

# 자율조직 CMAC 신경망에 의한 비선형 시계열 예측

## Prediction of Nonlinear Sequences by Self-Organized CMAC Neural Network

이 태 호  
Tai-Ho Lee

### 요 약

SOCMAC 신경망에 의하여 Mackey-Glass의 비선형 시계열 예측을 시도하였다. 다차원 연속 입력 변수를 가지는 문제는 요구되는 기억용량의 규모가 너무 커서 CMAC에서는 일반적으로 취급이 곤란한 대상이었으나 SOCMAC에서는 이것이 가능함을 보였다. 또한 학습과정에서 수용영역(receptive field)을 가변으로 하는 개선된 방법을 제시하였다. 예측오차는 TDNN(time-delayed neural network)이나 BP(back-propagation) 수준이었다.

### Abstract

An attempt of using SOCMAC neural network for the prediction of a nonlinear sequence, which is generated by Mackey-Glass equation, is reported. The report shows the SOCMAC can handle a system with multi-dimensional continuous inputs, which has been considered very difficult, if not impossible, task to be implemented by a CMAC neural network because of a huge amount of memory required. Also, an improved training method based on the variable receptive fields is proposed. The performance ranged somewhere around those of TDNN and BP neural networks.

Key words: CMAC, SOCMAC, Self-organization, Neural Network, Prediction

### 1. 서 론

Menozzi 와 Chow는 CMAC(Cerebella Model Articulation Controller)의 입력부에 Kohonen의 SOM기법<sup>1)</sup>을 도입해서 자율조직 CMAC (Self-Organized CMAC, SOCMAC) 구조를 제안하였다[1]. 그런데 이 제안자들은 그 후, SOM의 복잡성이 CMAC의 장점, 즉 속도와 선형성을 손상한다는 이유로 SOCMAC의 이용가치를 부인하고 다른 형태의 입력공간 조직기법을 제안하고 있으며[2], 이래 SOCMAC에 관한 연구는 발표되지 않고 있다. 그러나 [2]의 제안은 SOM보다 성능 면에서는 뒤떨어지는 것이며 입력조직에 관한 한 SOCMAC의 논리적 배경이 가장 탁월하므로 문제점이 있더라도 그 성격과 능력을 구명해 볼 가치가 있다. 본 연구에서는 SOCMAC의 시계열 예측 능력을 평가하여 CMAC의 활용영역을 확장하고자 하였으며, SOCMAC에 맞는 개선된 학습과정으로서 가변 수용영역 방식을 제안하였다.

원래 CMAC 신경망은 Albus가 소개한 이래[3], 단순성

과 국지적 일반화(local generalization) 특성으로 인하여 관심을 받아왔으나, 기억장치의 크기가 차원에 따라 지수함수적으로 증가하는 문제로 인하여 그 활용이 극히 제한되고 있다. 이 문제는 처음부터 제기되어 있었으며 해결책으로 Hashing 기법, 계층적방법, 입력 양자화 최적화 방법등이 제안되었으나 나름대로 문제를 안고 있다.[3][4][5] 한편 SOM은 입력공간의 조직 성능이 가장 우수하는데 이에 대하여 Principe[6]는 4 개항을 제시하고 있다. 요약하면 다음과 같다. 1) SOM은 입력공간을 효과적으로 대표한다. 2) SOM은 입력공간의 토폴로지를 잘 반영해서 '이웃(neighbor)' 관계가 출력공간에 자연스럽게 드러나게 해준다. 3) SOM은 통계적 성질을 반영한다. 즉, 입력밀도가 높은 영역은 상대적으로 넓은 출력영역을 차지한다. 4) 연속적인 입력공간이 낮은 차원의 이산공간으로 사상되므로 차원의 축소가 일어난다. Principe는 SOM을 보조수단으로 하고 별도의 선형망을 이용하여 시계열발생 모형을 형성하였는데, 본 연구에서는 SOM-CMAC 결합구조로 직접 시계열예측을 시도하였다.

II에서는 계통의 구성에 관하여 요약하고, III에서는 교육방법에 대하여, IV에서는 모의실험결과에 대하여 그리고 V에서는 결론을 제시한다.

1) SOFM (Self-Organization Feature Map)으로 더 잘 알려져 있으나 여기서는 'Feature'의 F를 빼고, SOM으로 표시하기로 한다.

## II. SOCMAC의 구성

1. CMAC의 기본이론: CMAC은 연상 기억장치의 한 형태로 볼 수 있으며, 동작 알고리즘은 다음 두 개의 사상으로 표현된다[3].

$$F: X \Rightarrow A \quad (1)$$

$$G: A \Rightarrow Y \quad (2)$$

여기서  $X$  와  $Y$ 는 각각 입, 출력 공간이며,  $A$ 는 가상적 기억장치의 주소들의 집합으로 연관(association)공간이라 한다. 동작의 실체는 기본적으로 다음과 같다. 개 개의 입력( $x \in X$ )은 미리 지정된 방식으로  $A$ 에 사상되어 한 조의 기억위치를 가리키는 연관벡터  $\alpha = F(x)$  를 발생한다. 둘째 함수는 연관 벡터를 실제 기억공간에 투영하여 출력을 만든다. 이 과정을 식으로 표현하면

$$y = G(\alpha) = W^T \alpha \quad (3)$$

이 되는데,  $y \in Y$  는 출력,  $W$  는 실제 기억장치의 내용이며 보통 가중벡터(weight vector)라고 부른다.  $\alpha$ 가 지시하는 한 조의 기억위치들을 입력  $x$ 에 대한 수용영역( receptive field)이라 하며 유사한 입력간에는 수용영역의 일부를 공유하게 되어있다. 그림 1은 이 관계를 나타낸 것이다. 입력  $x_i$ 가 주어지면 수용영역  $a_i$ 가 활성화되고 이에 해당하는 기억내용이 합성되어 출력  $y$ 가 발생되는 것이다.  $x_i$ 와  $x_j$ 의 수용영역  $a_i$ 와  $a_j$ 는 일부가 중첩되어있음을 볼 수 있다. 그 결과 기억공간  $W$ 의 일부를 공유하게 되며 입력이 유사하면 출력도 서로 유사성이 발생하며 국지적 일반화를 이루는 바탕이 된다.

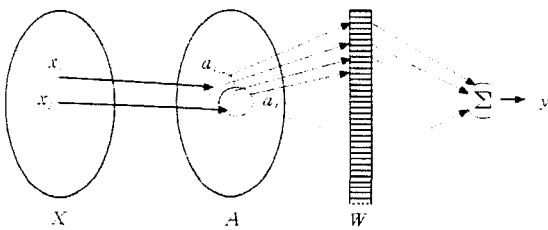


그림 1. CMAC의 동작  
Fig. 1. Operation of CMAC.

2. SOCMAC: 식 (1)은 입력을 이산화(discretize)하여 한 조의 기억주소를 발생하는 일련의 과정이다. 통상의 CMAC에서  $F(x)$ 는 일종의 고정된 A/D변환의 형태를 띄는데 반하여 SOCMAC에서는 입력 샘플들의 분포에 따라 최선의 이산화를 모색함으로써 기억공간을 절약할 수 있게 된다. SOCMAC의 구조는 위 (1), (2)식에 해당

하는 두 개의 층으로 구성되며 그림 2와 같다.

첫째 층, 또는 '입력층'은 Kohonen의 SOM의 확장된 형태를 사용하며 이 층의 출력이 연관공간( $A$ )를 형성한다. 입력층 전반부의 SOM은 구조적으로는 경쟁망(competitive network)이며, 입력에 가장 잘 반응하는 뉴런, 즉, 승자(winner)를 선택하여 활성화한다. 후반부는 수용영역 형성을 위한것이며, 승자,  $a$ 를 중심으로 그 이웃(neighbor)들에 해당하는 몇 개의 뉴런들을 활성화하여 연관벡터,

$$\alpha = N_d(a) \quad (4)$$

를 발생한다. 여기서  $N_d(a)$ 는 수용영역을 결정하는 이웃 선택함수로서 SOM 학습과정에서 사용하는 이웃결정 방법과 동일한 것이며,  $a$  와 그 주변의 거리  $d$  범위안에 있는 뉴런들의 출력을 '1'로 설정한다.

둘째 층, 또는 '출력층'은 식 (3)의 연산을 수행하는 선형망으로 표현할 수 있으므로 그림에는 LIN으로 나타내었다.  $\alpha$ 는 0과 1만으로 구성되어 있으므로 식 (3)은 실제로 단순한 합산을 의미한다.

3. SOCMAC의 학습: 통상의 CMAC에서는 식 (1)의 과정은 미리 정하여지므로 학습은 둘째층에만 적용되며 식 (3)의 결과가 만족스럽게 되도록 가중벡터  $W$ 를 수정하는 과정만으로 구성된다. 그러나 SOCMAC에서는 식 (1)의 사상도 학습에 의하므로 두 단계로 이루어진다.

학습의 1 단계는 입력층에 대한 것이다. 이 부분의 학습은 Kohonen의 SOM 알고리즘에 따라 입력공간을 자율 조직한다.

둘째층의 학습은 CMAC 학습의 전형적 형태를 따른다. 즉, 입력  $x$ 에 의한 학습 결과는

$$dW = r[y_d(x) - y(x)]\alpha(x) \quad (5)$$

이며, 여기서,

$dW$ : 가중벡터 ( $W$ )의 수정 벡터

$y_d$ : 희망 출력

$r$ : 학습율

이다.

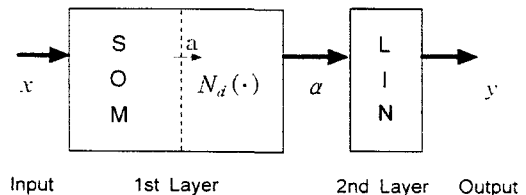


그림 2. SOCMAC의 구조  
Fig. 2. Structure of SOCMAC

### III. 가변 수용영역

CMAC에 있어서 학습은 사용된 입력 샘플 주변에만 이루어지므로 일반화는 국지적이며, 이들이 합성되어 전체적(global) 모형을 발생하는 것일 뿐, 전체적 규모의 일반화는 이루어지지 않는다. 그런데 만약 기억공간에 학습을 경험하지 못한 지역이 있다면 이 부분의 반응은 전혀 예측할 수가 없게 된다.[7] 즉, 기억의 규모가 작아도 문제이지만 너무 클 때도 문제가 생길 수 있으며 사용자는 이에 대한 대비를 하여야 한다. SOCMAC은 이점에 있어서 통상의 CMAC보다 상당히 유리한 위치에 있다. I 장에서 언급한 성질 3)에 표현된 바와 같이 입력공간을 효과적으로 재배치하기 때문이다. 그래도 역시 이 문제는 발생하며 이의 완화를 위하여 본 연구에서는 '가변 수용영역' 시스템을 도입하였다. 이것은 2 단계 학습, 즉, 식 (5)를 실행함에 있어서 연관벡터  $a$ 를 가변적으로 운용하는 것이다. 이 방법은 SOM 교육 때 이웃의 범위를 전체에서 점차 축소하여 최종적 값에 이르게하는 과정과 일치하는 것이며 식 (4)의  $d$  값을 처음에는 큰 값으로 하여 전체적 모형을 갖추게 하고 점차 값을 줄여나가서 최종의 국지적 모형을 발생하도록 한다.

### IV. 실험

실험 대상으로는 Mackey-Glass의 비선형 시계열의 모형을 택하여 여섯 샘플 뒤의 값을 예측하도록 하였다. 이 모형은 예측시스템의 벤치마크로 많이 이용되며 다음의 방정식으로 표현된다.

$$\frac{dx(t)}{dt} = -bx(t) + a \frac{x(t-\tau)}{1+x(t-\tau)^{10}} \quad (6)$$

파라미터들은 [8]에서와 마찬가지로  $a=0.2$ ,  $b=0.1$  및  $\tau=17$ 로 하였으며, 4-점 Runge-Kutta 방식으로 적분하여 시계열을 발생하였다.  $x(n)$ ,  $x(n-6)$ ,  $x(n-12)$ 와  $x(n-18)$ 의 네 개를 입력으로 하여  $x(n+6)$ 를 추정하도록 하였다. 입력은 4차원 벡터이지만 SOM 출력은 정방형의 2차원으로 구성하여 (10×10), (20×20)등의 규격으로 하였으며 이 형태는 그대로 기억공간의 구조가 되므로 식 (3)과 (5)의 가중벡터, W는 실험에서는 2차원 행렬로 구현되었다. 수용영역은 가중행렬 평면상에서 거리 1 이내의 이웃들을 범위로 잡았다. SOM의 구성과 훈련에는 'MATLAB Neural Network Toolbox'를 사용하였다.

1. 실험결과: 그림 3에는 추정결과를 실제 값과 함께 나타내었다. 표 1은 [8]-[10]에 보고된 자료와 비교한 것이다. SOCMAC의 성능은 중간정도로 TDNN이나 BP 수준에 있다. (표 1의 TDNN과 TDRNN자료는 [9]에 보고된 것 중 최선의 값을 택하였으며 이 값은 학습과 시험데이

터 모두에 대한 평균이므로 다른 자료들과 조건을 같이 하면 더 나쁘게 될 것이다) 여기서 NDEI(nondimensional error index)는 오차 지수의 하나로 다음 식에 의하였으며 [9]의 NMSE의 평방근에 해당한다.

$$NDEI = \frac{\text{rms 오차}}{\text{목표시계열의 표준편차}} \quad (7)$$

그림 4는 규격변화에 따른 NDEI의 변화를 나타낸 것이다. 기억공간이 커질수록 오차가 작아지는 것을 볼 수 있으나 (30×30)구조에서는 오히려 증가하는 것으로 나타났

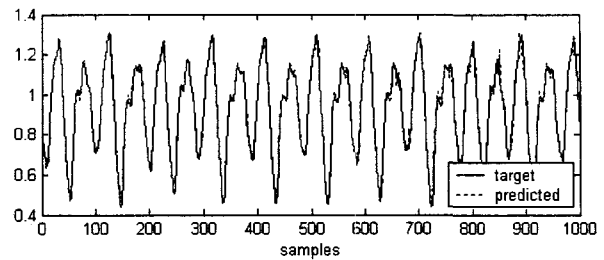


그림 3. Mackey-Glass 시계열 예측

(구조: 25×25, 앞쪽 500 샘플은 학습에 사용됨)

Fig. 3. Prediction of Mackey-Glass time series.

(Structure: 25×25. Front half of target samples were used for training)

표 I. NDEI(Nondimensional Error Index)의 비교  
TABLE I. Comparison of NDEI among various predictors.

방 식	NDEI	방 식	NDEI
ANFIS[8]	0.007	Back-Propagation[8]	0.06
AR Model[8]	0.19	TDNN[9]	0.075*
Cascaded-Correlation [8]	0.19	TDRNN[9]	0.0613*
6-th order Polynomial[8]	0.04	SOCMAC1(25×25)	0.0866
Linear Prediction[8]	0.55	SOCMAC2(25×25)	0.0947
Schiavo's[10]	0.004	SOCMAC1**	0.0576

\* 학습샘플을 시험에 포함시킨 결과임.

\*\* SOCMAC1을 학습샘플만으로 시험한 결과.

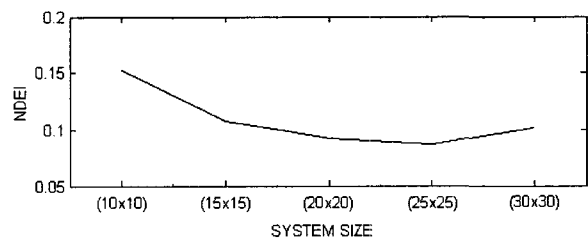


그림 4. 시스템 크기 대 NDEI 값

Fig. 4. NDEI values vs system size.

다. 이것은 III장에서 언급한 바와 같이 교육샘플 수에 비하여 기억공간이 매우 크므로 일반화가 충분히 이루어지지 못한 결과로 보이며, 가변수용영역이 (25×25)구조에 최적화 되어있음을 의미한다.

2. SOM의 조직: 그림 5는 입력공간이 학습에 의하여 조직된 결과의 한 예이다. 그림에는 15×15 개의 네모가 있으며 이들 각각은 입력공간을 벡터양자화한 세포들이다. 그림에서 아래로 내려간 부분은 학습용 입력 샘플들의 양자화에 사용되지 않은 영역을 나타낸 것이다. 이 부분의 면적은 (10×10) 구조에서는 16/100, (15×15) 구조에서는 58/225, (20×20) 구조에서는 152/400, 그리고 (25×25) 구조에서는 무려 328/625로 나타났다. 이런 영역의 발생은 기억공간의 효과적 활용이라는 관점에서는 손실로 보이지만 학습용 샘플들이 대표하지 못하는 공간을 수용하기 위해서는 필요한 결과이다.

3. 가변수용영역의 효과: 표 I의 SOCMAC1은 가변수용영역에 의한 것이고 SOCMAC2는 고정수용영역에 의한 것이다. 약간의 개선이 있음을 볼 수 있는데 이 효과는 시험 샘플영역 전체의 평균이므로 오차가 많은 영역에서는 상당한 차이를 의미한다. 가변수용영역의 적용은 다음과 같은 과정에 의하였다. 1) 먼저 수용영역의 크기를 결정하는 이웃 반경( $d$  값)의 초기치와 최종치를 설정한다. 2) 1차 수렴과정:  $d$ 를 초기치로부터 감소시켜 최종치에 이르기까지 식 (5)의 학습을 반복한다. 3) 2차 수렴과정:  $d$ 를 최종치로 고정하고 학습을 정한 회수만큼 반복한다. 실험에서는  $d$ 의 초기치는  $\sqrt{(\text{가중행렬 행의 크기})/2+2}$ 로 하였고 최종치는 1로 하였다.

그림 6은 학습과정에서 오차의 수렴 형태를 보인 것이다. 고정수용영역 방법(점선)에서는 단조감소의 빠른 수렴곡선을 나타내는데 반하여, 가변수용영역 방법(실선)에서는 수용영역의 크기가 감소될 때 마다 새로운 조건하에서 수렴이 다시 시작되므로 복잡한 과정을 거치는 것이 잘 들어나 있다. 2000회 부근에서 수용영역 최종치에 도달한 것을 알 수 있다. 수용영역을 고정한 편이 오히려 최종 오차가 작은 것으로 나타난 것은 흥미로운 일이다.(0.0576 대 0.0522) 이것은 학습용 샘플에 대하여서는 수용영역을 고정하는 것이 가변형 보다 오차가 작다는 것을 의미한다. 그림 7은 수용영역을 가변으로 했을 때의 가중행렬,  $[wv(i,j)]$ 와 고정했을 때의 가중행렬,  $[wf(i,j)]$ 와의 차를 나타낸 것이다. 같은 학습 데이터를 사용하였는데도 학습된 비중값에는 상당한 차이가 있음을 알 수 있다.

### V. 결 론

본 연구에서는 SOCMAC에 의한 비선형 시계열 예측을 시도하였다. 결과는 TDNN수준이며 정교하지는 못했지만 충분히 실용적인 범위에 들어오는 것으로 판단된다.

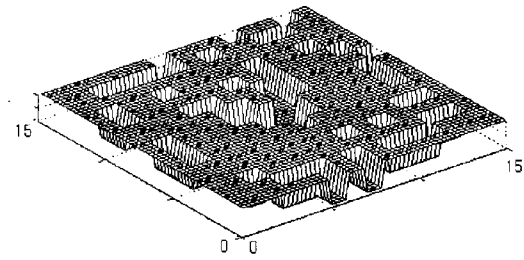


그림 5. 입력공간 조직의 예

(구조: 15×15, 내려간 부분은 SOM 상에서 학습샘플들에 의하여 선택되지 않은 위치들)

Fig. 5. An example of input-space organization. (Structure: 15×15. Sunken area represents the memory locations not addressed by the training samples)

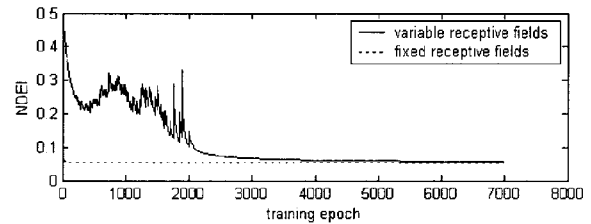


그림 6. 학습오차의 수렴과정

Fig. 6. Learning curves.

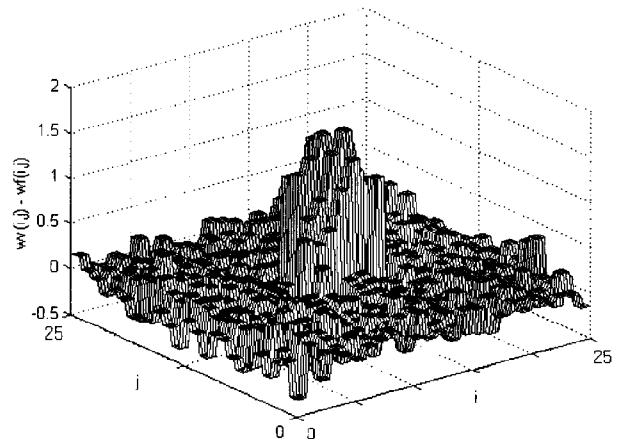


그림 7. 수용영역을 가변으로 할 때와 고정으로 할 때 비중행렬의 차 (구조: 25×25)

Fig. 7. Difference between weight matrices of variable and fixed receptive-field systems.

CMAC의 성질상 입력이 다차원 연속변수인 경우에는 실용적이 되지 못하므로 그 활용이 초기의 기대에 못미쳐왔으며 시계열 예측은 상상할 수도 없는 것이었다. 그러나 본 연구에서 보인바와 같이 SOM을 결합함에 의하여

CMAC의 활용이 상당히 자유롭게 되었음이 입증되었다. BP등 다차원 비선형 방정식의 배경을 가진 대표적 신경망들은 정교하게 조정됐을 경우 강력한 예측능력을 발휘할 수 있지만 일반적으로 국지적 극소점(local minima)의 문제, 과도정의(over-define)된 부분의 불안정문제등 실제로 활용하는데는 사실상 해결하기 어려운 상황에 봉착되기 쉽다. 이에 비하여 CMAC은 일종의 조건표(lookup table)이므로 거칠기는 하지만 쉽게 수렴하고 안정한 모형을 얻을 수 있다. 적정규모의 예측, 가변수용영역의 최적화를 포함한 학습 기법과 동작속도의 개선 등 아직 추가적 연구가 요구되지만 그동안 걸림돌이었던 입력차원 문제가 SOM에 의하여 어느 정도 완화됨에 따라 더 많은 활용을 기대할 수 있게 되었다.

접수일자 : 2002. 10. 1      수정완료 : 2002. 10. 19

### 참고문헌

- [1] Chow, M. and Menzzi, A, " A Self-Organized CMAC Controller", Proceedings of the IEEE International Conference on Industrial Technology, 1994, pp.68 -72, 1994
- [2] Chow, M. and Menzzi, A., "On the training of a multi-resolution CMAC neural network", IECON 97. 23rd International Conference on Industrial Electronics, Control and Instrumentation, Vol. 3, pp.1130 -1135, 1997
- [3] Albus, J. S. "A new approach to manipulator control: the Cerebella Model Articulation Controller(CMAC)", Transactions of ASME Series G, Journal of Dynamic Syatems, Measurement and Control, vol. 97, pp.220-227, 1975
- [4] Benavent, X., Domingo, J., Vegara, F. and Pelechano, J., "Two suggestions for efficient implementation of CMAC's", Proceedings of the 2000 IEEE International Symposium on Intelligent Control, 2000, pp.309 -314
- [5] Lee, H., Chen, C. and Lu, Y., "A self-organizing HCMAC neural network classifier", Proceedings. IJCNN '01. International Joint Conference on Neural Networks, 2001, vol. 3, pp.1960 -1965
- [6] Principe, J. C., Wang, L. and Motter, M. A., "Local dynamic modeling with self-organizing maps and applications to nonlinear system identification and control", Proceedings of the IEEE , Vol. 86, No. 11, pp.2240 -2258, Nov. 1998
- [7] Lee, T. and Lee, K. Y., "Application of CMAC for Short-Term Load Forecasting", Proceedings of International Conference on Intelligent System Applcation to Power Systems, Seoul, pp.372-377, 1997
- [8] Jang, J.-S.R. and Sun, C.-T., "Predicting chaotic time series with fuzzy if-then rules ", ., Second IEEE International Conference on Fuzzy Systems, 1993, vol. 2, pp.1079 -1084
- [9] Kim, S, S., et al, "Time-Delayed Recurrent Neural Network for Temporal Correlations and Prediction, ans its Application to Phone Recognition", Journal of Electrical Eng. and Information Science, vol. 4, No, 1, pp.97-104, Feb. 1999
- [10] Lo Schiavo, A. and Luciano A. M., "Powerful and Flexible Fuzzy Algorithm for Nonlinear Dynamic System Identification", IEEE Transactions on Fuzzy Systems, vol. 9, no. 6, pp.828-835, Dec. 2001



이 태 호(Tai-Ho Lee)

正會員

1966 한양대학교

1969 서울대학교(공학석사)

1975 연세대학교(공학박사)

1969-1972 한국과학기술연구소

1973-현재 울산대학교 전기전자정보시스템공학부 교수

관심분야 : 음성신호처리, 인공신경망, 통신시스템