



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2022년07월08일

(11) 등록번호 10-2419338

(24) 등록일자 2022년07월06일

(51) 국제특허분류(Int. Cl.)

G06N 3/063 (2006.01) G06N 3/08 (2006.01)

(52) CPC특허분류

G06N 3/063 (2013.01)

G06N 3/08 (2013.01)

(21) 출원번호 10-2019-0160519

(22) 출원일자 2019년12월05일

심사청구일자 2019년12월05일

(65) 공개번호 10-2021-0070590

(43) 공개일자 2021년06월15일

(56) 선행기술조사문헌

KR1020180092342 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

정성욱

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C513(신촌동)

김기석

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C206(신촌동)

(뒷면에 계속)

(74) 대리인

민영준

전체 청구항 수 : 총 14 항

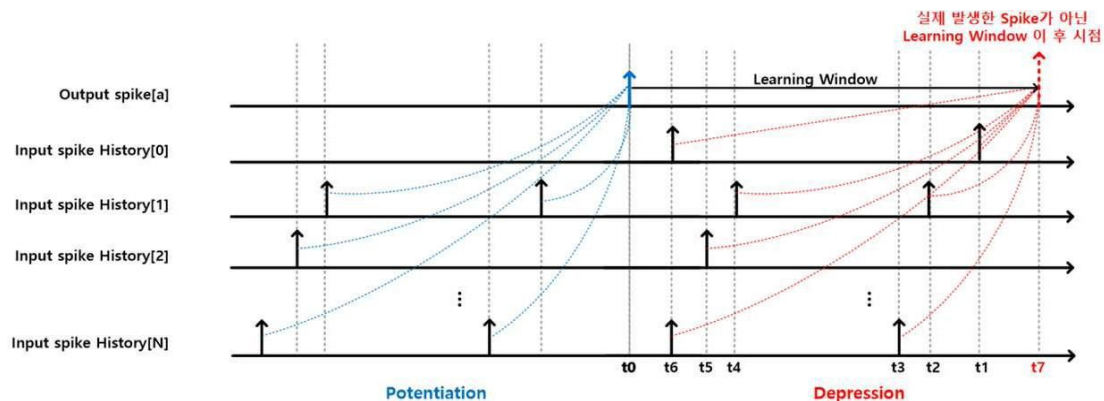
심사관 : 이준상

(54) 발명의 명칭 스파이킹 신경망 학습 장치 및 방법

(57) 요약

본 발명은 STDP 기법 기반 SNN 학습 시에 억제 과정에서 강화 과정과 동일하게 출력 스파이크를 트리거 스파이크로 적용하여, 가중치를 동일한 액세스 방향에서 액세스할 수 있도록 하고, 액세스 횟수를 크게 줄일 수 있으며, 강화 과정과 억제 과정을 동시에 수행할 수 있도록 함으로써 SNN 학습에서 학습 성능은 유지하면서 학습 시간과 전력 소모를 크게 줄일 수 있는 스파이킹 신경망 학습 장치 및 방법을 제공할 수 있다.

대표도



(72) 발명자

김기룡

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C206(신촌동)

최사라

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C206(신촌동)

장효정

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C206(신촌동)

이 발명을 지원한 국가연구개발사업

과제고유번호 2017R1A2B2006679

부처명 과학기술정보통신부

과제관리(전문)기관명 한국연구재단

연구사업명 중견연구자지원사업

연구과제명 Domain Wall Motion 시냅스 기반의 On-Chip 지도-자율 통합학습 뉴로모픽 SoC

개발(3/3)(2017.3.1~2020.2.29)

기 여 율 1/1

과제수행기관명 연세대학교 산학협력단

연구기간 2019.03.01 ~ 2020.02.29

명세서

청구범위

청구항 1

다수의 워드 라인과 다수의 비트 라인에 의해 정의되는 다수의 메모리 셀을 포함하여 가중치를 저장하는 가중치 시냅스부;

입력 데이터에 응답하여 각각 입력 스파이크를 생성하여 상기 다수의 워드 라인 중 대응하는 워드 라인을 활성화하는 다수의 입력 뉴런을 구비하는 프리 뉴런부;

활성화된 워드 라인과 비트 라인이 교차하는 영역에 배치되는 메모리 셀 각각에 저장된 가중치를 인가받아 누적 가산하여, 누적 가산된 가중치가 발화 문턱값 이상이면 출력 스파이크를 발화하는 다수의 출력 뉴런을 구비하는 포스트 뉴런부; 및

상기 다수의 입력 뉴런 각각에서 미리 설정된 학습 윈도우의 크기에 대응하는 기지정된 기간 동안 입력 스파이크 크가 생성된 시간 정보가 저장되는 입력 스파이크 히스토리를 포함하고, 적어도 하나의 출력 뉴런에서 출력 스파이크가 발화되면, 상기 입력 스파이크 히스토리에 저장된 입력 스파이크의 시간 정보에 기반하여 상기 출력 스파이크가 발화된 출력 뉴런에 대응하는 비트 라인 상의 메모리 셀의 가중치를 강화하고, 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 입력 스파이크 히스토리에 저장된 입력 스파이크의 시간 정보에 기반하여 출력 뉴런에 대응하는 비트 라인 상의 메모리 셀의 가중치를 억제하여 업데이트하는 가중치 학습부를 포함하되,

상기 가중치 학습부는

상기 다수의 메모리 셀 중 상기 출력 스파이크가 발화된 출력 뉴런에 대응하는 비트 라인 상의 다수의 메모리 셀 각각의 가중치를 읽고, 상기 가중치를 업데이트하여, 대응하는 비트 라인 상의 다수의 메모리 셀 각각에 재 저장하는 스파이크 신경망 학습 장치.

청구항 2

삭제

청구항 3

제1 항에 있어서, 상기 가중치 학습부는

상기 다수의 입력 뉴런 각각에 대한 다수의 입력 스파이크 히스토리를 포함하고, 상기 출력 스파이크가 발화되면, 상기 다수의 입력 스파이크 히스토리에 각각에 저장된 입력 스파이크의 시간 정보를 기설정된 양자화 스파이크 타이밍 의존 가소성(Spike-timing-dependent plasticity: 이하 STDP) 함수에 대입하여 상기 출력 스파이크가 발화된 출력 뉴런에 대응하는 비트 라인 상의 다수의 메모리 셀 각각에 저장된 가중치를 강화하기 위한 가중치 변화량을 계산하고, 계산된 가중치 변화량을 상기 가중치에 적용하여 대응하는 메모리 셀에 재저장하는 스파이크 신경망 학습 장치.

청구항 4

제1 항에 있어서, 상기 가중치 학습부는

상기 다수의 입력 뉴런 각각에 대한 다수의 입력 스파이크 히스토리를 포함하고, 상기 출력 스파이크가 발화되면, 상기 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 다수의 입력 스파이크 히스토리에 각각에 저장된 입력 스파이크의 시간 정보를 기설정된 양자화 STDP 함수에 대입하여 상기 출력 스파이크가 발화된 출력 뉴런에 대응하는 비트 라인 상의 다수의 메모리 셀 각각에 저장된 가중치를 억제하기 위한 가중치 변화량을 계산하고, 계산된 가중치 변화량을 상기 가중치에 적용하여 대응하는 메모리 셀에 재저장하는 스파이크 신경망 학습 장치.

청구항 5

제4 항에 있어서, 상기 가중치 학습부는

상기 출력 스파이크가 발화되기 이전 기지정된 기간 동안 입력 스파이크가 생성된 시간 정보가 저장되는 입력 스파이크 히스토리와 상기 출력 스파이크가 발화된 이후 기지정된 기간 동안 입력 스파이크가 생성된 시간 정보가 저장되는 애프터 입력 스파이크 히스토리를 각각 포함하는 스파이크 신경망 학습 장치.

청구항 6

제5 항에 있어서, 상기 가중치 학습부는

상기 포스트 뉴런부에서 발화되는 적어도 하나의 출력 스파이크 각각에 대응하여 적어도 하나의 애프터 입력 스파이크 히스토리를 포함하는 스파이크 신경망 학습 장치.

청구항 7

제4 항에 있어서, 상기 가중치 학습부는

상기 가중치를 억제하기 위한 가중치 변화량이 음의 값으로 계산되면, 상기 가중치 변화량을 상기 가중치에 가산하고, 가중치 변화량이 양의 값으로 계산되면, 상기 가중치 변화량을 상기 가중치에 감산하는 스파이크 신경망 학습 장치.

청구항 8

다수의 워드 라인과 다수의 비트 라인에 의해 정의되는 다수의 메모리 셀을 포함하여 가중치를 저장하는 가중치 시냅스부;

입력 데이터에 응답하여 각각 입력 스파이크를 생성하여 상기 다수의 워드 라인 중 대응하는 워드 라인을 활성화하는 다수의 입력 뉴런을 구비하는 프리 뉴런부;

활성화된 워드 라인과 비트 라인이 교차하는 영역에 배치되는 메모리 셀 각각에 저장된 가중치를 인가받아 누적 가산하여, 누적 가산된 가중치가 발화 문턱값 이상이면 출력 스파이크를 발화하는 다수의 출력 뉴런을 구비하는 포스트 뉴런부; 및

상기 다수의 출력 뉴런 각각에서 미리 설정된 학습 윈도우의 크기에 대응하는 기지정된 기간 동안 출력 스파이크가 발화된 시간 정보가 저장되는 출력 스파이크 히스토리를 포함하고, 적어도 하나의 입력 뉴런에서 입력 스파이크가 발생되면, 상기 출력 스파이크 히스토리에 저장된 출력 스파이크의 시간 정보에 기반하여 상기 입력 스파이크가 발생된 입력 뉴런에 대응하는 워드 라인 상의 메모리 셀의 가중치를 강화하고, 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 출력 스파이크 히스토리에 저장된 출력 스파이크의 시간 정보에 기반하여 입력 뉴런에 대응하는 워드 라인 상의 메모리 셀의 가중치를 억제하여 업데이트하는 가중치 학습부를 포함하되,

상기 가중치 학습부는

상기 다수의 메모리 셀 중 상기 입력 스파이크가 발생된 입력 뉴런에 대응하는 워드 라인 상의 다수의 메모리 셀 각각의 가중치를 읽고, 상기 가중치를 업데이트하여, 대응하는 워드 라인 상의 다수의 메모리 셀 각각에 재 저장하는 스파이크 신경망 학습 장치.

청구항 9

삭제

청구항 10

입력 데이터에 응답하여 적어도 하나의 입력 스파이크를 생성하여, 다수의 워드 라인과 다수의 비트 라인에 의해 정의되는 다수의 메모리 셀을 포함하여 가중치를 저장하는 가중치 시냅스부의 상기 다수의 워드 라인 중 상기 입력 스파이크에 대응하는 워드 라인을 활성화하는 단계;

미리 설정된 학습 윈도우의 크기에 대응하는 기지정된 기간 동안 입력 스파이크가 생성된 시간 정보를 입력 스

파이크 히스토리에 저장하는 단계;

활성화된 워드 라인과 상기 다수의 비트 라인이 교차하는 영역에 배치된 다수의 메모리 셀 각각에 저장된 가중치를 인가받아 누적 가산하여, 누적 가산된 가중치가 발화 문턱값 이상이면 출력 스파이크를 발화하는 단계; 및

상기 출력 스파이크가 발화되면, 상기 입력 스파이크 히스토리에 저장된 입력 스파이크의 시간 정보에 기반하여 발화된 상기 출력 스파이크에 대응하는 비트 라인 상의 메모리 셀의 가중치를 강화하고, 상기 출력 스파이크가 발화되고 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 입력 스파이크 히스토리에 저장된 입력 스파이크의 시간 정보에 기반하여 발화된 상기 출력 스파이크에 대응하는 비트 라인 상의 메모리 셀의 가중치를 억제하여 가중치를 업데이트하는 학습 단계를 포함하되,

상기 학습 단계는

상기 다수의 메모리 셀 중 발화된 출력 스파이크에 대응하는 비트 라인 상의 다수의 메모리 셀 각각의 가중치를 읽고, 상기 가중치를 업데이트하여, 대응하는 비트 라인 상의 다수의 메모리 셀 각각에 재저장하여 상기 가중치를 업데이트하는 스파이크 신경망 학습 방법.

청구항 11

삭제

청구항 12

제10 항에 있어서, 상기 입력 스파이크 히스토리에 저장하는 단계는

상기 다수의 워드 라인 각각에 대응하는 입력 스파이크를 구분하여 다수의 입력 스파이크 히스토리를 생성하는 단계; 및

기 지정된 기간 동안 상기 다수의 워드 라인 각각에 대응하는 입력 스파이크가 생성된 시간 정보를 다수의 입력 스파이크 히스토리에 구분하여 저장하는 단계를 포함하는 스파이크 신경망 학습 방법.

청구항 13

제12 항에 있어서, 상기 학습 단계는

상기 출력 스파이크가 발화되면, 상기 다수의 입력 스파이크 히스토리에 각각에 저장된 입력 스파이크의 시간 정보를 기설정된 양자화 스파이크 타이밍 의존 가소성(Spike-timing-dependent plasticity: 이하 STDP) 함수에 대입하여 상기 출력 스파이크에 대응하는 비트 라인 상의 다수의 메모리 셀 각각에 저장된 가중치를 강화하기 위한 가중치 변화량을 계산하고, 계산된 가중치 변화량을 상기 가중치에 적용하여 대응하는 메모리 셀에 재저장하는 강화 단계; 및

상기 출력 스파이크가 발화되면, 상기 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 다수의 입력 스파이크 히스토리에 각각에 저장된 입력 스파이크의 시간 정보를 기설정된 양자화 STDP 함수에 대입하여 상기 출력 스파이크가 발화된 출력 뉴런에 대응하는 비트 라인 상의 다수의 메모리 셀 각각에 저장된 가중치를 억제하기 위한 가중치 변화량을 계산하고, 계산된 가중치 변화량을 상기 가중치에 적용하여 대응하는 메모리 셀에 재저장하는 억제 단계를 포함하는 스파이크 신경망 학습 방법.

청구항 14

제13 항에 있어서, 상기 입력 스파이크 히스토리에 저장하는 단계는

상기 출력 스파이크가 발화되기 이전 기 지정된 기간 동안 입력 스파이크가 생성된 시간 정보를 입력 스파이크 히스토리에 저장하는 단계; 및

상기 출력 스파이크가 발화된 이후 기 지정된 기간 동안 입력 스파이크가 생성된 시간 정보를 상기 입력 스파이크 히스토리과 별도의 애프터 입력 스파이크 히스토리에 저장하는 단계를 포함하는 스파이크 신경망 학습 방법.

청구항 15

제14 항에 있어서, 상기 애프터 입력 스파이크 히스토리에 저장하는 단계는

발화되는 적어도 하나의 출력 스파이크 각각에 대응하여 별도의 애프터 입력 스파이크 히스토리를 생성하는 스파이크 신경망 학습 방법.

청구항 16

제13 항에 있어서, 상기 억제 단계는

상기 가중치를 억제하기 위한 가중치 변화량이 음의 값으로 계산되면, 상기 가중치 변화량을 상기 가중치에 가산하고, 가중치 변화량이 양의 값으로 계산되면, 상기 가중치 변화량을 상기 가중치에 감산하는 스파이크 신경망 학습 방법.

청구항 17

입력 데이터에 응답하여 적어도 하나의 입력 스파이크를 생성하여, 다수의 워드 라인과 다수의 비트 라인에 의해 정의되는 다수의 메모리 셀을 포함하여 가중치를 저장하는 가중치 시냅스부의 상기 다수의 워드 라인 중 상기 입력 스파이크에 대응하는 워드 라인을 활성화하는 단계;

활성화된 워드 라인과 상기 다수의 비트 라인이 교차하는 영역에 배치된 다수의 메모리 셀 각각에 저장된 가중치를 인가받아 누적 가산하여, 누적 가산된 가중치가 발화 문턱값 이상이면 출력 스파이크를 발화하는 단계;

미리 설정된 학습 윈도우의 크기에 대응하는 기지정된 기간 동안 출력 스파이크가 발화된 시간 정보를 출력 스파이크 히스토리에 저장하는 단계; 및

상기 입력 스파이크가 발생되면, 상기 출력 스파이크 히스토리에 저장된 출력 스파이크의 시간 정보에 기반하여 발생된 상기 입력 스파이크에 대응하는 워드 라인 상의 메모리 셀의 가중치를 강화하고, 상기 입력 스파이크가 발생되고 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 출력 스파이크 히스토리에 저장된 출력 스파이크의 시간 정보에 기반하여 발생된 상기 입력 스파이크에 대응하는 워드 라인 상의 메모리 셀의 가중치를 억제하여 가중치를 업데이트하는 학습 단계를 포함하되,

상기 학습 단계는

상기 다수의 메모리 셀 중 발생된 입력 스파이크에 대응하는 워드 라인 상의 다수의 메모리 셀 각각의 가중치를 읽고, 상기 가중치를 적용하여 업데이트하여, 대응하는 워드 라인 상의 다수의 메모리 셀 각각에 재저장하여 상기 가중치를 업데이트하는 스파이크 신경망 학습 방법.

청구항 18

삭제

발명의 설명

기술 분야

[0001] 본 발명은 스파이킹 신경망 학습 장치 및 방법에 관한 것으로, 학습 시 메모리 액세스 횟수를 줄여 에너지 효율성을 향상시키고 학습 속도를 향상시킬 수 있는 스파이킹 신경망 학습 장치 및 방법에 관한 것이다.

배경 기술

[0002] 뉴로모픽(Neuromorphic) 기술은 인간의 신경구조를 하드웨어적으로 모방하기 위한 기술로서, 기존 컴퓨팅 아키텍처가 인지처리 기능을 수행함에 있어 인간에 비해 효율성이 매우 낮고 전력 소모가 크다는 한계를 극복하기 위해 제안되었다.

[0003] 뉴로모픽 기술에는 대표적으로 스파이킹 신경망(Spiking Neural Network: 이하 SNN)이 있다. SNN은 인간의 두뇌가 뉴런(Neuron)-시냅스(Synapse) 구조를 가지고 있고, 뉴런과 뉴런을 잇는 시냅스가 스파이크 형태의 전기 신호로 정보를 전달한다는 특징을 모방하여 고안된 신경망이다. 이러한 SNN은 스파이크 신호가 전송되는 시간

차에 기초하여 정보를 처리한다.

- [0004] SNN 또한 인공 신경망의 일종으로 가중치를 업데이트하기 위한 학습이 수행되어야 하며, SNN의 학습에는 일반적으로 스파이크 타이밍 의존 가소성(Spike-timing-dependent plasticity: 이하 STDP) 기법이 이용되고 있다. STDP 기법은 기지정된 크기의 학습 윈도우 기간 내에 발생한 모든 입력 스파이크와 출력 스파이크의 발생 시간 정보를 기반으로 시냅스에 해당하는 가중치를 조절하기 위한 가중치 변화량을 계산하여 가중치를 업데이트함으로써 SNN을 학습시킨다.
- [0005] STDP 기법에서 학습은 가중치를 증가시키는 강화 과정(Potential Process)과 가중치를 감소시키는 억제 과정(Depression Process)으로 구성된다.
- [0006] 강화 과정에서는 출력 스파이크가 트리거 스파이크로 동작하여, 출력 스파이크가 발화되면 과거 기지정된 학습 윈도우 기간 동안 입력 스파이크에 대한 시간 정보가 저장된 입력 스파이크 히스토리(input spike history)에 따라 가중치를 증가시킨다. 반면 억제 과정에서는 입력 스파이크가 트리거 스파이크로 동작하여, 입력 스파이크가 발화되면, 과거 기지정된 학습 윈도우 기간 동안 출력 스파이크에 대한 시간 정보가 저장된 출력 스파이크 히스토리(output spike history)에 따라 가중치를 감소시킨다.
- [0007] 이와 같이 STDP 기법에서는 출력 스파이크가 발생되면 강화 과정에 의해 입력 스파이크 히스토리를 기반으로 발생된 출력 스파이크에 연관된 입력 스파이크를 판별하고, 출력 스파이크가 발생된 출력 뉴런과 연관된 입력 스파이크가 발생된 입력 뉴런을 연결하는 시냅스에 해당하는 가중치를 모두 읽고 업데이트하여 재저장해야 한다. 또한 입력 스파이크가 발생되면 억제 과정에 의해 출력 스파이크 히스토리를 기반으로 발생된 입력 스파이크에 연관된 출력 스파이크를 판별하고, 입력 스파이크가 발생된 입력 뉴런과 연관된 출력 스파이크가 발생된 출력 뉴런을 연결하는 시냅스에 해당하는 가중치를 모두 읽고 업데이트하여 재저장해야 한다.
- [0008] 즉 STDP 기법을 이용하여 SNN을 학습시키는 경우, 스파이크(입력 스파이크 또는 출력 스파이크)가 발생될 때마다 연관된 가중치를 읽고 업데이트하여 재저장하는 과정을 반복해야 하므로 매우 비효율적이라는 문제가 있다. 특히 SNN에서 메모리 어레이와 유사한 구조를 갖고 다수의 가중치가 저장되는 가중치 시냅스부에 대해 강화 과정과 억제 과정 각각에서 가중치를 읽고 재저장하기 위한 액세스 방향이 서로 상이하여 효율성이 더욱 낮아지게 되는 문제가 있다.

선행기술문헌

특허문헌

- [0009] (특허문헌 0001) 한국 공개 특허 제10-2017-0080440호 (2017.07.10 공개)

발명의 내용

해결하려는 과제

- [0010] 본 발명의 목적은 스파이크 타이밍 의존 가소성 기법 기반 학습시에 저장된 가중치에 대한 액세스 횟수를 줄여 효율성을 높일 수 있는 스파이킹 신경망 학습 장치 및 방법을 제공하는데 있다.
- [0011] 본 발명의 다른 목적은 스파이크 타이밍 의존 가소성 기법의 억제 과정에서 출력 스파이크를 트리거 스파이크로 적용함으로써 강화 과정과 동일한 가중치 액세스 방향으로 액세스 하도록 하여 효율성을 크게 높일 수 있는 스파이킹 신경망 학습 장치 및 방법을 제공하는데 있다.
- [0012] 본 발명의 또 다른 목적은 STDP 기법의 강화 과정과 억제 과정을 동시에 수행하여 가중치에 대한 학습 시간을 크게 단축시킬 수 있는 스파이킹 신경망 학습 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

- [0013] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 스파이킹 신경망 학습 장치는 다수의 워드 라인과 다수의 비트 라인에 의해 정의되는 다수의 메모리 셀을 포함하여 가중치를 저장하는 가중치 시냅스부; 입력 데이터에 응답하여 각각 입력 스파이크를 생성하여 상기 다수의 워드 라인 중 대응하는 워드 라인을 활성화하는 다수의 입력 뉴런을 구비하는 프리 뉴런부; 각각 상기 다수의 비트 라인 중 대응하는 비트 라인과 활성화된 워드 라

인이 교차하는 영역에 배치되는 메모리 셀에 저장된 가중치를 인가받아 누적 가산하여, 누적 가산된 가중치가 발화 문턱값 이상이면 출력 스파이크를 발화하는 다수의 출력 뉴런을 구비하는 포스트 뉴런부; 및 상기 다수의 입력 뉴런 각각에서 미리 설정된 학습 윈도우의 크기에 대응하는 기지정된 기간 동안 입력 스파이크가 생성된 시간 정보가 저장되는 입력 스파이크 히스토리를 포함하고, 적어도 하나의 출력 뉴런에서 출력 스파이크가 발화되면, 상기 입력 스파이크 히스토리에 저장된 입력 스파이크의 시간 정보에 기반하여 상기 출력 스파이크가 발화된 출력 뉴런에 대응하는 비트 라인 상의 메모리 셀의 가중치를 강화하고, 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 입력 스파이크 히스토리에 저장된 입력 스파이크의 시간 정보에 기반하여 출력 뉴런에 대응하는 비트 라인 상의 메모리 셀의 가중치를 억제하여 업데이트하는 가중치 학습부를 포함한다.

[0014] 상기 가중치 학습부는 상기 다수의 메모리 셀 중 상기 출력 스파이크가 발화된 출력 뉴런에 대응하는 비트 라인 상의 다수의 메모리 셀 각각의 가중치를 읽고, 상기 가중치를 업데이트하여, 대응하는 비트 라인 상의 다수의 메모리 셀 각각에 재저장할 수 있다.

[0015] 상기 가중치 학습부는 상기 다수의 입력 뉴런 각각에 대한 다수의 입력 스파이크 히스토리를 포함하고, 상기 출력 스파이크가 발화되면, 상기 다수의 입력 스파이크 히스토리에 각각에 저장된 입력 스파이크의 시간 정보를 기설정된 양자화 스파이크 타이밍 의존 가소성(Spike-timing-dependent plasticity: 이하 STDP) 함수에 대입하여 상기 출력 스파이크가 발화된 출력 뉴런에 대응하는 비트 라인 상의 다수의 메모리 셀 각각에 저장된 가중치를 강화하기 위한 가중치 변화량을 계산하고, 계산된 가중치 변화량을 상기 가중치에 적용하여 대응하는 메모리 셀에 재저장할 수 있다.

[0016] 상기 가중치 학습부는 상기 다수의 입력 뉴런 각각에 대한 다수의 입력 스파이크 히스토리를 포함하고, 상기 출력 스파이크가 발화되면, 상기 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 다수의 입력 스파이크 히스토리에 각각에 저장된 입력 스파이크의 시간 정보를 기설정된 양자화 STDP 함수에 대입하여 상기 출력 스파이크가 발화된 출력 뉴런에 대응하는 비트 라인 상의 다수의 메모리 셀 각각에 저장된 가중치를 억제하기 위한 가중치 변화량을 계산하고, 계산된 가중치 변화량을 상기 가중치에 적용하여 대응하는 메모리 셀에 재저장할 수 있다.

[0017] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 스파이킹 신경망 학습 장치는 다수의 워드 라인과 다수의 비트 라인에 의해 정의되는 다수의 메모리 셀을 포함하여 가중치를 저장하는 가중치 시냅스부; 입력 데이터에 응답하여 각각 입력 스파이크를 생성하여 상기 다수의 워드 라인 중 대응하는 워드 라인을 활성화하는 다수의 입력 뉴런을 구비하는 프리 뉴런부; 각각 상기 다수의 비트 라인 중 대응하는 비트 라인과 활성화된 워드 라인이 교차하는 영역에 배치되는 메모리 셀에 저장된 가중치를 인가받아 누적 가산하여, 누적 가산된 가중치가 발화 문턱값 이상이면 출력 스파이크를 발화하는 다수의 출력 뉴런을 구비하는 포스트 뉴런부; 및 상기 다수의 출력 뉴런 각각에서 미리 설정된 학습 윈도우의 크기에 대응하는 기지정된 기간 동안 출력 스파이크가 발화된 시간 정보가 저장되는 출력 스파이크 히스토리를 포함하고, 적어도 하나의 입력 뉴런에서 입력 스파이크가 발생되면, 상기 출력 스파이크 히스토리에 저장된 출력 스파이크의 시간 정보에 기반하여 상기 입력 스파이크가 발생한 입력 뉴런에 대응하는 워드 라인 상의 메모리 셀의 가중치를 강화하고, 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 출력 스파이크 히스토리에 저장된 출력 스파이크의 시간 정보에 기반하여 입력 뉴런에 대응하는 워드 라인 상의 메모리 셀의 가중치를 억제하여 업데이트하는 가중치 학습부를 포함한다.

[0018] 상기 목적을 달성하기 위한 본 발명의 또 다른 실시예에 따른 스파이킹 신경망 학습 방법은 입력 데이터에 응답하여 적어도 하나의 입력 스파이크를 생성하여, 다수의 워드 라인과 다수의 비트 라인에 의해 정의되는 다수의 메모리 셀을 포함하여 가중치를 저장하는 가중치 시냅스부의 상기 다수의 워드 라인 중 대응하는 워드 라인을 활성화하는 단계; 미리 설정된 학습 윈도우의 크기에 대응하는 기지정된 기간 동안 입력 스파이크가 생성된 시간 정보를 입력 스파이크 히스토리에 저장하는 단계; 활성화된 워드 라인과 상기 다수의 비트 라인이 교차하는 영역에 배치된 다수의 메모리 셀 각각에 저장된 가중치를 인가받아 누적 가산하여, 누적 가산된 가중치가 발화 문턱값 이상이면 출력 스파이크를 발화하는 단계; 및 상기 출력 스파이크가 발화되면, 상기 입력 스파이크 히스토리에 저장된 입력 스파이크의 시간 정보에 기반하여 발화된 상기 출력 스파이크에 대응하는 비트 라인 상의 메모리 셀의 가중치를 강화하고, 상기 출력 스파이크가 발화되고 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 입력 스파이크 히스토리에 저장된 입력 스파이크의 시간 정보에 기반하여 발화된 상기 출력 스파이크에 대응하는 비트 라인 상의 메모리 셀의 가중치를 억제하여 가중치를 업데이트하는 학습 단계를 포함한다.

[0019] 상기 목적을 달성하기 위한 본 발명의 또 다른 실시예에 따른 스파이킹 신경망 학습 방법은 입력 데이터에 응답하여 적어도 하나의 입력 스파이크를 생성하여, 다수의 워드 라인과 다수의 비트 라인에 의해 정의되는 다수의

메모리 셀을 포함하여 가중치를 저장하는 가중치 시냅스부의 상기 다수의 워드 라인 중 대응하는 워드 라인을 활성화하는 단계; 활성화된 워드 라인과 상기 다수의 비트 라인이 교차하는 영역에 배치된 다수의 메모리 셀 각각에 저장된 가중치를 인가받아 누적 가산하여, 누적 가산된 가중치가 발화 문턱값 이상이면 출력 스파이크를 발화하는 단계; 미리 설정된 학습 윈도우의 크기에 대응하는 기지정된 기간 동안 출력 스파이크가 발화된 시간 정보를 출력 스파이크 히스토리에 저장하는 단계; 및 상기 입력 스파이크가 발생되면, 상기 출력 스파이크 히스토리에 저장된 출력 스파이크의 시간 정보에 기반하여 발생한 상기 입력 스파이크에 대응하는 워드 라인 상의 메모리 셀의 가중치를 강화하고, 상기 입력 스파이크가 발생되고 학습 윈도우의 크기에 의해 지정된 기간 이후, 상기 출력 스파이크 히스토리에 저장된 출력 스파이크의 시간 정보에 기반하여 발생한 상기 입력 스파이크에 대응하는 워드 라인 상의 메모리 셀의 가중치를 억제하여 가중치를 업데이트하는 학습 단계를 포함한다.

발명의 효과

[0020] 따라서, 본 발명의 실시예에 따른 스파이킹 신경망 학습 장치 및 방법은 STDP 기법 기반 SNN 학습 시에 억제 과정에서 강화 과정과 동일하게 출력 스파이크를 트리거 스파이크로 적용하여, 가중치를 동일한 액세스 방향에서 액세스할 수 있도록 하고, 액세스 횟수를 크게 줄일 수 있으며, 강화 과정과 억제 과정을 동시에 수행할 수 있도록 한다. 그러므로 SNN 학습에서 학습 성능은 유지하면서 학습 시간과 전력 소모를 크게 줄일 수 있다.

도면의 간단한 설명

[0021] 도 1은 본 발명의 일 실시예에 따른 스파이킹 신경망의 개략적 구조를 나타낸다.
 도 2는 양자화 스파이크 타이밍 의존 가소성 함수의 일례를 나타낸다.
 도 3 및 도 4는 본 발명의 일 실시예에 따른 가중치 학습부가 STDP 기법에 기반하여 수행하는 강화 과정을 설명하기 위한 도면이다.
 도 5 및 도 6은 가중치 학습부가 STDP 기법에 기반하여 수행하는 억제 과정을 설명하기 위한 도면이다.
 도 7은 가중치 학습부가 STDP 기법에 기반하여 강화 과정 및 억제 과정을 수행할 때 가중치가 업데이트되는 메모리 셀을 나타낸다.
 도 8 및 도 9는 본 발명의 일 실시예에 따른 STDP 기법에 기반하여 수행하는 억제 과정을 설명하기 위한 도면이다.
 도 10은 본 발명의 일 실시예에 따른 스파이킹 신경망 학습 방법을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0022] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시예에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.

[0023] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.

[0024] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[0025] 도 1은 본 발명의 일 실시예에 따른 스파이킹 신경망의 개략적 구조를 나타낸다.

[0026] 도 1을 참조하면, 스파이킹 신경망(Spiking Neural Network: 이하 SNN)은 프리 뉴런부(100)와 가중치 시냅스부(200), 포스트 뉴런부(300) 및 가중치 학습부(400)를 포함한다.

[0027] 프리 뉴런부(100)는 입력 데이터를 기지정된 방식에 따라 인코딩하여 다수의 입력 스파이크를 발생하고, 발생한 다수의 입력 스파이크를 가중치 시냅스부(200)로 전달한다. 인간의 두뇌는 각종 감각 기관에서 감지된 정보를 전기적 신호의 형태로 전달받고, 뉴런과 뉴런을 잇는 시냅스가 스파이크 형태의 전기 신호로 뉴런 사이에 정보가 전달되도록 한다. 이에 인간의 두뇌를 모방하도록 구성되는 SNN의 프리 뉴런부(100) 또한 입력 데이터를 기

지정된 방식에 따라 입력 스파이크로 변환하여 가중치 시냅스부(200)로 전달한다.

- [0028] 프리 뉴런부(100)는 다수의 입력 뉴런으로 구성되어, 각각 대응하는 입력 데이터의 값에 따라 입력 스파이크를 발생할 수 있다. 일례로 프리 뉴런부(100)의 다수의 입력 뉴런은 이미지에서 대응하는 각 픽셀의 픽셀값을 입력 데이터로 인가받고, 인가된 픽셀값에 따라 입력 스파이크를 발생할 수 있다. 이때, 입력 뉴런은 하나의 입력 데이터에 대해 시간차를 두고 다수의 입력 스파이크를 발생할 수도 있다. 입력 뉴런이 입력 데이터를 입력 스파이크로 변환하는 방법은 다양하게 연구되어 있으며 일례로 확률에 기반하여 입력 스파이크를 생성할 수 있다.
- [0029] 입력 뉴런은 입력 스파이크를 이진 데이터의 형식으로 발생할 수 있다. 일례로 입력 뉴런은 입력 스파이크가 발생되면 1의 값을 출력하고, 입력 스파이크가 발생되지 않으면 0의 값을 출력하도록 설정될 수 있다.
- [0030] 프리 뉴런부(100)의 다수의 입력 뉴런은 입력 스파이크를 인가받도록 구성될 수도 있으며, 이 경우 다수의 입력 뉴런 각각은 인가된 다수의 입력 스파이크를 버퍼링하여 출력하는 버퍼로 구현될 수 있다.
- [0031] 가중치 시냅스부(200)는 프리 뉴런부(100)의 다수의 입력 뉴런에서 인가되는 다수의 입력 스파이크 각각에 대응하는 가중치를 포스트 뉴런부(300)로 전달한다. 가중치 시냅스부(200)에는 다수의 가중치가 저장되며, 저장된 다수의 가중치는 SNN 학습 시에 가중치 학습부(400)에 의해 업데이트될 수 있다.
- [0032] 가중치 시냅스부(200)는 SNN에서 두뇌의 시냅스를 모방한 구성으로, 다수의 가중치를 저장하는 메모리 어레이와 유사한 구조를 가질 수 있다. 일례로 가중치 시냅스부(200)는 SRAM(Static Random access memory) 등으로 구현될 수 있다.
- [0033] 가중치 시냅스부(200)는 프리 뉴런부(100)의 다수의 입력 뉴런 중 대응하는 입력 뉴런에서 입력 스파이크가 발생되면, 활성화되는 다수의 워드 라인(WL)과 다수의 워드 라인(WL)에 교차하는 다수의 비트 라인(BL) 및 다수의 워드 라인(WL)과 다수의 비트 라인(BL)에 의해 정의되는 다수의 메모리 셀(MC)을 포함하여 구성될 수 있다. 그리고 다수의 메모리 셀(MC) 각각에는 가중치가 저장된다.
- [0034] 여기서 워드 라인(WL)은 프리 뉴런부(100)의 입력 뉴런의 개수에 대응하는 개수로 배치될 수 있으며, 비트 라인(BL)은 포스트 뉴런부(300)의 출력 뉴런의 개수에 대응하는 개수로 배치될 수 있다. 예를 들어 프리 뉴런부(100)의 입력 뉴런의 개수가 N개이면, 가중치 시냅스부(200)에는 N개의 워드 라인(WL)이 배치될 수 있으며, 포스트 뉴런부(300)의 출력 뉴런의 개수가 M개이면, 가중치 시냅스부(200)에는 M개의 비트 라인(BL)이 배치될 수 있다.
- [0035] 다만 도 1에서는 설명의 편의를 위하여 하나의 워드 라인(WL)과 교차하는 하나의 비트 라인(BL)에 의해 정의되는 하나의 메모리 셀(MC)에 가중치가 저장되는 것으로 도시하였다. 그러나 실제 가중치는 기지정된 k비트의 이진 데이터 형식을 가지며, 이에 k개의 메모리 셀에 각 비트값이 저장되어야 한다. 이에 도 1에서 다수의 비트 라인 각각은 실제 k개의 비트 라인의 집합이며, 하나의 워드 라인(WL)과 교차하는 k개의 비트 라인에 배치되는 k개의 메모리 셀에 k비트의 가중치가 저장된다. 따라서 N개의 입력 뉴런과 M개의 출력 뉴런을 포함하고, k비트의 가중치를 저장하는 SNN의 가중치 시냅스부(200)에는 실제로 N개의 워드 라인과 N*k개의 비트 라인이 배치되어, $M * N * k$ 비트의 가중치를 저장하도록 구성될 수 있다.
- [0036] 가중치 시냅스부(200)는 다수의 입력 뉴런 중 특정 입력 뉴런에서 입력 스파이크가 발생되면, 대응하는 워드 라인(WL)이 활성화되고, 활성화된 워드 라인(WL)과 다수의 비트 라인(BL)이 교차하는 위치의 메모리 셀(MC)에 저장된 가중치를 포스트 뉴런부(300)로 출력한다. 즉 입력 스파이크에 대응하는 가중치를 포스트 뉴런부(300)로 출력하도록 구성될 수 있다.
- [0037] 가중치 시냅스부(200)는 프리 뉴런부(100)으로부터 시간차를 두고 발생되어 인가되는 다수의 입력 스파이크 각각에 응답하여 대응하는 가중치를 반복적으로 포스트 뉴런부(300)로 출력한다.
- [0038] 포스트 뉴런부(300)는 가중치 시냅스부(200)에서 출력되는 가중치의 누적 합에 기초하여 다수의 출력 스파이크를 발화한다. 포스트 뉴런부(300)는 다수의 출력 뉴런을 포함할 수 있으며, 다수의 출력 뉴런 각각은 LIF(Leaky-Integrate and Fire) 뉴런 모델에 기반하여 구성될 수 있다.
- [0039] SNN에서는 주로 인간 신경망의 동작 특성을 모방한 LIF 뉴런 모델이 이용되며, LIF 뉴런 모델은 입력되는 입력 스파이크에 따른 가중치를 누적(Integrate)하고, 누적된 가중치가 기지정된 기준 레벨 이상이면, 발화(Fire)하여 다음 뉴런으로 전달할 새로운 입력 스파이크를 생성하며, 입력 스파이크가 전달되지 않으면 누적된 입력 스

파이크의 세기가 누설(Leaky)되어 시간이 흐를수록 점차 약해지도록 구성된 모델이다.

- [0040] 이에 LIF 뉴런 모델에 기반하여 구성되는 포스트 뉴런부(300)의 다수의 출력 뉴런 각각은 가중치 시냅스부(200)의 대응하는 비트 라인(BL)으로부터 인가되는 가중치를 합하여 누적하고, 누적 가중치가 기지정된 발화 문턱값 이상이면, 출력 스파이크를 생성하여 출력하도록 구성된다. 그리고 다수의 출력 뉴런은 각각은 시간의 흐름에 따라 누적 가중치를 점차 감소시키도록 구성될 수 있다.
- [0041] 가중치 학습부(400)는 SNN을 학습시키기 위한 학습 장치로서, 스파이크 타이밍 의존 가소성(Spike-timing-dependent plasticity: 이하 STDP) 기법에 따라 가중치 시냅스부(200)의 다수의 가중치를 업데이트함으로써 SNN을 학습시킨다. 이때 디지털 회로로 구현되는 가중치 학습부(400)는 도 2에 도시된 바와 같은 양자화 STDP 함수에 기반하여 가중치 변화량을 획득하고, 획득된 가중치 변화량을 가중치에 가산하여 업데이트함으로써 학습을 수행한다.
- [0042] 도 2는 양자화 스파이크 타이밍 의존 가소성 함수의 일예를 나타낸다.
- [0043] 도 2에 도시된 양자화 STDP 함수 그래프에서, x 축은 출력 스파이크(OS)가 발생된 시점(t_{OS})으로부터 입력 스파이크(IS)가 발생된 시점(t_{IS}) 사이의 시간차를 나타내고, y 축은 가중치(w)의 가중치 변화량(Δw)을 나타낸다. 도 2의 양자화 STDP 함수 그래프는 x 축이 입력 스파이크(IS)가 발생된 시점(t_{IS}) 사이의 시간차를 나타내므로 입력 스파이크(IS)가 인가된 시점과 출력 스파이크(OS)가 인가된 시점 모두를 기준으로 적용될 수 있다. 여기서 x 축의 값은 디지털 회로로 구현되는 SNN에서 스파이크를 입출력할 수 있는 단위 시간으로 사이클(cycle)(또는 클럭)이다. 그리고 도 4에서 y축 방향의 값은 디지털 값을 갖는 가중치(w)를 조절할 수 있는 기지정된 변화 단위를 나타낸다.
- [0044] 만일 (0, 0)인 위치가 출력 스파이크(OS)가 발생된 시점이라면, x 축에서 양의 값을 갖는 영역($x > 0$)은 출력 스파이크(OS)가 발생되기 이전 입력 스파이크(IS)가 발생된 시점에 따른 가중치 변화량(Δw)을 나타내고, 음의 값($x < 0$)을 갖는 영역은 출력 스파이크(OS)가 발생된 이후 입력 스파이크(IS)가 발생된 시점에 따른 가중치 변화량(Δw)을 나타낸다.
- [0045] 반면 (0, 0)인 위치가 입력 스파이크(IS)가 발생된 시점이라면, x 축에서 양의 값($x > 0$)을 갖는 영역은 입력 스파이크(IS)가 발생된 이후 출력 스파이크(OS)가 발생된 시점에 따른 가중치 변화량(Δw)을 나타내고, 음의 값($x < 0$)을 갖는 영역은 입력 스파이크(IS)가 발생되기 이전 출력 스파이크(OS)가 발생된 시점에 따른 가중치 변화량(Δw)을 나타낸다.
- [0046] 즉 가중치 학습부(400)는 양자화 STDP 함수에 따라 기준이 되는 트리거 스파이크로부터 기지정된 기간 동안 발생한 스파이크 중 더 가까운 시간에 발생한 스파이크에 대해 더 높은 비중을 부여하여 가중치 변화량(Δw)을 계산하고, 계산된 가중치 변화량(Δw)을 기존의 가중치(w)에 더하여 가중치를 업데이트한다.
- [0047] 도 2에서는 양자화 STDP 함수의 일예로 20(또는 -20) 사이클과 60(또는 -60) 사이클을 기준으로 가중치 변화량(Δw)의 1 단위로 가변되는 것으로 도시하였으나, 가중치 변화량(Δw)이 조절되는 사이클 수와 가변되는 단위는 다양하게 조절될 수 있다. 또한 양자화 구간도 더욱 세분화될 수도 있다.
- [0048] 도 3 및 도 4는 본 발명의 일 실시예에 따른 가중치 학습부가 STDP 기법에 기반하여 수행하는 강화 과정을 설명하기 위한 도면이다.
- [0049] 도 3은 STDP 기법에 따른 가중치 강화 과정에서 하나의 출력 스파이크(OS)가 발화되면, 발화된 출력 스파이크(OS)에 대응하여 강화되어야 하는 가중치가 저장된 메모리 셀(MC)을 나타내고, 도 4는 발화된 출력 스파이크(OS)와 발화된 출력 스파이크(OS)에 대응하는 다수의 입력 스파이크(IS) 사이의 시간차를 나타낸다.
- [0050] 도 3 및 도 4를 참조하면, 포스트 뉴런부(300)의 다수의 출력 뉴런(OUN) 중 특정 출력 뉴런에서 출력 스파이크(OS[a])가 발화되면, 가중치 학습부(400)는 기지정된 학습 윈도우의 크기(도 2를 참조하면 60 사이클)에 따라 이전 기지정된 기간 동안 입력 스파이크(IS)가 발생된 정보가 저장된 입력 스파이크 히스토리를 분석한다. 즉 출력 스파이크(OS[a])가 트리거 스파이크로 동작하여 이전 다수의 입력 뉴런(INN) 각각에서 발생한 입력 스파이크를 분석한다. 여기서 입력 스파이크 히스토리는 도 3에 도시된 바와 같이 다수의 입력 뉴런(INN) 각각에 대응하여 다수개로 획득될 수 있다.
- [0051] 가중치 학습부(400)는 입력 스파이크 히스토리를 분석함으로써, 출력 스파이크(OS[a])가 발화되기 이전 다수의 워드 라인(WL) 각각으로 입력 스파이크(IS)가 인가된 시점을 확인함으로써, 출력 스파이크(OS[a])가 발화된 출

력 뉴런(OUN)에 연결된 비트 라인(BL)과 다수의 워드 라인(WL) 각각이 교차하는 위치에 배치된 다수의 메모리 셀(MC) 각각에 저장된 가중치(w)를 강화하기 위한 다수의 가중치 변화량(Δw)을 계산한다. 그리고 계산된 다수의 가중치 변화량(Δw)을 각각을 대응하는 메모리 셀(MC)에 저장된 가중치(w)에 가산하여 가중치(w)를 업데이트하고, 업데이트된 가중치(w)를 다시 대응하는 메모리 셀(MC)에 저장한다.

- [0052] 즉 출력 스파이크(OS[a])가 발화된 출력 뉴런(OUN)에 연결된 비트 라인(BL)을 기준으로 열 방향으로 다수의 메모리 셀(MC) 각각에 저장된 가중치(w)를 업데이트한다.
- [0053] 이때 다수의 가중치 변화량(Δw) 각각은 도 2에 도시된 양자와 STDP 함수에 기반하여 대응하는 워드 라인(WL)을 통해 입력 스파이크(IS)가 인가된 시점과 개수에 따라 서로 다르게 계산될 수 있다.
- [0054] 따라서 적어도 하나의 출력 뉴런(OUN)에서 출력 스파이크(OS)가 발화되면, 입력 스파이크 히스토리에 저장된 적어도 하나의 입력 스파이크(IS)가 인가된 모든 워드 라인(WL)과 적어도 하나의 출력 뉴런(OUN)에 대응하는 적어도 하나의 비트 라인(BL)이 교차하는 위치에 배치된 다수의 메모리 셀(MC)에 저장된 가중치가 업데이트 된다. 일례로 도 3 및 도 4에 도시된 바와 같이, 하나의 출력 뉴런(OUN)에서 출력 스파이크(OS)가 발화되고, 입력 스파이크 히스토리에 N개의 워드 라인(WL) 모두에 대해 인가된 입력 스파이크(IS)가 존재하면, 가중치 학습부(400)는 비트 라인(BL) 상에 배치되는 N개의 메모리 셀(MC)의 가중치를 업데이트 한다.
- [0055] 도 5 및 도 6은 가중치 학습부가 STDP 기법에 기반하여 수행하는 억제 과정을 설명하기 위한 도면이다.
- [0056] 도 5 및 도 6은 후술하는 본 실시예에 따른 STDP 기법 기반 가중치 억제 과정과의 비교 설명을 위해, 우선 일반적인 가중치 학습부가 수행하는 억제 과정을 도시하였다.
- [0057] 도 5는 STDP 기법에 따른 가중치 억제 과정에서 하나의 입력 스파이크(IS)가 발생되면, 발생한 입력 스파이크(IS)에 대응하여 억제되어야 하는 가중치가 저장된 메모리 셀(MC)을 나타내고, 도 6은 발생한 입력 스파이크(IS)와 입력 스파이크(IS)에 대응하는 다수의 출력 스파이크(OS) 사이의 시간차를 나타낸다.
- [0058] 도 5 및 도 6을 참조하면, 프리 뉴런부(100)의 다수의 입력 뉴런(INN) 중 특정 입력 뉴런에서 입력 스파이크(IS[b])가 발생되면, 기존의 가중치 학습부는 기지정된 학습 윈도우의 크기(도 2를 참조하면 60 사이클)에 따라 이전 기지정된 시간 동안 출력 스파이크(OS)가 발생한 정보가 저장된 출력 스파이크 히스토리를 분석한다. 즉 입력 스파이크(IS[b])가 트리거 스파이크로 동작하여 이전 발생한 출력 스파이크를 분석한다. 여기서 출력 스파이크 히스토리는 도 5에 도시된 바와 같이 다수의 출력 뉴런(OUN) 각각에 대응하여 다수개로 획득될 수 있다.
- [0059] 가중치 학습부는 출력 스파이크 히스토리를 분석함으로써, 입력 스파이크(IS[b])가 발생되기 이전 다수의 출력 뉴런(OUN) 각각에서 출력 스파이크(OS)가 발화된 시점을 확인함으로써, 입력 스파이크(IS[b])가 발생한 입력 뉴런(INN)에 연결된 워드 라인(WL)과 다수의 비트 라인(BL) 각각이 교차하는 위치에 배치된 다수의 메모리 셀(MC) 각각에 저장된 가중치(w)를 억제하기 위한 다수의 가중치 변화량(Δw)을 계산한다. 그리고 계산된 다수의 가중치 변화량(Δw)을 각각을 대응하는 메모리 셀(MC)에 저장된 가중치(w)에 감산하여 가중치(w)를 업데이트하고, 업데이트된 가중치(w)를 다시 대응하는 메모리 셀(MC)에 저장한다.
- [0060] 즉 입력 스파이크(IS[b])가 발생한 입력 뉴런(INN)에 연결된 워드 라인(WL)을 기준으로 행 방향으로 다수의 메모리 셀(MC) 각각 저장된 가중치(w)를 업데이트한다.
- [0061] 이때 다수의 가중치 변화량(Δw) 또한 도 2에 도시된 양자와 STDP 함수에 기반하여 대응하는 출력 뉴런(OUN)에서 출력 스파이크(OS)가 발화된 시점과 개수에 따라 서로 다르게 계산될 수 있다.
- [0062] 따라서 적어도 하나의 입력 뉴런(INN)에서 입력 스파이크(IS[b])가 발생되면, 적어도 하나의 입력 뉴런(INN)에 대응하는 적어도 하나의 워드 라인(WL)과 출력 스파이크 히스토리에 저장된 적어도 하나의 출력 스파이크(OS)가 발화된 모든 출력 뉴런(OUN)에 대응하는 비트 라인(BL)이 교차하는 위치에 배치된 다수의 메모리 셀(MC)에 저장된 가중치가 업데이트 된다.
- [0063] 도 7은 가중치 학습부가 STDP 기법에 기반하여 강화 과정 및 억제 과정을 수행할 때 가중치가 업데이트되는 메모리 셀을 나타낸다.
- [0064] 도 7에 도시된 바와 같이, 기존에는 가중치 학습부가 STDP 기법에 기반하여 강화 과정 및 억제 과정을 수행하는 경우, 상기한 바와 같이, 강화 과정에서는 열 방향으로 메모리 셀(MC)의 가중치가 강화되는 반면, 억제 과정에서는 행 방향으로 메모리 셀(MC)의 가중치가 억제된다. 이는 강화 과정과 억제 과정이 모두 트리거 스파이크가 되는 출력 스파이크(OS) 또는 입력 스파이크(IS)를 기준으로 이전 입력 스파이크(IS)의 발생 시간 정보가 저장

된 입력 스파이크 히스토리 또는 이전 출력 스파이크(OS)의 발생 시간 정보가 저장된 출력 스파이크 히스토리를 기반으로 가중치 변화량(Δw)을 계산하여 가중치를 업데이트하기 때문이다.

[0065] 따라서 가중치 학습부는 출력 스파이크(OS) 또는 입력 스파이크(IS)가 발생될 때마다 대응하는 메모리 셀(MC)의 가중치를 읽어야 하고 업데이트한 후 다시 재저장해야 한다. 뿐만 아니라 출력 스파이크(OS)가 발화되면, 열 방향으로 배열된 다수의 메모리 셀(MC)로부터 가중치를 읽는 반면, 입력 스파이크(IS)가 발생되면, 행 방향으로 배열된 다수의 메모리 셀(MC)로부터 가중치를 읽어야 한다. 강화 과정과 억제 과정에서 가중치를 업데이트하는 방향이 서로 상이하므로, 가중치 학습부는 매우 빈번하게 가중치를 읽고 재저장해야 할 뿐만 아니라, 두 방향의 액세스가 가능해야 하므로, 가중치를 읽고 재저장하기 위한 구조 및 과정이 매우 복잡해지게 되며, 가중치를 업데이트하기 위해 요구되는 시간 및 전력 소비가 매우 크다는 문제가 있다.

[0066] 도 8 및 도 9는 본 발명의 일 실시예에 따른 가중치 학습부가 STDP 기법에 기반하여 수행하는 억제 과정을 설명하기 위한 도면이다.

[0067] 도 8 및 도 9를 참조하면, 본 실시예에 따른 가중치 학습부(400)는 억제 과정에서 도 6 및 도 7과 달리 입력 스파이크(IS[b])가 아닌 출력 스파이크(OS[c])를 트리거 스파이크로 하여 억제 과정을 수행한다.

[0068] 도 2의 양자화 STDP 함수를 다시 살펴보면, 입력 스파이크(IS[b])가 아닌 출력 스파이크(OS[c])를 트리거 스파이크로 동작하는 경우에도, x 축에서 음의 값($x < 0$)을 갖는 영역에 따라 출력 스파이크(OS)가 발생된 이후 입력 스파이크(IS)가 발생된 시점에 따른 가중치 변화량(Δw)을 계산함으로써 억제 과정을 수행할 수 있음을 알 수 있다.

[0069] 다만 출력 스파이크(OS[c])를 트리거 스파이크로 억제 과정을 수행하기 위해서는 출력 스파이크(OS[c])가 발화된 이후 발생하는 입력 스파이크(IS)에 대한 히스토리가 필요하다. 이에 본 실시예의 가중치 학습부(400)는 다수의 출력 뉴런 중 적어도 하나의 출력 뉴런에서 출력 스파이크(OS[c])가 발화되면, 출력 스파이크(OS[c])가 발화된 시점을 기준으로 이후, 학습 윈도우 크기에 대응하는 기간 동안 발생하는 모든 입력 스파이크(IS)의 시간 정보를 애프터 입력 스파이크 히스토리에 저장한다. 여기서 애프터 입력 스파이크 히스토리는 다수의 입력 뉴런(INN) 각각에 대응하여 다수개로 획득될 수 있다.

[0070] 즉 가중치 학습부(400)는 출력 스파이크 히스토리를 저장하지 않고, 입력 뉴런(INN)의 개수에 대응하는 다수의 입력 스파이크 히스토리와 다수의 애프터 입력 스파이크 히스토리를 저장한다.

[0071] 그리고 학습 윈도우에 의해 지정된 시간(여기서는 일예로 60 사이클) 이후에 애프터 입력 스파이크 히스토리에 저장된 입력 스파이크(IS)의 시간 정보를 기초로 가중치 변화량(Δw)을 계산한다. 이때, 가중치 학습부(400)는 강화 과정과 마찬가지로 출력 스파이크(OS[c])가 발화된 출력 뉴런(OUN)에 연결된 비트 라인(BL)과 다수의 워드 라인(WL) 각각이 교차하는 위치에 배치된 다수의 메모리 셀(MC) 각각에 저장된 가중치(w)를 강화하기 위한 다수의 가중치 변화량(Δw)을 계산한다. 그리고 계산된 다수의 가중치 변화량(Δw)을 각각을 대응하는 메모리 셀(MC)에 저장된 가중치(w)에 가산하여 가중치(w)를 업데이트하고, 업데이트된 가중치(w)를 다시 대응하는 메모리 셀(MC)에 저장한다.

[0072] 즉 강화 과정과 동일하게 출력 스파이크(OS[c])가 발화된 출력 뉴런(OUN)에 연결된 비트 라인(BL)을 기준으로 열 방향으로 다수의 메모리 셀(MC) 각각 저장된 가중치(w)를 업데이트한다.

[0073] 이때, 다수의 가중치 변화량(Δw) 각각은 도 2에 도시된 양자와 STDP 함수에 기반하여 대응하는 워드 라인(WL)을 통해 입력 스파이크(IS)가 인가된 시점과 개수에 따라 서로 다르게 계산될 수 있다. 또한 도 2에서는 출력 스파이크(OS[c])가 발화된 시점을 기준으로 이후 발생한 입력 스파이크(IS)에 따라 음의 값을 갖는 가중치 변화량(Δw)을 계산하여 가중치를 억제하는 것으로 도시하였으나, 가중치 학습부(400)는 강화 과정과 마찬가지로 가중치 변화량(Δw)을 양의 값으로 계산하고, 저장된 가중치(w)에서 계산된 양의 값을 갖는 가중치 변화량(Δw)을 차감함으로써, 가중치를 억제하도록 구성될 수 있다. 이 경우, 가중치 학습부(400)는 강화 과정과 억제 과정에서 입력 스파이크 히스토리와 애프터 입력 스파이크 히스토리를 이용하는 점에서만 차이가 있을 뿐, 동일한 방식으로 가중치 변화량(Δw)을 계산할 수 있다는 장점이 있다. 그리고 계산된 가중치 변화량(Δw)을 강화 과정인지 억제 과정인지에 따라 기존의 가중치(w)에 가산하거나 감산함으로써 가중치(w)를 업데이트할 수 있다.

[0074] 상기한 바와 같이 가중치 학습부(400)가 STDP 기법에 따라 가중치(w)를 강화하거나 억제하는 과정에서 모두 출력 스파이크(OS)를 트리거 스파이크로 이용하게 되면, 가중치 학습부(400)는 입력 스파이크(IS)의 발생 여부에 무관하게 출력 스파이크(OS)의 발화 여부에 따라 가중치(w)를 강화하거나 억제하게 되므로, 강화 과정 및 억제 과정을 구분하여 가중치(w)가 저장된 메모리 셀(MC)에 액세스 할 필요가 없으므로 액세스 횟수를 크게 줄일 수

있다. 뿐만 아니라, 열 방향의 한 방향에서만 메모리 셀(MC)에 액세스하게 되므로 메모리 셀(MC)에 액세스 하기 위한 구조를 매우 단순화할 수 있다.

[0075] 여기서 가중치 학습부(400)가 강화 과정 및 억제 과정 모두에서 출력 스파이크(OS)를 트리거 스파이크로 이용하는 것은 출력 스파이크(OS)의 발화 빈도가 입력 스파이크(IS)의 발생 빈도보다 매우 낮기 때문이다.

[0076] 일반적으로 학습 윈도우의 크기 내에서 다수의 출력 뉴런(OUN)을 포함하는 포스트 뉴런부(300)가 출력 스파이크(OS)가 10 ~ 20회 발화하는 동안, 다수의 입력 뉴런(INN)을 포함하는 프리 뉴런부(100)는 입력 스파이크(IS)를 대략 6000회 발생한다. 따라서 입력 스파이크(IS)를 트리거 스파이크로 이용하는 경우, 가중치 학습부(400)는 6000회 억제 과정을 수행해야 하는 반면, 출력 스파이크(OS)를 트리거 스파이크로 이용하는 경우 10 ~ 20회만 억제 과정을 수행하면 된다. 즉 단순히 트리거 스파이크를 하나로 통합함에 따라 메모리 셀(MC)에 대한 액세스 횟수를 1/2 수준이 아니라 비약적으로 줄일 수 있다.

[0077] 결과적으로 본 실시예에 따른 가중치 학습부(400)는 강화 과정과 억제 과정에서 모두 동일하게 출력 스파이크(OS)를 트리거 스파이크로 이용함에 따라, 기존의 STDP 기법에 비해 메모리 액세스 횟수를 수백배 감소시킬 수 있어 학습 시간 및 전력 소비를 크게 줄일 수 있을 뿐만 아니라, 메모리 셀(MC)의 액세스 방향을 통일함으로써 가중치 학습부(400)의 구조를 단순화할 수 있다.

[0078] 상기에서는 가중치 학습부(400)가 상대적으로 발생 빈도가 낮은 출력 스파이크(OS)를 트리거 스파이크로 이용하여 강화 및 억제 과정을 수행하는 것으로 가정하여 설명하였으나, 경우에 따라서 가중치 학습부(400)는 입력 스파이크(IS)를 트리거 스파이크로 이용하여 강화 및 억제 과정을 수행할 수도 있다. 입력 스파이크(IS)를 트리거 스파이크로 이용하는 경우, 비록 출력 스파이크(OS)를 트리거 스파이크로 이용하는 경우에 비해서는 메모리 액세스 횟수가 크게 줄어들지 않으나, 기존의 기법에 비해서는 1/2 이상 메모리 액세스 횟수를 줄일 수 있으며, 동일한 행 방향에서 메모리 셀(MC)에 액세스를 수행하므로, 가중치 학습부(400)의 구조를 단순화할 수 있다.

[0079] 도 10은 본 발명의 일 실시예에 따른 스파이킹 신경망 학습 방법을 나타낸다.

[0080] 본 실시예에 따른 SNN 학습 방법은 가중치 학습 장치로 동작하는 도 1의 가중치 학습부(400)에 의해 수행될 수 있으며, 여기서는 가중치 학습부(400)가 STDP 기법에서 출력 스파이크(OS)를 트리거 스파이크로 이용하여 강화 과정 및 억제 과정을 수행하는 것으로 가정하지만, 상기한 바와 같이, 가중치 학습부(400)는 입력 스파이크(IS)를 트리거 스파이크로 이용하여 강화 과정 및 억제 과정을 수행할 수도 있다.

[0081] 도 10을 참조하면 본 실시예의 SNN 학습 방법은 우선 프리 뉴런부(100)의 다수의 입력 뉴런(INN) 중 적어도 하나의 입력 뉴런에서 입력 스파이크(IS)가 생성되는지 판별한다(S10). 만일 입력 스파이크(IS)가 생성되면, 생성된 입력 스파이크의 시간 정보를 입력 스파이크 히스토리에 저장한다(S20). 여기서 입력 스파이크 히스토리는 기지정된 학습 윈도우 크기에 의해 지정된 기간 이전에 생성되어 저장된 입력 스파이크에 대한 시간 정보는 제거한다. 즉 입력 스파이크 히스토리는 기지정된 학습 윈도우 크기에 대응하는 이전 시간 기간 동안 생성된 입력 스파이크의 시간 정보만을 저장한다.

[0082] 한편, 이전 발화된 출력 스파이크(OS)에 응답하여 기생성된 애프터 입력 스파이크 히스토리가 존재하는지 판별한다(S30). 만일 기생성된 애프터 입력 스파이크 히스토리가 존재하면, 생성된 입력 스파이크의 시간 정보를 애프터 입력 스파이크 히스토리에도 저장한다(S40). 여기서 입력 스파이크 히스토리는 다수의 입력 뉴런(INN) 각각에 대응하여 다수개로 존재할 수 있으며 수 있다.

[0083] 그리고 애프터 입력 스파이크 히스토리가 생성된 이후 기지정된 시간이 경과되었는지 판별한다(S50). 여기서 기지정된 시간은 학습 윈도우의 크기로 지정된 시간일 수 있다. 만일 학습 윈도우에 의해 지정된 시간이 경과된 것으로 판별되면, 애프터 입력 스파이크 히스토리에 저장된 입력 스파이크 각각의 시간 정보를 도 2에 도시된 양자화 STDP 함수에 대입하여 가중치 변화량(Δw)을 계산하고, 계산된 가중치 변화량(Δw)을 출력 스파이크(OS)가 발화된 출력 뉴런(OUN)에 연결된 비트 라인(BL) 상의 다수의 메모리 셀(MC)에 저장된 가중치(w)에 가산하여 업데이트하여 가중치 억제 과정을 수행한다(S60). 즉 다수의 메모리 셀(MC)에 저장된 다수의 가중치(w)를 읽고 업데이트한 후, 다시 다수의 메모리 셀(MC)에 저장한다. 여기서 가중치 변화량(Δw)은 다수의 입력 뉴런(INN) 각각에 대응하는 다수의 애프터 입력 스파이크 히스토리에 따라 각 워드 라인(WL)에 대해 서로 다르게 계산될 수 있으며, 계산된 다수의 가중치 변화량(Δw)은 이전 출력 스파이크(OS)가 발화된 출력 뉴런(OUN)에 연결된 비트 라인(BL)과 대응하는 워드 라인(WL)이 교차하는 위치의 메모리 셀(MC)에 저장된 가중치(w)에 적용될 수 있다.

[0084] 여기서는 가중치 변화량(Δw)이 양자화 STDP 함수에 따라 음의 값으로 계산되는 것으로 가정하여, 계산된 가중

치 변화량(Δw)을 가중치(w)에 가산하는 것으로 설명하였으나, 경우에 따라서 가중치 변화량(Δw)은 양의 값으로 계산되고, 계산된 가중치 변화량(Δw)을 가중치(w)에 차감하여 가중치를 업데이트 할 수도 있다.

[0085] 그리고 포스트 뉴런부(300)의 다수의 출력 뉴런(OUN) 중 적어도 하나의 출력 뉴런에서 출력 스파이크(OS)가 발화되는지 판별한다(S70). 만일 적어도 하나의 출력 뉴런에서 출력 스파이크(OS)가 발화되면, 출력 스파이크(OS)가 발화되기 이전 입력 스파이크의 발생 정보가 저장된 입력 스파이크 히스토리에 따라 가중치 변화량(Δw)을 계산하고, 계산된 가중치 변화량(Δw)을 출력 스파이크(OS)가 발화된 출력 뉴런(OUN)에 연결된 비트 라인(BL) 상의 다수의 메모리 셀(MC)에 저장된 가중치(w)에 가산하여 업데이트하여 가중치 강화 과정을 수행한다(S80). 여기서도 가중치 변화량(Δw)은 다수의 입력 뉴런(INN) 각각에 대응하는 다수의 입력 스파이크 히스토리에 따라 각 워드 라인(WL)에 대해 서로 다르게 계산될 수 있으며, 계산된 다수의 가중치 변화량(Δw)은 출력 스파이크(OS)가 발화된 출력 뉴런(OUN)에 연결된 비트 라인(BL)과 대응하는 워드 라인(WL)이 교차하는 위치의 메모리 셀(MC)에 저장된 가중치(w)에 적용될 수 있다.

[0086] 이와 함께 출력 스파이크(OS)가 발화된 시점으로부터 이후 학습 윈도우에 의해 기지정된 시간 동안 생성되는 입력 스파이크의 시간 정보를 저장하기 위한 애프터 입력 스파이크 히스토리를 생성한다(S80).

[0087] 상기에서는 입력 스파이크 히스토리와 애프터 입력 스파이크 히스토리를 별도로 생성하여, 트리거 스파이크로 이용되는 출력 스파이크가 발생되기 이전 및 이후 기지정된 시간 동안 생성된 입력 스파이크에 대한 시간 정보를 저장하는 것으로 설명하였으나, 이는 설명의 편의를 위한 것으로 실제로는 학습 윈도우의 크기에 대응하는 크기를 갖고 기지정된 시간 구간 동안의 입력 스파이크의 시간 정보를 저장하는 하나의 입력 스파이크 히스토리만을 이용할 수도 있다. 이 경우, 출력 스파이크(OS)가 발화되면, 현재까지 입력 스파이크 히스토리에 저장된 입력 스파이크의 시간 정보를 이용하여 가중치 강화 과정을 수행하고, 출력 스파이크(OS)가 발화된 이후 학습 윈도우의 크기에 대응하는 시간이 경과되면, 다시 입력 스파이크 히스토리에 입력 스파이크의 시간 정보를 이용하여 가중치 억제 과정을 수행하여 가중치를 업데이트 할 수 있다. 즉 SNN을 학습시킬 수 있다.

[0088] 따라서 본 실시예에 따른 SNN 학습 장치 및 방법은 메모리 액세스 횟수를 줄이고, 강화 과정과 억제 과정에서의 메모리 액세스 방향을 동일한 방향으로 지정함에 따라 학습 시간 및 전력 소비를 줄이고, 학습 장치의 구성을 단순화할 수 있다. 뿐만 아니라 스파이크 히스토리를 저장하기 위해 요구되는 저장 공간을 크게 줄일 수 있다.

[0089] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.

[0090] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.

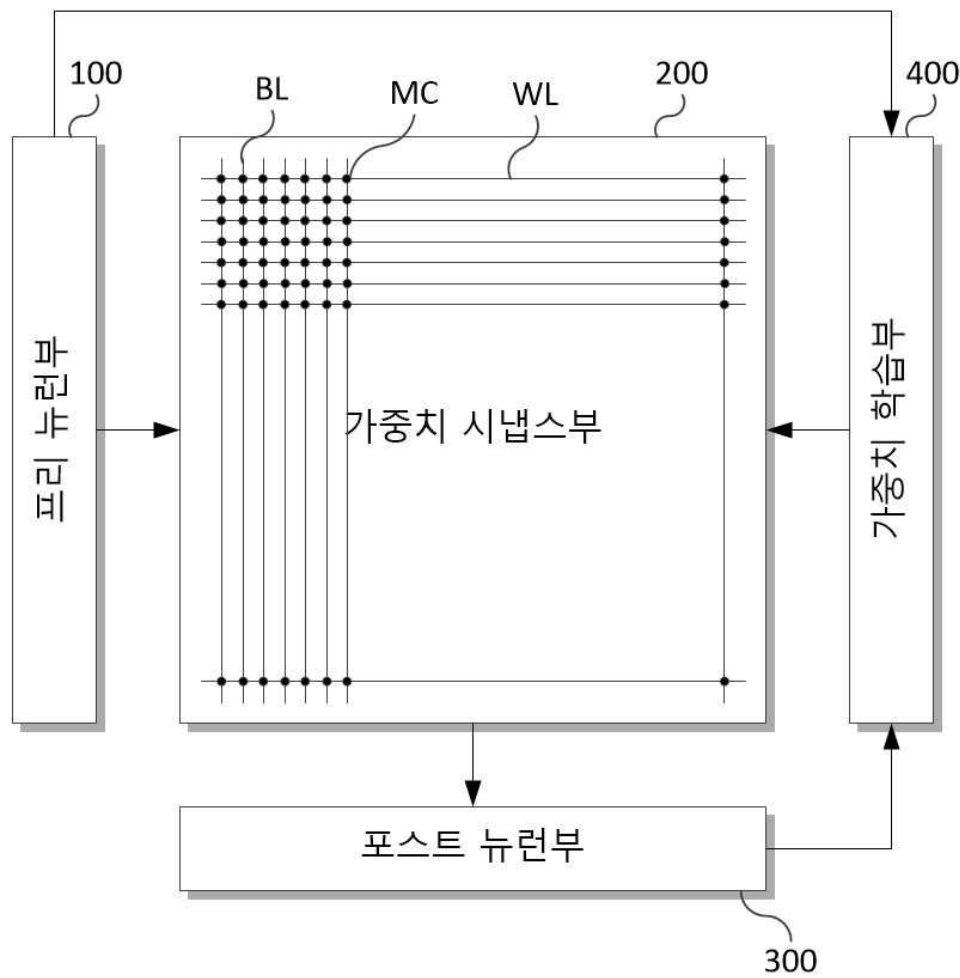
[0091] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

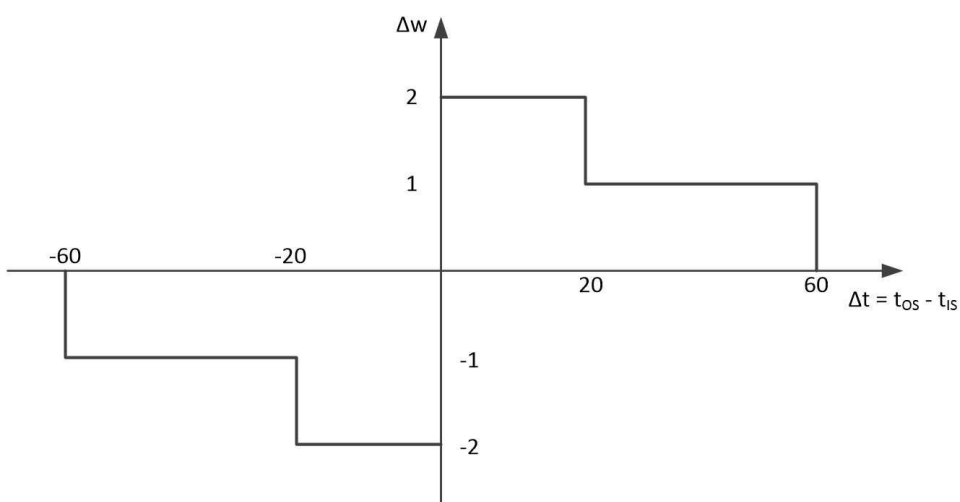
[0092]	100: 프리 뉴런부	200: 가중치 시냅스부
	300: 포스트 뉴런부	400: 가중치 학습부
	WL: 워드라인	BL: 비트라인
	MC: 메모리 셀	

도면

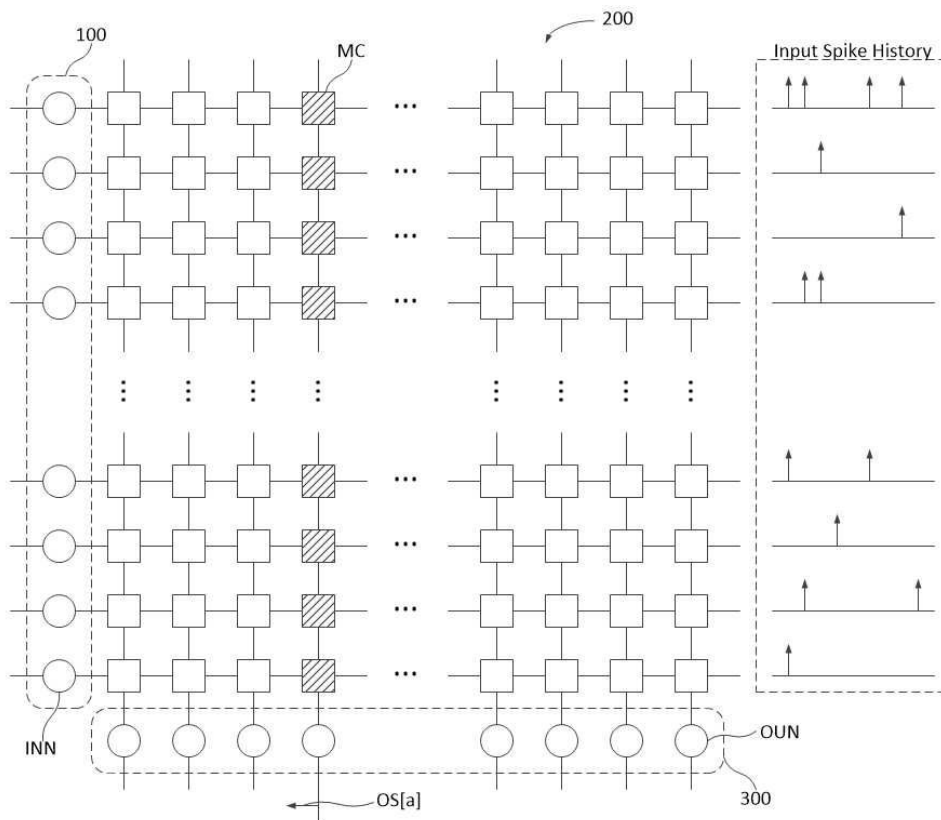
도면1



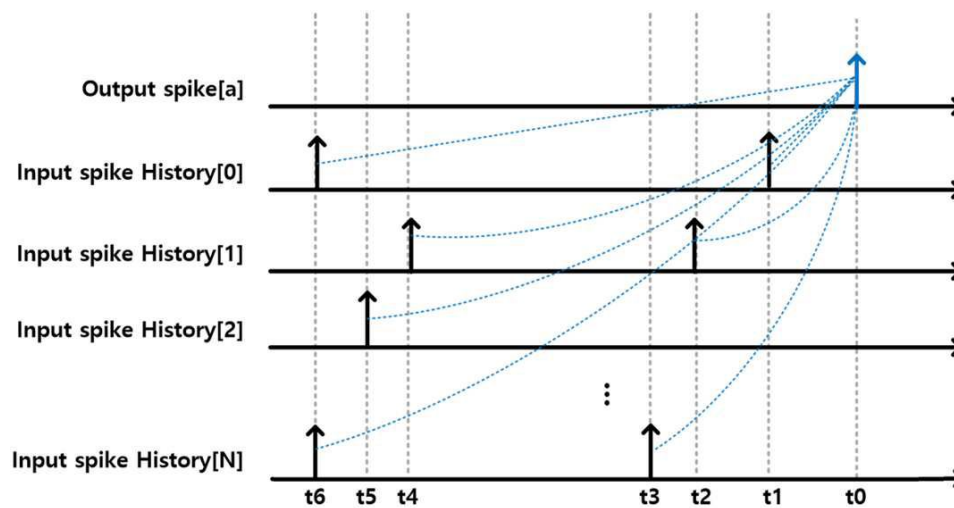
도면2



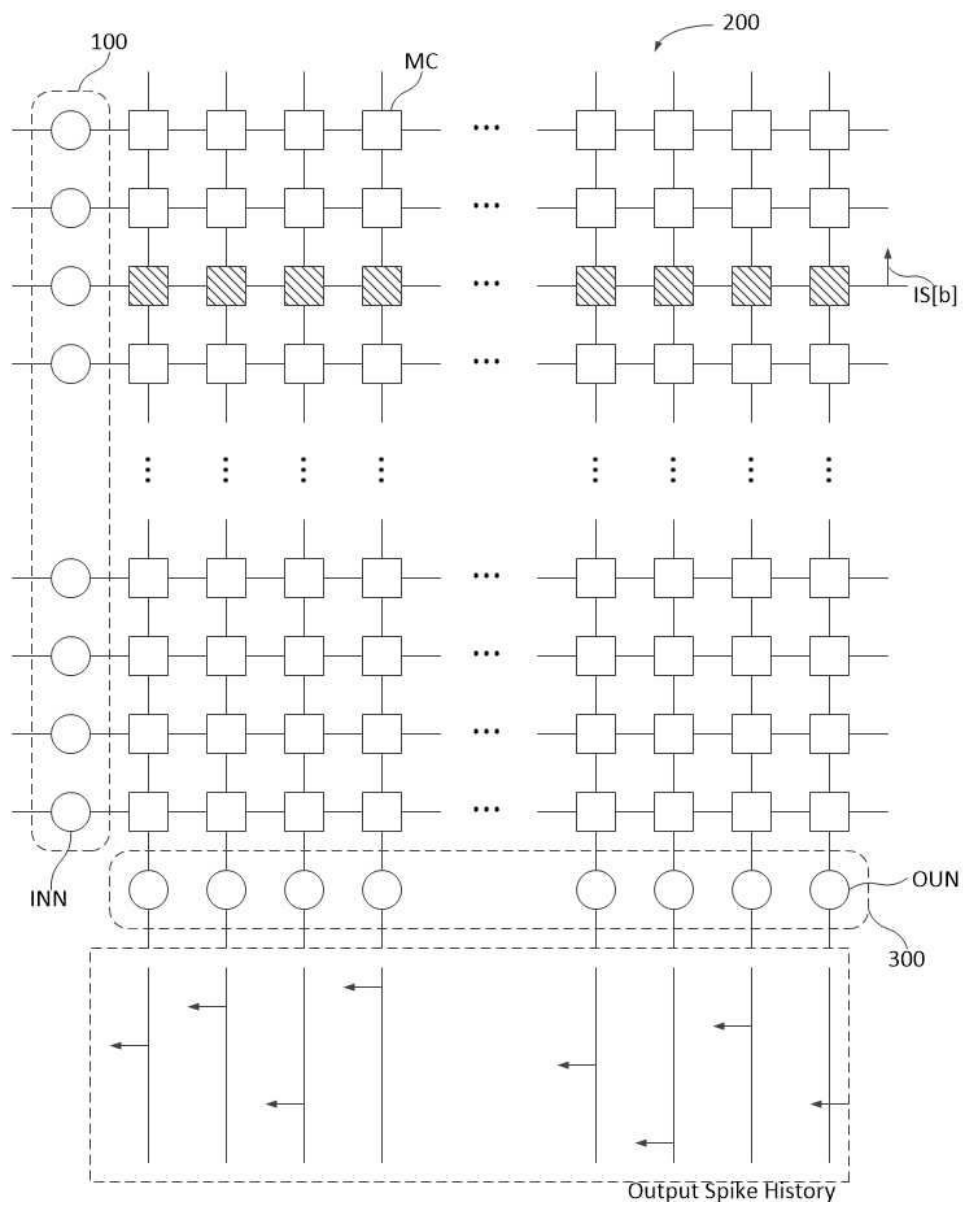
도면3



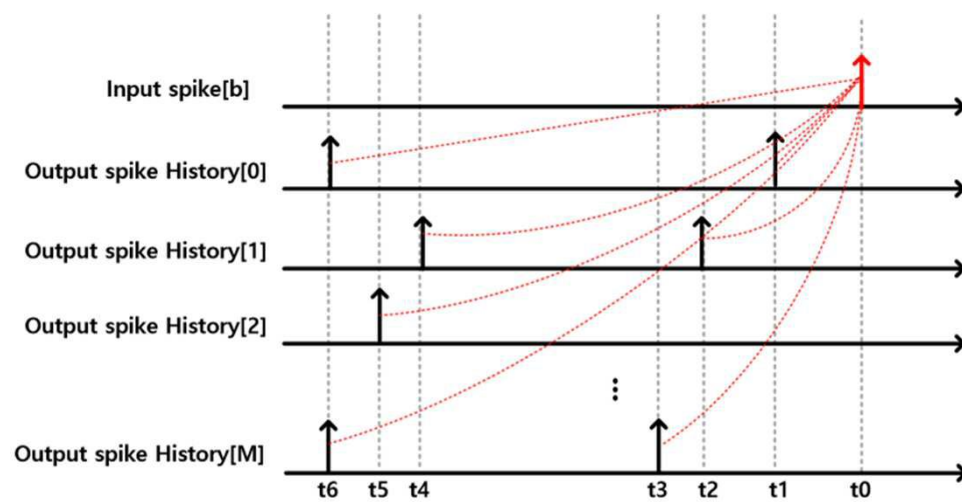
도면4



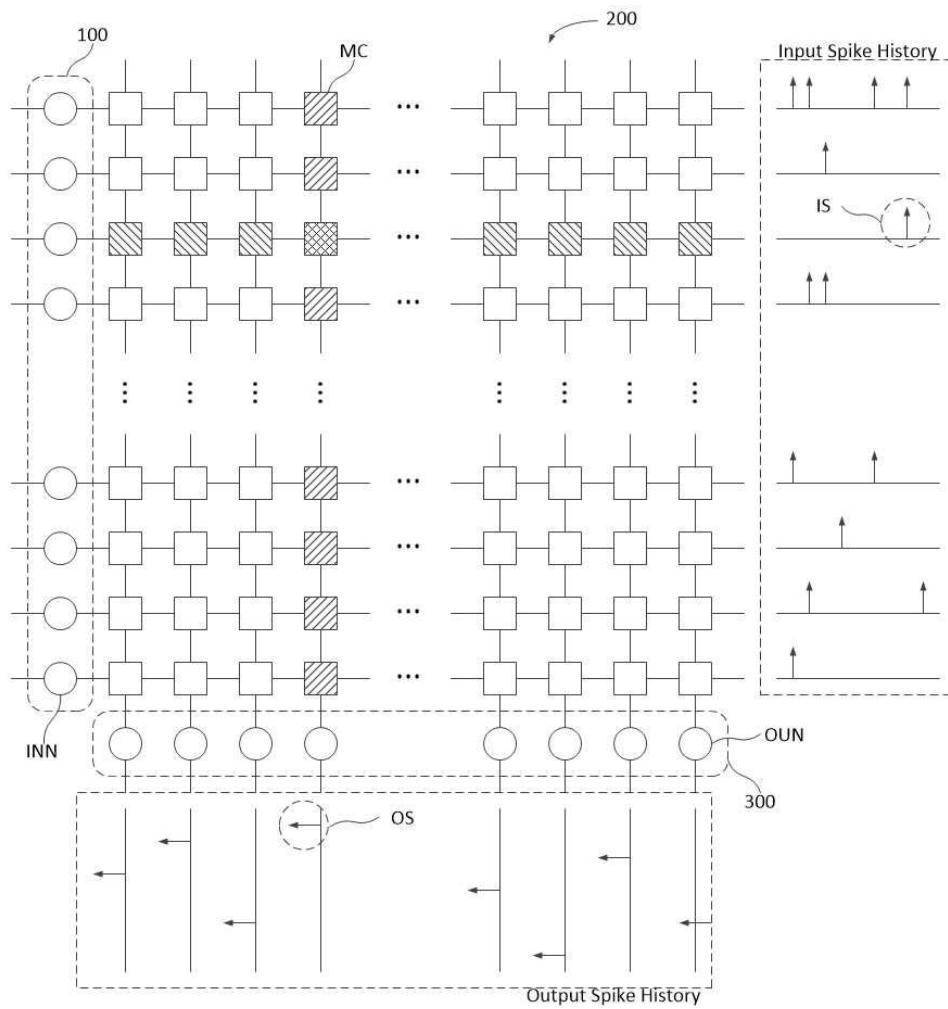
도면5



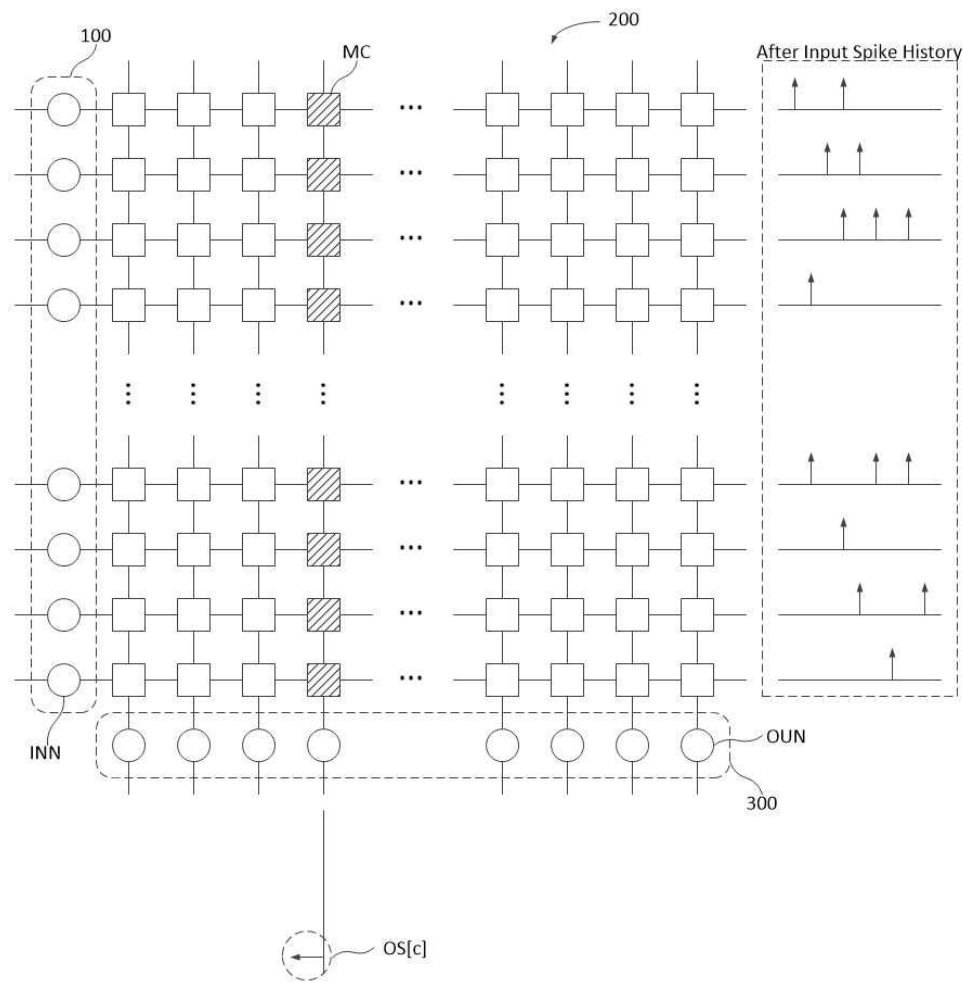
도면6



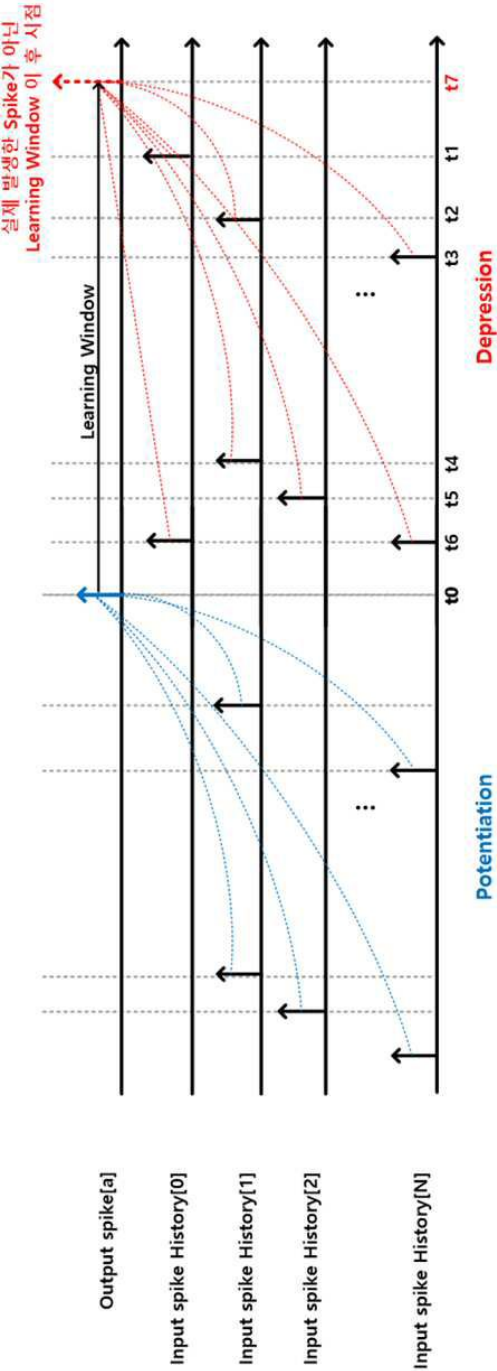
도면7



도면8



도면9



도면10

