

GA기반 TSK 퍼지 분류기의 설계 및 응용

The Design of GA-based TSK Fuzzy Classifier and Its application

곽근창, 김승석, 유정웅, 전명근

충북대학교 전기전자 및 컴퓨터 공학부

Keun-Chang Kwak, Seung-Suk Kim, Jeong-Woong Ryu, Myung-Geun Chun
School of Electrical and Electronics Engineering, Chungbuk University.
(kckwak@power.chungbuk.ac.kr)

ABSTRACT

In this paper, we propose a TSK-type fuzzy classifier using PCA(Principal Component Analysis), FCM(Fuzzy C-Means) clustering and hybrid GA(genetic algorithm). First, input data is transformed to reduce correlation among the data components by PCA. FCM clustering is applied to obtain a initial TSK-type fuzzy classifier. Parameter identification is performed by AGA(Adaptive Genetic Algorithm) and RLSE(Recursive Least Square Estimate). we applied the proposed method to Iris data classification problems and obtained a better performance than previous works.

Keywords : GA, TSK Fuzzy Classifier, FCM clustering, PCA, Iris data classification

I. 서론

패턴 분류는 많은 공학적인 문제에 있어서 중요한 분야로 영상처리, 음성인식, 수중 음파 탐지기, 레이더, 지진학, 진단 시스템에서 어떤 상황을 정확하게 분류하기 위한 능력을 요구하는 학문에 적합한 분야이다. 자동 제어와 예측 시스템은 종종 입출력 관계를 결정하기 위해 분류기를 이용한다. 이런 폭 넓은 응용 가능성 때문에 패턴 분류는 지금까지도 많이 연구가 진행되어지고 있다. 지금까지 연구되어 왔던 패턴 분류 알고리즘 중 전형적인 통계적 기법은 불확실성과 애매함을 가진 데이터를 다룰 때 부적절한 결과를 보이며 이 단점은 분류 문제에서 좀 더 유연한 기법에 대한 연구로 한층 고조되었다. 퍼지 집합의 이론은 이러한 단점을 해결하기 위한 방법으로서 제안되었으며 패턴 분류와의 결합은 많은 연구가들에 의해 연구되어왔다. 또한 신경회로망은 결합극복 능력, 병렬계산 능력, 학습능력 때문에 분류문제에 대해 통계적 기법에 비해 좋은 성능을 보였다. 그러나 신경회로망은 역전파 학습방법이 최대 경사법이기 때문에 전역적인 최소점이 아닌 지역적인 최소점에 빠지는 문제점이 발생할 수 있다. 또한 학습률과 노드의 수 등 초기 변수의 설정

이 어려운 단점을 가지고 있다. 이런 문제점을 해결하기 위해 퍼지 집합의 유연성과 신경회로망의 계산 능력의 융합을 통해 보다 좋은 성능을 보이고 있다. 그러나 이 기법은 오차를 최소화하도록 하여 파라미터를 조정할 수는 있지만 잘못 분류된 개수에 대한 정보를 첨가하지 않기 때문에 최종적으로 오차가 적을지라도 잘못 분류된 개수가 많아질 수 있는 문제점을 가지고 있다.

따라서, 본 논문에서는 유전알고리즘과 FCM[1] 클러스터링을 이용한 TSK 형태의 퍼지 분류기를 구축하여 Iris 데이터 분류 문제에 적용하여 유용성을 보이고자 한다[4-6].

II. 본론

본 절에서는 유전알고리즘과 FCM 클러스터링을 이용한 TSK 퍼지 분류기를 구축하기 위해 두 과정인 구조 동정과 파라미터 동정이 수행된다. 먼저, 구조동정에 대하여 살펴보면 다음과 같다. 먼저, 수치적인 입출력 데이터가 주어질 때 0과 1사이로 정규화를 수행한 다음, 주 성분 분석기법을 이용한 입력 데이터 성분간의 상관 관계를 제거하여 입력공간을 변환한다. 주 성분 분석기법은 통신이론과 영상처리 분야에

서 K-L변환으로서 알려진 기법으로 입력을 선형적으로 결합하고 입력공간의 차원을 줄이기 위한 접근방법으로 알려져 있다. 우선, 입력 데이터 집합을 다음 식과 같이 공분산 행렬을 계산하고 고유치와 고유벡터를 구한다.

$$C = \frac{1}{N} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T \quad i=1,2,\dots,n \quad (1)$$

여기서, $\mu = \frac{1}{N} \sum_{i=1}^n x_i$ 은 각 입력데이터 성분들의 평균이며 x_i 는 i 번째 입력 데이터이다. 구해진 고유벡터와 입력데이터의 선형결합에 의해 새로운 입력 데이터로 변형되어진다.

주성분 분석기법에 의해 변환된 입력공간을 FCM 클러스터링에 의해 스캐터 분할하여 적은 규칙의 수를 생성하도록 초기 퍼지 시스템 (FCM-ANFIS)을 구축한다[2].

그림 1은 FCM-ANFIS의 변형된 구조로 1층 이전에 새로운 층을 추가하여 상관 관계를 제거 할 수 있도록 나타내고 있다.

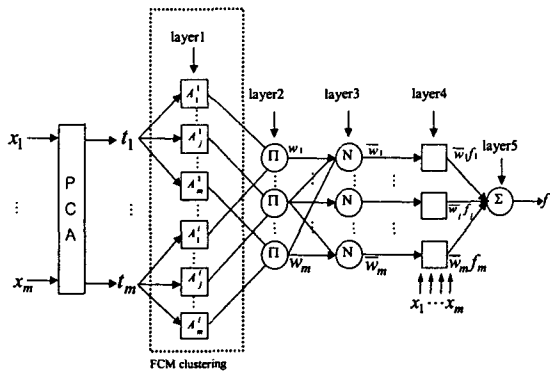


그림 1. FCM-ANFIS의 구조

FCM 클러스터링에 의해 얻어진 각 클러스터는 가우시안 소속함수의 중심이 된다. 따라서, 각 클러스터는 시스템의 특성을 기술하는 퍼지 규칙으로 되며, 다음 식과 같이 TSK 형태의 퍼지 규칙이 사용된다.

$$R^i = \text{If } t_1 \text{ is } A_1^i \text{ and } \dots t_m \text{ is } A_m^i \text{ then } f^i = a_0^i + a_1^i t_1 + \dots + a_m^i t_m \quad (2)$$

여기서, $\{a_0^i, a_1^i, \dots, a_m^i\}$ 는 결론부 파라미터이며 t_1, \dots, t_m 은 주성분 분석기법에 의해 변환된 입력이다. 퍼지 모델의 추론 값 f 는 가중 평균법에 의해서 구해지며 결론부 파라미터는 전체부의 가우시안 소속함수가 정해지면 RLSE에 의해 추정되어진다.

위와 같이 초기 FCM-ANFIS을 구축한 후에 유전알고리즘과 RLSE에 의한 하이브리드 GA를 이용하여 파라미터 동정을 수행한다. 여

기서 사용된 유전 알고리즘[3]은 적합도 값에 근거해서 교차 확률과 돌연변이 확률을 적응적으로 변화시켜 집단의 다양성과 전역적인 최적해의 수렴 속도를 가속화 할 수 있다. 이때 구축된 FCM-ANFIS의 전체부 파라미터인 가우시안 소속함수의 중심과 폭을 초기치 집단으로 발생시켜 유전알고리즘에 의해 최적에 가깝도록 탐색되어진다. 결론부 파라미터는 선형방정식을 이루기 때문에 전체부 파라미터가 결정된 후에 RLSE에 의해 추정되어진다. 그로 인해 유전 알고리즘의 탐색공간이 줄어들어 빨리 최적에 가깝도록 원하는 해를 찾아낼 수 있다.

유전알고리즘에 의해 퍼지 시스템이 구축되는 수행순서는 다음과 같다.

[단계 1] 세대수, 집단수, 교차확률, 돌연변이확률, 비트 수, 범위설정, 적합함수를 문제 고유의 특성에 알맞게 선택한다.

[단계 2] 초기 FCM-ANFIS에서 얻어진 전체부 파라미터를 초기치 집단으로 발생시킨다. 이때 결론부 파라미터는 RLSE에 의해 추정되며 가중평균법에 의해 퍼지 모델의 출력 값인 f 을 계산한다.

[단계 3] 각 개체에 대한 적합함수 F 는 다음 식과 같이 계산한다.

$$F1 = \frac{1}{1 + \text{MSE}_{tm} + \text{MSE}_{chk}} \quad (3)$$

$$F2 = \frac{1}{(\text{misclass}_{tm} + \text{misclass}_{chk})/N} \quad (4)$$

$$F = F1 + F2 \quad (5)$$

여기서 misclass_{tm} 과 misclass_{chk} 은 각각 잘못 분류된 학습데이터와 검증데이터의 개수이며 N 은 총 데이터의 수이다.

[단계 4] GA 탐색결과와 수렴성을 보장하는 방법으로 엘리티즘 선택법을 이용하여 최상위의 두 개체를 선택하여 새로운 집단에 유지시킨다.

[단계 5] 적합도의 값이 양의 값을 유지하도록 $F^* = F - \min(F)$ 와 같이 스케일링을 한다.

[단계 6] 선택된 부모개체로부터 교배 연산을 수행한다. 이때 교배확률은 다음 식과 같이 적합도의 값에 따라 적응적으로 변한다.

$$P_c = k_1(f_{\max} - f_i') / (f_{\max} - \bar{f}), \quad f_i' \geq \bar{f} \quad (6)$$

$$P_c = k_2, \quad f_i' < \bar{f} \quad k_1, k_2 \leq 1.0 \quad (7)$$

[단계 7] 돌연변이는 엘리티즘 선택법에 의해 선택된 두 개체에는 영향을 주지 않도록 수행한다. 돌연변이 확률도 적응적으로 변화하도록 다음 식과 같이 계산되어진다.

$$P_m = k_3(f_{\max} - f_i') / (f_{\max} - \bar{f}), \quad f_i' \geq \bar{f} \quad (8)$$

$$P_m = k_4, f_i' < \bar{f} \quad k_3, k_4 \leq 1.0 \quad (9)$$

[단계 8] 설정한 세대수에 도달하거나 스케일링된 값의 합이 0이 되면 하나의 값에 수렴하기 때문에 종료한다. 만약에 그렇지 않을 경우에는 [단계3]으로 가서 반복 수행한다.

III. 시뮬레이션 및 결과

이 절에서는 제안된 퍼지 분류기를 Iris 데이터 분류 문제에 적용하고 컴퓨터 시뮬레이션을 통해 제안된 퍼지 분류기가 이전의 다른 논문들과 비교하여 좋은 성능을 보이고자 한다. Fisher의 Iris 데이터는 전형적인 패턴 인식과 분류 문제의 예로서 많은 연구가들에 의해 활발히 연구되어지고 있다. 데이터의 집합은 총 150개의 패턴으로 이루어졌으며 네 개의 특징은 다음과 같다.

$$\begin{aligned} &\text{sepal length(SL), sepal width(SW),} \\ &\text{petal length(PL), petal length(PW)} \end{aligned} \quad (10)$$

이 네 특징들은 Iris 꽃의 모양과 크기를 기술하고 있으며, 데이터 집합에서 각 패턴들은 다음과 같이 세 가지 클래스로 분류되어진다.

$$\begin{aligned} &\text{iris setosa(class 1), iris versicolor(class 2),} \\ &\text{iris virginica(class 3)} \end{aligned} \quad (11)$$

각 클래스는 50개의 패턴을 가지고 있으며 클래스 1은 나머지 두 클래스로부터 선형적으로 분리가능하며 클래스 2와 클래스 3는 각각으로부터 선형적으로 분리하지 못하는 특성을 가지고 있다.

그림 2는 네 개의 특징에 대해 Iris 데이터의 분포도이며 그림에서 알 수 있듯이 클래스 1은 클래스 2와 클래스 3으로부터 쉽게 분리할 수 있지만 클래스 2와 클래스 3는 서로 중첩되어 있기 때문에 분류하기가 어렵다는 것을 알 수 있다. 이 Iris 데이터 분류문제에 있어서 기존의 논문들과 비교하기 위해 총 150개의 데이터 중 임의로 75개의 데이터를 학습데이터로 선택하여 분류기를 구축하고 나머지 임의의 75개 데이터를 검증데이터로 선택하여 검증에 사용되어 분류성능을 알아본다. 이때 학습 데이터와 검증 데이터 모두 각 클래스에 대해 25개의 데이터가 선택되도록 한다. 그림 3은 주성분 분석기법에 의해 변환된 입력 데이터 분포도이다. 상관관계 제거와 정규화 단계를 수행한 후 FCM 클러스터링에 클러스터를 구한다. 여기서 최소의 규칙의 수는

각 클래스에 대해 하나의 규칙을 선택함으로써 얻을 수 있기 때문에 세 개의 퍼지 규칙을 생성하도록 한다. 클러스터가 선택되면 전체부 파

라미터를 고정시키고 RLSE에 의해 초기치가 0인 파라미터를 추정한다.

패턴 분류를 수행하기 위해 TSK 퍼지 분류기의 출력 클래스 Class k 는 다음과 같은 분류영역을 가진다.

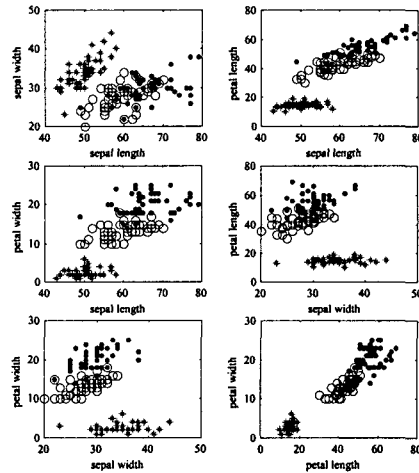


그림 2. Iris 데이터 분포도

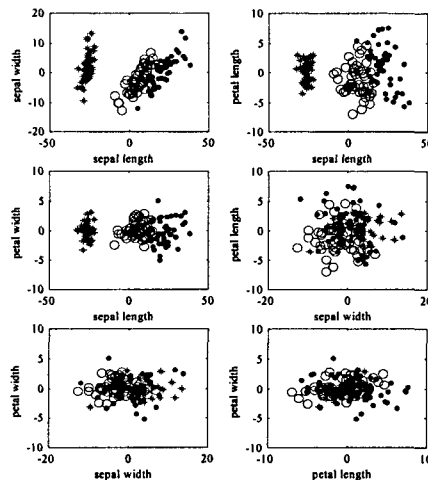


그림 3. PCA에 의해 변형된 입력 데이터 분포도

$$\text{Class } k = \begin{pmatrix} 1, & \text{if } f < 1.5 \\ 2, & \text{if } 1.5 \leq f < 2.5 \\ 3, & \text{if } 2.5 \leq f \end{pmatrix} \quad (12)$$

초기 TSK 퍼지 분류기를 가지고 유전알고리즘을 적용할 경우 제어 파라미터는 세대 100, 초기 교차확률 0.95, 초기 돌연변이 확률 0.01, 집단 30, 비트 16로 설정하였다.

본 논문에서는 20번의 시뮬레이션을 수행하였으며 각 경우의 잘못 분류된 개수는 표 1에서 기존의 논문과 비교되어진다. 표에서 알 수 있듯이 이전의 다른 연구보다 우수한 분류 성능을 보여주고 있음을 알 수 있었다.

표 1. 이전 연구와의 분류 성능 비교

	사용된 방법	잘못 분류된 개수	평균(%)	비고
Hong	fuzzy expert system	평균3.3개	95.57%	6.21 rules
Chen	multiclass neural network	1~5(평균3.2)	95.7%	5.7 hidden node(평균)
Simpson	fuzzy min-max neural network	2	97.3%	48 hidden nodes
Nauck	neuro-fuzzy	학습: 3개 검증: 2개	학습: 96% 검증:97.3%	7 rules
		학습: 2개 검증: 3개	학습:97.3% 검증: 96%	3 rules
Wu	fuzzy system	평균2.8개	96.21%	3 rules
Gabrys	general FMMN	0~6개	100~92%	5 hyperboxes
Meesad	Neuro-Fuzzy	2~9(평균 4.3)	88~97.3%	11.8rules(평균)
Yen	Neuro-Fuzzy	1~5개	학습:96.7% 검증:96.2%	6.1 hidden nodes
Lee	Fuzzy system	0~5(평균2.48)	96.7%	all 4 features
		0~4(평균2.16)	97.12%	2 selected features
제안된 방법	Initial TSK fuzzy classifier	학습: 0~3 (평균0.9) 검증: 0~5 (평균2.05)	학습:98.8% 검증:97.3%	3 rules
	GA-TSK fuzzy classifier	학습:0~1 (평균0.15) 검증:0~1 (평균0.5)	학습:99.8% 검증:99.3%	3 rules

III. 결론

본 논문에서는 FCM 클러스터링과 유전 알고리즘을 이용하여 TSK 퍼지 분류기를 구축하였다. 제안된 분류기는 초기 FCM-ANFIS를 구축하고 얻어진 전체부 및 결론부 파라미터를 초기 집단으로 발생시켜 AGA와 RLSE에 의한 하이브리드 GA를 이용하여 파라미터 동정을 수행한다.

시뮬레이션 결과 제안된 방법은 이전의 다른 연구에 비해 좋은 분류 성능을 보임을 알 수 있었다.

감사의 글 : 본 연구는 정보통신부 대학기초연구 지원사업에 의해 수행되었습니다

IV. 참고문헌

[1] J. Bezdek, R. Hathaway, M. Sabin *Convergence theory for fuzzy c-means : Counter examples and repairs*, CRC Press, 1987.

[2] K. C. Kwak, M. G. Chun, J. W. Ryu, "FCM-based Adaptive Fuzzy Inference System for Coagulant Dosing Process in a Water Purification Plant", *Journal of Knowledge-Based Intelligent Engineering Systems*, Vol.4, No.4, pp. 230-236, 2000.

[3] M. Srinivas, L. M. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms", *IEEE Trans. on Systems, Man, Cybern.*, Vol.24, No.4, 1994.

[4] T. P. Hong, C. Y. Lee, "Induction of fuzzy rules and membership functions from training examples", *Fuzzy Sets and Systems*, Vol. 84, pp. 33-47, 1996.

[5] K. H. Chen, H. L. Chen, H. M. Lee, "A multiclass neural network classifier with fuzzy teaching inputs", *Fuzzy Sets and Systems*, Vol. 91, pp. 15-35, 1997.

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