(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년03월07일
(11) 등록번호 10-2778708
(24) 등록일자 2025년03월05일

(51) 국제특허분류(Int. Cl.)
G06N 3/065 (2023.01) *G11C 16/08* (2006.01)
G11C 27/02 (2021.01)
(52) CPC특허분류
G06N 3/065 (2023.01)
G11C 16/08 (2013.01)
(21) 출원번호 10-2023-7039451(분할)
(22) 출원일자(국제) 2019년01월24일
심사청구일자 2023년11월15일
(85) 번역문제출일자 2023년11월15일
(65) 공개번호 10-2023-0163572
(43) 공개일자 2023년11월30일
(62) 원출원 특허 10-2023-7010513
원출원일자(국제) 2019년01월24일
심사청구일자 2023년03월27일
(86) 국제출원번호 PCT/US2019/015022
(87) 국제공개번호 WO 2019/177698
국제공개일자 2019년09월19일
(30) 우선권주장
62/642,884 2018년03월14일 미국(US)
15/991,890 2018년05월29일 미국(US)
(56) 선행기술조사문헌
JP2009032362 A
JP2004077743 A

(73) 특허권자
실리콘 스토리지 테크놀로지 인크
미국 캘리포니아주 95134 산호세 홀거 웨이 450
(72) 발명자
트랜 휴 반
미국, 95135 캘리포니아, 산 호세, 게일리 플레이스 2642
홍 스탠리
미국, 95131 캘리포니아, 산 호세, 브리스톨 베이 카먼 1848
(뒷면에 계속)
(74) 대리인
강명구, 최정윤

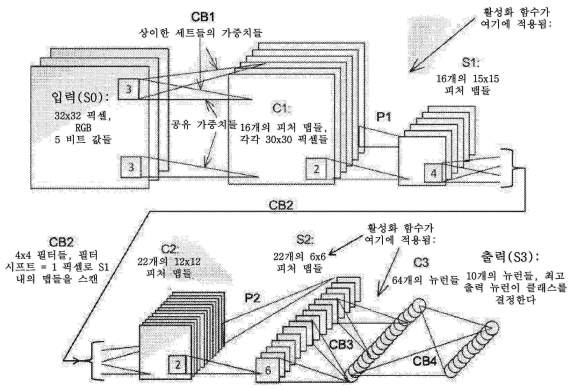
전체 청구항 수 : 총 3 항 심사관 : 이준상

(54) 발명의 명칭 **심층 학습 인공 신경망에서 아날로그 신경 메모리를 위한 디코더들**

(57) 요 약

인공 신경망에서 벡터-매트릭스 승산(vector-by-matrix multiplication, VMM) 어레이와 함께 사용하기 위한 디코더들의 다수의 실시예들이 개시된다. 디코더들은 비트 라인 디코더들, 워드 라인 디코더들, 제어 게이트 디코더들, 소스 라인 디코더들, 및 소거 게이트 디코더들을 포함한다. 소정 실시예들에서, 디코더의 고전압 버전 및 저전압 버전이 사용된다.

대 표 도 - 도6



(52) CPC특허분류
G11C 27/02 (2013.01)
G11C 2216/14 (2013.01)
(72) 발명자
리 안
미국, 95121 캘리포니아, 산 호세, 린드무어 드라이브 3385
부 투안
미국, 95138 캘리포니아, 산 호세, 대너 코트 431
팜 히엔
베트남, 호 치 민, 고 밥 디스트릭트, 워드 15, 르 두 토 스트릿 736/163/18
느구옌 카
베트남, 호 치 민, 고 밥 디스트릭트, 워드 16, 느구옌 반 루옹 스트릿 350/160
트란 한
베트남, 호 치 민, 고 밥 디스트릭트, 워드 14, 엘 티. 아. 캉지아 아파트먼트 룸 502

명 세 서

청구범위

청구항 1

전류-전압(current-to-voltage) 회로로서,

워드 라인 단자, 제어 게이트 단자, 및 비트 라인 단자를 포함하는 기준 메모리 셀, 그리고 상기 기준 메모리 셀의 비트 라인 단자에 결합되는 제1 단자와, 입력 전류를 수신하고, 스위치를 통해 상기 기준 메모리 셀의 상기 워드 라인 단자 또는 제어 게이트 단자에 선택적으로 결합되는 제2 단자를 포함하는 캐스코딩 바이어스 트랜지스터를 포함하는 기준 회로; 그리고

접지에 결합되는 제2 단자 및 제1 단자를 포함하는 커패시터와, 상기 커패시터의 제1 단자와 상기 기준 메모리 셀의 비트 라인 단자 사이에 결합되는 샘플링 스위치를 포함하는 샘플-앤드-홀드 회로 - 샘플링된 전압이 상기 커패시터의 제1 단자에 제공됨 - 를 포함하는, 전류-전압 회로.

청구항 2

제1항에 있어서, 상기 샘플링된 전압을 수신하고 증폭된 전압을 출력하기 위한 증폭기를 추가로 포함하는, 전류-전압 회로.

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

전류-전압 회로로서,

워드 라인 단자, 제어 게이트 단자, 및 비트 라인 단자를 포함하는 기준 메모리 셀, 그리고 상기 기준 메모리 셀의 비트 라인 단자에 결합되는 제1 단자와, 입력 전류를 수신하여 스위치를 통해 상기 기준 메모리 셀의 상기 워드 라인 단자 또는 제어 게이트 단자에 선택적으로 결합되는 제2 단자를 포함하는 캐스코딩 바이어스 트랜지스터를 포함하는 기준 회로;

상기 기준 메모리 셀의 비트 라인 단자로부터 전압을 수신하고, 출력 단자 상에 증폭된 전압을 출력하기 위한 증폭기; 그리고

접지에 결합되는 제2 단자 및 제1 단자를 포함하는 커패시터와, 상기 커패시터의 제1 단자와 상기 증폭기의 출력 단자 사이에 결합되는 샘플링 스위치를 포함하는 샘플-앤드-홀드 회로 - 상기 커패시터의 제1 단자에 출력 전압이 제공됨 - 를 포함하는, 전류-전압 회로.

청구항 8

삭제

청구항 9

삭제

발명의 설명

기 술 분 야

[0001] 우선권 주장

[0002] 본 출원은 2018년 3월 14일자로 출원되고 발명의 명칭이 "Decoders for Analog Neuromorphic Memory in Artificial Neural Network"인 미국 가특허 출원 제62/642,884호, 및 2018년 5월 29일자로 출원되고 발명의 명칭이 "Decoders For Analog Neural Memory In Deep Learning Artificial Neural Network"인 미국 특허 출원 제15/991,890호에 대한 우선권을 주장한다.

[0003] 기술분야

[0004] 인공 신경망에서 벡터-매트릭스 승산(vector-by-matrix multiplication, VMM) 어레이와 함께 사용하기 위한 디코더들의 다수의 실시예들이 개시된다.

배 경 기 술

[0005] 인공 신경망은 생물학적 신경망(동물의 중추신경계, 특히 뇌)을 모방하는데, 이들은 다수의 입력에 의존할 수 있고 일반적으로 알려져 있지 않은 기능들을 추정하거나 근사화하는 데 이용된다. 인공 신경망은, 일반적으로, 서로 메시지들을 교환하는 상호연결된 "뉴런(neuron)"의 층들을 포함한다.

[0006] 도 1은 인공 신경망을 예시하며, 여기서 원은 뉴런의 층 또는 입력을 나타낸다. 연접부(시냅스(synapse)들로 지칭됨)는 화살표로 표현되며, 경험에 기초하여 튜닝될 수 있는 수치 가중치를 갖는다. 이는 신경망을 입력에 적응시키고 학습할 수 있게 한다. 전형적으로, 신경망은 다수의 입력들의 층을 포함한다. 전형적으로 뉴런의 하나 이상의 중간 층, 및 신경망의 출력을 제공하는 뉴런의 출력 층이 있다. 각각의 레벨의 뉴런은 개별적으로 또는 집합적으로 시냅스들로부터의 수신된 데이터에 기초하여 결정을 행한다.

[0007] 고성능 정보 처리를 위한 인공 신경망의 개발에서의 주요 과제들 중 하나는 적절한 하드웨어 기술의 결여이다. 사실상, 실제 신경망은 매우 많은 수의 시냅스들에 의존하여, 뉴런들 사이의 높은 연결성, 즉 매우 높은 계산 병렬성(computational parallelism)을 가능하게 한다. 원칙적으로, 그러한 복잡성은 디지털 슈퍼컴퓨터들 또는 특수 그래픽 처리 유닛 클러스터들로 달성될 수 있다. 그러나, 고비용에 더하여, 이들 접근법은 또한 주로 저-정밀 아날로그 계산을 수행하기 때문에 훨씬 적은 에너지를 소비하는 생물학적 망(biological network)과 비교하여 평범한 에너지 효율을 겪는다. CMOS 아날로그 회로가 인공 신경망에 사용되어 왔지만, 대부분의 CMOS-구현된 시냅스들은 많은 수의 뉴런들 및 시냅스들이 주어지면 너무 부피가 커졌다.

[0008] 출원인은, 참고로 포함되는, 미국 특허 출원 제15/594,439호에서 하나 이상의 비휘발성 메모리 어레이를 시냅스로서 이용하는 인공 (아날로그) 신경망을 이전에 개시하였다. 비휘발성 메모리 어레이들은 아날로그 뉴로모픽 메모리(analog neuromorphic memory)로서 동작한다. 신경망 디바이스는 제1 복수의 입력들을 수신하고 그로부터 제1 복수의 출력들을 생성하도록 구성된 제1 복수의 시냅스들, 및 제1 복수의 출력들을 수신하도록 구성된 제1 복수의 뉴런들을 포함한다. 제1 복수의 시냅스들은 복수의 메모리 셀들을 포함하는데, 여기서 메모리 셀들 각각은 반도체 기판 내에 형성되고 채널 영역이 사이에 연장되는 이격된 소스 영역과 드레인 영역, 채널 영역의 제1 부분 위에 배치되고 그로부터 절연되는 플로팅 게이트, 및 채널 영역의 제2 부분 위에 배치되고 그로부터 절연되는 비-플로팅 게이트를 포함한다. 복수의 메모리 셀들 각각은 플로팅 게이트 상의 전자들의 수에 대응하는 가중치 값을 저장하도록 구성된다. 복수의 메모리 셀들은 제1 복수의 입력들을 저장된 가중치 값들과 승산하여 제1 복수의 출력들을 생성하도록 구성된다.

[0009] 아날로그 뉴로모픽 메모리 시스템에서 사용되는 각각의 비휘발성 메모리 셀은 플로팅 게이트에서 매우 특정적이고 정확한 양의 전하를 보유하도록 소거되고 프로그램되어야 한다. 예를 들어, 각각의 플로팅 게이트는 N개의 상이한 값들 중 하나를 보유해야 하며, 여기서 N은 각각의 셀에 의해 표시될 수 있는 상이한 가중치들의 수이다. N의 예들은 16, 32 및 64를 포함한다.

[0010] 종래의 플래시 메모리 어레이들에서 사용되는 종래 기술의 디코딩 회로들(예컨대, 비트 라인 디코더들, 워드 라

인 디코더들, 제어 게이트 디코더들, 소스 라인 디코더들, 및 소거 게이트 디코더들)은 아날로그 뉴로모픽 메모리 시스템에서 VMM과 함께 사용하기에 적합하지 않다. 이에 대한 한 가지 이유는, VMM 시스템에서, 프로그래밍 및 검증 동작의 (판독 동작인) 검증 부분이 단일 선택된 메모리 셀에 대해 동작하는 반면, 판독 동작이 어레이 내의 모든 메모리 셀들에 대해 동작한다는 것이다.

[0011] 아날로그 뉴로모픽 메모리 시스템에서 VMM과 함께 사용하기에 적합한 개선된 디코딩 회로들이 필요하다.

발명의 내용

[0012] 인공 신경망에서 VMM 어레이와 함께 사용하기 위한 디코더들의 다수의 실시예들이 개시된다.

도면의 간단한 설명

[0013] 도 1은 인공 신경망을 예시하는 도면이다.

도 2는 종래의 2-게이트 비휘발성 메모리 셀의 측단면도이다.

도 3은 종래의 4-게이트 비휘발성 메모리 셀의 측단면도이다.

도 4는 종래의 3-게이트 비휘발성 메모리 셀의 측단면도이다.

도 5는 다른 종래의 2-게이트 비휘발성 메모리 셀의 측단면도이다.

도 6은 비휘발성 메모리 어레이를 활용하는 예시적인 인공 신경망의 상이한 레벨들을 예시하는 도면이다.

도 7은 벡터 승산기 매트릭스를 예시하는 블록도이다.

도 8은 다양한 레벨의 벡터 승산기 매트릭스를 예시하는 블록도이다.

도 9는 벡터 승산기 매트릭스의 일 실시예를 도시한다.

도 10은 벡터 승산기 매트릭스의 다른 실시예를 도시한다.

도 11은 벡터 승산기 매트릭스의 다른 실시예를 도시한다.

도 12는 벡터 승산기 매트릭스의 다른 실시예를 도시한다.

도 13은 벡터 승산기 매트릭스의 다른 실시예를 도시한다.

도 14는 벡터 승산기 매트릭스를 위한 비트 라인 디코더의 일 실시예를 도시한다.

도 15는 벡터 승산기 매트릭스를 위한 비트 라인 디코더의 다른 실시예를 도시한다.

도 16은 벡터 승산기 매트릭스를 위한 비트 라인 디코더의 다른 실시예를 도시한다.

도 17은 벡터 승산기 매트릭스를 동작시키기 위한 시스템을 도시한다.

도 18은 벡터 승산기 매트릭스를 동작시키기 위한 다른 시스템을 도시한다.

도 19는 벡터 승산기 매트릭스를 동작시키기 위한 다른 시스템을 도시한다.

도 20은 벡터 승산기 매트릭스와 함께 사용하기 위한 워드 라인 드라이버의 일 실시예를 도시한다.

도 21은 벡터 승산기 매트릭스와 함께 사용하기 위한 워드 라인 드라이버의 다른 실시예를 도시한다.

도 22는 벡터 승산기 매트릭스와 함께 사용하기 위한 워드 라인 드라이버의 다른 실시예를 도시한다.

도 23은 벡터 승산기 매트릭스와 함께 사용하기 위한 워드 라인 드라이버의 다른 실시예를 도시한다.

도 24는 벡터 승산기 매트릭스와 함께 사용하기 위한 워드 라인 드라이버의 다른 실시예를 도시한다.

도 25는 벡터 승산기 매트릭스와 함께 사용하기 위한 워드 라인 드라이버의 다른 실시예를 도시한다.

도 26은 벡터 승산기 매트릭스와 함께 사용하기 위한 워드 라인 드라이버의 다른 실시예를 도시한다.

도 27은 벡터 승산기 매트릭스와 함께 사용하기 위한 소스 라인 디코더 회로를 도시한다.

도 28은 벡터 승산기 매트릭스와 함께 사용하기 위한 워드 라인 디코더 회로, 소스 라인 디코더 회로, 및 고전

압 레벨 시프터를 도시한다.

도 29는 벡터 승산기 매트릭스와 함께 사용하기 위한 소거 게이트 디코더 회로, 제어 게이트 디코더 회로, 소스 라인 디코더 회로, 및 고전압 레벨 시프터를 도시한다.

도 30은 벡터 승산기 매트릭스와 함께 사용하기 위한 워드 라인 디코더 회로를 도시한다.

도 31은 벡터 승산기 매트릭스와 함께 사용하기 위한 제어 게이트 디코더 회로를 도시한다.

도 32는 벡터 승산기 매트릭스와 함께 사용하기 위한 다른 제어 게이트 디코더 회로를 도시한다.

도 33은 벡터 승산기 매트릭스와 함께 사용하기 위한 다른 제어 게이트 디코더 회로를 도시한다.

도 34는 벡터 승산기 매트릭스 내의 워드 라인을 제어하기 위한 전류-전압(current-to-voltage) 회로를 도시한다.

도 35는 벡터 승산기 매트릭스 내의 워드 라인을 제어하기 위한 다른 전류-전압 회로를 도시한다.

도 36은 벡터 승산기 매트릭스 내의 제어 게이트 라인을 제어하기 위한 전류-전압 회로를 도시한다.

도 37은 벡터 승산기 매트릭스 내의 제어 게이트 라인을 제어하기 위한 다른 전류-전압 회로를 도시한다.

도 38은 벡터 승산기 매트릭스 내의 제어 게이트 라인을 제어하기 위한 다른 전류-전압 회로를 도시한다.

도 39는 벡터 승산기 매트릭스 내의 워드 라인을 제어하기 위한 다른 전류-전압 회로를 도시한다.

도 40은 벡터 승산기 매트릭스 내의 워드 라인을 제어하기 위한 다른 전류-전압 회로를 도시한다.

도 41은 벡터 승산기 매트릭스 내의 워드 라인을 제어하기 위한 다른 전류-전압 회로를 도시한다.

도 42는 도 9의 벡터 승산기 매트릭스에 대한 동작 전압들을 도시한다.

도 43은 도 10의 벡터 승산기 매트릭스에 대한 동작 전압들을 도시한다.

도 44는 도 11의 벡터 승산기 매트릭스에 대한 동작 전압들을 도시한다.

도 45는 도 12의 벡터 승산기 매트릭스에 대한 동작 전압들을 도시한다.

발명을 실시하기 위한 구체적인 내용

[0014] 본 발명의 인공 신경망은 CMOS 기술과 비휘발성 메모리 어레이들의 조합을 이용한다.

[0015] 비휘발성 메모리 셀들

[0016] 디지털 비휘발성 메모리들이 잘 알려져 있다. 예를 들어, 미국 특허 제5,029,130호("'130 특허")가 분리형 게이트 비휘발성 메모리 셀들의 어레이를 개시하고 있으며, 모든 목적을 위해 본 명세서에 참고로 포함된다. 그러한 메모리 셀이 도 2에 도시되어 있다. 각각의 메모리 셀(210)은 반도체 기판(12) 내에 형성된 소스 영역(14) 및 드레인 영역(16)을 포함하며, 채널 영역(18)이 그들 사이에 있다. 플로팅 게이트(20)가 소스 영역(16)의 일부분 위에, 그리고 채널 영역(18)의 제1 부분 위에 형성되고 그로부터 절연된다(그리고 그의 전도율을 제어한다). 워드 라인 단자(22)(전형적으로 워드 라인에 커플링됨)가 채널 영역(18)의 제2 부분 위에 배치되고 그로부터 절연되는(그리고 그의 전도율을 제어하는) 제1 부분, 및 위쪽으로 그리고 플로팅 게이트(20) 위로 연장되는 제2 부분을 갖는다. 플로팅 게이트(20) 및 워드 라인 단자(22)는 게이트 산화물에 의해 기판(12)으로부터 절연된다. 비트 라인(24)이 드레인 영역(16)에 커플링된다.

[0017] 메모리 셀(210)은 워드 라인 단자(22) 상에 높은 포지티브 전압을 배치함으로써 소거되는데(여기서 전자들이 플로팅 게이트로부터 제거됨), 이는 플로팅 게이트(20) 상의 전자들이 파울러-노드하임 터널링(Fowler-Nordheim tunneling)을 통해 중간 절연체를 통과하여 플로팅 게이트(20)로부터 워드 라인 단자(22)로 터널링하게 한다.

[0018] 메모리 셀(210)은 워드 라인 단자(22) 상에 포지티브 전압을, 그리고 소스(16) 상에 포지티브 전압을 배치함으로써 프로그래밍된다(여기서 전자들이 플로팅 게이트 상에 배치됨). 전자 전류가 소스(16)로부터 드레인(14)을 향해 흐를 것이다. 전자들은 그들이 워드 라인 단자(22)와 플로팅 게이트(20) 사이의 갭에 도달할 때 가속되고 가열될 것이다. 가열된 전자들 중 일부는 플로팅 게이트(20)로부터의 정전 인력으로 인해 게이트 산화물(26)을 통과하여 플로팅 게이트(20) 상으로 주입될 것이다.

[0019] 메모리 셀(210)은 드레인(14) 및 워드 라인 단자(22) 상에 포지티브 판독 전압들을 배치함(이는 워드 라인 단자 아래의 채널 영역을 턴 온시킴)으로써 판독된다. 플로팅 게이트(20)가 포지티브로 대전되면(즉, 전자들이 소거되고 드레인(16)에 포지티브로 커플링되면), 플로팅 게이트(20) 아래의 채널 영역의 부분이 또한 턴 온되고, 전류가 채널 영역(18)을 가로질러 흐를 것이며, 이는 소거 또는 "1" 상태로 감지된다. 플로팅 게이트(20)가 네거티브로 대전되면(즉, 전자들로 프로그래밍되면), 플로팅 게이트(20) 아래의 채널 영역의 부분은 대부분 또는 완전히 턴 오프되고, 전류가 채널 영역(18)을 가로질러 흐르지 않을 것이며(또는 흐름이 거의 없을 것이며), 이는 프로그래밍된 또는 "0" 상태로 감지된다.

[0020] 표 1은 판독, 소거, 및 프로그래밍 동작들을 수행하기 위해 메모리 셀(210)의 단자들에 인가될 수 있는 전형적인 전압 범위들을 도시한다:

[0021] [표 1]

도 2 의 플래시 메모리 셀(210)의 동작

	WL	BL	SL
판독	2 내지 3 V	0.6 내지 2 V	0 V
소거	대략 11 내지 13 V	0 V	0 V
프로그래밍	1 내지 2 V	1 내지 3 μA	9 내지 10 V

[0022]

[0023] 다른 분리형 게이트 메모리 셀 구성들이 알려져 있다. 예를 들어, 도 3은 소스 영역(14), 드레인 영역(16), 채널 영역(18)의 제1 부분 위의 플로팅 게이트(20), 채널 영역(18)의 제2 부분 위의 선택 게이트(28)(전형적으로 워드 라인에 커플링됨), 플로팅 게이트(20) 위의 제어 게이트(22), 및 소스 영역(14) 위의 소거 게이트(30)를 포함하는 4-게이트 메모리 셀(310)을 도시한다. 이러한 구성은, 모든 목적을 위해 본 명세서에 참고로 포함되는, 미국 특허 제6,747,310호에 기재되어 있다. 여기서, 모든 게이트들은 플로팅 게이트(20)를 제외한 비-플로팅 게이트들이며, 이는 그것들이 전압원에 전기적으로 접속되어 있거나 접속가능하다는 것을 의미한다. 프로그래밍은 채널 영역(18)으로부터의 가열된 전자들이 플로팅 게이트(20) 상으로 자신들을 주입하는 것에 의해 나타난다. 소거는 전자들이 플로팅 게이트(20)로부터 소거 게이트(30)로 터널링하는 것에 의해 나타난다.

[0024] 표 2는 판독, 소거, 및 프로그래밍 동작들을 수행하기 위해 메모리 셀(310)의 단자들에 인가될 수 있는 전형적인 전압 범위들을 도시한다:

[0025] [표 2]

도 3 의 플래시 메모리 셀(310)의 동작

	WL/SG	BL	CG	EG	SL
판독	1.0 내지 2 V	0.6 내지 2 V	0 내지 2.6 V	0 내지 2.6 V	0 V
소거	-0.5 V/0 V	0 V	0 V/-8 V	8 내지 12 V	0 V
프로그래밍	1 V	1 μA	8 내지 11 V	4.5 내지 9 V	4.5 내지 5 V

[0026]

[0027] 도 4는 분리형 게이트 3-게이트 메모리 셀(410)을 도시한다. 메모리 셀(410)은, 메모리 셀(410)이 별개의 제어 게이트를 갖지 않는다는 점을 제외하고는, 도 3의 메모리 셀(310)과 동일하다. 소거 동작(소거 게이트를 통하여 소거함) 및 판독 동작은, 제어 게이트 바이어스가 없다는 점을 제외하고는, 도 3의 것과 유사하다. 프로그래밍 동작은 또한 제어 게이트 바이어스 없이 행해지고, 따라서 소스 라인 상의 프로그래밍 전압은 제어 게이트 바이어스의 결여를 보상하기 위해 더 높다.

[0028] 표 3은 판독, 소거, 및 프로그래밍 동작들을 수행하기 위해 메모리 셀(410)의 단자들에 인가될 수 있는 전형적

인 전압 범위들을 도시한다:

[0029] [표 3]

도 4 의 플래시 메모리 셀(410)의 동작

	WL/SG	BL	EG	SL
판독	0.7 내지 2.2 V	0.6 내지 2 V	0 내지 2.6 V	0 V
소거	-0.5 V/0 V	0 V	11.5 V	0 V
프로그래밍	1 V	2 내지 3 μA	4.5 V	7 내지 9 V

[0030]

[0031] 도 5는 스택형 게이트 메모리 셀(510)을 도시한다. 메모리 셀(510)은, 절연 층에 의해 분리되어, 플로팅 게이트(20)가 전체 채널 영역(18) 위로 연장되고, 제어 게이트(22)가 플로팅 게이트(20) 위로 연장된다는 점을 제외하고는, 도 2의 메모리 셀(210)과 유사하다. 소거, 프로그래밍, 및 판독 동작은 메모리 셀(210)에 대해 이전에 설명된 것과 유사한 방식으로 동작한다.

[0032] 표 4는 판독, 소거, 및 프로그래밍 동작들을 수행하기 위해 메모리 셀(510)의 단자들에 인가될 수 있는 전형적인 전압 범위들을 도시한다:

[0033] [표 4]

도 5 의 플래시 메모리 셀(510)의 동작

	CG	BL	SL	P-sub
판독	2 내지 5 V	0.6 내지 2 V	0 V	0 V
소거	-8 내지 -10 V/0 V	FLT	FLT	8 내지 10 V/15 내지 20 V
프로그래밍	8 내지 12 V	3 내지 5 V	0 V	0 V

[0034]

[0035] 인공 신경망에서 위에서 설명된 비휘발성 메모리 셀들의 유형들 중 하나를 포함하는 메모리 어레이들을 이용하기 위해, 두 가지 수정이 이루어진다. 첫째, 라인들은, 하기에서 추가로 설명되는 바와 같이, 각각의 메모리 셀이 어레이 내의 다른 메모리 셀들의 메모리 상태에 악영향을 미치지 않으면서 개별적으로 프로그래밍, 소거, 및 판독될 수 있도록 구성된다. 둘째, 메모리 셀들의 연속적인 (유사한) 프로그래밍이 제공된다.

[0036] 구체적으로, 어레이 내의 각각의 메모리 셀들의 메모리 상태(즉, 플로팅 게이트 상의 전하)는, 독립적으로 그리고 다른 메모리 셀들의 교란을 최소화시킨 상태로, 완전 소거 상태로부터 완전 프로그래밍 상태로 연속적으로 변경될 수 있다. 다른 실시예에서, 어레이 내의 각각의 메모리 셀의 메모리 상태(즉, 플로팅 게이트 상의 전하)는, 독립적으로 그리고 다른 메모리 셀들의 교란을 최소화시킨 상태로, 완전 프로그래밍 상태로부터 완전 소거 상태로 연속적으로 변경될 수 있고, 그 역도 성립한다. 이는 셀 저장소가 유사하거나 또는 적어도 많은 개별 값들(예컨대, 16개 또는 64개의 상이한 값들) 중 하나를 저장할 수 있음을 의미하며, 이것은 메모리 어레이 내의 모든 셀들의 매우 정밀하고 개별적인 튜닝을 허용하고, 메모리 어레이를 신경망의 시냅스 가중치들에 대한 미세 튜닝 조정들을 저장하고 행하는 데 이상적인 것으로 되게 한다.

[0037] 비휘발성 메모리 셀 어레이들을 채용한 신경망들

[0038] 도 6은 비휘발성 메모리 어레이를 활용하는 신경망의 비제한적인 예를 개념적으로 예시한다. 이 예는 얼굴 인식 애플리케이션에 대해 비휘발성 메모리 어레이 신경망을 이용하지만, 비휘발성 메모리 어레이 기반 신경망을 이용하여 임의의 다른 적절한 애플리케이션이 구현될 수 있다.

[0039] S0은, 이 예에 대해, 5 비트 정밀도를 갖는 32x32 픽셀 RGB 이미지(즉, 각각의 색상 R, G 및 B에 대해 하나씩인 3개의 32x32 픽셀 어레이들, 각각의 픽셀은 5 비트 정밀도임)인 입력이다. S0으로부터 C1로 가는 시냅스들

(CB1)은 가중치들 및 공유 가중치들의 상이한 세트들 양측 모두를 가지며, 입력 이미지를 3x3 픽셀 중첩 필터들(커널(kernel))로 스캔하여, 필터를 1 픽셀(또는 모델별로 지시되는 바와 같이 1 초과 픽셀)만큼 시프트시킨다. 구체적으로, 이미지의 3x3 부분 내의 9개 픽셀들(즉, 필터 또는 커널로 지칭됨)에 대한 값들이 시냅스들(CB1)에 제공되고, 이에 의해, 이들 9개의 입력 값들이 적절한 가중치들과 승산되고, 그 승산의 출력들을 합산한 후, 단일 출력 값이 결정되고, 피처 맵(feature map)(C1)의 층들 중 하나의 층의 픽셀을 생성하기 위해 CB1의 제1 뉴런에 의해 제공된다. 이어서, 3x3 필터가 하나의 픽셀씩 우측으로 시프트되고(즉, 우측에 3개 픽셀들의 컬럼(column)을 추가하고, 좌측에서 3개 픽셀들의 컬럼을 뺌), 이에 의해 이러한 새롭게 위치된 필터에서의 9개 픽셀 값들이 시냅스들(CB1)에 제공되고, 이에 의해 이들은 동일한 가중치들과 승산되고, 제2 단일 출력 값이 연관된 뉴런에 의해 결정된다. 이러한 프로세스는, 3개의 모든 색상들에 대해 그리고 모든 비트들(정밀도 값들)에 대해, 3x3 필터가 전체 32x32 픽셀 이미지를 가로질러서 스캔할 때까지 계속된다. 이어서, 프로세스는, 층(C1)의 모든 피처 맵들이 계산될 때까지, 가중치들의 상이한 세트들을 사용하여 반복되어 C1의 상이한 피처 맵을 생성한다.

[0040] C1에서, 본 예에서, 각각 30x30 픽셀들을 갖는 16개 피처 맵들이 있다. 각각의 픽셀은 입력들과 커널을 승산한 것으로부터 추출된 새로운 피처 픽셀이고, 따라서 각각의 피처 맵은 2 차원 어레이이고, 따라서, 이러한 예에서, 시냅스들(CB1)은 2차원 어레이들의 16개 층들을 구성한다(본 명세서에서 언급된 뉴런 층들 및 어레이들은 반드시 물리적 관계인 것이 아니라 논리적 관계임 - 즉, 어레이들은 반드시 물리적으로 2차원 어레이들로 배향되지는 않음 - 에 유념한다). 16개 피처 맵들 각각은 필터 스캔에 적용되는 시냅스 가중치들의 상이한 16개 세트들 중 하나의 세트에 의해 생성된다. C1 피처 맵들은 모두, 경계 식별과 같은 동일한 이미지 피처의 상이한 태양들에 관한 것일 수 있다. 예를 들어, (제1 맵을 생성하는 데 사용되는 모든 스캔을 위해 공유되는 제1 가중치 세트를 사용하여 생성된) 제1 맵은 원형 에지들을 식별할 수 있고, (제1 가중치 세트와는 상이한 제2 가중치 세트를 사용하여 생성된) 제2 맵은 직사각형 에지들, 또는 특정 피처들의 종횡비 등을 식별할 수 있다.

[0041] 활성화 함수(P1)(풀링(pooling))는 C1로부터 S1로 가기 전에 적용되는데, 이는 각각의 피처 맵 내의 연속적인 비중첩 2x2 영역들로부터의 값들을 풀링한다. 풀링 스테이지의 목적은, 예를 들어 에지 위치의 의존성을 감소시키고 다음 스테이지로 가기 전에 데이터 크기를 감소시키기 위해 인근 위치를 평균하는 것이다(또는 최대 함수가 또한 사용될 수 있음). S1에는, 16개 15x15 피처 맵들(즉, 각각 15x15 픽셀들의 상이한 16개 어레이들)이 있다. S1로부터 C2로 가는 CB2 내의 시냅스들 및 연관된 뉴런들은 1 픽셀의 필터 시프트를 갖는 4x4 필터들로 S1 내의 맵들을 스캔한다. C2에는, 22개 12x12 피처 맵들이 있다. 활성화 함수(P2)(풀링)가 C2로부터 S2로 가기 전에 적용되는데, 이는 각각의 피처 맵 내의 연속적인 비중첩 2x2 영역들로부터의 값들을 풀링한다. S2에는, 22개 6x6 피처 맵들이 있다. 활성화 함수가 S2로부터 C3으로 가는 시냅스들(CB3)에서 적용되며, 여기서 C3 내의 모든 뉴런은 S2 내의 모든 맵에 연결된다. C3에는, 64개 뉴런들이 있다. C3으로부터 출력(S3)으로 가는 시냅스들(CB4)은 S3을 C3에 완전히 연결한다. S3에서의 출력은 10개 뉴런들을 포함하고, 여기서 최고 출력 뉴런이 분류를 결정한다. 이러한 출력은, 예를 들어, 원래의 이미지의 콘텐츠의 식별 또는 분류를 나타낼 수 있다.

[0042] 시냅스들의 각각의 레벨은 비휘발성 메모리 셀들의 어레이 또는 그들의 어레이의 일부분을 사용하여 구현된다. 도 7은 비휘발성 메모리 셀들을 포함하고 입력 층과 다음 층 사이의 시냅스들로서 활용되는 VMM 어레이의 블록도이다. 구체적으로, VMM(32)은 비휘발성 메모리 셀들(33)의 어레이, 소거 게이트 및 워드 라인 게이트 디코더(34), 제어 게이트 디코더(35), 비트 라인 디코더(36) 및 소스 라인 디코더(37)를 포함하며, 이들은 메모리 어레이(33)에 대한 입력들을 디코딩한다. 이 예에서의 소스 라인 디코더(37)는 또한 메모리 셀 어레이의 출력을 디코딩한다. 대안으로, 비트 라인 디코더(36)는 메모리 어레이의 출력을 디코딩할 수 있다. 메모리 어레이는 두 가지 목적들을 제공한다. 첫째, 그것은 VMM에 의해 사용될 가중치들을 저장한다. 둘째, 메모리 어레이는 입력들을 메모리 어레이에 저장된 가중치들과 유효하게 승산하고 이들을 출력 라인(소스 라인 또는 비트 라인)마다 합산하여 출력을 생성하며, 이는 다음 층으로의 입력 또는 최종 층으로의 입력일 것이다. 승산 및 합산 기능을 수행함으로써, 메모리 어레이는 별개의 승산 및 합산 로직 회로들에 대한 필요성을 무효화하고, 또한 인-시츄(in-situ) 메모리 계산으로 인해 전력 효율적이다.

[0043] 메모리 어레이의 출력은 차동 합산기(예컨대, 합산 연산 증폭기)(38)에 공급되고, 이는 메모리 셀 어레이의 출력들을 합산하여 그 컨볼루션(convolution)에 대한 단일 값을 생성한다. 차동 합산기는 예를 들어 양의 가중치 및 음의 가중치와 양의 입력의 합산을 실현하기 위한 것이다. 이어서 합산된 출력 값들은 출력을 정류하는 활성화 함수 회로(39)에 공급된다. 활성화 함수는 시그모이드, tanh 또는 ReLU 함수들을 포함할 수 있다. 정류된 출력 값들은 다음 층(예를 들어, 위의 설명의 C1)으로서 피처 맵의 요소가 되고, 이어서 다음 시냅스에 적용

되어 다음 피처 맵 층 또는 최종 층을 생성한다. 따라서, 이 예에서, 메모리 어레이는 복수의 시냅스들(이들은 이전 뉴런 층으로부터 또는 이미지 데이터베이스와 같은 입력 층으로부터 그들의 입력들을 수신함)을 구성하고, 합산 연산 증폭기(38) 및 활성화 함수 회로(39)는 복수의 뉴런들을 구성한다.

[0044] 도 8은 다양한 레벨들의 VMM의 블록도이다. 도 14에 도시된 바와 같이, 입력은 디지털-아날로그 변환기(31)에 의해 디지털로부터 아날로그로 변환되고, 입력 VMM(32a)에 제공된다. 입력 VMM(32a)에 의해 생성된 출력은 다음 VMM(은닉 레벨 1)(32b)으로의 입력으로서 제공되고 다음 VMM으로의 입력은 이어서 다음 VMM(은닉 레벨 2)(32b)으로의 입력으로서 제공되는 출력을 생성하고, 등등이다. VMM(32)의 다양한 층들은 컨볼루션 신경망(convolutional neural network, CNN)의 시냅스들 및 뉴런들의 상이한 층들로서 기능한다. 각각의 VMM은 독립형 비휘발성 메모리 어레이일 수 있거나, 또는 다수의 VMM들이 동일한 비휘발성 메모리 어레이의 상이한 부분들을 활용할 수 있거나, 또는 다수의 VMM들이 동일한 비휘발성 메모리 어레이의 중첩 부분들을 활용할 수 있다. 도 8에 도시된 예는 5개의 층들(32a,32b,32c,32d,32e), 즉 하나의 입력 층(32a), 2개의 은닉 층들(32b, 32c), 및 2개의 완전 접속 층들(32d, 32e)을 포함한다. 당업자는, 이것이 단지 예시적인 것이고, 시스템이 대신에, 2개 초과의 은닉 층들 및 2개 초과의 완전 접속 층들을 포함할 수 있다는 것을 이해할 것이다.

[0045] VMM 어레이들

[0046] 도 9는 뉴런 VMM(900)을 도시하며, 이는 도 2에 도시된 타입의 메모리 셀들에 특히 적합하고, 입력 층과 다음 층 사이의 시냅스들 및 뉴런들의 부분들로서 활용된다. VMM(900)은 비휘발성 메모리 셀들의 메모리 어레이(903), 기준 어레이(901) 및 기준 어레이(902)를 포함한다. 기준 어레이들(901, 902)은 단자들 BLR0 내지 BLR3 내로 흐르는 전류 입력들을 전압 입력들 WL0 내지 WL3으로 변환하는 역할을 한다. 도시된 바와 같은 기준 어레이들(901, 902)은 컬럼 방향으로 있다. 일반적으로, 기준 어레이 방향은 입력 라인들에 직교한다. 사실상, 기준 메모리 셀들은 전류 입력들이 그들 안으로 흐르는 멀티플렉서들(멀티플렉서 및 기준 비트 라인을 바이어싱하기 위한 캐스코딩 트랜지스터(VBLR)를 포함하는, 멀티플렉서(914))을 통해 다이오드 접속된다. 기준 셀들은 타깃 기준 레벨들로 튜닝된다.

[0047] 메모리 어레이(903)는 두 가지 목적들을 제공한다. 첫째, 그것은 VMM(900)에 의해 사용될 가중치들을 저장한다. 둘째, 메모리 어레이(903)는 입력들(단자들 BLR0 내지 BLR3에 제공되는 전류 입력들; 기준 어레이들(901, 902)은 이들 전류 입력들을 워드 라인들 WL0 내지 WL3에 공급할 입력 전압들로 변환함)을 메모리 어레이에 저장된 가중치들과 유효하게 승산하여 출력을 생성하며, 이는 다음 층으로의 입력 또는 최종 층으로의 입력일 것이다. 승산 함수를 수행함으로써, 메모리 어레이는 별개의 승산 로직 회로들에 대한 필요성을 무효화하고, 또한 전력 효율적이다. 여기서, 전압 입력들은 워드 라인들 상에 제공되고, 출력은 판독(추론) 동작 동안 비트 라인 상에서 나온다. 비트 라인 상에 배치된 전류는 비트 라인에 접속된 메모리 셀들로부터의 모든 전류들의 합산 함수를 수행한다.

[0048] 도 42는 VMM(900)에 대한 동작 전압들을 도시한다. 표 내의 컬럼들은 선택된 셀들에 대한 워드 라인들, 비선택된 셀들에 대한 워드 라인들, 선택된 셀들에 대한 비트 라인들, 비선택된 셀들에 대한 비트 라인들, 선택된 셀들에 대한 소스 라인들, 및 비선택된 셀들에 대한 소스 라인들 상에 배치된 전압들을 나타낸다. 로우(row)들은 판독, 소거 및 프로그래밍의 동작들을 나타낸다.

[0049] 도 10은 뉴런 VMM(1000)을 도시하며, 이는 도 2에 도시된 타입의 메모리 셀들에 특히 적합하고, 입력 층과 다음 층 사이의 시냅스들 및 뉴런들의 부분들로서 활용된다. VMM(1000)은 비휘발성 메모리 셀들의 메모리 어레이(1003), 기준 어레이(1001) 및 기준 어레이(1002)를 포함한다. VMM(1000)은, VMM(1000)에서 워드 라인들이 수직 방향으로 이어진다는 점을 제외하고는, VMM(900)과 유사하다. 2개의 기준 어레이(1001)(상부에 있음, 이는 입력 전류를 짝수 로우들에 대한 전압으로 변환하는 기준을 제공함) 및 기준 어레이(1002)(하부에 있음, 이는 입력 전류를 홀수 로우들에 대한 전압으로 변환하는 기준을 제공함)가 존재한다. 여기서, 입력들은 워드 라인들 상에 제공되고, 출력은 판독 동작 동안 소스 라인 상에서 나온다. 소스 라인 상에 배치된 전류는 소스 라인에 접속된 메모리 셀들로부터의 모든 전류들의 합산 함수를 수행한다.

[0050] 도 43은 VMM(1000)에 대한 동작 전압들을 도시한다. 표 내의 컬럼들은 선택된 셀들에 대한 워드 라인들, 비선택된 셀들에 대한 워드 라인들, 선택된 셀들에 대한 비트 라인들, 비선택된 셀들에 대한 비트 라인들, 선택된 셀들에 대한 소스 라인들, 및 비선택된 셀들에 대한 소스 라인들 상에 배치된 전압들을 나타낸다. 로우들은 판독, 소거 및 프로그래밍의 동작들을 나타낸다.

[0051] 도 11은 뉴런 VMM(1100)을 도시하며, 이는 도 3에 도시된 타입의 메모리 셀들에 특히 적합하고, 입력 층과 다음

층 사이의 시냅스들 및 뉴런들의 부분들로서 활용된다. VMM(1100)은 비휘발성 메모리 셀들의 메모리 어레이(1101), 기준 어레이(1102)(입력 전류를 짝수 로우들에 대한 입력 전압으로 변환하는 기준을 제공함), 및 기준 어레이(1103)(입력 전류를 홀수 로우들에 대한 입력 전압으로 변환하는 기준을 제공함)를 포함한다. VMM(1100)은, VMM(1100)이 메모리 셀들의 로우의 제어 게이트들에 커플링된 제어 라인(1106) 및 메모리 셀들의 인접한 로우들의 소거 게이트들에 커플링된 제어 라인(1107)을 추가로 포함한다는 점을 제외하고는, VMM(900)과 유사하다. 여기서, 워드 라인들, 제어 게이트 라인들, 및 소거 게이트 라인들은 동일한 방향을 갖는다. VMM은, 기준 비트 라인을 선택된 기준 메모리 셀의 비트 라인 콘택트에 선택적으로 커플링하는 기준 비트 라인 선택 트랜지스터(1104)(mux(1114)의 일부), 및 기준 비트 라인을 특정 선택된 기준 메모리 셀에 대한 제어 라인(1106)에 선택적으로 커플링하는 스위치(1105)(mux(1114)의 일부)를 추가로 포함한다. 여기서, 입력들은 (메모리 어레이(1101)의) 워드 라인들 상에 제공되고, 출력은 판독 동작 동안 비트 라인(1109)과 같은 비트 라인 상에 나타난다. 비트 라인 상에 배치된 전류는 비트 라인에 접속된 메모리 셀들로부터의 모든 전류들의 합산 함수를 수행한다.

[0052] 도 44는 VMM(1100)에 대한 동작 전압들을 도시한다. 표 내의 컬럼들은 선택된 셀들에 대한 워드 라인들, 비선택된 셀들에 대한 워드 라인들, 선택된 셀들을 위한 비트 라인들, 비선택된 셀들에 대한 비트 라인들, 선택된 셀들에 대한 제어 게이트들, 선택된 셀들과 동일한 섹터 내의 비선택된 셀들에 대한 제어 게이트들, 선택된 셀들과는 상이한 섹터 내의 비선택된 셀들에 대한 제어 게이트들, 선택된 셀들에 대한 소거 게이트들, 비선택된 셀들에 대한 소거 게이트들, 선택된 셀들에 대한 소스 라인들, 및 비선택된 셀들에 대한 소스 라인들 상에 배치된 전압들을 나타낸다. 로우들은 판독, 소거 및 프로그래밍의 동작들을 나타낸다.

[0053] 도 12는 뉴런 VMM(1200)을 도시하며, 이는 도 3에 도시된 타입의 메모리 셀들에 특히 적합하고, 입력 층과 다음 층 사이의 시냅스들 및 뉴런들의 부분들로서 활용된다. VMM(1200)은, VMM(1200)에서 소거 게이트 라인(1201)과 같은 소거 게이트 라인들이 수직 방향으로 이어진다는 점을 제외하고는, VMM(1100)과 유사하다. 여기서, 입력들은 워드 라인들 상에 제공되고, 출력은 소스 라인들 상에 나타난다. 비트 라인 상에 배치된 전류는 비트 라인에 접속된 메모리 셀들로부터의 모든 전류들의 합산 함수를 수행한다.

[0054] 도 45는 VMM(1200)에 대한 동작 전압들을 도시한다. 표 내의 컬럼들은 선택된 셀들에 대한 워드 라인들, 비선택된 셀들에 대한 워드 라인들, 선택된 셀들을 위한 비트 라인들, 비선택된 셀들에 대한 비트 라인들, 선택된 셀들에 대한 제어 게이트들, 선택된 셀들과 동일한 섹터 내의 비선택된 셀들에 대한 제어 게이트들, 선택된 셀들과는 상이한 섹터 내의 비선택된 셀들에 대한 제어 게이트들, 선택된 셀들에 대한 소거 게이트들, 비선택된 셀들에 대한 소거 게이트들, 선택된 셀들에 대한 소스 라인들, 및 비선택된 셀들에 대한 소스 라인들 상에 배치된 전압들을 나타낸다. 로우들은 판독, 소거 및 프로그래밍의 동작들을 나타낸다.

[0055] 도 13은 뉴런 VMM(1300)을 도시하며, 이는 도 3에 도시된 타입의 메모리 셀들에 특히 적합하고, 입력 층과 다음 층 사이의 시냅스들 및 뉴런들의 부분들로서 활용된다. VMM(1300)은 비휘발성 메모리 셀들의 메모리 어레이(1301) 및 (어레이의 상부에 있는) 기준 어레이(1302)를 포함한다. 대안으로, 다른 기준 어레이가, 도 10의 것과 유사하게, 하부에 배치될 수 있다. 다른 면에서, VMM(1300)은, VMM(1300)에서 제어 게이트 라인(1303)과 같은 제어 게이트 라인들이 수직 방향으로 이어지고(따라서, 입력 제어 게이트 라인들에 직교하는, 로우 방향의 기준 어레이(1302)), 소거 게이트 라인(1304)과 같은 소거 게이트 라인들이 수평 방향으로 이어진다는 점을 제외하고는, VMM(1200)과 유사하다. 여기서, 입력들은 제어 게이트 라인들 상에 제공되고, 출력은 소스 라인들 상에서 나온다. 일 실시예에서 짝수 로우들만이 사용되고, 다른 실시예에서 홀수 로우들만이 사용된다. 소스 라인 상에 배치된 전류는 소스 라인에 접속된 메모리 셀들로부터의 모든 전류들의 합산 함수를 수행한다.

[0056] 신경망에 대해 본 명세서에 설명된 바와 같이, 플래시 셀들은 바람직하게는 서브-임계 영역에서 동작하도록 구성된다.

[0057] 본 명세서에 기술된 메모리 셀들은 하기와 같은 약 반전(weak inversion)으로 바이어싱된다:

[0058] $Ids = Io * e^{(Vg- Vth)/kVt} = w * Io * e^{(Vg)/kVt}$

[0059] $w = e^{(- Vth)/kVt}$

[0060] 메모리 셀을 사용하는 I-V(I-to-V) 로그 변환기가 입력 전류를 입력 전압으로 변환하도록 하기 위해:

[0061] $Vg= k*Vt*\log [Ids/wp*Io]$

[0062] 벡터-매트릭스 승산기(VMM)로서 사용되는 메모리 어레이의 경우, 출력 전류는 하기와 같다:

[0063] $Iout = wa * Io * e^{(Vg)/kVt}$, 즉

[0064] $Iout = (wa/wp) * Iin = W * Iin$

[0065] $W = e^{(Vthp - Vtha)/kVt}$

[0066] 워드 라인 또는 제어 게이트가 입력 전압을 위해 메모리 셀에 대한 입력으로서 사용될 수 있다.

[0067] 대안으로, 플래시 메모리 셀들은 선형 영역에서 동작하도록 구성될 수 있다:

[0068] $Ids = beta* (Vgs-Vth)*Vds$; $beta = u*Cox*W/L$

[0069] $W \alpha (Vgs-Vth)$

[0070] I-V 선형 변환기에 대해, 선형 영역에서 동작하는 메모리 셀은 입력/출력 전류를 입력/출력 전압으로 선형으로 변환하는 데 사용될 수 있다.

[0071] ESF 벡터 매트릭스 승산기에 대한 다른 실시예들은, 본 명세서에 참고로 포함되는 미국 특허 출원 제15/826,345호에 기재된 바와 같다. 소스 라인 또는 비트 라인이 뉴런 출력(전류 합산 출력)으로서 사용될 수 있다.

[0072] 도 14는 비트 라인 디코더 회로(1400)의 일 실시예를 도시한다. 비트 라인 디코더 회로(1400)는 컬럼 디코더(1402) 및 아날로그 뉴로모픽 뉴런(analog neuromorphic neuron, "ANN") 컬럼 디코더(1403)를 포함하며, 이들 각각은 VMM 어레이(1401)에 커플링된다. VMM 어레이는 앞서 논의된 VMM 설계(예컨대, VMM(900, 1000, 1100, 1200, 1300)) 또는 다른 VMM 설계들 중 임의의 것에 기초할 수 있다.

[0073] 아날로그 뉴로모픽 시스템들에서의 하나의 난제는, 시스템이 개별 선택된 셀들을 프로그래밍 및 검증할 수 있어야 하고(이는 판독 동작을 수반함), 그것이 또한 어레이 내의 모든 셀들이 선택 및 판독되는 ANN 판독을 수행할 수 있어야 한다는 것이다. 다시 말하면, 비트 라인 디코더는 때때로 하나의 비트 라인만을 선택해야 하고, 다른 경우들에서는, 모든 비트 라인들을 선택해야 한다.

[0074] 비트 라인 디코더 회로(1400)는 이러한 목적을 달성한다. 컬럼 디코더(1402)는 종래의 컬럼 디코더(프로그래밍 및 소거, 즉 PE, 디코딩 경로)이고, 예컨대 프로그래밍 및 프로그래밍 검증(감지 동작)을 위해, 개별 비트 라인을 선택하는 데 사용될 수 있다. 컬럼 디코더(1402)의 출력들은 프로그래밍, 프로그래밍/소거(PE) 검증, 및 소거를 제어하기 위해 PE 컬럼 드라이버 회로(도 14에 도시되지 않음)에 커플링된다. ANN 컬럼 디코더(1403)는 매 비트 라인에 대한 판독 동작을 동시에 인에이블시키도록 특별히 설계된 컬럼 디코더이다. ANN 컬럼 디코더(1403)는 비트 라인(여기서는, BL0)에 커플링된 예시적인 선택 트랜지스터(1405) 및 출력 회로(예컨대, 전류 합산기 및 활성화 함수, 예를 들어 tanh, S상(sigmoid), ReLU)(1406)를 포함한다. 동일한 디바이스들의 세트가 다른 비트 라인들 각각에 부착된다. 선택 트랜지스터들(1405)과 같은 선택 트랜지스터들 모두는 선택 라인(1404)에 커플링된다. ANN 판독 동작 동안, 선택 라인(1404)이 인에이블되는데, 이는 선택 트랜지스터(1405)와 같은 선택 트랜지스터들 각각을 턴온시키고, 이들 선택 트랜지스터들 각각은, 이어서, 각각의 비트 라인으로부터의 전류가 회로(1406)와 같은 출력 회로에 의해 수신되고 출력되게 한다.

[0075] 도 15는 비트 라인 디코더 회로(1500)의 일 실시예를 도시한다. 비트 라인 디코더 회로(1500)는 VMM 어레이(1501)에 커플링된다. VMM 어레이는 앞서 논의된 VMM 설계(예컨대, VMM(900, 1000, 1100, 1200, 1300)) 또는 다른 VMM 설계들 중 임의의 것에 기초할 수 있다.

[0076] 선택 트랜지스터들(1502, 1503)은 한 쌍의 상보적 제어 신호들 V0 및 VB_0에 의해 제어되고, 비트 라인 BL0에 커플링된다. 선택 트랜지스터들(1504, 1505)은 다른 쌍의 상보적 제어 신호들 V1 및 VB_1에 의해 제어되고, 다른 비트 라인 BL1에 커플링된다. 선택 트랜지스터들(1502, 1504)은, 예컨대 프로그래밍을 인에이블시키기 위해 동일한 출력에 커플링되고, 선택 트랜지스터들(1503, 1505)은, 예컨대 프로그래밍을 금지하기 위해 동일한 출력에 커플링된다. 트랜지스터들(1502/1503/1504/1505)의 출력 라인들(프로그래밍 및 소거 PE 디코딩 경로)은, 예컨대, 프로그래밍, PE 검증, 및 소거를 제어하기 위한 PE 컬럼 드라이버 회로(도시되지 않음)에 커플링된다.

[0077] 선택 트랜지스터(1506)는 비트 라인 BL0에 그리고 출력 및 활성화 함수 회로(1507)(예컨대, 전류 합산기 및 활성화 함수, 예를 들어 tanh, S상, ReLU)에 커플링된다. 선택 트랜지스터(1506)는 제어 라인(1508)에 의해 제어된다.

[0078] BL0만이 활성화되어야 할 때, 제어 라인(1508)이 디어서트(de-assert)되고 신호 V0이 어서트되어, 그에 따라 BL0만을 판독한다. ANN 판독 동작 동안, 제어 라인(1508)이 어서트되고, 선택 트랜지스터(1506) 및 유사한 트랜지스터들이 턴온되고, 모든 비트 라인들이, 예컨대 모든 뉴런 프로세싱을 위해 판독된다.

[0079] 도 16은 비트 라인 디코더 회로(1600)의 일 실시예를 도시한다. 비트 라인 디코더 회로(1600)는 VMM 어레이(1601)에 커플링된다. VMM 어레이는 앞서 논의된 VMM 설계(예컨대, VMM(900, 1000, 1100, 1200, 1300)) 또는 다른 VMM 설계들 중 임의의 것에 기초할 수 있다.

[0080] 선택 트랜지스터(1601)는 비트 라인 BL0에 그리고 출력 및 활성화 함수 회로(1603)에 커플링된다. 선택 트랜지스터(1602)는 비트 라인 BL0에 그리고 공통 출력(PE 디코딩 경로)에 커플링된다.

[0081] BL0만이 활성화되어야 할 때, 선택 트랜지스터(1602)가 활성화되고, BL0이 공통 출력에 부착된다. ANN 판독 동작 동안, 선택 트랜지스터(1601) 및 유사한 트랜지스터들이 턴온되고, 모든 비트 라인들이 판독된다.

[0082] 도 14, 도 15 및 도 16에서의 디코딩의 경우, 선택되지 않은 트랜지스터에 대해, 트랜지스터 누설이 메모리 셀 성능에 영향을 미치는 것을 감소시키도록 네거티브 바이어스가 인가될 수 있다. 또는 네거티브 바이어스는 어레이가 ANN 동작 중인 동안 PE 디코딩 경로에 인가될 수 있다. 네거티브 바이어스는 -0.1 V 내지 -0.5 V 이상일 수 있다.

[0083] 도 17은 VMM 시스템(1700)을 도시한다. VMM 시스템(1700)은, VMM 어레이(1701) 및 기준 어레이(1720)(이는 앞서 논의된 VMM 설계, 예컨대 VMM(900, 1000, 1100, 1200, 1300) 또는 다른 VMM 설계들 중 임의의 것에 기초할 수 있음), 저전압 로우 디코더(1702), 고전압 로우 디코더(1703), 기준 셀 저전압 컬럼 디코더(1704)(컬럼 방향으로의 기준 어레이에 대해 도시되며, 이는 로우 방향으로의 입력-출력 변환을 제공하는 것을 의미함), 비트 라인 PE 드라이버(1712), 비트 라인 멀티플렉서(1706), 활성화 함수 회로 및 합산기(1707), 제어 로직(1705), 및 아날로그 바이어스 회로(1708)를 포함한다.

[0084] 도시된 바와 같이, 기준 셀 저전압 컬럼 디코더(1704)는 컬럼 방향으로의 기준 어레이(1720)에 대한 것이며, 이는 로우 방향으로의 입력-출력 변환을 제공하는 것을 의미한다. 기준 어레이가 로우 방향으로의 것인 경우, 기준 디코더는 컬럼 방향으로의 입력-출력 변환을 제공하기 위해 어레이의 상부 및/또는 하부에서 이행될 필요가 있다.

[0085] 저전압 로우 디코더(1702)는 판독 및 프로그래밍 동작들을 위한 바이어스 전압을 제공하고, 고전압 로우 디코더(1703)에 대한 디코딩 신호를 제공한다. 고전압 로우 디코더(1703)는 프로그래밍 및 소거 동작들을 위한 고전압 바이어스 신호를 제공한다. 기준 셀 저전압 컬럼 디코더(1704)는 기준 셀들에 대한 디코딩 기능을 제공한다. 비트 라인 PE 드라이버(1712)는 프로그래밍, 검증, 및 소거 시에 비트 라인에 대한 제어 기능을 제공한다. 바이어스 회로(1705)는 다양한 프로그래밍, 소거, 프로그래밍 검증, 및 판독 동작들에 필요한 다수의 전압들을 제공하는 공유 바이어스 블록이다.

[0086] 도 18은 VMM 시스템(1800)을 도시한다. VMM 시스템(1800)은, VMM 시스템(1800)이 적색 어레이(1801), 비트 라인 PE 드라이버 BLDRV(1802), 고전압 컬럼 디코더(1803), NVR 섹터들(1804), 및 기준 어레이(1820)를 추가로 포함한다는 점을 제외하고는, VMM 시스템(1700)과 유사하다. 고전압 컬럼 디코더(1803)는 수직 디코딩 라인들에 대한 고전압 바이어스를 제공한다. 적색 어레이(1802)는 결함 어레이 부분을 대체하기 위한 어레이 리던던시를 제공한다. NVR(비휘발성 레지스터, 정보 섹터로도 알려짐) 섹터들(1804)은, 사용자 정보, 디바이스 ID, 패스워드, 보안 키, 트리밍비트(trimbit)들, 구성 비트들, 제조 정보 등을 저장하는 데 사용되는 어레이 섹터들인 섹터들이다.

[0087] 도 19는 VMM 시스템(1900)을 도시한다. VMM 시스템(1900)은, VMM 시스템(1900)이 기준 시스템(1999)을 추가로 포함한다는 점을 제외하고는, VMM 시스템(1800)과 유사하다. 기준 시스템(1999)은 기준 어레이(1901), 기준 어레이 저전압 로우 디코더(1902), 기준 어레이 고전압 로우 디코더(1903), 및 기준 어레이 저전압 컬럼 디코더(1904)를 포함한다. 기준 시스템은 다수의 VMM 시스템들에 걸쳐서 공유될 수 있다. VMM 시스템은 NVR 섹터들(1905)을 추가로 포함한다.

[0088] 기준 어레이 저전압 로우 디코더(1902)는 기준 어레이(1901)를 수반하는 판독 및 프로그래밍 동작들을 위한 바이어스 전압을 제공하고, 또한, 기준 어레이 고전압 로우 디코더(1903)에 대한 디코딩 신호를 제공한다. 기준 어레이 고전압 로우 디코더(1903)는 기준 어레이(1901)를 수반하는 프로그래밍 및 동작들을 위한 고전압 바이어스를 제공한다. 기준 어레이 저전압 컬럼 디코더(1904)는 기준 어레이(1901)에 대한 디코딩 기능을 제공한다.

기준 어레이(1901)는, 예컨대 프로그래밍 검증 또는 셀 마지닝(margining)(마진 셀들을 검색함)에 대한 기준 타깃을 제공하기 위한 것이다.

[0089] 도 20은 워드 라인 드라이버(2000)를 도시한다. 워드 라인 드라이버(2000)는 (여기에 도시된 예시적인 워드 라인들(WL0, WL1, WL2, WL3)과 같은) 워드 라인을 선택하고, 그 워드 라인에 바이어스 전압을 제공한다. 각각의 워드 라인은 제어 라인(2001)에 의해 제어되는 선택 iso(격리) 트랜지스터(2002)와 같은 선택 트랜지스터에 부착된다. iso 트랜지스터(2002)는, 예컨대 소거로부터의 고전압(예컨대, 8 내지 12 V)을, IO 트랜지스터들(예컨대, 1.8 V, 3.3 V)로 구현될 수 있는 워드 라인 디코딩 트랜지스터들로부터 격리시키는 데 사용된다. 여기서, 임의의 동작 동안, 제어 라인(2001)이 활성화되고, 선택 iso 트랜지스터(2002)와 유사한 모든 선택 트랜지스터들이 턴온된다. 예시적인 바이어스 트랜지스터(2003)(워드 라인 디코딩 회로의 일부)가 워드 라인을 제1 바이어스 전압(예컨대, 3 V)에 선택적으로 커플링시켰고, 예시적인 바이어스 트랜지스터(2004)(워드 라인 디코딩 회로의 일부)가 워드 라인을 제2 바이어스 전압(제1 바이어스 전압보다 낮음, 이는 접지, 사이의 바이어스, 사용되지 않은 메모리 로우들로부터의 누설을 감소시키기 위한 네거티브 전압 바이어스를 포함함)에 선택적으로 커플링시켰다. ANN 판독 동작 동안, 모든 사용된 워드 라인들이 선택될 것이고 제1 바이어스 전압과 관련될 것이다. 모든 사용되지 않은 워드 라인들이 제2 바이어스 전압과 관련된다. 예컨대 프로그래밍 동작을 위한 다른 동작들 동안, 하나의 워드 라인만이 선택될 것이고, 다른 워드 라인들은 어레이 누설을 감소시키기 위해 네거티브 바이어스(예컨대, -0.3 내지 -0.5 V 이상)일 수 있는 제2 바이어스 전압과 관련될 것이다.

[0090] 도 21은 워드 라인 드라이버(2100)를 도시한다. 워드 라인 드라이버(2100)는, 바이어스 트랜지스터(2103)와 같은 상부 트랜지스터가 바이어스 전압에 개별적으로 커플링될 수 있고, 모든 그러한 트랜지스터들이 워드 라인 드라이버(2000)에서와 같이 함께 관련되지 않는다는 점을 제외하고는, 워드 라인 드라이버(2000)와 유사하다. 이는, 모든 워드 라인이 상이한 독립적인 전압들을 동시에 병렬로 가질 수 있게 한다.

[0091] 도 22는 워드 라인 드라이버(2200)를 도시한다. 워드 라인 드라이버(2200)는, 바이어스 트랜지스터들(2103, 2104)이 디코더 회로(2201) 및 인버터(2202)에 커플링된다는 점을 제외하고는, 워드 라인 드라이버(2100)와 유사하다. 따라서, 도 22는 워드 라인 드라이버(2200) 내의 디코딩 서브회로(2203)를 도시한다.

[0092] 도 23은 워드 라인 드라이버(2300)를 도시한다. 워드 라인 드라이버(2300)는, 바이어스 트랜지스터들(2103, 2104)이 시프트 레지스터(2301)의 스테이지(2302)의 출력들에 커플링된다는 점을 제외하고는, 워드 라인 드라이버(2100)와 유사하다. 시프트 레지스터(1301)는 데이터를 직렬 시프트-인(serial shift in)하여 (레지스터들을 직렬로 클록킹함) 각각의 로우를 독립적으로 제어할 것을 허용하여, 예컨대, 시프트-인된 데이터 패턴에 따라 하나 이상의 로우들이 동시에 인에이블될 수 있게 한다.

[0093] 도 24는 워드 라인 드라이버(2400)를 도시한다. 워드 라인 드라이버(2400)는, 각각의 선택 트랜지스터가 커패시터(2403)와 같은 커패시터에 추가로 커플링된다는 점을 제외하고는, 워드 라인 드라이버(2000)와 유사하다. 커패시터(2403)는, 트랜지스터(2401)에 의해 라인(2440) 상의 샘플 전압으로 인에이블되는 동작의 시작 시에, 워드 라인에 사전 충전 또는 바이어스를 제공할 수 있다. 커패시터(2403)는 각각의 워드 라인에 대한 입력 전압을 샘플-앤드-홀드(sample and hold, S/H)하도록 작용한다. 트랜지스터들(2401)은 VMM 어레이의 ANN 동작(어레이 전류 합산기 및 활성화 함수) 동안 오프되는데, 이는 S/H 커패시터 상의 전압이 워드 라인에 대한 (플로팅) 전압원으로서의 역할을 할 것이라는 것을 의미한다. 대안으로, 커패시터(2403)는 메모리 어레이로부터의 워드 라인 커패시턴스에 의해 제공될 수 있다.

[0094] 도 25는 워드 라인 드라이버(2500)를 도시한다. 워드 라인 드라이버(2500)는, 바이어스 트랜지스터들(2501, 2502)이 스위치들(2503, 2504)에 각각 접속된다는 점을 제외하고는, 앞서 기술된 워드 라인 드라이버들과 유사하다. 스위치(2503)는 opa(연산 증폭기)(2505)의 출력을 수신하고, 스위치(2504)는 기준 입력을 opa(2505)의 네거티브 입력에 제공하는데, 이는, 본질적으로, opa(2505), 트랜지스터(2501), 스위치들(2503, 2504)에 의해 제공되는 폐루프의 작용에 의해, 커패시터(2403)에 의해 저장된 전압을 제공한다. 이러한 방식으로, 스위치들(2503, 2504)이 폐쇄될 때, 입력(2506) 상의 전압은 트랜지스터(2501)에 의해 커패시터(2403) 상에 중첩된다. 대안으로, 커패시터(2403)는 메모리 어레이로부터의 워드 라인 커패시턴스에 의해 제공될 수 있다.

[0095] 도 26은 워드 라인 드라이버를 도시한다. 워드 라인 드라이버(2600)는, 전압을 워드 라인 WL0 내로 구동하기 위해 커패시터(2604) 상의 전압에 대한 전압 버퍼로서 작용할 증폭기(2601)의 추가를 제외하고는, 앞서 기술된 워드 라인 드라이버들과 유사하며, 이는 S/H 커패시터 상의 전압이 워드 라인에 대한 (플로팅) 전압원으로서의 역할을 할 것이라는 것을 의미한다. 이는, 예를 들어, 워드 라인-워드 라인 커플링이 커패시터 상의 전압에 영향을 미치는 것을 회피시키기 위한 것이다.

[0096] 도 27은 고전압 소스 라인 디코더 회로(2700)를 도시한다. 고전압 소스 라인 디코더 회로는 도시된 바와 같이 구성된 트랜지스터들(2701, 2702, 2703)을 포함한다. 트랜지스터(2703)는 소스 라인을 저전압에 대해 선택해제하는 데 사용된다. 트랜지스터(2702)는 어레이의 소스 라인 내로 고전압을 구동하는 데 사용되고, 트랜지스터(2701)는 소스 라인 상의 전압을 모니터링하는 데 사용된다. 트랜지스터들(2702, 2701) 및 드라이버 회로(예컨대, opa)는, PVT(프로세스, 전압, 온도) 및 변동되는 전류 부하 조건에 걸쳐서 전압을 유지하도록 폐루프 방식(힘/감지)으로 구성된다. SLE(구동된 소스 라인 노드) 및 SLB(모니터링된 소스 라인 노드)가 소스 라인의 일 단부에 있을 수 있다. 대안으로, SLE는 소스 라인의 일 단부에 있을 수 있고, SLN은 소스 라인의 다른 단부에 있을 수 있다.

[0097] 도 28은 도 2에 도시된 타입의 메모리 셀들과 함께 사용하는 데 적합한, 워드 라인 디코더 회로(2801), 소스 라인 디코더 회로(2804), 및 고전압 레벨 시프터(2808)를 포함하는 VMM 고전압 디코드 회로들을 도시한다.

[0098] 워드 라인 디코더 회로(2801)는 도시된 바와 같이 구성된 PMOS 선택 트랜지스터(2802)(신호 HVO_B에 의해 제어됨) 및 NMOS 선택해제 트랜지스터(2803)(신호 HVO_B에 의해 제어됨)를 포함한다.

[0099] 소스 라인 디코더 회로(2804)는 도시된 바와 같이 구성된 NMOS 모니터 트랜지스터들(2805)(신호 HVO에 의해 제어됨), 구동 트랜지스터(2806)(신호 HVO에 의해 제어됨), 및 선택해제 트랜지스터(2807)(신호 HVO_B에 의해 제어됨)를 포함한다.

[0100] 고전압 레벨 시프터(2808)는 인에이블 신호 EN을 수신하였고, 고전압 신호 HV 및 그의 상보물 HVO_B를 출력한다.

[0101] 도 29는 도 3에 도시된 타입의 메모리 셀들과 함께 사용하는 데 적합한, 소거 게이트 디코더 회로(2901), 제어 게이트 디코더 회로(2904), 소스 라인 디코더 회로(2907), 및 고전압 레벨 시프터(2911)를 포함하는 VMM 고전압 디코드 회로들을 도시한다.

[0102] 소거 게이트 디코더 회로(2901) 및 제어 게이트 디코더 회로(2904)는 도 28에서의 워드 라인 디코더 회로(2801)와 동일한 설계를 사용한다.

[0103] 소스 라인 디코더 회로(2907)는 도 28에서의 소스 라인 디코더 회로(2804)와 동일한 설계를 사용한다.

[0104] 고전압 레벨 시프터(2911)는 도 28에서의 고전압 레벨 시프터(2808)와 동일한 설계를 사용한다.

[0105] 도 30은 예시적인 워드 라인들 WL0, WL1, WL2, 및 WL3에 대한 워드 라인 디코더(300)를 도시한다. 예시적인 워드 라인 WL0은 풀업 트랜지스터(3001) 및 풀다운 트랜지스터(3002)에 커플링된다. 풀업 트랜지스터(3001)가 활성화될 때, WL0은 인에이블된다. 풀다운 트랜지스터(3002)가 활성화될 때, WL0은 디스에이블된다. 도 30의 기능은 격리 트랜지스터들이 없는 도 21의 것과 유사하다.

[0106] 도 31은 예시적인 제어 게이트 라인들 CG0, CG1, CG2, 및 CG3에 대한 제어 게이트 디코더(3100)를 도시한다. 예시적인 제어 게이트 라인 CG0은 풀업 트랜지스터들(3101) 및 풀다운 트랜지스터(3102)에 커플링된다. 풀업 트랜지스터(3101)가 활성화될 때, CG0은 인에이블된다. 풀다운 트랜지스터(3102)가 활성화될 때, CG0은 디스에이블된다. 도 31의 선택 및 선택해제 기능은 제어 게이트들의 경우에 도 30의 것과 유사하다.

[0107] 도 32는 예시적인 제어 게이트 라인들 CG0, CG1, CG2, 및 CG3에 대한 제어 게이트 디코더(3200)를 도시한다. 제어 게이트 디코더(3200)는, 제어 게이트 디코더(3200)가 각각의 제어 게이트 라인에 커플링된 커패시터(3203)와 같은 커패시터를 포함한다는 점을 제외하고는, 제어 게이트 디코더(3100)와 유사하다. 이들 샘플-앤드-홀드(S/H) 커패시터들은 동작 전에 각각의 제어 게이트 라인 상에 사전 충전 바이어스를 제공할 수 있는데, 이는 S/H 커패시터 상의 전압이 제어 게이트 라인들에 대한 (플로팅) 전압원으로서의 역할을 할 것이라는 것을 의미한다. S/H 커패시터는 메모리 셀로부터의 제어 게이트 커패시턴스에 의해 제공될 수 있다.

[0108] 도 33은 예시적인 제어 게이트 라인들 CG0, CG1, CG2, 및 CG3에 대한 제어 게이트 디코더(3300)를 도시한다. 제어 게이트 디코더(3300)는, 제어 게이트 디코더(3300)가 버퍼(3301)(예컨대, opa)를 추가로 포함한다는 점을 제외하고는, 제어 게이트 디코더(3200)와 유사하다.

[0109] 도 34는 전류-전압 회로(3400)를 도시한다. 회로는 구성된 다이오드 접속된 기준 셀 회로(3450) 및 샘플-앤드-홀드 회로(3460)를 포함한다. 회로(3450)는 입력 전류원(3401), NMOS 트랜지스터(3402), 캐스코딩 바이어스 트랜지스터(3403), 및 기준 메모리 셀(3404)을 포함한다. 샘플-앤드-홀드 회로는 스위치(3405), 및 S/H 커패시터(3406)로 이루어진다. 메모리(3404)는 다이오드 접속된 구성에서, 입력 전류를 전압으로 변환하기 위해, 예컨

대 워드 라인을 공급하기 위한 그의 비트 라인 상의 바이어스로 바이어싱된다.

[0110] 도 35는 S/H 커패시터 뒤에 증폭기(3501)를 추가한 전류-전압 회로(3400)와 유사한 전류-전압 회로(3500)를 도시한다. 전류-전압 회로(3500)는 구성된 다이오드 접속된 기준 셀 회로(3550), 샘플-앤드-홀드 회로(3470), 및 증폭기 스테이지(3562)를 포함한다.

[0111] 도 36은 다이오드 접속된 구성에서 제어 게이트에 대한 전류-전압 회로(3400)와 동일한 설계인 전류-전압 회로(3600)를 도시한다. 전류-전압 회로(3600)는 구성된 다이오드 접속된 기준 셀 회로(3650) 및 샘플-앤드-홀드 회로(3660)를 포함한다.

[0112] 도 37은 버퍼(3790)가 기준 회로(3750)와 S/H 회로(3760) 사이에 배치되는 전류-전압 회로(3700)를 도시한다.

[0113] 도 38은 다이오드 접속된 구성으로 접속된 제어 게이트를 갖는 도 35와 유사한 전류-전압 회로(3800)를 도시한다. 전류-전압 회로(3800)는 구성된 다이오드 접속된 기준 셀 회로(3550), 샘플-앤드-홀드 회로(3870), 및 증폭기 스테이지(3862)를 포함한다.

[0114] 도 39는 도 2에서의 메모리 셀에 적용된 바와 같은, 도 34와 유사한 전류-전압 회로(3900)를 도시한다. 전류-전압 회로(3900)는 구성된 다이오드 접속된 기준 셀 회로(3950) 및 샘플-앤드-홀드 회로(3960)를 포함한다.

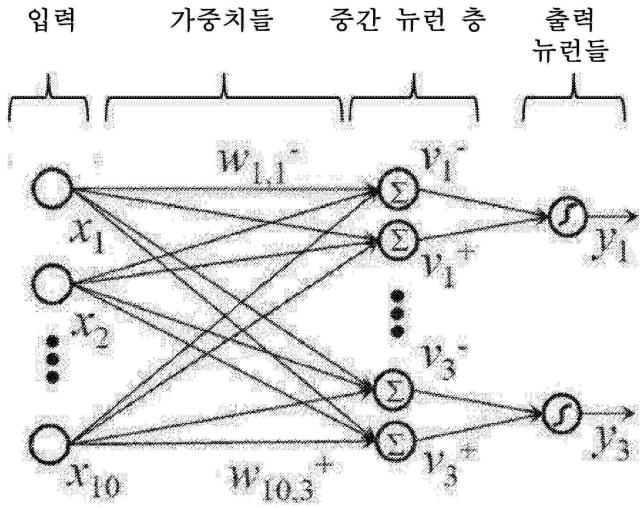
[0115] 도 40은 도 2에서의 메모리 셀에 적용되는 바와 같은, 도 37과 유사한 전류-전압 회로(4000)를 도시하며, 여기서 기준 회로(4050)와 S/H 회로(4060) 사이에 버퍼(4090)가 배치된다.

[0116] 도 41은 도 2에서의 메모리 셀에 적용된 바와 같은, 도 38과 유사한 전류-전압 회로(4100)를 도시한다. 전류-전압 회로(4100)는 구성된 다이오드 접속된 기준 셀 회로(4150), 샘플-앤드-홀드 회로(4170), 및 증폭기 스테이지(4162)를 포함한다.

[0117] 본 명세서에 사용된 바와 같이, 용어들 "~ 위에" 및 "~ 상에" 둘 모두는 "직접적으로 ~ 상에"(어떠한 중간 재료들, 요소들 또는 공간도 사이에 배치되지 않음)와 "간접적으로~ 상에"(중간 재료들, 요소들 또는 공간이 사이에 배치됨)를 포괄적으로 포함한다는 것에 유의하여야 한다. 마찬가지로, 용어 "인접한"은 "직접적으로 인접한"(어떠한 중간 재료들, 요소들 또는 공간도 사이에 배치되지 않음)과 "간접적으로 인접한"(중간 재료들, 요소들 또는 공간이 사이에 배치됨)을 포함하고, "~에 실장되는"은 "직접적으로 ~에 실장되는"(어떠한 중간 재료들, 요소들 또는 공간도 사이에 배치되지 않음)과 "간접적으로 ~에 실장되는"(중간 재료들, 요소들 또는 공간이 사이에 배치됨)을 포함하고, "전기적으로 커플링되는"은 "직접적으로 ~에 전기적으로 커플링되는"(요소들을 함께 전기적으로 접속시키는 어떠한 중간 재료들 또는 요소들도 사이에 없음)과 "간접적으로 ~에 전기적으로 커플링되는"(요소들을 함께 전기적으로 접속시키는 중간 재료들 또는 요소들이 사이에 있음)을 포함한다. 예를 들어, "기판 위에" 요소를 형성하는 것은 어떠한 중간 재료들/요소들도 사이에 갖지 않고서 직접적으로 기판 상에 요소를 형성하는 것뿐만 아니라, 하나 이상의 중간 재료들/요소들을 사이에 갖고서 간접적으로 기판 상에 요소를 형성하는 것을 포함할 수 있다.

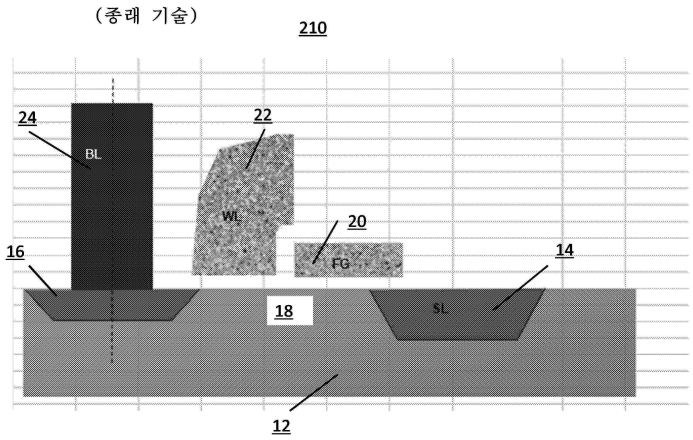
도면

도면1



(종래 기술)

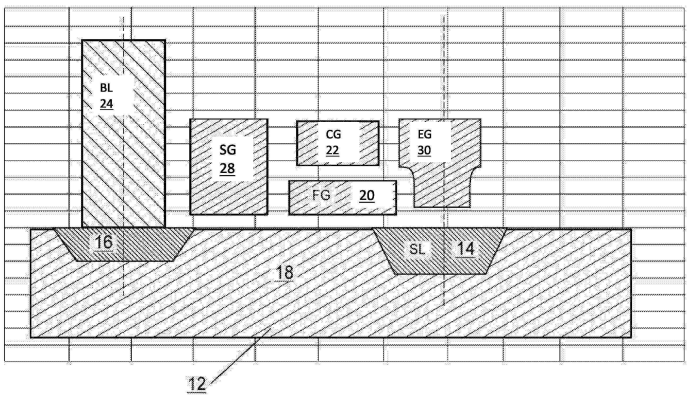
도면2



도면3

(종래 기술)

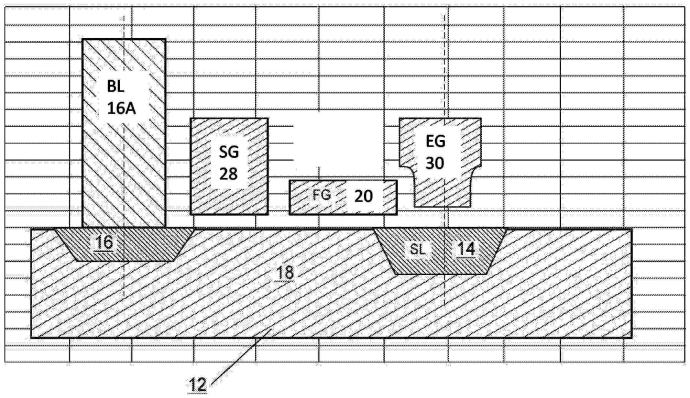
310



도면4

(종래 기술)

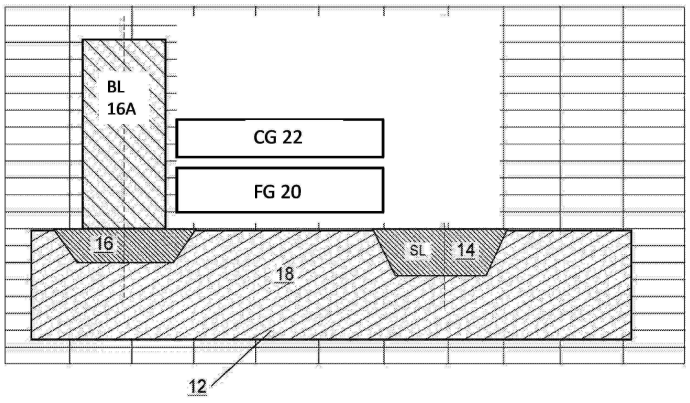
410



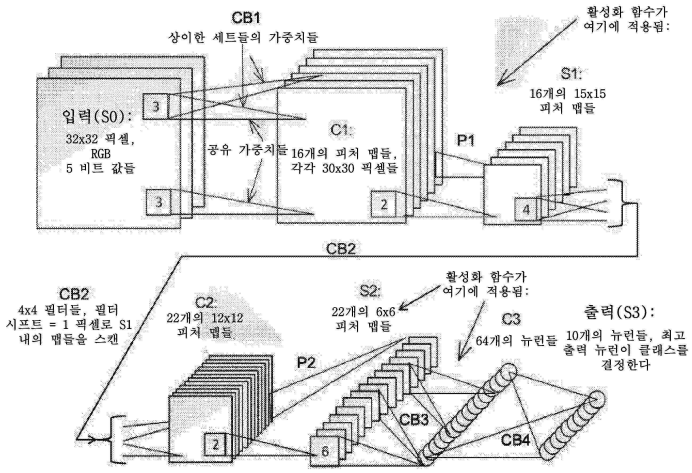
도면5

(종래 기술)

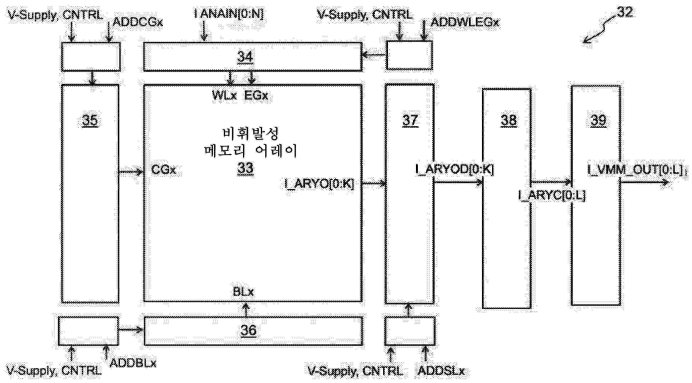
510



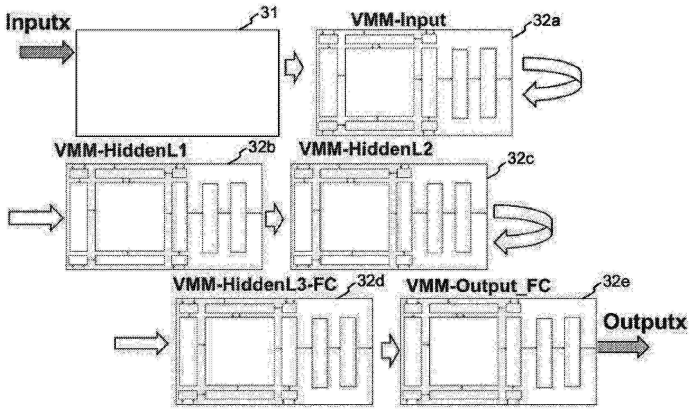
도면6


CB1
상이한 세트들의 가중치들
활성화 함수가 여기에 적용됨:
S1:
16개의 15x15 피처 맵들
입력(S0):
32x32 픽셀, RGB 5 비트 값들
공유 가중치들
C1:
16개의 피처 맵들, 각각 30x30 픽셀들
P1
CB2
CB2
4x4 필터들, 필터 시프트 = 1 픽셀로 S1 내의 맵들을 스캔
C2:
22개의 12x12 피처 맵들
S2:
22개의 6x6 피처 맵들
활성화 함수가 여기에 적용됨:
C3
64개의 뉴런들
출력(S3):
10개의 뉴런들, 최고 출력 뉴런이 클래스를 결정한다
P2
CB3
CB4

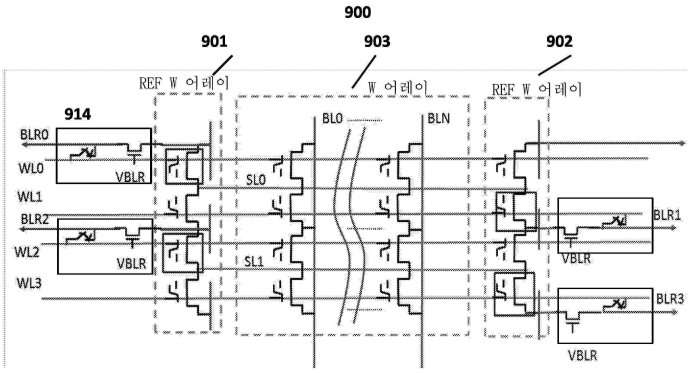
도면7


V-Supply, CNTRL
ADDCGx
I ANAIN[0:N]
V-Supply, CNTRL
ADDWLEGx
32
34
WLx EGx
35
비휘발성 메모리 어레이
33
CGx
I_ARYO[0:K]
37
I_ARYOD[0:K]
38
I_ARYC[0:L]
39
I_VMM_OUT[0:L]
BLx
36
V-Supply, CNTRL
ADDBLx
V-Supply, CNTRL
ADDSLx

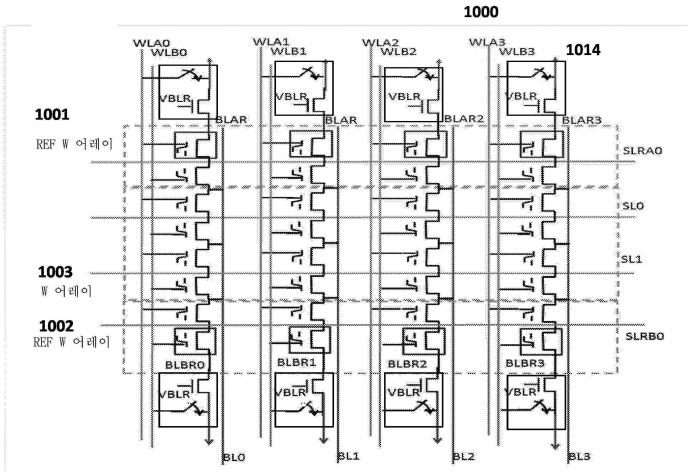
도면8


Inputx
31
VMM-Input
32a
VMM-HiddenL1
32b
VMM-HiddenL2
32c
VMM-HiddenL3-FC
32d
VMM-Output_FC
32e
Outputx

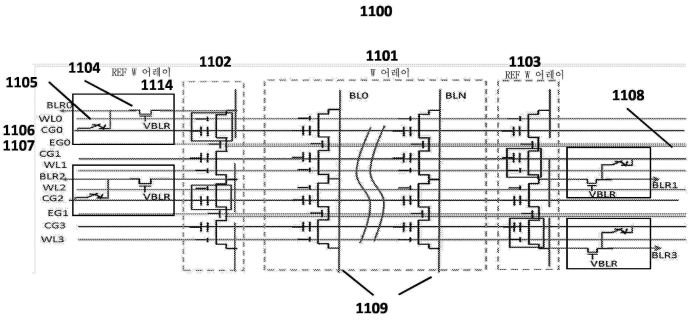
도면9



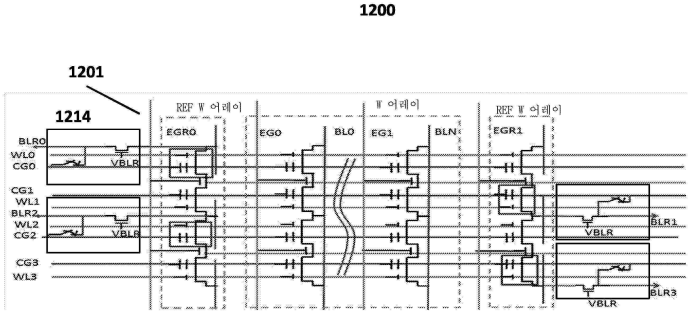
도면10



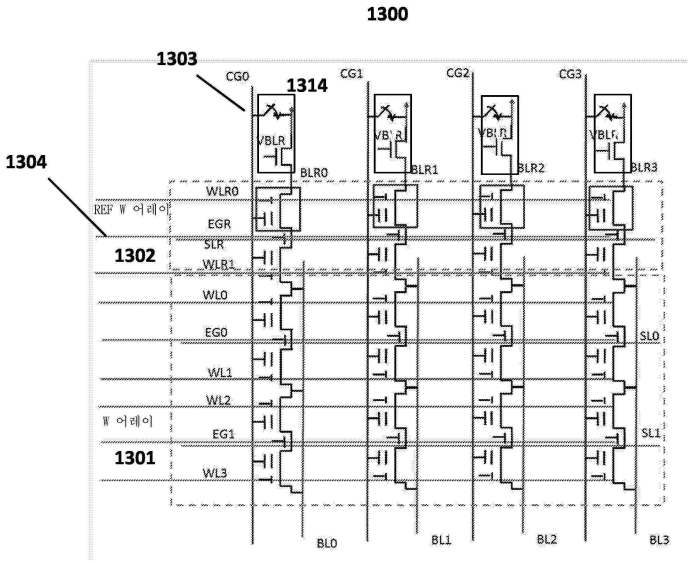
도면11



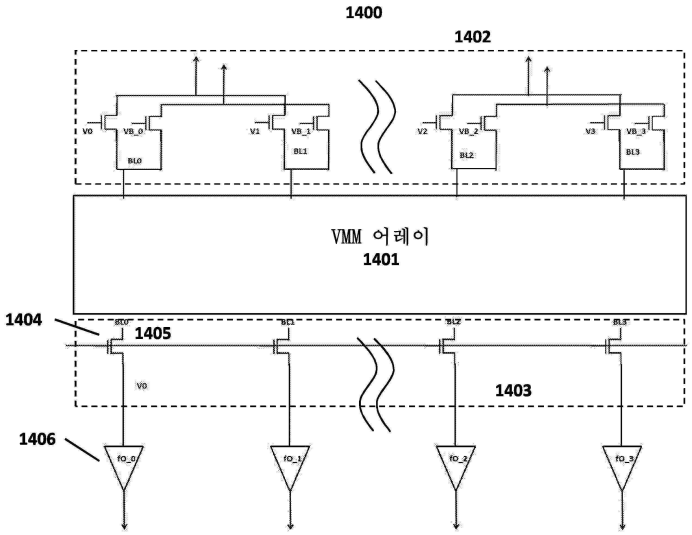
도면12



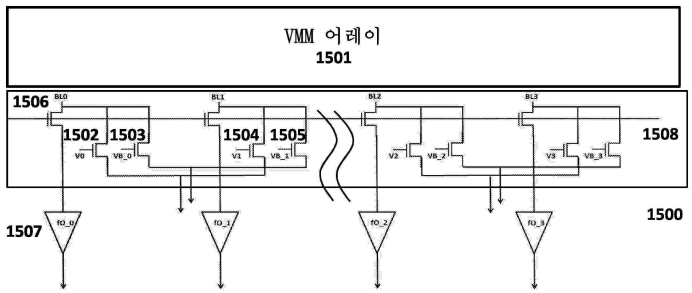
도면13



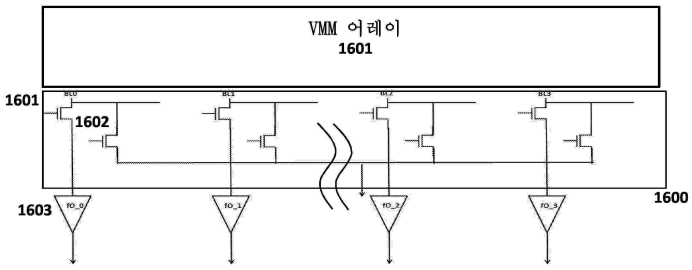
도면14



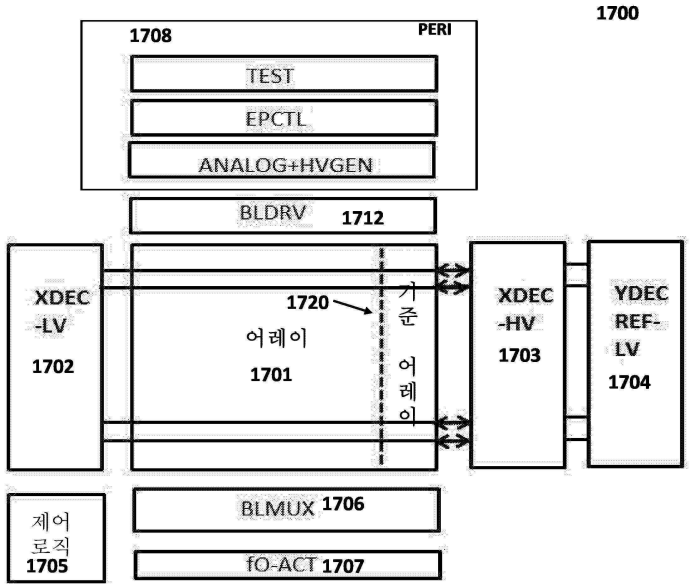
도면15



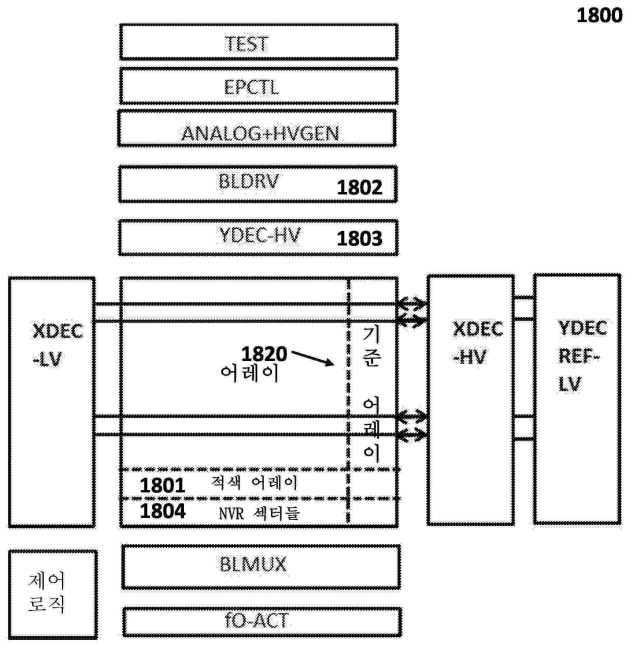
도면16



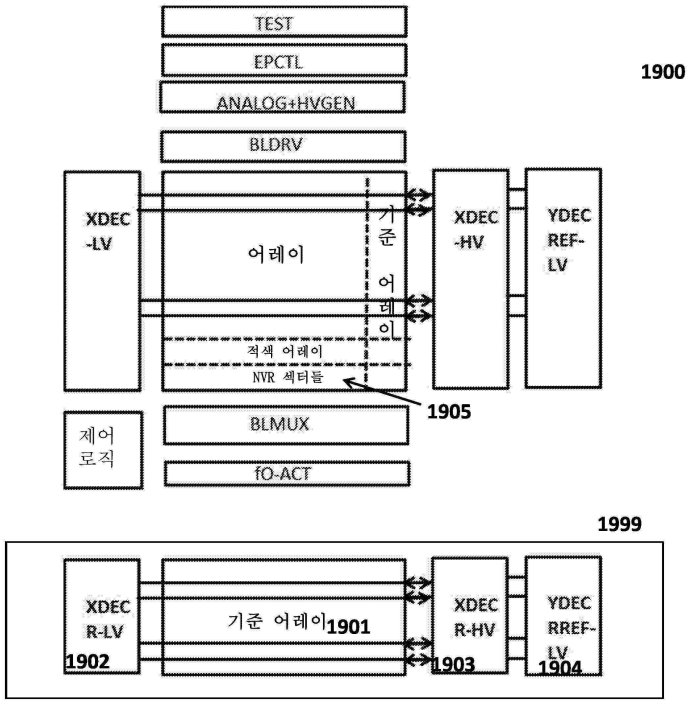
도면17



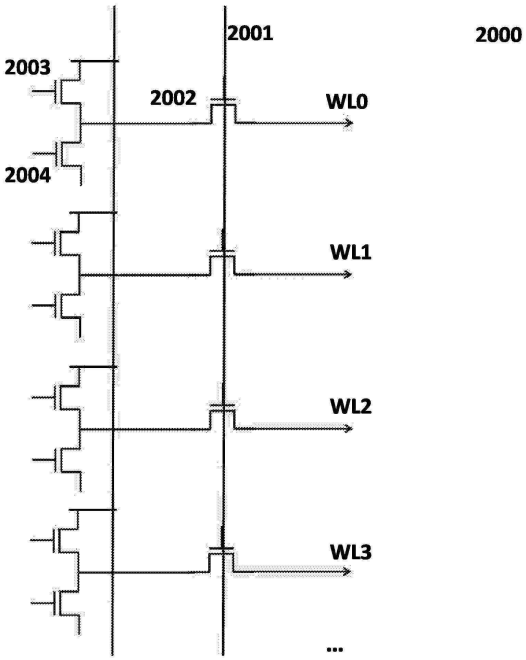
도면18



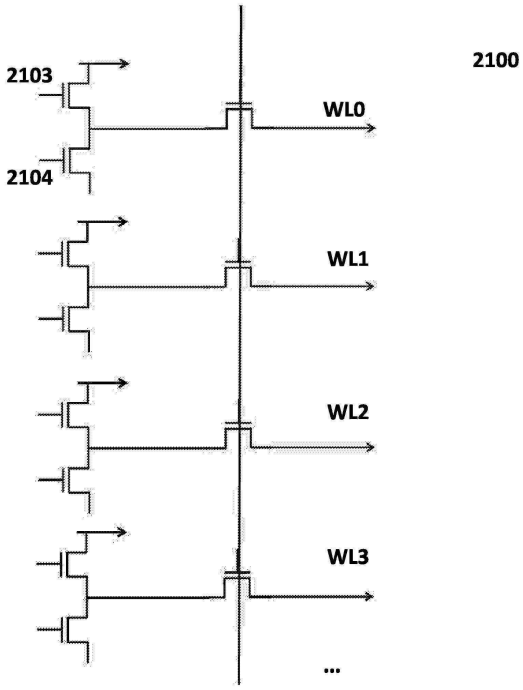
도면19


TEST
EPCTL
ANALOG+HVGEN
BLDRV
XDEC -LV
어레이
기준 어레이
XDEC -HV
YDEC REF-LV
적색 어레이
NVR 섹터들
1905
제어 로직
BLMUX
fO-ACT
1900
1999
XDEC R-LV
1902
기준 어레이
1901
XDEC R-HV
1903
YDEC RREF-LV
1904

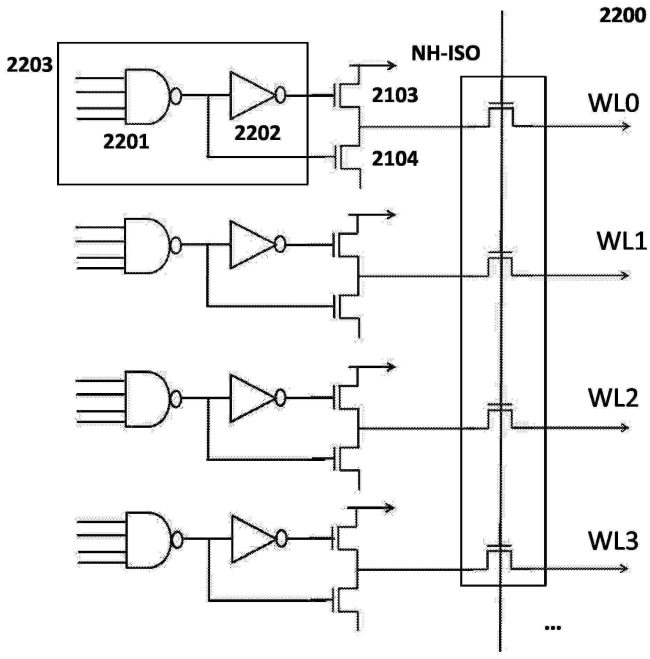
도면20


2001
2000
2003
2002
WL0
2004
WL1
WL2
WL3
...

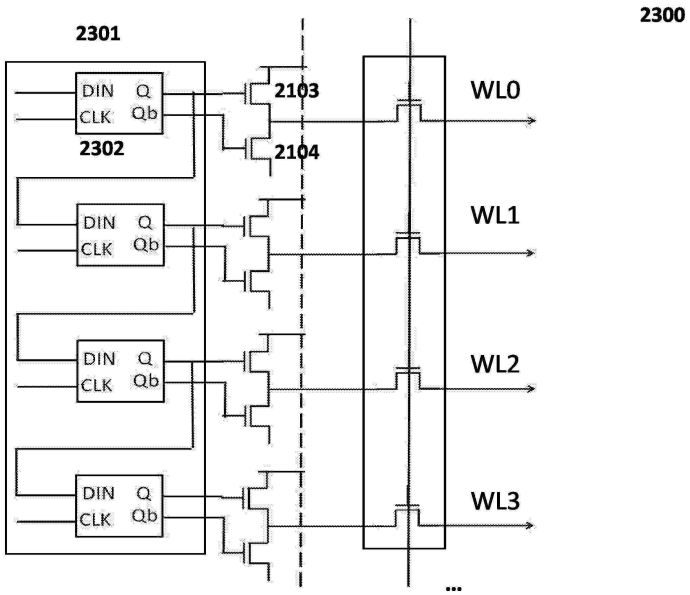
도면21


2100
2103
2104
WL0
WL1
WL2
WL3
...

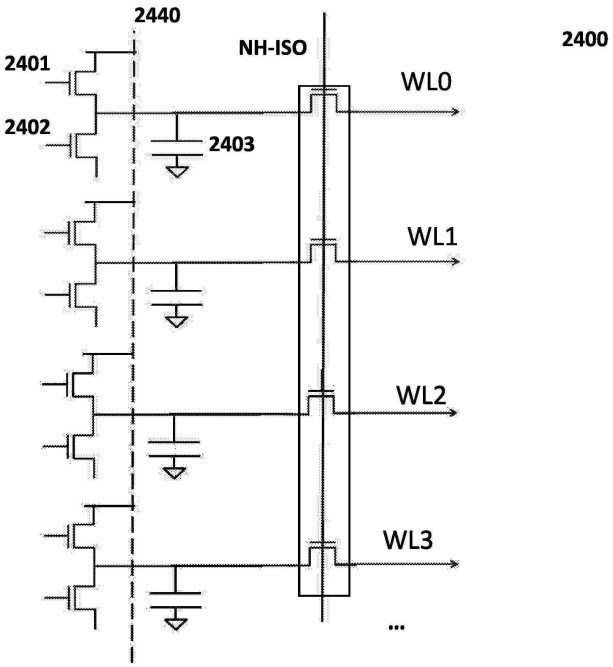
도면22


2200
2203
2201
2202
NH-ISO
2103
2104
WL0
WL1
WL2
WL3
...

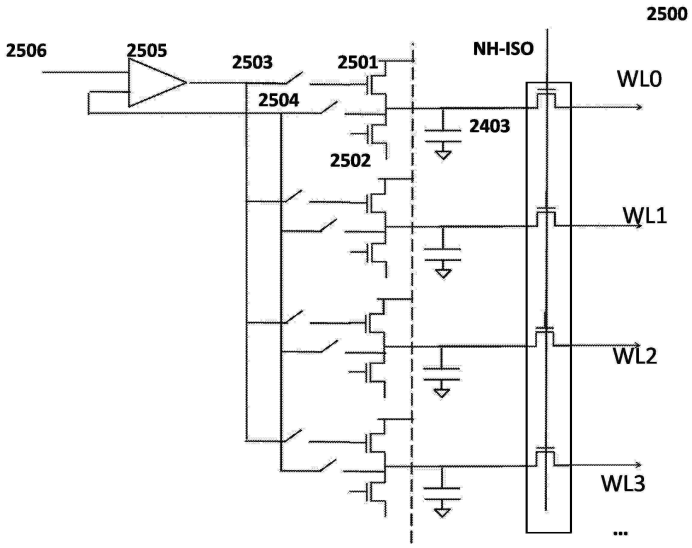
도면23


2300
2301
DIN
CLK
Q
Qb
2302
2103
2104
WL0
WL1
WL2
WL3
...

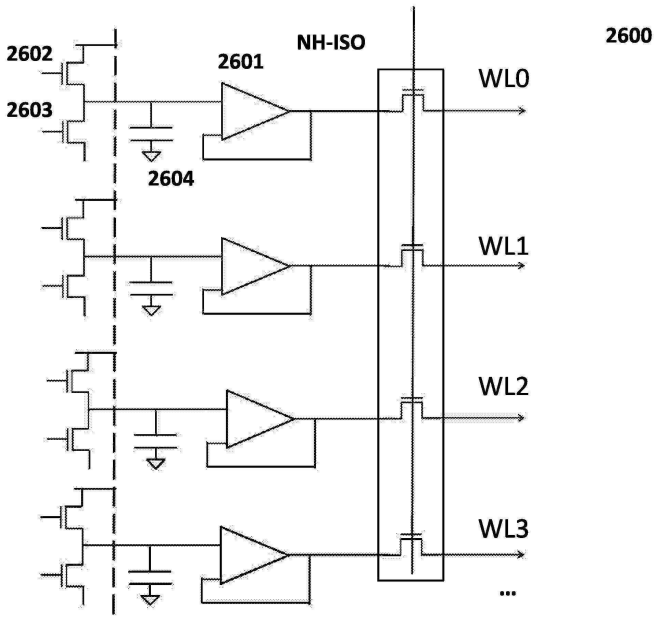
도면24


2400
2440
NH-ISO
2401
2402
2403
WL0
WL1
WL2
WL3
...

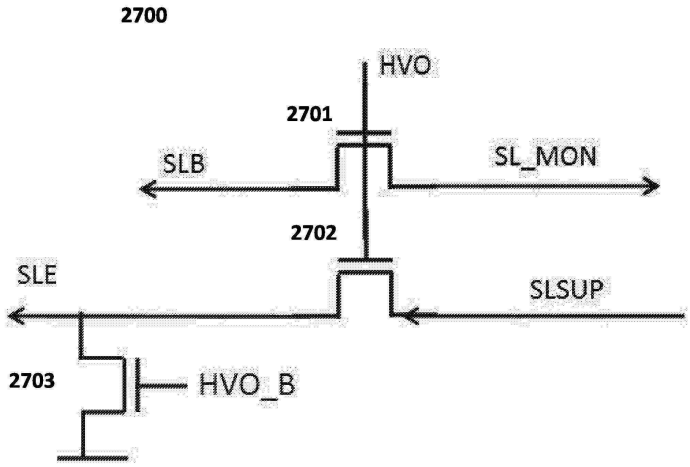
도면25


2500
2506
2505
2503
2501
NH-ISO
WL0
2504
2403
2502
WL1
WL2
WL3
...

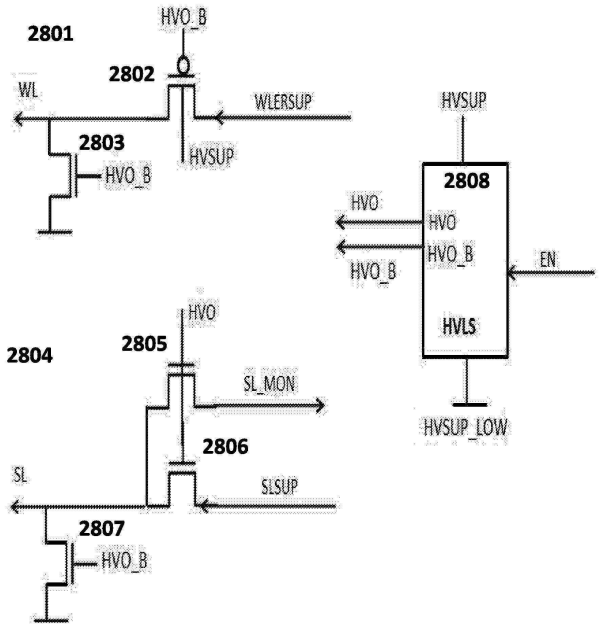
도면26


2600
2602
NH-ISO
2601
WL0
2603
2604
WL1
WL2
WL3
...

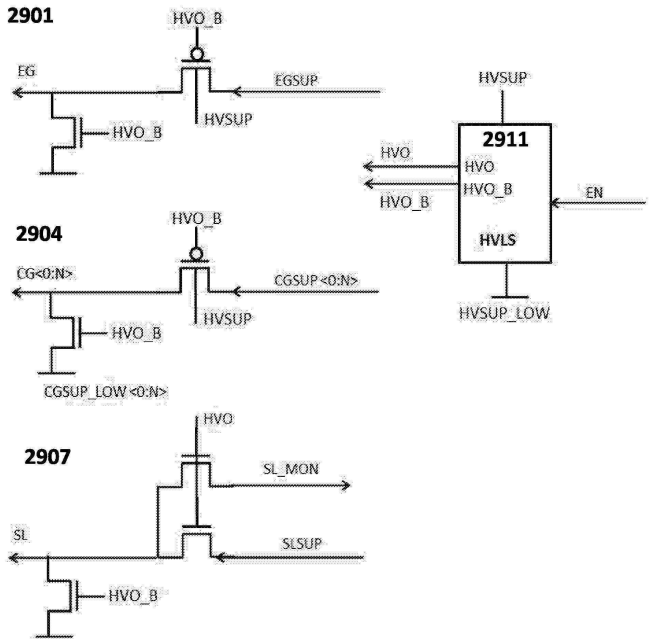
도면27


2700
HVO
2701
SLB
SL_MON
2702
SLE
SLSUP
2703
HVO_B

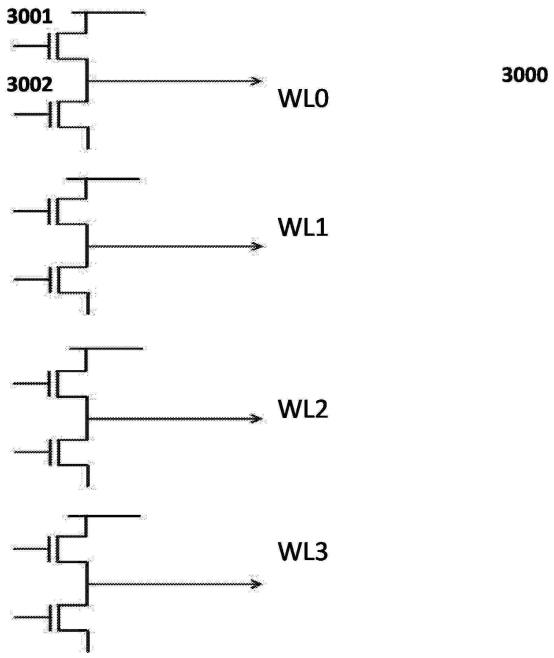
도면28


2801
HVO_B
2802
WL
WLERSUP
2803
HVO_B
HVSUP
HVSUP
2808
HVO
HVO
HVO_B
HVO_B
EN
HVLS
HVSUP_LOW
2804
HVO
2805
SL_MON
2806
SL
SLSUP
2807
HVO_B

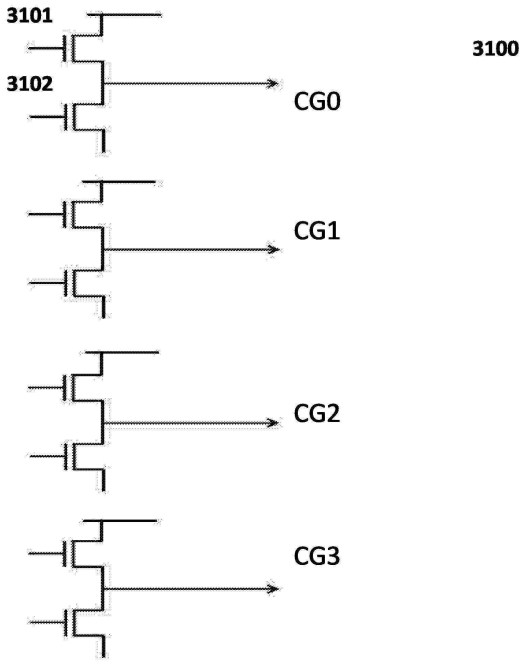
도면29


2901
HVO_B
EG
EGSUP
HVSUP
HVO_B
HVSUP
2911
HVO
HVO
HVO_B
HVO_B
EN
HVLS
HVSUP_LOW
2904
HVO_B
CG<0:N>
CGSUP<0:N>
HVSUP
HVO_B
CGSUP_LOW<0:N>
2907
HVO
SL_MON
SL
SLSUP
HVO_B

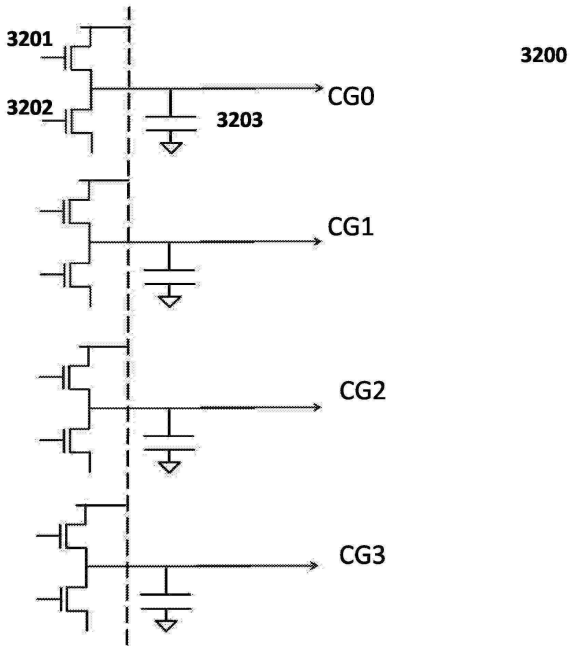
도면30


3001
3002
WL0
3000
WL1
WL2
WL3

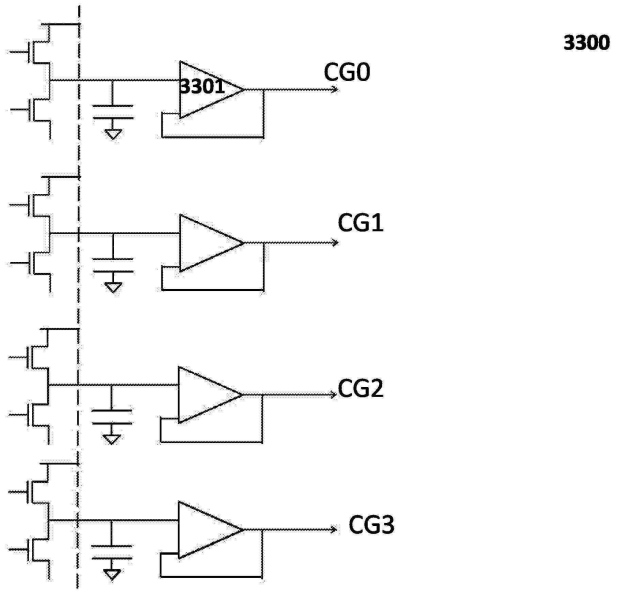
도면31


3101
3102
3100
CG0
CG1
CG2
CG3

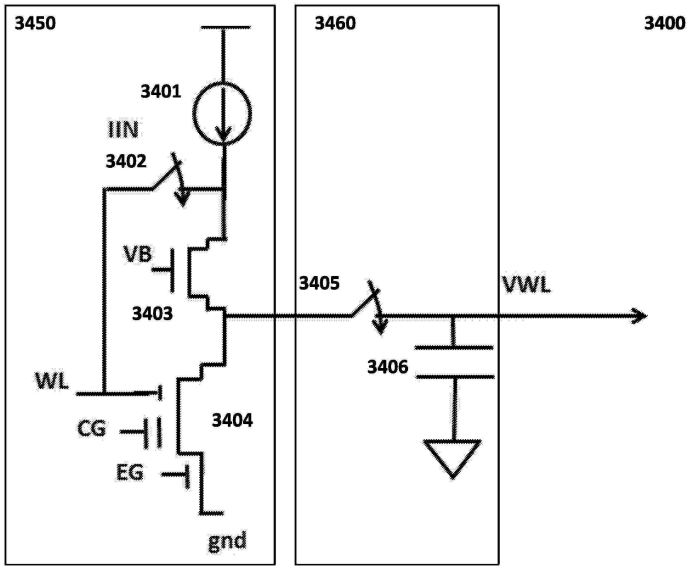
도면32


3201
3202
3203
3200
CG0
CG1
CG2
CG3

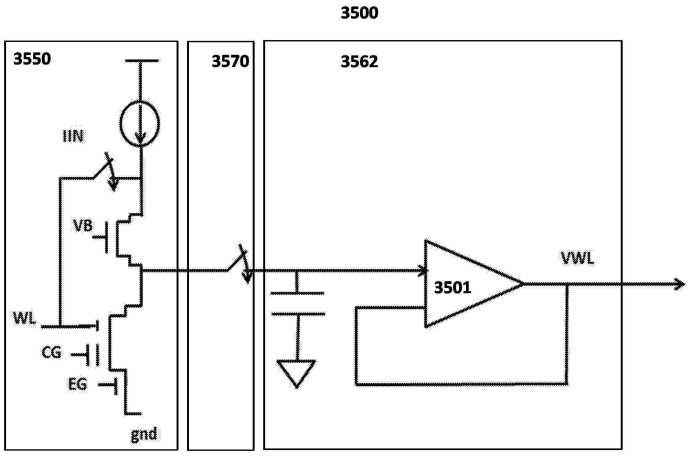
도면33


3300
3301
CG0
CG1
CG2
CG3

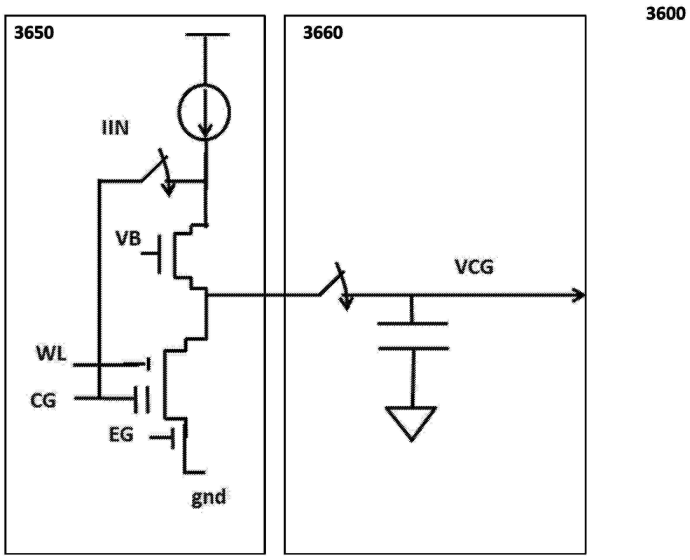
도면34


3450
3460
3400
3401
IIN
3402
VB
3403
3405
VWL
3406
WL
CG
EG
3404
gnd

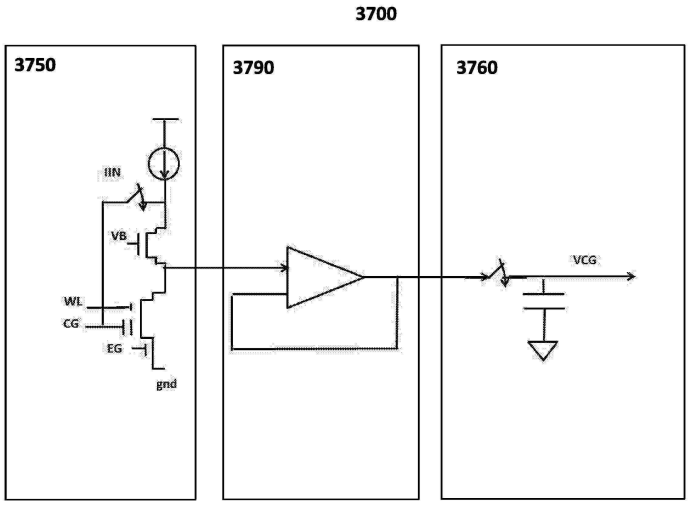
도면35



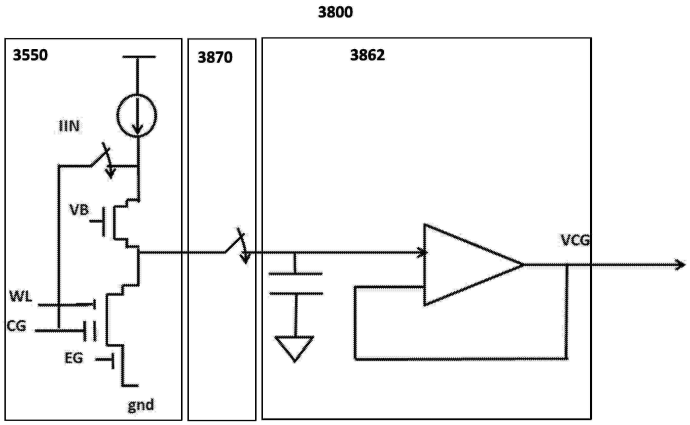
도면36



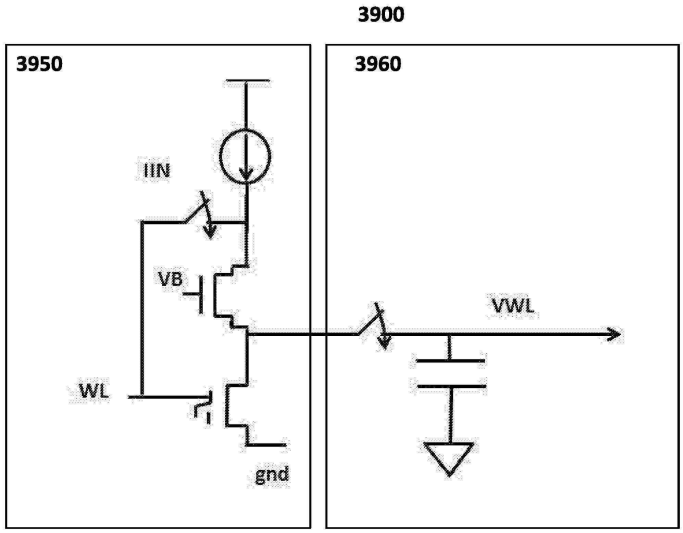
도면37



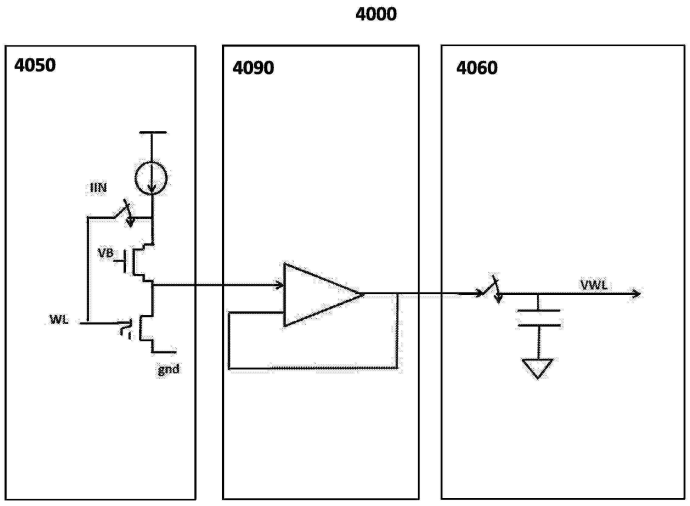
도면38



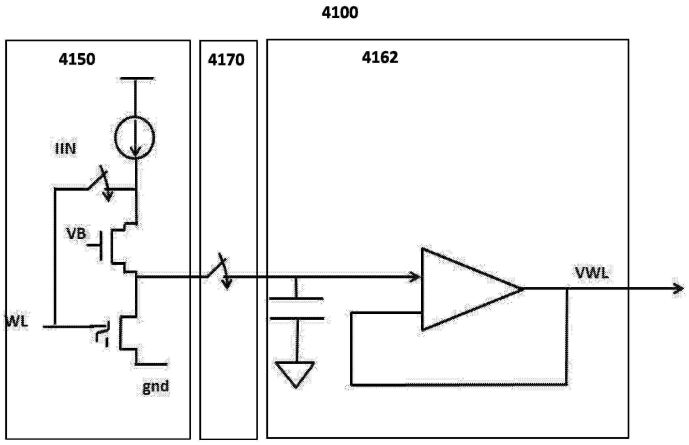
도면39



도면40



도면41



도면42

	WL	WL-비선택	BL	BL-비선택	SL	SL-비선택
판독	1-3.5 V	-0.5 V/0 V	0.6 - 2 V (Ineuron)	0.6 V-2 V/0 V	0 V	0 V
소거	~5-13 V	0 V	0 V	0 V	0 V	0 V
프로그래밍	1-2 V	-0.5 V/0 V	0.1-3 uA	Vinh 대략 2.5 V	4-10 V	0 - 1 V/FLT

도면43

	WL	WL-비선택	BL	BL-비선택	SL	SL-비선택
판독	1-3.5 V	-0.5 V/0 V	0.6-2 V	0.6 V-2 V/0 V	대략 1 V (Ineuron)	0 V
소거	~5-13 V	0 V	0 V	0 V	0 V	SL-금지 (대략 4 - 8 V)
프로그래밍	1-2 V	-0.5 V/0 V	0.1-3 uA	Vinh 대략 2.5 V	4-10 V	0 - 1 V/FLT

도면44

	WL	WL-비선택	BL	BL-비선택	CG	CG-비선택 동일 섹터	CG-비선택	EG	EG-비선택	SL	SL-비선택
판독	1.0-2 V	-0.5V / 0 V	0.6 - 2 V (Ineuron)	0 V	0-2.6 V	0-2.6 V	0-2.6 V	0-2.6 V	0-2.6 V	0 V	0 V
소거	0 V	0 V	0 V	0 V	0 V	0-2.6 V	0-2.6 V	5-12 V	0-2.6 V	0 V	0 V
프로그래밍	0.7-1 V	-0.5 V/ 0 V	0.1-1 uA	Vinh (1 - 2 V)	4-11 V	0-2.6 V	0-2.6 V	4.5-5 V	0-2.6 V	4.5-5 V	0-1 V

도면45

	WL	WL-비선택	BL	BL-비선택	CG	CG-비선택 동일 섹터	CG-비선택	EG	EG-비선택	SL	SL-비선택
판독	1.0-2 V	-0.5 V/0 V	0.6 - 2 V (Ineuron)	0 V	0-2.6 V	0-2.6 V	0-2.6 V	0-2.6 V	0-2.6 V	0 V	0 V
소거	0 V	0 V	0 V	0 V	0 V	CGINH (4 - 9 V)	0-2.6 V	5-12 V	0-2.6 V	0 V	0 V
프로그래밍	0.7-1 V	-0.5 V/ 0 V	0.1-1 uA	Vinh (1 - 2 V)	4-11 V	0-2.6 V	0-2.6 V	4.5-5 V	0-2.6 V	4.5-5 V	0-1 V