(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0156040
(43) 공개일자 2024년10월29일

(51) 국제특허분류(Int. Cl.)
G06N 3/063 (2023.01) *G06F 17/16* (2006.01)
G11C 11/16 (2006.01) *G11C 11/54* (2006.01)
(52) CPC특허분류
G06N 3/063 (2013.01)
G06F 17/16 (2013.01)
(21) 출원번호 10-2023-0052465
(22) 출원일자 2023년04월21일
심사청구일자 2023년04월21일

전체 청구항 수 : 총 5 항

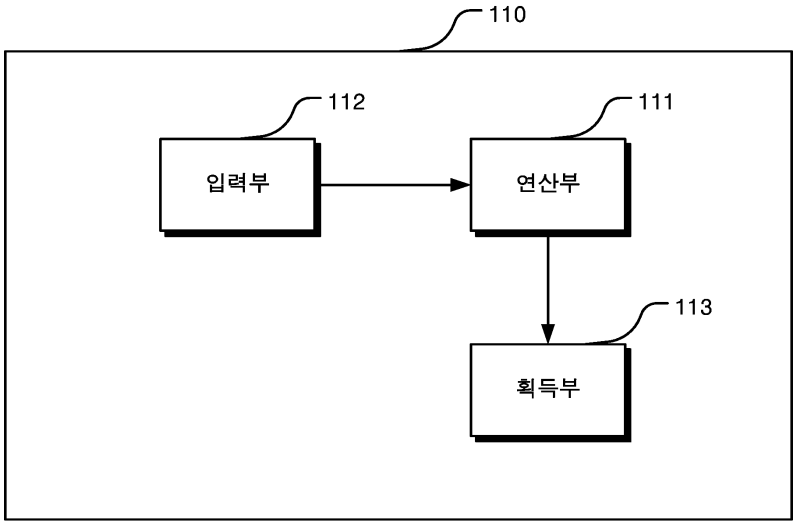
(71) 출원인
인천대학교 산학협력단
인천광역시 연수구 갯벌로 27, 인천대학교 이노베이션센터(송도동)
(72) 발명자
나태희
인천광역시 연수구 아트센터대로97번길 75 송도더샵하버뷰13단지 1303동 401호
(74) 대리인
특허법인시공

(54) 발명의 명칭 **2진 신경망의 출력 연산을 위한 MRAM 기반의 연산 장치**

(57) 요 약

2진 신경망의 출력 연산을 위한 MRAM 기반의 연산 장치가 개시된다. 본 발명은 MRAM의 각 메모리 셀을 2진 신경망의 가중치 저장을 위한 시냅스로 활용하여 상기 2진 신경망에 대한 출력을 연산할 수 있는 연산 장치에 대한 것이다.

대 표 도 - 도1



(52) CPC특허분류
G11C 11/1655 (2013.01)
G11C 11/1657 (2013.01)
G11C 11/54 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711173032
과제번호	2022M3I7A2079267
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	원천기술개발사업
연구과제명	자구벽 제어 기반 초저전력 PIM 단위소자 및 아키텍처 개발
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2022.04.20 ~ 2024.12.31

명 세 서

청구범위

청구항 1

2진 신경망의 출력 연산을 위한 MRAM 기반의 연산 장치에 있어서,

하나의 트랜지스터와 하나의 MTJ(Magnetic Tunnel Junction)로 구성된 메모리 셀들이, n(n은 2이상의 자연수임)개의 행들과 m(m은 2이상의 자연수임)개의 열들로 구성되는 어레이(array) 구조로 배치되어 있는 상태에서, 각 행을 구성하는 m개의 메모리 셀들은 동일한 워드 라인(word line)에 연결되어 있고, 각 열을 구성하는 n개의 메모리 셀들은 동일한 비트 라인(bit line)에 연결되어 있으며, 상기 m개의 열들 각각의 비트 라인은, 각 열에 대응되어 존재하는 감지 증폭기(sense amplifier)의 입력과 연결되고, 상기 m개의 열들 각각에 대응되는 감지 증폭기의 출력은, 각 열에 대응되어 존재하는 카운터의 입력과 연결되어 있는 연산부;

상기 n개의 행들 각각의 워드 라인을 통해, 2진 신경망(Binary Neural Network: BNN)의 입력을 인가하는 입력부; 및

상기 m개의 열들 각각에 대응되는 카운터의 출력으로부터, 상기 2진 신경망의 출력을 획득하는 획득부

를 포함하는 연산 장치.

청구항 2

제1항에 있어서,

상기 2진 신경망의 입력은 '1'과 '0'의 이진 값을 성분으로 갖는 n차원의 임베딩 벡터로 구성되어 있고, 이때, 상기 입력부는 상기 임베딩 벡터를 구성하는 n개의 성분들 각각을, 상기 n개의 행들 각각에 하나씩 할당한 후, 상기 n개의 행들 각각의 워드 라인을 통해, 각 행이 할당된 성분에 대응되는 제어 전압을, 첫 번째 행부터 n번째 행까지 순차적으로 입력으로 인가하는 것을 특징으로 하는 연산 장치.

청구항 3

제2항에 있어서,

상기 입력부는

상기 n개의 행들 중 '1'의 성분이 할당된 행에 대해서는, 사전 설정된 제1 크기의 제어 전압을 입력으로 인가하고, 상기 n개의 행들 중 '0'의 성분이 할당된 행에 대해서는, '0V'의 제어 전압을 입력으로 인가하는 것을 특징으로 하는 연산 장치.

청구항 4

제2항에 있어서,

상기 n개의 행들과 상기 m개의 열들을 구성하는 각 메모리 셀의 MTJ는, 내부 저항 값이, 제1 저항 값 또는 제2 저항 값 - 상기 제2 저항 값은, 상기 제1 저항 값보다 작은 값임 - 중 어느 한 저항 값으로 스위칭될 수 있는 소자로서, 상기 제1 저항 값 또는 상기 제2 저항 값으로의 스위칭을 기반으로, 상기 2진 신경망에 대한 기계학습을 통해 지정된, '+1' 또는 '-1'이라는 이진 가중치를 저장하는 데에 사용되는 소자이고, 이때, '+1'에 대응되는 가중치를 저장하기 위한 MTJ는, 내부 저항 값이 상기 제1 저항 값으로 스위칭되어 있고, '-1'에 대응되는 가중치를 저장하기 위한 MTJ는, 내부 저항 값이 상기 제2 저항 값으로 스위칭되어 있는 것을 특징으로 하는 연산 장치.

청구항 5

제4항에 있어서,

상기 획득부는

상기 m개의 열들 각각에 대응되는 카운터의 출력으로 산출되는 비트열에서, MSB(Most Significant Bit)를 확인하여, MSB가 '1'로 확인되는 카운트로부터 '1'을 출력 값으로 선택하고, MSB가 '0'으로 확인되는 카운트로부터 '0'을 출력 값으로 선택한 후, 상기 m개의 열들 각각에 대응되는 카운터로부터 선택된 출력 값을 성분으로 갖는 m차원의 벡터를, 상기 2진 신경망의 출력으로 획득하는 것을 특징으로 하는 연산 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 2진 신경망의 출력 연산을 위한 MRAM 기반의 연산 장치에 대한 것이다.

배 경 기 술

[0002] 최근, 인공지능 기술에 대한 관심이 증폭되면서, 다양한 종류의 신경망 모델이 제시되고 있다.

[0003] 이 중 2진 신경망(Binary Neural Network: BNN)은 신경망의 가중치를 '+1'과 '-1'의 이진 값으로 표현하여 구성한 신경망으로써, 다른 심층 신경망과 대비하여 연산의 복잡도가 낮아서, 시스템 자원이 제한된 엣지 컴퓨팅 장치에서 주요하게 활용될 수 있다.

[0004] 이러한 2진 신경망은 XNOR(exclusive NOR) 동작을 지원해야 한다는 점에서, 2진 신경망의 출력을 연산할 수 있도록 지원하는 새로운 아키텍처에 대한 연구가 필요하다.

[0005] 이와 관련해서, 최근에는 MTJ(Magnetic Tunnel Junction)를 저장 소자로 활용하는 M(Magnetic)RAM이 등장하고 있는데, MRAM의 각 메모리 셀을 2진 신경망의 가중치 저장을 위한 시냅스로 활용함으로써, 2진 신경망의 출력을 연산하는 아키텍처를 구성하는 방안의 도입을 고려할 수 있다.

[0006] 여기서, MTJ는 상단 전극(top electrode)에서 하단 전극(bottom electrode)으로 전류가 흐르는 경우, 낮은 저항(R_L)을 갖게 되고, 하단 전극에서 상단 전극으로 전류가 흐르는 경우, 높은 저항(R_H)을 갖게 되는 특성을 갖는 소자로서, 이러한 MTJ는, MTJ에 인가되는 전류의 방향을 조정함으로써, MTJ의 저항을 낮은 저항(R_L) 또는 높은 저항(R_H)으로 스위칭시킬 수 있다는 점에서, 소정의 데이터를 저장하는 데 사용될 수 있다.

[0007] 이와 관련해서, 소정의 2진 신경망에 대한 기계학습을 통해, 시냅스를 구성하는 각 가중치가 결정된 경우, 이 가중치를, MTJ의 저항 스위칭 특성을 이용하여, 각 메모리 셀의 MTJ에 저장해 둔 후, 소정의 입력에 따른 상기 2진 신경망의 출력 연산이 필요할 때, 각 메모리 셀의 MTJ에 저장된 가중치를 읽어들여, 상기 이진 신경망에 대한 출력을 연산하는 MRAM 기반의 연산 아키텍처에 대한 도입을 고려할 수 있다.

발명의 내용

해결하려는 과제

[0008] 본 발명은 MRAM의 각 메모리 셀을 2진 신경망의 가중치 저장을 위한 시냅스로 활용하여 상기 2진 신경망에 대한 출력을 연산할 수 있는 새로운 아키텍처를 제시하고자 한다.

과제의 해결 수단

[0009] 본 발명의 일실시예에 따른 2진 신경망의 출력 연산을 위한 MRAM 기반의 연산 장치는, 하나의 트랜지스터와 하나의 MTJ(Magnetic Tunnel Junction)로 구성된 메모리 셀들이, n(n은 2이상의 자연수임)개의 행들과 m(m은 2이상의 자연수임)개의 열들로 구성되는 어레이(array) 구조로 배치되어 있는 상태에서, 각 행을 구성하는 m개의 메모리 셀들은 동일한 워드 라인(word line)에 연결되어 있고, 각 열을 구성하는 n개의 메모리 셀들은 동일한 비트 라인(bit line)에 연결되어 있으며, 상기 m개의 열들 각각의 비트 라인은, 각 열에 대응되어 존재하는 감지 증폭기(sense amplifier)의 입력과 연결되고, 상기 m개의 열들 각각에 대응되는 감지 증폭기의 출력은, 각 열에 대응되어 존재하는 카운터의 입력과 연결되어 있는 연산부, 상기 n개의 행들 각각의 워드 라인을 통해, 2진 신경망(Binary Neural Network: BNN)의 입력을 인가하는 입력부 및 상기 m개의 열들 각각에 대응되는 카운터의 출력으로부터, 상기 2진 신경망의 출력을 획득하는 획득부를 포함한다.

발명의 효과

[0010] 본 발명은 MRAM의 각 메모리 셀을 2진 신경망의 가중치 저장을 위한 시냅스로 활용함으로써, 상기 2진 신경망에 대한 출력을 연산할 수 있는 새로운 아키텍처를 제시할 수 있다.

도면의 간단한 설명

[0011] 도 1은 본 발명의 일실시예에 따른 연산 장치의 구조를 도시한 도면이다.

도 2 내지 도 3은 본 발명의 일실시예에 따른 연산 장치를 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0012] 이하에서는 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명하기로 한다. 이러한 설명은 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였으며, 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 본 명세서 상에서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 사람에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다.

[0013] 본 문서에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있다는 것을 의미한다. 또한, 본 발명의 다양한 실시예들에 있어서, 각 구성요소들, 기능 블록들 또는 수단들은 하나 또는 그 이상의 하부 구성요소로 구성될 수 있고, 각 구성요소들이 수행하는 전기, 전자, 기계적 기능들은 전자회로, 집적회로, ASIC(Application Specific Integrated Circuit) 등 공지된 다양한 소자들 또는 기계적 요소들로 구현될 수 있으며, 각각 별개로 구현되거나 2 이상이 하나로 통합되어 구현될 수도 있다.

[0014] 도 1은 본 발명의 일실시예에 따른 2진 신경망의 출력 연산을 위한 MRAM 기반의 연산 장치의 구조를 도시한 도면이다.

[0015] 도 1을 참조하면, 본 발명에 따른 연산 장치(110)는 연산부(111), 입력부(112) 및 획득부(113)를 포함한다.

[0016] 연산부(111)는, 하나의 트랜지스터와 하나의 MTJ(Magnetic Tunnel Junction)로 구성된 메모리 셀들이, n(n은 2 이상의 자연수임)개의 행들과 m(m은 2이상의 자연수임)개의 열들로 구성되는 어레이(array) 구조로 배치되어 있는 상태에서, 각 행을 구성하는 m개의 메모리 셀들은 동일한 워드 라인(word line)에 연결되어 있고, 각 열을 구성하는 n개의 메모리 셀들은 동일한 비트 라인(bit line)에 연결되어 있으며, 상기 m개의 열들 각각의 비트 라인은, 각 열에 대응되어 존재하는 감지 증폭기(sense amplifier)의 입력과 연결되고, 상기 m개의 열들 각각에 대응되는 감지 증폭기의 출력은, 각 열에 대응되어 존재하는 카운터의 입력과 연결되어 있는 상태로, 존재하는 구성이다.

[0017] 관련해서, 각 메모리 셀은, 도 2에 도시된 그림과 같이, 하나의 트랜지스터(211)와 하나의 MTJ(212)로 구성되어 있을 수 있다. 여기서, 트랜지스터(211)는 NMOS가 사용될 수 있으며, 이때, MTJ(212)의 일측이 트랜지스터(211)의 드레인에 연결되어 있을 수 있고, MTJ(212)의 타측이 비트 라인에 연결되어 있을 수 있으며, 트랜지스터(211)의 게이트가 워드 라인에 연결되어 있을 수 있다.

[0018] 그리고, 각 메모리 셀은 도 3의 도면부호 320에 도시된 그림과 같이, n개의 행들과 m개의 열들로 구성되는 어레이 구조로 배치되어 있을 수 있다. 이때, 도면부호 320에 도시된 그림과 같이, 각 행을 구성하는 m개의 메모리 셀들은 동일한 워드 라인에 연결되어 있고, 각 열을 구성하는 n개의 메모리 셀들은 동일한 비트 라인에 연결되어 있을 수 있다.

[0019] 그리고, 상기 m개의 열들 각각의 비트 라인은, 도면부호 320에 도시된 그림과 같이, 각 열에 대응되어 존재하는 감지 증폭기의 입력과 연결되고, 상기 m개의 열들 각각에 대응되는 감지 증폭기의 출력은, 각 열에 대응되어 존재하는 카운터의 입력과 연결되어 있을 수 있다.

[0020] 이때, 본 발명의 다른 일실시예에 따르면, 감지 증폭기는 도면부호 320에 도시된 그림과 같이, 각 비트 라인마다 하나씩 존재할 수도 있고, 둘 이상의 비트 라인당 하나씩 존재하는 형태로 구성될 수도 있다. 구체적으로, 둘 이상의 비트 라인이, 멀티플렉서(multiplexer)와 같은 소정의 스위칭 소자를 매개로 하여, 하나의 감지 증폭기의 입력과 연결될 수 있다. 예컨대, 2개의 비트 라인이 멀티플렉서를 매개로 하여, 하나의 감지 증폭기의 입력과 연결되는 구조라고 하는 경우, 특정 열의 비트 라인이 멀티플렉서를 통해 감지 증폭기의 입력과 먼저 연결

되어 2진 신경망의 연산에 활용된 후, 다른 열의 비트 라인이 멀티플랙서를 통해 감지 증폭기의 입력과 연결되어 2진 신경망의 연산에 활용되는 형태로 동작할 수 있다.

[0021] 입력부(112)는 상기 n개의 행들 각각의 워드 라인을 통해, 2진 신경망(Binary Neural Network: BNN)의 입력을 인가한다.

[0022] 획득부(113)는 상기 m개의 열들 각각에 대응되는 카운터의 출력으로부터, 상기 2진 신경망의 출력을 획득한다.

[0023] 이때, 본 발명의 일실시예에 따르면, 상기 2진 신경망의 입력은 '1'과 '0'의 이진 값을 성분으로 갖는 n차원의 임베딩 벡터로 구성되어 있을 수 있다.

[0024] 이때, 입력부(112)는 상기 임베딩 벡터를 구성하는 n개의 성분들 각각을, 상기 n개의 행들 각각에 하나씩 할당한 후, 상기 n개의 행들 각각의 워드 라인을 통해, 각 행이 할당된 성분에 대응되는 제어 전압을, 첫 번째 행부터 n번째 행까지 순차적으로 입력으로 인가할 수 있다.

[0025] 또한, 본 발명의 일실시예에 따르면, 입력부(112)는 상기 n개의 행들 중 '1'의 성분이 할당된 행에 대해서는, 사전 설정된 제1 크기의 제어 전압을 입력으로 인가하고, 상기 n개의 행들 중 '0'의 성분이 할당된 행에 대해서는, '0V'의 제어 전압을 입력으로 인가할 수 있다.

[0026] 또한, 본 발명의 일실시예에 따르면, 상기 n개의 행들과 상기 m개의 열들을 구성하는 각 메모리 셀의 MTJ는, 내부 저항 값이, 제1 저항 값 또는 제2 저항 값(상기 제2 저항 값은, 상기 제1 저항 값보다 작은 값임) 중 어느 한 저항 값으로 스위칭될 수 있는 소자로서, 상기 제1 저항 값 또는 상기 제2 저항 값으로의 스위칭을 기반으로, 상기 2진 신경망에 대한 기계학습을 통해 지정된, '+1' 또는 '-1'이라는 이진 가중치를 저장하는 데에 사용되는 소자일 수 있다. 이때, '+1'에 대응되는 가중치를 저장하기 위한 MTJ는, 내부 저항 값이 상기 제1 저항 값으로 스위칭되어 있고, '-1'에 대응되는 가중치를 저장하기 위한 MTJ는, 내부 저항 값이 상기 제2 저항 값으로 스위칭되어 있을 수 있다.

[0027] 이때, 본 발명의 일실시예에 따르면, 획득부(113)는 상기 m개의 열들 각각에 대응되는 카운터의 출력으로 산출되는 비트열에서, MSB(Most Significant Bit)를 확인하여, MSB가 '1'로 확인되는 카운트로부터 '1'을 출력 값으로 선택하고, MSB가 '0'으로 확인되는 카운트로부터 '0'을 출력 값으로 선택한 후, 상기 m개의 열들 각각에 대응되는 카운터로부터 선택된 출력 값을 성분으로 갖는 m차원의 벡터를, 상기 2진 신경망의 출력으로 획득할 수 있다.

[0028] 이하에서는, 도 3을 참조하여, 본 발명에 따른 연산 장치(110)의 동작을 보다 상세히 설명하기로 한다.

[0029] 먼저, 도 3의 도면부호 310에 도시된 그림과 같이, n개의 성분들(In_1, In_2, In_3, ..., In_n)로 구성된 n차원의 임베딩 벡터가 입력으로 인가될 때, 이에 대응되는 m개의 성분들로 구성되는 m차원의 벡터가 출력으로 산출되는 2진 신경망이 존재한다고 가정하자.

[0030] 이러한 상황이라면, 도면부호 310에 도시된 그림과 같이, n개의 성분들(In_1, In_2, In_3, ..., In_n) 각각이, 2진 신경망의 시냅스를 구성하는 각 이진 가중치(W_{11}, W_{12}, W_{13}, W_{1n}, ..., W_{mn})와 곱해진 후, 이의 합산 값을, 시그모이드 함수와 같은 소정의 활성화 함수에 통과시킨 값이, 상기 2진 신경망의 출력인, m차원의 벡터를 구성하는 각 성분이 될 것이다.

[0031] 이때, 도면부호 310에 도시된 그림과 같은 2진 신경망을 본 발명의 메모리 셀 구조로 매핑시킨 것이, 도면부호 320에 도시된 그림이다.

[0032] 이러한 상황에서, 상기 2진 신경망은 이미 기계학습이 완료되어, 상기 2진 신경망을 구성하는 시냅스의 각 이진 가중치(W_{11}, W_{12}, W_{13}, W_{1n}, ..., W_{mn})가 이미 결정된 상태임을 가정하자.

[0033] 이로 인해, 상기 n개의 행들과 상기 m개의 열들을 구성하는 각 메모리 셀의 MTJ에는, 각 이진 가중치에 대한 정보가 저장되어 있을 수 있다. 이때, MTJ는 높은 저항 값인 제1 저항 값과, 낮은 저항 값인 제2 저항 값 중, 어느 한 저항 값으로 스위칭될 수 있는 특성을 가지고 있을 수 있다. 이 경우, '+1'에 대응되는 가중치를 저장하기 위한 MTJ는, 내부 저항 값이 상기 제1 저항 값으로 스위칭되어 있을 수 있고, '-1'에 대응되는 가중치를 저장하기 위한 MTJ는, 내부 저항 값이 상기 제2 저항 값으로 스위칭되어 있을 수 있다.

[0034] 예컨대, W_{11}이라는 가중치가 '+1'로 결정되었다고 하는 경우, W_{11}이라는 가중치를 저장하기 위한 MTJ는, 내부 저항 값이 상기 제1 저항 값으로 스위칭되어 있을 수 있다. 이러한, 각 메모리 셀의 MTJ의 내부 저항 값 스위칭

을 통한 이진 가중치의 저장은, 개발자에 의해 사전 수행될 수 있다. 즉, 상기 2진 신경망에 대한 기계학습이 완료되어, 각 이진 가중치가 결정되면, 개발자는 소정의 데이터 쓰기 과정을 통해, 각 메모리 셀의 MTJ의 내부 저항 값을, 각 MTJ에 대응되는 이진 가중치에 맞게 스위칭시켜둘 수 있다.

[0035] 또한, 상기 m개의 열들 각각에 대응되는 카운터는, 도면부호 320에 도시된 그림과 같이, 7비트 카운터가 사용될 수 있는데, 이때, m개의 열들 각각에 대응되는 카운터의 초기 상태는 중간 값, 즉, '0111111'로 설정되어 있을 수 있다.

[0036] 이러한 상황에서, n개의 성분들(In_1, In_2, In_3, ..., In_n)로 구성되는 임베딩 벡터가 입력으로 인가되어, 상기 2진 신경망의 출력이 산출되어야 하는 상황이 되는 경우, 입력부(112)는 상기 임베딩 벡터를 구성하는 n개의 성분들(In_1, In_2, In_3, ..., In_n) 각각을, 상기 n개의 행들 각각에 하나씩 할당한 후, 도면부호 320에 도시된 그림과 같이, 상기 n개의 행들 각각의 워드 라인을 통해, 각 행이 할당된 성분에 대응되는 제어 전압을, 첫 번째 행부터 n번째 행까지 순차적으로 입력으로 인가할 수 있다.

[0037] 관련해서, 입력부(112)는 첫 번째로, 첫 번째 행의 워드 라인을 통해, 'In_1'에 대응되는 제어 전압을 입력으로 인가할 수 있다. 이때, 'In_1'이 '1'이라고 하는 경우, 입력부(112)는 첫 번째 행의 워드 라인을 통해, 사전 설정된 제1 크기의 제어 전압을 입력으로 인가할 수 있다. 그러면, 첫 번째 행의 메모리 셀들을 구성하는 트랜지스터가 도통 상태가 되어, 상기 첫 번째 행의 메모리 셀들을 구성하는 MTJ의 저항 상태에 따라, m개의 열들 각각에 대응되는 카운터의 출력이 변경될 수 있다.

[0038] 그러고 나서, 입력부(112)는, 두 번째 행의 워드 라인을 통해, 'In_2'에 대응되는 제어 전압을 인가할 수 있다. 이때, 'In_2'가 '0'이라고 하는 경우, 입력부(112)는 두 번째 행의 워드 라인을 통해, '0V'의 제어 전압을 입력으로 인가할 수 있다. 그러면, 두 번째 행의 메모리 셀들을 구성하는 트랜지스터는 턴-오프가 되어, 상기 두 번째 행의 메모리 셀들을 구성하는 MTJ로 전류가 흐르지 않게 되기 때문에, m개의 열들 각각에 대응되는 카운터의 출력이 그대로 유지된다.

[0039] 이러한 방식으로, 입력부(112)는 도면부호 320에 도시된 그림과 같이, 상기 n개의 행들 각각의 워드 라인을 통해, 각 행이 할당된 성분에 대응되는 제어 전압을, 첫 번째 행부터 n번째 행까지 순차적으로 입력으로 인가할 수 있다.

[0040] 이렇게, 입력부(112)를 통해, 각 행이 할당된 성분에 대응되는 제어 전압이, 첫 번째 행부터 n번째 행까지 순차적으로 입력으로 인가되게 되면, m개의 열들 각각에 대응되는 카운터의 출력이, 각 열의 메모리 셀을 구성하는 MTJ의 저항 상태에 따라 변화하게 된다. 이로 인해, m개의 열들 각각에 대응되는 카운터의 출력은, 도면부호 330으로 표시한 부분과 같이, 각 열의 메모리 셀을 구성하는 n개의 입력들과 n개의 이진 가중치들이 곱해진 후, 이들이 합해진 값과 동일한 값이 될 수 있다.

[0041] 이때, 획득부(113)는 상기 m개의 열들 각각에 대응되는 카운터를 통해 최종 출력이 산출되면, 상기 m개의 열들 각각에 대응되는 카운터의 출력으로 산출되는 비트열에서, MSB를 확인하여, MSB가 '1'로 확인되는 카운트로부터 '1'을 출력 값으로 선택하고, MSB가 '0'으로 확인되는 카운트로부터 '0'을 출력 값으로 선택한 후, 상기 m개의 열들 각각에 대응되는 카운터로부터 선택된 출력 값을 성분으로 갖는 m차원의 벡터를, 상기 2진 신경망의 출력으로 획득할 수 있다.

[0042] 관련해서, 카운터의 초기 값이 '0111111'과 같은 중앙 값으로 시작하였기 때문에, MSB가 '1'이라는 의미는, 해당 카운터를 통해서 산출된 출력이 '0'을 초과하는 값이라고 볼 수 있기 때문에, 획득부(111)는 해당 카운터로부터 '1'을 출력 값으로 선택할 수 있다.

[0043] 이러한 방식으로, 획득부(113)가 상기 m개의 열들 각각에 대응되는 카운터로부터 출력 값을 선택한 결과, 첫 번째 열에 대응되는 카운터로부터 'H_1=1'이 선택되었고, 두 번째 열에 대응되는 카운터로부터 'H_2=1'이 선택되었으며, 세 번째 열에 대응되는 카운터로부터 'H_3=0'이 선택되는 방식으로, m개의 열들 각각에 대응되는 카운터로부터 소정의 출력 값이 선택되었다고 하는 경우, 획득부(113)는 각 카운터의 출력 값을 성분으로 갖는 m차원의 벡터인 '[1 1 0 .... H_m]'을 생성함으로써, 이 m차원의 벡터를 상기 2진 신경망의 출력으로 획득할 수 있다.

[0044] 본 발명의 일실시예에 따르면, 본 발명에서 제안된 2진 신경망의 출력을 연산하기 위한 아키텍처는, 2진 신경망의 출력 연산에 활용될 수 있을 뿐 아니라, 데이터를 저장하기 위한 메모리 저장 장치로서도 활용될 수 있다.

[0045] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다.

[0046] 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

부호의 설명

[0047] 110: 연산 장치

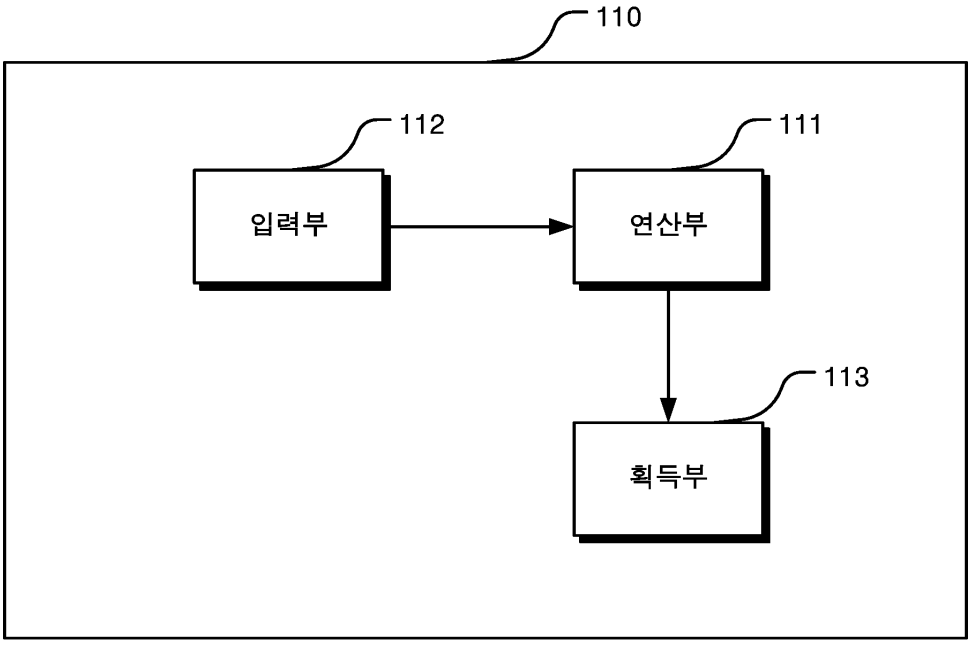
111: 연산부

112: 입력부

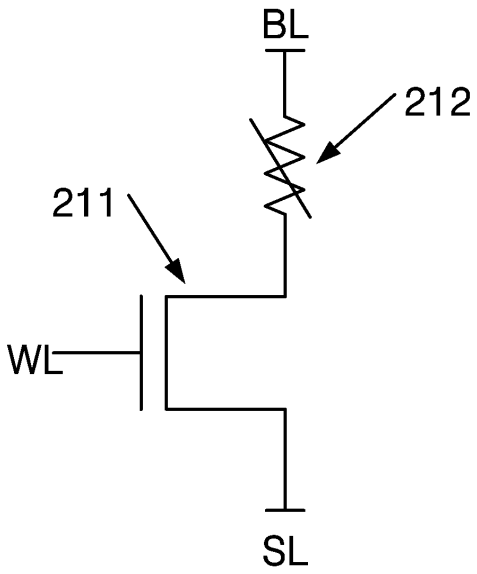
113: 획득부

도면

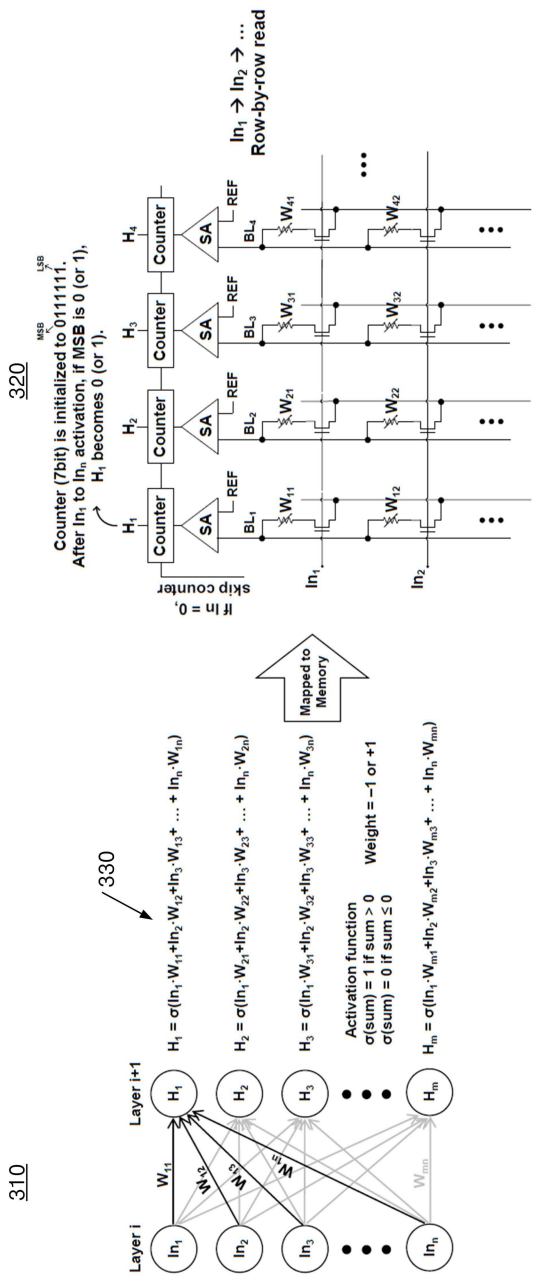
도면1



도면2



도면3


310
320
330
Layer i
Layer i+1
Activation function
σ(sum) = 1 if sum > 0
σ(sum) = 0 if sum ≤ 0
Weight = −1 or +1
Mapped to Memory
Counter (7bit) is initialized to 0111111.
After In1 to Inn activation, if MSB is 0 (or 1),
H1 becomes 0 (or 1).
MSB
LSB
If In = 0,
skip counter
In1 → In2 → ...
Row-by-row read
Counter
SA
REF