International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:6, No:12, 2012

Control Chart Pattern Recognition using Wavelet based Neural Networks

Jun Seok Kim, Cheong-Sool Park, Jun-Geol Baek, and Sung-Shick Kim

Abstract—Control chart pattern recognition is one of the most important tools to identify the process state in statistical process control. The abnormal process state could be classified by the recognition of unnatural patterns that arise from assignable causes. In this study, a wavelet based neural network approach is proposed for the recognition of control chart patterns that have various characteristics. The procedure of proposed control chart pattern recognizer comprises three stages. First, multi-resolution wavelet analysis is used to generate time-shape and time-frequency coefficients that have detail information about the patterns. Second, distance based features are extracted by a bi-directional Kohonen network to make reduced and robust information. Third, a back-propagation network classifier is trained by these features. The accuracy of the proposed method is shown by the performance evaluation with numerical results.

Keywords—Control chart pattern recognition, Multi-resolution wavelet analysis, Bi-directional Kohonen network, Back-propagation network, Feature extraction.

I. Introduction

IN manufacturing processes, control charts are widely used for monitoring and diagnosis. [1], [2] Every control chart has decision criteria such as upper and lower control limits. If one or more points in the control chart fall outside control limits, a process is considered as out-of control. However, the shape of control chart pattern (CCP) within control limits could be abnormal because of assignable causes such as tool wearing, lack of resources, vibration, machine troubles, and so on. Therefore, the recognition of these CCPs is very important to identify the process state and diagnose the causes of abnormal situation. The typical CCPs are illustrated in Fig. 1 and described with well-known causes as follows. [1]

1) Cyclic patterns (CYC): These patterns may result from systematic environmental changes such as temperature, regular rotation of operators or machines, or fluctuation in voltage or pressure or other variable in the equipment.

2) Systematic patterns (SYS): Variations of normal patterns (NOR) that are unsystematic, unexpected, and unpredictable. Such patterns may result from systematic environmental changes. The systematic pattern moves up and down (or down and up) consecutively.

Jun Seok Kim is with the School of Industrial Management Engineering, Korea University, Seoul, 136-713, Republic of Korea.

Cheong-Sool Park is with the Graduate School of Information Management and Security, Korea University, Seoul, 136-713, Republic of Korea.

Jun-Geol Baek is an associate professor in the School of Industrial Management Engineering, Korea University, Seoul, 136-713, Republic of Korea (Corresponding author to provide phone: +82-2-3290-3396; fax: +82-2-929-5888; e-mail: jungeol@korea.ac.kr).

Sung-Shick Kim is a professor emeritus in the School of Industrial Management Engineering, Korea University, Seoul, 136-713, Republic of Korea.

3) Drift patterns (UD, DD): Possible causes of trends may result from a gradual wearing out or deterioration of a tool or some other critical process component.

4) Shift patterns (US, DS): These patterns are usually due to changes in operators, methods, raw materials, or machines.

5) Stratification patterns (STA): One potential cause is incorrect calculation of control limits by one or more samples from several different underlying distributions within each subgroup.

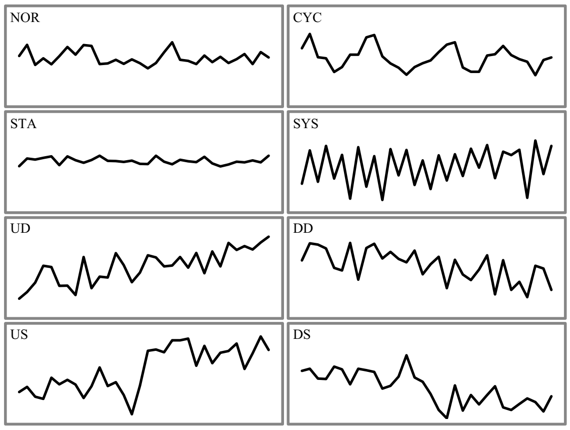

Fig. 1 Typical control chart patterns.

To recognize CCPs simply in a uniform way, Western Electric used a heuristic approach for detecting abnormal patterns. However, the false-alarm rate rises as the number of rules increases, and the implementation was accomplished manually. This resulted in high computational costs and inflexibility due to the difficulty of incorporating new heuristics. [3]

As computing technology develops, artificial neural networks (ANN) such as Learning Vector Quantization (LVQ), Back-propagation network (BPN), and Self-organizing map (SOM) are used for CCP recognition. However, it is hard to classify CCPs by ANN directly because noises are inherent and means or variances of patterns are very similar. Therefore, many researches about the feature extraction for CCPs have been conducted.

Hassan et al. [4] proposed the method that extracts statistical features such as mean, standard deviation, skewness, autocorrelation, and so on. These statistical features are used as inputs of multi-layer perceptron (MLP). However, similar problems like heuristic approaches could arise when one or more type of CCPs are added or removed.

Yang et al. [5] used correlation coefficients between reference patterns and sample patterns. Correlation coefficients could be distorted if there is a distinct time shift between reference patterns and sample patterns.

Multi-resolution wavelet analysis (MRWA) is used in several researches to denoise or extract appropriate features because CCPs could be regarded as signal-type data. Al-Assaf [6] computed the absolute mean of detail coefficients at each resolution level. Wang et al. and Cheng [7] et al. used MRWA to denoise samples of CCPs. Assaleh and Al-Assaf [8] modified the MRWA algorithm to induce multi-resolution discrete cosine transform and improve ANN performance. However, insufficient use of wavelet coefficients may cause the loss of information. Therefore, it is necessary to use wavelet coefficients without loosing information.

SOM is useful tool for visualizing low-dimensional cluster views of high dimensional data. Cheng et al. [9] and Pham et al. [10] used SOM to make clusters in their researches. However, it could be difficult to make distinguished clusters for similar CCPs having different classes because SOM is a kind of ANN which is trained by unsupervised learning. Therefore, a kind of supervised SOM such as bi-directional Kohonen network (BDK) is necessary to make distinguished clusters.

In this paper, we proposed Wavelet-BDK-BPN (WBB) method for CCP recognition precisely. MRWA is used to generate detail and meaningful information from CCPs. BDK are used for extracting robust and reduced features and BPN is used to classify CCPs.

The procedure of WBB method is illustrated in Fig. 2. In training phase, wavelet coefficients are generated by wavelet decomposition from historical CCPs. After a BDK is trained by wavelet coefficients, distance-based features are extracted from the input map of BDK. Finally, BPN is trained by these distance-based features.

In monitoring phase, wavelet coefficients are generated when a new CCP occurs. After a distance-based feature is generated by trained BDK, BPN predict the state of process. When a new CCP is recognized as one of abnormal patterns, the diagnosis of the process state would be accomplished.

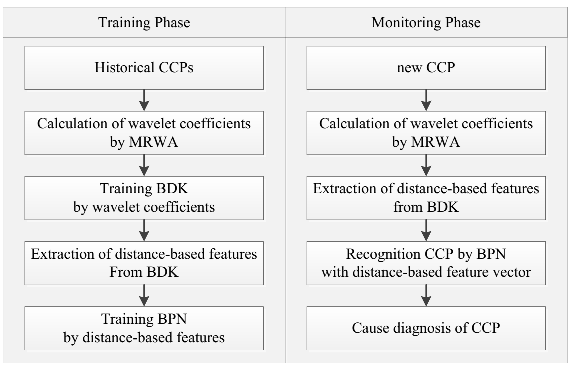

Fig. 2 The procedure of WBB method

In section II, MRWA and the way of generating coefficients for WBB method are mentioned. BDK and distance-based features are described in Section III. BPN and its structure are explained in Section IV. In Section V, the experimental description and the result are shown. Finally, Section VI contains conclusions and future works.

II. Multi-Resolution Wavelet Analysis (MRWA)

Signal-type data could be decomposed or reconstructed by wavelet transform with the scale basis function (father wavelet, ϕ) and the detail basis function (mother wavelet, ψ). [11] The scale function is described in (1) and the detail function is described in (2).

$$\phi_{j,k} = 2^{\frac{j}{2}}\phi(2^j x - k) \tag{1}$$

$$\psi_{j,k} = 2^{\frac{j}{2}}\psi(2^j x - k) \tag{2}$$

where j is a decomposition-level index and k is a translation index. In discrete wavelet transform, original signal is represented by linear combination of scale and detail functions in $L^2(R)$ as shown in (3).

$$f(x) = \sum_k c_{j,k}\phi_{j,k}(x) + \sum_j \sum_k d_{j,k}\psi_{j,k}(x) \tag{3}$$

where $c_{j,k}$ is the scale coefficient at j, k and $d_{j,k}$ is the detail coefficient at j, k. Scale and detail coefficients could be computed by (4) and (5).

$$c_{j+1,k} = \sum_n h_{n-2k} c_{j,n} \tag{4}$$

$$d_{j+1,k} = \sum_n (-1)^h h_{-n+2k+1} c_{j,n} \tag{5}$$

where n is a sequence index and h is a scale or detail coefficient of basis function for the discrete wavelet transform. MRWA is possible because scale and detail coefficients are produced by wavelet decomposition at each level as illustrated in Fig. 3. The original signal is regarded as a scale coefficient vector at level 0. Scale and detail coefficients are generated with half-length scale coefficients of previous level by the wavelet decomposition.

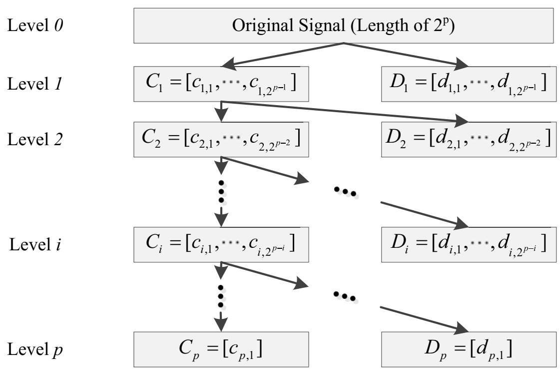

Fig. 3 Wavelet transform

Scale coefficients reflect the weighted average scheme of the corresponding section at previous level, and detail coefficients reflect the weighted difference scheme. That is, scale coefficients describe shapes and trends of the pattern, and detail coefficients represent directions and amounts of the change.

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:6, No:12, 2012

In this paper, the Haar wavelet is used for MRWA as the basis function. The Haar wavelet function is widely used because it is simple and easy to interpret. The Haar wavelet function could be calculated as (6) and (7).

$$\phi(x) = \begin{cases} 1, & 0 \leq \text{x} < 1 \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

$$\psi(x) = \begin{cases} 1, & 0 \leq \text{x} < \frac{1}{2} \\ -1, & \frac{1}{2} \leq \text{x} < 1 \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

If $X = [x_1, x_2, \cdots, x_{2^p}]$, we could generate the coefficient vector G through (8).

$$\begin{aligned} C_i &= [c_{i,1}, \cdots, c_{i,2^{p-i}}], \qquad i = 1, \cdots, p \\ C &= [C_1^T, \cdots C_p^T]^T \\ D_i &= [d_{i,1}, \cdots, d_{i,2^{p-i}}], \qquad i = 1, \cdots, p \\ D &= [D_1^T, \cdots D_p^T]^T \end{aligned}$$

$$G = [C^T, D^T]^T \tag{8}$$

Not all coefficients are used as features in other approaches. In general, the decomposition level i is determined arbitrarily and coefficients at level i are used as features. However, the loss of information could be possible because all coefficients contain the pattern information. For instance, detail coefficients at the first level are often regarded as noise in spite of systematic patterns are expressed. Therefore, all coefficients of all decomposition levels are used as features in this paper.

III. Bi-Directional Kohonen Network (BDK)

A kind of unsupervised learning, SOM is useful tool for visualizing low-dimensional cluster views of high dimensional data. However, because SOM could not reflect the relationship between the input and output, BDK is used for feature extraction in this paper. BDK is a type of supervised Kohonen network to overcome the problem of scaling and balancing the input and output data. [12] In a BDK network, the input map (X-map) and the output map (Y-map) are connected each other as illustrated in Fig. 4.

In the first updating step of a BDK, only the weights of the nodes in the X-map are adapted. The location of the winning node is determined by the dominating similarity measure between an object Y and the Y-map units as shown in (9).

$$S^X_{winner}(i,k) = (1-\alpha_t)S(X_i, W^X_k) + \alpha_t S(Y_i, W^Y_k) \tag{9}$$

where $S(X_i, W_k^X$ and $S(Y_i, W_k^Y)$ are similarity measures for the object pair (X_i, Y_i) and the node k in the X-map and Y-map. W_k^X and W_k^Y are weight vectors at node k in X-map and Y-map. α_t is the relative weight between X-map and Y-map and t refers to the epoch number of training. In the reverse step, the Y-map nodes of BDK are updated by the winner

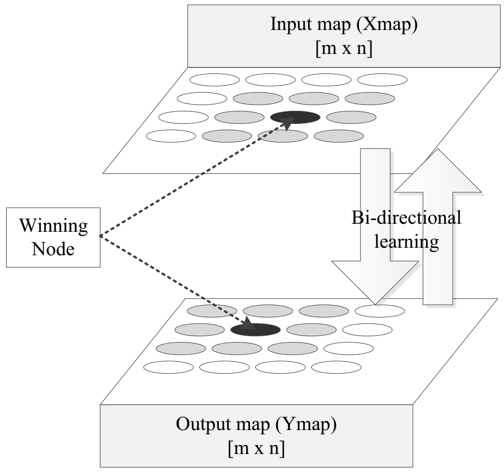

Fig. 4 The structure of BDK

which is determined by the dominating similarity between the object X and the corresponding X-map units according to (10).

$$S^Y_{winner}(i,k) = \alpha_t S(X_i, W^X_k) + (1-\alpha_t)S(Y_i, W^Y_k) \tag{10}$$

Nodes in X-map and Y-map are updated both directions as shown in Fig. 4. However, misclassification could be possible because BDK classifies by the distance measure between the input vector and the weight at each node in X-map. For example, a trained 7×7 X-map is illustrated in Fig. 5.

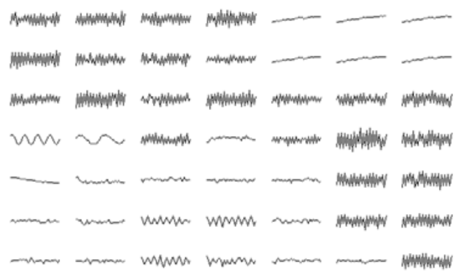

Fig. 5 An example of a trained X-map of BDK (7×7)

Some nodes have distinguished shape with others, but some nodes are similar to other nodes such as stratification and systematic. Therefore, we use this trained X-map to generate new distance-based features in this paper.

After BDK training, a similarity feature between X_i and all nodes in X-map could be calculated as shown in (11).

$$d(X_i, k) = S(X_i, W^X_k), k = 1, \cdots, mn$$

$$D_i = \begin{bmatrix} d(X_i, 1) \\ d(X_i, 2) \\ \vdots \\ d(X_i, mn) \end{bmatrix} \tag{11}$$

where D_i is a similarity vector for X_i. D_i means the amount of similarity between X_i and trained X-map of the BDK. These similarity vectors are used as input data of BPN.

The length of D_i is determined by the size of X-map size. Therefore, it could be possible to shrink the number of wavelet coefficients.

IV. Back-Propagation Network (BPN)

BPN is one of the widely used ANNs and proven for multi-class classification tasks in CCP recognition. [13] The network structure is composed with the input, hidden, and output layer. Each layer has one or more nodes that are connected with other nodes as illustrated in Fig. 6.

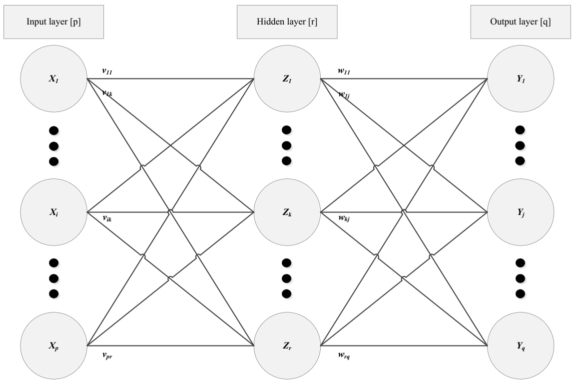

Fig. 6 The structure of BPN

The number of nodes in the input layer and the output layer is determined according to dimensions of input and output data. It is important to decide the number of nodes in the hidden layer because the network performance could be affected. Similarity vectors calculated by BDK are used as input data and class labels are used as output data.

In this paper, all wavelet coefficients of all decomposition levels are used as statistical features in order to reflect process characteristics. After training BDK by these wavelet coefficients, distance-based features are extracted from X-map of BDK. Finally, the CCPs recognition is conducted by BPN which is trained by distance-based features.

V. Experiment

A. Experimental Description

Eight patterns are generated to evaluate the CCP recognition performance as described in Table I. r_t is a random number at time t from a standard normal distribution ($N(0,1)$). The window size (ws) is 64 and every pattern is generated 500 observations at each experiment. The 5-fold cross-validation and 20 iterations of experiments are executed.

To compare the performance of WBB method, the approach of Cheng et al. [9] and Yoon et al. [14] are used. The approach of Cheng et al. uses raw data and wavelet coefficients of second decomposition level as input vectors of SOM. The approach of Yoon et al. uses raw data as input of BDK.

TABLE I
Experimental Configuration

Pattern	Parameter	Value	Amount at time t
NOR	Mean(μ)	0	$y_t^N = \mu + r_t\sigma$
	Stdev.(σ)	1	
SYS	Departure(d)	[1,3]σ	$y_t^N + (-1)^t$
CYC	Amplitude(a)	[0.5,3]σ	$y_t^N + a\sin(\frac{2\pi}{T}t)$
	Period(T)	[1/8, 1/4]ws	
	Position(t_0)	{1, 1/8ws, 1/4ws}	
UD	Gradient(g)	[0.05, 0.15]σ	$y_t^N + g(t-t_0)$
	Position(t_0)	{1, 1/8ws, 1/4ws}	
DD	Gradient(g)	[-0.15, -0.05]σ	$y_t^N + g(t-t_0)$
	Position(t_0)	{1, 1/8ws, 1/4ws}	
US	Magnitude(s)	[1,3]σ	$y_t^N + \frac{1}{2}\{sgn(t-t_0)+1\}$
	Position(t_0)	{1, 1/8ws, 1/4ws}	
DS	Magnitude(g)	[1,3]σ	$y_t^N + \frac{1}{2}\{sgn(t-t_0)+1\}$
	Position(t_0)	{1, 1/8ws, 1/4ws}	
STA	Stdev.(σ)	[0.2, 0.4]σ	$\mu + r_t\sigma'$

B. Experimental Results

The number of hidden nodes in BPN and the map size of SOM and BDK are determined by 5-cross validation as shown in Table II.

TABLE II
Map Size and Hidden Nodes

Methods	SOM/BDK map size	Hidden nodes
Cheng	6×6	20
Yoon	5×5	15
proposed	6×6	25

The results of the CCP recognition accuracy are summarized in Table III. The overall accuracy of WBB method is better than other methods. It seems that the MRWA with entire coefficients could provide precise and detail information about CCPs.

The NOR accuracy of Cheng's approach is too lower than other methods. Thus, we could think that the method with BDK have better performance than the method with SOM because of the weakness of the unsupervised learning.

TABLE III
Comparison of CCPR Accuracies (%)

CCP	Cheng [9]	Yoon [14]	WBB
NOR	6.92	95.12	96.23
SYS	99.98	99.88	99.92
CYC	91.88	95.42	96.23
UD	94.83	98.23	98.40
DD	97.58	98.08	98.13
US	99.17	97.95	98.08
DS	98.50	97.55	97.63
STA	99.00	100	100
(Avg)	**85.98**	**97.78**	**98.08**

The confusion matrices of CCP recognition methods are described in Table IV, V, VI. Almost of normal patterns are misclassified to stratification patterns because the shape of their patterns are very similar and SOM is a kind of unsupervised learning. Also, cyclic patterns are misclassified to stratification patterns while other methods misclassified cyclic patterns to normal patterns.

The performance of methods of Yoon and WBB are similar, but WBB method is slightly better for almost patterns. Therefore, we could think that wavelet coefficients can provide more detail information about CCPs. The overall performance of WBB methods is better than other methods.

TABLE IV
CONFUSION MATRIX (CHENG)

-	NOR	SYS	CYC	UD	DD	US	DS	STA
NOR	**6.92**	0	0	0	0	0.88	1.82	90.38
SYS	0	**99.98**	0	0.02	0	0	0	0
CYC	1.28	0	**91.88**	0	0	0	1.02	5.82
UD	0.02	0	0	**94.83**	0	5.05	0	0.10
DD	0.07	0	0	0	**97.58**	0	2.25	0.10
US	0.15	0	0	0.40	0	**99.17**	0	0.28
DS	0.10	0	0	0	0	0	**98.50**	1.40
STA	1.00	0	0	0	0	0	0	**99.00**

TABLE V
CONFUSION MATRIX (YOON)

-	NOR	SYS	CYC	UD	DD	US	DS	STA
NOR	**95.12**	0.07	1.78	0.12	0.10	1.48	1.22	0.12
SYS	0.02	**99.88**	0	0.07	0	0	0.03	0
CYC	3.82	0	**95.42**	0.05	0.08	0.13	0.50	0
UD	0.52	0	0.02	**98.23**	0	1.23	0	0
DD	0.55	0	0.05	0.02	**98.08**	0	1.30	0
US	1.48	0	0.10	0.45	0.02	**97.95**	0	0
DS	1.52	0	0.35	0.05	0.53	0	**97.55**	0
STA	0.00	0	0.00	0	0	0	0	**100**

TABLE VI
CONFUSION MATRIX (WBB)

-	NOR	SYS	CYC	UD	DD	US	DS	STA
NOR	**96.23**	0	1.33	0.10	0.15	1.08	0.97	0.13
SYS	0.07	**99.92**	0	0	0	0	0.02	0
CYC	3.13	0	**96.23**	0.02	0.03	0.15	0.43	0
UD	0.33	0	0.03	**98.40**	0	1.22	0.02	0
DD	0.42	0	0.07	0	**98.13**	0.02	1.37	0
US	1.30	0	0.12	0.50	0	**98.08**	0	0
DS	1.43	0	0.43	0	0.50	0	**97.63**	0
STA	0.00	0	0.00	0	0	0	0	**100**

VI. CONCLUSION

In this paper, we used entire wavelet coefficients to reflect the precise information about CCP. The distance-based feature from X-map of the BDK network is generated. The performance of the proposed method is better than other competitive methods.

In the future, research will be enhanced in three ways. First, the feature extraction for more precise and robust features for cyclic, drift and shift patterns will be conducted. Second, more realistic CCPs are supplemented such as concurrent patterns. Finally, the proposed method will be extended to multivariate CCP recognition problems.

ACKNOWLEDGMENT

This research was supported by the MKE (Ministry of Knowledge Economy), Korea, under the IT R&D Infrastructure Program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2012-(B1100-1101- 0002)).

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012-0008332).

REFERENCES

[1] D. C. Montgomery, *Introduction to Statistical Quality Control*, 5th ed. John Wiley & Sons, 2004.
[2] J. Guo, S. Guo, and X. Yu, "Monitoring and diagnosis of manufacturing process using extreme learning machine," *Advanced Science Letters*, vol. 4, no. 6-7, pp. 6–7, 2011.
[3] I. Masood and A. Hassan, "Statistical features-ann recognizer for bivariate process mean shift pattern recognition," in *Intelligent and Advanced Systems (ICIAS), 2010 International Conference on*. IEEE, 2010, pp. 1–6.
[4] A. Hassan, M. Baksh, A. Shaharoun, and H. Jamaluddin, "Improved spc chart pattern recognition using statistical features," *International Journal of Production Research*, vol. 41, no. 7, pp. 1587–1603, 2003.
[5] J. Yang and M. Yang, "A control chart pattern recognition system using a statistical correlation coefficient method," *Computers & Industrial Engineering*, vol. 48, no. 2, pp. 205–221, 2005.
[6] Y. Al-Assaf, "Recognition of control chart patterns using multi-resolution wavelets analysis and neural networks," *Computers & Industrial Engineering*, vol. 47, no. 1, pp. 17–29, 2004.
[7] C. H. Wang and W. Kuo, "Identification of control chart patterns using wavelet filtering and robust fuzzy clustering," *Journal of Intelligent Manufacturing*, vol. 18, no. 3, pp. 343–350, 2007.
[8] K. Assaleh and Y. Al-assaf, "Features extraction and analysis for classifying causable patterns in control charts," *Computers & industrial engineering*, vol. 49, no. 1, pp. 168–181, 2005.
[9] H. P. Cheng and C. S. Cheng, "Control chart pattern recognition using wavelet analysis and neural networks," *Journal of Quality Vol*, vol. 16, no. 5, p. 311, 2009.
[10] D. Pham and A. Chan, "Control chart pattern recognition using a new type of self-organizing neural network," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 212, no. 2, pp. 115–127, 1998.
[11] R. T. Ogden, *Essential Wavelets for Statistical Applications and Data Analysis*. Philadelphia: SIAM, 1992.
[12] W. Melssen, R. Wehrens, and L. Buydens, "Supervised kohonen networks for classification problems," *Chemometrics and Intelligent Laboratory Systems*, vol. 83, no. 2, pp. 99–113, 2006.
[13] L. Fausett, *Fundamentals of Neural Networks*. Prentice Hall, 1993.
[14] J.-J. Yoon, C.-S. Park, J. S. Kim, and J.-G. Baek, "Recognition of control chart pattern using bi-directional kohonen network and artificial neural network," *Journal of the Korea Society for Simulation*, vol. 20, no. 4, pp. 115–125, 2011.