

Bi-Directional Kohonen Network와 인공신경망을 사용한 관리도 패턴 인식

윤재준¹ · 박정술¹ · 김준석¹ · 백준걸^{1†}

Recognition of Control Chart Pattern using Bi-Directional Kohonen Network and Artificial Neural Network

Jae-Jun Yun · Cheong-Sool Park · Jun-Seok Kim · Jun-Geol Baek

ABSTRACT

Manufacturing companies usually manage the process to achieve high quality using various types of control chart in statistical process control. When an assignable cause occurs in a process, the data in the control chart changes with different patterns by the specific causes. It is important in process control to classify the CCP (Control Chart Pattern) recognition for fast decision making. In former research, gathered data from process used to apply as raw data, leads to degrade the performance of recognizer and to decrease the learning speed. Therefore, feature based recognizer, employing feature extraction method, has been studied to enhance the classification accuracy and to reduce the dimension of data. We propose the method to extract features that take the distances between CCP data and reference vector generated from BDK (Bi-Directional Kohonen Network). We utilize those features as the input vectors in ANN (Artificial Neural Network) and compare with raw data applied ANN to evaluate the performance.

Key words : Control Chart Pattern Recognition, Feature Extraction, Bi-Directional Kohonen Network, Artificial Neural Network

요약

제품의 품질 수준 제고를 위해 통계적 공정 관리(SPC : Statistical Process Control)의 다양한 관리도가 기업의 생산 공정을 관리하는데 사용된다. 관리도에 기록되는 공정 데이터는 특정 요인(Assignable Cause)에 의한 이상이 발생했을 때 그 요인에 따라 서로 다른 패턴(Pattern)으로 변화한다. 이러한 패턴을 구별하는 관리도 패턴(CCP : Control Chart Pattern) 인식(Recognition)은 공정에 대한 관리자의 빠른 의사 결정을 위해 매우 중요하다. 앞 선 연구들은 수집되는 원 데이터를 가공 하지 않고 그대로 사용하였기 때문에 인식기(Recognizer)의 성능과 학습 속도가 저하되는 문제점이 있었다. 따라서 최근 데이터의 차원 축소와 인식기의 성능 향상을 위해 특징 추출법(Feature Extraction)을 적용한 특징 기반 인식기(Feature based Recognizer)에 대한 연구가 활발히 진행 중이다. 본 논문은 BDK(Bi-Directional Kohonen Network)를 사용하여 CCP의 참조 벡터(Reference Vector)를 생성하고 참조 벡터와 CCP 데이터의 거리를 기반으로 하는 특징을 추출하였다. 추출된 특징을 인공 신경망 기반 인식기의 입력 벡터로 사용하여 학습하였으며 원 데이터를 사용하여 학습하는 인공신경망 인식기와 예측 정확도 비교를 통해 제안 알고리즘의 성능을 평가하였다.

주요어 : 관리도 패턴 인식, 특징 추출, Bi-Directional Kohonen Network, 인공신경망

1. 서론

다양한 제품이 생산되고 기업 간의 경쟁이 치열해질수록 높은 품질 수준의 제품 생산은 기업의 수익 창출과 경쟁력 확보를 위해 매우 중요하다. 제품 생산 과정에서 품질을 높은 수준으로 유지하기 위해서는 생산의 안정된 상

접수일(2011년 9월 26일), 심사일(1차 : 2011년 11월 6일),
게재 확정일(2011년 11월 6일)

¹⁾ 고려대학교 산업경영공학과

주저자 : 윤재준

교신저자 : 백준걸

E-mail: jungeol@korea.ac.kr

태를 관찰하기 위한 지속적인 모니터링과 불안정한 상태에 대한 빠른 탐지가 필요하다. 이를 위해 오랫동안 많은 연구가 진행되어 왔으며 그 중에서 통계적 공정 관리(SPC : Statistical Process Control)가 현재까지 가장 널리 사용되고 있는 공정 관리 기법이다. 통계적 공정 관리에서 공정의 변동성과 이상을 탐지하기 위해 다양한 관리도가 사용되는데 슈하트 차트(Shewhart Chart)가 대표적으로 사용된다. 대부분의 관리도는 중심선과 관리 한계선으로 구성되며 관리 한계선을 넘었을 때 공정에 이상이 발생한 것으로 판단하고 있다(Montgomery, 2005).

안정된 상태의 공정 데이터는 자연적 잡음을 동반한다. 하지만 공정 내, 외부의 특정 요인(Assignable Cause)에 의해 이상이 발생 하였을 때 공정 데이터는 자연적 잡음을 가지는 형태를 벗어나 어떤 패턴(Pattern)을 가지는 형태로 변하게 되는데 이를 관리도 패턴(CCP : Control Chart Pattern)이라 한다. 일반적으로 사용되는 관리도는 공정의 이상 탐지를 목적으로 하기 때문에 다양한 원인에 의해 발생하는 CCP에 대한 분석 및 분류 정보를 관리자에게 제공하지 못한다. CCP에 대한 정보를 신속하고 정확하게 분석 및 분류하는 것은 공정 상태에 대한 관리자의 정확하고 빠른 의사 결정을 위해 매우 중요하다(Hassan 등, 2003). 따라서 CCP의 정확한 분류를 목적으로 하는 관리도 패턴 인식(Control Chart Pattern Recognition)분야에 대한 연구가 지속적으로 이루어지고 있다.

Western Electric Handbook(1958)에 의해 정의된 다양한 CCP 중 대표적으로 관찰되는 CCP는 표 1과 같다(Wang 등, 2008).

그림 1은 대표적 CCP의 예를 나타낸다.

표 1에서 정의된 CCP의 인식을 위하여 기존 연구에서는 다양한 기법들이 활용되었다. 초창기 Western Electric은 휴리스틱 접근(Heuristic approach)으로 관리도 패턴

인식을 수행하였으며 이러한 방법은 다양한 규칙(Rule)들이 추가 될 때 마다 가성 알람(False-alarm) 발생 비율이 증가하는 문제점을 야기시켰다(Perry 등, 2007). 이 후 인공신경망(ANN : Artificial Neural Network)의 발달과 더불어 LVQ(Learning Vector Quantization), MLP(Multi-Layer Perceptron), 역 전파 네트워크(BPN : Back Propagation Network)와 같은 다양한 인공신경망을 활용한 연구들이 수행되었다(Pham 등, 1994; Ho 등, 1999; Hwang 등, 1993). 인공신경망 기반의 인식기들은 관리도 시간 창(Time window)범위에서 기록되는 원 데이터를 가공하지 않고 인공신경망의 입력 벡터(Input Vector)로 사용하였다. 그러나 원 데이터를 그대로 사용하는 인식기는 관리도 시간 창의 범위가 넓어질수록 구조가 복잡해져 신경망의 학습 속도가 오래 걸리게 되고 다양한 형태의 CCP에 대한 신경망의 분류 성능을 저하시킨다(Hassan 등, 2003).

이러한 한계점을 극복하기 위해 최근 특징 추출(Feature Extraction)을 적용한 특징 기반 인식기(Feature based Recognizer)에 대한 연구가 활발히 진행되고 있다. 특징 기반의 인식기는 노이즈가 포함된 데이터에 강건하며 입력 벡터의 차원 축소로 인해 학습 시간이 줄어든다는 장점을 가진다(Wang 등, 2008).

특징 기반 인식기에 대한 연구의 핵심은 각 CCP의 특징을 가장 잘 표현하고 여러 CCP를 효과적으로 구별해 낼 수 있는 특징을 관리도에 기록된 원 데이터로부터 추출하는 것에 있다. 인식기의 성능을 향상 시킬 수 있는 특징을 추출하기 위해 Hassan 등(2003)은 통계적 특징(Statistical Feature) 추출법을 제안하였다. 정상, 상향/하향 추세, 상향/하향 시프트, 주기 패턴으로 정의된 6 종류의 CCP 데이터를 구분 짓기 위해 평균, 표준 편차, 왜도(Skewness), 래그(lag) 1,2 자기 상관(Average

표 1. 대표적 관리도 패턴(Control Chart Pattern)

패턴	설명	가능한 원인
시프트(Shift)	특정 시점에서의 갑작스러운 평균의 변화	원자재의 변경, 공정 세팅(Setting)의 변화, 새로운 작업자 투입 등
추세(Trend)	값의 지속적인 증가/감소	작업자의 피로, 장비의 노후화 등
주기(Cyclic)	특정 주기로 정점과 저점이 반복적으로 발생	주기적 작업자의 교체, 공정 장비 변동 등
규칙성(Systematic)	작은 값, 큰 값이 차례차례 반복적 발생	샘플링 하는 생산라인의 변화 등

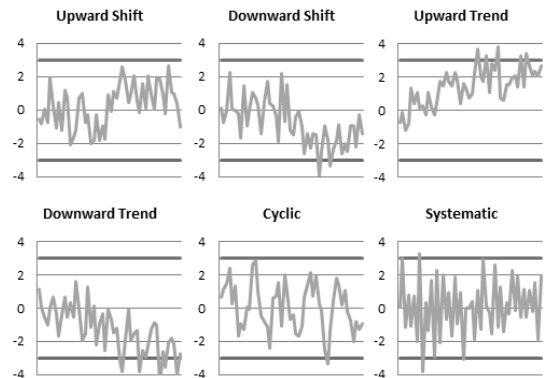


그림 1. 관리도 패턴의 예

Autocorrelation)의 평균, 제곱 합의 평균(Mean-Square Value), CUSUM 통계량을 특징으로 추출하였다. 추출된 특징은 MLP기반 인식기의 입력 벡터로 사용되었으며 그 인식기의 성능은 원 데이터를 입력 벡터로 사용하는 인식기의 성능보다 향상된 것을 보여준다. 이 밖에도 여러 연구들을 통해 다양한 형태의 통계적 특징이 정의되고 활용되었다(Gauri 등, 2007, 2009). 그러나 통계적 특징기반의 인식기는 새롭게 정의된 CCP를 구분하기 위하여 기존 특징에 새롭게 정의된 CCP의 특성을 반영할 수 있는 또 다른 특징을 추가해야만 인식기의 성능을 유지할 수 있기 때문에 비효율적이다.

Yang 등(2005)은 상관관계 계수(Correlation Coefficient)를 사용한 특징 추출법을 제안하였다. 6종류의 CCP로 구성된 샘플 패턴 데이터의 시점별 평균으로 참조 벡터(Reference Vector)를 생성하고 6종류의 참조 벡터와 CCP 데이터의 상관관계 계수를 구하여 특징으로 사용하였다. 참조 벡터는 각 CCP의 대표적 모습을 의미하고 상관관계 계수는 참조 벡터와의 유사도를 측정하는 도구로서의 역할을 한다. 같은 패턴을 가진 관리도 데이터라도 그 형태가 다양하다. 예를 들어 추세 패턴을 가진 관리도 데이터는 크고 작은 다양한 기울기를 가지며 시프트 패턴을 가진 데이터는 다양한 시프트 폭을 가진다. 따라서 참조 벡터를 생성함에 있어 평균으로 하나의 참조 벡터만을 생성하게 되면 다양한 형태를 가진 CCP의 특징을 제대로 반영할 수 없게 된다.

이에 본 논문은 다양한 형태를 가진 CCP의 특성을 효과적으로 반영하기 위하여 BDK(Bi-Directional Kohonen Network)를 사용하여 하나의 CCP당 여러 개의 참조 벡터(Reference Vector)를 생성하였다. 생성된 여러 참조 벡터와 CCP 데이터의 유클리드 거리(Euclidean Distance)로 유사도를 측정하였으며 모든 참조 벡터와의 거리를 각 CCP 특징으로 추출하였다. 인공신경망 기반의 인식기의 입력 벡터로 추출된 특징을 사용하여 학습하였고 원 데이터를 사용하여 학습하는 인공신경망 인식기와 예측 정확도 비교를 통해 제안 알고리즘의 성능을 평가하였다.

본 논문의 구성은 다음과 같다. 2장에서는 2.1장 제안한 알고리즘에서 사용될 BDK의 설명과 적용 방식, 2.2장 학습된 BDK에서 CCP 데이터의 특징 추출 방법, 2.3장 제안된 특징을 입력 벡터로 하는 인공신경망 기반 인식기의 구조 및 설계 방식을 제시한다. 3장에서는 실험 및 결과 분석을 통해 제안하는 알고리즘의 효율성을 보이고 4장에서는 결론을 기술한다.

2. BDK와 인공신경망을 사용한 관리도 패턴 인식

본 논문에서 제안하는 알고리즘은 그림 2와 같은 절차(Procedure)를 통해 진행된다. 학습 과정은 크게 특징 추출 단계와 CCP 인식 모델 생성 단계로 나뉜다. 특징 추출 단계는 CCP 원 데이터를 BDK로 학습하여 참조 벡터를 생성하는 과정을 거친 후 BDK를 통해 생성된 모든 참조 벡터와 학습 데이터와의 유클리드 거리를 기반으로 하는 특징을 생성하는 단계이다. 본 논문은 특징 추출 단계를 통해 생성된 특징을 BDK기반 거리 특징(Distance Feature)이라 정의한다.

CCP 인식 모델 생성 단계는 추출된 특징을 입력 벡터로 하는 인공신경망 기반 CCP 인식 모델을 생성하는 단계이다. 인공신경망 구조를 결정하고 생성된 거리 특징을 입력 벡터로 사용하며 학습 과정을 통해 인식 모델을 생성한다. 테스트 과정에서는 학습과정을 통해 생성된 CCP 인식 모델을 바탕으로 새로운 CCP 데이터에 대한 패턴 인식을 수행하여 알고리즘의 성능을 평가한다.

2.1 BDK(Bi-Directional Kohonen Network)

본 논문은 CCP의 각 패턴을 대표하는 참조 벡터를 생성하고 참조 벡터와 CCP 데이터의 유사도를 특징으로 추출한다. 참조 벡터는 각 CCP의 특성을 가장 잘 반영하는 대

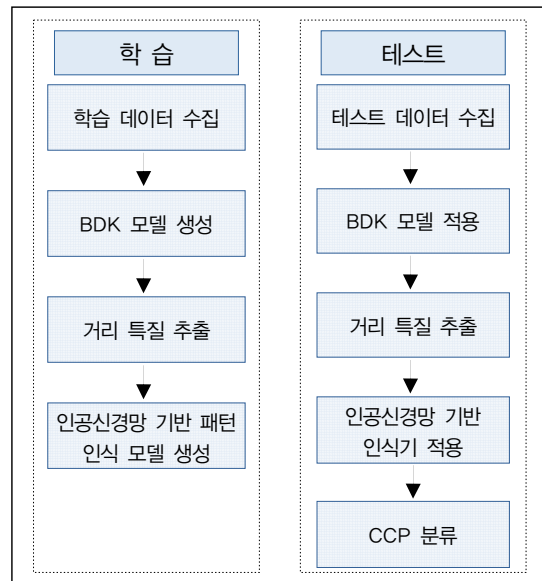


그림 2. 제안 알고리즘 절차(Procedure)

표 벡터로써 Yang 등(2005)은 각 패턴의 시점별 평균을 참조 벡터로 사용하였다. 서론에서 기술했듯이 CCP의 변화의 폭은 같은 패턴 내에서도 다양한 크기로 발생한다. 따라서 평균으로 이루어진 하나의 CCP 참조 벡터는 CCP의 특성을 온전히 반영하는데 한계가 있다. 본 논문은 CCP의 다양한 참조 벡터 생성을 위해 BDK 각 노드(Node)의 가중치 벡터(Weight Vector)가 데이터의 특징을 반영하는 대표하는 모습으로 학습되는 특성을 활용하였다.

BDK는 SKN(Supervised Kohonen Network)의 한 종류이다. 일반적인 Kohonen Network는 인공지능망의 자율학습(Unsupervised Learning)방식 중 하나로 학습과정에서 자율적으로 군집(Cluster)을 형성하고 그 과정에서 생성된 가중치 벡터들이 데이터의 특성을 고스란히 반영하여 데이터의 대표적 모습들로 학습되는 특징을 지닌다. 또한 다차원의 데이터를 낮은 차원(2차원 또는 3차원)의 맵 상으로 표현하므로 다차원 데이터의 시각화에 매우 유용한 알고리즘이다(Kohonen, 2001). 반면 CCP와 같이 클래스 정보가 명확하게 정의되고 같은 클래스 내에서 변화 폭이 심한 데이터의 학습 시 클래스 정보를 반영하지 못하기 때문에 가중치 벡터가 데이터의 대표적 특징을 제대로 반영하지 못하는 경우가 빈번히 발생한다. Kohonen (2001)에 의해 제안된 SKN는 Kohonen Network의 장점을 그대로 유지하면서 클래스 정보를 활용하지 못하는 단점을 보완하여 클래스 정보를 학습 과정에 활용하는 알고리즘이다. 하지만 SKN은 입력 변수 X와 출력변수 Y의 적절한 스케일링(Scaling)이 학습에 매우 중요한 영향을 미치며 X, Y의 맵을 공유하기 때문에 X, Y의 비율에 따라 모델이 입, 출력 변수 중 하나에 크게 지배(Dominate)되므로 잘못된 모델을 생성할 가능성을 가지는 단점이 있다 (Melssen 등, 2006). BDK는 위에서 설명한 Kohonen Network와 SKN의 단점들을 보완한 알고리즘이다. 입, 출력 변수가 서로 다른 맵을 통해 양방향(Bi-Direction) 학습을 하기 때문에 학습에 영향을 미치는 정도가 하나의 변수에 크게 지배되지 않으며 학습 시 유사도 측정 과정에서 스케일링을 수행하기 때문에 전처리 과정에서의 스케일링 없이 학습이 가능하다(Melssen 등, 2006). 따라서 클래스 내에서도 큰 폭의 변화를 가지는 데이터의 학습 시 별도의 스케일링이 필요치 않으며 클래스 정보와 입력 변수의 정보를 균형적으로 활용하기 때문에 데이터의 특징을 절절히 반영한 맵을 매우 효과적으로 생성한다.

BDK의 구조와 학습과정은 다음과 같다. BDK는 그림 3과 같이 2개의 2차원 맵을 가진 형태의 구조를 가진다.

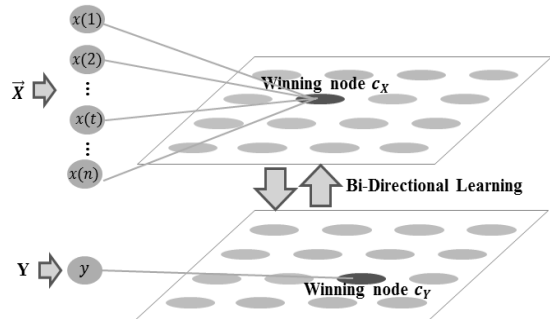


그림 3. 4×4 BDK 구조

각 맵은 (\vec{X}, Y) 한 쌍으로 구성된 데이터의 \vec{X} 를 입력 하는 X맵, Y를 입력 하는 Y맵으로 나뉜다. X맵과 Y맵은 여러 개의 노드로 구성되며 맵의 노드 수는 사용자가 정하는 맵의 구조에 의해 결정된다.

X맵의 입력 데이터를 식 (1)과 같이 벡터로 정의한다. 관리되는 n 개의 시점의 시간 창 범위에서 관리 된다. 따라서 CCP 데이터는 n 개의 시점으로 이루어져 있으며 시점 t 의 데이터 값은 $x(t)$ 이다.

$$\vec{X} = \begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(t) \\ \vdots \\ x(n) \end{bmatrix}, \quad t = 1, \dots, n \quad (1)$$

Y맵의 입력 데이터는 식 (2)와 같이 정의한다. CCP 데이터의 경우 Y는 정상, 상/하향 추세, 상/하향 시프트, 주기, 규칙성 패턴으로 구성된 총 7가지 패턴으로 정의 된다.

$$Y = y, \quad y = \{Set\ of\ 7\ patterns\} \quad (2)$$

X맵과 Y맵의 노드는 동일한 m 개의 노드를 가지며 i 노드는 식 (3), 식 (4)와 같은 가중치 벡터를 가진다.

$$\vec{W}_{Xmap_i} = \begin{bmatrix} w_{Xmap_{i1}} \\ w_{Xmap_{i2}} \\ \vdots \\ w_{Xmap_{iu}} \\ \vdots \\ w_{Xmap_{in}} \end{bmatrix}, \quad i = 1, \dots, m \quad (3)$$

$$W_{Ymap_i} = w_{Ymap_i} \quad (4)$$

BDK는 입력 데이터를 X맵과 Y맵의 특정 노드에 할당 후 특정 노드의 가중치 벡터와 이웃의 가중치 벡터를 갱신 하는 방식으로 학습된다. 데이터가 할당 되는 특정 노드를 승리 노드(Winning Node)라 하는데 승리 노드는 식 (5)~(8)과 같은 방식으로 결정된다.

$$S_{winnerX}(i) = (1 - \alpha(k))S(\vec{X}, \vec{W}_{Xmap_i}) + \alpha(k)S(Y, W_{Ymap_i}) \quad (5)$$

$$S_{winnerY}(i) = \alpha(k)S(\vec{X}, \vec{W}_{Xmap_i}) + (1 - \alpha(k))S(Y, W_{Ymap_i}) \quad (6)$$

$$c_X = \arg \min_i S_{winnerX}(i) \quad (7)$$

$$c_Y = \arg \min_i S_{winnerY}(i) \quad (8)$$

\vec{X} 와 X맵 노드의 가중치 벡터와의 유사도를 $s(\vec{X}, \vec{W}_{Xmap_i})$ 로 정의하고 Y와 Y맵 노드의 가중치 벡터와의 유사도를 $S(Y, W_{Ymap_i})$ 로 정의 한다. $s(\vec{X}, \vec{W}_{Xmap_i})$ 는 유클리디안 거리를 측정 도구로 사용하며 $S(Y, W_{Ymap_i})$ 는 분류 문제의 경우 Tanimoto 거리(Rogers 등, 1960)를 사용한다. 두 가지 유사도는 각 노드의 최대 유사도 값으로 나누어 줌으로써 최댓값이 1이 되도록 스케일링한다. 두 가지 유사도의 가중 조합(Weighted Combination) 형태로 표현되는 $S_{winnerX}(i)$, $S_{winnerY}(i)$ 가 가장 작은 값을 가지는 노드에 입력 데이터가 할당되며 이를 입력 데이터의 승리 노드(c_X , c_Y)로 결정한다. k 는 학습 반복(Iteration) 횟수를 의미하고 $\alpha(k)$ 는 유사도의 가중치를 의미한다. $\alpha(k)$ 는 0~1 사이 값을 가지며 학습 반복 초기에 큰 값으로 설정되었다가 학습이 진행될수록 선형적으로 감소하여 0.5로 수렴한다. 따라서 학습 초기에 X맵의 승리 노드를 결정하는데 $S(Y, W_{Ymap_i})$ 가 크게 영향을 미치며 반대로 Y맵의 승리 노드를 결정하는데 $s(\vec{X}, \vec{W}_{Xmap_i})$ 가 큰 영향을 미친다는 것을 알 수 있다. 이는 X맵에서 같은 클래스 정보를 가진 데이터끼리 가까운 노드로 강제적으로 모아주는 역할을 하며 또한 Y맵의 승리 노드 결정 방식을 통해 같은 추세 패턴이라도 다양한 기울기를 가지는 데이터와 같이 클래스 내에서도 다양한 형태를 가지는 데이터를 서로 다른 노드로 분산 시켜주는 역할을 한다.

k 번째의 입력 데이터의 승리 노드로 할당이 끝난 다음 식 (9)~(12)와 같은 방식으로 두 맵의 노드 i 의 $k+1$ 번째의 가중치 벡터를 갱신한다.

$$W_{Xmap_i}(k+1) = W_{Xmap_i}(k) + F_X l(k) h_{ci}(k) [\vec{X}(k) - W_{Xmap_i}(k)] \quad (9)$$

$$W_{Ymap_i}(k+1) = W_{Ymap_i}(k) + F_Y l(k) h_{ci}(k) [Y(k) - W_{Ymap_i}(k)] \quad (10)$$

$$F_X = 2 - S_{winnerX}(i) \quad (11)$$

$$F_Y = 2 - S_{winnerY}(i) \quad (12)$$

$l(k)$ 는 시간에 따라 감소하는 학습 율(Learning rate)에 관한 함수이다. F_X , F_Y 는 학습 율의 가중치 함수로써 잘못된 입력 데이터가 학습에 미치는 영향을 줄여주는 역할을 한다. 학습 율 가중치는 1~2사이 값을 가지며 입력 데이터와 가중치 벡터가 일치 할 경우 거리를 기반으로 하는 유사도는 0이 되고 그때의 학습 율 가중치는 최댓값인 2가 된다. 또한 완전히 불일치 할 경우 가중치는 1이 된다. $h_{ci}(k)$ 는 승리 노드 c 와 이웃하는 노드 i 의 갱신 비율을 결정하는 함수이다. 즉, 승리 노드의 가중치 벡터만을 갱신 하는 것이 아니라 이웃하는 노드까지 갱신하여 서로 이웃하는 노드의 가중치 벡터의 모습을 유사한 형태로 갱신하는 것을 의미한다. 모든 데이터의 승리 노드가 결정 되면 학습은 종료된다.

2.2 BDK기반 거리 특징(Distance Feature) 추출

BDK 학습 특성상 같은 클래스 정보를 가진 데이터가 서로 이웃하는 노드로 할당되고 같은 클래스 내에서 서로 다른 형태를 보이는 데이터는 분산 되어 이웃하지만 서로 다른 노드로 할당된다는 것을 알 수 있다. 또한 X맵의 가중치 벡터 갱신 과정을 살펴보면 각 노드의 가중치 벡터는 할당되는 데이터의 대표적인 모습으로 변화 되도록 학습된다는 것을 알 수 있다. CCP 데이터의 경우 이는 X맵의 모든 노드가 각 패턴을 대표하는 여러 개의 참조 벡터를 생성한다는 의미이며 따라서 BDK는 CCP 데이터 분석에 매우 효과적인 학습 방법이다. 본 논문은 이러한 BDK의 특성을 활용하여 BDK X맵의 가중치 벡터를 CCP의 참조 벡터로 활용한다.

CCP 데이터를 4×4 BDK를 통해 학습하여 얻어지는 X맵 노드의 참조 벡터 예는 그림 4와 같다. 학습 과정에서 생성되는 참조 벡터의 수는 BDK 맵의 노드 수에 따라서 결정된다. 따라서 다양한 참조 벡터 생성이 가능하고 CCP 패턴이 추가되더라도 맵의 노드 수 변화를 통해 참조 벡터의 수를 조절함으로써 변화에 유연(Flexible)하게 적용 가능한 장점을 가진다.

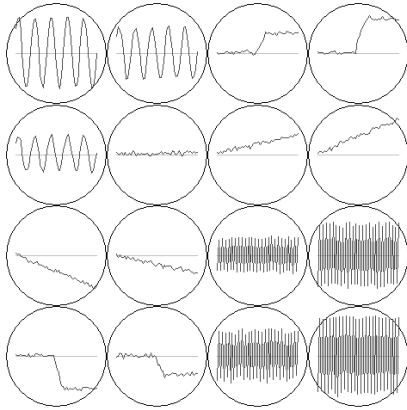


그림 4. 4×4 BDK X맵 가중치 벡터

BDK는 그 자체로 뛰어난 인식기의 성능을 보이지만 입력 벡터와 각 노드의 가중치 벡터와의 유클리디안 거리가 가장 가까운 특정 승리 노드의 클래스로 예측하므로 유클리디안 거리의 특성상 작은 변화를 보이는 시프트, 추세 패턴에 대해 오 분류(Misclassification)의 가능성이 있다. 따라서 본 논문은 가장 가까운 승리 노드의 가중치 벡터와의 거리만이 아닌 식 (13)과 같이 X맵 상 모든 노드의 가중치 벡터와 데이터 사이의 거리를 데이터의 새로운 특징으로 추출하여 인공지능망 기반의 인식기의 입력 벡터로 사용하는 방식을 제안한다. 즉, 새로운 데이터가 주어졌을 때 식 (14)에서 표현되는 벡터 \vec{D} 데이터의 새로운 특징이 된다. \vec{D} 는 d_i 의 최댓값으로 나누어 0과 1사이 값으로 스케일링하며 이를 인공지능망 인식기의 입력 벡터로 사용한다.

$$d_i = \| \vec{X} - \vec{W}_{Xmap} \| \quad (13)$$

$$\vec{D} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_i \\ \vdots \\ d_m \end{bmatrix} \quad (14)$$

BDK 기반의 거리 특징 \vec{D} 는 BDK X맵 각 노드의 가중치 벡터로 특정 지어지는 참조 벡터와 원 데이터가 얼마나 닮아있는지에 대한 유사도 측정치를 의미한다.

2.3 인공지능망(ANN)기반의 CCP 인식 모델

인공지능망은 입력 변수와 출력 변수 간의 비선형 함수를 추정 하는 것에 매우 강건(Robust)하다. 특히 실제 공

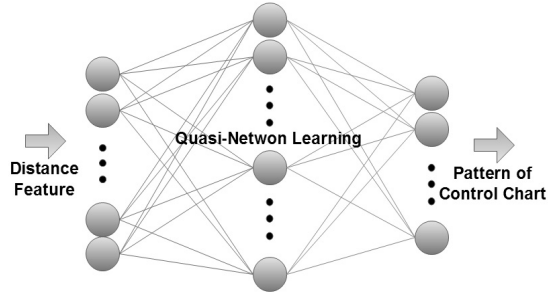


그림 5. 인공지능망 구조

정의 다양하고 복잡한 데이터를 해석하고 분류하는 매우 효과적이다(Mitchell, 1997). 따라서 본 논문은 추출된 특징을 입력하여 CCP 인식 모델을 생성하는데 인공지능망을 활용하였다. 그림 5에서 보듯이 인공지능망의 구조는 노드와 층(Layer), 각 노드를 연결하는 가중치로 이루어져 있다. 각 층은 입력 층, 은닉 층, 출력 층으로 구성되고 각 층은 여러 개의 노드를 가진다.

입력 층의 노드 수는 입력되는 특징의 수와 같고 출력 층의 노드 수는 학습하는 CCP 패턴의 수에 따라 결정되며 은닉 층의 개수와 은닉 층의 노드 수는 사용자가 설계한 구조에 의해 결정된다. 인공지능망 학습 과정에서 가중치의 조정을 위해 다양한 활성화 함수가 사용된다. Sigmoid 함수가 대표적인 활성화 함수이다. 인공지능망 학습은 실제 출력 데이터와 모델에서 생성된 출력 데이터 사이의 오차를 최소화 하도록 가중치를 갱신하는 역 전파(Backpropagation) 알고리즘이 주로 사용되는데 역 전파 알고리즘에는 기울기 하강 법(Gradient Descent), Levenberg-Marquardt, Quasi-Newton 등이 있다.

본 논문은 입력 벡터로 BDK기반의 거리 특징을 사용하며 1개의 은닉 층과 7개의 출력 층을 가진 인공지능망 구조를 생성하였다. 역 전파 알고리즘 중 비선형 최적화 문제 풀이에 매우 효과적인 Quasi-Newton(Setiono 등, 1995)을 사용하였으며 활성화 함수는 시그모이드(Sigmoid) 함수를 사용하였다. 완성된 구조의 인공지능망에 추출된 특징을 학습하여 CCP 인식 모델을 생성하였다.

3. 실험 및 결과 분석

3.1 CCP 데이터 생성 시뮬레이션

제안 인식기의 학습에 필요한 실제 공정의 다양한 이상 CCP 데이터를 다량으로 수집하기 위해서는 많은 시간과 자원을 소모하게 된다. 또한 일반적으로 기업은 정보

보안 유지 차원에서 데이터의 반출을 허용하지 않는다. 따라서 본 논문은 실제 공정 CCP 데이터의 특징을 모방한 시뮬레이션을 통해 실험 CCP 데이터를 생성하였다. 제안 인식기의 성능 평가를 위하여 본 논문은 Swift(1987)가 처음 제안하였고 이후 많은 관리도 패턴 인식 연구에서 일반적으로 사용되고 있는 CCP 데이터 생성 시뮬레이션 기법을 사용하였다. Swift(1987)가 제안한 CCP 데이터의 생성 함수는 식 (15)~(20)과 같다.

정상 데이터 :

$$x(t) = n(t) \sim N(0,1) \quad (15)$$

$x(t)$: t 시점의 데이터 값

상향/하향 시프트 패턴 데이터 :

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

s : 시프트 크기

t_h : 시프트 패턴 발생 시점

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

상향/하향 추세 패턴 데이터 :

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

d : 추세 패턴 기울기

주기 패턴 데이터 :

$$x(t) = n(t) \pm a \sin\left(\frac{2\pi t}{T}\right) \quad (19)$$

a : 주기 패턴의 진폭

T : 한 주기의 길이

규칙성 패턴 데이터 :

$$x(t) = n(t) \pm (-1)^t sy \quad (20)$$

sy : 규칙성 패턴의 진폭

데이터 생성을 위한 함수의 모수들은 표 2와 같은 값으로 설정하였다. 관리도 패턴 인식은 관리도가 이상 패턴으로 진행되는 현상을 관리 한계선을 벗어나기 전에 미리 감지하여 공정 상태에 대한 관리자의 빠른 의사 결정을 돕는 것을 목적으로 연구되고 있다. 따라서 시프트의 크기, 추세 패턴의 기울기, 주기 패턴의 진폭, 규칙성 패턴의

표 2. CCP 데이터 생성 함수 모수

모수	설정 값
a	1~3
d	0.02~0.06
s	1~3
sy	1~3
T	10
t_h	25~30

진폭의 값은 이상 패턴의 데이터가 관리도의 관리 한계선으로 설정된 3σ 이하로 발생하도록 설정하였다. 또한 정상 데이터가 가지는 1σ 의 화이트 노이즈(White Noise)로 인한 잘못된 패턴 생성을 방지하기 위해 최소 1σ 이상으로 설정하였다. 따라서 a , d , s , sy 값은 이상 패턴 데이터가 $1\sigma \sim 3\sigma$ 의 범위 안의 랜덤(Random)한 변화 폭을 가진 데이터로 생성되도록 설정하였다.

시간 창은 관리도에 기록되는 공정 데이터를 한 번에 모니터링 할 수 있는 시간의 범위를 의미한다. CCP 데이터는 n 개의 시점의 시간 창 범위에서 관리되는데 본 논문은 n 을 50으로 가정하였다. 따라서 CCP 데이터는 50개의 시점을 가진 데이터로 생성된다. 또한 주기 패턴의 한 주기의 길이는 10으로 가정하였고 시프트는 총 50개의 시점 중 25~30 시점 사이에서 발생한다고 가정하였다. 정상 데이터 및 상향 시프트/하향 시프트, 상향 추세/하향 추세, 주기, 규칙성 패턴의 6가지 이상 CCP 데이터 600개씩 총 4200개의 데이터를 생성하여 실험하였다.

3.2 실험 설계

본 논문의 실험은 BDK 학습을 통한 특징 추출, 인공신경망 기반 인식기 생성 및 성능 평가의 순서로 진행하였다. 정상과 이상 패턴에 대한 예측 정확도를 성능 평가의 척도로 활용하였으며 알고리즘의 정확한 성능 평가를 위해 10-집단 교차 검증(10-Fold Cross Validation)을 사용하였다. 10-집단 교차 검증은 4200개로 구성된 데이터셋을 10등분하여 10개의 군으로 형성한 후 하나의 군을 테스트 군으로 사용하고 나머지를 학습 군으로 사용하는 실험을 총 10번에 걸쳐 진행하여 모든 군이 한번 씩 테스트 군으로 사용되는 성능 평가 방법으로 특이한 테스트 군에 의한 영향을 최소화시키기 위해 사용되는 알고리즘 검증방법이다(Mitchell, 1997).

BDK 맵의 크기는 7개의 패턴을 대표하는 참조 벡터의 생성 개수를 의미한다. 본 논문은 4×4 , 4×5 , 5×5 , $5 \times$

6으로 구성된 맵을 생성하여 16, 20, 25, 30개의 BDK 기반 거리 특질을 생성하였다. BDK 학습의 구체적인 사양은 표 3과 같다. $\alpha(k)$ 값은 초기 0.75를 시작으로 학습이 진행 될수록 선형적으로 감소되어 학습 종료 시 0.5로 수렴되도록 설정하였으며 $h_{ci}(k)$ 는 초기 노드간의 거리의 2/3가 되는 값에서 노드간의 거리의 -2/3로 감소하도록 설정하였다. $h_{ci}(k)$ 값이 1이하로 감소되면 승리 노드의 가중치 벡터만 갱신한다. 승리 노드의 가중치 벡터와 학습 데이터의 평균거리가 50이하로 수렴(Convergence)되거나 최대 학습 반복 횟수가 200회에 도달했을 때 학습을 종료시켰다.

학습이 종료된 BDK X맵으로부터 추출된 거리 특질은 인공신경망 기반 인식기의 입력 벡터로 사용된다. 인공신경망 인식기 학습 과정의 세부 사항은 표 4와 같다. 실험 시 학습 군 중 1개의 집단을 검증(Validation) 군으로 활용하였다. 은닉 층은 1개로 고정 시켰으며 그 노드의 수는 반복 실험을 하여 가장 예측 정확도가 높은 모델의 은닉 층의 수로 결정 하였다. 또한 과적합(Overfitting)의 방지를 위해 학습 반복을 늘려가면서 검증 군에 대한 예측 정확도가 낮아지는 시점에서 학습을 종료하였다. 또한 최대 학습 반복 횟수를 200회로 제한하였다. BDK 학습을 통해 추출된 특질을 기반으로 하는 인공신경망 인식기와 원 데이터를 입력 벡터로 하는 인공신경망 기반 인식기의 패턴 예측 정확도 비교를 통해 제안 알고리즘의 성능을

평가하였다.

3.3 실험 결과 분석

10-집단 교차 검증을 수행하여 도출된 테스트 CCP 데이터에 대한 성능 비교 평가는 표 5와 같다. 원 데이터를 입력 벡터로 하는 인공신경망 인식기와 4종류의 특질을 입력 벡터로 하는 인공신경망 인식기의 특질 수, 예측 정확도, 신경망 구조에 대한 비교 분석을 수행하였다. 패턴 예측 정확도 시험결과를 살펴보면 제안 거리 특질 기반 인식기의 성능이 원 데이터 기반 인식기에 비해 향상되었다는 것을 알 수 있다. BDK 맵의 크기에 따라 서로 다른 개수의 특질이 추출되었는데 특질 수에 따른 예측 정확도는 큰 차이를 보이지 않는다. 따라서 인공신경망의 구조에 따른 학습 시간과 인식기의 효율성을 고려해 볼 때 4×4 BDK로부터 추출된 16개의 특질을 기반으로 하는 인식기가 가장 효율적이라는 것을 알 수 있다.

원 데이터 기반의 인식기와 4×4 BDK 거리 특질 기반 인식기 성능 차이에 대한 정확한 검증을 위해 같은 샘플을 대상으로 하는 두 가지 처리간의 평균에 대한 검증을 위해 주로 사용되는 쌍대 t-검정(Paired t-test)을 실시하였다(Walpole 등, 2007). 귀무가설(H_0)은 ‘두 인식기의 예측 정확도 평균이 같다’이며 대립가설(H_1)은 ‘제안 인식기의 예측 정확도 평균이 높다’이다. 쌍대 t-검정 결과는 표 6과 같다. 검정 결과 P-Value의 값이 0.01 이하이므로 귀무가설을 기각 한다. 따라서 제안 거리 특질 기반 인식기의 성능이 원 데이터기반의 인식기의 성능과 비교해 우수하다는 것을 확인하였다.

표 7, 8은 원 데이터의 인식기와 4×4 BDK 거리 특질 기반의 인식기의 Confusion Matrix이다. 총 10번의 실험을 통해 생성되는 테스트 데이터에 대한 예측치를 종합하여 하나의 Confusion Matrix를 생성하였다. 표 7, 8을 통해 확인할 수 있듯이 제안 거리 특질 기반 인식기는 원 데이터 기반 인식기에 비해 정상 데이터에 대한 예측 성능

표 3. BDK 학습 사양(Specification)

사양	설정 값
$\alpha(k)$	0.75 ~ 0.5
$h_{ci}(k)$	노드간 거리의 2/3 ~ 노드간의 거리의 -2/3
격자(Grid)	사각형
최대 반복 횟수	200

표 4. 인공신경망 학습 사양

사양	세부 사항
역 전파 알고리즘	Quasi-Newton
입력 층 노드 수	생성된 거리 특질의 수
은닉 층 노드 수	테스트 예측 정확도가 가장 높을 때의 노드 수
출력 층 노드 수	7(패턴의 개수)
활성화 함수	시그모이드 함수
학습 종료 기준	검증 군의 예측 정확도가 줄어드는 시점/ 최대 학습 반복횟수 도달
최대 반복 횟수	200

표 5. 인식기 예측 정확도 비교

	특질 수	인공신경망 구조	예측 정확도	표준 편차
원 데이터	-	50×7×7	91.93	1.61
4×4 BDK	16	16×7×7	95.95*	1.20
4×5 BDK	20	20×9×7	95.62	1.01
5×5 BDK	25	25×12×7	95.74	0.99
5×6 BDK	30	30×14×7	95.88	0.96

표 6. 쌍대 t-검정($\alpha=0.01$)

가설	t-통계량	t_α	P-Value	검정 결과
$H_0 : \mu_{acc\ of\ bdk4 \times 4} - \mu_{acc\ of\ raw} = 0$	8.42	0.03	0.00	H_0 기각
$H_1 : \mu_{acc\ of\ bdk4 \times 4} - \mu_{acc\ of\ raw} > 0$				

표 7. 원 데이터 기반 인식기 Confusion Matrix

	정상	상향 시프트	하향 시프트	상향 추세	하향 추세	주기	규칙성
정상	93.19	1.19	1.19	1.36	1.19	1.19	0.68
상향 시프트	0.69	91.02	0.00	8.12	0.00	0.00	0.17
하향 시프트	1.48	0.00	89.16	0.00	9.20	0.16	0.00
상향 추세	2.60	11.36	0.00	86.04	0.00	0.00	0.00
하향 추세	2.64	0.00	9.59	0.00	87.60	0.17	0.00
주기	2.15	0.00	0.17	0.17	0.00	97.35	0.17
규칙성	0.33	0.00	0.00	0.00	0.00	0.33	99.33

표 8. 4×4 BDK 거리 특질 기반 인식기 Confusion Matrix

	정상	상향 시프트	하향 시프트	상향 추세	하향 추세	주기	규칙성
정상	98.17	0.17	0.17	0.67	0.33	0.50	0.00
상향 시프트	0.50	92.67	0.00	6.83	0.00	0.00	0.00
하향 시프트	0.00	0.00	94.17	0.00	5.83	0.00	0.00
상향 추세	0.50	6.17	0.00	93.33	0.00	0.00	0.00
하향 추세	0.00	0.00	6.17	0.00	93.67	0.00	0.17
주기	0.17	0.00	0.00	0.00	0.00	99.83	0.00
규칙성	0.00	0.17	0.00	0.00	0.00	0.00	99.83

이 뛰어나며 제 1 종 오류(α)와 제 2종 오류(β) 또한 상당히 감소된 것을 확인하였다. 이상 패턴에 대한 성능을 살펴보면 원 데이터 기반 인식기는 시프트 패턴과 추세 패턴에 대하여 오 분류 확률이 매우 높게 나타났다. 이러한 결과는 원 데이터 기반 인식기가 시프트와 추세 패턴의 다양한 변화를 학습과정에 제대로 반영하지 못한다는 것을 의미한다. 반면 제안 거리 특질 기반 인식기는 원 데이터 기반 인식기에 비해 시프트와 추세 패턴에 대한 오 분류 확률이 상당히 낮게 나타났다. 이는 BDK 학습을 통해 생성되는 참조 벡터가 시프트와 추세 패턴의 특징을 적절히 반영하는 형태로 생성되며 이를 통해 추출된 거리 특질을 인식기의 입력 벡터로 사용하면 다양한 변화를 보이는 시프트와 추세 패턴의 분류에 매우 효과적이라는 것을 의미한다.

4. 결 론

본 논문은 제품의 품질 향상과 유지를 위해 중요한 이슈인 관리도 패턴 인식을 위하여 BDK 기반의 거리 특질 추출과 거리 특질을 입력 벡터로 하는 인공신경망 인식기를 제안하였다. BDK 학습을 통해 패턴을 대표하는 다양한 참조 벡터를 생성하였으며 원 데이터와 참조 벡터와의 유클리디안 거리를 특질로 추출하였다. 원 데이터를 기반으로 하는 인공신경망 인식기와 제안 거리 특질을 기반으로 하는 인공신경망 인식기의 예측 정확도 비교 평가를 위하여 실험을 수행하였다. 실험 결과를 통해 확인할 수 있듯이 제안된 알고리즘은 인공신경망 인식기의 입력 벡터의 차원을 축소하여 인공신경망 구조를 간결하게 하였으며 적절한 특질 추출로 인식기의 패턴 예측 정확도를 향상시켰다.

본 논문에서 제안하는 알고리즘의 성능은 BDK 학습에 많은 영향을 받는다. BDK의 학습을 통해 CCP 데이터의 적절한 참조 벡터를 생성하기 위해서는 학습 과정에서 이상 데이터의 특징을 충분히 반영해야 하며 이를 위해 많은 이상 데이터가 요구된다. 실제 공정에서 이상 패턴 데이터는 정상 데이터에 비해 매우 낮은 비율로 생성된다. 따라서 생성 비율의 차이가 심한 데이터의 적절한 샘플링을 적용시켜 BDK 학습에 효과적인 데이터 셋을 생성하는 것이 추후 연구과제로 요구된다. 또한 BDK, 인공신경망의 구조와 모수들을 적절한 선택 방법이 필요하며 과적합을 방지하기 위한 다양한 학습 방법에 대한 연구가 필요하다.

감사의 글

이 논문 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2011-0002761).

본 과제는 정보통신산업진흥원의 SW공학 요소기술 개발과 전문인력 양성사업의 결과물임을 밝힙니다.

이 논문은 2011년도 2단계 두뇌한국(BK)21사업에 의하여 지원되었음.

참고 문헌

- Gauri, S. K. and Chakraborty, S. (2007), "A study on the various features for effective control chart pattern recognition", *Int J Adv Manuf Technol*, Vol. 34, pp. 385-398.
- Gauri, S. K. and Chakraborty, S. (2009), "Recognition of control chart patterns using improved selection of features", *Computer & Industrial Engineering*, Vol. 56, pp. 1577-1588.
- Hwang, H. B. and Hubele, N. F. (1993), "Back-propagation pattern recognizers for \bar{x} control charts: methodology and performance", *Computer & Industrial Engineering*, Vol. 24, pp. 219-235.
- Ho, E. S. and Chang, S. I. (1999), "An integrated neural network approach for simultaneous monitoring of process mean and variance shifts a comparative study", *International Journal of Production Research*, Vol. 37, No. 8, pp. 1881-1901.
- Hassan, A., Shariff Nabi Baksh, M., Shaharoun, A. M. and Jamaluddin, H. (2003), "Improved SPC chart pattern recognition using statistical features", *International Journal of Production Research*, Vol. 41, No. 7, pp. 1587-1603.
- Kohonen T. (2001), *Self-organizing maps*, 3th Edition, Springer.
- Mitchell, T. M. (1997), *Machine Learning*, McGraw-Hill Science.
- Montgomery, D. C. (2005), *Introduction to Statistical Quality Control*, 5th Edition, John Wiley & Sons, Inc.
- Melssen, W., Wehrens R. and Buydens, L. (2006), "Supervised Kohonen networks for classification problems", *Chemometrics and Intelligent Laboratory System*, Vol. 83, pp. 99-113.
- Pham, D. T. and Oztemel, E. (1994), "Control chart pattern recognition using learning vector quantization networks", *International Journal of Production Research*, Vol. 32, No. 3, pp. 721-729.
- Perry, M. and Pignatiello, J. (2007), "A Review of Artificial Neural Network Applications in Control Chart Pattern Recognition".
- Swift, J. A. (1987), *Development of a knowledge based expert system for control chart pattern recognition and analysis*, Ph.D Thesis, Oklahoma State University.
- Setiono, R. and Hui L. C. K. (1995), "Use of a Quasi-Newton Method on a Feedforward Neural Network Construction Algorithm", *IEEE Transactions on Neural Networks*, Vol. 6, No. 1. pp. 273-277.
- Western Electric Company (1958), *Statistical Quality Control Handbook*, Western Electric, Indiana.
- Walpole, R. E., Myers, R. H., Myers, S. L. and Ye, K. (2007), *Probability & Statistics for Engineers & Scientists*, 8th Edition, Pearson Education International.
- Wang, C. H., Guo, R. S. Chiang, M. H. and Wong J. Y. (2008), "Decision tree based control chart pattern recognition", *International Journal of Production Research*, Vol. 46, No. 17, pp. 4889-4901.
- Yang, J. H. and Yang, M. S. (2005), "A control chart pattern recognition system using a statistical correlation coefficient method", *Computer & Industrial Engineering*, Vol. 48, pp.205-221.

**윤재준** (junyeel7@korea.ac.kr)

2010 고려대학교 산업시스템정보공학과 학사
2010~현재 고려대학교 산업경영공학과 석사

관심분야 : Advanced Process Control, Data Mining, Control Chart Pattern Recognition

**박정술** (dumm97@korea.ac.kr)

2003 고려대학교 산업시스템정보공학과 학사
2006 고려대학교 산업시스템정보공학과 석사
2005~2006 삼성경제연구소 6시그마 Research Analyst
2006~2007 고등기술연구원 로봇생산기술센터 연구원
2008~현재 고려대학교 정보경영공학전문대학원 박사과정

관심분야 : Advanced Process Control, Data Mining

**김준석** (blihs@korea.ac.kr)

2004 고려대학교 산업시스템정보공학과 학사
2006 고려대학교 산업시스템정보공학과 석사
2007~현재 고려대학교 정보경영공학전문대학원 박사과정

관심분야 : Advanced Process Control, Data Mining, Control Chart Pattern Recognition

**백준걸** (jungeol@korea.ac.kr)

1993 고려대학교 산업공학과 학사
1995 고려대학교 산업공학과 석사
2001 고려대학교 산업공학과 박사
2001~2002 고려대학교 정보통신기술연구소 연구조교수
2002~2007 인덕대학교 산업시스템경영학과 조교수
2007~2008 광운대학교 경영학부 조교수
2008~현재 고려대학교 산업경영공학과 부교수

관심분야 : Advanced Process Control, Intelligent Fault Detection, Data Mining Applications