

데이터 마이닝 기반의 품질설계지원시스템*

지 원 철**

Quality Design Support System based on Data Mining Approach*

Won Chul Jhee**

■ Abstract ■

Quality design in practice highly depends on human designer's intuition and past experiences due to lack of formal knowledge about the relationship among IO variables. This paper represents an data mining approach for developing quality design support system that integrates Case Based Reasoning (CBR) and Artificial Neural Networks (ANN) to effectively support all the steps in quality design process. CBR stores design cases in a systematic way and retrieve them quickly and accurately. ANN predicts the resulting quality attributes of design alternatives that are generated from CBR's adaptation process. When the predicted attributes fail to meet the target values, quality design simulation starts to further adapt the alternatives to the customer's new orders. To implement the quality design simulation, this paper suggests (1) the data screening method based on ξ - δ Ball to obtain the robust ANN models from the large production data bases, (2) the procedure of quality design simulation using ANN and (3) model management system that helps users find the appropriate one from the ANN model base. The integration of CBR and ANN provides quality design engineers the way that produces consistent and reliable design solutions in the remarkably reduced time.

Keyword : Intelligent Quality System, Data Mining, Artificial Neural Network, Case Base Reasoning, Data Screening, Quality Design Simulation, Model Management System

논문접수일 : 2003년 7월 12일 논문게재확정일 : 2003년 7월 25일

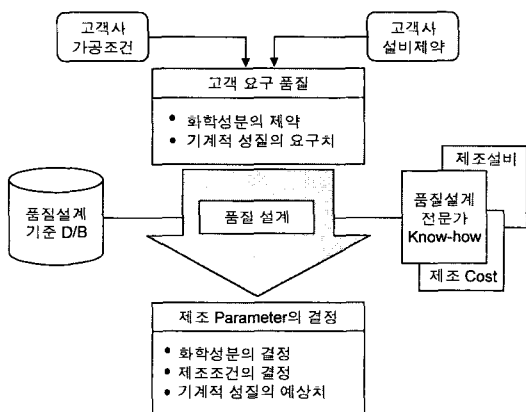
* 본 논문은 홍익대학교 교수 연구년 기간(2001. 8. 1~2002. 7. 31)중 연구되었음.

** 홍익대학교 정보산업공학과 교수

1. 서 론

고객의 다양한 품질요구를 신속히 만족시키기 위해서는 품질기준에 대한 객관적 검증 및 지속적 보완이 필요하다. 따라서 제품설계 및 생산에 관련된 지식들을 체계적으로 수집, 분석하여 공유가 가능하도록 하는 지능형 품질시스템(Intelligent Quality System, IQS)의 필요성이 커지고 있다. 최근 IQS의 개발 사례들이 많이 소개되고 있는 것은 급격한 경영환경의 변화에 능동적으로 대처하면서 고객의 품질요구에 신속히 대응하기 위함이다. 효과적인 IQS의 구축 및 활용을 위해서는 제품 수주, 설계 및 생산의 전 과정에서 품질의 설계, 추적, 판정 및 해석이 가능하도록 하여야 한다.

프로세스 산업의 경우 총생산량의 20~30% 정도가 고객의 주문에 의해 생산되지만, 주문생산 요구에의 신속한 대응여부는 해당 기업의 능력 및 대외적 평판을 결정하는 가장 중요한 요소이다. [그림 1]에서 볼 수 있듯이 제품에 대한 고객의 품질 요구에 맞추어 원재료의 화학적 성분 구성 및 생산공정의 조정여부를 결정하는 품질설계는 프로세스 산업에서도 가장 중요한 기능이다.



[그림 1] 프로세스산업의 품질설계 과정

하지만, 고객의 품질 요구조건과 제조파라미터들 사이의 관계를 단순 통계적 분석에 의존하는 경우가 많다. 특히 신제품 또는 유사제품의 품질특성

에 대한 설계는 개인 품질설계전문가들의 경험과 판단에 의존하는 경우가 대부분으로, 품질설계 결과의 일관성 부족 및 고객요구에 대한 신속한 대응이 어려운 상황이다. 따라서 품질설계지원시스템의 개발은 IQS의 성공적 구축을 위한 핵심 전제조건이다.

품질설계전문가들이 신제품이나 유사제품의 품질설계시 과거의 품질설계 사례들을 많이 이용한다는 사실을 감안할 때, 사례기반추론(Case Based Reasoning, CBR)을 활용한 지능형 품질설계지원시스템의 구축은 바람직한 접근방법으로 판단된다. 과거 설계사례들의 체계적인 수집 및 검색을 통하여 신속하면서도 일관성 있는 품질설계를 가능하게 한다는 장점이 있지만, 검색된 유사사례로부터 고객이 요구한 목표품질을 달성할 수 있는 설계대안을 생성하는 CBR의 적응과정(Adaptation)을 자동화하는데 어려움이 많아 여전히 품질설계전문가들의 의존도가 높아진다는 단점이 있다. CBR의 적응과정을 위한 규칙기반시스템을 구축하는 경우에도 규칙베이스의 생성 및 유지보수에 많은 노력을 투입하여야 한다. 따라서 본 논문에서는 CBR과 신경망(Artificial Neural Networks, ANN)을 상호보완적으로 사용함으로써 효과적인 품질설계지원시스템을 구축할 수 있는 방안을 제시하고자 한다.

생산설비의 자동화를 위한 지속적인 투자 결과로 최근 생산현장에서는 프로세스 컴퓨터에 의한 대용량 생산실적 자료의 수집이 용이해졌으며, 자료의 수집양도 생산현장에서는 3개월 이상을 보관하기 어려운 경우가 많다. 하지만 수집된 실적자료의 체계적인 분석 및 활용 시스템의 구축은 아직도 미진한 상황이어서 대부분의 자료가 공정 및 품질관리를 위한 유용한 자료로 활용되지 못하고 사장되고 있는 실정이다. 따라서 본 연구에서는 데이터 마이닝 접근 방법에 의한 생산실적자료의 분석 결과를 품질설계 과정에 효과적으로 활용할 수 있도록 다음 사항에 중점을 두었다.

첫째, 데이터 마이닝에 의한 생산공정의 모형화 과정에서 데이터 스크리닝(Data Screening) 작업의

중요성이다. 본 연구에서 생산 공정의 모형화란 품질설계에 관련된 입출력 요소들 간의 관계를 신경망을 통해 모형화하는 것이다. 수립된 신경망 모형은 품질설계 결과로서 도출된 제조파라미터들이 실제 생산 현장에서 목표한 품질특성치들을 달성할 수 있는지 예측하는데 1차적으로 활용된다. 데이터 마이닝에서 자료 전처리(Preprocessing)과정의 중요성은 이미 강조되어 왔다. 하지만 자동화된 생산 실적 자료의 분석에서는 결측치나 이상치의 처리보다는 상반된 자료의 처리가 수립된 모형의 강건성(Robustness)을 결정하는 중요한 요인이다. 상반된 자료란 입력값의 동질성에도 불구하고 출력값이 크게 다른 경우로서 생산현장의 다양한 요인에 의해 최종제품의 품질특성치들이 왜곡되는 경우가 적지 않기 때문이다. 따라서 프로세스 산업에서 활용 가능한 데이터 스크리닝 방법에 대해서 중점적으로 설명한다.

둘째, 수립된 신경망 모형을 CBR의 적응과정에 활용하는 방안이다. 대용량 생산실적 데이터베이스로부터 수립된 신경망 모형은 신뢰성이 높으며, 품질설계의 입출력 요소들 사이에서 흔히 관측되는 비선형 관계를 효과적으로 반영할 수 있으므로, 고객의 요구품질을 만족시키기 위한 제조파라미터의 정교한 조정에 유리하다. 따라서 CBR에 의해 생성된 품질설계 대안에 대해 목표 품질의 달성여부를 확인하고, 필요시 입출력 변수들 간의 관련 유효범위 안에서 고객요구 품질을 달성할 수 있는 제조파라미터의 값들을 역추적하는데 학습된 신경망을 활용하는 방안을 제시하였다. 본 연구에서는 이를 품질설계시물레이션이라고 명명하였다.

마지막으로 학습된 신경망 모형들의 체계적 관리 필요성이다. 다양한 제품 및 그에 따른 품질특성들이 존재하므로 이를 정확히 반영하기 위해서는 다양한 입출력변수 조합에 따른 신경망 모형들이 개발되어야 하며, 필요시 주어진 상황에 가장 적합한 신경망 모형을 찾아내 품질설계 과정에 활용할 수 있는 시스템적 뒷받침이 필요하다. 이와 같은 요구사항에 대처하기 위하여 본 연구에서는 모델관리

시스템을 구현하였다.

본 논문의 구성은 다음과 같다. 2장에서는 지능형품질시스템 및 데이터 마이닝 관련 문헌을 간단히 소개한 후, 품질설계지원시스템의 개발방향에 대해 설명한다. 3장에서는 데이터 마이닝 기반의 생산실적 분석 과정을 데이터 스크리닝을 중심으로 설명한 후, 신경망을 이용한 생산 공정 모형을 철강 산업에 적용함으로써 제안한 데이터 스크리닝 방법의 유효성을 검증한다. 4장에서는 학습된 신경망을 이용한 품질설계시물레이션의 이론적 배경 및 방법에 대해 설명한다. 5장에서는 구현된 품질설계지원시스템을 모형관리시스템을 중심으로 설명한다. 마지막으로 6장에서는 연구결과의 요약 및 향후 연구 방향에 대해 논의한다.

2. 품질설계지원시스템

2.1 지능형 품질시스템

인공지능을 이용한 품질관리 기법의 개발에 관한 연구는 이미 많이 이루어져 왔으며(Rao and Lingaraj, 1988 ; Spur and Specht, 1992), 1990년대 들어서부터 관련 시스템의 개발 사례들이 많이 소개되고 있다(Pham and Oztemal, 1996). DEFT, FAITH, FIS 및 FALCON과 같은 고장진단용 전문가 시스템(Braun, 1990)을 제외하더라도, 샘플링 계획 수립에 관한 ASASP(Fard and Sabuncuoglu, 1990), 기계 부품의 품질관리를 위한 RA-IQSE(Irgens, 1991), 화학공장의 실시간 품질관리용인 RESCU(Leitch *et al*, 1991), 생산과정에서 최종제품의 품질예측을 가능케 해주는 TRIOS(Ogai *et al*, 1991), FMEA 기법에 근거한 SYNOPSIS(Owen, 1988) 등은 대표적인 품질관리용 지식기반 시스템들이다. 최근에는 생산 공정과 온라인으로 연결하여 SPC(Statistical Process Control)의 자동화를 추구하기 위한 전문가 시스템들이 개발되고 있는 추세이다(Pham and Oztemal, 1992). 신경망의 패턴 인식 능력을 사용하여 불량품의 선별이나 SPC 차

트를 분석하는 연구(Hruska and Kuncicky, 1991) 등과 같이 신경망의 품질관리에의 응용도 활발히 이루어지고 있으며, 전문가 시스템과 신경망을 결합하여 사용한 품질관리 시스템으로는 PCB 기관의 품질관리에 사용된 ESQC(Calabrese et al., 1991), 신일본제철의 QDES(Iwata and Obama, 1991) 및 공정관리 및 품질관리를 통합한 XPC(Oztemel, 1992) 등이 있다.

이와 같이 지능형 품질시스템의 개발이 보편화 되고 있는 이유는 세 가지로 요약할 수 있다.

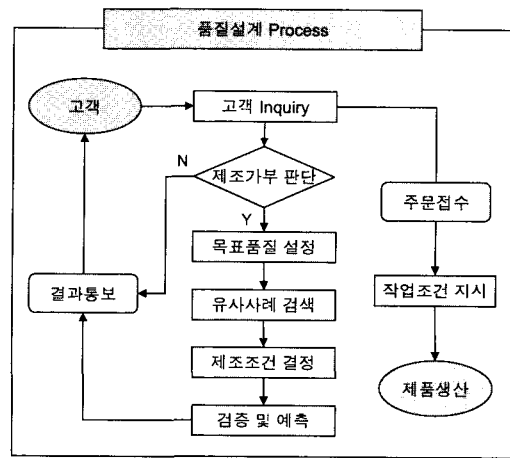
- (1) 품질관련 각종 기준들을 실적자료의 분석을 통해 객관적으로 검증하고 지속적으로 보완해 갈 수 있는 체제를 확립함으로써 생산현장에 명확한 조업기준을 제시한다.
- (2) 품질설계 관련 지식들을 체계적으로 수집, 정리, 공유함으로써 다양한 고객의 주문에 신속히 대처할 수 있도록 한다.
- (3) 지능형 품질설계지원시스템을 구축함으로써 전문 인력 부족에 사전 대비한다.

이와 같은 필요성에도 불구하고 지능형 품질설계지원시스템에 대한 체계적 연구는 Suh et al. (1998)의 '사례기반추론을 이용한 품질설계 전문가 시스템'에 관한 연구를 제외하고는 아직도 미진한 상태이다. 따라서 본 연구에서는 데이터 마이닝 기법을 활용하여 선행 연구를 확장함으로써 생산현장에 보다 효과적으로 적용 가능한 품질설계지원시스템의 구축 방안을 제시한다. 데이터 마이닝에 대해서는 Han and Kamber(2001), Pyle(1999)을 신경망에 대해서는 Rao and Rao(1995)을 참고하기 바란다.

2.2 품질설계지원시스템 구축방향

[그림 2]는 현재 일반적으로 이루어지는 품질설계 과정을 나타낸 것이다. [그림 1]에서도 볼 수 있듯이 품질설계 과정에 있어 품질설계전문가의 노하우가 매우 중요한 역할을 한다. 목표품질의 설정

이나 과거 유사사례의 검색을 원활히 수행할 수 있는 설계전문가의 양성에 상당기간이 소요되며, 동일 제품의 설계시에도 설계결과가 큰 차이가 발생하는 경우가 많다. 그 결과 고객의 주문조회에 신속한 응답이 어렵고, 생산현장에서도 조업조건의 불필요한 변경이 잦게 되어 많은 낭비 요소들이 발생하게 된다.



[그림 2] 일반적인 품질설계 프로세스

따라서 CBR을 이용한 품질설계지원시스템의 설계는 훌륭한 대안으로 생각된다. CBR에 대한 간단한 설명을 하기에 앞서 필요한 용어들을 정의하도록 한다.

[정의 1] 주요출력변수(KOV, Key Output Variable)는 최종 생산제품의 품질특성을 나타내는 값들이다. KOV는 고객들의 일차적 관심대상으로 철강재의 경우 탄성(TS), 인장강도(YP) 등이 대표적으로 고객사의 가공조건에 따라 제약조건이 주어진다. 특정 주문(O_i)에 주어진 KOV의 집합을 $Skov(i)$ 로 표시한다.

[정의 2] 주요입력변수(KIV, Key Input Variable)는 최종 생산제품의 품질특성, 즉 KOV에 영향을 미칠 수 있는 설계변수이다. 철강산업의 경우 탄소(C), 실리콘(S) 망간(Mn) 등과 같이 철강재를 구성하는 화학적 성분량과 소둔 압연 등과 같은 기계적

가공조건들이 전형적인 KIV를 구성한다. 특정 주문(O_i)을 충족시키기 위해 사용된 KIV의 집합을 S_{KIV}(i)로 표시한다. 편의상 KOV와 KIV를 함께 지칭하는 경우에는 KIOV를 사용한다.

[정의 3] 목표특성치(Target Attribute, TA)는 고객에 의해 주어진 제약조건들로서 품질설계 과정에서 반드시 만족시켜야 한다. TA는 KIOV일 수 있으며 특정 주문에 주어진 TA의 집합을 S_T(i) = {TA(i, j) | TA(i, j) ∈ S_{KOV}(i) or TA(i, j) ∈ S_{KIV}(i)}로 표현한다.

[정의 4] 설계특성치(Design Attribute, DA)는 TA를 만족시키기 위해 품질설계자에 의해 결정되어야 할 KIV들이다. 특정 주문을 충족시키기 위하여 필요한 DA의 집합을 S_D(i) = {DA(i, k) | DA(i, k) ∈ S_{KIV}(i)}로 표현한다.

[정의 5] 결과특성치(Resulting Attribute, RA)는 주문에 대한 실제 생산이 이루어진 후에 수집된 KIOV에 대한 실적 통계이다. 따라서 RA 집합의 KOV들은 TA 집합에 포함되지 않았던 KOV들을 포함하게 되며, RA의 값들도 범위로 주어지는 것이 일반적이다. 특정 주문에 대한 RA 집합을 S_R(i) = {RA(i, p) | RA(i, p) ∈ S_{KIV}(i) or RA(i, p) ∈ S_{KOV}(i)}로 표현한다.

Suh *et al.*(1998)은 이상의 정의를 사용하여 [그림 3]과 같이 과거의 생산실적을 나타내는 사례를 표현한 후, 계층적 구조를 갖는 사례베이스 및 유사도 기반의 사례 검색시스템을 설계하였다. 특히 사례간 유사도를 결정함에 있어 TA 기반의 유사도와 RA 기반의 유사도를 동시에 사용하여 유효한 사례의 검색효율을 높였으며 검색된 사례로부터의 적응과정을 위하여 규칙베이스를 활용하는 방안을 제시하였다.

이상과 같은 CBR 기반의 품질설계지원시스템은 유사사례의 검색 및 조정을 전적으로 품질설계전문가들의 노하우에 의존하여 발생하였던 문제들을 해결할 수 있는 유력한 대안임에 틀림없지만 다음

과 같은 보완이 필요하다. 첫째, 품질설계 결과로 결정된 KIV들이 실제 생산결과 TA들을 만족시킬 수 있는지를 정확히 예측할 수 있는 수단이 필요하며, 둘째, 검색된 사례로부터 원하는 DA 값들을 결정하는 적응과정을 개선하는 것이다. 규칙베이스를 이용한 적응방법은 지식획득의 어려움을 감안할 때 계속적으로 변화해가는 생산현장의 모든 여건들을 정확히 반영할 수 있는 규칙베이스의 유지관리 자체가 시스템 운영의 장애요인으로 등장할 가능성이 크기 때문이다.

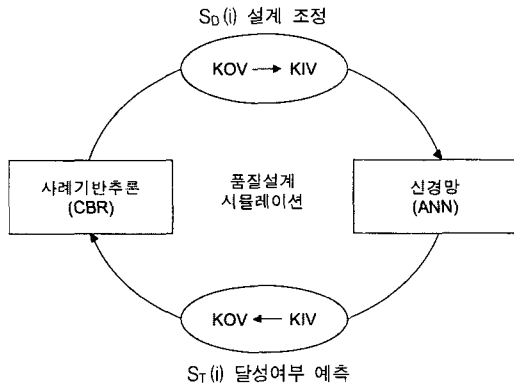
```

{{ Case #101
  강종 : 후판
  두께 : 6
  너비 : 1524
  용도 : C03
  주문처 : ○○○주식회사
  주문번호 : Oi
  KOV : TS, YP, YL
  KIV : C, Mn, Si, S, P
  TA : (TS 55±5), (YP ≥ 30), (C ≤ 0.2),
        (Mn ≤ 1.6)
  DA : (C 0.139), (Mn 1.3), (Si 0.387), (S 0.004),
        (P 0.017)
  RA : TS 54±0.3), (YP 39.7±1.2), (EL 25±3)
        (C 0.139±0.002), (Mn 1.33±0.07), (Si 0.39±
        0.01), (S 0.004±0.0005), (P 0.016±0.003)
}}
    
```

[그림 3] 사례의 표현 예

따라서 본 연구에서는 [그림 4]와 같이 CBR과 신경망이 결합된 품질설계지원시스템을 설계 구현하였다. 시스템에서 제공된 기능들을 사용하여 최종설계에 이르는 과정을 품질설계시물레이션이라고 부른 이유는 양자가 서로 보완적인 기능을 수행해가면서 최종설계에 도달해 갈 수 있도록 시스템을 설계하였기 때문이다. 즉, 품질설계의 제약조건들이 결정되면 고객의 품질 요구특성치들을 KOV로 하여 유사사례를 검색하여 요구품질의 달성이 가능한 KIV의 값들을 정한다. 사례기반추론에 의해 결정된 KIV들이 고객의 요구품질을 만족시킬 수 있는가를 신경망의 품질예측 기능을 통하여 검증하고, 만약 목표품질의 달성이 불가능할 경우에는 사

례기반추론으로 되돌아가 다른 유사사례들을 검토하는 과정을 반복적으로 수행할 수 있도록 하였다.



[그림 4] CBR과 ANN의 결합에 의한 품질설계

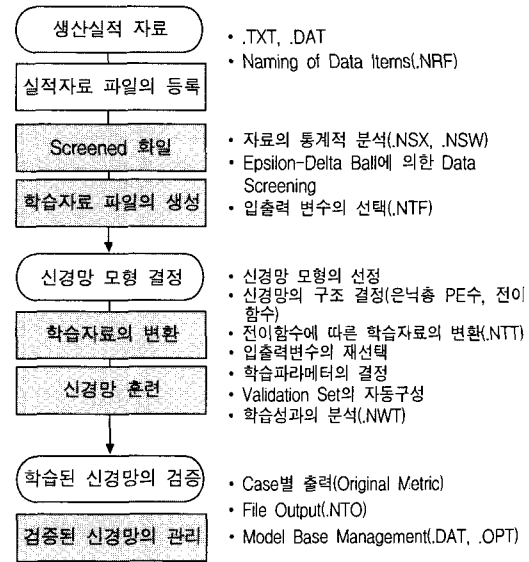
이와 같이 품질설계지원시스템의 주요 모듈인 사례기반추론과 신경망은 서로 보완적 기능을 수행할 수 있으며, 또 각기 독립적으로도 품질설계가 가능하도록 시스템을 구성하였다. 다음 장에서는 데이터 마이닝 접근방법에 의한 신경망 모형의 수립에 대해 설명한다.

3. 데이터 마이닝 접근방법

3.1 신경망을 이용한 생산실적 자료의 분석

대용량 생산 실적자료를 신경망을 이용하여 분석하는 과정은 [그림 5]와 같다. 생산현장의 프로세스 컴퓨터로부터 생산실적 자료가 수집되면 기초 통계분석 및 데이터 스크리닝을 통한 자료의 전처리 과정을 거친 후, 신경망의 학습자료를 생성하게 된다. TA를 구성하는 KOV들을 신경망의 출력으로 결정한 후 이에 상응하는 입력변수를 선정하면 신경망 구조에 관한 의사결정을 거쳐 신경망을 훈련시키게 된다. 학습된 신경망은 검증과정을 거쳐서 품질설계 결과가 생산현장에 적용되었을 경우의 예상되는 TA들을 예측하는데 사용되거나, 품질설계시뮬레이션 과정에 활용된다. 참고로 [그림 5]에서 괄호안의 약어는 분석과정에서 변환된 파일

들의 확장자명을 나타낸다.



[그림 5] 신경망을 이용한 데이터 마이닝 과정

효과적인 품질설계를 위해서는 다양한 제품군들에 대해 생산실적 자료를 이용하여 많은 신경망 모형들을 필요하다. 따라서 신경망 학습의 효율성을 높이고 학습된 신경망의 활용이 용이하도록 하기 위해서는 다음과 같은 기능들이 필요하다.

(1) 데이터 스크리닝 기능

생산 실적자료를 신경망 학습자료로 직접 사용할 경우 좋은 결과를 얻기 어렵다. 전처리 과정을 거치지 않은 생산 실적자료는 상식적인 범위를 벗어나는 이상치나 결측치를 가진 자료들이 다수 포함되어 있기 때문이다. 또, 상반된 학습자료의 제거는 학습된 신경망의 일반화 능력을 높이기 위해 필수적이다. 상반된 자료의 발생원인은 두 가지로 볼 수 있다. 첫째는 생산현장에서 최종제품의 품질 특성치들이 왜곡되는 경우이고, 두 번째는 모형의 입출력 변수를 결정했을 경우, 입력변수의 제한으로 인하여 입출력 변수간의 대응에 이상상태들이 발생하는 경우이다. 특히, 두 번째 유형은 흔히 Modeling Problem으로 알려져 있는 문제로 이의 해결을 위해서는 데이터 스크리닝 과정이 필요하다.

(2) 변수선정 지원기능

효율적인 신경망 모델을 구성하는 데 있어서, 원하는 품질 목표치 KOV에 영향을 주는 품질인자 KIV의 선정이 어렵다는 것이다. 생산라인에서 많은 경험을 쌓은 품질 실무자라 하더라도 해당 KOV에 대한 KIV의 영향도를 정확히 파악하지 못하는 경우가 많다. 효율적인 신경망 학습이 이루어지기 위해서는 상호 영향도가 큰 KIV와 KOV로 이루어진 최소한의 학습파일의 구성이 중요한 요소이므로, 이를 지원하기 위하여 변수별 특성치의 파악, 상관분석, 다변량 회귀분석 등의 통계적 분석방법들이 필요하다.

(3) 신경망 학습자동검증 기능

학습중인 신경망 모형의 학습정도를 파악하여 최적의 학습이 이루어진 시점에서 학습을 멈추게 하는 기능이다. 학습집합에 대한 정확도만을 고려하여 신경망의 학습 멈춤조건을 설정하는 것은 학습된 신경망의 일반화 능력을 떨어뜨리는 결과를 가져올 수도 있다. 이는 품질설계전문가 뿐만 아니라 신경망 학습의 전문가에게도 어려운 문제이다. 따

라서 신경망의 학습시 하나 이상의 검증용 파일을 함께 사용하여 학습과정에서 수시로 학습중인 신경망의 일반화 능력을 검증하여 적절한 학습 멈춤 시점을 결정하는 기능이 필요하다.

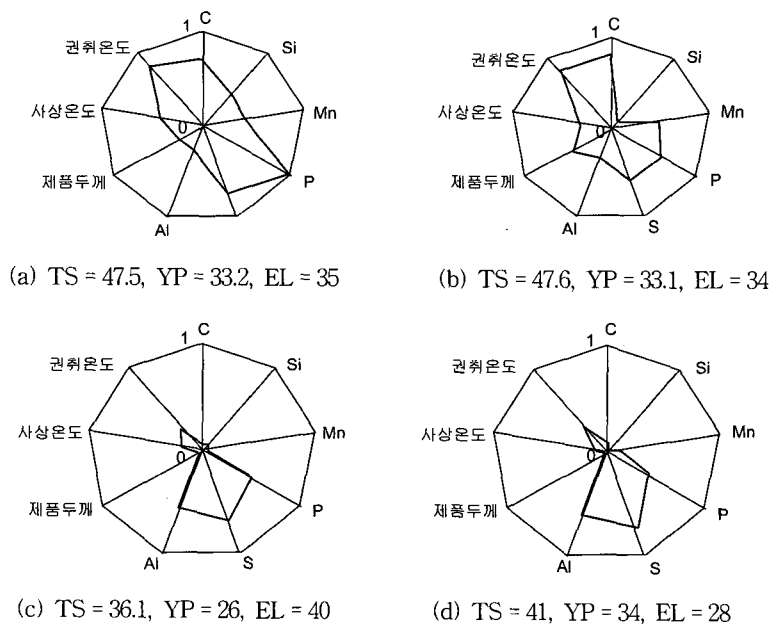
(4) 모델관리 기능

학습을 마친 신경망 모델을 검증하고 이의 추후 활용을 용이하도록 모델관리 기능이 개발되었다. MMS(Model Management System)의 주요기능은 검증결과 유효한 신경망 모델들을 상세한 사용조건과 함께 모델베이스에 등록시키는 기능, 품질설계자가 원하는 조건의 신경망을 모델베이스로부터 쉽게 찾아주는 기능 및 선택한 신경망을 이용하여 품질예측 및 품질설계시물레이션을 수행할 수 있도록 해주는 기능 등이다.

3.2 데이터 스크리닝

3.2.1 상반된 학습자료 패턴

[그림 6]은 특정 제품군을 위한 신경망 학습집합에 포함된 생산실적자료를 발췌한 것으로 TS, YP,



[그림 6] KOV와 KIV의 샘플 패턴

EL 등의 KOV와 학습집합에 선정된 화학성분 및 권취온도, 사상온도 등과 같은 기계적 가공조건 등 KIV의 관계를 레이더 다이어그램으로 나타낸 것이다. [그림 6](a)와 [그림 6](b)는 서로 상이한 KIV 패턴을 가지고 있지만 KOV 패턴은 유사함을 볼 수 있다. 하지만, [그림 6](c)와 [그림 6](d)는 아주 유사한 KIV 패턴을 가지고 있음에도 불구하고 서로 다른 KOV 패턴을 보여주고 있다. 이와 같은 학습자료가 동일한 학습집합 내에 존재할 경우 상반된 자료라고 한다. 이는 일반적 함수의 정의에 위배되므로 신경망의 학습을 크게 저해시키고, 학습된 신경망의 예측능력 즉, 일반화 능력을 크게 떨어뜨리는 요인이 된다. 따라서 상반된 학습자료의 제거는 신경망의 효과적 학습을 위하여 필수적으로 필요한 과정이다.

3.2.2 Epsilon-Delta Ball

상반된 학습자료를 제거하여 모형의 일반화 능력을 향상시킨 연구는 이미 존재한다. Versaggi (1995)는 개인 신용평가 문제에 대해 입력 패턴들 간의 유사도를 판단하는 Epsilon Ball을 사용하여 상반된 자료들을 제거한 결과 예측 모형의 정확도가 크게 향상되었다는 연구결과를 발표하였다. 이 연구는 분류 문제를 다루었기 때문에 이산형 단일 출력변수만을 고려하면 되므로 입력패턴의 유사성 여부만이 관심의 대상이었다. 하지만, 품질설계 문제에 적용하기 위해서는 KIV 패턴뿐만 아니라

다수의 연속형 변수에 의해 구성되는 KOV 패턴들 간의 유사성도 함께 판단할 수 있도록 확장되어야 한다.

[그림 7]은 본 연구에서 상반된 학습자료 제거를 위해 사용한 데이터 스크리닝 방법의 개념을 나타낸 것이다. [그림 7](a)와 같이 KIV 패턴들 간의 유사도는 ξ -Ball을 KOV 패턴들 간의 유사도는 δ -Ball을 사용하여 판단한다. 만일 동일한 ξ -Ball에 소속된 KIV 패턴들에 대응된 KOV 패턴들이 서로 다른 δ -Ball에 소속될 경우 해당 학습자료는 제거 대상이 되는 것이다. 따라서 ξ - δ Ball을 적용하기 위해서는 두 패턴 x_1 과 x_2 의 유사도를 판단하기 위한 거리척도가 필요하며, 다음과 같은 벡터놈(Vector Norms)들이 사용가능하다.

(1) L_1 - Norm :

$$\|x_1 - x_2\|_1 = \sum_i |x_1^i - x_2^i|$$

(2) L_2 - Norm :

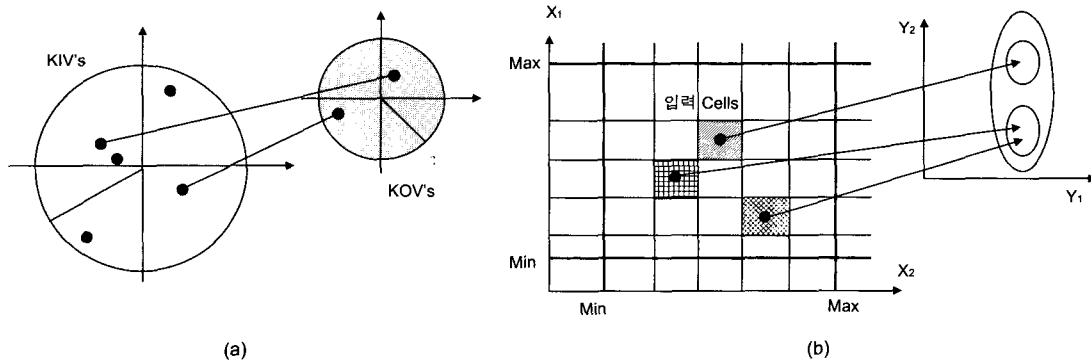
$$\|x_1 - x_2\|_2 = \sqrt{\sum_i (x_1^i - x_2^i)^2}$$

(3) L_∞ - Norm :

$$\|x_1 - x_2\|_\infty = \text{Max}_i |x_1^i - x_2^i|$$

이상과 같은 거리 척도를 사용하여 ξ , δ 의 범위를 결정할 때 다음과 같은 사항들에 유의하여야 한다.

첫째, 패턴을 구성하는 KIOV들이 갖는 값의 크



[그림 7] ξ - δ Ball의 개념도

기차이로 인한 유사도 왜곡 현상을 방지하는 것이다. 둘째, KOV와 KIV들 간의 영향도를 사전에 파악하여 가중치를 사용함으로써 보다 현실적인 유사도 판정이 가능하도록 하는 것이다. 셋째, $\xi - \delta$ Ball들의 중심점을 결정하는 문제이다. 특히, 모든 학습자료들의 KIV 패턴은 반드시 하나의 ξ -Ball에 속하도록 중심점을 결정하여야 한다. 마지막으로 $\xi - \delta$ Ball들의 경계에 위치하는 패턴들 간의 입출력 대응관계에서 문제가 발생할 수 있다는 것이다. 즉, 사실상 매우 가까운 위치에 존재하는 두 패턴입에도 불구하고 서로 다른 Ball에 소속됨으로써 잘못된 입출력 대응관계를 가지게 되어 해당 학습자료가 제거되는 현상이 발생하는 것이다.

이를 위해 본 연구에서는 각각의 KIOV들별로 분포를 점검한 후, [그림 7](b)에서와 같이 구간을 분할하도록 하였으며, 유사도를 판정할 때는 정규화된 값을 사용하였다. 이는 거리적으로서 L_{∞} -Norm을 사용한 것과 같은 효과가 있으며, $\xi - \delta$ Ball들의 중심점을 결정하는 복잡한 문제를 피할 수 있다는 장점이 있다. $\xi - \delta$ Ball의 경계를 설정함으로써 발생하는 문제를 해결하기 위해서는 ξ, δ 값들에 대한 Perturbation Analysis가 필요하다. 이 때 KOV 수에 KIV 수를 곱한 수자의 두 배에 해당하는 유사도 검증 절차를 필요로 하게 된다. 따라서 많은 수의 KIOV를 갖는 대용량 학습집합에 대해 데이터 스크리닝 작업을 할 경우 많은 시간이 소요되므로 본 연구에서는 perturbation 작업 수를 효과적으로 제한하는 방법을 제시하였다.

단변량 분석을 실시하여 결측치 및 이상치(Outliers)를 제거한 학습자료에 대해 $\xi - \delta$ Ball을 적용하기 위하여 KIOV들의 구간 분할에 의해 형성되는 다차원 공간들을 각기 KIV Cell(C_{KIV}), KOV Cell(C_{KOV})이라고 한다. 또, 주어진 perturbation 작업회수를 N 이라고 하고, $\xi - \delta$ 값의 조정율을 η 라고 하면 본 연구에서 사용한 데이터 스크리닝 절차는 다음과 같다.

(1) 학습집합내의 모든 KIOV에 대해 다음을 반복

한다.

- (1.1) 해당 KIOV의 분포를 검토하여 적당한 수의 구간으로 분할한다.
- (1.2) 해당 KIOV 값들을 정규화한다.
- (2) KIV 집합 및 KOV 집합을 중요도에 따라 정렬한 후, 가중치(ω)를 부여한다.
- (3) 조정대상이 되는 KIOV 수 m_1, m_2 를 결정한 후, 다음을 N 회 반복한다.
 - (3.1) ω 를 확률로 사용하여 조정대상이 되는 KIOV들을 각각 m_1, m_2 개씩 임의 추출한다.
 - (3.2) 선정된 KIOV들에 대해 ξ, δ 값들을 $\eta\%$ 만큼 이동시킨다. KIOV별 이동방향은 임의로 정한다.
 - (3.3) 모든 C_{KIV} 에 대해 다음을 반복한다.
 - (3.3.1) 해당 C_{KIV} 에 속한 모든 학습자료들에 대해 다음을 반복한다.
 - (3.3.1.1) 각 학습자료들의 소속 C_{KOV} 를 결정한다.
 - (3.3.1.2) 최대빈도를 갖는 C_{KOV} 를 대표 $C_{KOV}(C_{KOV}^*)$ 로 결정한다.
 - (3.3.1.3) C_{KOV}^* 를 갖는 학습자료는 '포함' 아닌 경우는 '불포함'으로 표시한다.
- (4) '포함'이 임계치 미만인 학습자료들을 학습집합으로부터 제거한다.

3.2.3 실험 결과

본 연구에서 제안한 데이터 스크리닝 방법의 유용성을 검증하기 위하여 사용된 신경망은 다계층 퍼셉트론으로 하나의 은닉층만을 사용하였으며 백프로파게이션 알고리즘에 의해 학습시켰다. 일관 제철소의 후판 생산 공정으로부터 자료를 얻어 실험한 결과를 <표 1>과 <표 2>에 나타내었다. <표 1>은 4개의 강종에 대하여 데이터 스크리닝을 실시하기 전의 학습집합에 대해 다중회귀 분석 및 신경망(ANN)을 적용한 결과와 스크리닝 작업을 실시한 후 신경망(DS)을 적용한 결과를 요약한 것이다. 예측의 정확도는 KOV들 중에서 TS와 YP에 대하여 R^{*2} (R-Square)와 MAPE(Mean Abso-

〈표 1〉 4개 제품군에 대한 모형간 예측성능 비교

품질특성치		TS			YP			
사용된 모형		다중회귀	ANN	DS	다중회귀	ANN	DS	
DQ001	학습집합	R**2	0.85	0.95	0.98	0.69	0.83	0.86
	검증집합	MAPE	1.98	1.58	0.75	3.99	2.37	1.69
		R**2	0.82	0.89	0.96	0.58	0.77	0.80
DQ051	학습집합	R**2	0.51	0.61	0.79	0.35	0.53	0.68
	검증집합	MAPE	2.01	1.81	1.12	3.29	2.31	1.29
		R**2	0.39	0.48	0.72	0.20	0.42	0.63
DQ052	학습집합	R**2	0.63	0.72	0.84	0.41	0.61	0.73
	검증집합	MAPE	2.00	2.29	1.06	3.24	2.95	1.35
		R**2	0.56	0.52	0.78	0.29	0.39	0.67
DQ053	학습집합	R**2	0.35	0.63	0.81	0.24	0.52	0.68
	검증집합	MAPE	1.75	1.36	0.93	3.40	2.42	1.51
		R**2	0.38	0.53	0.75	0.12	0.37	0.59

〈표 2〉 DQ052에 대한 10-fold Cross-Validation 결과

	TS			YP		
	학 습 집 합	검 증 집 합		학 습 집 합	검 증 집 합	
	R**2	MAPE	R**2	R**2	MAPE	R**2
NDS	0.701 (0.0935)	2.277 (0.2897)	0.516 (0.1295)	0.592 (0.0911)	2.910 (0.2739)	0.377 (0.1619)
DS	0.825 (0.0562)	1.137 (0.1563)	0.818 (0.0797)	0.716 (0.0558)	1.355 (0.1482)	0.660 (0.0844)
t-value*	-5.977	14.387	-6.309	-4.348	17.147	-3.919
p-value**	0.0002	0.000	0.0001	0.0019	0.0000	0.0035

주) 1. * 유의수준 5%에서의 paired t-test 결과임.

2. ** 양측검증시 p-value임.

lute Percentage Error)에 의해 측정하였다.

〈표 1〉에서 볼 수 있듯이 데이터 스크리닝 결과 6~9%의 학습자료들이 제거되었다. 이는 Versaggi(1995)의 연구에서 35% 정도의 학습자료가 제거된 것이 비하면 적은 양이지만 데이터 스크리닝 전후 신경망의 학습집합 및 검증집합에 대한 R**2 값을 비교해 보면 데이터 스크리닝 후에 R**2값의 하락률이 크게 향상되었음을 관측할 수 있다.

〈표 2〉는 데이터 스크리닝에 의한 신경망 모형의 예측성능 향상을 통계적으로 검증하기 위하여

제품 DQ052의 학습집합을 10개로 분할하여 Cross-Validation을 시행한 결과이다. 이 때 검증집합은 본래의 실적자료를 사용하고, 학습집합에 해당되는 것들만 데이터 스크리닝을 적용하였다. 데이터 스크리닝 전후의 신경망의 예측 결과에 대해 쌍체 비교에 의한 t-검증을 실시한 결과 양자의 차이는 유의하였다. 특히 KOV인 YP에 대한 예측성능 개선은 주목할만 하였다. 따라서 본 연구에서 제시한 데이터 스크리닝 방법은 신경망의 일반화 능력을 제고시키는 수단으로 매우 유용함을 알 수 있다.

4. 품질설계시뮬레이션

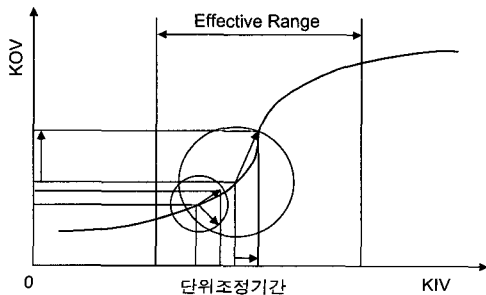
4.1 학습된 신경망의 활용

품질설계지원시스템에서 학습된 신경망의 활용은 다음과 같다. 고객의 주문 O_i 를 위하여 CBR에 의해 결정된 $S_D(i)$ 들이 고객의 요구품질인 $S_T(i)$ 를 만족시킬 수 있는가를 신경망의 품질예측 기능을 통하여 검증한다. 즉, $S_D(i)$ 및 $S_T(i)$ 에 속한 KIV들을 입력으로 $S_T(i)$ 에 속한 KOV들을 출력으로 갖는 신경망 모형을 이용하는 것으로 이를 Quick-Test 기능이라고 한다.

만약 예측결과 $S_T(i)$ 에 의해 주어진 목표품질을 달성하지 못했다면 CBR의 적응과정이 실패한 것으로 간주할 수 있다. 하지만 $S_T(i)$ 의 값과 신경망 예측치 사이의 차이가 크지 않다면 품질설계시뮬레이션(QDSim) 기능을 이용하여 $S_D(i)$ 에 속한 KIV들의 값들을 미세 조정함으로써 목표품질을 달성할 수 있다. 물론 주어진 제약조건하에서 QDSim에 의해 목표품질을 달성할 수 없다면 다시 CBR 시스템으로 돌아가 다른 유사사례를 검색하여 적응과정을 거친 후, 신경망을 이용한 품질예측 및 KIV 값들의 조정과정을 반복한다.

4.2 신경망에 의한 KIV 값의 조정

신경망을 이용하여 $S_T(i)$ 에 속한 KOV들이 주어졌을 경우, 목표 KOV 값들을 얻을 수 있는 KIV의 값들을 결정하는 문제는 [그림 8]과 같이 나타낼



[그림 8] 신경망기반 품질설계시뮬레이션

수 있다. 즉, 원하는 출력과 유사한 출력을 주는 입력패턴의 선택이 가능하다면 선택된 입력패턴으로부터 원하는 출력패턴을 얻을 수 있는 입력패턴을 정확히 찾아갈 수 있다는 것이다.

[그림 8]에서 곡선은 신경망을 나타내며, KIV의 미세조정에 따른 KOV의 변화량을 추적해감으로써 원하는 KIV 값들을 얻고자 하는 문제로서 정식화하면 다음과 같다.

$$y = f(w, x), \quad (1)$$

$$\text{where, } y \in \mathbb{R}^m, x \in \mathbb{R}^n, w \in \mathbb{R}^p$$

식 (1)과 같이 표현된 학습된 신경망이 주어졌을 때, 원하는 출력 y_d 를 줄 수 있는 입력 x_d 를 찾는 것이다. 여기서 w 는 신경망의 가중치벡터를 나타낸다.

만약, 목표출력과 유사한 출력을 주는 입력 패턴의 선택이 가능하다면 선택된 입력 패턴 (x_c, y_c) 은 식 (2)를 만족한다.

$$y_c \approx y_d, \text{ and } y_c = f(w, x_c) \quad (2)$$

식 (1)의 (x_c, y_c) 주위에 대한 Taylor 급수 전개를 하면 식 (3), 식 (4)와 같이 된다.

$$y - y_c = J(w, x_c)(x - x_c) \quad (3)$$

where,

$$J(w, x_c) = \left[\begin{array}{ccc} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{array} \right]_{x=x_c} \quad (4)$$

식 (3)을 이용하여, $y = y_d$ 가 주어졌을 때, 상응하는 입력 $x = x_d$ 를 구하면 다음과 같다.

$$y_d - y_c = J(w, x_c)(x_d - x_c) \quad (5)$$

일반적으로 $n > m$ 이므로, 식 (5)는 Under-determined Case이고 많은 해들이 존재한다. 따라서

많은 해들 중 가장 크기가 작은 해를 최적의 해라고 하면 식 (6)과 같이 원하는 입력값들을 결정할 수 있다.

$$\begin{aligned} (x_d - x_c) &= J^T(JJ^T)^{-1}(y_d - y_c) \\ x_d &= x_c + J^T(JJ^T)^{-1}(y_d - y_c) \end{aligned} \quad (6)$$

식 (4) $J(w, x_c)$ 의 요소들을 구해야 식 (6)을 사용할 수 있다. 이를 위해 i 번째 입력값만을 증분시킨 입력벡터 x^i 를 다음과 같이 결정할 후

$$x^i = x_c + [0, 0, \dots, \Delta x, \dots, 0]^T \quad (7)$$

x^i 를 입력으로 신경망 출력 $y^i = f(w, x^i)$ 를 계산하면, $\Delta y^i = y^i - y_c$ 를 얻을 수 있다. 따라서 다음과 같이 식 (4)의 구성요소들을 얻을 수 있다.

$$\frac{\Delta y^i}{\Delta x} = \left[\frac{\partial f_1}{\partial x_1}, \dots, \frac{\partial f_m}{\partial x_1} \right]^T \quad (8)$$

QDSim에서는 이상의 절차를 사용하여 $S_T(i)$ 에 속한 KOV의 목표 값들을 얻을 수 있는 KIV의 조합을 찾아 갈 수 있도록 하였다. 즉, 조정된 KIV의 조합을 찾아가는 매 단계에서 KOV 목표치와 예측치의 차이가 가장 큰 KOV에 영향력이 가장 큰 KIV값을 조정한다. QDSim에서는 시스템이 자동적으로 찾아갈 수 있도록 하였지만, 설정된 제약조건하에서 KIV의 조합을 찾아가는 과정에서 실패할 수도 있으므로 수동으로 한 단계씩 입력의 변화에 따른 출력의 변화를 확인할 수 있도록 하여, 각 단계에서 출력값에 가장 큰 영향을 미치는 KIV를 확인할 수 있도록 하였다. 이 기능은 Quick Test와 연계해 사용함으로써 시뮬레이션 과정을 효과적으로 수행할 수 있도록 도와준다.

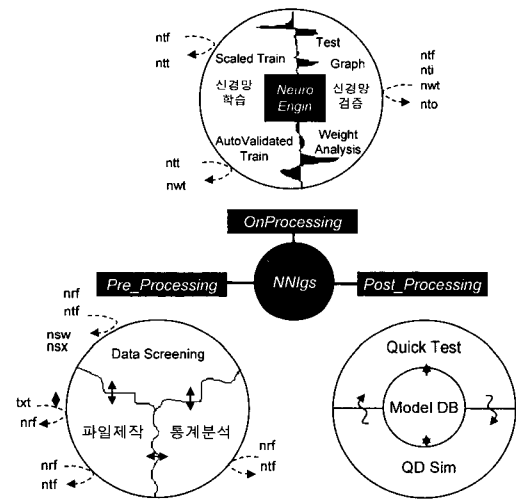
5. 품질설계지원시스템의 구현

5.1 시스템 구조

데이터 마이닝 기반의 품질설계지원시스템의 구

현에 있어 신경망 모형의 개발자나 사용자에게 최대한 편의를 제공하기 위해서 [그림 9]와 같은 구조를 갖는 신경망 관련 모듈들이 개발되었으며, 신경망 학습의 각 단계에서 발생하는 관련 학습자료의 변형 및 가공절차를 자동화하였다.

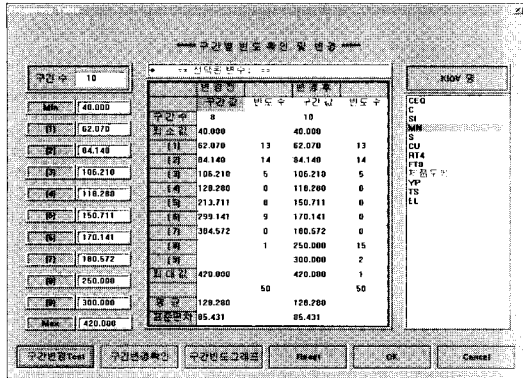
본 논문에서는 3장, 4장에서 언급하였던 4가지 중요 기능 및 QDSim의 구현에 대해 설명한다. 신경망 입력변수선정 지원 기능을 상관분석 및 중회귀분석 기능이 구현되었지만 구체적 설명은 생략하고, 데이터 스크리닝, 신경망 학습 자동검증기능 및 모델관리 시스템을 화면 중심으로 설명한다.



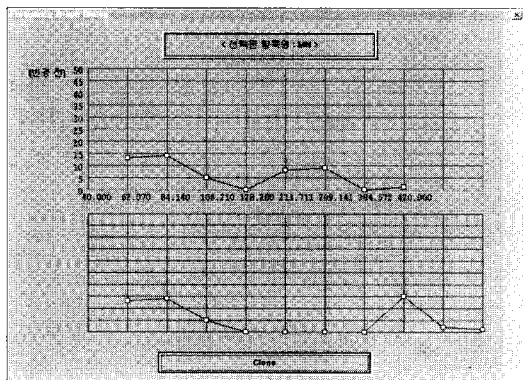
[그림 9] 품질설계지원을 위한 신경망학습시스템

5.2 데이터 스크리닝

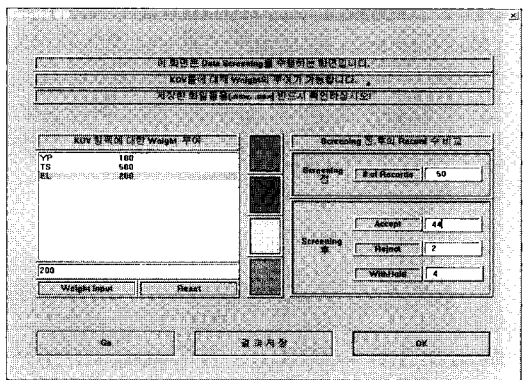
[그림 10]은 $\xi - \delta$ Ball 개념에 의한 데이터 스크리닝을 지원하기 위하여 학습집합에 포함된 KIOV들의 분포를 확인하고 구간분할을 지원하는 화면이다. 기본적으로 화면에서 선택된 KIOV의 최대 최소 값 및 평균 및 표준편차를 이용하여 $(\mu \pm n\sigma)$ 8개의 구간으로 나누어 보여준다. 사용자는 [그림 11]에서와 같이 분포를 확인한 후, 구간의 수 및 분할을 조정할 수 있다. 이 때 최대 최소 값을 변경하면 범위를 벗어나는 학습자료들은 자동적으로 제거된다.



[그림 10] KIOV의 구간분할



[그림 11] KIOV의 분포확인



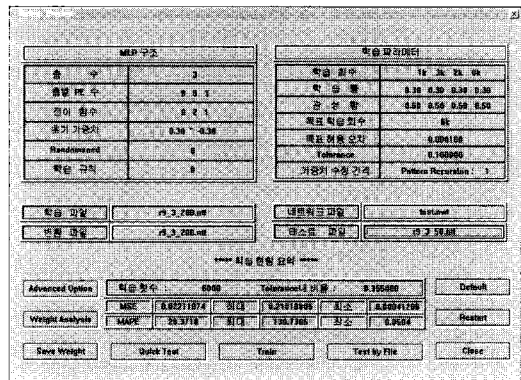
[그림 12] 데이터 스크리닝 실행 화면

[그림 12]는 데이터 스크리닝을 시행하는 화면이다. KIOV에 대한 가중치 부여가 가능하며, 스크리닝 결과는 학습자료별로 수락, 제거, 보류의 세 가지로 구분하며, 각 경우에 속한 학습자료의 수를

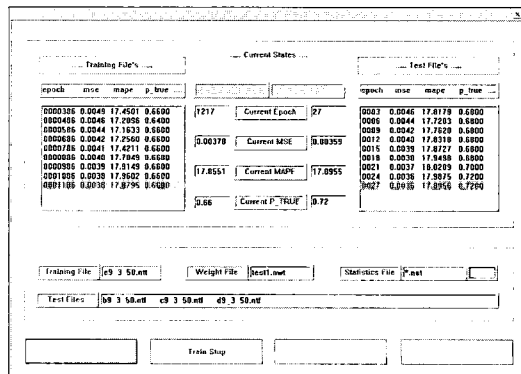
보여주며, 각각 별도 파일에 저장된다. 시스템 구현시 보류 항목을 추가한 것은 포함·불포함에 대한 임계치의 $\pm 5\%$ 범위에 있는 학습자료들에 대해서는 사용자가 학습집합에의 포함여부를 최종적으로 판단하도록 하기 위함이다.

5.2 신경망 학습 자동검증 기능

[그림 13]은 품질설계지원을 위해 본 연구에서 구현한 신경망 학습용 화면이다. 신경망의 구조 및 학습파라미터들은 상세한 설정은 부화면에서 처리되도록 하였다. 학습 자동검증 기능은 'Advanced Option'의 부화면에 구현되어 있는데 [그림 14]는 학습 자동검증 기능이 설정되어 신경망 학습이 진행되는 상황을 보여주는 화면이다.



[그림 13] 신경망 학습용 주화면



[그림 14] 자동 검증기능 수행화면

학습된 신경망의 일반화 능력을 높이기 위하여 학습과정에서 별도의 검증용 집합들(Validation Sets)을 사용하여 신경망의 학습정도를 판단하여 예측능력이 가장 좋을 것이라고 판단되는 학습회수에서 신경망의 학습을 자동으로 멈추는 기능이다. 따라서 본 연구에서는 검증용 집합을 동시에 5개까지 사용할 수 있도록 하였으며, 사용자 지정한 일정한 학습회수 간격마다 신경망 모형을 저장하고, 검증용 집합들에 대한 예측능력을 MAPE를 기준으로 확인한다.

학습 자동 멈춤을 위하여 다음과 같은 기준을 적용하였다. 첫째, 사용자에게 부여된 의무 학습회수를 수행한 후, 가장 좋은 MAPE 값을 갖는 학습회수를 찾는 것으로 현재의 MAPE가 가장 좋았던 MAPE 보다 사용자에게 부여된 임계치(Default Threshold = 0.2)보다 크면 학습을 멈춘다. 둘째는 MAPE 값의 변동이 계속적으로 미미할 경우로, 현행 MAPE와 가장 좋았던 MAPE간의 차이가 0.01 이하인 상태가 사전에 정의된 학습회수 이상 진행될 경우 자동으로 신경망 학습을 정지시키는 것이다.

5.3 모델관리시스템

신경망 모형의 예측 정확성을 높이기 위해서는 하나의 제품군에 대해서도 다양한 KIOV 조합에 의한 품질예측 모형들을 개발하여야 한다. 예측모형의 수가 증가하면 필요시 적합한 예측모형을 찾아낼 수 있도록 모형관리시스템(Model Management System, MMS)을 개발하였다.

[그림 15]는 MMS의 주화면으로 학습된 신경망 모형들을 등록하거나 품질설계시물레이션에 적합모형을 찾는데 사용된다. 등록되는 모형들의 주 키는 제품군과 등록일자(연 월 일-시 분 초)만으로 구성하였다. 하지만 원하는 신경망 모형을 찾을 경우에는 제품군만이 필수 항목이며, 등록일을 입력하거나 화면상에 주어진 KIOV들 중에서 하나이상만 선택하면 입력내용에 대응될 수 있는 모든 모형들

을 찾아서 보여준다. 화면의 Grid 상에 표시된 신경망 모형을 더블클릭하면 별도 파일에 저장된 모형의 상세내역을 볼 수 있다. 신경망 학습에 사용된 데이터의 특성, 과거 사용이력 및 평가 등 모형의 신뢰성에 대한 정보를 얻을 수 있으므로, 해당 신경망이 자신이 원하는 모형인가의 여부를 판단할 수 있다.

KIOV	KIOV ID	Model ID	TS ID	Weight	Model Name										
VP	13	EL	CE0	SI	P	F10	delta	SI	40-50	1.4	2.0	50	L	1H	199705

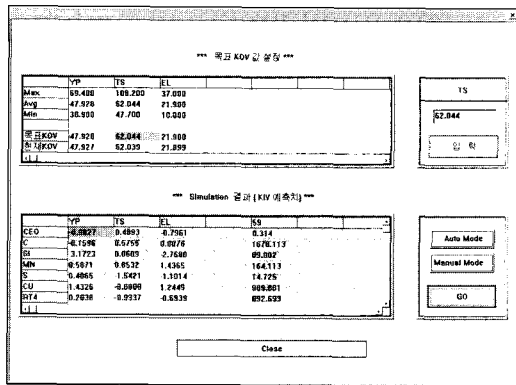
[그림 15] 신경망 모형관리시스템(MMS)

5.4 품질설계시물레이션의 구현

신경망과 CBR의 연계는 MMS를 통하여 이루어진다. MMS 호출시 CBR에서 생성된 설계대안들이 갖는 $S_D(i)$ 및 $S_T(i)$ 의 값들을 MMS에 넘겨주면 적합한 신경망 모형을 찾아서 품질설계시물레이션을 시작할 수 있다. 우선 MMS 내의 Quick Test 기능을 이용하여 $S_D(i)$ 및 $S_T(i)$ 의 KIV들이 목표 KOV 값들을 달성할 수 있는지 최종 생산제품에 대한 품질예측작업을 한다.

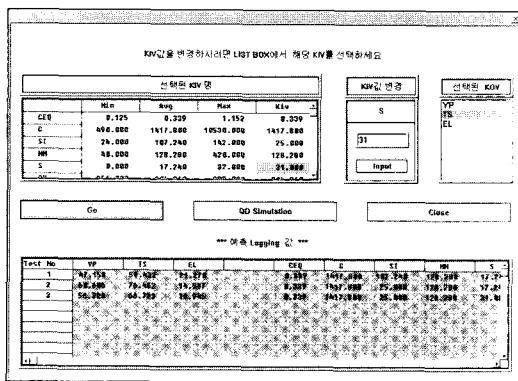
[그림 16]는 QuickTest 화면을 나타낸 것이다. 이 화면에서는 CBR에서 넘겨받은 KIOV들을 이용하여 예측 작업을 수행하면서, 필요시 사용자가 KIV 값들을 변경 입력함으로써 대응되는 KOV 값들의 변화를 확인할 수 있다.

만약, CBR에서 생성된 설계 대안이 목표 KOV 값을 달성하지 못한다면 QDSim 버튼을 이용하여 목표 KOV를 달성할 수 있도록 $S_D(i)$ 내의 KIV 값들을 미세 조정한다.



[그림 16] QDSim 화면

[그림 17]은 QDSim 화면으로 자동과 수동의 두 가지 작업형태를 지원한다. 자동 작업이란 4장에서 논의된 내용대로 KIV 값들을 조정하는데, 만약 조정과정에서 신경망 모형의 특정 KIV가 자신이 학습되었던 학습집합이 가졌던 값의 범위를 벗어날 경우 더 이상의 조정 작업을 수행하지 못하도록 하였다. 이는 신경망의 비선형을 감안할 때 결과의 신뢰성을 보장할 수 없기 때문이다. 이를 보완하기 위하여 수동으로 KIV 값의 미세 조정으로 인한 KOV 값의 변동을 확인할 수 있도록 하였으며, 또 다른 설계 대안을 찾을 수 있도록 목표 KOV 값을 일시 변경한 후, 조정 작업을 계속할 수 있도록 하였다.



[그림 17] QuickTest 화면

MMS를 이용하여 원하는 신경망 모형을 찾아서, Quick Test의 품질예측기능을 통하여 현 품질설계 대안의 실행여부를 검증한 후, QDSim을 이용하여

실제 생산 현장에서 보다 실행가능성이 높은 설계 대안을 찾을 수 있다. 물론 QDSim이 갖고 있는 한계로 인하여 조정과정이 실패할 경우에는 다시 CBR로 돌아가 새로운 설계 대안을 생성한 후, 앞의 과정을 반복 수행함으로써 원하는 품질설계 대안을 얻을 수 있다.

6. 결론 및 향후 연구방향

연구에서는 지능형 품질시스템의 핵심 구성요소인 품질설계지원시스템에 대하여 설명하였다. 사례기반추론을 이용한 선행 연구를 확장하여, 신경망과 사례기반추론을 결합하여 효과적인 품질설계시물레이션 기능을 구현함으로써 프로세스 산업 현장에서 직접적으로 활용 가능하도록 하였다. 특히, 강건한 신경망 모형을 수립할 수 있도록 데이터 마이닝 기반의 신경망 학습시스템을 구축하였는데, $\xi - \delta$ Ball에 의한 데이터 스크리닝, 자동 학습검증기능, 모형관리시스템 기능 등에 대해 중점 설명하였다. 사례기반추론에 의해 생성된 품질설계 대안의 유효성을 검증하고 실제 생산현장에 적합한 대안으로 조정해가는 과정인 신경망을 활용한 품질설계시물레이션에 대해 중점적으로 설명하였다.

품질설계지원시스템의 지속적인 발전을 위하여 보완할 사항들은 다음과 같다. 첫 번째, 신경망의 역할을 확장하는 것이다. 본 연구에서는 신경망을 생산 공정의 모델링 수단으로 이용하여 품질설계의 결과인 화학성분 및 공정조건을 입력으로 하여 사전 품질예측을 수행하는 기능으로 사용하였다. 이 외에도 고객의 주문조회에 대한 일차적 대응수단으로서 고객이 요구하는 품질특성치를 입력으로 하여 조회된 제품의 생산가능 여부를 사전 검토하는 기능에 사용함으로써 필요 없는 품질설계 작업을 사전 차단하는 역할을 하게 할 수 있다. 또, 신경망의 연상기억기능을 이용하여 사례기반추론에서의 유사사례 검색 효율을 강화하는데 사용할 수 있다.

최근 데이터 마이닝의 연구 동향은 Bagging,

Ensemble, Committee 모형 등과 같이 신경망이나 의사결정수를 여러 개 결합하여 하나의 예측 모형을 구성하는 것이 많은 관심을 끌고 있다(Breiman, 1996 ; Ratsch *et al*, 2000 ; Trespo, 2001), 이러한 데이터 마이닝 기법들은 예측의 정확성을 높이는 장점 대신에 모형의 구조에 대한 정보를 잃게 되는 단점이 있다. 따라서 본 연구에서 제시한 품질설계 시뮬레이션을 위해서는 사용이 어렵다.

마지막으로 본 연구의 한계는 품질설계에 있어 비용 측면에 대한 고려가 없었다는 점이다. 현재 구현된 품질설계지원시스템의 사용자들이 품질설계전문가들이라 활용에 큰 문제가 없으나, 앞으로 전문성이 부족한 설계자들에 의해서도 사용되려면 비용최적화도 반드시 해결해야할 향후 연구 과제이다. 이를 위해서는 제약만족기법(CSP, Constraint Satisfaction Programming) 등을 활용하여 품질설계시뮬레이션의 목표 해를 찾아가는 과정을 개선할 필요가 있다.

참 고 문 헌

- [1] Barletta, B., "An Introduction to Case-Based Reasoning," *AI Expert*, Vol.6, No. 8(1991), pp.42-49.
- [2] Bigus, J.P., *Data Mining with Neural Networks*, McGraw Hill, New York, 1996.
- [3] Braun, R.J., "Turning Computers into Experts," *Quality Progress*, (Feb. 1990), pp. 71-75.
- [4] Breiman, L., "Bagging Predictors," *Machine Learning*, Vol.24(1996), pp.123-140.
- [5] Calabrese, C., E. Gnerre and E. Fratesi, "An Expert System for Quality Assurance based on Neural Networks," *Parallel Architectures and Neural Networks, 4th Workshop, Int'l Institute for Advanced Scientific Studies*, Salerno, Italy, (1991), pp.296-300.
- [6] Fard, N.S. and I. Sabuncuoglu, "An Expert System for Selecting Attribute Sampling Plans," *Int'l J. of Computer Integrated Manufacturing*, Vol.3, No.6(1990), pp.364-372.
- [7] Hruska, S.I. and D. Kuncicky, "Application of Two-stage learning to an Expert Network for Control Chart Selection," *Proc. of ANNIE*, (1991), pp.915-920.
- [8] Irgens, C., "RA-IQSE : A System for On-line Quality Support for the Designer of Machined Parts and Products," *Computer Integrated Manufacturing Systems*, Vol.3, No.4(1990), pp.246-251.
- [9] Iwata, Y. and N. Obama, "QDES : Quality Design Expert System for Steel Products," *Innovative Applications of AI*, (1991), pp. 177-191.
- [10] Han, J., and Kamber, M., *Data Mining : Concepts and Techniques*, Morgan Kaufman, 2001.
- [11] Leitch, R.R., R. Kraft and R. Luntz, "RESCU : A Realtime Knowledge-based System for Process Control," *IEE Proceedings, Part D*, Vol.138, No.3(1991), pp.217-227.
- [12] Ogai, H., T. Ueyama, H. Sato, Y. Mishima and S. Itonaga, "An Expert System for Quality Diagnosis of Glass Films on Silicon Steels," *ISIJ International*, Vol.30, No.2 (1990), pp.173-181.
- [13] Owen, B.D., "An Expert System to aid Quality Planning," *4th Int'l Conf. on Computer-aided Production Engineering*, Edinburgh, UK, (Nov. 1988), pp.529-534.
- [14] Oztemel, E., *Integrating Expert Systems and Neural Networks for Intelligent On-line Statistical Process Control*, PhD Thesis, University of Wales, 1992.

- [15] Pham, D.T. and E. Oztemel, "XPC : An On-line Expert System for Statistical Process Control," *Int'l J. of Production Research*, Vol.30, No.12(1992), pp.2857-2873.
- [16] Pham, D.T. and E. Oztemel, *Intelligent Quality Systems*, Springer, 1996.
- [17] Pyle, D., *Data Preparation for Data Mining*, Morgan Kaufman, 1999.
- [18] Rao, H.R. and B.P. Lingaraj, "Expert Systems in Production and Operation Mgt : Classification and Prospects," *Interface*, Vol.18, No.6(1988), pp.80-91.
- [19] Rao V. and Rao, H., *C++ Neural Networks and Fuzzy Logic*, 2nd ed., MIS : Press, 1995.
- [20] Ratsch, G., Schkopf, B., Smola, A., Mller, K.-R. and Mika, S., "nu-arc : Ensemble learning in the presence of outliers," In *Advances in Neural Information Processing Systems*, Vol.12(2000), MIT Press.
- [21] Smith, M., *Neural Networks for Statistical Modeling*, VNR, 1994.
- [22] Suh, M.S., Jhee, W.C., Ko, Y.K., Lee, A., "A Case-based Expert System Approach for Quality Design," *Expert Systems With Applications*, Vol.15(1998), pp.181-190.
- [23] Spur, G. and D. Specht, "Knowledge Engineering in Manufacturing," *Robotics and Computer Integrated Manufacturing*, Vol. 9, No.4/5(1992), pp.303-309.
- [24] Versaggi, M., "Understanding Conflicting Data," *AI Expert*, (1995), pp.21-25.
- [25] Trespo, V., "Committee machines," In *Handbook on neural Network Signal Processing*. CRC Press, 2001.
- [26] Watson, I. and Perera, S., "Case-based Design : A Review and Analysis of Building Design Applications," *AI for Engineering, Design and Manufacturing*, Vol.11, No.1 (1997), pp.59-87.