

# Digital 신경회로망을 위한 비선형함수의 구현

김 진태\*, 정 덕진  
인하대학교 전자재료공학과

Design of Nonlinear(Sigmoid) Activation Function for Digital Neural Network

Jintae Kim\*, Duckjin Chung  
Department of Electronic Materials & Devices Engineering  
Inha University

**Abstract** - A circuit of sigmoid function for neural network is designed by using Piecewise Linear (PWL) method. The slope of sigmoid function can be adjusted to 2 and 0.25. Also the circuit presents both sigmoid function and its differential form. The circuits is simulated by using ViewLogic. Theoretical and simulated performance agree with 1.8 percent.

## 1. 서론

신경회로망에서 각 뉴런의 출력특성으로 비선형함수가 사용되고있다. 이들 비선형함수에는 하이퍼볼릭 탄젠트(Hypobolic Tangent, tanh), Hard Limiter, Pseudolinear 함수와 시그모이드(Sigmoid)함수등이 있으나 다층 퍼셉트론(Multilayer Perceptron) 모델에서는 시그모이드 함수가 자주 사용되고 있다[1].

시그모이드 함수의 표현식과 그 미분의 형태는 다음과 같다.

$$y = \frac{1}{1 + e^{-kx}}$$
$$y' = ky(1 - y)$$

이와 같은 시그모이드 함수를 구현하는 방법에는 디지털에 의한 방법과 아날로그에 의한 방법이 있다. 아날로그 방식은 각각의 뉴런의 동작에 따른 비선형성에 대한 회로의 구성이 용의하나 drift와 잡음에 대한 문제로 인하여 비선형함수의 환경을 조절하기가 힘들다[2].

반면에 디지털방식은 drift에 무관하고 잡음에 대해 강한 저항력을 보이지만 이를 하드웨어로 구현하는 방법에는 여러가지 문제점이있다[2]. 이의 구현 방법에는 첫째로 테일러 전개 (Taylor Series)를 이용하여 함수를 네개내지 다섯개의 항으로 분류하고 이를 계산하는 방법으로 구현하는데 많은 비용이 드는 단점이 있다[3]. 두번째 방법은 RAM이나 ROM과 같은 메모리를 사용하여 입력값에 대한 출력값을 미리 저장시켜 놓고 이를 사용하는 예측표(Look-Up Table) 방법으로 입력과 출력값이 16비트인 경우 128k바이트의 메모리 용량이 필요하게 됨으로써 비현실적이다[4]. 세번째 방법은 Piecewise Linear (PWL)간략화를 이용하는 방법으로 이 방법은 입력영역을 임의의 영역으로 구분하고 이들 각 영역에서의 입력력간계를 연결시켜 출력하는데 적은비용으로 구현이 가능하면서 함수의 특성

을 유사하게 표현할 수 있다는 장점을 지니고 있다[4,5].

본 논문에서는 구현방법과 비용면에서 우수한 PWL간략화 구조를 사용하고 신경회로망의 학습속도에 큰 영향을 미치는 함수의 기울기의 변화에 대해 제한적으로 대응할수 있는 방법을 제시하고 이를 회로로 구현하며 이를 토대로 결론을 맺는다.

## 2. PWL 간략화

Power of Two PWL간략화를 사용하여 활성화 함수를 구현하는 방법에는 두가지 방법이 있다. 그중 하나는 Myers 와 Hutchinson에의해 제안된 7세그먼트 간략화이며 [4] 다른 하나는 Alippi 와 Storti-Gajani 가 제안한 15세그먼트에 의한 방법이다[5].

Myers 와 Hutchinson의 간략화는 펄스 부호화 방법(Pulse Coded Modulation, PCM)에 의해 함수를 구성하였으나 함수의 미분형에 대해서는 표현하지 못하였고 이를 회로를 구성하여 설계하였다. Alippi와 Storti-Gajani는 계단(step)간략화와 Power of Two 승산기에의한 PWL간략화를 제안하고 승산기를 사용하여 함수의 미분값을 계산하고 이를 시뮬레이션을 사용하여 학습을 시켰다[5].

Murtagh와 Tsoi는 위의 두 논문으로 부터 시그모이드 함수를 구현하는데 있어 가장 적절한 모델을 도출해 내었는데 여기서 나타난 결과는 15세그먼트 PWL간략화를 사용하여 활성화 함수를 구현하고 15세그먼트 계단 간략화를 사용하여 미분함수를 구현하는 것이 회로망의 수렴과 학습에 가장 적절하다는 것이었다[2].

## 3. 본 논문에서의 제안된 방법

본 논문에서는 Murtagh와 Tsoi의 결과에 기초를 두되 함수의 미분값을 표현하는데 있어서도 15세그먼트 PWL 간략화를 사용하였으며 한정된 기울기의 변화에 대응할 수 있도록 회로를 구성하였다. 이를 위하여 각각의 기울기에 대한 구분을 위하여 1비트 이상이 입력값에 추가하였다. 이와 같이 임의의 비트를 입력값에 추가함으로써 Power of Two의 비로 여러가지 기울기를 표현할 수 있으며 이들을 디코딩하여 각각의 기울기에대한 출력값을 구할수있다. 이 방법을 사용하여 시그모이드 함수값과 미분값을 차례로 구하는 회로를 구성하고 모의실험을 하여 결과를 나타낸다.

표1부터 표4까지는 본 논문에서 정의한 입력값 영인 지점에서의 기울기가 0.25인 함수와 기울기 2인 함수의 각

영역에서의 함수값과 미분값의 입출력관계를 나타내었다.

표1. 기율기 0.25인 함수의 함수값 입출력 관계

입력	출력
0.XYZabcdefgijkl	0.10XYZabcdefgij

표2. 기율기 0.25인 함수의 미분값의 입출력 관계

입력	출력
0.XYZabcdefgijkl	0.0011X'Y'Z'a'b'c'd'e'f'g'h'i'

표3. 기율기 2인 함수의 함수값 입출력 관계

영역	입력	출력
1	0.000abcdefgijkl	0.10abcdefgijkl1
2	0.001abcdefgijkl	0.110abcdefgijkl
3	0.010abcdefgijkl	0.1110abcdefgijjk
4	0.011abcdefgijkl	0.11110abcdefgij
5	0.100abcdefgijkl	0.111110abcdefghi
6	0.101abcdefgijkl	0.1111110abcdefgh
7	0.110abcdefgijkl	0.11111110abcdefg
8	0.111abcdefgijki	0.11111111abcdefg

표4. 기율기 2인 함수의 미분값 입출력 관계

영역	입력	출력
1	0.000abcdefgijkl	1.1a'b'c'd'e'f'g'h'i'j'k'l'l'
2	0.001abcdefgijkl	1.0 a'b'c'd'e'f'g'h'i'j'k'l'
3	0.010abcdefgijkl	0.(a'b')1a'b'c'd'e'f'g'h'i'j'k'
4	0.011abcdefgijkl	0.001a'b'c'd'e'f'g'h'i'j'k'
5	0.100abcdefgijkl	0.0001a'b'c'd'e'f'g'h'i'j'
6	0.101abcdefgijkl	0.00001a'b'c'd'e'f'g'h'i'
7	0.110abcdefgijkl	0.0000001a'b'c'd'e'f'g'
8	0.111abcdefgijki	0.0000000a'1b'c'111

위의 표에서 볼 수 있듯이 기율기 0.25인 함수는 기율기의 변화가 거의 없기때문에 전체 입력영역에 대해서 영역분할을 하지않았다. (Myers와 Alippi의 논문에서는 입력값의 범위가 -8에서 8까지였다) 그러나 기율기가 2인 함수의 경우에는 전체 입력영역 내에서의 함수값의 변화가 크게 일어나기 때문에 각각의 영역에 대한 입출력 상관관계가 다르게 주어졌다.

이들 함수에 대한 입출력관계의 결정은 각각의 기율기에 대한 각 영역에서의 함수값의 변화를 파악하고 이로부터 실제값과 결과값과의 차이를 최소화할 수 있는 방법이 선택되어야 한다. 표5부터 표8까지는 본 논문에서 정의한 입출력관계로부터 각 경계영역에서의 실제값과의 차이를 나타낸다.

표5. 기율기 0.25인 함수와 회로의 함수값 비교

입력범위	시그모이드 계산	회로 출력
0-0.999755	0.5 - 0.7310105	0.5 - 0.7499389

표6. 기율기 0.25인 함수와 회로의 기율기 비교

입력범위	미분값 계산	회로 출력
0-0.999755	0.25 - 0.1966341	0.25 - 0.1875

표7. 기율기 2인 함수와 회로의 함수값 비교

입력범위	시그모이드 계산	회로 출력
0-0.125	0.5 - 0.731059	0.5 - 0.75
0.125-0.25	0.731059 - 0.880797	0.75 - 0.875
0.25-0.375	0.880797 - 0.952574	0.875 - 0.9375
0.375-0.5	0.952574 - 0.982014	0.9375 - 0.96875
0.5-0.625	0.982014 - 0.993307	0.96875 - 0.984375
0.625-0.75	0.993307 - 0.997527	0.984375 - 0.992187
0.75-0.875	0.997527 - 0.999089	0.992187 - 0.996094
0.875-1.0	0.999089 - 0.999665	0.996094 - 1

표8. 기율기 2인 함수와 회로의 기율기 비교

입력범위	미분값 계산	회로 출력
0-0.125	2.0 - 1.572895	2.0 - 1.5
0.125-0.25	1.572895 - 0.839949	1.5 - 1.0
0.25-0.375	0.839949 - 0.361413	1.0 - 0.25
0.375-0.5	0.361413 - 0.141302	0.25 - 0.125
0.5-0.625	0.141302 - 0.053184	0.125 - 0.0625
0.625-0.75	0.053184 - 0.019732	0.0625 - 0.03125
0.75-0.875	0.019732 - 0.007282	0.015625 - 0.007812
0.875-1.0	0.007282 - 0.002682	0.007812 - 0.002197

그림1은 회로도의 블럭다이어그램이다. 이 그림에서 디코더는 입력값으로 부터 임의의 기율기를 판단하고 기율기의 각 입력에 대한 실제 입출력값의 상관관계로 부터 출력값을 결정하는 일을 하는 부분이며 19비트 레지스터는 디코더의 값으로 부터 입력값을 자리가동시킬 수 있도록 직병렬입력-병렬출력, 적절 자리가동이 가능하여야 한다. 15비트 레지스터는 출력값을 저장한다.

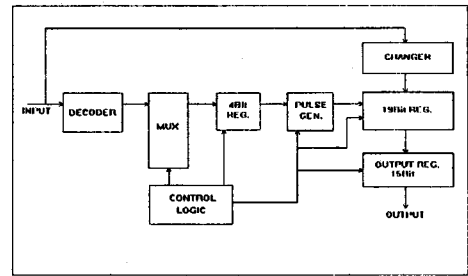


그림1.회로도의 블럭다이어그램

#### 4. 모의실험 결과

본 논문에서는 앞에서 언급한 두 논문[2,4]에서 기율기 값으로 사용한 입력값 영역에서의 기율기가 0.25인 함수와 기율기 2인 함수에 대하여 연구를 수행하였으며 기율기가 0.25인 함수에 대해서는 입력값의 범위가 -1에서 1까지의 범위내에서 출력값의 변화가 비선형함수라기 보다는 선형함수에 더 가깝기 때문에 이를 임의의 영역으로 분할하지 않았다.

그림2는 위에서 제안한 방법을 사용하여 회로를 모의실험한 결과이다. 이 그림2로 부터 위에서 제시한 입출력 상관관계와 일치함을 알수있다. 그림3은 실제 함수값과 모의실험 결과를 나타낸 그림으로 모의실험한 결과가 실제 함수와 아주 유사함을 나타내며 오차의 발생은 입력값 0.125에서 출력값의 차이가 0.018로 가장크게 나타났다. 그림4는 함수의 기율기의 변화를 나타낸다.

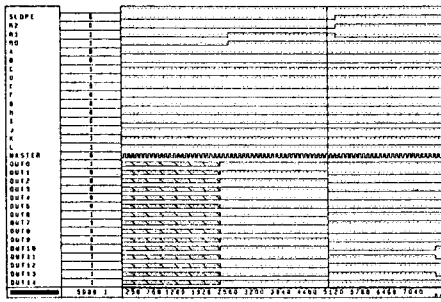


그림2. 구현된 회로의 모의실험 결과

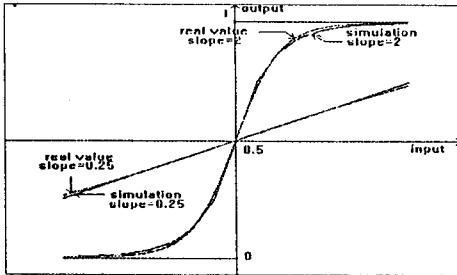


그림3. 실제값과 모의실험 결과와의 함수값 출력특성

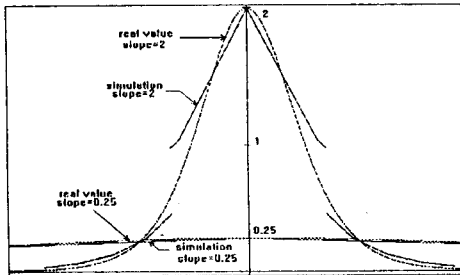


그림4. 실제값과 모의실험 결과와의 기울기 출력 특성

## 5. 결론

본 연구에서는 시그모이드 함수를 디지털방법을 사용하여 구성하는데 있어 임의의 기울기의 변화에 대응할 수 있도록 회로를 구성하는데 초점을 맞추었다.

위의 회로는 각 비트의 추가시 마다 Power of Two의 비율로 서로다른 기울기의 종류를 표현할 수 있을 뿐 아니라 회로의 구성시에도 각 기울기에 대한 디코딩기 제어회로만을 수정함으로써 쉽게 다시 회로를 구성할수있다.

그러나 PWL 간략화를 이용하는데 있어 가장 중요한 과정은 입출력 상관관계를 정의하는 과정으로 출력값의 연속성과 오차의 감소에 큰 영향을 미칠것이기 때문이다.

앞으로는 기울기의 변화를 4가지로 늘리고 입출력 상관관계를 좀더 자세히 정의하여 출력값의 연속성과 정확성을 개선한 회로를 구성할 계획이다.

## 6. 참고문헌

1. P.Murtagh and A.C.Tsoi, 'Implementation Issue of Sigmoid Function and Its Derivative for VLSI Digital Neural Network', IEE Proc. Vol.139, No.3, pp.207-214, 1992

2. D.J.Myers and R.A.Hutchinson, 'Efficient Implementation of Piecewise Linear Activation Function for Digital VLSI Neural Network', Electronics Letter, 23rd Vol.25, No.24, pp.1663-1664, 1989

3. J.Hertz, A.Krogh and R.G.Palmer, 'Introduction to the Theory of Neural Computation', (Addison-Wesley,1991)

4. C.Alippi and G.Storti-Hajani, 'Simple Approximation of Sigmoid Functions:realistic design of digital neural networks capable of learning', Proc. IEEE Int. Symp. Circuit and System, pp.1505-1508, 1991

5. J.Ouali and G.Gaucier, 'Fast Generation of Neuro - ASIC's', Proc. Int. Neural Networks Conf., pp.563-567, 1990