

불명확한 공정정보 처리를 위한 퍼지-통계적 관리도의 개발

김경환*, 하성도*

Development of Fuzzy-Statistical Control Chart for Processing Uncertain Process Information

Kyunghwan Kim*, Sungdo Ha*

ABSTRACT

Process information is known to have the continuous distribution in many manufacturing processes. Generalized p-chart has been developed for controlling processes by classifying the information characteristics into several groups. But it is improper to describe continuous processes with the classified process information, which is based on the classical set concept. Fuzzy control chart has been developed for the control of linguistic data, but it is also based on the dichotomous notion of classical set theory.

In this paper, fuzzy sampling method is studied in order to process the uncertain data properly. The method is incorporated with the fuzzy control chart. Statistical characteristics of the fuzzy representative value are utilized to device the fuzzy-statistical control chart. The fuzzy-statistical control chart is compared with the generalized p-chart and both the sensitivity to the process information distribution change and robustness against the noise on the process information of the fuzzy-statistical control chart are shown to be superior.

Key Words : SPC(통계적 공정관리), Fuzzy Control Chart(퍼지 공정 관리도), Fuzzy Representative Value(퍼지 대표값), Fuzzy Sampling(퍼지 추출법)

1. 서 론

생산 공정의 품질 정보는 다양한 특성을 가지고 존재하며 이들 정보를 관리하기 위한 대표적인 기법은 통계적 공정관리 (SPC : Statistical Process Control)방법이다.⁽¹⁾ 그러나 품질정보는 항상 이러한 통계적인 기법만을

가지고 표현하기에는 어려운 특성을 가지고 있다. 가장 대표적인 특성으로는 품질정보는 연속적이며 또한 수량적으로 표현하기 힘든 주관적인 특성이 많다는 것이다. 따라서 이들 품질정보를 고전 집합 논리를 바탕으로 하는 통계적인 기법으로 표현하는 경우 이분적인 개념에 의한 정보의 왜곡이 발생하게 되어 이러한 정보의 손실을 줄일

* 한국과학기술연구원 CAD/CAM 연구센터

수 있는 기법의 개발이 필요하다.⁽²⁾

퍼지 이론은 연속적이며 주관적인 정보를 처리하기 위해 개발되었으며, 그 특성상 품질 특성을 표현하기에 적합한 많은 장점이 있다. 그러나 기존의 퍼지관리도에 대한 연구에서는 품질 정보의 퍼지성에 대한 연구의 부족으로 퍼지 기법의 특성이 충분히 적용되지 않았으며, 여전히 이분적인 개념을 벗어나지 못하는 한계가 있다.⁽³⁾

본 연구에서는 고전 집합 논리를 바탕으로 하는 이분적인 데이터 분류방식에서 탈피하여 연속적인 데이터 분류를 위한 퍼지 추출(Fuzzy Sampling)분류기법을 개발하여 측정 데이터로부터 퍼지 대표값을 구하고, 퍼지기법을 적용하여 얻어지는 퍼지대표값의 특성은 퍼지변환 이전의 데이터 특성과 동일한 점에 착안하여, 퍼지대표값에 통계적인 기법을 적용시킨 Fuzzy-Statistical Control Chart를 개발하였다. 정규분포의 데이터로부터 얻어진 퍼지 대표값에 콜모고로프 검정을 적용하여 그 정규분포성 여부를 파악하고, 관리한계선 설정에 정규분포의 특성을 이용하였다. 그리고 개발된 품질 정보처리 기법의 성능을 분석하기 위해서, 측정 데이터에 포함되는 노이즈의 증가가 관리도의 형태와 공정변화(Process Shift)에 미치는 영향을 분석하고, 각 품질 분류에 속하는 측정 데이터의 분포에 따른 민감성을 비교 분석하였다.

2. 통계적 개념과 퍼지 개념에서의 품질정보 표현

2.1 통계적 개념의 품질정보 표현

고전 논리를 바탕으로 하고있는 확률이론은 품질 분류의 정의에서 집합 개념을 바탕으로 하는 원소의 존재성 여부를 그 근거로 하고 있다.⁽⁴⁾ 따라서 통계적인 관리도에서는 각 품질 분류 집합을 설정한 후 그 집합에 속하는 원소의 개수를 이용하여 확률적으로 데이터를 분석한다.

통계적 기법의 관리도 중의 하나인 Generalized p-chart는 다양한 품질분류의 관리를 위하여 P-chart를 확장한 것으로서 그 내용은 다음과 같다.

세 등급으로 제품 품질을 분류하는 공정의 경우 공정을 감시하기 위해서 각 관리주기마다 독립 무작위 표본을 추출하면 관리품질 수치는 다음과 같이 계산된다.

$$Y_i^2 = \sum_{j=1}^3 \frac{(X_{ij} - n_i \pi_j)^2}{n_i \pi_j} \quad (2.1)$$

여기에서

n_i : 주기 i 의 표본의 크기

X_{ij} : 관리 주기 i 에서 j 분류의 표본수

π_j : 표준공정비율

식 (2.1)은 $\{(관측치-기대치)/기대치\}$ 의 형태를 띠고 있으므로 공정이 정상 제어상태에 있을 때에 Y_i^2 는 자유도 2의 카이제곱 분포 χ^2 를 이룬다⁽²⁾. 따라서 카이제곱 분포표를 사용하여 관리 상한선을 결정한다. 그러나 이러한 Generalized p-chart에서도 각각 분류에 속하는 데이터는 이분적인 개념으로 다루어지므로 정보의 손실이 발생하게 된다.

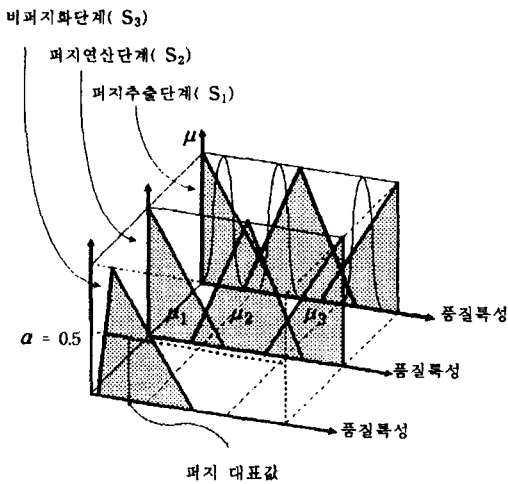
2.2 퍼지 개념의 품질정보 표현

품질정보특성은 수치적인 표현보다 직관적 언어로 표현할 때 더 적합한 경우가 있다. 그러나 언어가 가진 주관적인 특성으로 인해 기존의 이론으로 구현하기에 어려움이 있다. 이분적인 집합개념을 사용하는 통계이론에 비해 퍼지이론의 큰 특징중의 하나는 이러한 경계가 불명확한 집합의 표현에 언어변수를 사용할 수 있다는 것이다. 즉 공정에서 발생하는 품질 데이터는 거의 모두가 연속성을 내재하고 있으며, 퍼지 이론은 이러한 연속성으로 인해 발생하는 불명확한 정보를 표현하기 위한 체계적인 기법이다. 따라서 품질 관리와 같이 전문가의 추론, 인식, 감정 등이 개입될 여지가 많은 영역에서의 효율을 높이기 위해서 퍼지논리의 도입이 필요하며, 이는 시스템 인자의 퍼지화로서 구현된다. 예를 들면 실제 공정 상에서 발생하는 많은 품질특성 중에 “어느 정도 관리상태에 있는가”, “어느 정도 비관리 상태에 있는가”와 같이 잘 정의되거나 결정되지 않는 인간의 사고 과정이나, 수치적인 표현보다는 기술자의 언어적인 표현이 더욱 적합한 경우에 퍼지이론은 품질 정보를 잘 표현할 수 있는 표현력을 가지고 있다. 그러나 시스템과 공정품질정보가 어떠한 퍼지특성을 가지고 있는가에 대한 연구가 선행되어야 앞에서 언급한 퍼지특성을 이용할 수 있다.

3. 품질 정보의 퍼지화

3.1 기존 퍼지관리도⁽⁴⁾

기존 퍼지관리도에서는 공정데이터에 퍼지연산을 수행하여 각 품질 정보 비율에 대한 퍼지 소속함수를 생성하고(Fig.1의 S_2), 이 소속함수의 비퍼지화를 통해서 관리도에 나타내기 위한 퍼지 대표값을 산출한다(Fig.1의 S_3).



$S_1 \rightarrow S_2 \rightarrow S_3$: 퍼지추출기법을 사용한 퍼지데이터 처리과정
 $S_2 \rightarrow S_3$: 이분적 품질정보를 가지는 기존의 퍼지 데이터 처리과정

Fig. 1 Concept of fuzzy processing of quality characteristics

즉, 각 공정관리 주기에 속하는 제품의 비율(p)에 대해서 품질의 정도를 나타내는 언어 변수에 해당하는 삼각소속함수 \tilde{T} 를 적용하여, 식 3.1과 같이 연산을 통해서 각 주기 공정의 품질을 대표하는 단일 퍼지소속함수 ($\tilde{T}\tilde{P}$)를 생성한다. 여기서 제품의 비율은 고전 집합 개념을 바탕으로 구한 비율을 사용한다.

$$\tilde{T}\tilde{P} = [\tilde{T}_1 \cdots \tilde{T}_c] \begin{bmatrix} P_1 \\ \vdots \\ P_c \end{bmatrix} = \tilde{T}_1 P_1 + \cdots + \tilde{T}_c P_c \quad (3.1)$$

그리고 생성된 단일 퍼지소속함수를 비퍼지화하여 관리도에 표현하기 위한 퍼지 대표값을 구한다. 이때 사용되는 비퍼지화 기법은 α -level의 중심값을 퍼지 대표값으로 설정하는 α -level Fuzzy Midrange로서 그 계산은 다음과 같다.

$$\text{퍼지대표값} = \left[\frac{1-\alpha}{2} \quad \alpha \quad \frac{1-\alpha}{2} \right] \begin{bmatrix} t_{11} \cdots t_{c1} \\ t_{12} \cdots t_{c2} \\ t_{13} \cdots t_{c3} \end{bmatrix} \begin{bmatrix} P_1 \\ \vdots \\ P_c \end{bmatrix} \quad (3.2)$$

여기서 c : 품질 분류의 개수

p : 각 주기에서 품질 비율
 이다. Fig.1의 S_3 는 $\alpha = 0.5$ 인 경우의 α -level Fuzzy Midrange를 표현한 것이다.

퍼지관리도의 중심선은 표준품질비율을 식 (3.2)에 적용한 퍼지 대표값으로 결정하며, 관리 한계선은 퍼지 대표값의 계산에 이용된 단일퍼지함수의 Mean Deviation을 사용하여 설정한다. 여기서 Mean Deviation은 삼각 퍼지함수의 넓이와 같다.

$$\begin{aligned} MLCL &= \max(0, CL - k\delta(T^*)) \\ MUCL &= \min(1, CL + k\delta(T^*)) \end{aligned} \quad (3.3)$$

여기서 k: 설정된 유의수준에 대해 Simulation과 전문가의 경험에 의해 선택되는 상수

$\delta(T^*)$: Mean Deviation

MLCL: Membership Lower Control Limit

MUCL: Membership Upper Control Limit

이다. k값 설정을 위한 경험적인 데이터가 부족한 상태에서 이의 결정에는 많은 모의 실험이 필요하다.

3.2 퍼지 추출법 (Fuzzy Sampling Method)

앞서 서론에서 간략하게 언급하였듯이, 기존의 품질 정보의 퍼지 구현 과정에서 가장 큰 문제점은 정보 자체가 내포하고 있는 연속성으로 인한 불명확이 어느 단계에서 발생하는 지에 대해서 오류가 있었다는 점이다. 공정으로부터 측정되는 품질 특성 데이터는 연속적인 분포를 이루므로 퍼지 이분적인 개념을 사용하여 등급으로 분류하는 경우에 각 등급에 속하는 비율의 산출에 있어 많은 정보의 손실이 발생된다. 즉 분류에 속하는 원소의 갯수를 고려하므로 각 품질 분류내의 품질의 분포를 알 수 없다. 따라서 좋은 품질의 경우에도 아주 좋은 품질과 좋은 품질 내에서 저급에 속하는 품질들에 대한 정보를 얻을 수 없다. 따라서 각각의 품질 분류에 공정의 특성에 따라 연속적으로 그 값이 변화하는 삼각퍼지함수를 설정하여 이용하면, 분류 경계에서의 작은 차이로 다른 등급에 속하게 되는 불합리성이 보상되며 각 등급 안에 위치하는 정보의 분포특성 파악도 가능하게 된다.

Fig.2에서 가, 나, 다는 품질특성을 등급으로 분류하기 위한 구간이며 각 품질 구간에 대해 삼각 퍼지 함수를 설정한 뒤 측정된 데이터에 각 퍼지소속함수를 적용시켜 식 (3.4)와 같이 소속함수를 계산한다.

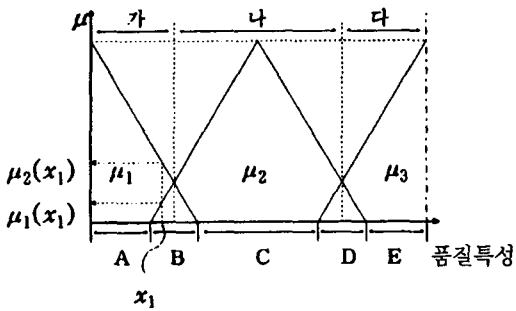


Fig. 2 Concept of fuzzy sampling method

$$\mu_{total} = \sum \mu_{A1} + \sum \mu_{B1} + \sum \mu_{B2} + \sum \mu_{C2} + \sum \mu_{D2} + \sum \mu_{D3} + \sum \mu_{E3} \quad (3.4)$$

μ_{Xi} : X구간의 데이터들에 대한 μ_i 소속함수의 값
 (X = A, B, ..., E, i = 1, 2, 3) 이때 품질비율은 다음과 같다.

$$\left[\frac{\sum \mu_{A1} + \sum \mu_{B1}}{\mu_{total}}, \frac{\sum \mu_{B2} + \sum \mu_{C2} + \sum \mu_{D2}}{\mu_{total}}, \frac{\sum \mu_{D3} + \sum \mu_{E3}}{\mu_{total}} \right] \quad (3.5)$$

퍼지 추출 기법을 포함한 전체 데이터 처리도는 Fig.1에서 $S_1 \rightarrow S_2 \rightarrow S_3$ 로서 표현된다. 이러한 퍼지추출은 기존의 퍼지관리도와 구별되는 것으로서 퍼지관리도에서 사용되던 이분적 구분에 의하여 발생하는 정보의 손실을 방지할 수 있다.

4. 통계적 기법의 적용

기존의 퍼지 기법에서 사용되는 관리 한계선 설정방법에는 k값 설정에 어려움이 있다는 단점을 이해하여, 본 논문에서는 이러한 단점을 Hybrid 기법을 통해서 개선하였다. 퍼지 추출과 퍼지 연산과정을 거친 후 비퍼지화 결과치인 퍼지대표값은 원래의 데이터에 형태상의 변화가 발생하였으나 기본적인 정보에는 변화가 없다. 따라서 퍼지대표값에 통계적인 기법을 적용할 수 있다. 그리고 통계적인 관리도는, 가설 검정과 같은 객관적인 성능 평가를 위한 다양한 기법이 존재하므로, 이분적인 개념보다 많은 양의 정보를 가지고 있는 데이터에 통계적인 기법을

적용하여 기존의 퍼지 관리도에서 문제점으로 인식되어 왔던 관리한계 설정의 어려움을 해결할 수 있다.

Fig.3은 이러한 개념을 표현한 것으로, 불명확한 데이터가 퍼지처리 과정을 거치면서 다루기 쉬운 데이터로 변환되며, 이러한 불명확한 데이터와 변환된 다루기 쉬운 데이터는 같은 특성을 가지게 된다. 이러한 기법을 적용하기 위해서 퍼지 연산을 거친 퍼지대표값의 정규 분포성이 퍼지연산 이전 데이터의 분포와 동일성 여부를 검정할 필요가 있으며, 본 논문에서는 비모수 통계기법인 콜모고로프 검정을 이용하여 이러한 가정을 검정하였다.

5. 성능비교분석

품질정보는 연속적인 개념이므로 각 품질 분류 집합에 각 공정 데이터가 포함되는가의 여부만을 가지고 판단하는 이분적인 통계적인 개념으로는 이러한 품질 정보 특성을 표현하는데 어려움이 있다. Fig. 4와 같이 통계적인

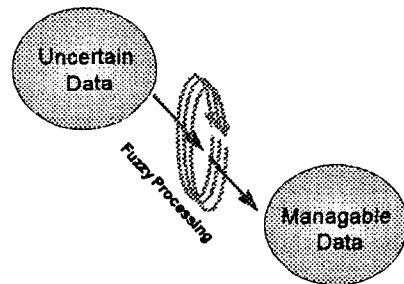


Fig. 3 Basic concept of applying statistical method to fuzzy data

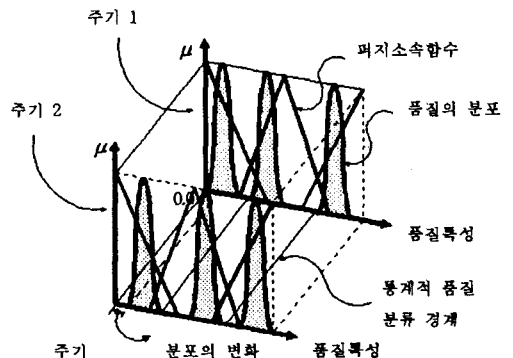


Fig. 4 Shift of quality distribution as change process period

품질 분류법에서는 공정 주기가 변함에 따라 품질정보에 변화가 발생하더라도 이들 변화가 품질 분류 내에서 변한다면 이러한 변화를 인식하지 못한다. 따라서 품질 정보의 손실이 발생하게 된다. 그러나 퍼지 개념을 적용하여 품질 특성을 평가하는 경우 연속적인 특성을 효과적으로 표현할 수 있다.

데이터의 분류 구간 안에서의 분포 변화에 대한 민감성 비교를 위하여 각각의 분류 구간에 1개씩의 데이터가 있는 두쌍의 데이터 세트를 고려한다.

Fig.5에서 A와 B의 데이터는 상이한 품질 경향을 가짐에도 불구하고 종래의 이분적인 방법으로는 동일한 비율(0.33, 0.33, 0.33)로 계산된다. 그러나 퍼지추출기법을 사용하면 아래와 같은 결과를 얻어 구간 내에서의 분포의 차이를 인식할 수 있다.

Table 1의 소속함수값에 대해서 퍼지 추출 기법을 적

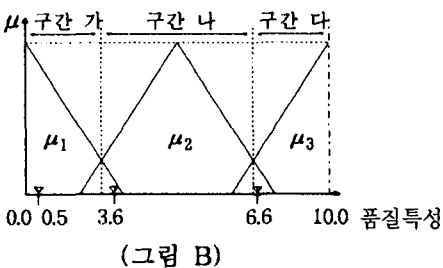
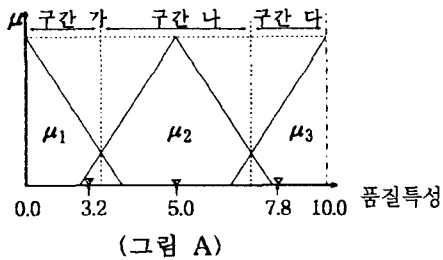


Fig. 5 Sensitivity comparison for information distribution changes

Table 1 Result of fuzzy sampling using fig.5 data

위 치	A			B		
	3.20	5.00	7.80	0.50	3.60	6.60
μ_1	0.25			0.88	0.10	
μ_2	0.10	1.00			0.30	0.20
μ_3			0.45			0.15

용하면 데이터 A의 경우 품질 비율은

$$\left[\frac{0.25}{1.8} \frac{1.1}{1.8} \frac{0.45}{1.8} \right] = [0.138 \quad 0.611 \quad 0.25]$$

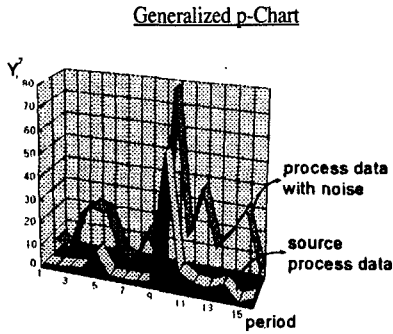
이며, B의 경우에는 (0.53 0.246 0.215)이다.

A에서 3개 품질데이터는 '나'구간에 치우치는 경향이 있으며 B에서는 '가'구간에 치우치는 경향이 있는데 이러한 경향은 종래의 비율 계산 방법으로는 구분할 수 없으나 퍼지추출을 통하여 그 경향이 파악됨을 알 수 있다.

본 연구에서는 노이즈의 증가에 따른 Process Shift의 크기를 비교하여 퍼지-통계적 관리도의 강인성을 비교 분석한다. 여기서 비교 대상으로 설정한 통계적인 관리도는 비교의 공정성을 위해 다양한 품질 분류의 관리를 위해서 개발된 Generalized p-Chart이다.⁽⁵⁾ 종래의 데이터 이분적인 추출법에서는 측정데이터에 포함되는 노이즈의 크기가 증가하면 분류 구간의 경계 부근에 위치하는 데이터는 경계를 넘어 이동할 가능성이 있다. 따라서 Generalized p-Chart 혹은 퍼지관리도에서는 노이즈가 증가하게 되면 공정상태가 변화한 것으로 판단하는 제1종 오류의 가능성이 커지게 된다. 그러나 퍼지 추출법을 적용한 퍼지-통계적 관리도에서는 노이즈의 증가에 따른 영향이 종래의 방법에 비하여 크지 않다.

본 논문에서 모의 실험의 대상으로 선정한 데이터는 기존의 퍼지 연구에서 공통적으로 많이 사용되어진 유가차 Manufacturing Process에 대한 벽돌 품질 분류 데이터를 적용하였다.⁽⁵⁾ 품질 분류 구간에 대해서 각 데이터를 Random으로 분포를 시킨 후 각 데이터의 실제 형태에 대한 기존의 고전집합개념의 경우와 퍼지 추출법을 적용한 각각의 경우에 노이즈를 첨가하여 데이터 자체에 변화를 주었다. Fig.6은 벽돌 분류 데이터에 적용한 노이즈의 크기가 증가함에 따라서 관리도가 변화하는 모습을 보여 준다. Generalized p-Chart의 경우에는 기본 데이터에 비하여 관리도의 Process Shift의 크기가 크지만, 퍼지 추출법을 적용한 퍼지-통계적 관리도의 경우에는 Process Shift가 상대적으로 작고 데이터의 경향도 안정되어 있음을 알 수 있다.

각 주기의 데이터를 정규화하여 도시한 Fig.7은 퍼지추출법을 사용한 퍼지-통계적 관리도의 노이즈에 대한 강인성을 잘 나타내고 있다. 즉 노이즈의 크기가 증가함에 따라 Generalized p-Chart는 Process Shift가 커지나 퍼지관리도는 상대적으로 그 값이 안정되어 있음을 알 수 있다.



Fuzzy Control Chart Using Fuzzy Sampling

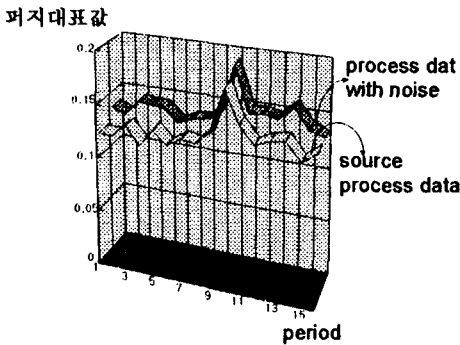


Fig. 6 Comparison of control chart performance for noise

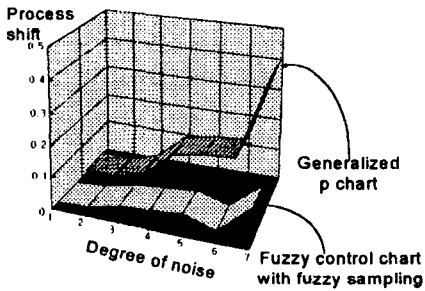


Fig. 7 Robustness against noise effect

6. 결 론

품질 특성의 구간을 기준으로 데이터를 분류하여 관리하는 경우, 종래의 관리도는 각각의 데이터가 어느 구간에 속하는가에 따라 비율을 계산하여 품질을 관리하는 이분적인 기법이다. 그러나 연속성을 가지는 품질 특성을 특정 경계를 중심으로 이분적으로 나누는 것이 불합리하

여 품질정보의 손실이 발생하게 된다.

본 연구에서 개발한 퍼지-통계적 관리도는 퍼지추출을 통하여 품질데이터를 이분적이 아닌 연속적인 개념으로 분석하므로 더욱 신뢰성 있는 품질정보를 얻을 수 있으며, 퍼지 추출법은 연속적인 각 표본에 대해서 그 분포변화에 대하여 민감하게 반응하여, 공정상에서 발생하는 하드웨어 및 주변환경의 변화 영향 등의 노이즈에도 강인한 특성을 가진다. 그리고 퍼지연산과정을 거친 정보에 통계적인 기법을 적용하여, 퍼지관리도에서 존재하는 관리 한계선 설정의 어려움을 개선하였다.

참 고 문 헌

1. Harrison M. Wadsworth, Kenneth S. Stephens, and A Blanton Godfrey, "Modern Methods for Quality Control and Improvement", JOHN WILEY & SONS, pp. 3-12, 1986.
2. Piero P. Bonissone, "Discussion: Fuzzy Logic Control Technology : A Personal Perspective", Technometrics, Vol. 37, No. 3, August, 1995.
3. Abraham Kandel, Alejandro Martins and Roberto Pacheco, "Discussion : On the Very Real Distinction Between Fuzzy and Statistical Methods", Technometrics, Vol. 37, No. 3, August, 1995.
4. Michael Laviolette, John W. Seaman, Jr, J. Douglas Barrett and William H. Woodall, "A Probabilistic and Statistical View of Fuzzy Methods", Technometrics, Vol. 37, No. 3, August, 1995.
5. Mark Marcucci, "Monitoring Multinomial Processes", Journal of Quality Technology, Vol. 17, No. 2, pp. 86-91, April, 1985.
6. Arnold Kaufmann and Madan M. Gupta, "Introduction to Fuzzy Arithmetic", International Thomson Computer Press.
7. Jyh-Hone Wang and Tzvi Raz, "Applying Fuzzy Set Theory in the development of Quality Control Chart", International Industrial Engineering Conference Proceedings, pp. 30-35, 1988.