

論文2000-37CI-1-4

시스템 모델링을 위한 일반화된 RBF 신경회로망의 온라인 구성

(An On-line Construction of Generalized RBF Networks for System Modeling)

權五臣*, 金炯奭**, 崔鐘洙***

(Oh Shin Kwon, Hyong Suk kim, and Jong Soo Choi)

요 약

이 논문에서는 비선형 시스템 모델링을 위한 일반화된 RBF 신경회로망(GRBFN)을 순차적으로 구성하기 위한 온라인 학습 알고리즘을 제안한다. 상수 연결강도를 갖는 표준 RBF 신경회로망의 확장형인 GRBFN은 여러 개의 국부 선형모델을 결합하여 비선형 시스템을 표현할 수 있는 구조이다. 제안한 학습 알고리즘은 구조 학습과 파라미터 학습을 수행하는 두 단계의 학습으로 구성된다. 구조 학습은 주어진 훈련 데이터로부터 새로운 은닉 유닛 및 선형 국부모델을 할당하기 위하여 훈련 오차와 Mahalanobis 거리에 기초한 두 개의 생성 조건을 이용하여 GRBFN 모델을 구성한다. 파라미터 학습은 경사강하 법칙을 기반으로 기존 네트워크의 파라미터 벡터를 갱신한다. 제안한 알고리즘의 모델링 성능을 평가하기 위해서 잘 알려진 두 예제에 대한 시뮬레이션 및 결과를 제시한다.

Abstract

This paper presents an on-line learning algorithm for sequential construction of generalized radial basis function networks (GRBFNs) to model nonlinear systems from empirical data. The GRBFN, an extended form of standard radial basis function (RBF) networks with constant weights, is an architecture capable of representing nonlinear systems by smoothly integrating local linear models. The proposed learning algorithm has a two-stage learning scheme that performs both structure learning and parameter learning. The structure learning stage constructs the GRBFN model using two construction criteria, based on both training error criterion and Mahalanobis distance criterion, to assign new hidden units and the linear local models for given empirical training data. In the parameter learning stage the network parameters are updated using the gradient descent rule. To evaluate the modeling performance of the proposed algorithm, simulations and their results applied to two well-known benchmarks are discussed.

* 正會員, 群山大學校 電氣電子制御工學部

(School of Electrical Eng., Kunsan National Univ)

** 正會員, 全北大學校 情報電子工學部 및 메카트로닉스 연구센터

(School of Electrical Eng and Mechatronics Research Center, Chonbuk National Univ)

*** 正會員 (前) 浦項産業科學院

(Research Institute of Industrial Science and Technology)

接受日字: 1999年3月29日, 수정완료일: 2000年1月3日

I. 서론

제어, 신호처리, 패턴인식 및 시계열 예측 등 많은 실용적인 분야에서 널리 이용되고 있는 RBF(radial basis function) 신경회로망은 생물학 분야에 유래를 두고 있는 다층 신경회로망 모델과 달리 다차원 공간상에서 다변수 함수 보간 이론에 근거를 두고 있으며^[1,2], 임의의 연속함수를 근사화 할 수 있다^[3].

RBF 유닛과 출력단 사이의 상수 연결강도(constant

weights)로 구성된 RBF 신경회로망은 데이터 수가 매우 많을 경우 많은 RBF 유니트를 필요로 하는데, 이러한 문제점을 극복하기 위해서 여러 연구들이 발표되었다. Poggio등^[4]은 저차원 공간에서 탐색을 통한 준최적의 해를 찾는 근사화법을 제안하였고, Moody등^[5]은 RBF 유니트가 자기조직(self-organized) 기법으로 선정되는 방법을 제안하였다. Platt^[6]은 온라인으로 훈련 데이터의 신규성을 기반으로 새로운 RBF 유니트를 할당하는 RAN (resource-allocating network)을 제안하였고, Kadirkamanathan등^[7]은 RAN의 성능을 개선시킨 연구를 발표하였다. 최근 Kuncheva^[8]는 유전 알고리즘을 이용하여 RBF 유니트의 프로토타입을 초기화하는 방법을 제안하고 패턴분류 문제에 적용하였다. 이러한 연구들은 명백히 RBF 중심의 선정 및 계산에 중점을 두고 있다.

한편, Langari등^[9]은 기존 RBF 신경회로망의 복잡성을 저감하는 방안으로 기존 RBF 신경회로망 상수 연결강도를 입력변수의 국부 선형모델로 구성되는 회귀 연결강도(regression weights)로 대체한 일반화된 RBF 신경회로망 모델을 제안하였다. 이 모델은 동작영역에 기저함수를 할당하고 선형모델을 결합하는 LMN(local model networks)^[10] 및 퍼지 규칙의 후건부가 입력 변수의 선형모델로 구성되는 Takagi-Sugno 퍼지 모델^[11]과 유사성이 있는 모델로서 일반화된 RBF 신경회로망으로 생각할 수 있다. 다양한 실험을 통해서 Langari등은 일반화된 RBF 신경회로망을 사용하게 되면 기존 RBFN 신경회로망보다 RBF 유니트 수를 현저히 줄일 수 있음을 밝혔다. 그러나 일괄학습(batch learning)에 의한 K-means 클러스터링 및 Nearest-Neighbor Heuristic을 사용하여 RBF 유니트의 중심과 폭을 결정하고 EM(expectation-maximization) 학습 알고리즘을 통해서 회귀 연결강도만 학습시키는 오프라인 방식이어서 비선형 시스템의 실시간 모델링 및 제어에 적용하기에는 적합하지 않다. 또한 초기에 클러스터의 수를 인위적으로 결정해주어야 하고, 가중지수(weighting exponent)의 최적 선정을 위한 이론적 근거가 결여되어 있다.

이 논문에서는 위의 단점들을 개선하기 위해서 순차적 학습에 의한 일반화된 RBF 신경회로망의 온라인 구성 알고리즘을 제안한다. 제안하는 알고리즘은 훈련 데이터를 입력으로 받아 오차를 기준으로 한 오차 판별기준과 입력과 RBF 중심 사이의 최소 유클리드 거

리를 이용한 거리 판별기준을 통하여 관측 데이터의 신규성을 검출한 후 RBF 유니트의 생성여부를 결정하는 RAN^[6,7]을 기반으로 한다. 그러나 RAN의 신규성 검출에서 유클리드 거리를 이용하는 것은 과도한 RBF 유니트를 생성하여 네트워크 크기를 증가시킬 뿐만 아니라 over-fitting을 유발하여 모델링 성능을 저하시킨다. 따라서 이 논문에서는 이러한 문제점을 개선하기 위하여 Mahalanobis 거리를 이용한 판별기준을 도입한다. 온라인 구성 알고리즘은 RBF 유니트를 생성하고 선형모델 파라미터를 초기화하는 구조 학습 루틴과 RBF 유니트 및 선형모델 파라미터를 최적화하는 파라미터 학습 루틴으로 이루어진다. 제안한 알고리즘의 성능을 검증하기 위해서 시스템 모델링 문제에 적용하고 기존 모델링 결과와 성능을 비교한다.

II. 일반화된 RBF 신경회로망

1. 일반화된 RBF 신경회로망 모델

RBF 신경회로망 모델을 정의하기 위해 모델은 m 개의 입력과 한 개의 출력을 갖는다고 가정하자. 그러면 RBF 신경회로망은 입력공간 R^m 의 입력 패턴을 출력공간 R^1 로의 비선형 사상을 다음 식과 같이 수행한다.

$$y(x) = \sum_{i=1}^L y_i(x) = \sum_{i=1}^L w_i(x) \phi_i(x) \quad (1)$$

여기에서 $x \in R^m$ 은 입력벡터이고, $w_i(\cdot)$ ($i=1, 2, \dots, L$)는 RBF 유니트와 출력 유니트 사이의 연결강도이고, $\phi_i(\cdot)$ 는 RBF 유니트의 응답이다. RBF 유니트는 다음과 같은 다양한 함수들이 고려될 수 있다^[12].

- Thin-plate-spline 함수 : $\phi(z, \sigma) = z^2 \log(z)$
- Multiquadric 함수 : $\phi(z, \sigma) = (z^2 + \sigma^2)^{1/2}$
- Inverse multiquadric 함수 : $\phi(z, \sigma) = \frac{1}{(z^2 + \sigma^2)^{1/2}}$
- Gaussian 함수 : $\phi(z, \sigma) = \exp(-z^2/\sigma^2)$

식 (1)에서 연결강도 w_i 가 상수값을 갖는 경우를 표준 RBF 신경회로망(standard RBF networks, 이하 RBFN으로 표기)이라 정의한다. 이 논문에서 고려하는 일반화된 RBF 신경회로망(generalized RBF networks, 이하 GRBFN으로 표기)은 w_i 가 상수 연결강도를 갖는 RBFN과는 달리 입력벡터의 선형결합으로 이루어진 국

부모델을 갖는다. 즉 입력변수의 국부 선형모델 함수인 w_i 는 다음과 같이 표현된다.

$$w_i(\mathbf{x}) = a_{i0} + a_{i1}x_1 + \dots + a_{im}x_m = \mathbf{a}_i^T \mathbf{x} \quad (2)$$

여기에서, $\mathbf{a}_i^T = [a_{i0}, a_{i1}, \dots, a_{im}]$ 은 회귀 파라미터 벡터이고, $\mathbf{x} = [1, x_1, \dots, x_m]^T$ 는 확장된 입력벡터이다. 따라서 GRBFN의 네트워크 구조는 그림 1과 같은 형태를 갖으며 출력은 다음 식으로 주어진다.

$$y(\mathbf{x}) = \sum_{i=1}^L w_i(\mathbf{x}) \phi_i(\mathbf{x}; c_i, \sigma_i) = \sum_{i=1}^L (\mathbf{a}_i^T \mathbf{x}) \phi_i(\mathbf{x}; c_i, \sigma_i) \quad (3)$$

GRBFN의 정규화된 형태는 다음과 같이 표현된다.

$$y(\mathbf{x}) = \frac{\sum_{i=1}^L (\mathbf{a}_i^T \mathbf{x}) \phi_i(\mathbf{x}; c_i, \sigma_i)}{\sum_{i=1}^L \phi_i(\mathbf{x}; c_i, \sigma_i)} \quad (4)$$

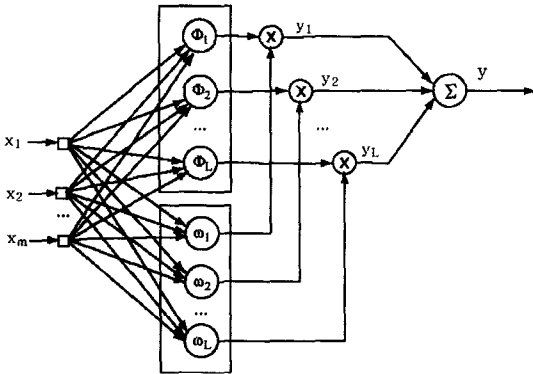


그림 1. 일반화된 RBF 신경회로망(GRBFN)의 구조
Fig. 1. An architecture of the GRBFN.

이 논문에서 고려하는 RBF 유니트는 Gaussian 함수로 각 입력 차원에 대하여 동일한 폭을 갖는다면, 즉 $\sigma_1 = \sigma_2 = \dots = \sigma_m = \sigma_i$ 라 하면 각 유니트의 활성화 함수는 다음과 같이 간략하게 표현할 수 있다.

$$\phi_i(\mathbf{x}; c_i, \sigma_i) = \exp\left[-\frac{\|\mathbf{x} - c_i\|^2}{2\sigma_i^2}\right] \quad (5)$$

여기에서 c_i 와 σ_i 는 각각 i 번째 RBF 유니트에 대한 중심벡터와 폭을 나타내고, $\|\cdot\|$ 는 유클리드 norm

을 의미한다.

GRBFN은 복잡한 비선형 시스템을 국부적으로 동작하는 여러 개의 모델로 분해하는 특징이 있어 표준 RBFN보다 은닉 유니트의 수를 현저히 줄일 수 있는 장점을 갖고 있다. Langari등^[9]은 GRBFN을 구성하기 위해서 일괄학습의 K-means 알고리즘을 사용한다. 즉 오프라인으로 RBF의 중심 및 폭을 결정한 후 회귀 함수의 파라미터들만 EM 알고리즘을 사용하여 학습시킨다. 따라서 이 모델 구성 방법은 대상 시스템에 대한 모델을 구하는데 있어서 전적으로 설계자의 판단에 의존해야 하며 온라인 모델링 및 제어에는 적합하지 않다.

2. 다른 비선형 모델과의 연관성

1) LMN(Local Model Networks)

Murray-Smith등^[10]에 의해 표현된 LMN은 입력공간을 동작점에 따라 여러 영역으로 나누어 각 영역에 기저함수를 할당하고 이에 연결되는 여러 개의 단순 모델들을 결합하는 모델이다. 전체적으로 복잡한 모델을 생성하기 위해 다음과 같이 결합된다.

$$y = \sum_{i=1}^L f_i(\mathbf{x}) \rho_i(v) \quad (6)$$

여기에서 v 는 시스템의 동작점(operating point)을 정의한다. 그림 2는 2차원 공간상에서의 동작영역을 나타낸다. 그림 2에서 각 국부모델 $f_i(\mathbf{x})$ 는 동작영역과 결합되고, 영역들간의 중첩은 모델 사이의 보간을 제공한다. 기저함수 $\rho_i(v)$ 는 입력공간에 대하여 partition of unity를 형성한다. 기저함수는 partition of unity를 얻기 위해서 다음과 같이 정규화된다.

$$\rho_k(v) = \frac{\rho(d(v; c_k, \sigma_k))}{\sum_{i=1}^L \rho(d(v; c_i, \sigma_i))} \quad (7)$$

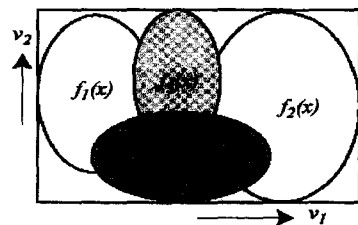


그림 2. LMN의 동작 영역
Fig. 2. Operating boundary of LMN.

여기에서 $\rho(\cdot)$ 는 정규화되지 않은 기저함수이고, $d(v; c_k, \sigma_k)$ 는 σ_k 의 폭을 가진 기저함수의 중심 c_k 로부터 현재의 동작점 v 간의 거리를 측정하는 거리함수이다. 따라서 국부모델 $f_i(\mathbf{x})$ 가 입력의 선형모델로 구성되고, 위에 언급한 RBF 함수를 기저함수로 사용한다면, LMN은 GRBFN으로 해석될 수 있다.

2) Takagi-Sugeno-Kang 퍼지 모델

적은 수의 퍼지 규칙으로 비선형 시스템을 잘 표현할 수 있는 Takagi-Sugeno-Kang (TSK) 퍼지 모델은 규칙의 후건부가 입력변수의 선형회귀 함수로 구성되어 있다. TSK 퍼지 모델의 i 번째 규칙은 다음과 같이 표현된다.

$$R_i: \text{IF } x_1 \text{ is } A_{i1} \text{ and, } \dots, \text{ and } x_m \text{ is } A_{im} \\ \text{THEN THEN } f_i = a_0 + a_{i1}x_1 + \dots + a_{im}x_m \quad (8)$$

여기에서 A_{ij} 는 퍼지 집합이고, a_{ij} 는 조정되는 후건부 파라미터이며, f_i 는 i 번째 규칙의 출력이다. L 개의 규칙으로 구성되는 TSK 퍼지 모델의 전체 출력은 정규화된 형태로 다음과 같이 나타낼 수 있다.

$$y(\mathbf{x}) = \frac{\sum_{i=1}^L h_i(\mathbf{x}) f_i(\mathbf{x})}{\sum_{i=1}^L h_i(\mathbf{x})} \quad (9)$$

여기에서 $h_i(\mathbf{x})$ 는 입력에 대한 i 번째 규칙의 전건부 발화강도(firing strength)를 의미한다. 입력변수에 대한 멤버십 함수로 Gaussian 함수를 사용하고, 각 규칙의 발화강도를 계산하기 위해 곱(multiplication) 연산자를 사용하면 i 번째 규칙의 발화강도는 다음과 같다.

$$h_i(\mathbf{x}) = \prod_{j=1}^m \mu_{ij}(x_j) \quad (10)$$

Hunt^[11]이 기술한 TSK 퍼지 모델과 GRBFN과의 등가조건에 의거하면 식 (10)의 발화강도는 GRBFN의 비선형 활성화 함수인 식 (5)와 기능적으로 등가이다. 즉 $h_i(\mathbf{x}) = \phi_i(\mathbf{x})$ 인 관계가 성립한다. 결과적으로 몇 가지 전제조건 하에서 TSK 퍼지 모델과 GRBFN은 등가임을 알 수 있다.

III. GRBFN의 온라인 구성 알고리즘

미지의 구조 및 파라미터 정보를 갖는 실시간 시스템의 제어는 신경회로망과 같은 비파라미터화된 기법을 이용하여 온라인 또는 순차 모델링(또는 식별 (identification))을 수행하는 것을 기반으로 할 수 있다. 시스템의 온라인 모델링 문제는 요구되는 모델링 오차와 이전의 모델링 구조 및 온라인 관측 데이터가 주어지면, 모델 파라미터 적응 법칙 또는 신경회로망 근사화를 얻기 위해 이러한 정보들을 결합하는 것이다.

이 절에서는 주어진 관측 데이터로부터 시스템의 GRBFN 모델링을 온라인으로 구성하기 위한 학습 알고리즘을 기술한다. GRBFN의 온라인 구성 알고리즘은 순차적으로 입력되는 관측 데이터의 신규성을 판단하여 RBF 유닛을 생성시키고 새로운 유닛에 대응하는 선형모델을 할당한다. 데이터의 신규성 판단 기준은 RAN 알고리즘^[6,7]에 기반을 두고 있다. RAN은 관측된 훈련 데이터에 대한 오차와 RBF 유닛 사이의 거리를 신규성 판단 기준으로 사용한다. 그러나, RAN의 거리 판단기준은 각 RBF 유닛들의 동작영역을 고려하지 않고 단지 입력벡터와 RBF 유닛 중심간의 유클리드 거리를 측정하기 때문에 과도한 유닛을 생성하는 단점이 있고, 이로 인하여 필요이상으로 네트워크 크기가 증가하게 되어 over-fitting을 유발시킬 수 있다. 따라서 RAN의 거리 판별기준의 단점을 개선하는 방법을 제안하고, GRBFN의 온라인 학습 알고리즘에 대하여 상세하게 기술한다.

1. 관측 데이터의 신규성 판별

RAN은 네트워크에 새로운 RBF 유닛을 할당하기 위해 다음의 두 조건을 신규성 판별기준으로 이용한다^[6,7].

$$|e_n| = y_n^d(\mathbf{x}_n) - y_n(\mathbf{x}_n) > e_{\min} \quad (11)$$

$$\|\mathbf{x}_n - \mathbf{c}_{nr}\| > \delta_n \quad (12)$$

여기에서, n 번째 관측된 패턴에 대하여 e_n 은 목표 출력 y_n^d 와 GRBFN의 출력 y_n 사이의 오차를 나타내고, \mathbf{c}_{nr} 은 입력벡터 \mathbf{x}_n 과 가장 가까운 RBF 유닛의 중심벡터이다. 또한 e_{\min} 은 모델의 정확도를 나타내는 각 패턴에 대한 허용오차이고, δ_n 은 입력공간의 분해

능 스케일을 나타내는 거리로서 적절히 선택되어야 할 문턱값들이다. RAN 알고리즘은 위의 두 판별기준을 만족하는 경우에만 새로운 유니트를 할당하고, 그렇지 않으면 기존 네트워크의 파라미터를 최적화한다. 식 (12)는 관측된 입력과 RBF 유니트 사이의 거리를 판단하여 새로운 유니트의 생성여부를 결정하는 것으로 거리 δ_n 은 non-zero의 확률분포를 갖는 입력공간의 스케일 크기인 δ_{max} 로 출발하여 관측한 패턴의 수가 증가함에 따라 점점 줄어들게 된다. 즉, 초기에 $\delta_n = \delta_{max}$ 로 시작하여 다음 식과 같이 δ_{min} 에 도달할 때까지 감소한다.

$$\delta_n = \text{MAX} \{ \delta_{max} \gamma^n, \delta_{min} \} \quad (13)$$

여기에서 $\gamma(0 < \gamma < 1)$ 는 감쇄상수이다.

식 (12)에서와 같이 유클리드 거리를 신규성 판별에 사용할 경우 과도한 RBF 유니트를 생성시키는 단점이 있다. 즉, 입력벡터와 유니트 중심사이의 유클리드 거리를 고려하기 때문에 현재의 입력에 대하여 충분히 동작하는 유니트가 있음에도 불구하고 새로운 유니트를 생성시키는 결과를 초래한다. 예를 들면, 그림 3과 같은 1차원 입력의 경우 임의의 시간 t 에서 입력공간에 이미 할당된 RBF 유니트가 3개라고 가정하자. 입력 $x(t)$ 에 대하여 각 유니트의 중심과의 유클리드 거리를 $d_i(t) (i=1, 2, 3)$ 라 하면,

$$0 < d_2(t) < d_1(t) < d_3(t) \quad (14)$$

인 관계가 성립한다. 거리 상으로 입력 $x(t)$ 와 가장 가까운 유니트는 $\phi_2(x)$ 이지만, 활성화 정도를 고려하면

$$0 < \phi_2(x(t)) < \phi_1(x(t)) \quad (15)$$

임을 알 수 있다. 이 시점에서 최소 유클리드 거리 $d_2(t)$ 와 $\delta_n(t)$ 의 관계가

$$d_2(t) < \delta_n(t) < d_1(t) \quad (16)$$

이라면, 입력 $x(t)$ 에 대하여 충분히 동작하는 유니트 $\phi_1(x)$ 이 있음에도 불구하고 새로운 유니트를 추가해야 한다. 이 경우, RAN 알고리즘은 과도한 유니트를 생성하게 되고, 이로 인하여 over-fitting을 초래한다.

유클리드 거리 판별의 단점을 해결하기 위해 이

논문에서는 Mahalanobis 거리^[13]를 도입하여 RBF 유니트의 동작영역(operating boundary)을 정의한다. Mahalanobis 거리 d_M 은 RBF 유니트가 식 (5)와 같은 형태를 갖는 경우 i 번째 유니트에 대하여 다음과 같이 정의된다.

$$d_M^i(x) = \left(\frac{\| \mathbf{x} - \mathbf{c}_i \|^2}{2\sigma_i^2} \right)^{1/2} \quad (17)$$

위 식에서 보는 바와 같이 Mahalanobis 거리는 유클리드 거리와 달리 입력과 RBF 유니트 사이의 거리 및 폭을 고려하고 있음을 알 수 있다. δ_M 을 RBF 유니트의 동작영역을 결정하는 Mahalanobis 문턱값으로 정의하면, 입력공간에서 입력 x 에 대하여 L 개의 RBF 유니트에 대한 각 유니트의 동작영역은 초구체(hyper-sphere) $H_i(\delta_M)$ 로서 다음 식과 같이 정의할 수 있다.

$$H_i(\delta_M) = \left\{ x \mid (d_M^i(x))^2 = \frac{\| \mathbf{x} - \mathbf{c}_i \|^2}{2\sigma_i^2} \leq (\delta_M)^2 \right\} \quad (18)$$

for $(i=1, \dots, L)$

입력 x 에 대하여 i 번째 유니트가 식 (18)을 만족한다면 유니트의 동작영역 범위 내에 x 가 포함되어 있음을 의미한다. 결과적으로 식 (18)을 만족하는 것은

$$\phi_i(x) = \exp \left(- \frac{\| \mathbf{x} - \mathbf{c}_i \|^2}{2\sigma_i^2} \right) \leq \exp \left(- (\delta_M)^2 \right) \quad (19)$$

를 만족하는 것과 같다. 따라서 δ_M 이 커지면 각 유니트의 동작영역이 넓어지고, 반대로 δ_M 이 작아지면 동작영역이 좁아지는 효과가 있다.

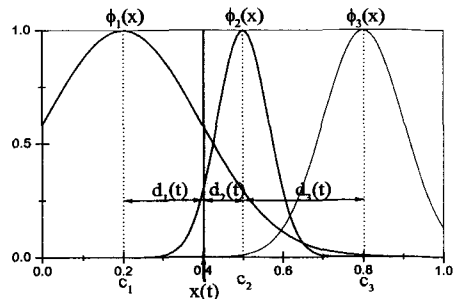


그림 3. 유클리드 거리 판단기준의 단점을 설명한 예 Fig. 3. A description for a shortcoming of Euclidean distance criterion.

제안하는 Mahalanobis 거리를 기반으로한 훈련 데이터의 신규성 판별기준은 다음과 같이 정의된다.

$$\epsilon_M > 0 \quad (20)$$

여기에서 ϵ_M 은 다음과 같이 표현된다.

$$\epsilon_M = \begin{cases} 0, & \text{if } x \in H_i(\delta_M) \text{ for } i=1, \dots, L \\ 1, & \text{Otherwise} \end{cases} \quad (21)$$

따라서, 식 (12)로 정의되는 RAN의 판별기준은 제안하는 식 (20)으로 대체된다. Mahalanobis 거리를 도입한 판별기준은 입력공간 상에서 이미 존재하는 RBF 유니트의 동작영역을 고려하여 새로운 유니트의 생성을 조절하므로 과도한 유니트가 생성되는 것을 방지하여 over-fitting을 피할 수 있고, 더욱 압축된 크기의 GRBFN 모델을 구성할 수 있는 장점이 있다.

2. 학습 알고리즘

GRBFN의 온라인 구성 알고리즘은 훈련 데이터의 신규성을 판별하여 RBF 유니트를 생성하고 이에 대응하는 선형모델 파라미터를 초기화하는 모델구조 학습과 RBF 유니트 및 선형모델 파라미터를 학습규칙에 의해서 최적화하는 파라미터 학습으로 구성된다.

GRBFN은 초기에 RBF 유니트를 갖지 않고 훈련을 시작한다. 모델구조 학습을 통해서 관측 데이터들이 입력됨에 따라 새로운 유니트를 추가함으로써 자동으로 네트워크 구조가 결정된다. 관측 데이터 (\mathbf{x}_n, y_n^d)에 대하여 다음의 두 판별기준을 만족하면 새로운 RBF 유니트를 생성한다.

$$\text{조건 1 : } |e_n| = y_n^d(\mathbf{x}_n) - y_n(\mathbf{x}_n) > e_{\min} \quad (22)$$

$$\text{조건 2 : } \epsilon_M > 0 \quad (23)$$

현재의 네트워크이 L 개의 RBF 유니트로 구성되어 있다면 새로 생성되는 $L+1$ 번째 유니트의 파라미터들은 다음과 같이 초기화된다.

$$\mathbf{c}_{L+1} = \mathbf{x}_n \quad (24)$$

$$\sigma_{L+1} = x \|\mathbf{x}_n - \mathbf{c}_{nr}\| \quad (25)$$

$$\mathbf{a}^{(L+1)i} = \begin{cases} e_n, & \text{if } i=0 \text{ for } i=0,1,\dots,m \\ 0, & \text{Otherwise} \end{cases} \quad (26)$$

여기에서 x 는 입력공간에서 RBF 유니트 응답의 중첩 정도를 결정하는 요소이고, $\|\mathbf{x}_n - \mathbf{c}_{nr}\|$ 는 입력 벡터와 가장 가까운 거리에 있는 유니트 중심 사이의 유클리드 거리이다.

관측된 데이터에 대하여 새로운 RBF 유니트 생성을 위한 판별기준을 만족하지 못하면 네트워크는 파라미터 학습 과정을 통해서 최적화된다. n 번째 패턴에 대한 오차함수는 다음 식으로 정의된다.

$$E_n = \frac{1}{2} [y_n^d - y_n]^2 \quad (27)$$

편의상 이하의 파라미터 표기에서 아래첨자 n 은 생략한다. 학습 목적은 식 (27)의 오차함수를 최소화하기 위해 네트워크 파라미터 벡터를 온라인으로 최적화하는 것이다. 관측 데이터 (\mathbf{x}_n, y_n^d)가 식 (22) 및 (23)의 판별기준을 만족하지 않는 경우에는 새로운 유니트를 생성하지 않고 파라미터 벡터 $\mathbf{v} = [\mathbf{c}_1^T, \dots, \mathbf{c}_L^T, \sigma_1^T, \dots, \sigma_L^T, \mathbf{a}_1^T, \dots, \mathbf{a}_L^T]^T$ 를 다음과 같이 갱신한다.

$$\mathbf{v}(t+1) = \mathbf{v}(t) + \eta \Delta \mathbf{v}(t) \quad (28)$$

여기에서 η 는 학습률이다. 모델이 L 개의 RBF 유니트를 갖는다면, j 번째 입력에 대한 i 번째 RBF 유니트의 중심, 폭 및 k 번째 선형모델 파라미터 벡터에 대한 변화량 $\Delta \mathbf{v} = [\Delta c_{ji}, \Delta \sigma_{ji}, \Delta a_{ji}]$ 는 식 (27)에 정의된 오차함수에 대하여 경사강하 알고리즘을 적용하여 다음과 같이 계산된다.

$$\Delta c_{ji} = - \frac{\partial E}{\partial c_{ji}} = (y^d - y) (w_i - y) \frac{\phi_i}{\sum_{i=1}^L \phi_i} \frac{(x_j - c_{ji})}{(\sigma_{ji})^2} \quad (29)$$

$$\Delta \sigma_{ji} = - \frac{\partial E}{\partial \sigma_{ji}} = (y^d - y) (w_i - y) \frac{\phi_i}{\sum_{i=1}^L \phi_i} \frac{(x_j - c_{ji})^2}{(\sigma_{ji})^3} \quad (30)$$

$$\Delta a_{ji} = - \frac{\partial E}{\partial a_{ji}} = (y^d - y) \frac{\phi_i}{\sum_{i=1}^L \phi_i} x_j \quad (31)$$

제안한 GRBFN의 전체 학습 알고리즘을 단계별로 요약하면 다음과 같다.

STEP 1 : 초기화

모델 설계시 사용되는 문턱값 및 파라미터 e_{\min} ,

δ_M, x, σ_0 와 학습률 η 를 초기화한다.

STEP 2 : 구조 및 파라미터 학습

STEP 2.1 : RBF 유닛 생성여부 판별

입력되는 훈련 데이터에 대하여 RBF 유닛 생성을 위한 판별기준인 식 (22), (23)을 만족하면 STEP 2.2로 가고, 그렇지 않으면 STEP 2.3으로 간다.

STEP 2.2 : RBF 유닛 생성

새로운 RBF 유닛을 생성하고 파라미터 벡터를 식 (24), (25) 및 (26)으로 초기화한 다음 STEP 2.4로 간다.

STEP 2.3 : 파라미터 학습

식 (29), (30) 및 (31)의 학습 규칙을 이용하여 파라미터 벡터를 갱신하고 STEP 2.4로 간다.

STEP 2.4 : 훈련 데이터 평가

모든 훈련 데이터가 입력되었으면 STEP 3으로 가고, 그렇지 않으면 STEP 2로 되돌아간다.

STEP 3 : 성능 평가

성능을 평가하여 목표 성능지수에 도달하거나 설정한 반복 학습회수를 모두 수행했으면 학습을 종료하고, 그렇지 않으면 STEP 2로 되돌아간다.

IV. 모델링 실험

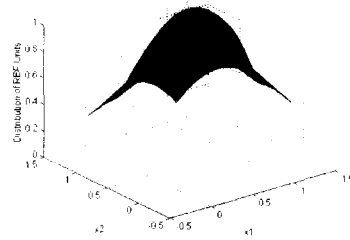
이 절에서는 제안한 GRBFN의 온라인 구성 알고리즘의 유용성을 설명하기 위해서 Zimmerman & Zysno의 실험 데이터^[14-18]와 Box & Jenkins의 가스로(gas furnace) 실험 데이터^[18-24]에 대한 모델링 실험을 수행하고 그 결과에 대하여 기술한다.

1. Zimmermann & Zysno의 실험 데이터

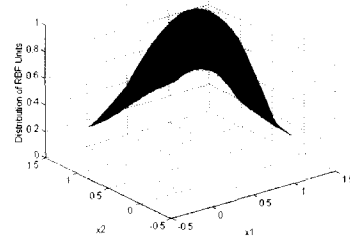
Zimmerman & Zysno의 실험 데이터^[14]는 Dyckhoff & Pedrycz^[15], Krishnapuram & Lee^[16], Langari & Wang^[17] 및 Delgado등^[18]의 많은 문헌에서 모델링 실험 예제로 사용되었다. 훈련에 사용되는 데이터는 두 개의 입력 (x_1, x_2)과 한 개의 목표출력 (y^d)으로 구성된 24개의 데이터 집합이다. 성능 평가지표 PI(performance index)는 다른 실험들과 동일하게 다음과 같이 표현되는 제곱오차를 사용한다.

$$PI = \sum_{j=1}^{24} [y^d(j) - y(j)]^2 \quad (32)$$

여기에서 $y(j)$ 는 GRBFN의 출력이다. 실험에 사용된



(a)



(b)

그림 4. Zimmermann & Zysno의 데이터에 대한 RBF 유닛의 분포
(a) 유닛이 2개인 경우
(b) 유닛이 3개인 경우

Fig. 4. Distribution of RBF units for Zimmermann & Zysno's data.
(a) Case of two unit.
(b) Case of three unit.

표 1. Zimmermann & Zysno의 실험 데이터에 대한 성능 비교

Table 1. Performance for the Zimmermann & Zysno's empirical data.

	규칙 수	PI
Zimmermann & Zysno[14]	-	0.080
Dyckhoff & Pedrycz[15]	-	0.066
Krishnapuram & Lee[16]	-	0.064
Langari & Wang[17]	4	0.043
Delgado etc.[18]	4	0.040
제안한 모델(GRBFN)	2	0.044
	3	0.033

실제 파라미터들은 $e_{min} = 0.02$, $\delta_M = 1.75$, $x = 0.87$ 및 학습률 $\eta = 0.02$ 로 설정하였다. 실험결과는 표 1과 그림 4에 제시하였다. 최초로 생성되는 RBF 유닛의 폭 σ_0 를 0.25로 한 경우, 최종적으로 생성된 유닛 수

는 2개이고 그 때의 PI는 0.044로 Langari 등의 결과인 0.043과 유사하다. 이 결과는 표 1에 제시한 바와 같이 Langari 등이 4개의 유니트(퍼지 규칙의 수)를 사용한 것에 비해 제안한 GRBFN는 2개의 유니트를 가지고 유사한 결과를 보임을 알 수 있다. 이 실험에 대한 GRBFN의 입력공간 상의 RBF 유니트 분포는 그림 4(a)에 나타냈다. 유니트의 폭 σ_0 를 0.3으로 한 경우 생성된 RBF 유니트 수는 3개이고 이 때의 PI는 0.033 이었고, 입력공간 상에서 RBF 유니트는 그림 4(b)와 같이 분포하였다. 결과적으로 제안한 GRBFN은 다른 모델에 비해 더 적은 유니트(또는 퍼지 규칙)를 가지고 더 정밀한 성능을 얻을 수 있음을 확인하였다.

2. Box & Jenkins의 가스로 데이터

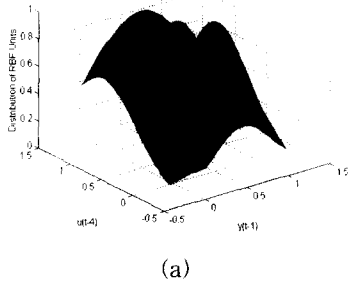
Box & Jenkins에 의해 비선형 시스템 모델링 예제로 널리 알려진 가스로(gas furnace) 모델링^[18, 24]에 적용하고 제안한 모델과 기존 모델과의 성능을 비교한다. 가스로 데이터는 gas flow rate인 입력 $u(t)$ 와 CO_2 concentration인 출력 $y(t)$ 로 구성되며, 훈련에 사용되는 데이터는 9초 간격으로 수집된 296개의 입출력 관측 데이터 쌍으로 구성된다. 모델링 입력은 실험에서 많이 사용하는 $y(t-1)$, $u(t-4)$ 로 선정하였다. 성능 평가를 위해서 N 개의 데이터 포인트에 대하여 다음 식으로 표현되는 MSE(mean square error)를 사용한다.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i^d - y_i)^2 \tag{33}$$

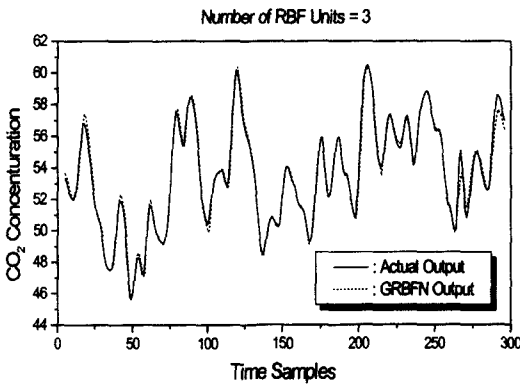
여기에서 y_i^d 는 목표 출력이고, y_i 는 GRBFN 모델 출력이다. 실험의 편의를 위해 이 논문에서는 입력 및 출력 데이터를 [0, 1]로 정규화하여 사용한다. 가스로 모델링에 사용된 설계 파라미터들은 $e_{min} = 0.02$, $\alpha = 0.87$, $\eta = 0.02$ 로 설정하고, Mahalanobis 거리의 문턱값 δ_M 은 1.75와 2.0에 대하여 실험을 수행하였다. 최초로 생성되는 유니트의 폭 σ_0 은 모두 0.2로 하였다. δ_M 이 1.75인 경우, 2000회 동안 학습시킨 후 생성된 GRBFN 모델은 3개의 RBF 유니트를 갖는 네트워크를 구성하였으며 이때의 MSE는 0.149였다. 모델링 실험 결과는 그림 5에 나타냈다. 그림 5(a)는 입력공간 상에서의 RBF 유니트 분포를 나타냈고, 그림 5(b)는 가스로 데이터와 GRBFN에 의해서 모델링된 결과를 각각 도시하였다.

그림 6은 δ_M 이 2.0인 경우에 대한 RBF 유니트의 분포 및 모델링 결과로서 생성된 GRBFN 모델은 2개의 유니트를 갖는 네트워크를 구성하였으며, 이때의 MSE는 0.158이었다. 실험으로부터 δ_M 값이 증가하면 RBF 유니트의 동작영역이 넓어지기 때문에 더 작은 개수의 유니트가 생성됨을 알 수 있다.

제안한 GRBFN과의 모델링 성능을 비교하기 위해 표 2에 기존의 모델링 실험 결과를 나타냈다. 인용된 모델은 퍼지 모델을 사용한 것으로서 A 형태를 갖는 모델은 퍼지 모델의 후건부가 퍼지 집합으로 표현되는 모델이고, B 형태를 갖는 모델은 TSK 퍼지 모델로서 2절에서 언급한 GRBFN과 TSK 퍼지 모델의 동가성을 고려해볼 때 동일한 모델로 생각할 수 있다. 제안한



(a)



(b)

그림 5. 가스로 데이터에 대한 모델링 결과($\delta_M=1.75$ 인 경우)

- (a) RBF 유니트의 분포
- (b) 모델링 출력

Fig. 5. Modeling results for gas furnace data(Case to $\delta_M=1.75$)

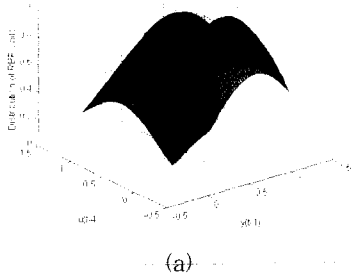
- (a) Distribution of RBF unit.
- (b) Output of Modeling.

안한 GRBFN은 어느 정도의 성능을 보장하면서 빠른 시간안에 비선형 시스템에 대한 모델링을 수행하고 있음을 확인하였다.

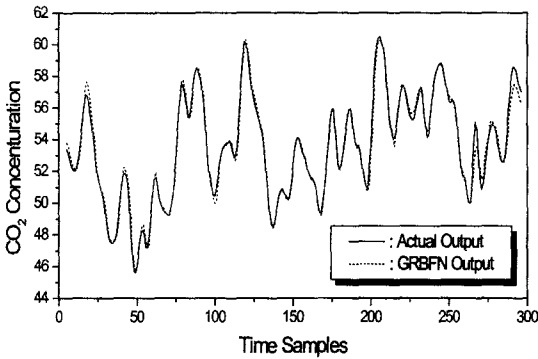
표 2. Box-Jenkins 가스로 데이터에 대한 성능 비교

Table 2. Performance for the Box-Jenkins gas furnace data.

	모델형태	입력 수	규칙 수	MSE
Tong[19]	A	2	19	0.469
Pedrycz[20]	A	2	81	0.320
Sugeno[21]	B	2	2	0.359
Sugeno[22]	A	3	6	0.190
Wang[23]	B	2	5	0.158
Delgado[18]	B	2	4	0.410
Farag[24]	A	2	37	0.111
제안한 모델 (GRBFN)	B	2	2	0.158
			3	0.149



(a) Number of RBF Units = 2



(b)

그림 6. 가스로 데이터에 대한 모델링 결과 ($\delta_M=2.0$ 인 경우)

- (a) RBF 유닛의 분포
- (b) 모델링 출력

Fig. 6. Modeling results for gas furnace data(Case to $\delta_M=2.0$)

- (a) Distribution of RBF unit.
- (b) Output of Modeling.

GRBFN의 성능은 동일한 입력 수를 갖고 B 형태를 취하고 있는 Wang등^[23]의 모델링 결과보다 규칙 수(또는 RBF 유닛 수) 및 MSE면에서 우수함을 알 수 있다. Farag등^[24]의 결과는 A 형태를 갖는 퍼지 모델로써 신경회로망과 유전 알고리즘을 사용한 복잡 퍼지 모델로서 37개의 퍼지 규칙으로 0.111의 MSE를 보이고 있다. 이 결과는 모델링 성능면에서 GRBFN보다 우수하지만 매우 많은 규칙(또는 네트워크 파라미터)을 필요로 한다. 또한 이 결과는 Pentium 166MHz 상에서 C++ 코드로 구현된 프로그램을 약 124분간 수행한 결과로서 학습에 매우 많은 시간이 소요되므로 실시간 모델링에는 적합하지 않음을 알 수 있다. 반면에 제안한 모델은 Pentium 150MHz 상에서 C++ 코드로 구현된 프로그램으로 80초 동안에 모델링을 수행하였다. 결과적으로 제

IV. 결론

이 논문에서는 선형모델을 갖는 일반화된 RBF 신경회로망(GRBFN)을 온라인으로 구성하는 알고리즘을 제안하였다. 제안한 알고리즘은 모델 설계시 인위적으로 클러스터 수(또는 네트워크 크기)를 결정해 주어야 하고 오프라인으로 모델을 구성해야 하는 Lagari 등^[9]이 제안한 네트워크의 단점을 개선한 것이다. 제안한 GRBFN의 학습 알고리즘은 RAN에 기초를 두고 있으나 훈련 데이터의 신규성을 판별하기 위해 유클리드 거리를 사용함으로써 인해서 발생하는 과도한 유니트 생성 문제를 Mahalanobis 거리를 이용한 판별기준을 제안하여 개선하였다. GRBFN은 훈련 데이터로부터 실시간으로 RBF 유닛을 생성하고 선형모델 파라미터를 초기화하는 구조 학습과 네트워크 파라미터 벡터를 최적화하는 파라미터 학습을 통해서 모델이 자동으로 구성된다. 제안한 알고리즘의 유효성을 평가하기 위해서 Zimmerman & Zysno의 실험 데이터 및 Box & Jenkins의 가스로 데이터 모델링에 적용하였다. 실험 결과로부터 제안한 GRBFN의 성능은 기존의 모델보다 더 적은 네트워크 크기로 더 우수한 모델링 성능을 보였고 모델 구성에 소요되는 시간도 매우 빠름을 확인하였다. 따라서 제안한 알고리즘은 비선형 시스템의 실시간 모델링 및 제어 문제에 유용하게 적용될 수 있을 것으로 기대된다.

참 고 문 헌

- [1] D. S. Broomhead and D. Lowe, "Multivariate functional interpolation and adaptive networks," *Complex Systems*, vol. 2, pp. 321-355, 1988.
- [2] M. J. D. Powell, "Radial basis functions for multivariable interpolation: A review," in *IMA Conf. Approximation Functions Data*, pp. 143-167, Shrivvenham, UK, 1985.
- [3] J. Park and J. W. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Computation*, vol. 3, pp. 246-257, 1991.
- [4] T. Poggio and F. Girosi, "Regularization algorithms for learning that are equivalent to multilayer networks," *Science*, vol. 247, pp. 978-982.
- [5] J. Moody and C. Darken, "Fast learning in networks of locally-tuned processing units," *Neural Computation*, vol. 1, pp. 281-294, 1989.
- [6] J. Platt, "A resource-allocating network for function approximation," *Neural Computation*, vol. 3, pp. 213-225, 1991.
- [7] V. Kadirkamanathan and M. Niranjan, "A function estimation approach to sequential learning with neural networks," *Neural Computation*, vol. 5, pp. 954-975, 1993.
- [8] L. I. Kuncheva, "Initializing of an RBF network by a genetic algorithm," *Neuro-computing*, vol. 14, pp. 273-288, 1997.
- [9] R. Langari, L. Wang, and J. Yen, "Radial basis function networks, regression weights, and the expectation-maximization algorithm," *IEEE Transactions on Systems, Man, and Cybernetics - Part A*, vol. 27, no. 5, pp. 613-622, 1997.
- [10] R. Murray-Smith and K. J. Hunt, "Local model architecture for nonlinear modelling and control," in *'Advances in Neural Networks for Control Systems'* (K. J. Hunt, G. R. Irwin, and K. Warwick, eds.), Advances in Industrial Control, Springer-Verlag, 1995.
- [11] K. J. Hunt, R. Haas, and R. Murray-Smith, "Expanding the functional equivalence of radial basis function networks and fuzzy inference systems," *IEEE Transactions on Neural Networks*, vol. 7, no. 3, pp. 776-781, 1996.
- [12] S. Chen and S. A. Billings, "Neural networks for nonlinear dynamic system modelling and identification," *International Journal of Control*, vol. 56, no. 2, pp. 319-346, 1992.
- [13] J. Schurmann, *Pattern classification: A unified view of statistical and neural approaches*, John Wiley & Sons, Inc., 1996.
- [14] H.-J. Zimmerman and P. Zysno, "Latent connectives in human decision making," *Fuzzy Sets and Systems*, vol. 4, pp. 37-51, 1980.
- [15] H. Dyckhoff and W. Pedrycz, "Generalized means as model of compensative connectives," *Fuzzy Sets and Systems*, vol. 14, pp. 143-154, 1984.
- [16] R. Krishnapurum and J. Lee, "Fuzzy-connective-based hierarchical aggregation networks for decision making," *Fuzzy Sets and Systems*, vol. 46, pp. 11-27, 1992.
- [17] R. Langari, and L. Wang, "Fuzzy models, modular networks, and hybrid learning," *Fuzzy Sets and Systems*, vol. 79, pp. 141-150, 1996.
- [18] M. Delgado, A. F. Gomez-Skarmeta, and F. Matin, "A fuzzy clustering-based rapid prototyping for fuzzy rule-based modeling," *IEEE Tr. on Fuzzy Systems*, vol. 5, no. 2, pp. 223-233, 1997.
- [19] R. M. Tong, "The evaluation of fuzzy models derived from experimental data," *Fuzzy Sets and Systems*, vol. 4, pp. 1-12, 1980.
- [20] W. Pedrycz, "An identification algorithm in fuzzy relational systems," *Fuzzy Sets and Systems*, vol. 13, pp. 153-167, 1984.
- [21] M. Sugeno and K. Tanaka, "Successive identification of a fuzzy model and its application to prediction of a complex system," *Fuzzy Sets and Systems*, vol. 42,

pp. 315-334, 1991.

- [22] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modeling," *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 1, pp. 7-31, 1993.
- [23] L. Wang and R. Langari, "Complex systems modeling via fuzzy logic," *IEEE Tr. on Systems, Man and Cybernetics*, vol. 26, No. 1, pp. 100-106, 1996.
- [24] W.A. Farag, V. H. Quintana, and G. Lambert-Torres, "A genetic-based neuro-fuzzy approach for modeling and control of dynamic systems," *IEEE Tr. on Neural Networks*, Vol. 9, No. 5, pp. 756-767, 1998.

저 자 소 개



權五臣(正會員)

1982년 전북대학교 전기공학과 졸업(학사). 1984년 전북대학교 대학원 전기공학과 졸업(석사). 1988년 전북대학교 대학원 전기공학과(박사). 1978년~1988년 전주공업전문대 전자계산학과 조교수. 1993년~

1994년 일본 Osaka Electro-Communication University 객원 교수. 1997년~1999 군산대학교 공학연구소 소장. 1989년~현재 군산대학교 제어계측공학과 교수 1999년~현재 군산대학교 공과대학장. 주관심분야는 지능 제어, 퍼지 제어, 신경회로망 응용



崔鐘洙(正會員)

1969년 1월 18일생 1990년 전북대학교 전기공학과 졸업(학사). 1992년 전북대학교 대학원 전기공학과 졸업(석사). 1996년 전북대학교 대학원 전기공학과 졸업(박사). 1994년 3월~1996년 7월 (주)포스콘 기

술연구소 선임연구원. 1996년 7월~1999년 9월 포항산업과학연구원 선임연구원. 1999년 3월~1999년 9월 위덕대학교 정보통신공학과 겸임부교수. 주관심분야는 신경회로망, 퍼지시스템, 지능제어

金炯奭(正會員) 第35卷 C編 第11號 參照