

패턴 인식을 위한 Interval Type-2 퍼지 집합 기반의 최적 다중출력 퍼지 뉴럴 네트워크

Optimized Multi-Output Fuzzy Neural Networks Based on Interval Type-2 Fuzzy Set for Pattern Recognition

박 건 준* · 오 성 권†
(Keon-Jun Park · Sung-Kwun Oh)

Abstract - In this paper, we introduce an design of multi-output fuzzy neural networks based on Interval Type-2 fuzzy set. The proposed Interval Type-2 fuzzy set-based fuzzy neural networks with multi-output (IT2FS-based FNNm) comprise the network structure generated by dividing the input space individually. The premise part of the fuzzy rules of the network reflects the individuality of the division space for the entire input space and the consequent part of the fuzzy rules expresses three types of polynomial functions with interval sets such as constant, linear, and modified quadratic inference for pattern recognition. The learning of fuzzy neural networks is realized by adjusting connections of the neurons in the consequent part of the fuzzy rules, and it follows a back-propagation algorithm. In addition, in order to optimize the network, the parameters of the network such as apexes of membership functions, uncertainty factor, learning rate and momentum coefficient were automatically optimized by using real-coded genetic algorithm. The proposed model is evaluated with the use of numerical experimentation.

Key Words : Interval type-2 fuzzy set, Fuzzy neural networks, Multi-output, Genetic optimization, Pattern recognition

1. 서 론

퍼지뉴럴네트워크는 복잡성과 불확실성을 나타내는 비선형 공정 및 패턴분류에서 다른 지능형 모델에 비해 보다 나은 성능을 가지고 있다. 퍼지논리는 1965년 Zadeh[1]에 의해 창안된 “퍼지 집합” 이론을 이용하여 비선형적이고 복잡한 실 시스템의 특성을 해석하는데 적용함으로써 수학적 모델보다 좋은 결과를 가져왔다. 이것은 기존의 0이나 1중 하나의 값만을 선택해야 하는 이분법과는 달리 0과 1사이의 폐구간내 임의의 실수값을 가질 수 있어 비선형 시스템에 내재되어 있는 불확실성을 효과적으로 나타내었으며, 이들 시스템의 정적 혹은 동적 특성을 언어 변수를 사용하여 묘사하였다. 이 언어적인 변수를 표현하는 입력변수, 즉 멤버쉽 함수는 주어진 데이터에 대하여 퍼지 집합으로 표현하여 소속정도를 나타낸다. 하지만 멤버쉽 함수와 규칙을 결정하는데 많은 어려움을 가지고 있다. 이러한 문제를 해결하기 위해 학습능력을 가진 신경망과의 결합이 시도되었다. 하지만 여전히 멤버쉽 함수의 개수, 멤버쉽 함수 중심 등과 신경회로망에서는 층의 개수, 각 층의 노드 개수, 학습율, 모멘텀 계수 등과 같은 초기값 결정에 체계적인 방법이 제시되어 있지 않아 이를 조정하기 위해서 시행착오 및 전문가의 경

험으로 결정해야하는 단점이 있다. 또한 이들의 잘못된 선택은 모델자체의 수렴성이나 정확성에 악영향을 미칠 수 있기 때문에 신중한 선택이 요구된다. 따라서, 이들의 선택시 좀 더 객관적인 방법을 이용한다면 모델의 신뢰성을 높일 수 있다[2-5].

퍼지 모델은 시스템을 이해하기 쉽도록 정성적인 표현이 용이하고 불확실한 정보를 가지고 있는 시스템에 강인성을 발휘할 수 있다. 그러나 학습능력을 가지고 있지 않고 있으며 퍼지규칙을 얻기 위해 전문가나 경험에 의존해야 하는 어려움이 있다. 반면에 신경회로망은 시스템의 입출력 관계를 학습을 통해 추출할 수 있으며 병렬처리 기능을 가지고 있어 빠른 처리 능력을 가진다. 그러나 신경회로망은 주어진 시스템에 대해 정성적 표현이 어려워 시스템을 이해하기에는 어려움을 가진다. 이러한 두 모델의 장점을 결합하고 단점을 해결하기 위해 퍼지 모델과 신경회로망의 유사성과 상호 보완관계를 이용하여 다양한 결합을 시도하고 있으며 이러한 구조가 퍼지뉴럴네트워크이다[6-11].

Type-1 퍼지 집합의 확장인 Type-2 퍼지 집합은 Zadeh [12]에 의해 소개되었으며, Mizumoto와 Tanaka [13]는 type-2 퍼지 집합의 이론적인 연산 및 멤버쉽함수의 특성에 대해 연구하였다. Mendel과 Karnik [14]은 퍼지 “if-then” 규칙으로 표현되는 type-2 퍼지 논리 시스템에 대해 연구하였다. type-2 퍼지 논리 시스템은 전반부 및 후반부에 type-2 퍼지 집합을 사용하였다. Interval type-2 퍼지 뉴럴 네트워크(IT2FNN)는 Interval type-2 퍼지 집합과 신경회로망의 상호 결합에 의한 네트워크이며, 효율적으로 대처하기 위해 연구되었다.

† Corresponding Author : Dept. of Electrical Engineering,
The University of Suwon, Korea
E-mail : ohsk@suwon.ac.kr

* Dept. of Electrical Engineering, The University of Suwon,
Korea

Received : October 17, 2012; Accepted : February 26, 2013

본 논문에서는 전체 입력 공간을 서로 분리된 개별적인 공간을 이용하여 입력 공간을 개별적으로 퍼지 규칙의 전반부를 형성하고 각 공간은 Interval Type-2 퍼지 집합으로 구성된다. 후반부 구조는 패턴 인식을 위한 다중 출력을 가지는 구조로서 Interval 집합을 이용하여 다항식으로 표현되고 오류역전과 알고리즘을 이용하여 다항식 계수를 학습한다. 퍼지 규칙 기반 퍼지뉴럴네트워크는 각 입력 변수에 대한 전반부 멤버쉽함수의 정점과 불확실성 계수 그리고, 학습률 및 모멘텀 계수와 같은 파라미터들을 최적으로 동조하기 위하여 실수코딩 유전자 알고리즘[16]을 이용한다. 또한, 제안된 퍼지뉴럴네트워크는 패턴 인식을 위해 표준 모델로서 널리 사용되는 수치적인 예를 통하여 평가한다.

2. Interval Type-2 퍼지 집합 기반의 다중 출력 퍼지 뉴럴 네트워크

Interval Type-2 퍼지 뉴럴 네트워크는 Interval Type-2 퍼지 이론과 신경회로망의 두 구조가 유사성과 상호 보완적인 특성으로부터 결합이 시도된다. 이 결합에 의한 퍼지 뉴럴 네트워크 구조는 퍼지이론의 언어적 논리를 규칙의 형태로 표현할 수 있는 능력과 신경회로망의 학습기능과의 상호간에 높은 융통성을 가지게 된다.

2.1 Interval Type-2 퍼지 집합

Type-2 퍼지 집합은 다음과 같은 형태의 Type-2 멤버쉽함수로 특징지어진다

$$\tilde{A} = \int_{x \in X} \mu_{\tilde{A}}(x)/x = \int_{x \in X} \left[\int_{u \in J_x} f_x(u)/u \right] / x \quad (1)$$

$J_x \subseteq [0,1]$

식 (1)에서, secondary 멤버쉽함수의 영역은 x에 대한 primary 멤버쉽함수이며, J_x 는 x에 대한 primary 멤버쉽함수이다. $f_x(u)$ 는 secondary 멤버쉽함수의 크기이다.

$f_x(u)/u = 1, \forall u \in J_x \subseteq [0,1]$ 인 경우에는 secondary 멤버쉽함수는 Interval이며 $\mu_{\tilde{A}}(x)$ 는 Interval Type-2 멤버쉽함수이다. 따라서, \tilde{A} 는 다음과 같이 표현된다.

$$\tilde{A} = \int_{x \in X} \mu_{\tilde{A}}(x)/x = \int_{x \in X} \left[\int_{u \in J_x} 1/u \right] / x, J_x \subseteq [0,1] \quad (2)$$

primary 멤버쉽 값에 대한 불확실성은 footprint of uncertainty (FOU)라 불리는 경계구간을 갖는 Type-2 퍼지 집합으로 표현된다.

$$FOU(\tilde{A}) = \cup_{x \in X} J_x \quad (3)$$

FOU에 대한 예는 그림 1에서 보여준다. FOU는 Type-2 퍼지 집합의 Interval 특성을 강조하기 위하여 균일하게 표현된다. 불확실성은 임의의 범위 내에서 정점을 기준으로 멤버쉽함수의 값을 표현한다. 여기서, a, b, c는 멤버쉽 파라미터이며 σ_a 와 σ_c 는 a와 c의 불확실성을 의미한다.

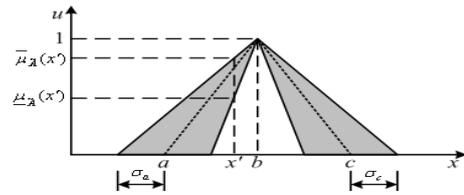


그림 1 Interval Type-2 퍼지 집합
Fig. 1 Interval Type-2 Fuzzy Set

Interval Type-2 퍼지 집합의 FOU에 대한 경계는 상위와 하위의 두 개의 Type-1 멤버쉽함수를 가진다. 상위 멤버쉽함수의 상위경계를 나타내며 하위 멤버쉽함수의 하위 경계를 의미한다.

2.2 퍼지 뉴럴 네트워크의 구조

입력 공간의 퍼지 분할은 개별적인 분할 공간에 의해 퍼지 규칙을 형성하며 후반부 구조는 간략 추론(Case 1), 선형 추론(Case 2), 변형된 2차식 추론(Case 3)에 의해 표현된다. 각 퍼지 규칙의 추론된 값은 다중 출력을 위해 연결가중치로 연결된 구조이다. 그림 2는 제안된 Interval Type-2 퍼지 집합 기반의 다중 출력 퍼지 뉴럴 네트워크의 구조를 보여준다.

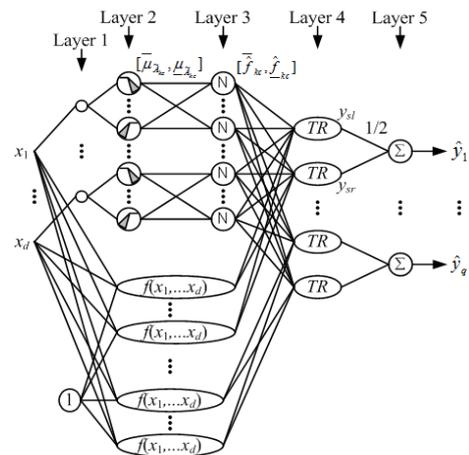


그림 2 Interval Type-2 퍼지 집합 기반의 다중 출력 퍼지 뉴럴 네트워크 구조
Fig. 2 Structure of Interval Type-2 Fuzzy Set-based multi-output Fuzzy Neural Networks

개별 입력 공간 분할에 의한 제안된 네트워크의 퍼지 규칙은 입력 공간별 퍼지 규칙을 형성하며 다음과 같이 표현된다.

$$R^{kc}: \text{If } x_k \text{ is } \tilde{A}_{kc} \text{ Then } y_{kc}^s = f(x_1, \dots, x_d) \quad (4)$$

간략 추론(Case 1) : $f = W_{kc}^{s0}$

선형 추론(Case 2) : $f = W_{kc}^{s0} + \sum_{i=1}^d W_{kc}^{si} x_i$

변형된 2차식 추론(Case 3) :

$$f = W_{kc}^{s0} + \sum_{i=1}^d W_{kc}^{si} x_i + \sum_{i=1}^d \sum_{j=i+1}^d W_{kc}^{sj} x_i x_j$$

여기서, $W_{kc}^{si} = [w_{kc}^{si} - s_{kc}^{si}, w_{kc}^{si} + s_{kc}^{si}]$, $i=0, \dots, d(d+1)/2$.

여기서, R^{kc} 는 k, c 번째 퍼지 규칙을 나타내며, A_{kc} 는 전 반부 퍼지변수로 멤버쉽함수로 표현된다. W_{kc}^{si} 는 구간 집합으로 후반부 계수를 의미하며, w_{kc}^{si} 는 W_{kc}^{si} 의 중심을, s_{kc}^{si} 는 W_{kc}^{si} 의 폭을 각각 의미한다.

전반부 멤버쉽함수는 이웃하는 멤버쉽함수와 보수적인 관계를 가진 삼각형 멤버쉽함수를 사용한다. 각 멤버쉽함수의 σ_a 와 σ_b 는 각 퍼지 부분 공간에 속한 데이터의 표준편차에 의해 구해지며, 불확실성 계수 ρ 를 이용하여 $(1+\rho)\sigma$ 에 의해 불확실성을 조절한다.

제안된 퍼지 뉴럴 네트워크 구조에서 각 층의 연산 과정은 다음과 같다.

[1층] 입력층

각 입력은 퍼지 입력 공간으로부터 분할된 멤버쉽함수로 출력된다.

[2층] 입력에 대한 적합도 계산

멤버쉽함수는 그림 3과 같이 삼각형 형태의 Interval Type-2 멤버쉽함수로, 각 입력에 의해 멤버쉽함수로부터 멤버쉽 값을 구한다.

$$\mu_{\tilde{A}_k}(x) = [\overline{\mu_{\tilde{A}_k}}(x), \underline{\mu_{\tilde{A}_k}}(x)] \quad (5)$$

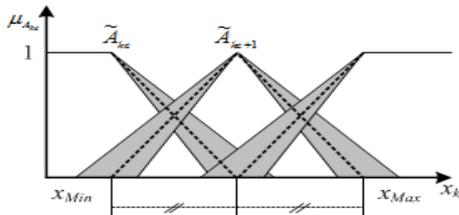


그림 3 Interval Type-2 멤버쉽함수
Fig. 3 Interval Type-2 Membership function

[3층] 정규화된 적합도 계산

각 규칙의 적합도가 구해지면 정규화된 적합도 \hat{f}_{kc} 를 계산한다.

$$\hat{f}_{kc} = [\overline{\hat{f}_{kc}}, \underline{\hat{f}_{kc}}] \quad (6)$$

$$\overline{\hat{f}_{kc}} = \frac{\overline{\mu_{kc}}}{\sum_{c=1}^q \overline{\mu_{kc}}}, \quad \underline{\hat{f}_{kc}} = \frac{\underline{\mu_{kc}}}{\sum_{c=1}^q \underline{\mu_{kc}}} \quad (7)$$

[4층] Type Reduction

가장 왼쪽 지점 y_l 과 가장 오른쪽 지점 y_r 은 \hat{f}_{kc} 에 의존하

며, 반복법에 따른 Karnik-Mendel (KM) 알고리즘[14]을 이용하여 구해진다.

$$y_{sl} = \frac{\sum_{k=1}^d \sum_{c=1}^q \hat{f}_{kc}^s y_{kc}^s}{\sum_{k=1}^d \sum_{c=1}^q \hat{f}_{kc}^s} \quad (8)$$

$$y_{sr} = \frac{\sum_{k=1}^d \sum_{c=1}^q \hat{f}_{kc}^r y_{kc}^r}{\sum_{k=1}^d \sum_{c=1}^q \hat{f}_{kc}^r} \quad (9)$$

여기서, \hat{f}_{kc}^s 와 \hat{f}_{kc}^r 는 s 번째 출력에 대한 각 규칙에서 가장 왼쪽 지점 y_l 과 가장 오른쪽 지점 y_r 에 영향을 준 상위 멤버쉽함수와 하위 멤버쉽함수이다.

[5층] 최종출력

IT2FSFNNm에서 최종 출력은 구간 집합이다. 따라서, 최종 출력은 가장 왼쪽 지점 y_l 과 가장 오른쪽 지점 y_r 의 평균에 의해 구해진다.

$$\hat{y}_s = \frac{y_{sl} + y_{sr}}{2} \quad (10)$$

2.3 퍼지 뉴럴 네트워크의 학습

네트워크의 학습은 후반부 다항식 계수인 연결가중치를 가장 적절한 값으로 조정해 나가면서 이루어진다. 학습방법은 오류역전파 알고리즘을 이용하며, 이는 원 출력과 모델 출력간의 차, 즉 오차를 바탕으로 이루어진다. 다중 출력에 의한 오차는 다음과 같이 정의한다.

$$E_p = \frac{1}{2} \sum_{s=1}^q (y_{ps} - \hat{y}_{ps})^2 \quad (11)$$

여기서, E_p 는 p 번째 데이터에 대한 오차이고, y_{ps} 와 \hat{y}_{ps} 는 p, s 번째 원 출력과 모델 출력을 나타낸다.

간략 추론의 경우 연결가중치 w_{kc}^{s0} 의 조절은 다음과 같다.

$$w_{kc}^{s0}(p+1) = w_{kc}^{s0}(p) + \Delta w_{kc}^{s0} \quad (12)$$

오차 E_p 를 최소화하기 위한 연결가중치의 변화량은 경사하강법을 이용하여 구할 수 있다.

$$\Delta w_{kc}^{s0} = \eta \left(- \frac{\partial E_p}{\partial w_{kc}^{s0}} \right) \quad (13)$$

여기서, η 는 학습률이고, 사슬법칙에 의해 다음과 같이 전개된다.

$$- \frac{\partial E_p}{\partial w_{kc}^{s0}} = - \frac{1}{2} \frac{\partial E_p}{\partial \hat{y}_{ps}} \left(\frac{\partial \hat{y}_{ps}}{\partial y_l} \frac{\partial y_l}{\partial w_{kc}^{s0}} + \frac{\partial \hat{y}_{ps}}{\partial y_r} \frac{\partial y_r}{\partial w_{kc}^{s0}} \right) \quad (14)$$

또한, 모멘텀을 추가해 학습시간을 단축시키고, 학습성능을 향상시킨다. 연결가중치의 변화량이 클 경우에는 변화량을 키워주고, 변화량이 작아지면 변화량을 줄여 주는 역할을 한다. 모멘텀은 현재의 연결가중치와 그 이전 연결가중치의 차이에 의해 계산된다. 이렇게 계산된 모멘텀은 모멘텀 계수 α 와 곱해져 연결 가중치 조절식에 추가된다. 모멘텀까지 포함하는 연결가중치 w_{kc}^{s0} 의 최종 변화량은 다음과 같다.

$$\Delta u_{kc}^{s0} = 0.25\eta(y_{ps} - \widehat{y}_{ps})(\widehat{f}_{kc}^f + \widehat{f}_{kc}^r)z + \alpha(w_{kc}^{s0}(p) - u_{kc}^{s0}(p-1)) \quad (15)$$

한편, 연결가중치 s_{kc}^{s0} 의 조절은 다음과 같다.

$$s_{kc}^{s0}(p+1) = s_{kc}^{s0}(p) + \Delta s_{kc}^{s0} \quad (16)$$

마찬가지로, 오차를 최소화하기 위한 연결가중치의 변화량은 경사하강법을 이용하여 구할 수 있다.

$$\Delta s_{kc}^{s0} = \eta \left(- \frac{\partial E_p}{\partial s_{kc}^{s0}} \right) \quad (17)$$

또한, 사슬법칙에 의해 다음과 같이 전개된다.

$$- \frac{\partial E_p}{\partial s_{kc}^{s0}} = - \frac{1}{2} \frac{\partial E_p}{\partial y_{ps}} \left(\frac{\partial y_{ps}}{\partial y_l} \frac{\partial y_l}{\partial s_{kc}^{s0}} + \frac{\partial y_{ps}}{\partial y_r} \frac{\partial y_r}{\partial s_{kc}^{s0}} \right) \quad (18)$$

마지막으로, 모멘텀까지 포함하는 연결 가중치 s_{kc}^{s0} 의 최종 변화량은 다음과 같다.

$$\Delta s_{kc}^{s0} = 0.25\eta(y_{ps} - \widehat{y}_{ps})(-\widehat{f}_{kc}^f + \widehat{f}_{kc}^r) + \alpha(s_{kc}^{s0}(p) - s_{kc}^{s0}(p-1)) \quad (19)$$

선형 추론인 경우와 변형된 2차식 추론인 경우에도 같은 방법으로 유도된다.

$$\Delta u_{kc}^{si} = 0.25\eta(y_{ps} - \widehat{y}_{ps})(\widehat{f}_{kc}^f + \widehat{f}_{kc}^r)x_i + \alpha(u_{kc}^{si}(p) - u_{kc}^{si}(p-1)) \quad (20)$$

$$\Delta s_{kc}^{si} = 0.25\eta(y_{ps} - \widehat{y}_{ps})(-\widehat{f}_{kc}^f + \widehat{f}_{kc}^r)|x_i| + \alpha(s_{kc}^{si}(p) - s_{kc}^{si}(p-1)) \quad (21)$$

$$\Delta u_{kc}^{sz} = 0.25\eta(y_{ps} - \widehat{y}_{ps})(\widehat{f}_{kc}^f + \widehat{f}_{kc}^r)x_i x_j + \alpha(u_{kc}^{sz}(p) - u_{kc}^{sz}(p-1)) \quad (22)$$

$$\Delta s_{kc}^{sz} = 0.25\eta(y_{ps} - \widehat{y}_{ps})(-\widehat{f}_{kc}^f + \widehat{f}_{kc}^r)|x_i x_j| + \alpha(s_{kc}^{sz}(p) - s_{kc}^{sz}(p-1)) \quad (23)$$

3. 실수 코딩 유전자 알고리즘

유전자 알고리즘[5,6]이란 적자생존의 생물학 원리에 바탕을 둔 최적화 기법중의 하나로 자연계의 생명체 중 환경에 잘 적응한 개체가 좀더 많은 자손을 남길 수 있다는 자연선택 과정과 자연계 생명체의 설계도와 같은 유전자의 변화를 통해서 좋은 방향으로 발전해 나간다는 자연진화 과정의 유전자 메카니즘에 바탕을 둔 탐색 알고리즘이다. 즉, 자연계에 있어서 생물학의 유전과 진화의 메카니즘을 공학적으로 모델화하는 것에 의해 생물의 환경 적응능력과 자연도태의

원리를 기반으로 만들어진 계산 모델로, 실세계의 문제를 풀기 위해 잠재적인 해들을 컴퓨터 상에서 코딩된 개체로 나타내고, 여러 개의 개체들은 모아 군집을 형성한 뒤, 세대를 거듭하면서 이들의 유전 정보를 서로 교환하거나 새로운 유전 정보를 부여하면서 적자 생존의 법칙에 따라 모의 진화를 시킴으로써, 주어진 문제에 대한 최적의 해를 찾는 계산 모델이다.

유전자 알고리즘은 기존의 다른 탐색방법들과 달리 유전자 알고리즘은 해가 될 가능성이 있는 개체집단을 유지하면서 그들 모두가 동시에 최적값을 찾아나가기 때문에 지역 극소에 빠질 위험을 해결할 수 있다는 점과 더불어 모델의 성능지수가 최소가 되는 전역 극소 영역을 찾는 능력을 갖고 있으며, 또한 기존의 방법들과는 달리 선형, 연속, 미분가능 등의 제한이 없기 때문에 다양한 분야에 적용할 수 있다는 장점을 가진다는 것이 중요한 특징이다.

유전자 알고리즘은 3개의 기본 연산자, 즉 재생산, 교배, 돌연변이를 가지고 있다. 생산(reproduction)은 다음 세대를 위해 교배 유전자가 선택되어지는 과정으로, 적합도에 따라 다음세대를 생산할 수 있는 기회를 부여한다. 각각의 문자열들은 그들의 적합도를 가지고 다음세대에 전달된다. 교배(crossover)는 교배 유전자로부터 염색체들은 짝짓고 짝지어진 문자열 쌍은 확률적으로 선택되어진다. 돌연변이(mutation)는 문자열 위치의 값을 임의로 변경한다. 2진 코딩에서 돌연변이는 0을 1로 또는 1을 0으로 바꾸는 것을 의미한다. 적절하게 정의된 적합도 함수와 함께 조합된 연산자는 유전자 알고리즘의 주요 부분이 된다. 유전자 알고리즘은 초기 집단 생성후 적합도를 평가하고, 재생산, 교배, 돌연변이의 기본 연산자를 통해 최적해를 탐색하며, 일반적인 수행과정은 그림 4와 같다.

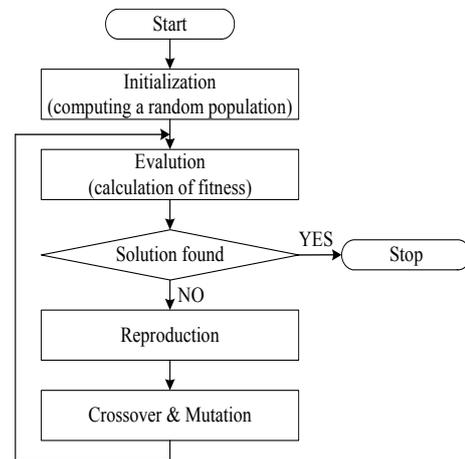


그림 4 유전자 알고리즘의 수행과정
Fig. 4 Flowchart of genetic algorithms

실세계의 문제들은 고정밀도의 해를 요구하고, 해에 대한 사전지식이 결여되거나, 복잡한 제약조건이 존재하는 경우가 종종 있다. 이진코딩을 채용할 때 정밀도를 높이거나 해에 대한 사전지식의 결여로 탐색구간을 확대하게 되면 탐색체 길이가 길어진다. 긴 탐색체는 매우 큰 탐색공간을 만들어 해를 찾는데 많은 시간을 요구하거나, 경우에 따라서는 실패

하는 원인을 제공한다. 이러한 문제점들은 근본적으로 염색체 표현을 좀더 해공간에 접근시켜 문제와 관련된 지식을 포함시키는 연산자 설계를 용이하게 하고, 제약조건을 다루기 위한 도구를 설계하는 것이 쉽도록 함으로써 해결할 수 있으며, 대표적인 표현법이 실수코딩이다. 실수코딩을 채용함으로써 염색체의 유전자와 벡터 요소는 정확히 일대일 일치하는 관계를 가지며 염색체의 길이는 벡터의 차원과 같게 된다.

본 논문에서는 멤버쉽함수의 정점과 불확실성 계수 및 학습률, 모멘텀 계수를 최적으로 설정해 주기 위하여 실수 코딩 유전자 알고리즘을 이용하여 네트워크의 최적화 과정을 수행하게 된다. 그림 5는 최적 동조를 위한 일련의 염색체 구조를 보여준다. 여기서, 괄호안의 수는 염색체 수를 나타낸다.

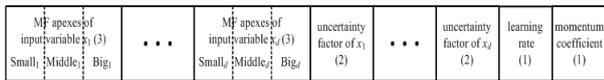


그림 5 염색체 구조
Fig. 5 Chromosomes structure

4. 시뮬레이션 및 결과 고찰

본 논문에서 제안된 네트워크를 적용하기 위하여 패턴 실험 데이터로 Fisher의 Iris 데이터를 이용하였다. 실험에 사용된 Fisher의 Iris 데이터는 3가지의 클래스(Iris Setosa, Iris Versicolor, Iris Virginica)로 구성되며, 각 패턴은 4가지의 속성(꽃받침 길이, 꽃받침 너비, 꽃잎 길이, 꽃잎 너비)으로 정의된다. 각 클래스마다 50개의 패턴으로 구성되며 총 150개의 패턴을 가진다.

실험은 random sub-sampling을 이용하여 각 데이터 집합을 7:3의 비율로 랜덤하게 학습 데이터와 테스트 데이터로 나누어 5번 반복 실험하였으며 제안된 네트워크를 이용하여 패턴 분류를 수행하였다. 성능평가는 인식률(Classification Ratio; CR) 및 MSE(Mean Square Error)에 의한 정확도(Performance Index; PI)의 평균과 표준편차에 의해 평가하였다. 여기서, K는 random sub-sampling에 의한 반복 횟수를 의미한다.

$$E_p = \frac{\text{No. of classification}}{\text{No. of examples}} \times 100 \quad (24)$$

$$CR = \frac{1}{K} \sum_{i=1}^K E_p \quad (25)$$

$$\Pi = \frac{1}{K} \sum_{i=1}^K MSE \quad (26)$$

실험에 앞서 표 1은 최적의 네트워크를 설계하기 위한 유전자 알고리즘 및 제안된 네트워크의 초기 파라미터를 보여준다.

실험은 표 1에서 상술된 파라미터를 가지고 파라미터 최적화를 수행하였다. 멤버쉽함수의 수는 증가시키면서 시행하였으며 간략 추론, 선형 추론, 변형된 2차식 추론에 대해 각각 시행하였다. 표 2는 각 입력 변수에 대해 멤버쉽함수의 수가 2, 3, 5개인 경우에 있어서 후반부 구조가 간략 추

표 1 제안된 네트워크의 유전자적 최적화를 위한 초기 파라미터

Table 1 Initial parameters of genetic optimization of the proposed networks

	파라미터	값
GAs	세대 수	100
	군집 크기	50
	교배율	0.65
	돌연변이율	0.1
IT2FSFNNm	불확실성 계수	$-1.0 \leq \rho \leq 1.0$
	학습률	$0.0 \leq \eta \leq 0.01$
	모멘텀 계수	$0.0 \leq \alpha \leq 0.001$

론, 선형 추론, 변형된 2차식 추론방법에 의한 최적화 전의 퍼지 뉴럴 네트워크에 대한 성능을 보여준다. 마찬가지로, 표 3은 최적화 후의 퍼지 뉴럴 네트워크에 대한 성능을 보여준다.

표 2 최적화 전의 네트워크 성능

Table 2 Performance of networks before optimization

MFs수	추론	인식률		정확도	
		학습	테스트	학습	테스트
2	1	85.14±2.98	83.56±3.72	0.092±0.01	0.089±0.01
	2	97.14±0.95	96.00±2.43	0.038±0.00	0.037±0.01
	3	96.00±0.43	97.78±1.57	0.036±0.00	0.038±0.01
3	1	97.90±0.80	96.00±2.43	0.026±0.00	0.027±0.00
	2	97.71±0.52	96.44±1.99	0.024±0.00	0.027±0.01
	3	98.29±1.04	97.78±1.57	0.024±0.00	0.029±0.00
5	1	96.38±1.24	96.89±1.99	0.020±0.00	0.019±0.01
	2	98.48±0.52	95.56±2.22	0.016±0.00	0.023±0.01
	3	97.71±0.52	95.11±2.90	0.015±0.00	0.025±0.01

표 3 최적화 후의 네트워크 성능

Table 3 Performance of networks after optimization

MFs수	추론	인식률		정확도	
		학습	테스트	학습	테스트
2	1	96.38±0.80	97.33±1.86	0.032±0.01	0.030±0.01
	2	99.24±0.80	96.89±1.22	0.025±0.00	0.033±0.01
	3	99.24±0.43	99.11±1.22	0.028±0.00	0.028±0.00
3	1	98.10±0.67	99.11±1.22	0.023±0.00	0.017±0.01
	2	99.81±0.43	98.22±0.99	0.021±0.00	0.019±0.00
	3	99.62±0.52	98.67±1.22	0.021±0.00	0.019±0.00
5	1	98.29±0.43	99.56±0.99	0.022±0.00	0.020±0.00
	2	99.05±0.67	99.56±0.99	0.020±0.00	0.014±0.00
	3	99.62±0.52	98.67±1.22	0.018±0.00	0.019±0.01

표 2와 표 3에서 살펴보면 유전자 알고리즘을 적용하여 전반부 멤버쉽함수의 정점과 불확실성 계수, 학습률, 모멘텀 계수를 동조한 경우 인식률과 정확도에 있어서 근사화 및 일반화 모두 성능이 향상된 것을 알 수 있다. 특히, 각 입력에 대해 멤버쉽함수의 수가 5개 이고, 규칙수가 20개 이며, 선형 추론인 경우 가장 좋은 성능을 보여준다. 이때의 인식

률은 학습 데이터에 대해 $99.05 \pm 0.67\%$ 이고 테스트 데이터에 대해 $99.56 \pm 0.99\%$ 이며, 정확도는 학습 데이터에 대해 0.020 ± 0.00 이고 테스트 데이터에 대해 0.014 ± 0.00 이다.

그림 6은 선각 입력에 대해 멤버쉽함수의 수가 5개 이고, 규칙수가 20개 이며, 선형 추론인 경우에 있어서 5번 반복에 따른 인식률과 정확도의 평균값에 대한 최적화 탐색과정을 보여준다.

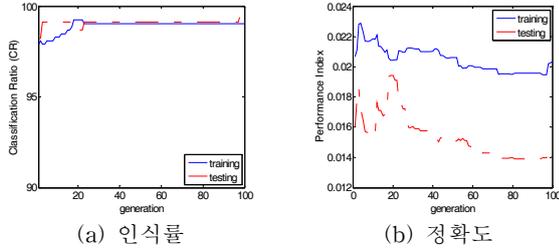


그림 6 최적화 과정
Fig. 6 Optimization process

표 4는 기존 모델과의 비교를 보여준다. 제안된 네트워크는 멤버쉽함수의 수가 5개이고 후반부 구조가 선형 추론인 경우에 있어서 학습 데이터와 테스트 데이터에 대한 평균 분류율을 나타내었다. 기존 모델과 비교해 보면 제안된 네트워크는 학습에 의한 근사화 및 테스트에 의한 일반화 성능에 있어서 모두 좋은 성능을 보여주는 것을 알 수 있다.

표 4 기존 모델과의 비교
Table 4 Comparison of previous models

모델	평균 분류율(%)		
	전체	학습	테스트
Simpson[17]	97.3		
Nauck[18]		96	97.3
Meesad[19]	88~97.3		
Yen[20]		96.7	96.2
Li[21]	96.67		
Proposed model		99.05	99.56

5. 결 론

본 논문에서는 Interval Type-2 퍼지 집합을 이용한 다중 출력을 가지는 다중 규칙 기반의 퍼지 뉴럴 네트워크를 설계하였다. 퍼지 규칙의 전반부는 전체 입력공간을 독립적으로 분할하여 형성하였으며, 각 퍼지 부분 공간은 Interval Type-2 퍼지 집합을 이용하여 구성하였다. 규칙의 후반부는 다중 출력을 위한 다중 규칙의 구조를 가지며 Interval 집합을 이용하여 다항식으로 표현하였다. 오류역전과 알고리즘을 이용하여 각 다항식의 계수를 학습하였으며, 실수 코딩 유전자 알고리즘을 이용하여 파라미터들을 최적화하였다. 실험에서 알 수 있듯이 제안된 네트워크는 상당히 좋은 성능을 보여주며, 인식률 정확도 측면에서도 상당히 개선되는 것을 알 수 있었다. 아울러, 제안된 네트워크는 퍼지 규칙의

특성상 입력 공간이 고차원으로 갈수록 차원의 문제로부터 다소 자유스러운 것을 알 수 있었다.

감사의 글

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2012-003568) 및 경기도의 경기도지역협력연구센터사업(GRRC 수원2012-B2, U-city 보안감시 기술협력센터)에서 지원하여 연구하였음.

References

- [1] L. A Zadeh, "Fuzzy sets", Inf. Control 8, 338-353, 1965.
- [2] H. Takagi and I. Hayashi, "NN-driven Fuzzy Reasoning", Int. J. of Approximate Reasoning, Vol. 5, No. 3, pp. 191-212, 1991.
- [3] N. Imasaki, J. Kiji, and T. Endo, "A Fuzzy Rule Structured Neural Networks", Journal of Japan Society for Fuzzy Theory and Systems, Vol. 4, No. 5, pp. 987-995, 1992(in Japanese).
- [4] H. Nomura and Wakami, "A Self-Tuning Method of Fuzzy Control by Descent Methods", 4th IFSA'91, pp. 155-159, 1991.
- [5] Shin-ichi Horikawa, Takeshi Furuhashi, and Yoshiaki Uchigawa, "On Fuzzy Modeling Using Fuzzy Neural Networks with the Back Propagation Algorithm", IEEE trans. Neural Networks, Vol. 3, No. 5, pp. 801-806, 1992.
- [6] Takeshi Yamakawa, "A Neo Fuzzy Neuron and Its Applications to System Identification and Prediction of the System Behavior", Proceedings of the 2nd International Conference on Fuzzy logic & Neural Networks, pp. 477-483, 1992.
- [7] Takeshi Yamakawa, "A New Effective Learning Algorithm for a Neo Fuzzy Neuron Model", 5th IFSA World Congress, pp. 1017-1020, 1993.
- [8] J. J. Buckley and Y. Hayashi, "Fuzzy neural networks: A survey", Fuzzy Sets and Systems, Vol. 66, pp. 1-13, 1994.
- [9] M. M. Gupta and D. H. Rao, "On the principles of fuzzy neural networks", Fuzzy Sets and Systems, Vol. 61, pp. 1-18, 1994.
- [10] J.-S. R. Jang, C. -T. Sun, E. Mizutani, Neuro-Fuzzy And Soft Computing, Prentice-Hall, 1997.
- [11] S.-K. Oh, "Computational Intelligence by Programming", Naeha Publishing Co., 2002
- [12] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-I," Information Science, Vol. 8, pp. 199-249, 1975.
- [13] Mizumoto, M. and K. Tanaka, "Some Properties of

Fuzzy Sets of Type-2," Information and Control, Vol. 31, pp. 312-340, 1976

[14] J. M. Mendel, Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions, Prentice-Hall: NJ, 2001.

[16] Gang-Gyoo Jin, "Genetic Algorithms and Their Applications", Kyowoo publishing Co., 2004

[17] P. K. Simpson, "Fuzzy Min-Max Neural Networks -Part 1 : Classification", IEEE Trans on. Neural Networks, Vol. 3, No. 5, 1992.

[18] D. Nauck, R. Kruse, "A neuro-fuzzy method to learn fuzzy classification rules from data", Fuzzy Sets and Systems, Vol. 89, pp. 277-288, 1997.

[19] P. Meesad, G. G. Yen, "Pattern classification by a neurofuzzy network : application to vibration monitoring", ISA Transactions, Vol. 39, pp. 293-308, 2000.

[20] G. G. Yen, P. Meesad, "An effective neuro-fuzzy paradigm for machinery condition health monitoring", IEEE Trans. on Systems, Man, and Cybern., : Part B, Vol. 31, No. 4, 2001.

[21] R. Li, M. Mukaidono and I.B. Turksen, "A fuzzy neural network for pattern classification and feature selection", Fuzzy Sets and Systems, Vol. 130, pp. 101-140, 2002.

저 자 소 개



박 건 준 (朴 建 俊)

2003년 원광대학교 전기전자공학부 졸업. 2005년 동 대학원 제어계측공학과 졸업 (공학석사). 2010년 수원대학교 전기공학과 박사졸업(공학박사). 관심분야는 컴퓨터 및 인공지능, 퍼지추론시스템, 신경망, 유전자 알고리즘 및 최적화이론, 자동화

시스템 및 제어 등.

Tel : 031-222-6544

Fax : 031-220-2667

E-mail : birds@suwon.ac.kr



오 성 권 (吳 聖 權)

1981년 연세대학교 전기공학과 졸업, 동 대학원 석사(1983), 박사(1993). 1983-1989년 금성산전연구소(선임연구원). 1996-1997년 캐나다 Manitoba 대학 전기 및 컴퓨터공학과 Post-Doc. 1993-2005년 원광대학교 전기전자 및 정보공학부 교수. 2005~현재 수원대학교 전기공학과 교수, 2002~현재 대한전기학회, 제어로봇시스템학회, 퍼지 및 지능시스템 학회 편집위원. 관심분야는 퍼지 시스템, 퍼지-뉴럴 네트워크, 자동화 시스템, 고급 computational intelligence, 지능 제어 등.

Tel : 031-229-8162

Fax : 031-220-2667

E-mail : ohsk@suwon.ac.kr