; 2008).

Normalization into a range between [0, 1] can be obtained using the following equation (Barghash and Santarisi, 2004):

$$Y_t = (Z_t - Z_{min}) / (Z_{max} - Z_{min}) \tag{2}$$

Normalization into a range between [–1, 1] can be obtained using the following equation (Gauri and Chakraborty, 2008):

$$Y_t = [2 \text{ x } (Z_t - Z_{min}) / (Z_{max} - Z_{min})] - 1 \tag{3}$$

where,

Y_t – normalization from standardized data or extracted features,

Z_t – standardized data or extracted features,

Z_{max} – maximum value of standardized data or extracted features,

Z_{min} – minimum value of standardized data or extracted features.

Hwarng and Hubele (1993) and Hwarng and Chong (1995) used binary encoding as input representation. Coding is a procedure to encode the standardized samples (Z_t) into coded forms, whereas zoning is a procedure to re-scale and divide the standardized *Shewhart* control chart into seven zones (i.e., $zone_{+3}$, $zone_{+2}$, $zone_{+1}$, $zone_0$, $zone_{-1}$, $zone_{-2}$ and $zone_{-3}$). For example, if a sample is plotted in $zone_{+2}$, a binary coding is represented as '0100000'. One sample requires seven input neurons. Therefore, it requires relatively large network size and increases computational effort.

On the other hand, Guh and Shiue (2005) and Guh (2007) reported another pre-processing technique. The samples were linearly transformed into a range between [–7.625, 7.625] which is differ from the practically standardization range, that is, [–3, 3]. Then, the 'transformed samples' was divided into 61 zones with an interval width of 0.25 standard deviations. They noted that a large range of transformed samples and zoning could allow for identifying large process variation, that is, up to 4.0 standard deviations.

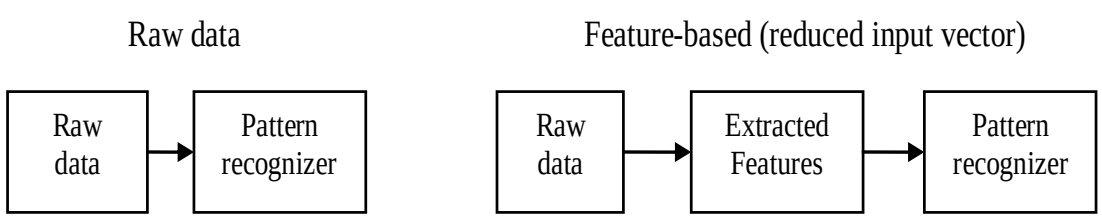
Raw data pre-processing techniques discussed above are well established. However, in order to ensure the practicality of the studies, it is important for researchers to ensure that the observation samples are standardized between [–3, 3] standard deviations.

4.2.2. Feature-Based

Raw data-based input representation yields large dimensional input vectors, computational efforts and time consuming for training ANN recognizer (Pham and Wani, 1997). In addressing this issue, feature-based input representation such as summary statistic features, frequency count features, shape features and statistical features have been proposed in developing univariate CCPR schemes. It involves features extraction procedure as shown in Figure 7 to extract the properties of the samples.

The presence of excessive or unnecessary features may burden the training process, while too few features may insufficient to represent effective pattern properties. Therefore, proper selection of features is important towards achieving a satisfied recognition performance. Features selection is an approach to determine the significant features set from features extraction process. For example, Hassan *et al.* (2006) reported features selection based on a resolution IV fractional factorial design of experiment (DOE). This method was utilized in identifying minimum set of statistical features that have substantial effect to the recognition performance of ANN. Generally, there has been limited work addressing features selection in CCPR.

Figure 7: Comparison between Raw Data-Based and Feature-Based Input Representation



4.2.2.1. Summary Statistic Features

Summary statistic features consist of 60 individual samples, mean and standard deviation of 15 statistical windows, 10 lags of autocorrelation, results of the computational *CUSUM* chart, and chi-square statistics (Tontini, 1996). However, the combination of all input would result in a large network size, increase computation time, and does not meet the aim of dimensionality reduction. The used of summary statistic features has not been reported in other research.

4.2.2.2. Frequency Count Features

Anagun (1998) used a set of frequency counts (histogram) as a simple and compact features representation. However, the robustness of the input representation seems to be rather limited since it loses the information on the order of the data. This ordering information can be a significant attribute to differentiate between upward trend and downward trend patterns, or between upward shift and downward shift patterns. He reported that ANN recognizer trained using frequency count input representation gave better classification accuracy compared to raw data.

4.2.2.3. Shape Features

Nine shape features was firstly developed by Pham and Wani (1997). It consisted of slope(s) for the least square line representing the pattern, number of mean crossings, number of least-square line crossings, cyclic membership, average slope of the line segments, slope difference, area between the pattern and the mean line, and area between the least-square line and the line segments. Later, Gauri and Chakraborty (2006; 2008) proposed two improved shape features. They noted that their latest shape features had low correlation among themselves and improved the recognition stability. Gauri and Chakraborty (2008) extracted shape features from non-standardized samples to overcome recognition problem for stratification patterns and this approach was applied for other CCPs. Then, the shape features were normalized between [−1, 1] for representing into ANN recognizer.

Based on centerline and control limit references, stratification patterns can be discriminated from normal patterns by the location of the samples. Stratification patterns may be viewed as normal patterns with unexpectedly lower variability. The distinction between normal patterns and stratification patterns is lost if the samples are standardized. Extracting shape features from non-standardized samples could alleviate this problem.

4.2.2.4. Statistical Features

Statistical features consist of mean, standard deviation, skewness, mean-square value, autocorrelation, and *CUSUM* (Hassan *et al.*, 2003; 2006). As shape features, the statistical features were also normalized between [−1, 1] for representing into ANN. They reported that the statistical features provided better recognition accuracy for ANN than raw data. Their study was focused on six CCPs, i.e., random, upward and downward shifts, upward and downward trends, and cyclic patterns.

Skewness and kurtosis features have also been used by Guh (2002) in representing non-normal distribution samples for abnormal CCPs. He reported that the non-normal distributed CCPs could be well-recognized using ANN.

Gauri and Chakraborty (2006) gave a comparison between shape features and statistical features. They noted that the shape features required only small amount of training examples without lost information on the order of the data. Inversely, statistical features require large amount of training examples and it will lost information on the order of the data.

4.2.3. Wavelet-Denoise

Multi-resolution wavelet analysis (MRWA) is another technique to reduce noise effects by denoising or filtering the samples in several decomposition levels. It is very appealing in many application areas such as for detecting discontinuity and/or abrupt change in signal processing and image processing (Wang *et al.*, 2007). In this paper, it is referred to wavelet-denoise.

Investigation on wavelet-denoise input representation in ANN-based CCPR schemes can be found in Al-Assaf (2004), Assaleh and Al-Assaf (2005), Chen *et al.* (2007) and Wang *et al.* (2007), among others. Al-Assaf (2004) used MRWA to denoise 32 samples of univariate CCPs (normal, upward shift, upward trend and cyclic patterns). The denoise samples were used as input signal to an ANN recognizer. Then, Assaleh and Al-Assaf (2005) modified the MRWA algorithm to ‘multi-resolution discrete cosine transform’ (MRDCT) algorithm for improving ANN performance in recognizing small magnitudes of trend and shift patterns. They referred the MRWA as a features

extraction technique. In this paper, however, wavelet is viewed different from features extraction because it did not change the sequence of the samples and the network size.

The beneficial performance of MRWA has also been reported in recognition of concurrent patterns (Chen *et al.*, 2007). It was used to decompose the concurrent patterns into different patterns (trend and cyclic patterns). It has reduced pattern examples and time for training ANN.

Wang *et al.* (2007) noted that the MRWA has played a crucial role for process control or monitoring. Insufficient denoising will distort waveforms and introduce errors. Beside, excessive denoising will over-smooth the sharp features of underlying signals by recognizing them as noise or outliers.

4.2.4. Input Representation for Multivariate Process

In recognition of multivariate (bivariate) shift patterns, raw data and features have been used as input representation to ANN. In developing the integrated multivariate SPC-ANN schemes, Cheng and Wang (2004) used individual samples and means from T^2 out-of-control signals ($X_{11,...}$, X_{15}, μ_1, $X_{21,...}$, X_{25}, μ_2), Niaki and Abbasi (2005) used means from T^2 out-of-control signals (μ_1, μ_2), whereas Cheng and Cheng (2007) used individual samples and variances from T^2-*variance* out-of-control signal ($X_{11,...}$, X_{15}, σ^2_1, $X_{21,...}$, X_{25}, σ^2_2) as input representation. On the other hand, in developing the ANN-based schemes, Zorriassatine *et al.* (2003) utilized samples of the source variables (X_{1i}, X_{2i}), Guh (2007) utilized samples and T^2 statistics (X_{1i}, X_{2i}, T^2_i), whereas Yu and Xi (2009) utilized samples and statistical features of the source variables for strengthening the pattern properties.

Since there have been limited works in multivariate process pattern recognition, input representation of multivariate process still open for further investigation. Better input representation technique is required towards achieving efficient recognition in dealing with small shift patterns and on-line quality control.

4.3. Recognizer Designs

The ANN recognizer is used to perform pattern recognition or classification. In order to achieve a satisfied recognition performance, the model of the recognizer has to be properly selected, designed and trained. Selection of the ANN models is depends on the nature of problem.

4.3.1. Isolated ANN Models

There are a few isolated ANN models have been investigated in developing univariate CCPR schemes, namely, multi-layer perceptrons (MLP), learning vector quantization (LVQ), radial basis function (RBF), adaptive resonance theory (ART) and Kohonen self-organizing mapping (SOM). The model architectures are highly related to the training paradigms to be discussed in Section 4.4. Briefly, MLP and LVQ are from supervised training, whereas ART and Kohonen SOM are from unsupervised training. Table 2 provides their comparison.

Table 2: Isolated ANN Models Used as Pattern Recognizers

Model	Description	Advantage	Limitation
MLP	• A feed forward and fully connected network. • Consists of neurons in an input layer, one or more hidden layer and an output layer.	• Proven effective for classification. • Fast during recalling.	• Time consuming for training. •
LVQ	• Supervised form of vector quantization (VQ). • Classification is based on clustering input samples around a predetermined number of reference vectors.	• Relatively faster in training compared to MLP. •	• Lack of stability in learning. • Some neurons tend to win too often while others are always inactive.
ART	• Comprising two layers: feature representation field and category representation field. • Enhanced version of competitive learning, LVQ. • Types: ART1, ART2.	• Suitable for continuous, incremental on-line learning.	• Complex architecture. • Considerable time is needed for processing input data. • Selection of values for its large numbers of parameters is difficult and not easily determined.
Kohonen SOM	• Consists of a two-dimensional array of neurons. • Similarity among the patterns is mapped into the closeness relationship on the competitive layer (clustering).	• Easy to implement.	• Training process is time consuming. • Manual labeling to represent different data classes at the end of training. • Suffers from the stability-and-plasticity dilemma.

In this paper, more attention is given to MLP model since it has been widely used and proven effective for classification tasks (see Pham and Oztemel, 1993; Hwarng and Hubele, 1993; Cheng, 1995; 1997; Guh *et al.*, 1999a; 1999b; Guh and Tannock, 1999; Guh and Hsieh, 1999; Perry *et al.*, 2001; Hassan *et al.*, 2003; Al-Assaf, 2004; Al-Assaf and Assaleh, 2005; Gauri and Chakraborty, 2006; 2008; Chen *et al.*, 2007). Other researchers have studied LVQ model (see Pham and Oztemel, 1993; 1994; Yang and Yang, 2002), ART model (see Hwarng and Chong, 1995; Al-Ghanim, 1997; Pham and Chan, 1999; 2001) and Kohonen SOM model (see Pham and Chan, 1998).

MLP structure basically comprises an input layer, one or more hidden layer(s) and an output layer. The number of layers and nodes/neurons in each layer could affect the network performance. Thus, it should be properly selected during the design stage. Cheng (1997) noted that as a general guideline, the network size should be as small as possible to allow for efficient computation.

There have been some disagreements among researchers in choosing the right number of hidden layers. Most of researchers have used one hidden layers. On the other hand, Guh *et al.* (1999a; 1999b), Guh and Tannock (1999), Perry *et al.* (2001) have used more than one hidden layers. Billing *et al.* (1991) and Bishop (1995) suggested that one hidden layer is sufficient to approximate arbitrarily well any continuous mapping from one finite-dimensional space to another, provided the number of hidden nodes is sufficiently large. Bishop (1995) further argued that one might wonder if there is anything to be gained by using more hidden layers. This is supported in Cheng (1997) who reported that his network with two hidden layers did not converge during training.

In many cases, network configuration and parameters have been selected empirically (see Cheng, 1997; Guh *et al.*, 1999b; Dedeakayogullari and Burnak, 1999; Gauri and Chakraborty, 2006).

However, there is a trend that researchers begin to optimize the ANN designs using systematic approach such as DOE. Hwarng (1992) implemented full factorial DOE to study the effect of the number of hidden nodes, the number of scale zones and their joint effect on the performance of ANN. Although his application of DOE was rather limited, he has provided a good beginning. Then, Khaw *et al.* (1995) and Packianather *et al.* (2000) investigated the feasibility of using Taguchi Method (TM) for selecting design parameters of ANN. TM leads to a systematic approach in designing process. They reported that TM offered considerable benefits in development time and accuracy compared to the traditional trial and error approach. Despite benefits offered by TM, Montgomery (1996) discouraged its usage which lacked provision to deal adequately with the potential interaction between the factors/parameters. Later, Barghash and Santarisi (2004) applied the resolution IV fractional factorial experiments in analyzing the effects of training parameters. In focusing on shift patterns, they reported that minimum shift magnitudes, range of shift magnitudes, shift percentage of partially developed patterns and number of pattern examples have significant effects on the ANN performance. Beside, recognition window size (number of input nodes) and network size (number of hidden nodes) have no major effect. This method has also been applied in features selection as described in Section 4.2.2.

4.3.2. Modular and Integrated ANN-Based Models

The performances of isolated ANN models are rather limited to a simple classification tasks. This disadvantage is being addressed through modular and integrated ANN-based models. The motivation has been to improve recognition performance through better design structures. The respective shortcomings of the individual recognizers can be avoided when multiple recognizers are applied (Haykin, 1999).

Pham and Oztemel (1993) were among the earliest to report the benefit of using synergistic-ANN through a composite pattern recognition scheme that combined MLP and LVQ models. They reported that the structure based on the combined ANN has better classification capabilities than individual ANN. Then, more elaborate recognizer designs have been investigated for univariate CCPR schemes (see Pham and Oztemel, 1995; Hwarng, 1997; Cheng, 1997; Tontini, 1996; 1998; Pham and Chan, 1999; Wani and Pham, 1999; Guh *et al.*, 1999b). Different terminologies such as combined, composite, integrated, specialized, modular, synergy, multiple, ensemble, multi-stage and hybrid recognizers were used to described the proposed design structures. Despite the different terms used, many of them have some similarities. Furthermore, they aimed to improve the overall performance of the CCPR schemes. Table 3 attempts to group and describe the proposed recognizer designs.

Table 3: Modular and Integrated ANN-Based Models Used as Pattern Recognizers

General design structure	Description
Isolated/generalized	• One generic recognizer for classifying all pattern types.
Specialized	• One specialized recognizer for classifying only one class of pattern.
Modular, Multiple	• It seeks to divide complex recognition problem into smaller modules. • Each of these tackled by a specialized recognizer. • Final decision is arrived by combining the individual outputs. • A modular network is attractive over a single neural network in terms of learning speed and input representation.
Hybrid, Integrated, Synergy, Combination, Composite, Ensemble	• Use more than one recognizer to solve complex problems. • The objective is to combine the strengths offered by different recognition techniques/paradigms and to avoid their respective shortcoming.
Multi-stages, Cascade	• Use more than one recognizer and functioning in a sequential manner. • The result of the later stage is dependant on the earlier stage.

Pham and Oztemel (1995) integrated an expert system with ANN to exploit the complementary features of ANN and expert system. The proposed recognizer has enhanced recognition performance and provided a good user interface.

Tontini (1996; 1998) proposed hybrid-ANN recognizer that combined RBF and Fuzzy-ARTMAP. They argued that the inclusion of RBF model has improved recognition performance and tremendously reduced the problem of sensibility to the presentation order of the training patterns.

On the unsupervised learning, Pham and Chan (1999) proposed a synergistic self-organizing recognizer that combined Kohonen SOM and ART2 models. Kohonen SOM is responsible as the pre-processor and the ART2 as the end classifier. The proposed recognizer provided good performance.

Hwarng (1997) proposed specialized-ANN recognizers to identify the behavior of cyclic patterns. It showed significant improvement over the corresponding generalized-ANN. A specialized-ANN requires smaller network size and simple training compared to a generalized-ANN.

Cheng (1997) proposed modular-ANN recognizer. It has better performance compared to a generalized-ANN when trained using back-propagation algorithm. The other modular-ANN designs were then proposed for addressing different issues (see Guh *et al.*, 1999b; Guh and Hsieh, 1999; Guh, 2005). Guh *et al.* (1999b) found that generalized-ANN performed poor classification for on-line quality control application. At early study, there was a challenge in discrimination of normal patterns and small magnitude of abnormal CCPs (shift and trend patterns). To overcome this issue, they proposed modular-ANN that divided a complex recognition task into several sub-tasks. A series of specialized-ANN dedicated to each sub-task was implemented. Discrimination algorithm was proposed to control and evaluate the outputs of modular-ANN. In another study, Guh and Hsieh (1999) proposed two-stage modular-ANN for identifying abnormal process and for predicting abnormal patterns magnitudes. The first stage module consisted of a generalized-ANN for identifying whether a process is stable or belonged to any of unstable patterns. The second stage module consisted of three specialized-ANNs for estimating specific parameters of abnormal CCPs (upward shift, upward trend and cyclic patterns). In related study, Guh (2005) proposed an enhanced two-stage modular-ANN. The first stage module integrated a generalized-ANN and decision tree learning to enhance the capability for identifying unstable patterns. Then, the second stage module consisted of seven specialized-ANNs for estimating specific parameters of various abnormal CCPs (upward and downward shifts, upward and downward trends, cyclic, systematic and mixture patterns).

Above literatures shows that various design strategies have been investigated to improve the recognition performance and to enhance the capability of the recognizers for solving complex recognition problems. Padya and Macy (1995) noted that the accuracy of a recognition system depends upon the characteristics of individual recognizer and also upon the manner in which they are combined.

4.3.3. Recognizer Designs for Monitoring and Diagnosing Multivariate Process Shifts

The CCPR scheme aims to monitor and diagnose process variation automatically. Monitoring refers to the identification of process status, i.e., either in statistically stable state or in statistically unstable states, whereas diagnosis refers to the identification of the source of variation. Recognizer designs for monitoring and diagnosing univariate process are straight forward compared to multivariate process. In univariate CCPR schemes, the ANN-based recognizers were designed to perform monitoring and diagnosis simultaneously. In multivariate schemes, the ANN-based recognizers were designed either to perform diagnosis only or to perform monitoring and diagnosis simultaneously.

In developing the integrated multivariate SPC-ANN schemes, multivariate SPC charts such as T^2 and T^2-*variance* control charts were applied for monitoring, whereas a generalized-ANN were utilized only for diagnosing out-of-control signal from the MSPC chart (see Chen and Wang, 2004; Niaki and Abbasi, 2005; Cheng and Cheng, 2007). Chen and Wang (2004) utilized MLP model with $6p$ x $6p$ x p network architecture, where p is the number of variables being monitored. They reported that for bivariate cases (when p = 2), ANN with 12 x 12 x 2 architecture performed poor recognition for small mean shifts. Niaki and Abbasi (2005) utilized a smaller MLP model with $2p$ x hidden nodes (HN) x ($2p$ – 1) network architecture. The ANN with 4 x HN x 3 architecture gave a better performance for bivariate cases. Then, Cheng and Cheng (2007) reported good recognition for variance shifts.

On the other hand, in developing the ANN-based schemes, ANN were utilized for monitoring and diagnosing process mean shifts simultaneously (see Zorriassatine *et al.*, 2003; Guh, 2007; Yu and Xi, 2009). Zorriassatine *et al.* (2003) proposed a generalized-ANN model, namely novelty detector (ND) for recognizing upward shift and downward shift in bivariate process cases. They reported high recognition results for moderate and large mean shifts (1.5 to 2.5 standard deviations) and low results for small mean shifts (0.5 to 1.0 standard deviations). The proposed recognizer requires large recognition window size (up to 40) and huge training examples (up to 15,000 patterns). Later, Guh (2007) proposed a modular-ANN that consisted of two sequential modules. Module 1 contents a generalized-ANN for both monitoring and diagnosing bivariate process mean shifts. Module 2 contents four specialized-ANNs for identifying the magnitudes of each shift classes. MLP model with 36 x 24 x 24 x 9 network architecture with back-propagation algorithm was utilized for all the recognizers. The proposed recognizer gave excellent recognition for moderate and large mean shifts (1.25 to 3.0 standard deviations) but gave low recognition for small mean shift (1.0 standard deviation). Then, Yu and Xi (2009) proposed an ensemble-ANN with 'discrete particle swarm optimization' (DPSOEN) algorithm. Based on raw data with statistical features input representation, they reported good recognition results in dealing with moderate and large mean shifts (1.5 to 3.0 standard deviations) and low recognition results in dealing with small mean shift (1.0 standard deviation).

Generally, the existing ANN-based recognizers indicated low recognition performance in dealing with small mean shifts in bivariate processes. The study is also limited to about shift patterns. Therefore, further investigation is strongly needed towards improving it's capability to deal with other causable patterns.

4.4. Recognizer Training

The ANN recognizer needs to be trained and tested before it can be put into application. The terms training and learning are interchangeably used in the literature. Training paradigm, training algorithm, input representation of training data and training stopping criteria are a few issues associated to the training process. However, discussion in this section is focused on training paradigm which is highly related to the recognizer designs as described in Section 4.3. Decision on training paradigm, i.e., either supervised training or unsupervised training will influence other design requirements.

4.4.1. Supervised Training

Supervised training requires pre-prepared data with each of the dataset need to be labeled with a known class. It is suitable when sufficient training examples are available and time consuming for training does not negatively affect it application.

In univariate CCPR schemes, back-propagation algorithm have been widely used for training MLP model (see Pham and Oztemel, 1993; Hwarng and Hubele, 1993; Hwarng, 1995; Cheng, 1995; 1997; Anagun, 1998; Guh and Tannock, 1999; Guh and Hsieh, 1999; Guh *et al.*, 1999a; 1999b; Perry *et al.*, 2001; Hassan *et al.*, 2003; Gauri and Chakraborty, 2006; 2008). This algorithm has also been applied for multivariate process pattern recognition schemes (see Niaki and Abbasi, 2005; Cheng and Cheng, 2007; Guh, 2007).

The learning process takes place through adjustment of the weight connections to minimize error between the actual and desired output. MLP with back-propagation training algorithm has also been widely used in other prediction and classification tasks (Haykin, 1999).

4.4.2. Unsupervised Training and Cumulative Learning

Unsupervised training is more suitable for application in which the training examples are insufficient, limited or process is too dynamic and do not warrant time-consuming for supervised learning. In this learning paradigm, input data is provided without any information on the desired output. Researches associated to unsupervised training ANN applied for univariate CCPR schemes can be referred as cumulative learning (see Hwarng and Chong, 1995), on-line incremental learning (see Tontini, 1996; 1998) and unsupervised self-organizing (see Al-Ghanim, 1997; Pham and Chan, 1999; 2001).

Hwarng and Chong (1995) and Al-Ghanim (1997) presented CCPR schemes based on ART1. Hwarng and Chong (1995) noted the limitations of basic ART included recording instability, inability to classify translated patterns and learned categories tend to outgrow have hindered its successful implementation for CCPR. Such limitations were overcome by incorporating a synthesis layer and adopting a quasi-supervised training. They claimed that their proposed scheme was capable of fast and cumulative learning. Beside, Al-Ghanim (1997) reported that his proposed unsupervised scheme was inferior compared to the supervised one.

Pham and Chan (1998; 1999; 2001) proposed CCPR schemes using unsupervised ART2 and/or Kohonen self-organizing mapping (SOM). They noted that Kohonen SOM suffered from the stability-and-plasticity dilemma in spite of its unsupervised training paradigm. Plasticity refers to the ability to keep learning new input, whereas stability refers to the ability to preserve previously learned patterns. It means that the new pattern categories can be learned without affecting or erasing one that have already been established. Pham and Chan (1998) developed a new firing rule for SOM model to recognize univariate CCPs. They claimed that the proposed scheme could inherently overcome stability-and-plasticity dilemma and achieve good performance. Pham and Chan (2001) noted that the main advantage of ART2 model is that they do not suffer from the stability-and-plasticity problem and suitable for continuous/incremental on-line learning.

Stability-and-plasticity problem has also been addressed in Tontini (1996; 1998) who proposed RBF Fuzzy-ARTMAP model. He claimed that the proposed scheme was capable for on-line incremental learning and could reduce the limitation posed by the presentation order of the training samples, which he defined as sensibility.

5. Conclusion

This paper reviews advances in the development of ANN-based CCPR schemes with respect to input data and process patterns, input representation, recognizer design and training, and multivariate process monitoring and diagnosis.

With respect to input representation, most of the early works used standardized and normalized samples (raw data) as input signals to the ANN recognizers. Raw data-based input representation commonly produce large network structures and are not very effective and efficient in dealing with complicated pattern recognition problems. In recent years, alternative input representation using features has been increasingly studied. Relatively, feature-based input representation has been reported to have better recognition performance compared to raw data.

The review also indicates that there is a need for investigation into a better approach for selecting the ANN design parameters. Beside DOE techniques, Genetic Algorithm may also be useful for optimizing the ANN design. There are trends to enhance the recognition performance using modular and integrated recognizers. The application of such designs has raised a new challenge in terms of the methodology to combine outputs from multiple recognizers.

Selection of training paradigm should be compatible with the design of the recognizer and the situation in hand. Most of the existing works have focused on supervised training and large amount of training examples. Such CCPR schemes are only applicable for long-run production. Beside, in-line with the miniaturization technology and products, nowadays many industries deals with small production lot size and flexible manufacturing. In such a situation, process data are limited which is insufficient for training the recognizers. Therefore, unsupervised training seems to be more relevant.

The review reveals that most of the existing works have focused on univariate process cases. On the other hand, there is an increasing trend to shift the focus into multivariate process cases. Lack of standard reference patterns for multivariate/bivariate processes seems to have hindered advances in multivariate process pattern recognition.

In conclusion, to serve the new generation of SPC tool, new theories, methodologies and technologies need to be investigated. Concurrently, the existing CCPR schemes need to be enhanced

and applied. The issues and research direction highlighted in this paper could be useful for new researchers as a starting point to facilitate further improvement in this area.

References

[1] Al-Assaf, Y., 2004. "Recognition of Control Chart Patterns Using Multi-Resolution Wavelets Analysis and Neural Networks", *Computers and Industrial Engineering* 47, pp. 17-29.

[2] Al-Ghanim, A.M., 1997. "An Unsupervised Learning Neural Algorithm for Identifying Process Behaviour on Control Charts and a Comparison with Supervised Learning Approaches", *Computers and Industrial Engineering* 32, pp. 627-639.

[3] Amin, A., 2000. "Recognition of Printed Arabic Text Based on Global Features and Decision Tree Learning Techniques", *Pattern Recognition* 33, pp. 1309-1323.

[4] Anagun, A.S., 1998. "A Neural Network Applied to Pattern Recognition in Statistical Process Control", *Computers and Industrial Engineering* (35:1 & 2), pp. 185-188.

[5] Assaleh, K., and Al-Assaf, Y., 2005. "Features Extraction and Analysis for Classifying Causable Patterns in Control Charts", *Computers and Industrial Engineering* 49, pp. 168-181.

[6] Barghash, M.A., and Santarisi, N.S., 2004. "Pattern Recognition of Control Charts Using Artificial Neural Networks - Analysis the Effects of the Training Parameters", *Journal of Intelligent Manufacturing* 15, pp. 635-644.

[7] Billings, S.A., Jamaluddin, H.B., and Chen, S., 1991. "A Comparison of the Back-Propagation and Recursive Prediction Error Algorithms for Training Neural Networks", *Mechanical System and Signal Processing* 5, pp. 233-255.

[8] Bishop, C., 1995. "Neural Network for Pattern Recognition", *Oxford University Press*, New York.

[9] Chen, L.H., and Wang, T.Y., 2004. "Artificial Neural Networks to Classify Mean Shifts from Multivariate χ^2 Chart Signals". *Computers and Industrial Engineering* 47, pp. 195-205.

[10] Chen, Z., Lu, S., and Lam, S., 2007. "A Hybrid System for SPC Concurrent Pattern Recognition", *Advanced Engineering Informatics* 21, pp. 303-310.

[11] Cheng, C.S., 1995. "A Multi-Layer Neural Network Model for Detecting Changes in the Process Mean", *Computers and Industrial Engineering* 28, pp. 51-61.

[12] Cheng, C.S., 1997. "A Neural Network Approach for the Analysis of Control Chart Patterns", *International Journal of Production Research* 35, pp. 667-697.

[13] Cheng, C.S., and Cheng, H.P., 2008. "Identifying the Source of Variance Shifts in the Multivariate Process Using Neural Networks and Support Vector Machines", *Expert Systems with Applications* 35, pp. 198-206.

[14] Christodoulou, C.I., and Pattichis, C.S., 1999. "Unsupervised Pattern Recognition for the Classification of EMG Signals", *IEEE Transactions on Biomedical Engineering* 46, pp. 169-178.

[15] Demuth, H., and Beale, M., 1998. "Neural network toolbox, user's guide". *The MathWorks*, Natick.

[16] Gauri, S.K., and Chakraborty, S., 2006. "Feature-Based Recognition of Control Chart Patterns", *Computers and Industrial Engineering* 51, pp. 726-742.

[17] Gauri, S.K., and Chakraborty, S., 2008. "Improved Recognition of Control Chart Patterns Using Artificial Neural Networks", *International Journal of Advanced Manufacturing Technology* 36, pp. 1191-1201.

[18] Guh, R.S., 2005. "A Hybrid Learning-Based Model for On-Line Detection and Analysis of Control Chart Patterns", *Computers and Industrial Engineering* 49, pp. 35-62.

[19] Guh, R.S., 2007). "On-Line Identification and Quantification of Mean Shifts in Bivariate Processes Using a NN-Based Approach", *Quality and Reliability Engineering International* 23, pp. 367-385.

[20] Guh, R.S., 2002. "Robustness of the Neural Network Based Control Chart Pattern Recognition System to Non-Normality", *International Journal of Quality and Reliability Management* 19, pp. 97-112.

[21] Guh, R.S., and Hsieh, Y.C., 1999. "A Neural Network Based Model for Abnormal Pattern Recognition of Control Charts", *Computers and Industrial Engineering* 36, pp. 97-108.

[22] Guh, R.S., and Shiue, Y.R., 2005. "On-Line Identification of Control Chart Patterns Using Self-Organizing Approaches", *International Journal of Production Research* 43, pp. 1225-1254.

[23] Guh, R.S., and Tannock, J.D.T., 1999. "Recognition of Control Chart Concurrent Patterns Using a Neural Network Approach", *International Journal of Production Research* 37, pp. 1743-1765.

[24] Guh, R.S., Tannock, J.D.T., and O'Brien, C., 1999a. "Intellispc: A Hybrid Intelligence Tool for On-Line Economical Statistical Process Control", *Expert System with Applications* 17, pp. 195-212.

[25] Guh, R.S., Zoriassatine, F., Tannock, J.D.T., and O'Brien, C., 1999b. "On-Line Control Chart Pattern Detection and Discrimination - a Neural Network Approach", *Artificial Intelligence in Engineering* 13, pp. 413-425.

[26] Hassan, A., Nabi Baksh, M.S., and Shaharoun, A.M., 2000. "Issues in Quality Engineering Research", *International Journal of Quality and Reliability Management* 17, pp. 858-875.

[27] Hassan, A., Nabi Baksh, M.S., Shaharoun, A.M., and Jamaludin, H., 2003. "Improved SPC Chart Pattern Recognition Using Statistical Features". *International Journal of Production Research* 41, pp. 1587-1603.

[28] Hassan, A., Nabi Baksh, M.S., Shaharoun, A.M., and Jamaluddin, H., 2006. "Feature Selection for SPC Chart Pattern Recognition Using Fractional Factorial Experimental Design". *Virtual International Conference on Innovative Production Machines and Systems, I*IPROM*.

[29] Hassan, A., and Nabi Baksh, M.S., 2008. "An Improved Scheme for On-Line Recognition of Control Chart Patterns". *Virtual International Conference on Innovative Production Machines and Systems*, pp. 171-175.

[30] Haykin, S., 1999. "Neural Networks: A Comprehensive Foundation", 2nd ed., *Prentice Hall*, New Jersey.

[31] Hwarng, H.B., 1995a. "Proper and Effective Training of a Pattern Recognizer for Cyclic Data", *IEEE Transactions* 27, pp. 746-756.

[32] Hwarng, H.B., 1995b. "Multilayer Perceptron for Detecting Cyclic Data on Control Charts", *International Journal of Production Research* 33, pp. 3101-3117.

[33] Hwarng, H.B., 1997. "A Neural Network Approach to Identifying Cyclic Behaviour on Control Charts", *International Journal of System Science* 28, pp. 99-112.

[34] Hwarng, H.B., and Chong, C.W., 1995. "Detecting Process Non-Randomness Through a Fast and Cumulative Learning ART-Based Pattern Recognizer", *International Journal of Production Research* 33, pp. 1817-1833.

[35] Hwarng, H.B., and Hubele, N. F., 1991. "X-Bar Chart Pattern Recognition Using Neural Nets", *ASQC quality congress transactions*, pp. 884-889.

[36] Hwarng, H.B, and Hubele, N.F., 1993. "Back-Propagation Pattern Recognizers for X-Bar Control Charts: Methodology and Performance", *Computers and Industrial Engineering* 24, pp. 219-235.

[37] Khaw, J.F.C., Lim, B.S., and Lim, L.E.N., 1995. "Optimal Design of Neural Networks Using the Taguchi Method", *Neurocomputing* 7, pp. 225-245.

[38] Loh, A.S., Woo, S.W., and Lim, C.P., 2000. "Classification of Speech Signals Using Wavelet and Artificial Neural Networks: A Preliminary Investigation", *Proceeding of the First Regional Symposium on Quality and Automation*, pp. 359-364.

[39] Montgomery, D.C., 1996. "Design and Analysis of Experiment", 3rd ed., *John Wiley & Sons*, New York.

[40] Montgomery, D.C., 2005. "Introduction to Statistical Quality Control", 5th ed., *John Wiley & Sons*, New York.
[41] Nelson, L.S., 1985. "Interpreting Shewhart X-Bar Control Chart", *Journal of Quality Technology* (17:2), pp. 114-116.
[42] Nelson, L.S., 1989. "Standardization of *Shewhart* Control Chart", *Journal of Quality Technology* (21:4), pp. 287-289.
[43] Niaki, S.T.A., and Abbasi, B., 2005. "Fault Diagnosis in Multivariate Control Charts Using Artificial Neural Networks", *Quality and Reliability Engineering International* 21, pp. 825-840.
[44] Packianather, M.S., Drake, P.R., and Rowlands, H., 2000. "Optimizing the Parameters of Multilayered Feedforward Neural Networks Through Taguchi Design of Experiments", *Quality and Reliability Engineering International* 16, pp. 461-473.
[45] Pandya, A.S., and Macy, R.B., 1995. "Pattern Recognition with Neural Network in C++". *CRC Press*, Boca Raton, Florida.
[46] Pao, Y.H., 1989. "Adaptive Pattern Recognition and Neural Networks". *Addison-Wesley*, Reading, MA.
[47] Patterson, D.W., 1996. "Artificial Neural Networks: Theory and Applications". *Prentice Hall*, Singapore.
[48] Perry, M.B., Spoerre, J.K., and Velasco, T., 2001. "Control Chart Pattern Recognition Using Back Propagation Artificial Neural Networks", *International Journal of Production Research* 39, pp. 3399-3418.
[49] Pham, D.T., and Chan, A.B., 1998. "Control Chart Pattern Recognition Using a New Type of Self Organizing Neural Networks", *Proc. Instn. Mech. Engrs* 212 (Part I), pp. 115-127.
[50] Pham, D.T., and Chan, A.B., 1999. "A Synergistic Self-Organizing System for Control Chart Pattern Recognition" in S. G. Tzafestas. Advances in Manufacturing: Decision, Control and Information Technology. *Springer-Verlag*, London.
[51] Pham, D.T., and Chan, A.B., 2001. "Unsupervised Adaptive Resonance Theory Neural Networks for Control Chart Pattern Recognition", *Proc. Instn. Mech. Engrs*. 215 (Part B), pp. 59-67.
[52] Pham, D.T., and Oztemel, E., 1993. "Control Chart Pattern Recognition Using Combinations of Multilayer Perceptrons and Learning Vector Quantization Neural Networks", *Proc. Instn. Mech. Engrs*. 207, pp. 113-118.
[53] Pham, D.T., and Oztemel, E., 1994. "Control Chart Pattern Recognition Using Learning Vector Quantization Networks", *International Journal of Production Research* 32, pp. 721-729.
[54] Pham, D.T., and Wani, M.A., 1997. "Feature-Based Control Chart Pattern Recognition", *International Journal of Production Research* 35, pp. 1875-1890.
[55] Pugh, G.A., 1989. "Synthetic Neural Networks for Process Control", *Computers and Industrial Engineering* (17:1-4), pp. 24-26.
[56] Ripley, B.D., 1994. "Neural Networks and Related Methods for Classification", *J. Royal Statistical Society* 56, pp. 409-456.
[57] Schalkoff, R.J., 1997. "Artificial neural networks", *McGraw-Hill*, New York.
[58] Tontini, G., 1996. "Pattern Identification in Statistical Process Control Using Fuzzy Neural Networks", *Proceedings of the 5th IEEE International Conference on Fuzzy System* 3, pp. 2065-2070.
[59] Tontini, G., 1998. "Robust Learning and Identification of Patterns in Statistical Process Control Charts Using a Hybrid RBF Fuzzy Artmap Neural Network", *The 1998 IEEE International Joint Conference on Neural Network Proceedings* (IEEE World Congress on Computational Intelligence) 3, pp. 1694-1699.
[60] Utku, H., 2000. "Application of the Feature Selection Method to Discriminate Digitized Wheat Varieties", *Journal of Food Engineering* 46, pp. 211-216.
[61] Velasco, T., 1993. "Back-Propagation Artificial Neural Networks for the Analysis of Quality Control Charts", *Computers and Industrial Engineering* (25:1-4), pp. 397-400.

[62] Wang, C.H., Kuo, W., and Qi, H., 2007. "An Integrated Approach for Process Monitoring Using Wavelet Analysis and Competitive Neural Network", *International Journal of Production Research* 45, pp. 227-244.

[63] Wani, M.A., and Pham, D.T., 1999. "Efficient Control Chart Pattern Recognition Through Synergistic and Distributed Artificial Neural Networks", *Proc. Instn. Mech. Engrs.* 213 (Part B), pp. 157-169.

[64] Yang, M.S., and Yang, J.H., 2002. "A Fuzzy-Soft Learning Vector Quantization for Control Chart Pattern Recognition", *International Journal of Production Research* 40, pp. 2721-2731.

[65] Yu, J.B., and Xi, L.F., 2009. "A Neural Network Ensemble-Based Model for On-Line Monitoring and Diagnosis of Out-of-Control Signals in Multivariate Manufacturing Processes", *Expert Systems with Applications* 36, pp. 909-921.

[66] Zeki, A.A., and Zakaria, M.S., 2000. "New Primitive to Reduce the Effect of Noise for Handwritten Features Extraction", *IEEE 2000 Tencon Proceedings: Intelligent Systems and Technologies for the new millenium*, pp. 24-27.

[67] Zorriassatine, F., and Tannock, J.D.T., 1998. "A Review of Neural Networks for Statistical Process Control", *Journal of Intelligent Manufacturing* 9, pp. 209-224.

[68] Zorriassatine, F., Tannock, J.D.T., and O'Brien, C, 2003. "Using Novelty Detection to Identify Abnormalities Caused by Mean Shifts in Bivariate Processes", *Computers and Industrial Engineering* 44, pp. 385-408.

[69] Zurada, J.M., 1992. "Introduction to artificial neural systems". *West Publishing Company*, St. Paul, MN.