

論文96-33A-4-4

신경회로망 콘볼루션 복호기의 최적 성능에 대한 확률적 근사화

(Stochastic Approximation to an Optimal Performance of the Neural Convolutional Decoders)

劉 哲 雨 *, 康 昌 彦 *, 洪 大 植 *

(Cheol Woo You, Chang Eon Kang, and Dae Sik Hong)

요 약

콘볼루션 부호의 복호방식으로써 제안된 비터비 알고리즘은 사실상 최대 확률 복호기이고 이 관점에서 최적이다. 그러나 비터비 복호기는 구속장 길이가 하나 증가함에 따라 시스템의 하드웨어적 복잡도가 지수적으로 증가하는 단점이 있다. 본 논문에서는, 비터비 복호기와 비슷한 성능을 보이면서 비용은 절감되고 실시간 처리가 항상 가능하며, 또한 성능을 향상시킬 경우 적어도 지수적인 비용 상승이 발생하지 않는 복호기를 구성하기 위해 세 계층으로 이루어진 역전파 신경회로망을 이용한다. 그리고, 제안된 신경회로망 콘볼루션 복호기가, 이상적인 경우, 확률적 관점에서 비터비 복호기와 마찬가지로 최대 확률 결정 과정을 근사적으로 수행할 수 있음을 보인다. 또한, 간단한 실험 결과를 제시하고 실험 결과가 이론적 성능에 미치지 못하는 원인을 살펴보고, 제안된 방식의 발전 가능성을 신경회로망에서 사용된 비선형 전달 함수의 특성에서 찾아본다.

Abstract

It is well known that the Viterbi algorithm proposed as a method of decoding convolutional codes is in fact Maximum Likelihood (ML) and therefore optimal. But, because hardware complexity grows exponentially with the constraint length, there will be severe constraints on the implementation of the Viterbi decoders.

In this paper, the three-layered backpropagation neural networks are proposed as an alternative in order to get sufficiently useful performance and deal successfully with the problems of the Viterbi decoders. This paper shows that the Neural convolutional decoder (NCD) can make a decision in the point of ML in decoding and describes simulation results. The cause of the difference between stochastic results and simulation results is discussed, and then the future prospect of the NCD is described on the basis of the characteristic of the transfer function.

I. 서 론

디지털 통신이 발달함에 따라 오류 정정 기술의 중요성이 점차로 커졌다. 특히, 콘볼루션 부호가 쉬운 구현과 몇 가지 매력적인 복호방식 때문에 위성 통신 시스템에서 중요한 역할을 차지하게 되었다. 비터비 알고리즘은 비메모리 채널상에서 전송된 콘볼루션 부호의

복호방식으로써 1967년에 비터비에 의해 제안되었다.

그 뒤 비터비 복호기는 사실상 최대 확률 복호기이고, 이 관점에서 최적이라는 것이 밝혀졌다^{1), 2)}. 그런데 비터비 복호기는 구속장 길이(K)가 증가함에 따라 성능이 개선되지만, 복호시에 필요한 기억 용량이 지수적으로 증가하게 된다. 이 때문에 시스템의 하드웨어 역시 복잡해져 비용이 상승하게 되고 실제 이용상에 제한이 생기게 된다¹³⁾.

신경회로망은 1950년대 중반 학문으로서의 모습을 갖추었으나, 여러 기술상의 난점과 인공 지능의 활성화에 의하여 약 30년간에 걸친 긴 암흑기에 빠져있었다.

* 正會員, 延世大學校 電子工學科

(Dept. of Electronic Eng., Yonsei University)

接受日字: 1995年4月4日, 수정완료일: 1996年2月16日

1982년 Hopfield가 발표한 한편의 논문¹⁴⁾은 이러한 신경회로망의 전성기를 불러 왔으며, 그 뒤 오늘에 이르기까지 신경회로망은 가히 폭발적인 성장을 거듭해 오고 있다. 그 가운데에서 역전파 신경회로망(Back-propagation Neural Network: BPN)은 대표적인 신경회로망으로써 일명 다층 퍼셉트론이라고 부른다. 다층 퍼셉트론의 능력에 대한 연구는 이미 여러 사람들에 의해 행해졌다. 세 계층으로 구성된 퍼셉트론은 볼록 또는 비볼록 결정 영역을 형성할 수 있으며, 임의의 결정 영역을 형성하여 유한개의 점들을 임의의 두 집합으로 나눌 수 있다¹⁵⁾. 또한 무한개의 은닉층을 가질 때, 우리가 의도하는 사상과 신경 노드들의 입출력 특성이 몇 가지 제한조건을 만족시키면 어떤 임의의 사상도 수행할 수 있음이 증명되었다¹⁶⁾. 게다가 역전파 신경회로망은 베이즈 최적 분류 함수(Bayes Optimal Discriminant Function)에 대한 점근적 근사화를 할 수 있음도 밝혀졌다¹⁷⁾.

본 논문에서는 세 계층으로 구성된 BPN을 사용하여 콘볼루션 복호기를 구성한다. 이렇게 구성된 신경회로망 콘볼루션 복호기(Neural Convolutional Decoders : 이하 NCD)의 복호 과정에 필요한 시간은, 기존 방식과는 달리 오직 입력단에 도달한 신호가 출력단까지 진행되는 동안 통과하는 게이트들의 지연 시간 뿐이므로, 복호 과정이 빠른 시간 내에 일정하게 행해진다. 그리고 다른 것들과 비교할 때 가장 큰 차이점이자 장점인 전달함수에 의하여 가지게 되는 비선형성은, 선형 시스템이 가질 수 없는 여러 가지 가능성과 능력을 제공해 준다¹⁸⁾. 제안된 NCD는 두 가지 측면에서 성능 분석이 행해진다. 이론적으로 최대 가능한 성능을 알아보기 위해 확률적 관점에서 최대 확률 복호기에 대한 근사화가 이루어지고, 실제 성능을 알아보기 위해 부가 백색 가우스 잡음 채널하에서 실험이 행해진다. 그리고 이론적 성능과 실험적 성능 사이의 차를 발생시키는 원인에 대해 분석한다. 또한, 신경회로망 전달 함수의 특성 고찰을 통하여 NCD가 잠재적으로 가지고 있는 장점에 대한 분석이 이루어진다.

본 논문 내용의 구성은 다음과 같다. 먼저 제 2 장에서는 NCD의 구조를 간단히 설명한다. 제 3 장에서는 NCD의 최대 확률 복호기에 대한 점근적 근사화가 확률적 관점에서 이루어지며, 전달 함수를 통과한 부가 백색 가우스 잡음의 확률 특성을 살펴본다. 제 4 장에서는 몇 가지 실험 결과를 나타내고 분석한다. 끝으로

제 5 장은 결론 부분이다.

II. 신경회로망 콘볼루션 복호기

1. 역전파 신경회로망과 알고리즘

본 논문에서 NCD를 구현하기 위해 사용된 신경회로망은 역전파 신경회로망이다. 전통적인 역전파 신경회로망에서는 평균제곱오차를 줄이기 위해 GD (gradient descent)방법을 사용하는데, 본 논문에서 사용한 역전파 신경회로망 역시 이 방식에 따른다. 그리고 수렴 속도 향상을 위해 이미 효과가 인정되어 널리 이용되는 모멘트 항이 사용되었다.

p번째 표본에 대하여 GD방식에 의해 최소화해야 할 오차 기준 함수는 식 (1)과 같이 표현할 수 있다.

$$\begin{aligned} E_p &= \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2 \\ &= \frac{1}{2} \sum_{k=1}^M (y_{pk} - o_{pk})^2 \end{aligned} \quad (1)$$

여기서 y_{pk} 는 기대 출력값을, o_{pk} 는 실제 출력값을 나타낸다. 본 논문의 목적에 맞게 수정한 전달 함수 f 는 다음과 같다.

$$f(x) = \frac{1 - e^{-\alpha x}}{1 + e^{-\alpha x}} \quad (2)$$

원래 시그모이드 함수의 출력범위는 '0'에서 '1'인 반면, 식 (2)의 출력값 범위는, 본 논문에서 사용할 양극 신호를 이용하는 통신 시스템에 적합한, '-1'에서 '1'이 된다. 그림 1은 식 (2)에서 여러 가지 α 값을 가지는 경우의 각 전달 함수에 대한 그림이다. α 에 의해 함수의 비선형성이 변화함을 알 수 있다.

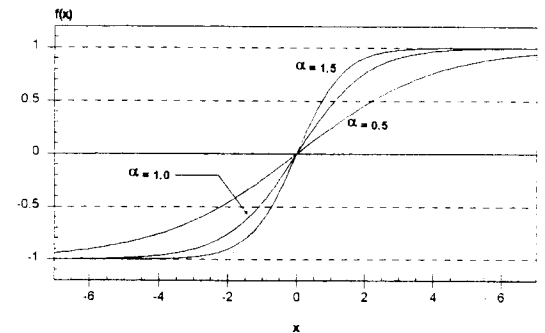


그림 1. 여러 가지 α 값을 가지는 경우의 각 전달 함수
Fig. 1. Comparison of the transfer functions with different α .

다음은 식 (2)의 일계 도함수이다.

$$f'(x) = \frac{\alpha}{2} (1 + f(x)) (1 - f(x)) \quad (3)$$

2. 한 부호어에 대한 출력 노드를 가진 NCD

그림 2는 부호율이 k/n인 콘볼루션 부호를 복호하기 위해, N개의 입력 노드와 k개의 출력 노드를 가진 역전과 신경회로망을 사용하여 NCD를 구현했을 때의 시스템도이다. 여기서 NCD의 입력단 노드들은 메모리 소자들로 구성된 이동 저장기(shift register)이며, 입력 노드수 N은 n의 정수배 이어야 한다. 관측된 신호열은 한 클럭에 n 심볼씩 NCD의 입력단에 가해지고 기존의 레지스터 내용은 그림의 화살표 방향으로 n 심볼씩 이동하게 된다. 이때, 입력단의 q개의 부호어 중에서, 학습시에 정한 어느 한 부호어의 복호된 k 심볼이 출력단에 나타나게 된다. 그림에서 \hat{m} 은 복호된 신호열이다. 이때 사용되는 NCD의 학습 표본들을 만드는 방법은 그림 3과 같다.

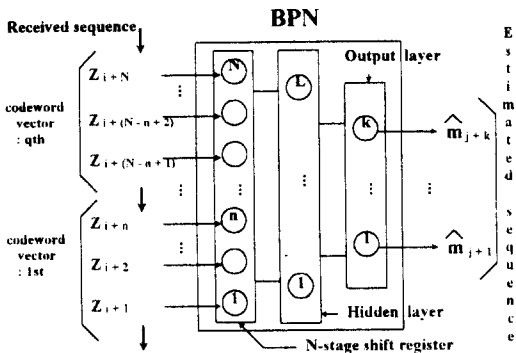


그림 2. 한 부호어에 대한 출력 노드를 가진 NCD
Fig. 2. NCD with k-output nodes according to one codeword.

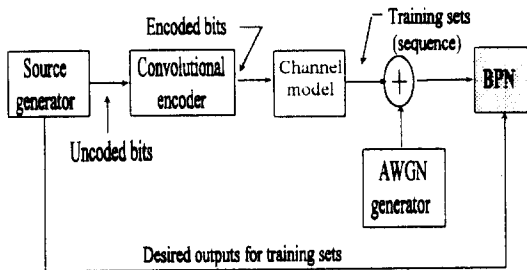


그림 3. NCD의 학습 샘플을 만드는 과정
Fig. 3. Data flow for the NCD training.

3. 여러 부호어에 대한 출력 노드를 가진 NCD

한 부호어에 대한 출력 노드가 k개인 NCD는 시스템은 간단하지만 잡음에 약한 단점이 생긴다. 더욱 정

확하게 클래스가 구분되어 오류에 강하게 하기 위해 여러 개의 부호어에 해당하는 출력 노드를 가진 NCD의 구현이 필요하다. 여러 개의 부호어에 해당하는 출력 노드를 가진 NCD는 그림 4와 같다. 입력단의 q개의 부호어 중에서, 학습시에 정한 여러 부호어의 복호된 M(=k*p) 심볼이 출력단에 나타난다는 점을 제외한 나머지는 한 부호어에 대한 출력 노드를 가진 NCD와 동일하다. 여기서 p는 q보다 작은 임의의 양의 정수이다. 학습시에만 M 개의 출력 노드를 이용하고, 실제 이용시에는 출력 노드중에서 임의로 정한 k 노드만을 이용하여, 한 부호어에 대한 출력 노드를 가진 NCD와 동일하게 이용할 수도 있다.

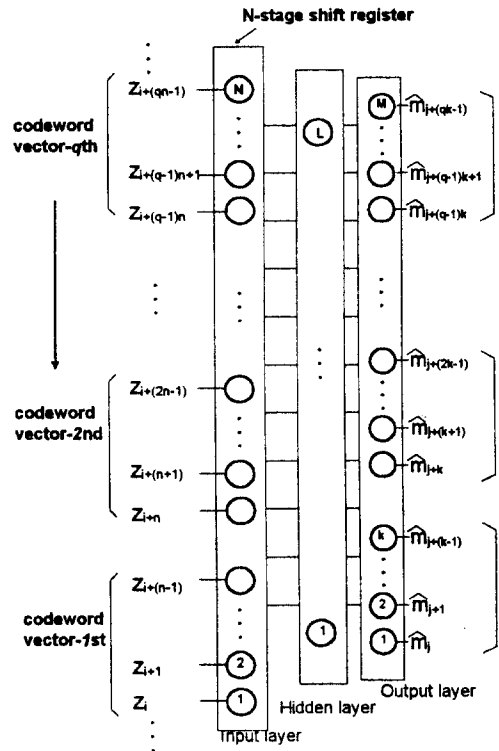


그림 4. 여러 개의 부호어에 대한 출력 노드를 가진 NCD의 구조도

Fig. 4. NCD with output nodes according to some codewords.

III. 신경회로망 콘볼루션 복호기의 확률적 특성 고찰

Kolmogorov 이론¹⁵⁾의 발견은 최초로 신경망의 사상에 관한 잠재능력을 보여주었으며, 실제세계의 많은

유용한 함수의 근사를 구현할 수 있으리라는 희망을 안겨주었다. 그 뒤 많은 논문들에서 세 계층으로 이루어진 신경망이면 실용적인 대부분의 함수를 원하는 정확도까지 근사화할 수 있다는 사실이 밝혀졌다. 또한, 다층 퍼셉트론은 블록 형태의 결정 영역 뿐만 아니라 일반 선형 분류화기는 할 수 없는 비블록 형태의 결정 영역도 형성할 수 있다. 그리고 주어진 몇 가지 제한 조건만 만족된다면, 은닉층에 무한개의 노드를 가진 세 계층으로 이루어진 역전파 신경회로망은 어떤 임의의 사상도 표현할 수 있다는 이론도 제시되었다 [6].

콘볼루션 부호화 및 복호 과정은 사상의 일종으로 간주될 수 있다. 그러므로 NCD는 복호 과정을 수행할 수 있으리라 짐작할 수 있다. 그러나, NCD가 과연 이론적으로 콘볼루션 복호 과정을 얼마나 정확하게 구현할 수 있는지에 대한 분석이 필요하다. 본 장에서는 NCD의 이론적 성능을 확률적 관점에서 근사화 해보기로 한다. 그리고 신경회로망의 특성을 알아보기 위해 전달 함수를 통과한 부가 백색 가우스 잡음의 확률 특성을 살펴본다.

1. 한 부호어에 대한 출력노드를 가진 NCD의 최대 확률 결정에 대한 확률적 근사화

부호율이 1/n인 콘볼루션 부호를 복호하기 위한 하나의 출력 노드를 가진 NCD는, N개의 노드로 이루어진 입력단에 가해진 Z의 일부분인 “ $z_i, z_{i+1} \dots z_{i+(N-1)}$ ” 중에서 “ $z_i, z_{i+1} \dots z_{i+(n-1)}$ ”에 해당하는 정보 비트 m_i 를 출력단에 출력시켜 주도록 구성된다. 이때 NCD는 길이 N인 모든 가능한 경로들과 그 경로에 해당하는 출력값을 학습하게 된다. 본 절에서는, NCD가 입력단에 가해진 “ $z_i, z_{i+1} \dots z_{i+(N-1)}$ ”에 대하여 m_i 를 결정할 때 이 결정이 최대 확률 관점에서 수행됨을 보인다.

어떤 관측된 열(sequence) $Z_i' \in Z'$ (Z' 는 길이가 N인 관측 가능한 모든 열 또는 벡터들의 집합)가 주어졌을 때 이 Z_i' 들을 두개의 집단으로 결정하는 경우의 최대 확률 결정 기준은 식 (4)과 같다 [9]. 여기서 m^1, m^2 는 Z_i' 이 각각 클래스 1과 클래스 2에 속했을 때의 결정되는 값이다. BPSK 경우, m^1 은 '1'로 m^2 는 '-1'으로써 각각 '1'과 '0'이 전송되었음을 의미한다. NCD의 경우에, Z' 는 입력단에 나타날 수 있는 모든 가능한 열(또는 벡터)들의 집합이 되고 m^1, m^2 는 Z_i' 가 입력단에 가해질 때 출력단에 나타나는 복호

된 정보 비트가 된다. 다시 말하자면, “ $z_i, z_{i+1} \dots z_{i+(N-1)}$ ”은 Z' 중의 어느 한가지이고, m_i 는 m^1, m^2 중의 어느 하나가 된다. m_i 가 m^1 인 경우의 Z_i' 들은 클래스 1에 속하게 되고 m_i 가 m^2 인 경우의 Z_i' 들은 클래스 2에 속하게 된다. 일반적으로 정보열을 구성하는 '1'과 '0'은 동일한 발생 확률을 가지기 때문에 $p(m^1)$ 과 $p(m^2)$ 는 0.5로써 동일하다. 그러므로 식 (4)는 식 (5)처럼 쓸 수 있다.

$$\Delta(Z') = \frac{p(Z'|m^2)}{p(Z'|m^1)} \begin{matrix} d_2 \\ > \\ < \\ d_1 \end{matrix} \quad 1 \quad (4)$$

$$\Delta(Z') = \frac{p(m^2|Z')}{p(m^1|Z')} \begin{matrix} d_2 \\ > \\ < \\ d_1 \end{matrix} \quad 1 \quad (5)$$

이때 식 (5)은 두 부문 분류 문제에서 베이지 결정 규칙(Bayes decision rule)에 의한 분별 함수(discriminant function)와 동일하다 [10].

$$g_o(Z') = [p(m^1|Z') - p(m^2|Z')] \begin{matrix} d_1 \\ > \\ < \\ d_2 \end{matrix} \quad 0 \quad (6)$$

그림 1에서 k가 '1'인 경우를 생각해보자. 임의의 Z_i' 가 NCD의 입력이라고 할 때, 이 입력에 대한 기대 출력값을 m_i 라고 하자. m_i 는 Z_i' 가 클래스 1일 때 '1', 클래스 2일 때 '-1'이다. 이와 같은 내용은 식 (5)에서 $g_o(Z')$ 가 '0'보다 크면 Z' 가 클래스 1이며 m^1 이 전송되었다고 결정하고 '0'보다 작으면 Z' 가 클래스 2이며 m^2 가 전송되었다고 결정하는 것과 동일하다. 역전파 신경회로망의 은닉 노드의 출력들을 벡터 I_i 라고 하면, 역전파 신경회로망의 입력과 출력과의 관계는 식 (6, 7, 8)처럼 표현할 수 있다. 여기서 O_i 는 역전파 신경회로망의 실제 출력값이고, 각 벡터 상단에 있는 h와 o는 은닉층과 출력단을 의미하는 표기이다.

$$W^h = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1N} \\ w_{21} & w_{22} & \dots & w_{2N} \\ \cdot & \dots & \cdot & \cdot \\ \cdot & \dots & \cdot & \cdot \\ \cdot & \dots & \cdot & \cdot \\ w_{L1} & w_{L2} & \dots & w_{LN} \end{pmatrix} \quad (7-1)$$

$$W^o = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1L} \\ w_{21} & w_{22} & \dots & w_{2L} \\ \cdot & \dots & \cdot & \cdot \\ \cdot & \dots & \cdot & \cdot \\ \cdot & \dots & \cdot & \cdot \\ w_{M1} & w_{M2} & \dots & w_{ML} \end{pmatrix} \quad (7-2)$$

$$I_i = (i_1 \ i_2 \ \dots \ i_L)^T \quad O_i = (o_1 \ o_2 \ \dots \ o_M)^T \quad (8)$$

식 (9-1)에서 f^h 와 f^o 는 각각 식 (2)과 같이 정의된 전달 함수 L개와 M개로 구성된 벡터 함수이고, 식 (9-2)의 f_{BPN} 과 W 는 식 (8)의 입출력 관계를 전부 합쳐서 표현하는 임의의 벡터 함수이다. 지금의 경우 ($k=1$)에서 식 (9)의 O_i 는 스칼라값이다.

$$I_i = f^h(W^h Z_i') \quad (9-1)$$

$$O_i = f^o[W^o f^h(W^h Z_i')] = f_{BPN}(W Z_i') \quad (9-2)$$

Z' 에 속하는 열의 수를 t 라고 하고, 그중 클래스 1($m^1=1$)의 개수를 t_1 , 클래스 2($m^2=-1$)의 개수를 t_2 라고 가정하자 ($t=t_1+t_2$). 이제 MSE 관점에서 역전파 신경회로망의 해를 구해보자.

식 (1)을 계산의 편의상 1/2를 빼고 전체 학습 샘플에 대해 다시 표현하면 식 (10)이 된다.

$$\begin{aligned} J_L(W) &= \sum_{i=1}^L E_i = \sum_{i=1}^L \sum_{k=1}^M (y_{ik} - o_{ik})^2 \\ &= \sum_{i=1}^L (o_{i1} - y_{i1})^2 \quad (M=1) \quad (10) \\ &= \sum_{i=1}^L \{ f_{BPN}(W Z_i') - m_i \}^2 \end{aligned}$$

다시 식 (10)을 클래스 1, 2에 따라 나누면 식 (11), (12)와 같이 된다.

$$\begin{aligned} J_L(W) &= \sum_{Z_i' \in \text{class1}} \{ f_{BPN}(W Z_i') - m^1 \}^2 \\ &\quad + \sum_{Z_i' \in \text{class2}} \{ f_{BPN}(W Z_i') - m^2 \}^2 \\ &= \sum_{Z_i' \in \text{class1}} \{ f_{BPN}(W Z_i') - 1 \}^2 \\ &\quad + \sum_{Z_i' \in \text{class2}} \{ f_{BPN}(W Z_i') + 1 \}^2 \\ &= t [\frac{t_1}{t} \frac{1}{t_1} \sum_{Z_i' \in \text{class1}} \{ f_{BPN}(W Z_i') - 1 \}^2 \\ &\quad + \frac{t_2}{t} \frac{1}{t_2} \sum_{Z_i' \in \text{class2}} \{ f_{BPN}(W Z_i') + 1 \}^2] \quad (11) \end{aligned}$$

$$\begin{aligned} \frac{J_L(W)}{t} &= \frac{t_1}{t} \frac{1}{t_1} \sum_{Z_i' \in \text{class1}} \{ f_{BPN}(W Z_i') - 1 \}^2 \\ &\quad + \frac{t_2}{t} \frac{1}{t_2} \sum_{Z_i' \in \text{class2}} \{ f_{BPN}(W Z_i') + 1 \}^2 \quad (12) \end{aligned}$$

t 가 아주 큰 경우 $\frac{t_1}{t} \rightarrow p(m^1)$, $\frac{t_2}{t} \rightarrow p(m^2)$ 이므로 $\frac{J_L(W)}{t} \underset{t \rightarrow \infty}{\approx} \overline{J_L(W)}$ 이라 하면 식 (12)는

식 (13)과 같이된다¹¹⁰⁾.

$$\begin{aligned} \overline{J_L(W)} &= p(m^1) E_1[\{ f_{BPN}(W Z') - 1 \}^2] \\ &\quad + p(m^2) E_2[\{ f_{BPN}(W Z') + 1 \}^2] \quad (13) \end{aligned}$$

식 (13)에서 $E_1[\{ f_{BPN}(W Z') - 1 \}^2]$ 는 $\int_{-\infty}^{+\infty} \{ f_{BPN}(W Z') - 1 \}^2 p(Z', m^1) dZ'$ 와 동일하므로 식 (14)과 같이 다시 쓸 수 있다.

$$\begin{aligned} \overline{J_L(W)} &= \int_{-\infty}^{+\infty} [f_{BPN}(W Z') - 1]^2 p(Z', m^1) dZ' \\ &\quad + \int_{-\infty}^{+\infty} [f_{BPN}(W Z') + 1]^2 p(Z', m^2) dZ' \\ &= \int_{-\infty}^{+\infty} f_{BPN}(W Z')^2 p(Z') dZ' \\ &\quad - 2 \int_{-\infty}^{+\infty} f_{BPN}(W Z') g_o(Z') p(Z') dZ' + 1 \\ &= \int_{-\infty}^{+\infty} [f_{BPN}(W Z') - g_o(Z')]^2 p(Z') dZ' \\ &\quad + [1 - \int_{-\infty}^{+\infty} g_o(Z')^2 p(Z') dZ']^2 \quad (14) \end{aligned}$$

윗식의 우변 둘째 항은 가중치와는 무관한 확률적인 값이므로, 역전파 신경회로망이 $J_L(W)$ 을 최소화하도록 학습한다는 것은

$\int_{-\infty}^{+\infty} \{ f_{BPN}(W Z') - g_o(Z') \}^2 p(Z') dZ' = E \{ \langle f_{BPN}(W Z') - g_o(Z') \rangle^2 \}$ 을 최소화한다는 것이므로, 결국 $f_{BPN}(W Z')$ 을 $g_o(Z')$ 으로 근사화하는 것이 된다 (식 15).

$$f_{BPN}(W Z') \rightarrow g_o(Z') \quad (15)$$

본 논문에서 이용된 역전파 신경회로망에선 $f_{BPN}(W Z')$ 이 '0'보다 큰지 작은지에 따라 최종 출력값(복호된 값)을 '1' 또는 '-1'로 결정하게 된다. 이것은 식 (15)에 의해 $g_o(Z')$ 가 '0'보다 큰지 작은지에 따라 최종 출력값을 '1' 또는 '-1'로 결정하는 것과 동일하다. 그러므로 NCD는 식 (6)와 마찬가지로 최대 확률 관점에서 출력값을 결정함을 알 수 있다. 부호율이 k/n 일 때에도 같은 방식으로 증명될 수 있다. 그리고 m^1 , m^2 가 각각 '1', '0'일 때의 증명은 참고 문헌¹⁷⁾에 설명되어 있다.

2. 여러 부호어에 대한 출력노드를 가진 NCD의 최대 확률 결정에 대한 확률적 근사화
본 절에서는 여러 개의 부호어에 해당하는 출력 노드를 가진 NCD에 대한 확률적 특성을 살펴본다.

수신단에서 복호되어야 할 전체 열 Z 가 송신단에서 보낸 q 개의 부호어로 이루어진 r 에 부가 백색 가우스 잡음이 더해진 모양이라고 가정하자. 그리고 NCD의 입력단은 Z 가 전부 입력에 가해질 수 있도록 $q*n$ 개의 신경 노드로 구성되고, 출력단 역시 $q*k$ 개의 신경 노드로 구성되었다고 하자. 완벽한 역전과 신경회로망의 성능을 위해 은닉층 노드는 무한개라고 가정한다. 이때 역전과 신경회로망의 MSE 해는 식 (16), (17)와 같다.

$$J_L(\mathbf{W}) = \sum_{i=1}^L \sum_{k=1}^M (y_{ik} - o_{ik})^2 \tag{16}$$

$$= \sum_{i=1}^L \sum_{k=1}^M (f_{BPN}^k(\mathbf{WZ}_i') - m_i)^2$$

$$J_L(\mathbf{W}) = \sum_{i=1}^L (f_{BPN}^1(\mathbf{WZ}_i') - m_i)^2 \tag{17}$$

$$+ \dots + \sum_{i=1}^L (f_{BPN}^M(\mathbf{WZ}_i') - m_i)^2$$

식 (16), (17)에서 f_{BPN}^k 는 식 (9-1), (18)의 관계를 가지는 함수이고, 식 (18)의 f^{ik} 의 k 는 f 가 역전과 신경회로망의 k 번째 출력 노드와 관련된 전달 함수임을 뜻한다.

$$O_{ik} = f^{ik}(\mathbf{W}^o Z_i') \quad O_{ik} = f_{BPN}^k(\mathbf{WZ}_i') \tag{18}$$

식 (17)의 좌변의 각 항은 식 (10)과 동일하다고 볼 수 있다. 그러므로 같은 방식으로 전개해보면 식 (19)과 같은 결과를 쉽게 얻을 수 있다. 이 식에서 g_o^k 는 역전과 신경회로망의 k 번째 출력 노드와 관련된 식 (6)의 기준 함수이다.

$$\overline{J_L(\mathbf{W})} = \int_{-\infty}^{+\infty} (f_{BPN}^k(\mathbf{WZ}') - g_o^k(Z'))^2 p(Z') dZ' \tag{19}$$

$$+ [1 - \int_{-\infty}^{+\infty} g_o^M(Z')^2 p(Z') dZ']^2 + \dots$$

$$+ \int_{-\infty}^{+\infty} (f_{BPN}^M(\mathbf{WZ}') - g_o^M(Z'))^2 p(Z') dZ'$$

$$+ [1 - \int_{-\infty}^{+\infty} g_o^M(Z')^2 p(Z') dZ']^2$$

그러므로 역전과 신경회로망이 식 (19)을 최소화시킨다는 것은 M 개의 $f_{BPN}^k(\mathbf{WZ}')$ 을 $g_o^k(Z')$ 으로 동시에 근사화한다는 것과 동일하다.

$$f_{BPN}^k(\mathbf{WZ}') \rightarrow g_o^k(Z') \quad (k = 1, 2, \dots, M) \tag{20}$$

Z 의 길이가 어느 정도(부호율이 1/2일 때 대략 구속장 길이의 4~5배)이상만 되면, 혹은 q 를 무한대로 보내면, 모든 k 에 대하여 식 (20)을 동시에 만족시키는 NCD 출력단의 출력열 \hat{m} 은 거의 유일하게 하나만 존

재하게 된다^{11, 31}. 그리고 이때 \hat{m} 은 송신단에서 송신이 가능한 길이 $q*n$ 을 가진 모든 경로들 중에서, Z 와 확률적으로 가장 가까운 경로에 해당하는 정보 비트열이 된다.

이상에 여러 식들에 의해서 '여러 개의 부호어 가치에 해당하는 출력 노드를 가진 NCD'는 최대 확률관점에서 Z 를 복호하는 것과 동일한 일을 한다고 말할 수 있다.

3. 전달 함수를 통과한 부가 백색 가우스 잡음의 확률 특성

통신 시스템에서 발생하는 잡음은 일반적으로 부가 백색 가우스 잡음이라고 가정한다. 즉, 가우스 잡음이다. 여기서는 부가 백색 가우스 잡음 채널하의 QPSK 신호 성상도를 살펴봄으로써 신경노드의 특성을 알아본다.

식 (21)은 전송 신호 m_x 에 평균이 '0'인 가우스 잡음이 더해진 후의 수신된 신호 x 의 확률 밀도 함수이다.

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma^2}} \tag{21}$$

위의 식에서 m_x 는 '1' 또는 '-1'의 값을 가진다. 수신된 신호가 신경회로망(특별히 시그모이드 함수)을 통과하게 되면 그 수신된 데이터들의 확률적 특성이 변하게 된다. 여기서 평균이 '1'이거나 '-1'인 가우스 확률 분포가 간단한 신경노드 하나를 통과한 뒤 어떤 확률 분포로 변하는지를 알아본다.

전송 신호를 x 라고 하고 신경노드의 가중치를 w 라고 할 때, 가중치가 곱해진 신호를 y 라고 하자. 이때 시그모이드 함수를 통과한 y 를 z 라고 하자.

$$y = wx \tag{22}$$

전달 함수를 식 (23)이라고 하면, z 는 식 (24)과 같이 표현할 수 있다.

$$g(x) = \frac{2}{1+e^{-ax}} - 1 \tag{23}$$

$$z = g(y) = \frac{1 - e^{-ay}}{1 + e^{-ay}} \tag{24}$$

결국 z 의 확률 밀도 함수는 식 (25)와 같다.

$$f_z(z) = \frac{2}{w\sigma\sqrt{2\pi}} \frac{1}{a|1-z^2|} e^{-\left(\frac{1}{a} \ln \frac{1+z}{1-z} - um_x\right)^2 / 2\sigma^2 w^2} \tag{25}$$

그림 5는 식 (21)을 확률 분포 함수로 가지는 수신

된 QPSK 신호의 성상도를 보여 주고 있다. 그림 6은 수신된 QPSK 신호가 신경노드 하나를 통과한 후에 나타나는 성상도이다. 이 신호의 확률 분포 함수는 식 (25)이다. w 와 α 값이 커질수록 그림 6의 점들은 더욱 더 경계쪽으로 이동하게 된다.

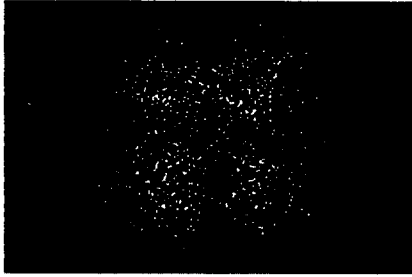


그림 5. 부가 백색 가우스 잡음이 더해진 QPSK 신호가 신경노드를 통과하기 전의 성상도
Fig. 5. The constellation of the QPSK signal before a sigmoid.

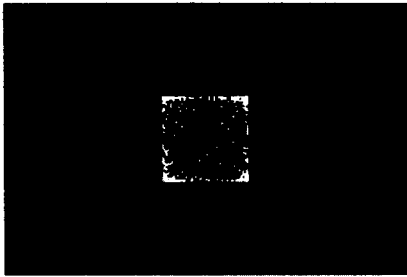


그림 6. 부가 백색 가우스 잡음이 더해진 QPSK 신호가 신경노드를 통과한 후의 성상도. ($w = 2.0$, $\alpha = 2.0$)
Fig. 6. The constellation of the QPSK signal after a sigmoid..

이들 그림에서 알 수 있듯이, 잡음에 의해 나빠진 QPSK 신호의 성상도가 단지 하나의 신경 노드를 통과함으로써 인하여 개선됨을 알 수 있다. 이런 사실은 신경회로망의 비선형적 사상 능력을 명확히 보여주고 있다. 실제로 사용되는 역전파 신경회로망은 여러 개의 신경노드로 구성되어 있으므로 좀 더 힘든 비선형 사상도 수행할 수 있다.

IV. 실험 결과

이 장에서는 NCD의 오류 성능을 컴퓨터 실험을 통하여 알아보고 그 결과를 나타낸다. 실험 환경은 부가

백색 가우스 잡음 채널이고, 구속장 길이(K)가 3, 부호율이 1/2인 체계적/비체계적 콘볼루션 부호가 이용된다. 실험에 이용된 복조 방식은 180° BPSK이며 동기 복조 방식을 채택했다.

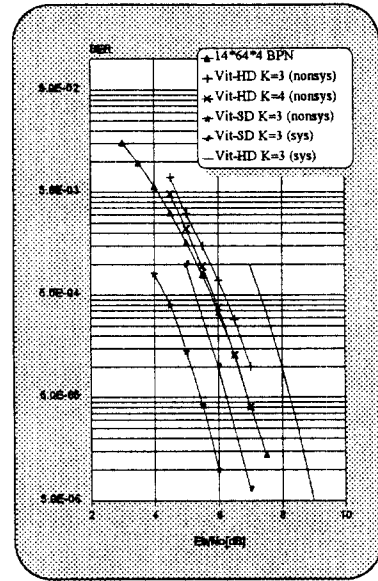


그림 7. 체계적 콘볼루션 부호에 의해 학습된 NCD와 비체계적 및 체계적 콘볼루션 부호를 사용한 32 비트 경로 메모리를 가진 비터비 복호기의 오류 성능 비교 곡선 (K=3, code rate=1/2)
Fig. 7. Bit error rate versus E_b/N_0 . The NCD trained with the K=3 systematic code compared to Viterbi decoder for the nonsystematic code with 32-bit paths.

모든 실험에서 한 부호어에 대한 출력노드를 가진 NCD보다 여러 부호어에 대한 출력노드를 가진 NCD가 더 나은 성능을 보여주었으므로, 본 장에서 제시하는 실험 결과들은 모두 두 번째 NCD에 의해서 행해진 것들이다. 그림 7에서 14*64*4 역전파 신경회로망은 K=3, 4의 비체계적 콘볼루션 부호를 사용한 경관정 비터비 복호기보다 나은 성능을 보임을 알 수 있다. 그러나 K=3인 비체계적 콘볼루션 부호를 사용한 연관정 비터비 복호기보다는 약 1.5 [dB] 정도 성능이 떨어진 것이다. 그러나, 같은 종류의 체계적 콘볼루션 부호를 사용한 비터비 복호기의 결과와 비교한 경우, 비터비 복호기와의 오류 성능차가 0.5 [dB] 이하인 것을 알 수 있다.

이와 같이 신경회로망 콘볼루션 복호기가 체계적 콘

루션 부호에 대해서 더 좋은 성능을 보여 주는 이유를 분석하기 위해 그림 8에 나타나 있는 간단한 시스템을 예로 들어보자. 편의를 위해서 입력단에 가해지는 정보 열은 r 에 부가 백색 가우스 잡음이 더해진 z 가 아니라 r 그 자체라고 가정한다. 그림 8의 시스템의 동작 원리는 표 1과 같다.

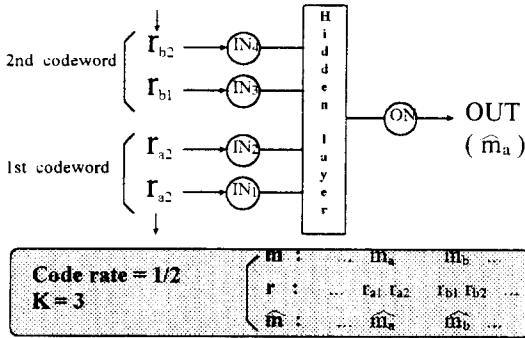


그림 8. 4개의 입력 노드와 1개의 출력 노드를 가진 NCD
Fig. 8. NCD with 4 input nodes and 1 output node.

표 1. 그림 8에 나타나 있는 시스템의 동작 원리
Table 1. The relationship between inputs and outputs.

노드 이름	시스템의 입력 패턴	기대 출력
	$IN_4 \quad IN_3 \quad IN_2 \quad IN_1$	ON (OUT)
노드 내용	$r_{b2} \quad r_{b1} \quad r_{a2} \quad r_{a1}$	m_a
	$r_{c2} \quad r_{c1} \quad r_{b2} \quad r_{b1}$	m_b
	$r_{d2} \quad r_{d1} \quad r_{c2} \quad r_{c1}$	m_c

우선 체계적 콘볼루션 부호에 대하여 살펴보자. 모든 가능한 입력열과 그때의 기대 출력을 표 2에 나타내었다. 이 표에서 알 수 있듯이 NCD의 입력단에 나타날 수 있는 가능한 패턴의 종류는 16가지이다. 이때 신경회로망의 출력 노드가 하나이므로, 이들 입력열에 대한 기대 출력값은 '-1' 혹은 '1'이 된다. 이런 경우 신경회로망의 학습이란, 16가지 입력 패턴을 두 개의 클래스로 분류해 줄 수 있는 가중치를 찾아내는 과정으로 볼 수 있다. 표 2를 살펴보면, 클래스를 구분할 때 $r_{b2} \quad r_{b1}$ 은 어떤 판단의 기준도 제시하지 못하고 단지 $r_{a2} \quad r_{a1}$ 에 의해 클래스가 결정됨을 알 수 있다. 즉, $r_{a2} \quad r_{a1}$ 가 (-1, -1) 또는 (1, -1)이면 클래스1이 되고

표 2. 체계적 콘볼루션 부호를 사용한 경우 그림 8의 NCD에서 가능한 입출력 패턴

Table 2. All possible input patterns in the case of the nonsystematic code.

$m_b \quad m_a \quad m_c$ 에 의한 부호어		$m_a \quad m_c \quad m_b$ 에 의한 부호어		Class 종류
r_{b2}	r_{b1}	r_{a2}	r_{a1}	
-1	-1	-1	-1	class1
1	-1	-1	-1	
1	1	-1	-1	
-1	1	-1	-1	
-1	-1	1	-1	class2
1	-1	1	-1	
-1	1	1	-1	
1	1	1	-1	
1	-1	-1	1	class2
-1	-1	-1	1	
-1	1	-1	1	
1	1	-1	1	

표 3. 비체계적 콘볼루션 부호를 사용한 경우 그림 8의 NCD에서 가능한 입출력 패턴
Table 3. All possible input patterns in the case of the systematic code.

$m_b \quad m_a \quad m_c$ 에 의한 부호어		$m_a \quad m_c \quad m_b$ 에 의한 부호어		Class 종류
r_{b2}	r_{b1}	r_{a2}	r_{a1}	
-1	-1	-1	-1	class1
1	-1	-1	-1	
1	1	-1	-1	
-1	1	-1	-1	
1	-1	-1	1	class2
-1	-1	-1	1	
-1	1	-1	1	
1	1	-1	1	

(1, 1) 또는 (-1, 1)이면 클래스2가 된다. 이를 이차원 평면상에서 살펴본 것이 그림 9이다. 여기서 알 수 있

듯이 클래스를 구분하는 것이 '선형 분리 가능성(linear separability)'을 가지고 있는 단순한 경우이므로, 이런 동작을 하는 가중치를 찾는 것은 간단하다. 다시 말하자면, 신경회로망의 학습이 쉽게 이루어진다는 것이다. 이와 같은 경우는 단순한 ADALINE (adaptive linear neuron)이나 퍼셉트론으로도 구현할 수 있다.

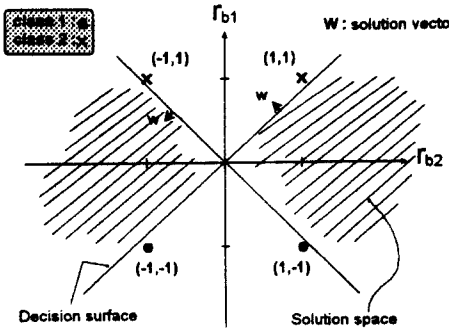


그림 9. 표 2의 결과를 이차원 평면으로 도시한 그림
Fig. 9. The patterns of Table 2.

그러나, 비체계적 콘볼루션 부호의 경우는 표 3 처럼 체계적 콘볼루션 부호와는 다른 양상을 보인다. 클래스를 구분하는 기준은 r_{a2} r_{a1} 이 아니라 r_{b2} r_{b1} 이다. r_{b2} r_{b1} 가 (-1, -1) 또는 (1, 1)이면 클래스1이 되고, (1, -1) 또는 (-1, 1)일 때 클래스2가 된다. 이를 마찬가지로 이차원 평면에서 살펴보자. 그림 10에서 알 수 있듯이 비체계적 콘볼루션 부호의 경우는 '선형 분리 가능성'을 가진 단순한 경우가 아니다. 이는 일종의 XOR 문제로서 최소한 2개층 이상과 셋 이상의 신경노드가 필요한 경우이다. XOR 문제는 신경회로망이 다룰 수 있는 것 중에서 전형적으로 까다로운 문제에 속한다. 게다가, 기본적인 네점 (-1, -1), (-1, 1), (1, -1), (1, 1)에 잡음까지 섞인 입력 패턴들에 대한 XOR 문제를 해결해 주는 신경회로망의 가중치를 구하는 것은 한층 어려운 문제이다. 물론 많은 노드를 가진 역전파 신경회로망은 이를 해결할 수 있지만, 역시 앞의 '선형 분리 가능성'을 가진 경우보다는 학습이 어렵다. 결국, 신경회로망 콘볼루션 부호의 입력 노드수가 늘어남에 따라 XOR문제가 중첩되어 발생하므로, 비체계적 콘볼루션 부호를 분류기 개념으로 푸는 것은 점차로 어려워질 것이라는 점은 쉽게 예측할 수 있다. 그러나, 이 말은 해결 불가능하다는 것을 의미하는 것이 아니라 학습이 힘들다는 것이다. 그러므로 좀더 나은 학

습 방법을 찾아낸다면 비체계적 콘볼루션 부호에 대한 신경회로망 복호기 역시 좋은 성능을 보여 줄 수 있을 것이다.

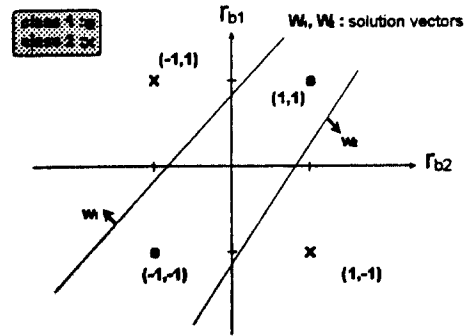


그림 10. 표 3의 결과를 이차원 평면으로 도시한 그림
Fig. 10. The patterns of Table 3.

V. 결 론

본 논문에서 제안된 NCD는, 비선형 사상의 일종인 콘볼루션 부호 과정을 최적화된 비터비 복호기와 마찬가지로 최대 확률 관점에서 행함을 알 수 있다. 그리고, K=3, 부호율이 1/2인 체계적 콘볼루션 부호에 대하여, 비교적 간단한 구조를 가지고 학습된 NCD의 성능이 K=3과 K=4인 비체계적 콘볼루션 부호에 의한 경관정 비터비 복호기의 성능보다 뛰어났다. 또한 이 NCD는 K=3인 체계적 콘볼루션 부호에 의한 연관정 비터비 복호기와는 유사한 성능을 보여 준다. 그러나, K=3인 비체계적 콘볼루션 부호에 의한 연관정 비터비 복호기의 성능보다는 떨어지는데, 이는 사용된 부호의 차이가 주된 이유이다.

그러나, 신경망은 학습방법이나 환경에 의하여 그 성능차가 생기기 때문에 이 논문에서 보인 NCD의 결과들이 주어진 환경하에서 나타낼 수 있는 최적의 성능이라고는 할 수 없다. 그러므로, 비체계적 콘볼루션 부호의 사상적 특성을 잘 파악하여 유전자 알고리즘을 혼합하거나 확률 신경회로망을 사용하여 적절히 학습을 할 수 있다면 최대 확률 관점에서 복호를 행하는 NCD는 비체계적 콘볼루션 부호에 대해서도 비터비 복호기의 성능을 쫓아갈 수 있으리라 기대된다. 또한 신경회로망의 특성을 살릴 수 있는 위성 채널 같은 비선형 특성이 부가된 경우, NCD는 더 나은 결과를 보여 주리라 예상된다.

참 고 문 헌

- [1] G. D. Forney, JR., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268-278, Mar. 1973.
- [2] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 751-772, Oct. 1971.
- [3] J. A. Heller and I. M. Jacobs, "Viterbi decoding for satellite and space communication," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 835-847, Oct. 1971.
- [4] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci.*, vol. 79, pp. 2554-2558, Apr. 1982.
- [5] A. N. Kolmogorov, "On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition," *Dokl. Akad. Nauk VSSR*, vol. 144, pp. 954-956, Nov. 1957.
- [6] M. Amari, "Mapping abilities of three-layer neural networks," *Int'l Joint Conf. Neural Networks*, Washington D.C., pp. 419-423, June 1989.
- [7] O. K. Ersoy and D. S. Hong, "Parallel, self-organizing, hierarchical neural networks-II," *IEEE Trans. Industrial Electronics*, vol. 40, no. 2, pp. 218-227, April 1993.
- [8] C. W. You and D. S. Hong, "Neural Convolutional Decoders," *Int'l Conf. on Neural Information Processing*, Seoul, pp. 987-992, Oct. 1994.
- [9] J. L. Melsa and D. L. Cohn, *Decision and Estimation Theory*, McGraw-Hill, 1978.
- [10] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley-Interscience, 1973.

저 자 소 개



劉哲雨(正會員)

1993년 2월 연세대학교 전자공학과 졸업(공학사). 1995년 2월 연세대학교 전자공학과 졸업(공학석사). 1995년 3월 ~ 현재 연세대학교 전자공학과 박사과정.

주관심분야는 디지털 통신 시스템, 신호 처리, 신경회로망 등임

洪大植(正會員)

1983년 2월 연세대학교 전자공학과 졸업(공학사). 1985년 2월 연세대학교 전자공학과 졸업(공학석사). 1985년 3월 ~ 1986년 8월 연세대학교 산업기술 연구소 연구원. 1990년 8월 Purdue University 졸업(Ph.D). 1990년 9월 ~ 1991년 8월 Purdue University Post-Doctoral Research Associate. 1991년 9월 ~ 1995년 9월 연세대학교 전자공학과 조교수. 1995년 9월 ~ 현재 연세대학교 전자공학과 부교수. 주관심분야는 디지털 통신 시스템 및 신경회로망 등임

康昌彦(正會員) 第31卷A編第4號 參照