

동적 가우시안 함수를 이용한 Kohonen 네트워크 수렴속도 개선

길민욱* 이극**

Improved Rate of Convergence in Kohonen Network using Dynamic Gaussian Function

Min-wook Kil* Geuk Lee**

요약

자기조직화 지도(self-organizing feature map)는 학습시 수렴하기 위하여 많은 입력패턴을 필요로 하는 단점이 있다. 본 논문에서는 자기조직화 지도 학습시 학습률이 일정한 이웃 상호작용 집합을 동적 가우시안 함수로 변환하여 수렴속도와 수렴도를 개선할 수 있는 방법을 제안한다. 제안한 방법은 이웃 상호작용 함수로 사용된 가우시안 함수의 편차와 폭을 학습 회수에 따라 감소하는 동적 성질과 승자 뉴런으로부터의 위상학적 위치에 따라 각기 다른 학습률을 갖도록 하였다. 따라서 본 논문에서는 자기조직화 지도의 수렴속도와 수렴도를 향상시켰다.

Abstract

The self-organizing feature map of Kohonen has disadvantage that needs too much input patterns in order to converge into the equilibrium state when it trains. In this paper, we proposed the method of improving the convergence speed and rate of self-organizing feature map converting the interaction set into Dynamic Gaussian function.

The proposed method provides us with dynamic properties that the deviation and width of Gaussian function used as an interaction function are narrowed in proportion to learning times and learning rates that varies according to topological position from the winner neuron. In this paper, we proposed the method of improving the convergence rate and the degree of self-organizing feature map.

* 문경대학 인터넷 정보계열 조교수

** 한남대학교 컴퓨터전자통신공학부 컴퓨터공학전공 교수

I. 서론

신경망의 자율학습(unsupervised learning) 모델들 중 Kohonen에 의해 제안된 자기조직화지도는 일종의 벡터양자화(vector quantization)를 위한 신경망으로 볼 수 있다[1,2,3]. 이 신경망의 장점은 다수의 입력벡터가 가지고 있는 특징들을 1차원 또는 2차원의 특징지도(feature map)로 형성하며 그에 따른 입력벡터들의 분류 문제를 자동적으로 해결한다. 또한 시간 지연(time delay)효과를 특징지도 자체에서 추적이 가능하므로 시간보정(time alignment)이 필요한 음성인식 같은 분야에서 응용이 가능하다. 실제적으로 이 모델은 음성인식의 하부모델로 구현되는 경우가 많으며 그러한 분야에서 아주 적합한 것으로 알려지고 있다[4]. 그러나 자기 조직화 지도의 학습 방법은 이웃 집합(neighborhood set)안에서 승자 뉴런과 다른 뉴런과의 거리에 관계없이 학습률이 일정하여 학습이 느리며 평형상태(equilibrium state)로 수렴하기 위한 다수의 입력벡터가 필요하다는 단점을 가지고 있다.

관련 분야의 최근 연구는 주로 학습 단계의 성능 향상에 관심이 모아지고 있다. Lenne, Paolo는 입력벡터를 그룹화한 후에 가중치를 조절하는 방법을 제시하였고, Behis, George는 기하학적인 해쉬 함수(geometric hash function)를 정의하여 학습하는 방법을 제안하였다[5,6]. 또한 B.Bavarian은 이웃 상호작용 함수(neighborhood interaction function)로 정적인 가우시안 함수(Gaussian function)를 정의하여 위상학적 위치에 따라 각기 다른 학습률(learning rate)을 갖도록 하였으나 기존의 자기조직화지도 보다는 높은 수렴도를 얻을 수 있었지만 자기조직화가 정밀하게 이루어지지 않는 경우도 발생을 한다[7,8,9].

본 논문에서는 동적 가우시안 함수를 제안하여 신경망의 수렴속도를 향상시키고, 소수의 입력벡터로도 훈련이 가능하며 보다 세밀한 자기조직화를 할 수 있는 방법을 제안한다

II. 자기조직화 지도

자기조직화지도는 여타 신경망과는 달리 입력층과 상호 경쟁하는 출력층인 2개의 층으로 구성되며 층 사이가 완전연결(fully connected) 되어 있어 입력패턴을 모든 출력층 뉴런이 소유하는 단순한 형태를 취하고 있다

출력층의 각 뉴런들은 연결 가중치 벡터와 입력 벡터의 차를 계산하여 승자 뉴런을 결정한다. 이 승자뉴런을 중심으로 인접한 이웃집합안에 있는 뉴런들만이 제시된 입력 벡터에 대하여 학습이 허용되는데 이웃의 크기는 시간이 경과함에 따라 서서히 축소된다.

승자 뉴런을 결정하는 법칙으로는 출력층의 뉴런과 입력층의 차원이 같은 점을 이용 유클리디안 거리로 측정하여 가장 거리가 짧은 뉴런을 승자뉴런으로 결정하며, 학습 법칙으로는 식 (1)과 같이 입력과 가중치간의 거리에 학습상수를 곱해줌으로써 반경안에 속하는 뉴런들의 가중치가 입력 패턴에 점진적으로 가까이 가는 형태를 취하고 있다.

$$M_i(t_{k+1}) = M_i(t_k) + \alpha(t_k)[X(t_k) - M_i(t_k)],$$

$$\text{for } i \in N_f(t_k)$$

$$M_i(t_{k+1}) = M_i(t_k), \text{ for } i \notin N_f(t_k) \quad (1)$$

여기에서 $X(t_k)$ 는 입력, $M_i(t_k)$ 는 출력층 i 뉴런의 가중치, $N_f(t_k)$ 는 이웃집합, $\alpha(t_k)$ 는 학습상수를 나타낸다.

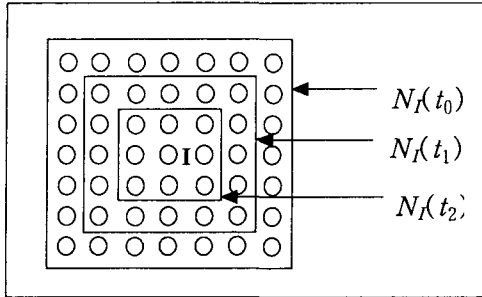


그림 1. 이웃 집합의 크기 조정

학습상수는 학습회수의 증가에 따라 단조감소함으로써 출력층 뉴런 가중치들의 입력을 향한 유동을 완화시켜 신경망이 평형상태에 도달하도록 하는 역할을 하며 그림1에서 보는바와 같이 이웃 반경 역시 시간이 지남에 따라 단조감소 된다.

2.1 Kohonen의 이웃 상호작용 집합

Kohonen의 자기조직화 지도는 승자뉴런이 결정되면 이 승자뉴런을 중심으로 이웃집합안에 있는 인접한 뉴런들만이 제시된 입력 벡터에 대하여 학습이 허용되는데 그림2에서 보는 것처럼 승자뉴런 I의 이웃 집합안에서는 승자 뉴런과의 거리에 관계없이 일정한 학습률을 갖고 학습하므로 평형상태로 수렴하기위한 다수의 입력 벡터가 필요하며 학습이 느리다[7].

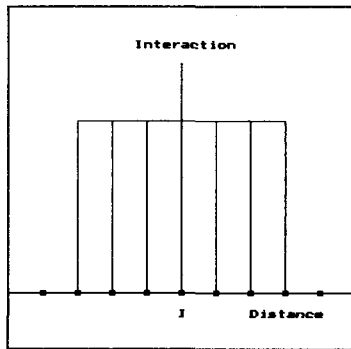


그림 2. 이웃 상호작용 집합

Kohonen의 자기조직화지도는 빠른 학습시간과 단순한 구조를 가지고 있으며 모호한 특징들 사이의 관계를 자기조직화 시켜주는 장점을 가지고 있다. 그러나 신경망의 크기가 충분히 크지 못할 경우 좋은 성능을 발휘할 수 없으며 다수의 입력 벡터를 필요로 하고 가중치들의 값을

임의의 값으로 초기화 시켜주어야 하는 단점을 가지고 있다[8,9].

2.2 B.Bavarian의 이웃 상호작용 함수

B.Bavarian에 의해 제안된 방법은 기존 신경망과는 달리 이웃 상호작용 함수를 정의하여 이웃 집합 안에서 승자 뉴런(winning neuron)과 다른 뉴런과의 거리에 따라 각기 다른 학습률을 갖게 하고 있다. 이웃 상호작용 함수 AI는 그림3과 같은 가우시안 함수를 사용하고 그 수식은 (2)와 같다.

$$A_f(i, t_k) = c + d \cdot e^{-\frac{H(i-I)^2}{2\sigma^2}}, \text{ for } i \in N_f(t_k)$$

$$\sigma = |N_f(t_k)| \quad (2)$$

여기서 c, d, H 는 상수. I는 승자 뉴런을 나타내고, σ 는 시간에 따라 감소하며 시간 t_k 에서의 이웃의 크기를 나타낸다.

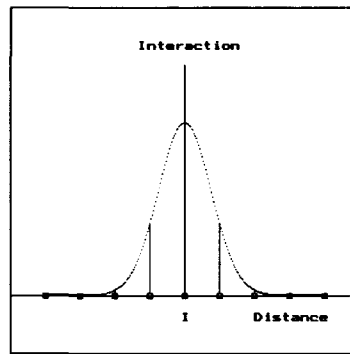


그림 3. 이웃 상호작용 함수

학습법칙은 자기조직화지도의 학습법칙에 상호 작용 함수 $A_f(i, t_k)$ 가 첨가한 식 (3)과 같다.

$$M_i(t_{k+1}) = M_i(t_k) + \alpha(t_k)A_f(i, t_k)[X(t_k) - M_i(t_k)], \text{ for } i \in N_f(t_k)$$

$$M_i(t_{k+1}) = M_i(t_k), \text{ for } i \notin N_f(t_k) \quad (3)$$

이 신경망은 자기조직화지도 보다 빠른 수렴속도를 얻을 수 있으며 그에 따른 입력 벡터의 양을 줄일 수 있는 장점이 있다. 그러나 신경망의 크기가 커지면 코드북(codebook)의 일부가 중첩되어 자기조직화가 완전히 이루어지지 않는 경우도 발생을 한다.

따라서 이웃 상호작용 함수를 분석하고 동적 이웃 상호작용 함수를 이용하는 방법을 제안한다.

III. 동적 이웃 상호작용 함수

3.1 이웃 상호작용 함수 분석

상호작용 함수로 사용된 가우시안 함수의 성질을 분석하는 것은 신경망의 성능을 향상시키기 위한 절차로서 변수들의 변화와 그로 인한 신경망의 성능을 알아볼 수 있다. 그림4는 파라미터 H 에 따른 그래프를 도시한 것으로 H 가 증가함에 따라 그래프의 폭이 줄어드는 것을 알 수 있다.

그리고 그림5는 변수 H 의 변화로 인한 신경망의 성능을 식 (4)를 이용하여 시간영역(time domain)에 대해서 평가한 결과이다. 단, 파라미터 H 가 단조증가(monotonic increase)시 오차(error)자체가 단조감소(monotonic decrease)하지는 않는다.

$$error(t_k) = \sum_i [(M_{i_1}(t_k) - \mu_{i_1})^2 + (M_{i_2}(t_k) - \mu_{i_2})^2] \quad (4)$$

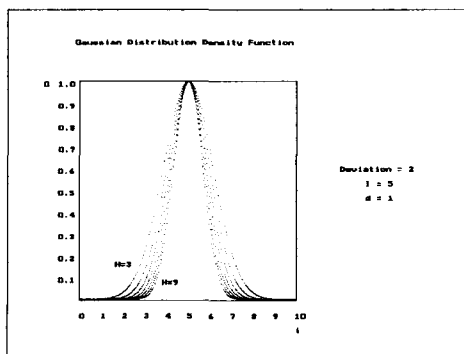


그림 4. H의 값에 따른 가우시안함수의 변화

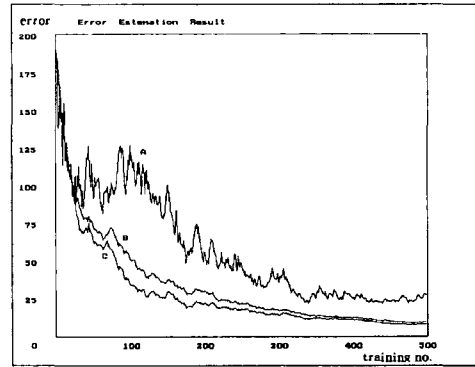


그림 5. H의 변화에 따른 신경망의 오차율 (A : H=5, B : H=8, C : H=7)

3.2 동적 이웃 상호작용 함수

본 논문에서는 그림 6처럼 자기조직화 지도의 이웃 상호작용 함수를 학습 회수에 따라 폭이 감소하는 동적 성질과 승자 뉴런으로부터의 위상학적 위치에 따라 각기 다른 학습률을 갖도록 하여 수렴속도를 높이고 자기 조직화 지도의 수렴도를 개선 시킬 수 있는 새로운 방법을 제안 한다.

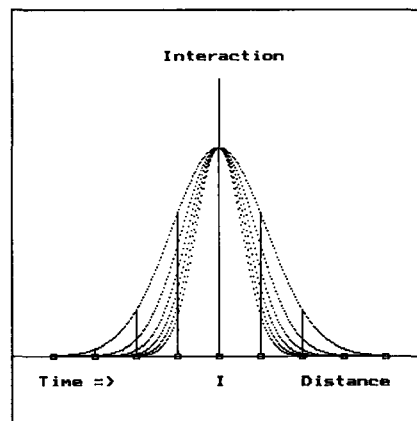


그림 6. 동적 가우시안 함수

B.Bavarian 방법에서 사용된 가우시안 함수 식 (2)를 앞장에서 분석한 결과를 기반으로 하여 파라미터 H 의 값을 학습 회수가 증가함에 따라 변화시켜 승자뉴런의 위상학적 위치에 따라 각기 다른 학습률을 갖도록 하고, 또한 초기의 시간영역에서 승자뉴런에 집중적으로 몰려있는 가중치(weight) 개선효과를 위하여 편차 σ 를 이웃 크기 보다 넓게 선택하여 시간이 지남에 따라 모든 이웃 집합

안의 뉴런에 완만하게 적용되도록 하였다. 제안한 방법의 이웃 상호작용 함수는 식 (5)와 같이 수식화되며 학습법칙은 식 (6)과 같다.

$$A_f(i, t_k) = c + d \cdot e^{-\frac{(H + \frac{t_k}{T})(i-1)^2}{2\sigma^2}},$$

for $i \in N_f(t_k)$

$$\sigma = |N_f(t_k)| + \beta \quad (5)$$

여기서 c,d,f,β,H는 상수이다.

$$M_f(t_{k+1}) = M_f(t_k) + \alpha(t_k)A_f(i, t_k)[X(t_k) - M_f(t_k)], \text{ for } i \in N_f(t_k)$$

$$M_f(t_{k+1}) = M_f(t_k), \text{ for } i \notin N_f(t_k) \quad (6)$$

그리고 식 (6)은 시간 t_k 에 대해 이웃 집합안에서 가중치 갱신의 정도를 높이면서 모든 뉴런에게 완만하게 적용시키는 역할을 하며 자기조직화 지도에서 대표 뉴런들을 균일하게 분포하여 세밀한 자기조직화를 함으로 코드북을 효율적으로 구성할 수 있다. 또한 위의 방법에 의한 신경망은 다음을 만족한다

정리 1 : 매 훈련단계에서 이웃 집합을 위상적 순서(topological order)가 되게 한다.

증명 : 주어진 승자 뉴런 I의 이웃 반경 안에 있는 임의의 뉴런 $i_1, i_2 \in N_f(t_k)$ 라 하자. 만약 i_1 이 i_2 보다 I에 근접해 위치한다면 $A_f(i_1, t_k) > A_f(i_2, t_k)$ 이고 가중치간의 유클리디안 거리 r_i 의 변화율은 $r_{i_1} > r_{i_2}$ 이다. 따라서 성립한다.

정리 2 : 만약 신경망이 위상적 순서를 가진다면 이 순서를 보존하게 된다

증명 : 주어진 승자 뉴런 I의 이웃 반경 안에 있는 임의의 뉴런 $i_1, i_2 \in N_f(t_k)$ 라 하자. 만약 i_1 이 i_2 보다 I에 근접해 위치한다면 $A_f(i_1, t_k) > A_f(i_2, t_k)$ 이고

입력 벡터로부터의 거리는 $d_{i_1} > d_{i_2}$ 이 된다. 따라서 성립한다.

이러한 증명을 통하여 가우시안 함수를 동적으로 변화하도록 제안된 알고리즘은 위상적 순서를 형성하고 보존하며 수렴함을 알 수 있다.

IV. 실험결과

Kohonen 방법, B.Bavarian 방법, 제안한 방법의 위상적 순서화의 평가와 자기조직화 정도를 분석하기 위하여 모두 32 X 32 신경망을 만들어 2차원 위상에 적용시켜 실험 하였다. Kohonen 방법은 승자 뉴런과의 거리에 관계없이 일정한 학습률을 가지고 학습하며, 거리에 따라 각기 다른 학습률을 갖도록 B.Bavarian 방법과 제안한 방법은 이웃 상호작용 함수로 가우시안 함수를 사용한다. B.Bavarian 방법에서는 정적인 가우시안 함수를 사용하며 제안한 방법에서는 동적인 가우시안 함수를 사용한다. 2차원의 입력벡터와 사각형 위상을 가진 2차원의 신경망을 동일하게 채택하였으며 입력벡터는 사각형 위상에서 평등분포(uniform distribution)되게 선택하였다. 각 경우의 초기 이웃 집합의 크기는 15로 하였고, 감소되는 반경 역시 동일하게 단조감소 되도록 하였다.

$$\alpha(t_k) = Z(1 - \frac{t_k}{T_1}) \quad (7)$$

여기서 Z는 초기 학습상수 이고 T1은 최대 학습 회수이다. 또한 학습상수 $\alpha(t_k)$ 도 시간의 경과에 따라 단조감소하게 했으며, 본 실험에서 사용된 수식은 다음과 같다. 본 실험에서는 사용된 수식은 식 (5)에서 c=0, d=1, β=1로 설정한 식 (8)과 같다. 여기에서 β값을 1로 한 이유는 편차 값을 이웃 크기보다 넓게 설정하여 승자 뉴런에 집중적으로 몰려 있는 가중치 개선효과를 위하여 실험결과 가장 좋은 값 1로 설정하였다.

$$A_f(i, t_k) = e^{-\frac{(H + \frac{t_k}{T})(i-D)^2}{2\sigma^2}},$$

for $i \in N_f(t_k)$

$$\sigma = |N_f(t_k)| + 1 \quad (8)$$

초기 H 의 값으로는 분석 과정에서 가장 좋은 성능을 나타냈던 값 7을 기준으로 하였으며 초기 학습상수 Z 는 0.3 으로 하였다. (그림 7)은 위의 조건을 토대로 식 (4)에 의한 평가 기법을 적용한 결과로서 시간 영역 (time domain)에 대해서 평가한 결과이다. 훈련 횟수를 독립 변수로 한 수렴도를 나타낸 것으로써 제안한 방법이 가장 적은 오차값을 가지고 수렴함을 알 수 있다.

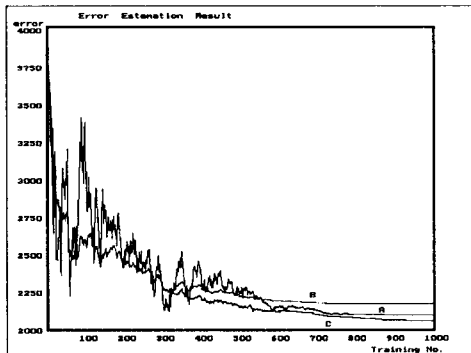


그림 7. 훈련 회수에 따른 신경망의 성능비교 (A: Kohonen, B: B.Bavarian, C: 제안된 방법)

또한 (그림 8)은 위와 동일한 조건으로 200회와 1000회를 훈련시킨 신경망의 상태를 2차원 평면상에 나타내었다. 입력 벡터와 같은 좌표상에 각 뉴런들의 가중치 벡터의 좌표를 나타낸 것으로 가장 근접한 뉴런의 가중치 벡터끼리 연결한 것이다. 다시 말해서 격자상의 선의 교점은 가중치 벡터의 좌표를 나타내고, 두 개의 교점을 연결하는 선은 그 교점의 가중치 벡터를 갖는 뉴런들이 인접하고 있는 것을 나타낸다. 학습결과 제안한 방법이 Kohonen의 방법과 Bavarian의 방법보다 세밀한 자기 조직화가 이루어 진다.

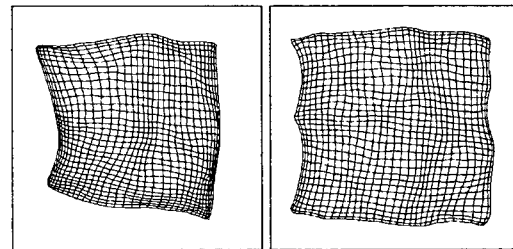
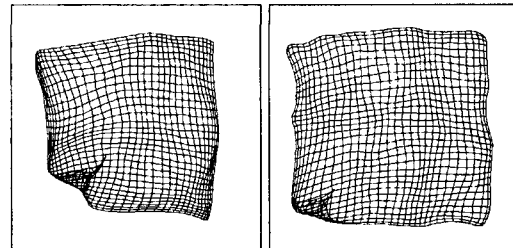
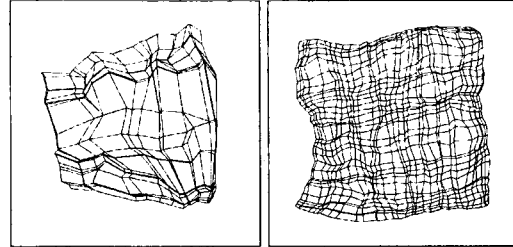


그림 8. 200회와 1000회를 훈련시킨 32x32 신경망

V. 결론

자기 조직화 지도의 학습 방법은 느린 학습률(learning rate)과 평형상태(equilibrium state)로 수렴하기 위하여 학습시 많은 입력패턴이 필요하다는 단점이 있다. 이러한 문제점을 해결하기 위해 본 논문에서는 B.Bavarian의 자기 조직화 지도의 학습시 이웃 상호작용 함수로 사용된 정적인 가우시안 함수를 학습 회수에 따라 가우시안 함수의 편차와 폭을 변화시켜 자기조직화 정도와 신경망의 수렴속도를 개선할 수 있는 동적 가우시안 함수를 제안하

였다.

실험결과 이 방법은 기존 방법들의 오차곡선 자체가 대부분 비선형(non-linear)인 경우와는 달리 거의 선형 곡선을 형성해 다른 방법보다 우수하면서도, 곡선 자체의 기울기가 작아 평형상태에 빠른 수렴 속도를 나타내는 장점을 가진다. 따라서 입력벡터가 평등한 분포를 가지는 경우에는 Kohonen의 방법보다 소수의 입력패턴을 가지고 고도 신경회로망을 학습 시킬 수 있다.

참고문헌

- [1] Teuvo Kohonen, "The Neural Phonetic Typewriter", computer, Vol.21, No. 3, pp.11-22, 1988
- [2] Teuvo Kohonen, Self-Organization and Associative Memory, Springer-Verlag, 1989.
- [3] Teuvo Kohonen, "The Self-Organizing Map", Proc. of the IEEE, Vol.78, pp.1464-1480, Sep. 1990.
- [4] Teuvo Kohonen, E.Oja, A.Visa, O.Simula, "Engineering applications of self-organizing map", Proc. of the IEEE, pp.1358-1384, 1996.
- [5] L.Paolo, T.Patrick, V. Nikolaos, "Modified self-organizing feature map algorithms for efficient digital hardware implementation", IEEE trans. on neural network, Vol.8, pp. 315-330, 1997.
- [6] B.George, Georgiopoulous, Michael, "Using self-organizing maps to learn geometric hash function for model-based recognition", IEEE trans. on neural network, Vol.9, pp. 560-570, 1998.
- [7] Z.P.Lo and B.Bavarian, "Improved Rate of Convergence in Kohonen Neural Network", Proc. of IJCNN Vol.2, pp. 201-206, July

1991.

- [8] Z.P.Lo and B.Bavarian, "Two Theorem for the Kohonen Mapping Neural Network", Proc. of IJCNN, Vol.4, pp.755-760, 1992.
- [9] Z.P.Lo Y.Yu and B.Bavarian, "Convergence Properties of Topology Preserving Neural Network", IEEE Trans. Neural Networks Vol.4, No. 2, March 1993.

저자 소개



길민욱

1989년 한남대학교 전자계산공학과(학사)
 1991년 한남대학교 전자계산공학과(석사)
 2000년 한남대학교 전자계산공학과(박사)
 1997년 ~ 현재 문경대학 인터넷 정보계열 조교수
 관심분야 : 인공지능, 음성인식, 멀티미디어, 정보보안



이극

1983년 경북대학교 전자공학과(전산모듈) 공학사
 1986년 서울대학교 컴퓨터공학과 공학석사
 1993년 서울대학교 컴퓨터공학과 공학박사
 1988년 ~ 현재 한남대학교 컴퓨터전자통신공학부 컴퓨터공학전공 교수
 관심분야 : 인공지능, 신경회로망, 멀티미디어, 보안시스템, 음성인식