

독립 성분 분석기법과 뉴로 퍼지를 이용한 비선형 시스템 모델링

Nonlinear System Modeling using Independent Component Analysis and Neuro-Fuzzy Method

김성수 · 곽근창* · 유정웅*

Sung-Soo Kim, Keun-Chang Kwak* and Jeong-Woong Ryu*

우석대학교 전기공학과, *충북대학교 전기전자공학부

요 약

본 논문에서는 적응 뉴로 퍼지 모델링을 위해 최근에 BBS(blind source separation)분야에서 발전된 독립 성분 분석기법(ICA)을 전처리로 이용하여 효과적인 퍼지 규칙을 생성하는 방법을 제안한다. 기존의 뉴로 퍼지 모델링은 입력 데이터 성분간의 상관관계를 고려하지 않고 입력공간을 분할하기 때문에 효과적으로 분할하지 못하는 단점이 있다. 이로 인해 과도한 규칙 수와 큰 오차를 가지고 있었다. 이에, 본 연구에서는 독립 성분 분석기법을 이용하여 입력 데이터 성분간의 상관관계를 제거함으로써 적은 규칙 수를 갖으면서도 효율적인 퍼지 규칙을 얻을 수 있도록 하였다. 시뮬레이션 예로서 Box-Jenkins의 가스 온도 데이터의 모델링에 적용하여 유용성과 제안된 방법이 이전의 연구보다 좋은 결과를 보임을 알 수 있었다.

ABSTRACT

In this paper, an efficient fuzzy rule generation scheme for adaptive neuro-fuzzy system modeling using the Independent Component Analysis(ICA) as a preprocessing is proposed. Correlation between inputs was not considered in the conventional neuro-fuzzy modeling schemes, such that enormous number of rules and large amount of error were unavoidable. The correlation between inputs is weakened by employing ICA so that the number of rules and the amount of error are reduced. In simulation, the Box-Jenkins furnace data is used to verify the effectiveness of the proposed method.

1. 서 론

퍼지이론은 1965년 Zadeh에 의해 제안된 이후로 오늘날 우리가 추구하는 인공지능 세계를 실현하는 하나의 방안으로 주목받고 있다[1]. 그 동안 이론적인 단계에서 연구되어오던 퍼지 이론은 1980년대 들어서면서 여러 분야에서 널리 응용되기 시작하면서 많은 연구가 활발히 진행되고 있다. 퍼지 기법은 정량적으로 표현하기 힘든 전문가의 경험적 지식을 소속함수와 퍼지 규칙으로 구성하여 효과적으로 시스템을 묘사할 수 있는 특성을 갖고 있다[2]. 그러나 고도의 비선형 특성을 갖는 시스템에 대한 퍼지 규칙은 전문가의 경험적 지식과 실험에 의해 구성되므로 상당히 주관적이고 경험에 의존하게 되어 최적의 규칙 베이스를 결정하기가 쉽지 않은 문제점을 가지고 있다. 이러한 문제점을 해결하기 위해 수치적인 입출력 데이터로부터 효과적인 규칙을 생성하기 위해 신경회로망과 퍼지의 융합인 뉴로 퍼지 모델링에 관한 연구가 활발

히 진행되고 있다[3-5]. 이러한 모델링은 신경회로망의 학습능력, 적응성과 퍼지의 if-then 규칙에 의한 근사 추론 능력으로 인해 비선형 시스템을 잘 묘사하지만 입력 데이터 성분간의 상관관계를 고려하지 않고 입력공간을 분할하기 때문에 효과적으로 분할하지 못하는 단점을 가지고 있다. 이런 이유로 과도한 규칙의 수와 큰 오차를 보이고 있다. 이러한 문제점을 해결하기 위해 Kim[18]은 주성분 분석기법(Principal Component Analysis)을 이용하여 입력공간을 변환하여 퍼지 모델링을 수행하였다. 그러나 주성분 분석기법은 상관관계를 어느정도 제거해주고 있지만 여전히 성분들이 종속적인 문제점을 가지고 있다.

따라서, 본 논문에서는 최근에 BBS(blind source separation)분야(예를 들면 speech separation, array antenna processing, multisensor biomedical records, financial market data analysis)에서 발전된 독립 성분 분석(Independent Component Analysis)기법[7,20]을 전처리로 이용하여 효율적인 퍼지 규칙을 생성하는 방

법을 제안한다. 독립 성분 분석기법은 주성분 분석기법의 확장으로서 신호들간의 상관관계뿐만 아니라 종속성까지도 최소가 되도록 한다. 이를 이용해 입력 데이터 성분간의 상관관계를 제거한 다음 그 데이터를 통해 이전의 논문인 FCM(Fuzzy c-means) 클러스터링을 이용한 ANFIS(Adaptive Neuro-Fuzzy Inference System)[8]의 소속함수와 규칙을 생성시킨다. 이 기법은 입력 데이터 성분간에 상관관계가 존재하는 시스템일 경우 상관관계를 제거할 수 있기 때문에 다른 기존의 뉴로-퍼지 모델링에도 적용될 수 있는 기법이다. 2장에서는 독립 성분 분석기법을 전처리로 한 뉴로-퍼지 모델링에 대한 알고리즘을 제안하며 3장에서는 이렇게 제안된 방법을 Box-Jenkins의 가스로 데이터의 모델링[8-18]에 적용하여 이전에 연구된 방법들에 비하여 좋은 결과를 얻음을 보이고자 한다.

2. 독립 성분 분석기법 (ICA) 을 전처리로 한 뉴로 - 퍼지 모델링

본 단원에서는 독립 성분 분석기법을 전처리로 하여 입력 데이터 성분간의 상관관계를 제거한 다음, FCM 클러스터링을 이용한 ANFIS에 적용하는 과정을 설명하면 다음과 같다.

먼저 독립 성분 분석기법의 메카니즘에 대해서 간략히 살펴보자[7]. 독립성분 분석기법은 선형적으로 혼합된 둘 이상의 신호들을 서로 독립적인 신호들로 분리하는 방법으로서 n 개의 독립 성분의 선형 결합인 x_1, x_2, \dots, x_n 을 아래의 식과 같이 관측하였다고 가정하자.

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \dots + a_{jn}s_n \quad (1)$$

식 (1)을 벡터-행렬식으로 변환하면

$$x = As \quad (2)$$

여기서 A 를 a_{ij} 요소를 가진 정방행렬이고 s 는 s_1, \dots, s_n

요소를 가진 독립성분 벡터이다. 독립성분 벡터 s 와 혼합 행렬 A 는 모두 미지의 것이며 오직 관측벡터 x 만을 이용할 수 있다. 관측벡터 x 가 주어질 때 독립 성분 벡터 s 가 식 (3)에서 정의되는 출력벡터 y 로부터 복원되도록 행렬 W 을 찾는 것이다.

$$y = Wx \quad (3)$$

여기서 $y = [y_1, y_2, \dots, y_n]^T$ 이다.

여기에서 기술되는 문제를 흔히 BBS라고 하며 blind라는 용어는 독립 성분 신호를 복원하기 위해 사용되는 정보는 관측벡터 x 에 포함되어진다는 사실을 의미한다. 이 문제를 해결하기 위해 사용된 기법이 독립 성분 분석기법(ICA)이며 이것은 주성분 분석(PCA)의 확장으로서 보여질 수 있다. PCA는 직교성(Orthogonality)을 가지는 방향벡터로 제한하며 오직 2차까지만 독립성을 부여할 수 있는 반면, ICA는 출력 벡터 y 의 개별적인 성분에 관해 통계적인 독립성을 부여하며 직교성의 제한조건을 가지지 않는다.

ICA에 대한 학습 알고리즘은 식 (4)와 같이 natural 경사법(Gradient Method)[20]을 사용하였으며, 이 알고리즘에 대한 일반적인 안정도 분석은 Amari[19]에 의해서 제시되어졌다.

$$W(t+1) = W(t) + \eta(t)[I - \phi(y(t))y^T(t)]W(t) \quad (4)$$

여기서 $\eta(t)$ 는 학습률이고 A 는 단위 행렬이며 $\phi(y(t))$ 는 활성화함수(Activation Function)이다.

그림 1에서는 BBS 문제에 대한 블록 다이어그램을 보여주고 있으며, BSS는 자율학습 신경회로망으로 이루어져 있으며 출력 데이터 없이 입력 데이터만을 가지고 학습을 수행한다. 그림 2에서는 ICA와 PCA를 이용해서 얻은 복원된 데이터의 분포를 비교하고 있다. 그림 2(b)의 데이터 분포는 관측벡터 x 를 의미하고 이것으로부터 그림 2(a)의 독립성분인 s 를 찾아야 한다. 그림(c)(d)에서 알 수 있는 바와 같이 PCA에

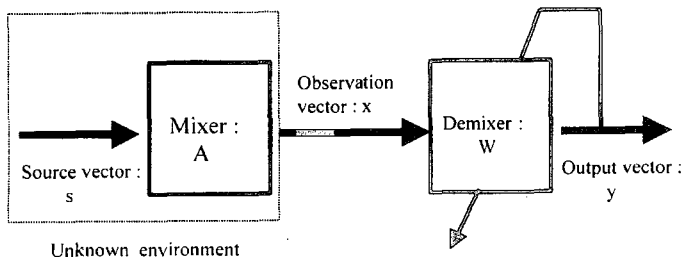


그림 1. Blind source separation에 대한 블록 다이어그램
Fig. 1. Block diagram for the blind source separation

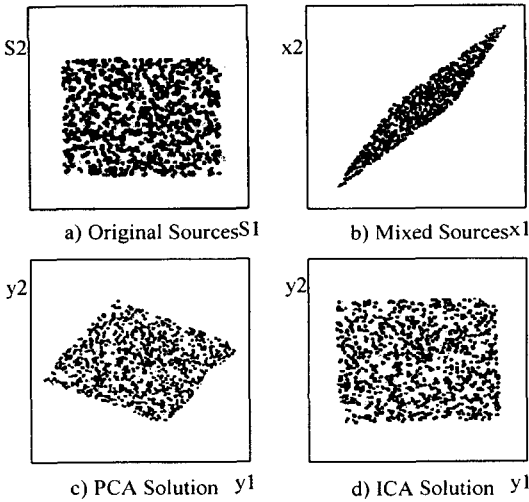


그림 2. PCA와 ICA의 비교
Fig. 2. The comparison of PCA and ICA

의해 복원된 데이터의 분포보다 ICA를 통한 데이터가 그림 2(a)와 매우 흡사하다는 것을 알 수 있다. 즉 PCA보다 ICA가 더 효과적으로 상관관계가 제거됨을 알 수 있다.

이렇게 ICA를 통해 상관관계가 제거된 다음 FCM을 이용한 ANFIS의 구조동정(Structure Identification)이 이루어진다[8].

FCM에 의해 구조동정이 결정된 후에 파라미터 동정(Parameter Identification)은 Jang[5]의 방법과 같이 최소자승법(Least Square Estimate)과 역전파 알고리즘(Back Propagation)을 이용한 하이브리드 학습에 의해 수행되어진다. 여기서 ANFIS 모델링과 학습 메카니즘을 간략히 살펴보자. 먼저 아래와 같은 두 개의 입력과 Sugeno-Takagi 형태의 퍼지 규칙을 갖는 추론 시스템에 대해 생각해보자.

$$R^1 : \text{If } y_1 \text{ is } A_1 \text{ and } y_2 \text{ is } B_1 \text{ then } f_1 = p_1 y_1 + q_1 y_2 + r_1$$

$$R^2 : \text{If } y_1 \text{ is } A_2 \text{ and } y_2 \text{ is } B_2 \text{ then } f_2 = p_2 y_1 + q_2 y_2 + r_2$$

(5)

여기서 y_1 와 y_2 는 입력이고 A_i, B_i 는 언어적 변수이며 $\{p_i, q_i, r_i\}$ 는 결론부 파라미터이다. 위와 같은 퍼지 규칙에 대한 학습절차는 다음과 같다. 우선 뉴로-퍼지 시스템을 통한 추론 값은 식 (6)과 같이 가중평균법에 의해 주어진다.

$$f_i^* = \frac{\sum w_i f_i}{\sum w_i}, i = 1, 2$$

(6)

여기서 \bar{w}_i 은 $w_i/w_1 + w_2$ 이고, w_i 은 $\mu_{A_i}(y_1) \times \mu_{B_j}(y_2)$ 이다.

전제부 소속함수로 다음과 같은 가우시안 소속함수를 선택하였다.

$$\mu_{A_i}(y) = \exp\left\{-\left(\frac{y_1 - c_i}{a_i}\right)^2\right\}$$

(7)

여기서 $\{a_i, c_i\}$ 는 전제부 파라미터이며, a_i 는 소속함수의 폭이고 c_i 는 소속함수의 중심이다. 전제부 파라미터에 관한 학습 알고리즘을 위해 유도하기 위해 최소화하기 위한 목적함수는 다음과 같다.

$$E = \frac{1}{2}(f^* - f_d)^2$$

(8)

여기서 f_d 는 입력벡터 $x = [y_1, y_2]^T$ 에 대한 원하는 데이터 출력이고 f^* 는 뉴로-퍼지 모델의 출력이다. 소속함수의 모양이 전제부 파라미터에 의해 정의되어지기 때문에 목적함수 E 는 결론부 파라미터가 고정된다면 $\{a_i, c_i\}$ 의 파라미터로 결정되어진다. 목적함수 E 의 값을 감소시키기 위해서는 전제부 파라미터에 대해서 아래와 같이 최대경사법에 의해 반복적으로 이루어진다.

$$a_j^i(t+1) = a_j^i(t) - \eta \frac{\partial E}{\partial a_j^i}$$

$$c_j^i(t+1) = c_j^i(t) - \eta \frac{\partial E}{\partial c_j^i}$$

(9)

다른 한편으로, 결론부 파라미터는 최소자승법에 의해 추정되어진다.

정리하면, 전방향 경로에서는 최소자승법으로 후방

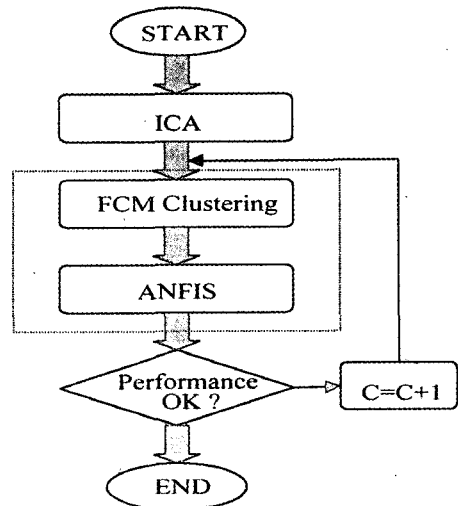


그림 3. ICA를 전처리로 한 FCM-ANFIS 모델링 과정
Fig. 3. The process of FCM-ANFIS modeling based on preprocessing by ICA

향 경로에서는 역전과 알고리즘으로 전제부 및 결론부 파라미터를 조정하여 실제 데이터 출력과 퍼지 모델 출력사이의 오차를 최소화하는 것이다. 그림 3은 ICA와 뉴로 퍼지를 이용한 모델링을 순서도로 나타내었다. 여기서 점선부분은 뉴로 퍼지의 구조 동정과 파라미터 동정이 수행되어진다.

3. 시뮬레이션 및 결과

본 논문에서 제안한 방법으로 Box-Jenkins의 가스로 데이터의 모델링에 적용하여 이전에 연구된 방법들에 비하여 보다 나은 결과를 얻을 수 있음을 보이고자 한다. 여기서는 가스로에서의 가스 흐름을 $u(t)$ 와 이산화탄소 농도 $y(t)$ 시계열 데이터인 296 $\{y(t)$,

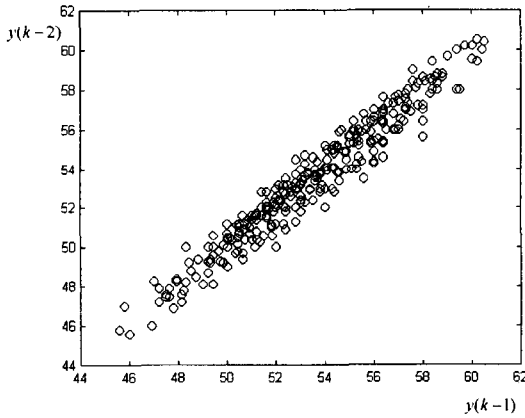


그림 4. 상관관계가 존재하는 입력 $y(k-1), y(k-2)$ 의 분포도
Fig. 4. The data distribution of $y(k-1)$ and $y(k-2)$ in the presence of the correlation

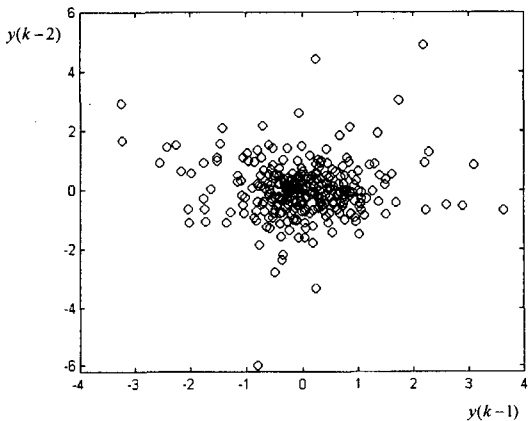


그림 5. ICA를 통해 상관관계가 제거된 입력 $y(k-1), y(k-2)$ 의 분포도
Fig. 5. The data distribution of $y(k-1)$ and $y(k-2)$ after ICA

$u(t)$ 개개의 쌍이 사용된다. 본 논문에서는 이전의 다른 논문과 비교하기 위해 입력을 $u(k), u(k-1), u(k-2)$ 와 $y(k-1), y(k-2), y(k-3)$ 으로 선택하고 출력 변수는 $y(t)$ 를 선택하였다.

그림 4는 입력데이터 $y(k-1), y(k-2)$ 의 분포도이며 상관관계가 존재함을 알 수 있고 그림 5는 ICA를 이용하여 입력데이터 성분간의 상관관계를 제거한 분포도이다.

그림 6은 제안된 방법을 이용하여 모델의 추론 값과 원하는 데이터 출력을 비교하고 있다. 그림에서 알 수 있듯이 적은 오차를 보이면서 예측하고 있음을 알 수 있다. 그림 7은 Box-Jenkins 가스로 데이터의 모델링 오차를 보여주고 있다.

표 1은 시뮬레이션 결과로 모델에 대한 규칙 수와 오차를 비교한 결과이며 제안된 방법이 이전의 다른

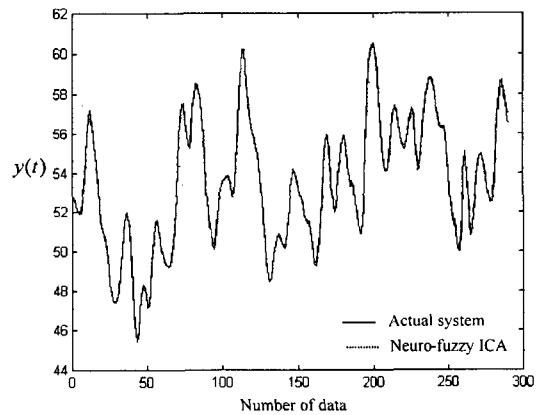


그림 6. 제안된 모델의 추론 값과 실제 데이터 출력과의 비교
Fig. 6. The comparisons of the proposed model inference and desired output value

표 1. 규칙 수와 오차의 비교

Table 1. The comparison of rule numbers and mean square error

| | 입력 수 | 규칙 수 | 평균제곱오차 |
|------------------------|------|------|--------|
| Box and Jenkins[9] | 6 | - | 0.71 |
| Tong[10] | 2 | 19 | 0.469 |
| Xu[11] | 2 | 25 | 0.328 |
| Pedrycz[12] | 2 | 20 | 0.320 |
| Sugeno and Yasukawa[2] | 3 | 6 | 0.190 |
| Joo[13] | 2 | 6 | 0.166 |
| Kwak[8] | 2 | 6 | 0.117 |
| Wang[14] | 6 | 2 | 0.066 |
| Sugeno and Tanaka[15] | 6 | 2 | 0.068 |
| Lin[16] | 5 | 4 | 0.071 |
| Kim[17] | 6 | 2 | 0.055 |
| Kim(PCA)[18] | 6 | 2 | 0.048 |
| This model | 6 | 2 | 0.041 |

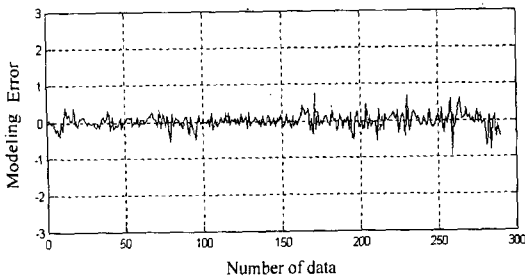


그림 7. 제안된 모델의 추론 값과 실제 데이터 출력과의 오차
 Fig. 7. The mean square error of the proposed model inference and desired output value

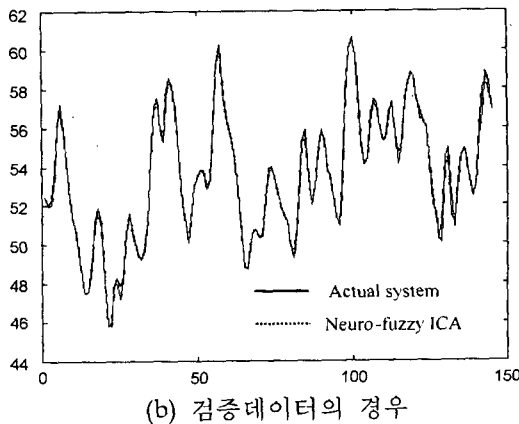
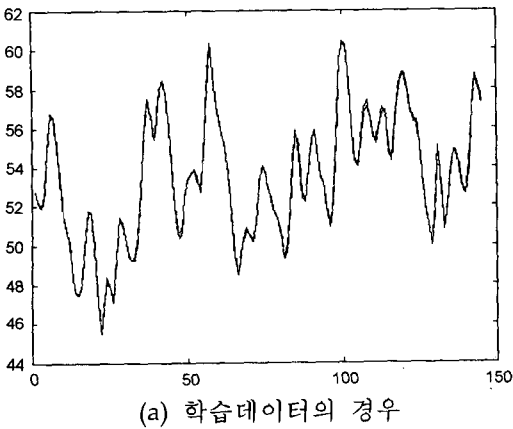


그림 8. 학습데이터와 검증데이터에 대한 예측결과 비교
 Fig. 8. The comparison of prediction result for training data and checking data

논문에 비해 적은 규칙의 수와 입력 수를 가지고도 작은 오차를 보임을 알 수 있었다. 표에서 제시된 평균제곱오차는 300epoch 동안 학습한 결과이고 오랜 학습 없이 1epoch일때의 오차도 0.0576이므로 다른 논문에 비해 유용함을 보여주고 있다. 그림 8은 시스

템의 출력이 모델 형성을 위한 학습데이터에 편중되지 않고 새로운 데이터에 대해서도 타당한 결과를 보일 수 있는가를 평가하기 위해 데이터 집합을 학습데이터와 검증데이터로 나누어 예측한 결과이다. 그림은 300epoch동안 학습한 결과이며 두 경우 모두 적은 오차를 보이면서 예측하고 있음을 알 수 있었다.

4. 결 론

비선형 특성을 갖는 시스템에 대한 퍼지 규칙은 전문가의 경험적 지식에 의해 구성되므로 주관적이고 경험에 의존하게 되어 최적의 규칙 베이스를 결정하기가 쉽지 않은 문제점을 가지고 있다. 이러한 문제점을 해결하기 위해 수치적인 입력력 데이터로부터 효과적인 규칙을 생성하기 위해 신경회로망과 퍼지의 융합인 뉴로-퍼지 모델링에 관한 연구가 활발히 진행되고 있다. 이 방법은 비선형 시스템을 잘 묘사하지만 입력 데이터 성분간의 상관관계를 고려하지 않고 입력공간을 분할하기 때문에 효과적으로 분할하지 못하는 단점을 가지고 있다.

이런 문제점을 해결하기 위해 본 논문에서는 주성분 분석기법의 확장인 독립성분 분석기법을 뉴로-퍼지 모델의 전처리로서 이용하여 입력 데이터 성분간의 상관관계를 제거하였다. 전처리된 데이터는 FCM 클러스터링에 근거한 ANFIS의 입력으로 사용되어 입력공간을 분할함으로써 효율적인 퍼지 규칙을 생성할 수 있도록 하였다. 또한 이 방법은 전처리뿐만 아니라 입력공간이 분할된 각 부분공간에 존재하는 상관관계를 제거하여 효과적으로 퍼지 규칙을 생성할 수 있다.

위의 방법을 모델링 기법의 적용 예로 널리 쓰이는 Box-Jenkins의 가스르 데이터에 적용하여 시뮬레이션해 본 결과 이전의 다른 논문에 비해 적은 규칙과 오차로 좋은 모델링 특성을 보임을 알 수 있었다. 따라서, 제안된 방법은 입력데이터 성분간에 상관관계가 존재하는 여러 분야에 유용하게 적용될 수 있으리라 생각한다.

참고문헌

- [1] L. A. Zadeh, "Fuzzy sets", *Information Control*, Vol. 8, pp. 338-353, 1965.
- [2] M. Sugeno, T. Yasukawa, "A fuzzy-logic based approach to qualitative modeling", *IEEE Trans. on Fuzzy Systems*, Vol. 3, No. 1, pp. 7-31, 1993.
- [3] J. S. R. Jang, C. T. Sun and E. Mizutani, "Neuro-Fuzzy and Soft Computing : A Computational Approach to Learning and Machine Intelligence", Prentice Hall, 1997.

- [4] C. T. Lin, C. S. G. Lee, "Neural Fuzzy Systems : A Neuro-Fuzzy Synergism to Intelligent Systems", Prentice Hall, 1996.
- [5] L. X. Wang, "A Course in Fuzzy Systems and Control", Prentice-Hall, 1997.
- [6] E. T. Kim, M. K. Park, S. W. Kim and M. G. Park, "A Transformed Input-Domain Approach to Fuzzy Modeling", *IEEE Trans. on Fuzzy Systems*, Vol. 6, No. 4, pp. 596-604, 1998.
- [5] J. S. R. Jang, "ANFIS : Adaptive-Networks-based Fuzzy Inference System", *IEEE Trans. on System, Man, and Cybern.*, Vol. 23, No. 3, pp. 665-685, 1993.
- [6] C. T. Lin, C. S. G. Lee, "Reinforcement structure parameter learning for neural network-based fuzzy logic control system", *IEEE Trans. on Fuzzy System*, Vol. 2, No. 1, pp. 46-63, 1994.
- [7] T. W. Lee, "Independent Component Analysis: Theory and Applications", Kluwer Academic Publishers, 1998.
- [8] 광근창, 한태환, 유정웅, 전명근, "정수장 응집제 주입공정의 적응네트워크기반 퍼지 시스템 모델링", 한국퍼지 및 지능 시스템학회 논문지, 제9권, 3호, pp. 313-322, 1999.
- [9] G. E. P. Box, G. M. Jenkins, "Time Series Analysis, Forecasting and Control", San Francisco, CA : Holden Day, 1970.
- [10] R. M. Tong, "The evaluation of fuzzy models derived from experimental data", *Fuzzy Sets and Systems*, Vol. 4, pp. 1-12, 1980.
- [11] C. W. Xu, Y. Z. Lu, "Fuzzy model identification and self-learning for dynamics systems", *IEEE Trans. on System, Man, and Cybern.*, Vol. 17, pp. 683-689, 1987.
- [12] W. Pedrycz, "An identification algorithm in fuzzy relational systems", *Fuzzy Sets and Systems*, Vol. 13, pp. 153-167, 1984.
- [13] Y. H. Joo, H. S. Hwang, K. B. Kim and K. B. Woo, "Fuzzy system modeling by fuzzy partition and GA hybrid schemes", *Fuzzy Sets and Systems*, Vol. 86, pp. 279-288, 1997.
- [14] L. Wang, R. Langari, "Building sugeno-type models using fuzzy discretization and orthogonal parameter estimation techniques", *IEEE Trans. on Fuzzy Systems*, Vol. 3, pp. 454-458, 1995.
- [15] M. Sugeno, K. Tanaka, "Successive identification of a fuzzy model and its applications to prediction of a complex system", *Fuzzy sets and Systems*, Vol. 42, pp. 315-334, 1991.
- [16] Y. Lin, G. A. Cunningham, "A new approach to fuzzy-neural modeling", *IEEE Trans. on Fuzzy Systems*, Vol. 3, pp. 190-197, 1995.
- [17] E. Kim, M. Park, S. Ji, and M. Park, "A new approach to fuzzy modeling", *IEEE Trans. on Fuzzy Systems*, Vol. 5, pp. 328-337, 1997.
- [18] E. Kim, M. Park, S. Ji, and M. Park, "A Transformed Input-Output Approach to fuzzy modeling", *IEEE Trans. on Fuzzy Systems*, Vol. 6, pp. 596-604, 1998.
- [19] S. Amari, T. P. Chen and A. Cichocki, "Stability analysis of learning algorithms for blind source separation", *Neural Networks*, Vol. 10, pp. 1345-1351, 1997
- [20] S. Amari, A. Cichocki and H. H. Yang, "A new learning algorithm for blind source separation", In *Advances in Neural Information Processing* Vol. 8, pp. 757-763, 1996.



김 성 수 (Sung-Soo Kim)

1983년 : 충북대 공대 전기공학과 졸업 (B.S.)
 1989년 : University of Arkansas-Fayetteville(M.S.)
 1997년 : University of Central Florida(Ph.D.)



광 근 창 (Keun-Chang Kwak)

제10권 3호 참조



유 정 웅 (Jeong-Woong Ryu)

제10권 3호 참조