

향상된 선형 신경 회로망 합성 방법

An Enhanced Method for Linear Binary Neural Network Synthesis

박 병 준, 이 정 훈

Computational Vision and Fuzzy Systems Laboratory

한양대학교 전자공학과

425-791 경기도 안산시 사 1 동 1271 번지

{bjpark,frhee}@fuzzy.hanyang.ac.kr

Abstract

본 논문에서는 선형 이진 신경회로망 (Linear Binary Neural Network) 을 최소화 하기 위하여, 입력 패턴의 그룹화 가능성을 측정하는 조건함수를 제시한다. 또한 이 조건식으로 그룹화 우선순위를 정하고 iteration 을 통해 신경회로망을 합성하는 MSP Term Grouping Algorithm 을 보인다. 여러 가지 예제에 대한 실제적 합성 실험을 통해 기존의 알고리즘과 제시된 알고리즘을 비교한 결과는 제시된 알고리즘이 기존의 알고리즘 보다 작은 크기의 선형 이진 신경회로망을 합성할 수 있는 향상된 방법임을 보여준다.

I. 서론

초창기의 신경회로망의 각 단위 뉴런 (Linear Threshold Unit) 들은 선형 가중 합 (weighted sum) 을 기본으로 한 퍼셉트론 (Perceptron) 이었다[1]. 이 퍼셉트론은 한 개의 소자를 가지고 AND, OR, NOT 과 같은 이진 논리를 가중치 (weight) 의 변경만으로 모두 표현할 수 있는 장점이 있기 때문에 하드웨어로 구현함에 있어서도 적은 개수의 소자로 빠른 처리 속도를 낼 수 있는 장점을 가지고 있다[2].

그러나 이러한 효율성에도 불구하고 회로망 합성과 가중치 결정에 대한 적절한 방법을 찾지 못해 실제로 사용되는 영역이 한정되어 있었다. 최근에는 기하학적 탐색기법을 이용한 Expand-and-Truncate Learning (ETL) 방법[3] 이나 Boolean-Like-Training Algorithm (BLTA) [4] 등의 알고리즘이 제시되었으나, 실제로 이러한 방법으로 합성된 선형 신경 회로망은 주어진 문제에 대하여 가장 최적의 결과라고 보기는 어렵다. 예를 들면, ETL 방법은 향상된 방법이라 할 수 있지만, 이 또한 주어진 인수 (parameter) 에 따라 알고리즘 시작점 (initial seed vertex) 가 달라지고 이에 따라 회로망 합성 결과가 달라지는 등의 한계를 가지고 있다[3].

본 논문에서는 디지털 논리 회로 합성에서 최소화된 2 단계 논리를 표현하는 한 가지 형태인 Minimal Sum of Product (MSP) 의 형태를 이

용하여 주어진 문제를 MSP 형태로 바꾸고, 이 MSP 형태의 각 항을 다시 몇 개로 묶어줌으로써 (grouping) 최소화된 회로를 합성해 낸다. 결과적으로 2진 공간을 두 개의 분할 공간으로 나누는 최소화된 3층 신경회로망을 합성하는 새로운 알고리즘을 제시한다.

기존에 제시된 방법에 비해 본 논문에서 제시한 MSP Term Grouping Algorithm 의 장점은 다음과 같다.

- 1) MSP 형태를 사용하므로 학습에 사용하는 입력 패턴의 수를 대폭 줄임으로써 학습 시간을 단축시킬 수 있다.
- 2) 학습을 시작하는데 초기 조건 설정이 필요치 않다.
- 3) 입력패턴의 순서에 상관없이 항상 같은 회로망이 합성된다.

II. 본론

2.1 선형 신경회로망 단위 뉴런의 구조

먼저 N 차원의 이진 공간을 생각해 보면 각 패턴 $x=(0,1)^N$ 은 N bit 의 이진수로 나타낼 수 있으며, 또 이것은 N 차원의 hypercube 의 vertex 로 생각할 수 있다. 선형 신경 회로망의 기본 단위인 Linear Threshold Unit(LTU)는 바로 N 개의 이진 입력을 받아 그것의 선형 가중합을 구하여, 그 결과가 임계치 보다 큰 영역과 임계치 보다 작은 영역, 이 두개의 영역으로 나누는 역할을 한다. 즉 N 차원의 이진 공간은 두개의 분할 공간으로 나뉘게 된다. 이것의 그림과 식은 다음과 같다.

Acknowledgement

이 논문은 1995년도 한국학술진흥재단의 대학부설연구소 연구과제 연구비에 의하여 연구 되었음.

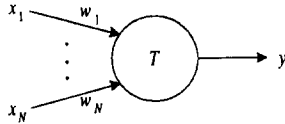


그림 1. Linear Threshold Unit 의 구조.

$$y = 1 \text{ if and only if } \sum_{i=1}^N w_i x_i - T \geq 0$$

$$y = 0 \text{ if and only if } \sum_{i=1}^N w_i x_i - T < 0 \quad (1)$$

이러한 공간에서 선형 가중합과 임계치가 서로 갈아지는 점들의 합수는 hyperplane 을 형성한다. 여기서 최종적인 목표는 임의의 주어진 이진 분할 공간을 최소의 hyperplane 을 가지고, 합성해 내는 것이므로 각 개개의 hyperplane 이 최대한의 분할을 갖도록 해야 할 필요가 있고, 이것을 찾기 위해 이미 이전의 여러 연구들에서 이 hyperplane 하나를 통해 나눌 수 있는 분할의 몇몇 충분조건과 필요 조건들이 밝혀진 바 있다[3][5]. 이중 본 논문에서 고려되어진 중요한 필요조건 중의 하나가 unate 특성이다.

Unate 특성은 주어진 이진함수를 Minimal Sum of Product 의 형태로 표현하였을 때, 어떠한 입력 변수도 positive 형태와 negative 형태가 동시에 존재해서는 한 개의 hyperplane 으로 분할할 수 없는 특성이다. 예를 들어 3 변수 2 진 함수 $f(x_1, x_2, x_3) = x_1 x_2' + x_2' x_3 + x_1 x_3$ 가 있다면 이 때는 모든 변수가 positive 이거나 negative 한가지의 형태로만 나타나 있다. 그러나, $g(x_1, x_2, x_3) = x_1 x_2' + x_2 x_3'$ 를 보면, x_2 와 x_2' 가 한 함수에서 동시에 나타난다. 이 때, $g(x_1, x_2, x_3)$ 는 Unate 특성을 만족시키지 못하고 이것은 한 개의 hyperplane 으로 분할할 수 없다. 그러나 이 성질은 필요 조건일 뿐 Unate 성질을 만족한다고 해서 항상 분할이 만족되는 것은 아니다[5][6].

이외에 PSP 형태[7] 등의 몇 가지 충분 조건들이 존재하지만 이것들은 충분조건일 뿐, 이 조건들을 만족하지 않고도 분할 가능한 것들이 많이 존재한다. 그러므로 보다 향상된 그룹화 조건이 필요하다.

2.2 가중치와 임계치의 결정

본 논문에서는 최소화된 그룹의 각 단위 뉴런에 대해서 가중치와 임계치를 구하기 위하여 몇 가지 제시된 방법 중 Reference HyperSphere (RHS) 방법을 사용하였다[8]. 이것은 두 개의 분할로 나누는 hyperplane 은 전체 패턴을 포함하는 hypersphere 와 true 패턴만을 포함하는 hypersphere 간의 접선을 지나는 평면과 평행하게 되는 원리를 이용하여 가중치를 구하고 구해진 가중치로부터 Minimum True 와 Maximum False 에 해당하는 패턴들의 가중합 사이의 중간 값으로 임계치를 정한다. 이 방법에 대한

자세한 사항은 [8]을 참조하면 된다.

2.3 최소화된 선형 신경회로망 합성

최소화된 선형 신경회로망을 합성하기 위하여 전체 함수를 한 개의 hyperplane 으로 분할 가능한 몇 개의 조각으로 나누어야 한다. 이 때 최소한의 조각으로 나누어야만 가장 빠른 계산 결과와 또 하드웨어로 구현하였을 때 최소한의 공간으로 만들 수 있다. 이를 위해 디지털 회로 설계에서 함수를 최소화하는 한가지 방법으로 알려져 있는 Minimal Sum of Product 의 형태로 전체 함수를 표시하고, 이것들의 각 항을 최소한의 그룹으로 나누기 위하여 그룹화를 시도한다. 또한 그룹화를 위해 탐색 우선순위를 정하는 함수를 제시함으로써 최소화된 선형 이진 신경회로망을 보다 빨리 합성할 수 있다.

A. MSP 항목 그룹화 조건

가장 적은 개수의 그룹으로 전체 함수를 표현하기 위해서는 한 개의 그룹에 최대한 많은 수의 MSP 항이 포함되도록 해야 한다. 이를 위해서는 모든 항의 조합을 탐색해 보는 것이 가장 좋은 결과를 낼 수 있겠지만, 그 복잡도는 패턴공간의 차원이 증가하면서 지수적으로 증가하기 때문에 실제 문제에 적용하기에는 많은 어려움이 있다. 여기서 이러한 전체 탐색을 거치지 않고도 한번 그룹화가 일어난 후에 또 계속해서 그룹화가 이뤄질 수 있는 조건이 두 개의 그룹 사이에서 서로 다른 변수의 개수에 의존함을 알 수 있다[7].

만약 A 와 B 두개의 항(Term) 그룹이 있을 때, 그 그룹의 변수들의 집합을 A'와 B'라고 정의하면 편차함수 d는 다음과 같이 정의된다.

$$d(A, B) = N \min\{|A'-B'|, |B'-A'|\} + \max\{|A'-B'|, |B'-A'|\} \quad (2)$$

이때 N 은 전체 함수의 변수의 개수이다. 그러므로 이 편차함수는 두 그룹의 변수집합들의 차집합의 원소의 개수 중 최소값에 보다는 존함을 알 수 있다. 예를 들어 함수 $f(x_1, \dots, x_5) = x_1 x_2' x_3 + x_1 x_2' x_4 x_5$ 이고 $A = x_1 x_2' x_3$, $B = x_1 x_2' x_4 x_5$ 라고 하면 $A' = \{x_1, x_2, x_3\}$, $B' = \{x_1, x_2, x_4, x_5\}$ 가 되고, $|A'-B'| = 1$, $|B'-A'| = 2$ 가 되며, $N = 5$ 이다. 그러므로 전체 $d(A, B) = 7$ 이 됨을 알 수 있다.

이렇게 구해진 편차 함수의 값이 가장 작은 쌍부터 그룹화를 하게 되면, 한번 그룹화에 참여했던 항들이 또 다시 그룹화할 수 있는 정도가 커지고, 결국 한 개의 그룹에 가장 많은 수의 항을 포함 시킬 수 있게 된다.

B. MSP 항목 그룹화 알고리즘

앞에서 주어진 편차함수를 이용하여 각 쌍의 그룹화 우선순위를 정한 그것들을 직접 그룹화 해주는 알고리즘은 다음과 같다.

MSP term Grouping Algorithm

Obtain the initial LTU set where its elements are terms of the MSP function;

REPEAT

FOR all existing set element pairs that are unate and available for grouping **DO**

Obtain pair that has the smallest d value;

END FOR

Apply the RHS method to the selected pair; (i.e., to check if pair is linear separable)

IF the pair can be grouped into a single threshold logic unit **THEN**

Partition the LTU set which include the pair; (i.e., the partition becomes an element of the LTU set)

ELSE

Set pair is not available for grouping;

ENDIF

UNTIL(there exists no available element pairs for grouping)

FOR all LTU set elements **DO**

Obtain a linear threshold unit using the RHS method;

END FOR

여기서 LTU set 은 전체 함수를 표현하는 그룹들의 집합이며, 초기의 LTU set 은 각 모든 MSP 항들을 하나의 그룹으로 정해줌으로써 만들어 진다. 각 그룹들은 앞에서 설명한 d 편차 함수가 가장 최소가 되는 쌍부터 그룹화를 하여 더 이상 그룹화를 할 수 없을 때까지 iteration 을 수행한다. 본 논문에서는 MSP 의 형태를 이용하였기 때문에 합성된 신경회로망의 출력단 소자는 모든 가중치와 임계치가 1 이 되는 OR 동작을 하게 된다. 이것은 실제 하드웨어로 구현할 때 단순한 Wired Logic 으로 표현이 가능하여 더욱 작은 공간에 하드웨어 구현을 할 수 있게 된다.

2.4 실험 결과

여기서는 본 논문에서 제시한 알고리즘과 기존의 알고리즘을 비교하기 위한 실제적 실험을 보이고 있다.

A. 양자화된 원의 분할

첫번째 예제는 [3][4]에서 사용되었던 6 bit 으로 양자화된 2 차원 원 문제를 보였다. 다음의 그림 2 에서도 알 수 있듯이 2 차원의 각 축을 3 bit 을 사용해 각각 양자화하면 전체 공간이 64 개의 사각형으로 구분되며, 이 가운데 중앙의 12 개에 대해서는 1 의 기대값을 할당하고 나머지 52 개의 사각형에 대해서는 0 의 기대값을 주었을 때 이를 구분하는 신경회로망을 합성하는 문제이다.

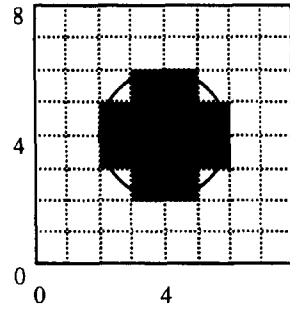


그림 2. 6 bit 으로 양자화된 원공간.

이 문제를 MSP 형태로 표시하면 8 개의 항을 갖는 형태가 되고 이를 다시 그룹화하여 4 개의 단위 뉴런을 갖는 신경회로망을 합성하였다. 결과적으로 합성된 신경회로망의 가중치 표와 회로망 그림을 아래에 보여 준다. 기존 알고리즘과 비교를 하면 ETL [3]방법에 의한 합성 결과는 뉴런이 5 개 필요하였고, BLTA[4]방법은 16 개의 뉴런이 필요하였다.

표 1. 6 bit 로 양자화된 원 공간의 분할을 위한 신경회로망의 가중치와 임계치.

unit	weights						T
1	-3	3	1	-3	3	1	6.5
2	-3	3	1	3	-3	-1	5.5
3	3	-3	-1	-3	3	1	5.5
4	3	-3	-1	3	-3	-1	4.5

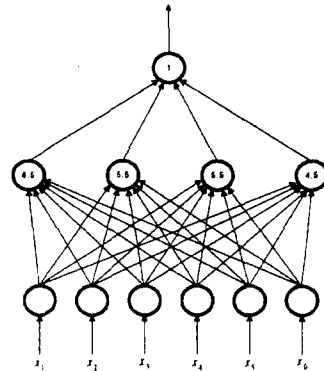


그림 3. MSP Term 그룹화 알고리즘으로 합성된 6 bit 로 양자화된 원공간의 분할의 신경회로망.

B. 네 개의 클래스를 가진 영상 문제

두 번째 예제는 그림 4 에서 보여주는 영상 분할문제이다. 그림에서는 32×32 의 해상도를 가지는 grayscale 영상이 각 클래스에 256 개의 점들을 포함하는 4 개의 클래스로 구성되어 있고, 각 축은 5 bit 로 양자화를 하였다니 전체 문제는 4 개의 출력과 10 개의 입력을 가지는 문제가 되었다. 이 문제를 각각의 클래스에 대해서 MSP 의 형식으로 표현하면 각 클래스에 대해서 5,7,7,5 개의 항을 가지는 함수가 되고 이를

다시 본 논문에서 제시된 그룹화 알고리즘에 적용하면, 2,3,3,2의 은닉층 뉴런 수를 가지는 신경회로망으로 합성이 된다. 표 2와 그림 5는 제시된 방법으로 합성한 신경회로망의 가중치와 임계치 그리고 회로망 그림을 보여준다.

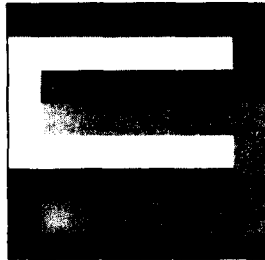


그림 4. 네 개의 클래스를 가진 grayscale 영상.

표 2. 합성된 신경회로망의 가중치와 임계치.

class	weights									T	
1	-4	-4	-4	0	0	60	4	60	0	0	110
	-8	-8	-8	0	0	8	8	0	0	0	12
2	4	4	4	0	0	-28	28	28	0	0	58
	4	4	4	0	0	28	28	-28	0	0	58
	8	8	8	0	0	8	-8	0	0	0	28
	-8	-8	-8	0	0	-8	8	0	0	0	4
3	-4	-4	-4	0	0	-28	-28	28	0	0	18
	-4	-4	-4	0	0	28	-28	-28	0	0	22
4	4	4	4	0	0	-60	-4	-60	0	0	-2
	8	8	8	0	0	-8	-8	0	0	0	20

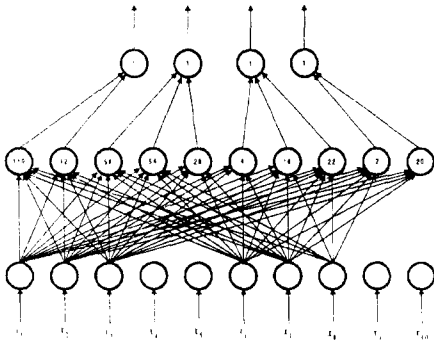


그림 5. 제시된 알고리즘으로 합성한 4개의 클래스를 가진 영상에 대한 회로망.

실험 결과 표와 그림에서 x_4, x_5 와 x_9, x_{10} 에 대한 모든 은닉층 뉴런으로의 가중치가 0이므로, 이 문제는 원래 각 축을 3bit으로 양자화 하여도 같은 결과를 얻을 수 있는 문제임을 알 수 있다. 이렇듯 본 논문에서 제시한 알고리즘은 redundancy를 밝힐 수 있어 적절한 수준의 양자화 정도까지도 알아 낼 수 있는 효율적인 알고리즘이다.

III. 결론

본 논문은 MSP 형태를 이용한 그룹화 알고리즘으로 선형 신경회로망을 최소화하는 새로운 방법을 제시하였다. 실험을 통해서도 기존의 알고리즘[3][4]보다 향상된 합성결과를 보여주고 있으며, Redundancy를 찾아낼 수 있는 진보된 방법임을 알 수 있다.

보다 향상된 신경회로망 합성을 위해 실수공간의 이진변환을 위한 적절한 양자화 방법에 관한 연구와 다중 출력함수를 표현하기 위한 Multiple Prime Implicant를 이용한 다중 출력 신경회로망 합성등에 대한 연구를 계속 하고 있다.

IV. 참고문헌

- [1] D. E. Rumelhart, J. M. McClelland, and the PDP Research Group, *Parallel Distributed Processing*, Cambridge, MA: MIT Press, 1986.
- [2] J. H. Kim, S. K. Park, Y. Han, H. Oh, and M. S. Han, "Efficient VLSI implementation of a 3-layer threshold network," *Proc. ICNN'97*, Houston, TX, vol. 2, pp. 888-893, Jun. 1997.
- [3] J. H. Kim and S. K. Park, "The geometrical learning of binary neural networks," *IEEE Trans. Neural Networks*, vol. 6, no. 1, pp.237-247, Jan. 1995.
- [4] D. L. Gray, and A. N. Michel, "A training algorithm for binary feedforward neural networks," *IEEE Trans. Neural Networks*, vol. 3, no. 2, pp. 176-194, Mar. 1992.
- [5] Z. Kohavi, *Switching and Finite Automata Theory*: 2nd ed., McGraw-HILL, 1970.
- [6] E. J. McCluskey, *Logic Design Principles*, Englewood Cliffs, NJ: Prentice Hall, 1986.
- [7] S. K. Park, J. H. Kim, and B. Ham, "Geometric analysis of feedforward networks for pattern classification problems," *Proc. IFSA'93*, vol. 1, pp. 131-134, 1993.
- [8] S. K. Park and J. H. Kim, "Geometrical learning algorithm for multilayer neural networks in a binary field," *IEEE Trans. Computers*, vol. 42, no. 8, pp. 988-992, Aug. 1993.