

퍼지 추론 메커니즘에 기반 한 다항식 네트워크 패턴 분류기의 설계와 이의 최적화

The Design of Polynomial Network Pattern Classifier based on Fuzzy Inference Mechanism and Its Optimization

김길성* · 박병준** · 오성권*

Gil-Sung Kim · Byoung-Jun Park · Sung-Kwun Oh

* 수원대학교 전기공학과

** (주)지엠테크 기술연구소

요 약

본 연구에서는 퍼지 추론 메커니즘에 기반 한 다항식 네트워크 패턴 분류기(Polynomial Network Pattern Classifier; PNC)를 설계하고 Particle Swarm Optimization 알고리즘을 이용하여 PNC 파라미터, 즉, 학습률, 모멘텀 계수, FCM 클러스터링의 퍼지화 계수(fuzzification Coefficient)를 최적화한다. 제안된 PNC 구조는 FCM 클러스터링에 기반 한 분할 함수를 활성화 함수로 사용하며, 다항식 함수로 구성된 연결가중치를 사용함으로써 기존 신경회로망 분류기의 선형적인 특성을 개선한다. PNC 구조는 언어적 해석관점에서 “If-then”의 퍼지 규칙으로 표현되며 퍼지 추론 메커니즘에 의해 구동된다. 즉 조건부, 결론부, 추론부 세 가지의 기능적 모듈로 나뉘어 네트워크 구조가 형성된다. 조건부는 FCM 클러스터링을 사용하여 입력 공간을 분할하고, 결론부는 분할된 로컬 영역을 다항식 함수로 표현한다. 마지막으로, 네트워크의 최종출력은 추론부의 퍼지추론에 의한다. 제안된 PNC는 다항식 기반 구조의 퍼지 추론 특성으로 인해 출력 공간상에 비선형 판별 함수(nonlinear discernment function)가 생성되어 분류기로서의 성능을 높인다.

Abstract

In this study, Polynomial Network Pattern Classifier(PNC) based on Fuzzy Inference Mechanism is designed and its parameters such as learning rate, momentum coefficient and fuzzification coefficient are optimized by means of Particle Swarm Optimization. The proposed PNC employs a partition function created by Fuzzy C-means(FCM) clustering as an activation function in hidden layer and polynomials weights between hidden layer and output layer. Using polynomials weights can help to improve the characteristic of the linear classification of basic neural networks classifier. In the viewpoint of linguistic analysis, the proposed classifier is expressed as a collection of “If-then” fuzzy rules. Namely, architecture of networks is constructed by three functional modules that are condition part, conclusion part and inference part. The condition part relates to the partition function of input space using FCM clustering. In the conclusion part, a polynomial function carries out the presentation of a partitioned local space. Lastly, the output of networks is gotten by fuzzy inference in the inference part. The proposed PNC generates a nonlinear discernment function in the output space and has the better performance of pattern classification as a classifier, because of the characteristic of polynomial based fuzzy inference of PNC.

Key Words : Pattern Recognition, Pattern Classifier, Neural Networks, Polynomial Neural Networks, Radial Basis Function Neural Networks, Polynomial Network Pattern Classifier.

1. 서 론

신경회로망을 기반으로 설계된 분류기는 학습능력과 일반화(generalization) 능력의 우수성으로 인해 다양한 분야의 패턴인식 시스템에서 가장 많이 사용되는 분류기 중 하나이다[1-2]. 기존에 신경회로망을 기반으로 한 많은 분류기는 연

결가중치가 상수항으로 구성되어 있다. 상수항으로 구성된 연결가중치는 은닉층 뉴런의 활성화 함수와의 선형 결합을 통해 출력 공간 내에 선형 판별 함수(linear discernment function)로 표현된다. 이는 네트워크가 출력 공간 내에 선형 판정 경계(초평면)를 생성하도록 만들며, 선형적인 특성을 가지게 한다[2].

Particle Swarm Optimization 알고리즘은 유전 알고리즘(Genetic Algorithm:GA)과 같은 진화 연산기법이다. PSO 알고리즘은 GA의 자연선택의 진화 메커니즘이 아닌 새 떼와 물고기 떼와 같은 생체군집의 사회적 행동양식을 바탕으로 한 최적화 알고리즘이다.

접수일자 : 2007년 11월 10일

완료일자 : 2007년 12월 5일

본 연구는 산업자원부의 지원에 의하여 기초전력연구원(R-2007-2-044) 주관으로 수행된 과제임.

본 논문에서는 퍼지 추론 메커니즘에 기반 한 다항식 네트워크 패턴 분류기(Polynomial Network Pattern Classifier; PNC)를 설계한다. 제안된 PNC 구조는 신경회로망 분류기의 성능 향상을 위해 퍼지 추론 메커니즘을 네트워크 구조에 적용하였으며, 은닉층과 출력층 사이의 상수항 연결가중치를 다항식으로 표현함으로써 출력 공간 내의 선형 판정 경계를 비선형 판정 경계로 나타내고자 한다. 또한 제안된 PNC 구조는 FCM 클러스터링에 기반 한 분할 함수를 활성 함수로 사용한다. 이는 시그모이드 함수, 방사형 기저 함수를 사용한 활성 함수보다 입력 공간의 데이터 분포 특성을 잘 반영함으로써 빠른 학습 수렴과 성능 개선의 특징을 가진다.

제안된 PNC 구조의 파라미터, 즉, 학습률, 모멘텀 계수, FCM 클러스터링의 퍼지화 계수의 최적화를 위해 생체 군집 이론에 바탕을 둔 Particle Swarm Optimization(PSO)을 이용한다.

본 논문은 5개의 장으로 구성된다. 2장에서는 제안된 PNC 구조의 기능적 모듈 수행을 위한 네트워크 구조와 학습방법을 설명한다. 3장에서는 학습률, 모멘텀 계수, 퍼지화 계수의 최적화를 위한 PSO 알고리즘에 대하여 설명한다. 4장에서는 모의데이터 집합과 Iris 데이터 집합에 응용하여 분류기로서의 성능을 분석하였고 5장에서 본 연구의 결론을 맺는다.

2. 다항식 네트워크 패턴 분류기의 구조

본 논문에서 제안한 다항식 네트워크 패턴 분류기(Polynomial Network Pattern Classifier; PNC)는 활성 함수로 FCM 클러스터링에 의한 분할 함수를 사용하며, 비선형 특성의 판정 경계를 위해, 상수항 연결가중치가 아닌 다항식 연결가중치 $w=f(\cdot)$ 를 이용한다. 뿐만 아니라 제안된 PNC 구조는 기능적 모듈로 동작되며, 이는 퍼지 추론 메커니즘의 특징을 나타낸다.

2.1 퍼지추론 메커니즘과 다항식 네트워크 패턴분류기

본 논문에서 제안한 PNC 구조는 FCM 클러스터링 분할 함수와 다항식의 적용으로부터 식 (1)의 퍼지 규칙 표현과 같이 언어적 관점에서 해석될 수 있다.

$$\text{If } x \text{ is } A_i \text{ then } f_{ij}(x) \tag{1}$$

여기서, x 는 입력 벡터, A_i 는 FCM 클러스터링에 의한 i ($=1, \dots, c$)번째 그룹의 소속 함수(membership function or partition matrix), $f_{ij}(x)$ 는 j ($=1, \dots, s$)번째 출력에 대한 i 번째 퍼지 규칙의 다항식이다. 제안된 PNC 구조는 퍼지 규칙에 기반 한 네트워크 구조를 가지며, 조건부, 결론부, 추론부의 세 가지 기능적 모듈로 분리되어 동작한다. 그림 1은 기능적 모듈로서의 PNC 구조를 보여준다.

2.1.1 PNC의 조건부

FCM 클러스터링을 사용하는 PNC의 조건부 기능은 학습 데이터(training data)의 특성 반영을 위해 입력 공간을 c 개의 클러스터 수(퍼지 규칙 수)만큼의 로컬 영역으로 분리하고 각 로컬 영역의 소속정도를 퍼지 집합으로서 출력한다. FCM 클러스터 방법은 각 데이터와 특정 클러스터 중심과의 거리에 소속 정도를 부여하고 이 소속정도에 따라 데이터를 분류하는 알고리즘이다[3]. FCM 클러스터링은 목적함수인 식 (2)를 최소화함으로써 입력데이터의 각 클러스터에 대한

소속 값을 구한다.

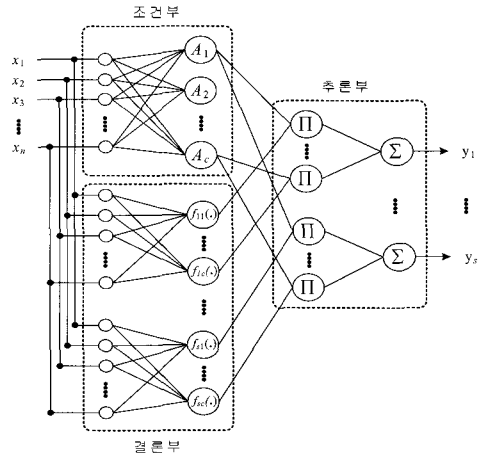


그림 1. 조건부, 결론부, 추론부의 세 가지 모듈로서 표현된 PNC의 구조

Fig. 1. Topology of PNC showing three modules of condition, conclusion, and inference phases

$$Q = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \| \mathbf{x}_k - \mathbf{v}_i \|^2 \tag{2}$$

여기서, c 는 클러스터의 개수(퍼지 규칙 수), N 은 입력 패턴 수, m 은 퍼지화 계수이고 m 은 1.0보다 크다. \mathbf{x}_k 는 k 번째 입력 벡터이고 \mathbf{v}_i 는 i 번째 클러스터의 중심이다. u_{ik} 는 k 번째 데이터가 i 번째 클러스터에 속하는 소속정도를 나타내는 0과 1사이의 실수이다. 식 (2)의 $\| \cdot \|$ 는 식 (3)으로 표현되는 가중 유클리디안 거리(weighted Euclidean distance)를 사용한다.

$$\| \mathbf{x}_k - \mathbf{v}_i \|^2 = \sum_{j=1}^n \frac{(x_{kj} - v_{ij})^2}{\sigma_j^2} \tag{3}$$

여기서, σ_j 는 입력 패턴들의 j 번째 입력 차원의 표준편차이다.

n 차원 유클리드 공간상의 N 개의 패턴으로 구성된 입력 벡터 집합 $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, $\mathbf{x}_k \in \mathbb{R}^n$, $1 \leq k \leq N$ 과 클러스터 중심 $\mathbf{v} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c\}$, $\mathbf{v}_i \in \mathbb{R}^n$, $1 \leq i \leq c$ 에 대한 소속정도를 소속행렬로 표현하면 $U = [u_{ik}]$ 로 표현되고 u_{ik} 와 \mathbf{v}_i 는 식 (4)와 식 (5)에 의해 각각 계산된다.

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{\| \mathbf{x}_k - \mathbf{v}_i \|^2}{\| \mathbf{x}_k - \mathbf{v}_j \|^2} \right)^{\frac{2}{m-1}}}, \quad 1 \leq k \leq N, 1 \leq i \leq c \tag{4}$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^N u_{ik}^m \mathbf{x}_k}{\sum_{k=1}^N u_{ik}^m}, \quad 1 \leq i \leq c \tag{5}$$

FCM 클러스터링은 식 (4)와 식 (5)를 반복적으로 수행하면서 소속행렬 U 와 각 클러스터의 중심 \mathbf{v}_i ($i=1, \dots, c$)를 수정하며 식 (2)의 목적함수 $Q(U, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c)$ 를 특정 값으로 수렴시킨다.

퍼지화 계수 “ m ”은 각 클러스터의 소속 함수의 형태를 결

정하는 매우 중요한 요소이다. 그림 2는 클러스터가 2개일 때 퍼지화 계수에 따른 FCM 클러스터링을 통하여 생성된 소속 함수의 형태를 보여준다. 이러한 퍼지화 계수의 특징은 PNC의 소속 함수의 형태를 결정하는 중요한 요소이다. 본 논문에서는 PSO알고리즘으로 입력 데이터 집합과 규칙 수에 따른 최적화된 퍼지화 계수를 동조한다.

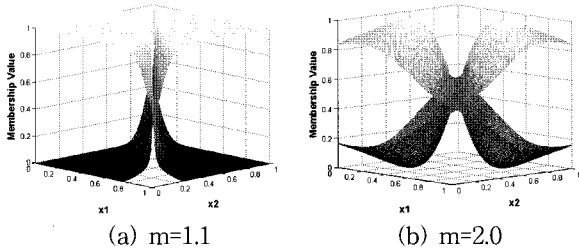


그림 2. 퍼지화 계수(m)에 따른 소속 함수 형태
Fig. 2. The form of membership function according to fuzzification coefficient

2.1.2 PNC의 결론부

PNC 구조의 결론부는 조건부에서 분리한 각 로컬 영역을 다항식 함수의 로컬 회기모델로서 표현하여 식 (1)의 “then” 이후의 규칙을 형성한다. 식 (1)의 $f_{ji}(x)$ 에서 출력 뉴런에 대한 첨자 $j(j=1, \dots, s)$ 를 생각한 $f_i(x)$ 는 식 (6)-(10)의 형태를 갖는 세 가지 타입의 함수 중 하나의 형태를 갖는다. 즉, 로컬 회기 모델은 상수항, 일차식 또는 이차식으로 표현되어 진다. $f_i(x)$ 가 식 (6)의 상수항인 경우 FCM 클러스터링을 통한 소속 함수를 활성 함수로 사용한 RBFNNs[1]와 동일한 구조를 갖는다.

Constant: $f_i(x) = a_{i0}$ (6)

Linear: $f_i(x) = a_{i0} + \sum_{j=1}^n a_{ij}x_j$ (7)

Quadratic: $f_i(x) = a_{i0} + \sum_{j=1}^n a_{ij}x_j + \sum_{j=1}^n \sum_{k=j}^n a_{ijk}x_jx_k$ (8)

식 (6)~(8)의 다항식 함수들은 조건부의 소속 함수(활성 함수)에 의해 활성화되며 각 퍼지 규칙의 로컬 회기 모델로 동작한다. 본 논문에서 제안한 PNC는 기존의 식 (6)의 형태를 갖는 상수항의 연결가중치를 식 (7)~(8)형태의 1차식과 2차식으로 확장함으로서 은닉층 뉴런의 출력간의 비선형 결합으로 인해 분류기로서의 성능을 향상시킨다.

2.1.3 PNC의 추론부

추론부에서는 “If-then” 퍼지 규칙 기반의 퍼지 추론에 의해 네트워크의 최종출력을 구하게 된다. 그림 1의 추론부의 “II”로 표기된 뉴런은 각 퍼지 규칙의 소속 함수와 다항식 로컬모델의 곱을 의미하며, “Σ”으로 표기된 뉴런에 의해 입력 신호들은 합하여져 그 결과를 출력층 뉴런의 최종 출력으로 내보낸다. 이와 같은 일련 과정은 퍼지 추론 과정과 같으며, Fuzzy Neural Networks[4]와 같은 형태를 갖게 된다. 결론적으로 제안된 PNC 구조를 보인 그림 1에서 $j(j=1, \dots, s)$ 번째 출력의 최종출력은 퍼지 추론에 의한 식 (9)와 같이 표현된다.

$$y_j = g_j(x) = \sum_{i=1}^c \frac{u_i f_{ji}(x)}{\sum_{k=1}^c u_k} = \sum_{i=1}^c u_i f_{ji}(x) \quad (9)$$

여기서, u_i 는 그림 1의 $A_i(x)$ 와 같고 모든 클러스터에 대한 총합은 1이다. 식 (9)의 $g_j(x)$ 는 PNC의 j 번째 클래스에 대한 판별 함수가 된다.

2.2 PNC의 학습

본 절에서는 PNC의 식 (9)의 판별 함수 $g_j(x)$ 를 생성하기 위한 PNC학습 과정을 설명 한다. PNC의 학습은 경사 하강법과 모멘텀 방법을 통해 이루어진다. 계수 동정을 위한 학습은 식 (10)으로 표현되는 오차를 모든 패턴에 대하여 최소화하도록 진행된다.

$$E_q = \frac{1}{2}(t_q - y_q)^2, \quad q=1, \dots, N \quad (10)$$

여기서, E_q 는 q 번째 패턴에 대한 오차, N 은 입력 패턴 수이다. t_q 는 q 번째 패턴에 대한 목적 값이다. y_q 는 q 번째 패턴에 대한 네트워크의 최종출력이며 식 (11)과 같다.

$$y_q = \sum_{i=1}^c u_i f_i \quad (11)$$

PNC의 학습은 f_i 의 네 가지 타입에 따라 3가지의 경우로 나뉜다.

i) 결론부 다항식 함수가 Constant 타입

다항식 함수가 Constant인 경우 $f_i = a_{i0}$ 이다. a_{i0} 를 구하기 위해 식 (10)의 E_q 와 a_{i0} 의 관계를 사슬법칙(Chain rule)으로 나타내면 식 (12)와 같다. 식 (12)의 우변의 첫째항, 둘째항, 셋째항을 미분하여 정리하면 각각 식 (13)-(15)와 같이 정리된다.

$$\frac{\partial E_q}{\partial a_{i0}} = \frac{\partial E_q}{\partial y_q} \cdot \frac{\partial y_q}{\partial f_i} \cdot \frac{\partial f_i}{\partial a_{i0}} \quad (12)$$

$$\frac{\partial E_q}{\partial y_q} = (-1)(t_q - y_q) \quad (13)$$

$$\frac{\partial y_q}{\partial f_i} = u_i \quad (14)$$

$$\frac{\partial f_i}{\partial a_{i0}} = 1 \quad (15)$$

식 (13)-(15)를 사용한 계수 a_{i0} 의 변화분(delta)은 식 (16)과 같고 모멘텀까지 고려한 최종 식은 식 (17)과 같다.

$$\Delta a_{i0} = -\eta \cdot \frac{\partial E_q}{\partial a_{i0}} = \eta(t_q - y_q)u_i \quad (16)$$

$$a_{i0}(p+1) = a_{i0}(p) + \eta(t_q - y_q)u_i + \alpha(a_{i0}(p) - a_{i0}(p-1)) \quad (17)$$

여기서 η 는 학습률, α 는 모멘텀 계수, $p+1$ 은 새로운 값, p 는 현재 값, $p-1$ 은 이전 값을 나타낸다.

ii) 결론부 다항식 함수가 Linear 타입

다항식 함수가 Linear인 경우 f_i 는 식 (18)과 같다. Constant 타입의 경우와 마찬가지로 체인 룰을 구성하고 미분하면 a_{i0} 와

a_{ij} 의 변화분은 식 (16)과 식 (19)와 같고 모멘텀까지 고려한 최종 식은 식 (17)과 식 (20)으로 표현된다.

$$f_i = a_{i0} + \sum_{j=1}^n a_{ij}x_j \quad (18)$$

$$\Delta a_{ij} = -\eta \cdot \frac{\partial E_q}{\partial a_{ij}} = \eta(t_q - y_q)u_i x_j \quad (19)$$

$$a_{ij}(p+1) = a_{ij}(p) + \eta(t_q - y_q)u_i x_j + \alpha(a_{ij}(p) - a_{ij}(p-1)) \quad (20)$$

iii) 결론부 다항식 함수가 Quadratic 타입

다항식 함수가 Quadratic인 경우 f_i 는 식 (21)과 같다. 앞선 경우와 마찬가지로 a_{ijk} 의 변화분에 대해 식 (22)와 같이 정리되며, 모멘텀을 고려한 최종 식은 식 (23)과 같다. a_{i0} 와 a_{ij} 의 학습에 대해 식 (17)과 식 (20)을 각각 이용한다.

$$f_i = a_{i0} + \sum_{j=1}^n a_{ij}x_j + \sum_{j=1}^n \sum_{k=j}^n a_{ijk}x_j x_k \quad (21)$$

$$\Delta a_{ijk} = -\eta \cdot \frac{\partial E_q}{\partial a_{ijk}} = \eta(t_q - y_q)u_i x_j x_k \quad (22)$$

$$a_{ijk}(p+1) = a_{ijk}(p) + \eta(t_q - y_q)u_i x_j x_k + \alpha(a_{ijk}(p) - a_{ijk}(p-1)) \quad (23)$$

3. Particle Swarm Optimization

Particle Swarm Optimization(PSO)는 Kennedy and Eberhart[5]에 의해 처음 소개된, 새 떼와 물고기 떼와 같은 생체군집의 사회적 행동양식을 바탕으로 한 최적화 알고리즘이다. 일반적으로 PSO는 이론의 간결성, 구현의 용이성, 연산의 효율성과 같은 특징을 보인다[6]. 본 절에서는 PSO 알고리즘에 관하여 설명하며 PSO의 연산자는 참고문헌 [5-6]에 자세히 설명되어있다.

3.1 PSO 알고리즘

PSO 알고리즘에서 swarm은 m 개의 particle을 가지며 이들은 주어진 문제의 해가 된다. 각 particle은 n 차원 실수 벡터로 구성된다. PSO 알고리즘은 다음과 같다.

[Step 1] *Initialization* - m 개의 particle을 탐색 공간 내에서 랜덤하게 생성한다. 이들은 초기 swarm이 된다. 마찬가지로 초기 particle velocity를 $[-v_{kmax}, v_{kmax}]$ 내에서 랜덤하게 생성한다. 초기 swarm의 각 particle은 목적함수를 이용하여 평가되고, 이들 모두는 $pbest$ 로 설정한다. 초기 particle 중 최적의 값은 다시 $gbest$ 로 선정된다.

[Step 2] *Inertia weight & velocity updating* - 관성하중 값을 계산하고, j 번째 particle velocity를 계산한다. 제한 조건 $[-v_{kmax}, v_{kmax}]$ 을 확인한다.

[Step 3] *Position updating* - 식 (24)에 의해 particle의 위치정보는 조절된다.

$$x_{jk}(t) = v_{jk}(t) + x_{jk}(t-1) \quad (24)$$

여기서, $x_k^{\min} \leq x_{jk}(t) \leq x_k^{\max}$ 이다.

[Step 4] *Individual & global best updating* - 새로운

위치정보를 가진 particle들은 목적함수에 의해 평가된다. 각 particle의 적합도는 이전 $pbest$ 의 적합도와 비교되며 $pbest$ 를 재설정한다. $pbest$ 의 particle 중 최적 해를 가지는 particle의 적합도와 이전 $gbest$ 의 적합도를 비교하여 최적 위치정보를 가진 particle을 $gbest$ 로 재설정한다.

[Step 5] *Stopping criteria* - 종료조건을 만족하면 탐색 과정을 종료하고, 그렇지 않으면 [Step 2]부터 반복한다.

[Step 6] *Optimal parameter* - 최종적으로 생성된 $gbest$ 는 최적의 위치정보를 가진다.

본 논문에서 제안된 PNC 구조의 파라미터인 학습률과 모멘텀 계수, 퍼지화 계수의 선택 문제는 경사 하강법과 FCM 클러스터링을 사용하는 PNC의 성능에 크게 영향을 미치는 중요한 문제이다. 이는 입력 데이터의 특성과 분류기 모델의 특성을 고려해야 하는 어려운 문제로서 기존 연구에서 명확한 해결책이 제시되지 못하고 있다. 이러한 이유로 본 논문에서는 PSO 알고리즘을 사용하여 학습률과 모멘텀 계수, 퍼지화 계수를 최적 동조하고 PNC의 분류기로서의 성능을 향상 시킨다.

4. 실험 및 결과

본 논문에서는 제안된 다항식 네트워크 패턴 분류기 (Polynomial Network Pattern Classifier; PNC)의 성능평가를 위해 모의 데이터 집합과 기계 학습 데이터 집합에 적용하여 분류기로서의 성능을 평가 및 분석한다. 모의데이터 집합에 대해 2-클래스 분류문제와 3-클래스 분류문제를 다루며, 기계 학습 데이터 집합은 UCI Machine Learning Repository(www.ics.uci.edu/~mllearn/MLRepository.html)로부터 획득한 Iris Dataset을 이용하여 기존 연구에서 제안된 분류기 기법들과 비교 평가 한다. 기계 학습 데이터 집합에 대한 실험은 각 데이터 집합을 7:3의 비율로 랜덤하게 학습 데이터와 테스트 데이터로 나누어 7번 반복 실험하며, 성능평가는 7-Fold Cross-Validation에 의한 분류 정확성의 평균과 표준편차에 의한다. 제안된 PNC 구조의 최적화를 위해 사용된 PSO 알고리즘의 파라미터들은 표 1과 같다.

표 1. PNC의 최적화에 사용된 PSO 파라미터
Table 1. Parameters of PSO

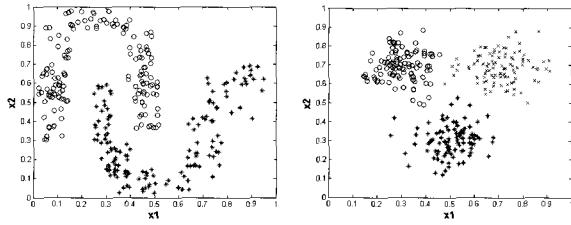
파라미터	값	
세대 수	60	
particle 수	50	
v_{max}	탐색공간의 20%	
$[wmin \ wmax]$	[0.4 0.9]	
c_1, c_2	2.0	
탐색 범위	학습률	[1e-8, 0.9]
	모멘텀 계수	[1e-8, 0.9]
	퍼지화 계수	[1.1, 3.0]

4.1 Synthetic dataset

모의 데이터 집합은 총 300개의 패턴을 갖고 있으며, 2-클래스 데이터 집합은 각 클래스가 150개 패턴으로 구성되고 3-클래스 데이터 집합은 각 클래스에 100개의 패턴으로 구성된다. 그림 3은 실험에 사용된 데이터 집합의 분포를 보여

준다. 그림 3의 Type A는 말발굽 형태의 2-클래스 모의 데이터 집합이며 Type B는 선형 분류가 어려운 3-클래스 모의 데이터 집합이다.

모의 데이터 집합에 대한 실험은 규칙 수(은닉층 뉴런 수)를 2-5개에 대하여 실험하였고 학습 횟수는 300번, 학습률 0.01, 모멘텀 계수 0.01, 퍼지화 계수 2.0을 사용하여 실험하였다. 표 2는 모의 데이터 집합에 대한 실험결과로 패턴 인식률을 보여준다. 괄호 안은 300개의 패턴 중 오 분류(error)한 패턴 수이고, RBFNNs 항은 연결가중치가 상수항으로 이루어진 RBF 신경회로망에 대한 결과이다.



(a) Type A (b) Type B

그림 3. 모의 데이터 집합

Fig. 3. Synthetic dataset

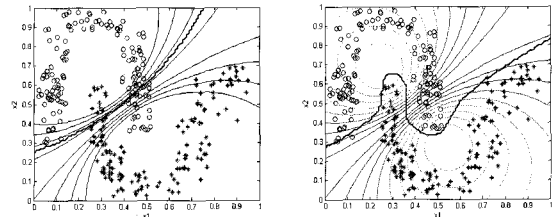
표 2의 대부분의 경우에 RBFNNs보다는 L-PNC와 Q-PNC의 패턴 분류결과가 더 좋게 나타났으며, L-PNC 보다는 Q-PNC가 조금 더 좋은 성능을 보였다. 표 2의 굵은 글씨체의 결과는 데이터 Type A, B에 대하여 가장 우수한 성능을 보인 패턴 인식률이며 Q-PNC가 가장 좋은 성능을 보임을 알 수 있다.

표 2. 모의 데이터 실험 결과

Table 2. The result for the experiment of the Synthetic dataset

데이터 타입	규칙수	RBFNNs	L-PNC	Q-PNC
Type A	2	87.0% (39)	88.3% (35)	99.3% (2)
	3	84.6% (46)	94.0% (18)	96.6% (1)
Type A	4	90.0% (30)	99.6% (1)	100% (0)
	5	91.6% (25)	98.0% (6)	100% (0)
Type B	2	78.3% (65)	99.3% (2)	99.3% (2)
	3	99.3% (2)	99.3% (2)	99.6% (1)
	4	99.0% (3)	99.3% (2)	99.6% (1)
	5	98.6% (4)	99.3% (2)	99.6% (1)

그림 4는 두 가지 타입의 데이터 집합에 대한 규칙 2개인 경우의 판정 경계와 소속 함수의 등고선을 보여준다. RBFNNs의 판별 함수는 소속 함수의 선형 결합으로 생성되며 그림 4(a)의 판정경계를 생성하고 그 결과로서 패턴 인식률이 87%로 300개중 39개의 패턴을 오 분류 하였다. 이에 반해 같은 규칙수를 갖고 있는 Q-PNC는 소속 함수의 비선형 결합으로 판별 함수가 생성되어 그림 4(b)와 같이 데이터 분포 특성에 적합한 판정 경계를 생성함으로써 99.3%의 패턴 인식률을 보이며 오 분류한 패턴은 2개에 그쳤다.



(a) RBFNNs(2 Rules) (b) Q-PNC(2 Rules)

그림 4. 규칙 2개인 경우의 소속 정도와 판정 경계

Fig. 4. Boundary of classification and membership value

4.2 Iris dataset

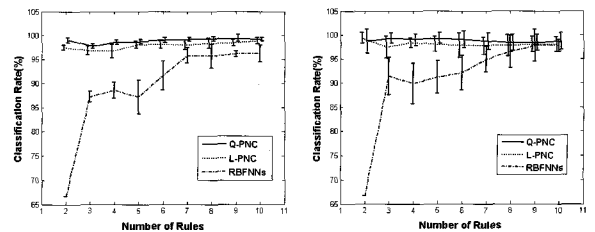
Iris 데이터 집합은 붓꽃의 세 가지 종(subspecies)을 판별하는 기계 학습 데이터 집합이다. Iris 데이터 집합은 각 클래스마다 50개의 패턴으로 구성되어 총 150개의 패턴으로 구성된다. Iris 데이터 집합의 분류 실험을 위해 PNC의 규칙 수를 2개에서 10개 까지 실험하였으며 각 규칙수마다 7-Fold Cross-Validation 시뮬레이션 방법을 통하여 평균과 표준편차를 구하였다. 또한 PNC의 최적화를 위해 PSO를 사용하여 학습률, 모멘텀 계수, 퍼지화 계수를 동정하였고 이때 사용된 PSO 파라미터는 앞에서 언급된 표 1과 같다. 그림 5는 실험한 모든 규칙 2-10개에 대한 패턴 인식률의 평균과 표준편차를 보여준다.

표 3. Iris 데이터 집합 패턴분류 결과

Table 3. The result of pattern classification

Number of Rules		Training dataset	Testing dataset
		AVG ±STD	AVG ±STD
2	RBFNNs	66.66 ± 0	66.66 ± 0
	L-PNC	97.55 ± 0.5	99.36 ± 1.08
	Q-PNC	99.04 ± 0.54	98.73 ± 2.51
9	RBFNNs	96.19 ± 0.54	97.77 ± 3.39
	L-PNC	98.5 ± 0.74	98.09 ± 1.53
	Q-PNC	99.45 ± 0.74	98.41 ± 1.67

단위: 패턴인식률 %



(a) 학습 데이터 결과 (b) 테스트 데이터 결과

그림 5. 규칙 수에 따른 패턴 분류 결과

Fig. 5. Classification Rate versus the number of rules

그림 6은 학습 데이터에 대한 패턴 분류 성능이 가장 좋은 Q-PNC(9 Rules)와 테스트 데이터에 대한 성능이 가장 좋은 L-PNC(2 Rules)의 PSO의 세대 수가 증가함에 따라서 더 적합한 학습률, 모멘텀 계수, 퍼지화 계수를 동조하여 PNC를 최적화 시키는 과정을 보여준다.

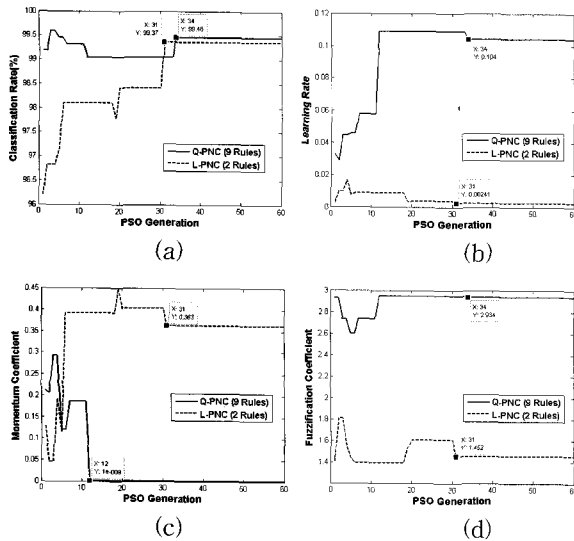


그림 6. PSO 세대수 증가에 따른 패턴인식률(a), 학습률(b), 모멘텀 계수(c), 퍼지화 계수(d)의 변화
 Fig. 6. The variance of Classification Rate(a), Learning Rate(b), Momentum Coefficient(c), Fuzzification Coefficient(d) according to PSO Generation

표 4는 제안된 PNC와 다양한 패턴인식 모델들과의 패턴 인식률의 평균 성능 비교표이다. 제안된 PNC를 기존에 연구 발표된 7가지 모델들과 비교하였다.

표 4. 다양한 패턴분류 모델의 평균 성능 비교
 Table 4. Comparison of the average performance of several classification models

Classifier Model	Classification Rate(%)	
C4.5 [7]	94.0	
FID3.1 [8]	96.0	
NEFCLASS [9]	96.0	
HNFB [10]	98.67	
HNFBQ [11]	98.67	
HNFB-1 [12]	98.67	
RBFNNs [13]	90.1	
Our Model	For Training	For Testing
L-PNC(2 Rules)	97.55±0.5	99.36±1.08
Q-PNC(6 Rules)	99.04±0.54	99.04±1.74

5. 결론

본 논문에서는 퍼지 추론에 기반 한 다항식 네트워크 패턴 분류기(Polynomial Network Pattern Classifier; PNC)를 설계하고 PSO 알고리즘을 이용하여 학습률, 모멘텀 계수, 퍼지화 계수를 동조하였다. 제안된 PNC 구조는 언어적 관점으로 부터 FCM 클러스터링을 통한 퍼지 공간 분할은 퍼지 규칙의 조건부로, 각 로컬 영역의 표현인 다항식 함수는 결론부로, 퍼지 추론에 의한 최종 출력은 추론부로 기능적 모듈 동작을 수행한다. 제안된 PNC의 패턴 분류 성능 분석을 위

해, 모의 데이터 집합과 기계학습 데이터 집합들에 적용하여 다양한 패턴 분류 모델들과 비교함으로써 제안된 PNC의 우수한 성능을 증명하였다. 제안된 PNC는 네트워크 구조적인 면에서 볼 때, 기존의 상수형 연결가중치를 입력 공간 변수들의 다항식으로 확장하여 기존의 신경회로망 기반 분류기들의 선형적 특성을 비선형 특성으로 개선하였다. 또한 FCM 클러스터링을 통해 입력 공간의 데이터 분포에 적합한 소속 함수를 은닉층의 활성화 함수로 사용함으로써 학습이 빠르고 패턴 분류 성능이 보다 좋은 분류기를 설계하였다. 제안된 PNC는 다항식 기반 구조의 퍼지 추론 특성으로 신경 회로망에 기반 한 기존 분류기들의 선형성을 개선하여 우수한 패턴 분류 성능을 가진다. 제안된 PNC는 여러 패턴 인식 시스템에서의 다양한 응용이 기대된다.

참고 문헌

- [1] W. Pedrycz, "Conditional fuzzy clustering in the design of radial basis function neural networks," *IEEE Trans. Neural Networks*, vol. 9, pp. 601 - 612, July 1998.
- [2] M.J. Er, S.Q. Wu, J.W. Lu, H.L. Toh, "Face recognition with radical basis function (RBF) neural networks," *IEEE Trans. Neural Networks*, vol. 13, No. 3, pp. 697-710, 2002.
- [3] A. Aiyer, K. Pyun, Y.Z. Huang, D.B. O'Brien, R.M. Gray, "Lloyd clustering of Gauss mixture models for image compression and classification," *Signal Processing: Image Communication*, vol. 20, pp. 459-485, 2005.
- [4] S.-K. Oh, W. Pedrycz, B.-J. Park, "Self-organizing neurofuzzy networks in modeling software data," *Fuzzy Sets and Systems*, vol. 145, pp. 165-181, 2004.
- [5] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proc. IEEE Int. Conf. Neural Networks*, vol. 4, pp. 1942-1948, 1995.
- [6] J. Kennedy, "The particle swarm: Social adaptation of knowledge," *Proc. IEEE Int. Conf. Evolutionary Comput.*, pp. 303-308, 1997.
- [7] J. R. Quinlan, "Improved use of continuous attributes in C4.5," *J. Artif. Intell. Res.*, vol. 4, pp. 77 - 90, 1996.
- [8] C. Z. Janikow and M. Faifer, "Fuzzy partitioning with FID3.1," in *Proc. IEEE 18th Int. Conf. North American Fuzzy Information Processing Soc.*, pp. 467 - 471, 1999.
- [9] F. Klawonn, D. Nauck, and R. Kruse, "Generating rules from data by fuzzy and neuro-fuzzy methods," in *Proc. Fuzzy-Neuro-Systeme*, pp. 223 - 230, 1995.
- [10] F. J. de Souza, M. M. B. R. Vellasco, and M. A. C. Pacheco, "Load forecasting with the hierarchical neuro-fuzzy binary space partitioning model," *Int. J. Comput. Syst. Signals*, vol. 3, no. 2, pp. 1 - 15, 2002.
- [11] F. J. de Souza, M. M. B. R. Vellasco, and M. A.

C. Pacheco, "Hierarchical neuro-fuzzy quadtree models," *Fuzzy Sets Syst.*, vol. 130, No. 2, pp.189 - 205, 2002.

[12] L.B. Gonçalves, M.M.B.R. Vellasco, M.A.C. Pacheco and F.J. de Souza, "Inverted hierarchical neuro-fuzzy BSP system: A novel neuro-fuzzy model for pattern classification and rule extraction in databases," *IEEE Transactions on Systems, Man & Cybernetics, Part C*, Vol. 26, No. 2, 2005.



오성권(Sung-Kwon Oh)
1981년 : 연세대학교 전기공학과 졸업
1983년 : 동 대학원 전기공학과(공학석사).
1983~1989년 : 금성산전연구소(선임연구원).
1993년 : 연세대 대학원 전기공학과 졸업 (공학박사).

1996~1997년 : 캐나다 Manitoba대학 전기 및 컴퓨터공학과 Post-Doc.

1993~2005년 : 원광대 전기전자 및 정보공학부 교수

2005년~현재 : 수원대 전기공학과 교수.

2002년~현재 : 대한전기학회, 퍼지및지능시스템학회 및 제어자동화시스템 공학회 편집위원.

관심분야 : 시스템 자동화, 퍼지이론, 신경회로망 응용 및 제어, 컴퓨터 지능 등.

E-mail : ohsk@suwon.ac.kr

저 자 소 개



김길성(Gil-Sung Kim)
2007년 : 수원대학교 컴퓨터학과 졸업
2007년~현재 : 동 대학원 전기공학과 석사 과정

관심분야 : 패턴인식, 퍼지이론, 최적화 이론 등

E-mail : gilsung@suwon.ac.kr



박병준(Byoung-Jun Park)
1998년 : 원광대 제어계측공학 졸업.
2000년 : 동 대학원 졸업(공학석사).
2003년 : 동 대학원 졸업(공학박사).
2004년~2005년 : 원광대 전기전자및정보공학부 강의교수.
2005년~2007년 : Postdoctoral fellow in Dept. of Electrical and Computer Eng., University of Alberta.

2007년~현재 : (주)지엠테크 기술연구소 소장.

관심분야 : CI, 뉴로퍼지컴퓨팅, 퍼지이론, 지능형 모델링 및 제어, 패턴인식 등.

E-mail : lcap2005@gmail.com