

論文93-30B-8-9

한국어 모음인식 신경회로망 집적회로의 제작

(Fabrication of a Neural Network IC for Korean Vowels Recognition)

蔡相勳*, 尹台焄*, 金在昌*

(Sang Hoon Chai, Tae Hoon Yoon and Jae Chang Kim)

要約

한국어 모음을 인식할 수 있는 신경회로망을 설계한 다음 집적회로로 제작하였다. 신경회로망은 3층 구조로 구성하였으며, Back Propagation 알고리즘에 의해 학습시켰다. 신경회로망을 집적회로로 구현시 뉴런몸체는 CMOS 인버터를, 시냅스는 폴리실리콘 저항을 사용하였다.

Abstract

This paper presents a neural network IC for Korean vowels recognition. The neural network is composed with three levels and which is learned by Back Propagation algorithm. In the neural network IC, the neuron bodys and synapses are implemented with CMOS inverters and ion-implanted polysilicon resistors.

1. 서론

신경회로망은 병렬분산처리(parallel distributed processing)를 하기 때문에 음성 또는 영상신호에 대해 실시간 처리가 가능할 뿐만 아니라 기존의 폰 노이만형 디지털 컴퓨터로 불가능하였던 많은 시간을 요하는 지능적인 문제를 집단적(emergent collective)으로 처리할 수 있다. 또한 결함 복구능력(fault tolerance), 적응능력(adaptivity) 등의 여러가지 장점을 갖고 있기 때문에 패턴인식, 영상처리, 음성인식, 오정보 정정, CAM(content addressable memory), 최적화 문제 등과 같은 분야에서 폭넓게 응

용되고 있다.^[1,2]

본 연구에서는 한국어 5개 기본모음(/아/, /애/, /어/, /우/, /이/)을 구별 인식할 수 있는 음성인식 신경망을 구성한 다음 이를 집적회로로 설계, 제작하여 음성인식 정도를 측정해 보았으며, 전기적 특성의 분석을 통해 신경회로망 집적회로를 최적화시켰다. 신경회로망의 구성과 학습에는 BP(Back Propagation) 알고리즘을 사용하였으며, 이를 단위뉴런은 증폭기로, 시냅스는 저항을 사용하여 전자회로로 설계하였다. 그리고 신경회로망을 집적회로로 구현하기 위하여 증폭기는 CMOS 인버터로, 저항은 폴리실리콘 저항을 사용하여 레이아웃하였다. 소비전력과 집적도, 음성인식능력 면에서 신경회로망 집적회로의 최적화된 조건을 구하기 위하여 폴리실리콘막의 면저항(sheet resistance)을 바꾸는 방법을 사용하여 회로 전체의 병렬저항을 8단계로 나누어서 8종류의 집적회로를 제작한 다음 음성인식 정도의 차이등을 관찰해 보았다.

*正會員, 釜山大學校 電子工學科

(Dept. of Elec. Eng., Pusan Nat'l Univ.)

(본 연구는 1991년도 교육부 학술연구 조성비에 의해 이루어졌음)

接受日字: 1993年 1月 15日

II. 뉴런과 뉴런의 모델화

인간의 뇌는 뉴런이라는 수많은 기본세포로 구성되어 있으며, 이러한 뉴런들에 의해 신체 한 부위에서 다른 부위로의 정보 전달과 입력된 정보에 대한 처리가 이루어진다. 뉴런은 뉴런몸체(neuron body)와 시냅스(synapse)로 이루어져 있으며, 이를 전자소자를 이용하여 집적회로로 구현할 때는 뉴런몸체는 증폭기로, 시냅스는 저항으로 구성한다. 뉴런증폭기의 전달함수를 일반적인 수식으로 나타내면

$$V_i = \frac{\tanh(\lambda \cdot V/2)}{2} \tag{1}$$

로 표시된다. 여기서 λ 는 전달함수의 기울기를 나타내는 상수로써 증폭기로 사용되는 전자소자의 종류에 따라 각기 다른 값을 갖는다. 또한, V 는 앞단 뉴런증폭기들의 출력 V_j 들이 흥분이나 억제작용을 거쳐 다음단으로 전달되는 총 입력전압으로써

$$V = R_i \cdot \sum_{j=1}^n \frac{V_j}{R_{ij}} \tag{2}$$

로 나타낼 수 있다. 이 때 R_{ij} 는 앞단의 뉴런증폭기와 다음단의 뉴런증폭기를 연결하는 각 저항값들을 나타낸다. 그리고 R_i 는 이 저항들의 전체 병렬저항값으로써 신경회로망 집적회로의 소비전력과 집적도 및 음성인식 정도를 결정지어 주는 중요한 변수로 작용하며

$$R_i = 1 / \left(\sum \frac{1}{R_{ij}} \right) \tag{3}$$

의 관계가 있다. 각 개별저항을 나타내는 R_{ij} 는 각 뉴런들 사이의 연결세기인 W_{ij} 값에 따라 저항의 값이 결정되며 이를 수식으로 표현하면 다음과 같다.

$$R_{ij} = \frac{1}{|W_{ij}|} \cdot R_i \tag{4}$$

즉, R_{ij} 는 전체 병렬저항 R_i 값과 뉴런사이의 가중치 W_{ij} 값에 의해 그 값이 결정된다.

III. 음성인식 신경회로망의 설계

신경회로망을 음성인식에 이용한 연구들로서 골드(Gold)는 홉필드(Hopfield) 모델을 사용하여 모음인식을 시도하였고, 코호넨(Kohonen)등은 음성의 포만트(formant) 주파수를 입력으로 하고 자율조직(self organization) 신경회로망을 구성하여 음성인

식에 대한 실험을 하였다. 또한, Ohotomo 등은 BP 학습 알고리즘을 이용하여 합성모음의 분포도를 신경회로망에 학습시켜 기본모음의 인식을 실험하였다.^[3]

본 연구에서는 그림 1과 같은 3층 구조를 갖는 모음인식 신경회로망을 구성하였으며, BP 학습 알고리즘을 이용하여 그림 2의 음운분포도에 나타나있는 한국어 5개 기본모음에 대한 제1, 제2 포만트 주파수 60개 중 40개를 신경회로망에 학습시켰다.^[4,5]

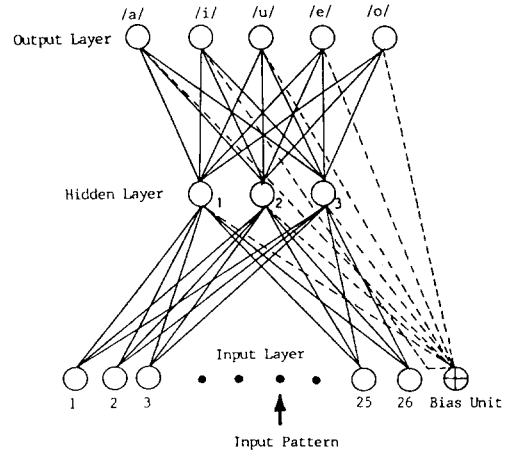


그림 1. 한국어 기본모음 인식을 위한 신경회로망 모델

Fig. 1. A neural network model for Korean vowels recognition.

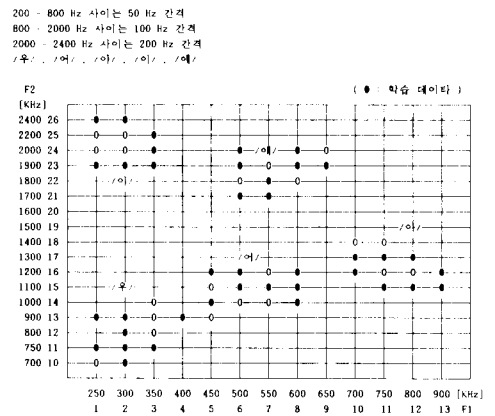


그림 2. 한국어 기본모음의 포만트 주파수 음운분포도

Fig. 2. Formant frequency distribution of Korean vowels.

구성된 신경회로망은 음성의 포먼트 주파수를 디지털화된 입력신호를 인가하기 위한 26개의 단자와, 바이어스 전압을 인가하기 위한 1개의 단자로 이루어진 입력층을 갖는다. 또한 인식된 5개의 음성을 구별해주는 5개의 단자를 갖는 출력층이 있으며, 입력층과 출력층 사이에는 한 층의 히든층(hidden layer)이 배치되어 3층 구조를 갖는다. 입출력층의 뉴런들은 제1, 제2 포먼트 주파수 검출을 위한 전처리 시스템 및 인식된 음성을 외부로 나타내주는 후처리 시스템과 각각 연결되며, 각층 간에는 각기 다른 연결세기에 의해 연결되어 있다. 히든층의 뉴런 수는 시뮬레이션 과정에서 소요시간과 학습율이 최적화 되는 3개로 결정하였으며, 입력층에 임의의 기본모음에 대한 제1, 제2 포먼트주파수를 입력해 주면 출력층에서는 5개의 단자 중 인식된 기본모음에 해당하는 하나의 단자만 나머지와 다른 응답하게 하였다. 그리고 신경망에서 뉴런의 전달함수 기울기를 결정하는 수식 (1)의 λ 값은 전자회로로 구현시 전자소자(CMOS 인버터)의 전달함수의 값과 비슷한 6으로 결정하였다. 그림 3은 제작된 신경망 모델을 전자회로로 구현한 것으로서 시냅스에 해당하는 각층 간의 연결세기를 저항으로, 뉴런의 몸체에 해당하는 각층은 증폭기로 구성하였다.

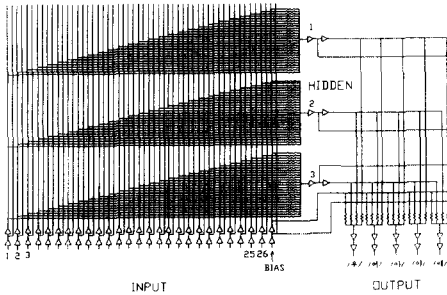


그림 3. 전자회로적 모음인식 신경회로망
Fig. 3. An electronic neural network for Korean vowels recognition.

IV. 신경회로망 집적회로의 제작 및 측정

설계된 전자회로적 신경회로망을 증폭기는 2.0 μ m 디자인 룰을 사용하여 CMOS 인버터로, 저항은 폴리실리콘 저항으로 레이아웃한 다음 CMOS 표준공정을 통하여 집적회로로 제작하였다. 이 과정에서 입력층과 히든층 사이의 전체 병렬저항 R_t 값에 따른 음성

인식 정도의 차이등을 알아보기 위해 이온 주입량을 변화시키는 방법을 사용하여 폴리실리콘의 면저항을 47 $\Omega/\square \sim 7 K\Omega/\square$ 사이에서 8단계로 나누어서 제작함으로써 집적회로 내의 R_t 값을 15.4 $\Omega \sim 5.90 K\Omega$ 으로 변화시켜주었다. 그림 4는 제작된 집적회로의 칩 사진이며, 표 1은 제작된 집적회로의 면저항 및 이에 따른 R_t 값을 나타낸 것이다. 표 2는 제작된 신경회로망 집적회로에 대해 각 모음당 11개씩 총 55개의 제1, 제2 포먼트 데이터를 입력하여 음성인식 정도를 측정한 것으로서 /우/, /아/, /에/, /이/ 음에 대해서는 대체적으로 잘 인식되었으나, /어/ 음에 대해서는 약간 불안한 인식이 이루어졌다. 이는 그림 1의 음운 분포도에 나타난 바와같이 /어/ 음이 /우/ 음과 매우 인접하여 분포되어 정확히 구분된 학습이 이루어지지 못하여 나타나며 차후 좀더 확실히 구분된 학습이 이루어진다면 해결될 것으로 본다. 한편, 전자회로의 소비전력과 집적도 및 음성인식률을 좌우하는 전체 병렬저항 R_t 값은 폴리실리콘의 면저항값이 1.12 $K\Omega/\square$ 인 373.3 Ω 이상의 크기부터 제대로 인식되는 결과가 나타났다. 여기서 R_t 값이 낮을 때 음성인식이 되지않는 것은 증폭기로 사용된 CMOS 인버터의 전류 구동능력이 작아서 낮은 R_t 값에서는 증폭기가 정상적인 동작을 하지 못하기 때문에 나타나는 현상이다. 일반적으로 신경회로망 내에서 R_t 값은 클수록 전류가 작게 흘러 소비전력 면에서는 유리하

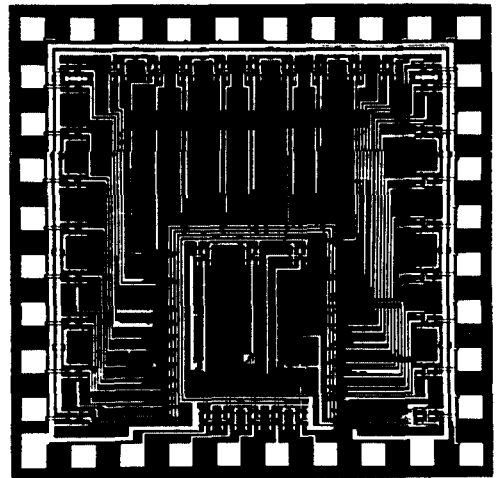


그림 4. 제작된 음성인식 신경회로망 집적회로의 사진
Fig. 4. Photography of the fabricated neural network IC for Korean vowels recognition.

나 큰값을 갖는 폴리실리콘 저항을 만들기 어렵기 때문에 집적도 면에서는 작을수록 유리하다. 따라서 이 신경회로망 집적회로에서는 음성인식 능력과 소비전력, 집적도 등을 동시에 고려할 때 표 2에 나타난 바와같이 R_t 를 1 K Ω ~2 K Ω 정도로 결정해 주면 회로적으로 최적화될 것으로 생각된다. 한편, 앞으로 개발될 좀더 복잡한 신경회로망 시스템은 저항의 갯수가 많아져서 전체 병렬저항값이 상대적으로 낮아질 것으로 예상된다. 이를 집적회로로 실현하기 위해서는 더 낮은 R_t 값에서도 동작이 가능한 CMOS 인버터보다 특성이 더 좋은 전자소자를 사용하는 것이 바람직할 것으로 본다.^[6]

표 1. 폴리실리콘의 면저항과 전체 병렬저항값의 관계

Table 1. Relation of the total parallel resistance vs polysilicon sheet resistance.

Ion-impant. Dose ($\times 10^{14}/\text{cm}^2$)	270	270	86	47	26	14	10	7.4
Polysilicon Sheet Resist (Ω/\square)	47	105	204	447	447	3.54K	7.06K	17.7K
Total Parall. Resistance (Ω)	15.7	35.0	68.0	149	149	1.18K	2.35K	5.90K

표 2. 신경회로망 집적회로의 모음인식 정도.
Table 2. Vowel recognition rates of the fabricated neural network IC.

(11개 음성 중 인식되는 숫자)

$R_t1(\Omega)$	/우/	/어/	/아/	/이/	/에/
15.7	0	0	0	0	0
35.0	0	0	0	0	0
68.0	0	0	0	0	0
149.0	0	0	0	0	0
373.3	7	4	11	11	11
1.18K	10	3	11	11	11
2.35K	10	3	11	11	11
5.90K	10	1	10	11	11

V. 결론

한국어 기본모음 5개를 인식할 수 있는 음성인식

신경회로망 집적회로를 제작한 다음 음성인식 정도를 관찰해 보았다. 이를 위해 입력층과 히든층 및 출력층을 갖는 3층 구조의 신경회로망 모델을 설정하였다. 이 신경회로망을 추출된 한국어 기본모음 5개에 대한 60개의 제1, 제2 포먼트 주파수 데이터 중 40개를 BP 학습 알고리즘을 이용하여 학습시켰으며, 이를 시냅스는 저항으로, 뉴런몸체는 증폭기를 사용하여 전자회로로 설계하였다. 또한 설계된 전자회로적 신경회로망을 증폭기는 CMOS 인버터로, 저항은 폴리실리콘 저항을 사용하여 2.0 μm 디자인 룰로 레이아웃하였으며, CMOS 표준공정을 이용하여 집적회로로 제작한 다음 55개의 음성 데이터를 입력하여 음성인식 정도를 관찰해 보았다. 그 결과 신경회로망 집적회로 내의 전체 병렬저항값이 373.3 Ω 이상일 때부터 일부 음을 제외하고는 음성인식이 양호하게 이루어졌다. 따라서 음성인식 능력과 소비전력, 집적도 등을 고려해 볼 때 신경회로망의 전체 병렬저항은 1 K Ω ~2 K Ω 일 때가 적당할 것으로 본다.

參考文獻

[1] Neural networks : Implementing Associative memory Models in Neurocomputers, R.K. Miller, S.E.A.I. Technical publications, (1987).

[2] R.P.Lippmann, "An introduction to computing with neural nets," IEEE ASSP Magazine, vol.4, pp.97-115, April 1987.

[3] T.Ohotomo, K.Takashai, and K.Hara, "Learning and discrimination of perceptual vowel distribution by a neural net model," vol. J71-D, no.11, pp. 2291-2299, Nov. 1988.

[4] 심재형, 이종혁, 윤태훈, 김재창, 이양성, "신경망을 이용한 모음의 학습 및 인식방법," 대한전자공학회 추계종합학술대회 논문지, 제12권 제2호, pp. 278 - 281, 1989. 11.

[5] 이종혁, 심재형, 윤태훈, 김재창, 이양성, "신경망을 이용한 모음의 인식 및 방법," 전자공학회 논문지, 제 27권, 제 12권, 제 1호, pp. 144-151, 1990

[6] 채상훈, 박용찬, 윤태훈, 김재창, "BiCMOS 인버터를 이용한 뉴런의 전자적 구현," 대한전자공학회 논문지, 제 28권, A편, 제 6호, pp. 43-51, 1991.

著 者 紹 介

蔡 相 勳(正會員) 第 28卷 A編 第 7號 參照

현재 한국전자통신연구소 집적회로
연구실 근무

尹 台 燾(正會員) 第 28卷 B編 第 4號 參照

현재 부산대학교 전자공학과 부교수

•

金 在 昌(正會員) 第 28卷 B編 第 4號 參照

현재 부산대학교 전자공학과 교수