2001년도 대한전기학회/대한전자공학회 시스템및제어분야 합동 추계학술대회 논문집 (2001.11.24)

NETLA를 이용한 이진 공간내의 패턴분류

성 상규* 박 두환* 정 종원* 이 준탁**
*동아대학교 전기공학과
**동아대학교 전기 전자 컴퓨터 공학부

Binary Neural Network in Binary Space using NETLA

Sang-Kyu Sung* Doo-Hwan Park* Jong-Won Jeong* Joo-Tark Lee**
*Department of Electrical Engineering, Dong-A University
**School of Electrical Electronics & Computer Engineering, Dong-A University

Abstract - 단층 퍼셉트론이 처음 개발되었을 때, 간단한 패턴을 인식하는 학습 기능을 가지고 있기 장점 때문에 학자들의 관심을 끌었다. 단층 퍼셉트론은 한 개의 소자를 이용해서 이진 논리를 가중치(weight)의 변경만으로 모두 표현할 수 있는 장점 때문에 영상처리, 패턴인식, 장면인식 등에 이용되어 왔다. 최근에, 역전파 학습(Back-Propagation Learning)알고리즘이 이진 공간내의 매핑 문제에 적용되고 있다. 그러나, 역전파 학습 알고리즘은 연속공간 내에서 긴 학습시간과 비효율적인 수행의 문제를 가지고 있다. 일반적으로 역전파 학습 알고리즘은 간단한 이진 공간에서 매핑하기 위해서 많은 반복과정을 요구한다. 역전파 학습 알고리즘에서는 은닉층의 뉴런의 수는 주어진 문제를 해결하기 위해서 우선순위(prior)를 알지 못하기 때문에 입력층과 출력층내의 뉴런의 수에 의존한다. 따라서, 3층 신경회로망의 적용에 있어 가장 중요한 문제중의 하나는 은닉층내의 필요한 뉴런수를 결정하는 것이고, 회로망 합성과 가중치 결정에 대한 적절한 방법을 찾지 못해 실제로 그 사용 영역이 한정되어 있었다.

본 논문에서는 패턴 분류를 위한 새로운 학습방법을 제시한다. 훈련입력의 기하학적인 분석에 기반을 둔 이진 신경회로망내의 은닉층내의 뉴런의 수를 자동적으로 결정할 수 있는 NETLA(Newly Expand and Truncate Learning Algorithm)라 불리우는 기하학적 학습알고리즘을 제시하고, 시뮬레이션을 통하여, 제안한 알고리즘의 우수성을 증명한다..

1. 서 론

기존의 이진 신경회로망의 학습방법으로는 MTGA(MSP Term Grouping Algorithm)방법이 있다.
MTGA는 디지털 논리 합성방법 중의 하나인 MSP(Minimal Sum of Product)형태와 필요조건인 Unate의 특성을 이용하여 합성하는 알고리즘이다. 디지털 논리 합성방법 중의 하나인 MSP(Minimal Sum of Product)형태로 표현하고, 필요조건인 Unate의 특성을 적용하여 Unate의 특성을 만족하는 항들만을 합성하는 알고리즘이다.
따라서, Unate의 특성을 적용하기 위해서 true 패턴들을 이진 함수의 표현 형태 중 MSP의 형태로 나타내어야 하고, Unate의 특성이 만족되지 못할 때에는 합성이 불가능한 단점을 지니고 있다.
본 논문에서 제안한 NETLA를 이용하여, 기존의 패턴 분류인식 학습 알고리즘과 비교 및 실험을 통해 본 논문에서 제안하는 NETLA의 우수성을 증명하였다.

2. 본 론

2.1 이진 신경회로망의 기본 구조

먼저 n-차원의 이진 공간을 생각해 보면 각 패턴 $x_i = \{0, 1\}$ 은 n-비트의 이진수로 나타낼 수 있으며, n-차원의 hypercube의 vertex로 가정할 수 있다. 여기 vertex는 아래의 net함수로 표현되어지는 (n-1) 차원의 hyperplane에 의해 두 클래스로 분리할 수 있다고 가정한다.
Net함수는

$$net(x,\ T) = w_1x_1 + w_2x_2 + \cdots + w_nx_n - T = 0 \qquad (1)$$

여기서, w , T : 상수
위 식을 만족할 경우 LS(Linearly Separable) 즉, 선형적으로 분리 가능하다고 하고, (n-1)차원의 hyperplane은 hard-limiter 활성함수를 가진 n-입력 뉴런에 의해 성립된다.

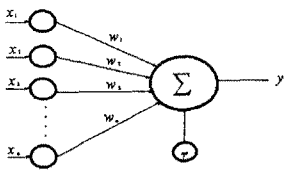

그림 1. 이진 신경회로망의 단위구조

$$y = \begin{cases} 1 : \sum_{i=1}^{n} w_i x_i - T \geq 0 \\ 0 : otherwise \end{cases} \qquad (2)$$

여기서, y : 뉴런의 출력
w_i : i 번째 입력과 뉴런 사이의 연결 가중치
x_i : i 번째 입력
T : 뉴런의 임계치

주어진 2진 매핑 함수가 선형분리 가능한 함수라면, 함수는 단지 1개의 뉴런에 의해 정의되지만, 주어진 함수가 선형 분리가능한 함수가 아니라면, 함수를 정의하기 위해 1개 이상의 뉴런이 요구된다. .

2.2 기존의 알고리즘

2.2.1 ETL

1990년대중반 이진 신경회로망(Binary Neural Network)에서 은닉층내의 필요한 뉴런수를 자동적으로 결정하는 학습 알고리즘인 ETL이 제안되었다. 주어진 훈련입력의 기하학적인 분석에 기반을 둔 ETL은 요구된 separating hyperplane의 함수를 찾아내고 뉴런의 모든 수의 임계치

와 가중치를 결정하는 것이다. 이들 hyperplane들은 다른 hyperplane으로부터 원하는 같은 출력을 가진 입력을 분할하는데 있다. 따라서, 두 이웃하는 hyperplane들 사이에 위치한 훈련입력은 원하는 출력과 같은 출력을 가지고, 언제나 이진 공간내의 매핑을 통해 수렴영역을 보상함으로서, 각 개개의 패턴으로부터 기하학적인 관계를 추적하여 이미 hyperplane에 포함되어 있는 패턴으로부터 가장 가까운 true 패턴을 한 vertex씩 포함시켜나가다가 더 이상의 포함이 불가능해지면 새로운 hyperplane을 추가시키는 학습 알고리즘이다. ETL은 별도의 전처리과정 없이 바로 패턴으로부터 학습을 시작하지만 가장 가까운 패턴을 찾는 과정이 매 학습 단계에서 필요로 하기 때문에 초기 패턴에 따라 학습결과가 달라지는 한계를 가지고 있다.

표 1. 각 학습 알고리즘의 비교

	ETL	MTGA	제안하는 NETLA
사용 목적	패턴 분류를 위한 다층 퍼셉트론 회로망의 결정과 가중치, 임계치의 결정에 대한 학습 알고리즘		
기존 방법	다층 퍼셉트론		
적용 범위	패턴인식, 장면인식, 영상처리		
가중치, 임계치의 결정	RHS 방법 → Hyperplane 방정식 → 가중치와 임계치 결정 (Hyperplane 수 = 은닉층의 뉴런의 수)		
합성시 사용 Boolean 함수	Sum of Product	MSP (Minimal Sum of Product)	ESP (Expanded Sum of Product)
입출력 필요조건	없음	Unate 특성	없음
합성의 원리	최소의 hyperplane내 최대한 많은 true 패턴 포함 → 최소화	Unate 특성 적용 후 → 최소화	패턴 확장 후 → 최소화
단 점	합성 후 → 회로망이 변함	Not-Unate시 → 은닉층 내 뉴런으로 부가	
합성 후 은닉층내 뉴런수	few (NETLA ≤ MTGA < ETL) many		
입력층과 출력층 사이의 노드수	few (NETLA < MTGA < ETL) many		
패턴인식 시간	long (ETL < MTGA < NETLA) short		

2.2.2 MTGA

MTGA는 주어진 이진 공간의 분할문제를 MSP(Minimal Sum of Product)형태로 바꾸고, 이진 함수 형태로 표현된 MSP에 필요조건인 Unate의 특성을 적용하여 합성하는 방법이다. 여기서, Unate의 특성은

$$f_1(x_1,x_2,x_3)=x_1x_2+x_2x_3+x_1x_3 \ : \ \text{Unate}$$

$$f_2(x_1,x_2,x_3)=x_1x_2+x_2x_3+x_1\overline{x}_3 \ : \ \text{Not-Unate}$$

함수 f_1은 어떠한 입력변수도 positive 형태와 negative 형태가 동시에 존재하지 않고, 함수 f_2는 x_2가 positive 형태와 negative 형태가 동시에 존재한다.
즉, positive 형태와 negative 형태가 동시에 존재하지 않는 함수의 형태일 때 함수를 Unate의 특성을 만족한다고 한다. MTGA는 MSP의 형태로 나타낸 함수에 Unate의 특성을 적용하여 Unate의 특성을 만족할 때만 합성이 가능하다.
따라서, Unate의 특성을 적용하기 위해서 true 패턴들을 이진 함수의 표현 형태 중 MSP의 형태로 나타내어야 하고, Unate의 특성이 만족되지 못할 때에는 합성이 불가능한 단점을 지니고 있다. Unate의 특성을 만족하지 못하는 입력 패턴들은 은닉층내의 뉴런으로 부가되는 한계를 가지고 있다.

2.3 제안하는 NETLA(Newly Expanded and Truncated Learning Algorithm)

본 논문에서 제안하는 NETLA는 다층 퍼셉트론에 기초한 학습 알고리즘으로서, 입력 패턴의 순서에 관계없이 최소화된 이진 신경회로망을 합성할 수 있는 학습 알고리즘으로서, 제안하는 알고리즘의 장점은 아래와 같다.

1) False, true 패턴들로부터 가중치와 임계치가 계산 가능함으로서, 반복적인 학습 불필요.
2) 기존의 학습 알고리즘에서 제안된 Unate의 특성과 같은 필요조건의 불필요.
3) Boolean 함수 표현 중 ESP(Expanded Sum of Product)의 형태를 사용함으로서 학습에 사용하는 학습조건의 수 최소화.
4) 은닉층 내의 뉴런의 수와 입력층과 은닉층 사이의 연결된 노드 수가 최소화됨으로서, 패턴인식시간 단축 가능.

그림.2는 제안하는 NETLA를 이용한 신경 회로망의 구성에 대한 블록 다이어그램이다.

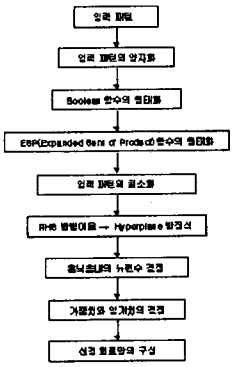

그림 2. NETLA 블록 다이어그램

2.3.1 ESP(Expanded Sum of Product)와 이진 신경회로망의 합성

Boolean 함수의 표현에는 ESP(Expanded Sum of Product) EPS(Expanded Product of Sum), MSP(Minimal Sum of Product), MPS(Minimal Product of Sum)의 대표적인 4개의 형태가 있다.
본 논문에서 입력 패턴을 최소화하기 위해서 Boolean 함수의 표현 형태 중 ESP의 형태를 사용하였다.
가. ESP(Expanded Sum of Product)
합의 확장된 곱의 형태로서,
예를 들어, 함수 $f(x_1,x_2,x_3,x_4,x_5,x_6)=x_1x_2\overline{x}_4x_5x_6+x_1\overline{x}_2x_4\overline{x}_5x_6$ 로 주어져 있을 때, 먼저 공통 변수를 찾는다. 공통변수 x_6이므로, x_6으로 묶어주면, 합의 곱의 형태인
$(x_1x_2\overline{x}_4x_5+x_1\overline{x}_2x_4\overline{x}_5)\cdot x_6$
가 되고, 입력 변수를 최소화하기 위해 x_6을 제거한다.
x_6을 제거하기 위해 합의 곱의 형태인
$(x_1x_2\overline{x}_4x_5+x_1\overline{x}_2x_4\overline{x}_5)\cdot 1$을 확장시킨다.

$$f(x_1,x_2,x_3,x_4,x_5,x_6)=(x_1x_2\overline{x}_4x_5+x_1\overline{x}_2x_4\overline{x}_5)\cdot x_6 + (x_1x_2\overline{x}_4x_5+x_1\overline{x}_2x_4\overline{x}_5)\cdot 1$$

이 된다.
여기서 부울 대수의 정의를 적용한다.
$f(x_1,x_2,x_3,x_4,x_5,x_6)=(x_1x_2\overline{x}_4x_5+x_1\overline{x}_2x_4\overline{x}_5)\cdot(x_6+1)$가 되고,
(x_6+1)은 "1"이 된다.
결과적으로,
$(x_1x_2\overline{x}_4x_5+x_1\overline{x}_2x_4\overline{x}_5)$의 최소항의 형태가 된다.

2.3.2 가중치와 임계치의 결정

이진 신경회로망의 기본 구조인 LTU의 가중치가 될 수 있는 조건은 n차원의 이진 공간에서 모든 true 패턴 x가 모든 false 패턴 y에 대해

$$\sum_{i=1}^{n} w_i x_i > \sum_{i=1}^{n} w_i y_i \qquad (3)$$

위 식을 만족시켜야 한다. 이 부등식을 풀기 위해 기하학적인 방법으로 LTU의 경우에 한해서 사용할 수 있는 RHS(Reference HyperSphere)방법이 있다.
이 RHS 방법은 선형 분리 가능함을 확인하는 동시에 가중치를 구해준다. 먼저 전체 패턴을 통과하는 RHS를 만든다.
즉, RHS식은

$$(x_1-\frac{1}{2})^2+(x_2-\frac{1}{2})^2+\cdots+(x_n-\frac{1}{2})^2=\frac{n}{4} \qquad (4)$$

이 된다.
또, SHS(Separating Hypersphere)를 RHS 안쪽에 중심점 $C=(\frac{c_1}{c_0},\frac{c_2}{c_0},\cdots,\frac{c_n}{c_0})$가 존재하며, 반지름 R이 되도록 정한다. 이때 R은 RHS와 SHS가 서로 교차하도록 만든다.
즉, SHS의 식은

$$(x_1-\frac{c_1}{c_0})^2+(x_2-\frac{c_2}{c_0})^2+\cdots+(x_n-\frac{c_n}{c_0})^2=R^2 \qquad (5)$$

이 된다.

여기서, RHS : 모든 패턴을 포함하는 hypersphere
SHS : true 패턴만을 포함하는 hypersphere

RHS방법에서는 앞의 식에서 c_0를 모든 true 패턴의 개수로 정의하고, c_i를 모든 true 패턴 x에서 x_i의 총합으로 정한다. 그러면, c는 모든 true 패턴들의 무게 중심에 위치하게 되고, 이렇게 만들어지는 hyperplane을 평행이동하면서 $\sum_{i=1}^{n} w_i x_i > \sum_{i=1}^{n} w_i y_i$을 만족시킬 수 있다.
이때, RHS와 SHS가 서로 교차하는 hyperplane 방정식은

$$(1-2\frac{c_1}{c_0})x_1+(1-2\frac{c_1}{c_0})x_2+\cdots+(1-2\frac{c_n}{c_0})x_n-\frac{T}{c_0}=0 \qquad (6)$$

이 되고, 이 식의 양변에 c_0를 곱하여 주면, 식(6)과 net 함수 $net(x,\ T)=w_1x_1+w_2x_2+\cdots+w_nx_n-T=0$를 같다고 하면
가중치 $w_i=c_0-2c_i$
임계치 $T=R^2c_0-\sum_{i=1}^{n} c_i^2/c_0$를 구할 수 있고, 선형 분리 가능한 hyperplane 방정식이 된다.
만약, 분할이 되지 않는 즉, 어떤 true 패턴 x와 어떤 false 패턴 y에 대해

$$\sum_{i=1}^{n} w_i x_i \le \sum_{i=1}^{n} w_i y_i \qquad (7)$$

의 관계가 되면 이 함수는 선형 분리 불가능하다. 이 방법은 '1'과 '0'의 합산으로 구해지므로, 모든 가중치와 임계치는 정수값을 가지게 된다. n차원의 이진 공간에서 모든 true 패턴 x가 모든 false 패턴 y에 대해 선형 분리 가능한 함수와 선형 분리 불가능한 함수를 분리하면 다음과 같다.

$$\sum_{i=1}^{n} w_i x_i > \sum_{i=1}^{n} w_i y_i \ : \text{선형 분리가능} \qquad (8)$$

$$\sum_{i=1}^{n} w_i x_i \le \sum_{i=1}^{n} w_i y_i \ : \text{선형 분리 불가능} \qquad (9)$$

2.4 네 개의 클래스를 가진 합성영상 분할문제

각 학습 알고리즘을 그림.3 에 보여지는 것처럼, 32*32 해상도를 가지는 gray scale 영상이 각 클래스에 256개의 점들을 양자화 한 결과 전제문제는 4개의 출력과 10개의 입력을 가지는 영상의 분할 문제에 적용하였다.

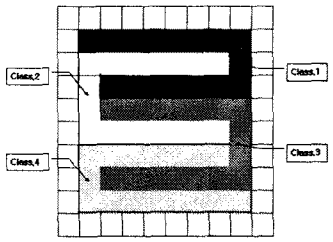

그림 3. 4개의 클래스를 가진 합성 영상

2.4.1 MTGA 합성 결과

MTGA의 필요조건이 MSP형태로 합성한 결과이다. MTGA를 사용하여 이진 신경회로망을 합성해 보면 각각의 클래스에 대해서 은닉층의 뉴런의 수가 각 클래스에 대해 2,3, 3, 2개로 구성된다. 아래는 각 클래스에 대해 MTGA로 합성한 결과이다.

$$f_1=x'_1x_6x_8+x'_2x_6x_8+x'_3x_6x_8+x_6x_7x_8+x'_1x'_2x'_3x_6x_7$$

$$f_2=x_3x'_6x_7x_8+x_2x'_6x_7x_8+x_1x'_6x_7x_8+x_3x_6x_7x'_8 +x_2x_6x_7x'_8+x_1x_6x_7x'_8+x_1x_2x_3x_6x'_7$$

$$f_3=x'_1x'_2x'_3x'_6x_7+x'_1x'_6x'_7x_8+x'_2x'_6x'_7x_8 +x'_3x'_6x'_7x_8+x'_1x_6x'_7x'_8+x'_2x_6x'_7x'_8+x_3x_6x'_7x'_8$$

$$f_4=x'_6x'_7x'_8+x_3x'_6x'_8+x_2x'_6x'_8+x_1x'_6x'_8+x_1x_2x_3x'_6x'_7$$

표.2 는 MTGA로 합성한 결과의 은닉층의 뉴런의 가중치와 임계치를 나타내었다. MTGA로 합성한 결과를 보면 w_4와 w_5 그리고, w_9와 w_{10}이 모두 '0'의 값을 가지고 있음을 알 수 있다. 그림. 4에 보여지는 것처럼 4개의 출력층과 은닉층 사이의 전체 뉴런이 10개로 구성된 회로망이 합성되었다.

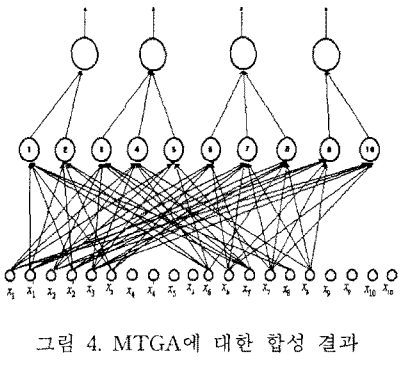

그림 4. MTGA에 대한 합성 결과

표 2. MTGA에 대한 가중치와 임계치

Layer		Unit	Weight										Threshold
Hidden	1	1	-4	-4	-4	0	0	60	4	60	0	0	110
		2	-8	-8	-8	0	0	8	8	0	0	0	12
	2	3	4	4	4	0	0	-28	28	28	0	0	58
		4	4	4	4	0	0	28	28	-28	0	0	58
		5	8	8	8	0	0	8	-8	0	0	0	28
	3	6	-8	-8	-8	0	0	-8	8	0	0	0	4
		7	-4	-4	-4	0	0	-28	-28	28	0	0	18
		8	-4	-4	4	0	0	28	-28	-28	0	0	22
	4	9	4	4	4	0	0	-60	-4	-60	0	0	-2
		10	8	8	8	0	0	-8	-8	0	0	0	20

2.4.2 제안하는 NETLA 합성 결과

아래는 NETLA로 합성한 결과이다. NETLA에서 제안한 ESP로 합성하고 각각의 최적의 합성을 위하여 각각 하나씩 확장하여 합성한 결과이다.

$$f_1=(x'_1x_8+x'_2x_8+x'_3x_8+x_7x_8+x'_1x'_2x'_3x_7)\cdot x_6$$
$$+(x'_1x_8+x'_2x_8+x'_3x_8+x_7x_8+x'_1x'_2x'_3x_7)\cdot 1$$

$$f_2=(x_3x_7x_8+x_2x_7x_8+x_1x_7x_8)\cdot x'_6$$
$$+(x_3x_7x_8+x_2x_7x_8+x_1x_7x_8)\cdot 1$$
$$+(x_3x_7x'_8+x_2x_7x'_8+x_1x_7x'_8+x_1x_2x_3x'_7)\cdot x_6$$
$$+(x_3x_7x'_8+x_2x_7x'_8+x_1x_7x'_8+x_1x_2x_3x'_7)\cdot 1$$

$$f_3=(x'_1x'_2x'_3x_7+x'_1x'_7x_8+x'_2x'_7x_8+x'_3x'_7x_8)\cdot x'_6$$
$$+(x'_1x'_2x'_3x_7+x'_1x'_7x_8+x'_2x'_7x_8+x'_3x'_7x_8)\cdot 1$$
$$+(x'_1x'_7x'_8+x'_2x'_7x'_8+x_3x'_7x'_8)\cdot x_6$$
$$+(x'_1x'_7x'_8+x'_2x'_7x'_8+x_3x'_7x'_8)\cdot 1$$

$$f_4=(x'_7x'_8+x_3x'_8+x_2x'_8+x_1x'_8+x_1x_2x_3x'_7)\cdot x'_6$$
$$+(x'_7x'_8+x_3x'_8+x_2x'_8+x_1x'_8+x_1x_2x_3x'_7)\cdot 1$$

표. 3은 NETLA로 합성한 결과의 은닉층의 뉴런의 가중치와 임계치를 나타내었다. 합성한 결과를 보면 w_4, w_5, w_6 그리고, w_9와 w_{10}이 모두 '0'의 값을 가지고 있음을 알 수 있다.

MSP 형태를 사용할 때 보다 ESP 형태를 사용함으로서 불필요하게 자세히 양자화 되어있는 패턴으로부터 더욱 상세하게 중복성을 찾음으로서, 각 입력층과 은닉층 사이의 노드의 수도 감소시킬 수 있음으로서, 입력 패턴의 인식속도도 감소시킬 수 있었다.

그림. 5에 보여지는 것처럼 4개의 출력층과 은닉층내의 전체뉴런이 10개로 구성된 회로망이 합성되었고, MTGA로 합성한 결과보다, 입력층과 은닉층 사이의 노드수도 감소됨을 보여주고 있다.

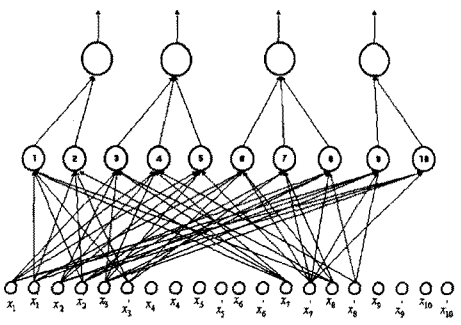

그림 5. MTGA에 대한 합성 결과

표 3. NETLA에 대한 가중치와 임계치

Layer		Unit	Weight										Threshold
Hidden	1	1	-4	-4	-4	0	0	0	60	4	0	0	50
		2	-8	-8	-8	0	0	0	8	0	0	0	20
	2	3	4	4	4	0	0	0	-28	28	0	0	-8
		4	4	4	4	0	0	0	28	28	0	0	50
		5	8	8	8	0	0	0	8	0	0	0	36
	3	6	-8	-8	-8	0	0	0	-8	8	0	0	4
		7	0	-4	-4	0	0	0	-28	-28	0	0	6
		8	-4	-4	4	0	0	0	-28	-28	0	0	-6
	4	9	4	4	4	0	0	0	-60	-4	0	0	70
		10	8	8	8	0	0	0	-8	0	0	0	28

3. 결 론

본 논문에서는 ESP 형태를 이용한 합성 알고리즘으로 이진 신경회로망을 최소 합성하는 방법을 제시했다. 시뮬레이션을 통해서도 기존의 이진 신경회로망합성 알고리즘들보다 향상된 합성 결과를 보여주고 있다. 시뮬레이션 결과에서 보여주고 있듯이 입력패턴의 불필요한 중복성을 찾음으로서 입력층과 은닉층사이의 노드수도 감소시킬 수 있었다. 따라서, 패턴인식의 속도도 단축시킬 수 있었다.

[참 고 문 헌]

[1] S. Park, J. H. Kim, and H. Chang, "A learning algorithm for discrete multilayer perceptron", in Proc. Int. Sysp. Cricuit Systems, Singapore, Jun.1991.

[2] 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, p. 988-992, Aug. 1993.

[3] J. H. Kim, S. K. Park, Han, H. Oh, and M. S. Han, "The geometrical learning of binary neural network", IEEE Trans. Neural Networks, vol. 6.no.1. pp.237-247, Jan. 1995.

[4] D.E Rumelhart, J.M. McClelland, and the PDP Research Group, Parallel Destributed Precessing. Cambridge, MA: MIT Press, 1986.

[5] J.H. Kim, S.K. Park, Y. Han, H. Oh, and M. S. Han, "Efficient VLSI implementation of a 3-layer thresshold network," Proc. ICNN '97, Houston, TX, vol. 2, pp. 888-893, jun. 1997.

[6] J. M. Zurada, Introduction to Artificial Neural Systems, Boston, MA:PWS Publishing Company, 1995.

[7] 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.

[8] D.L. Gray and A.N. Michel, "A training algorithm forbinary feedforward neural networks," IEEE trans. Neural Networks, vol. 3, no. 2, pp. 176-194, Mar. 1992.