

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7253563号
(P7253563)

(45)発行日 令和5年4月6日(2023.4.6)

(24)登録日 令和5年3月29日(2023.3.29)

(51)国際特許分類	F I			
G 1 1 C 16/08 (2006.01)	G 1 1 C	16/08	1 0 0	
G 0 6 G 7/60 (2006.01)	G 0 6 G	7/60		
G 0 6 N 3/063(2023.01)	G 0 6 N	3/063		
G 1 1 C 11/54 (2006.01)	G 1 1 C	11/54		
G 1 1 C 16/04 (2006.01)	G 1 1 C	16/04	1 2 0	
請求項の数 17 (全31頁)				

(21)出願番号	特願2020-548740(P2020-548740)	(73)特許権者	500147506 シリコン ストリージ テクノロジー インコーポレイテッド S I L I C O N S T O R A G E T E C H N O L O G Y , I N C . アメリカ合衆国 9 5 1 3 4 カリフォル ニア州 サンノゼ ホルガー ウェイ 4 5 0 110000626
(86)(22)出願日	平成31年1月24日(2019.1.24)	(74)代理人	弁理士法人英知国際特許商標事務所
(65)公表番号	特表2021-517705(P2021-517705 A)	(72)発明者	トラン、ヒュー バン アメリカ合衆国 9 5 1 3 5 カリフォル ニア州、サンノゼ、ゲイレイ プレイス 2 6 4 2
(43)公表日	令和3年7月26日(2021.7.26)	(72)発明者	ホン、スタンレー アメリカ合衆国 9 5 1 3 1 カリフォル ニア州、サンノゼ、ゲイレイ プレイス 2 6 4 2
(86)国際出願番号	PCT/US2019/015022		
(87)国際公開番号	WO2019/177698		
(87)国際公開日	令和1年9月19日(2019.9.19)		
審査請求日	令和4年1月14日(2022.1.14)		
(31)優先権主張番号	62/642,884		
(32)優先日	平成30年3月14日(2018.3.14)		
(33)優先権主張国・地域又は機関	米国(US)		
(31)優先権主張番号	15/991,890		
(32)優先日	平成30年5月29日(2018.5.29)		
	最終頁に続く		最終頁に続く

(54)【発明の名称】 深層学習人工ニューラルネットワークにおけるアナログニューラルメモリ用のデコーダ

(57)【特許請求の範囲】

【請求項 1】

ベクトルマトリクス乗算アレイに結合されたビット線デコーダ回路であって、前記ベクトルマトリクス乗算アレイは、行及び列に編成された不揮発性メモリセルのアレイを備え、各列はビット線に接続され、前記ビット線デコーダ回路は、

プログラム及び検証動作中に個々のビット線を有効にするための第1の回路と、読み出し動作中に全てのビット線を有効にするための第2の回路と、を備える、ビット線デコーダ回路。

【請求項 2】

前記第2の回路は、各ビット線に結合された、選択トランジスタと活性化関数回路と、を備える、請求項1に記載のビット線デコーダ回路。

【請求項 3】

各選択トランジスタのゲートは、同じ制御線に結合される、請求項2に記載のビット線デコーダ回路。

【請求項 4】

プログラム及び検証動作中、又は読み出し動作中に、負バイアスを各非選択メモリセルのワード線に与える、請求項1に記載のビット線デコーダ回路。

【請求項 5】

前記不揮発性メモリセルの各々は、スプリットゲートフラッシュメモリセルである、請求項1に記載のビット線デコーダ回路。

10

20

【請求項 6】

前記不揮発性メモリセルの各々は、積層ゲートフラッシュメモリセルである、請求項 1 に記載のビット線デコーダ回路。

【請求項 7】

前記不揮発性メモリセルの各々は、サブ閾値領域で動作するように構成される、請求項 1 に記載のビット線デコーダ回路。

【請求項 8】

前記不揮発性メモリセルの各々は、線形領域で動作するように構成される、請求項 1 に記載のビット線デコーダ回路。

【請求項 9】

プログラム及び検証動作中、又は読み出し動作中に、負バイアスを各非選択ビット線デコーダのゲートに与える、請求項 1 に記載のビット線デコーダ回路。

10

【請求項 10】

ベクトルマトリクス乗算アレイに結合されたビット線デコーダ回路であって、前記ベクトルマトリクス乗算アレイは、行及び列に編成された不揮発性メモリセルのアレイを備え、各列はビット線に接続され、前記ビット線デコーダ回路は、

多重化回路であって、第 1 のモードで、前記多重化回路は、プログラム及び検証動作中に個々のビット線を有効にし、第 2 のモードで、前記多重化回路は、読み出し動作中に全てのビット線を有効にする、多重化回路、を備える、ビット線デコーダ回路。

【請求項 11】

前記多重化回路は、各ビット線に結合された、選択トランジスタと活性化関数回路と、を備える、請求項 10 に記載のビット線デコーダ回路。

20

【請求項 12】

プログラム及び検証動作中、又は読み出し動作中に、負バイアスを各非選択メモリセルのワード線に与える、請求項 10 に記載のビット線デコーダ回路。

【請求項 13】

前記不揮発性メモリセルの各々は、スプリットゲートフラッシュメモリセルである、請求項 10 に記載のビット線デコーダ回路。

【請求項 14】

前記不揮発性メモリセルの各々は、積層ゲートフラッシュメモリセルである、請求項 10 に記載のビット線デコーダ回路。

30

【請求項 15】

前記不揮発性メモリセルの各々は、サブ閾値領域で動作するように構成される、請求項 10 に記載のビット線デコーダ回路。

【請求項 16】

前記不揮発性メモリセルの各々は、線形領域で動作するように構成される、請求項 10 に記載のビット線デコーダ回路。

【請求項 17】

プログラム及び検証動作中、又は読み出し動作中に、負バイアスを各非選択ビット線デコーダのゲートに与える、請求項 10 に記載のビット線デコーダ回路。

40

【発明の詳細な説明】

【技術分野】

【0001】

(優先権の主張)

本出願は、2018年3月14日出願の「Decoders for Analog Neuromorphic Memory in Artificial Neural Network」と題された米国仮特許出願第62/642,884号、及び、2018年5月29日出願の「Decoders For Analog Neural Memory In Deep Learning Artificial Neural Network」と題された米国特許出願第15/991,890号に対する優先権を主張する。

50

【 0 0 0 2 】

(発明の分野)

人工ニューラルネットワークにおけるベクトルマトリクス乗算 (V M M) アレイと共に使用するためのデコーダの多数の実施形態が開示される。

【 背景技術 】

【 0 0 0 3 】

人工ニューラルネットワークは、多数の入力に依存することができ、概ね未知である機能を推定する又は近似するために使用される生物学的ニューラルネットワーク (動物の中樞神経系、特に脳) を真似ている。人工ニューラルネットワークは、概して、メッセージを交換する相互接続した「ニューロン」の層を含む。

10

【 0 0 0 4 】

図 1 は、人工ニューラルネットワークを図示しており、ここで円は、入力又はニューロンの層を表す。接続 (シナプスと呼ばれる) は、矢印によって表され、経験に基づいて調整され得る数値の重みを有する。これは、ニューラルネットワークが入力に適応できるようにし、学習できるようにする。典型的には、ニューラルネットワークは、複数の入力の層を含む。典型的には、1つ以上のニューロンの中間層、及びニューラルネットワークの出力を提供するニューロンの出力層が存在する。それぞれのレベルでニューロンは、シナプスから受信されたデータに基づいて個々に又は合わせて決定を行う。

【 0 0 0 5 】

高性能情報処理用の人工ニューラルネットワークの開発における主要な課題の1つは、適切なハードウェア技術の欠如である。実際には、実用ニューラルネットワークは、非常に多数のシナプスに頼り、ニューロン間の高い接続性、すなわち、非常に高度な計算的並行処理を可能にする。原理的には、このような複雑性は、デジタルスーパーコンピュータ又は専門化したグラフィック処理ユニットクラスタによって実現され得る。しかしながら、高コストに加え、これらのアプローチはまた、主として低精度のアナログ計算を実施するのではるかに少ないエネルギーしか消費しない生物学的ネットワークに比べてあまり良くないエネルギー効率に悩まされている。C M O S アナログ回路は、人工ニューラルネットワークに使用されてきたが、ほとんどのC M O S 実装シナプス (CMOS-implemented synapses) は、多数のニューロン及びシナプスを考えると嵩高すぎている。

20

【 0 0 0 6 】

出願人は以前に、参照により組み込まれる米国特許出願第 1 5 / 5 9 4 , 4 3 9 号において、シナプスとして1つ以上の不揮発性メモリアレイを利用する人工 (アナログ) ニューラルネットワークを開示した。不揮発性メモリアレイは、アナログニューロモーフィックメモリとして動作する。ニューラルネットワークデバイスは、第 1 の複数の入力を受信するように、かつそれから第 1 の複数の出力を生成するように構成された第 1 の複数のシナプス、及び第 1 の複数の出力を受信するように構成された第 1 の複数のニューロンを含む。第 1 の複数のシナプスは、複数のメモリセルを含み、メモリセルのそれぞれは、半導体基板内に形成された、間にチャネル領域が延在している離間したソース領域及びドレイン領域と、チャネル領域の第 1 の部分の上方に配設され、チャネル領域の第 1 の部分から絶縁された浮遊ゲートと、チャネル領域の第 2 の部分の上方に配設され、チャネル領域の第 2 の部分から絶縁された非浮遊ゲートと、を含む。複数のメモリセルのそれぞれは、浮遊ゲートの多くの電子に対応する重み値を記憶するように構成される。複数のメモリセルは、第 1 の複数の入力に、記憶された重み値を乗じて第 1 の複数の出力を生成するように構成される。

30

40

【 0 0 0 7 】

アナログニューロモーフィックメモリシステムに使用される各不揮発性メモリセルは、消去され、浮遊ゲート内に非常に特異的かつ正確な量の電荷を保持するようにプログラムされなければならない。例えば、各浮遊ゲートはN個の異なる値のうちの1つを保持しなければならない。ここで、Nは、各セルによって示され得る異なる重みの数である。Nの例としては、16、32、及び64を含む。

50

【 0 0 0 8 】

従来技術のデコード回路（ビット線デコーダ、ワード線デコーダ、制御ゲートデコーダ、ソース線デコーダ、及び消去ゲートデコーダなど）は、アナログニューロモフィックメモリシステムにおけるVMMと共に使用するのに好適ではない。この理由の1つは、VMMシステムでは、プログラム及び検証動作の検証部分（読み出し動作である）が単一の選択メモリセルで動作する一方で、読み出し動作がアレイ内の全てのメモリセルで動作することである。

【 0 0 0 9 】

必要とされるものは、アナログニューロモフィックメモリシステムにおいてVMMと共に使用するのに好適な改良されたデコード回路である。

10

【 発明の概要 】

【 0 0 1 0 】

人工ニューラルネットワークにおけるベクトルマトリクス乗算（VMM）アレイと共に使用するための多数の実施形態が開示される。

【 0 0 1 1 】

【 0 0 1 2 】

【 0 0 1 3 】

【 0 0 1 4 】

【 0 0 1 5 】

【 0 0 1 6 】

20

【 0 0 1 7 】

【 0 0 1 8 】

【 0 0 1 9 】

【 0 0 2 0 】

【 0 0 2 1 】

【 0 0 2 2 】

【 0 0 2 3 】

【 0 0 2 4 】

【 0 0 2 5 】

【 0 0 2 6 】

30

【 0 0 2 7 】

【 0 0 2 8 】

【 0 0 2 9 】

【 0 0 3 0 】

【 0 0 3 1 】

【 0 0 3 2 】

【 0 0 3 3 】

【 0 0 3 4 】

【 0 0 3 5 】

【 0 0 3 6 】

40

【 0 0 3 7 】

【 0 0 3 8 】

【 0 0 3 9 】

【 0 0 4 0 】

【 0 0 4 1 】

【 0 0 4 2 】

【 0 0 4 3 】

【 0 0 4 4 】

【 0 0 4 5 】

【 0 0 4 6 】

50

【 0 0 4 7 】

【 0 0 4 8 】

【 0 0 4 9 】

【 0 0 5 0 】

【 0 0 5 1 】

【 0 0 5 2 】

【 0 0 5 3 】

【 0 0 5 4 】

【 図面の簡単な説明 】

【 0 0 5 5 】

10

【 図 1 】 人工ニューラルネットワークを示す図である。

【 図 2 】 従来の 2 ゲート不揮発性メモリセルの横断面図である。

【 図 3 】 従来の 4 ゲート不揮発性メモリセルの横断面図である。

【 図 4 】 従来の 3 ゲート不揮発性メモリセルの側面横断面図である。

【 図 5 】 別の従来の 2 ゲート不揮発性メモリセルの横断面図である。

【 図 6 】 不揮発性メモリアレイを利用する異なるレベルの例示的人口ニューラルネットワークを示す図である。

【 図 7 】 ベクトル乗算器マトリクスを示すブロック図である。

【 図 8 】 様々なレベルのベクトル乗算器マトリクスを示すブロック図である。

【 図 9 】 ベクトル乗算器マトリクスの一実施形態を示す。

20

【 図 1 0 】 ベクトル乗算器マトリクスの別の実施形態を示す。

【 図 1 1 】 ベクトル乗算器マトリクスの別の実施形態を示す。

【 図 1 2 】 ベクトル乗算器マトリクスの別の実施形態を示す。

【 図 1 3 】 ベクトル乗算器マトリクスの別の実施形態を示す。

【 図 1 4 】 ベクトル乗算器マトリクスのためのビット線デコーダの一実施形態を示す。

【 図 1 5 】 ベクトル乗算器マトリクスのためのビット線デコーダの別の実施形態を示す。

【 図 1 6 】 ベクトル乗算器マトリクスのためのビット線デコーダの別の実施形態を示す。

【 図 1 7 】 ベクトル乗算器マトリクスを動作させるためのシステムを示す。

【 図 1 8 】 ベクトル乗算器マトリクスを動作させるための別のシステムを示す。

【 図 1 9 】 ベクトル乗算器マトリクスを動作させるための別のシステムを示す。

30

【 図 2 0 】 ベクトル乗算器マトリクスと共に使用するためのワード線ドライバの一実施形態を示す。

【 図 2 1 】 ベクトル乗算器マトリクスと共に使用するためのワード線ドライバの別の実施形態を示す。

【 図 2 2 】 ベクトル乗算器マトリクスと共に使用するためのワード線ドライバの別の実施形態を示す。

【 図 2 3 】 ベクトル乗算器マトリクスと共に使用するためのワード線ドライバの別の実施形態を示す。

【 図 2 4 】 ベクトル乗算器マトリクスと共に使用するためのワード線ドライバの別の実施形態を示す。

40

【 図 2 5 】 ベクトル乗算器マトリクスと共に使用するためのワード線ドライバの別の実施形態を示す。

【 図 2 6 】 ベクトル乗算器マトリクスと共に使用するためのワード線ドライバの別の実施形態を示す。

【 図 2 7 】 ベクトル乗算器マトリクスと共に使用するためのソース線デコーダ回路を示す。

【 図 2 8 】 ベクトル乗算器マトリクスと共に使用するためのワード線デコーダ回路、ソース線デコーダ回路、及び高電圧レベルシフタを示す。

【 図 2 9 】 ベクトル乗算器マトリクスと共に使用するための消去ゲートデコーダ回路、制御ゲートデコーダ回路、ソース線デコーダ回路、及び高電圧レベルシフタを示す。

【 図 3 0 】 ベクトル乗算器マトリクスと共に使用するためのワード線デコーダ回路を示す。

50

【図 3 1】ベクトル乗算器マトリクスと共に使用するための制御ゲートデコーダ回路を示す。

【図 3 2】ベクトル乗算器マトリクスと共に使用するための別の制御ゲートデコーダ回路を示す。

【図 3 3】ベクトル乗算器マトリクスと共に使用するための別の制御ゲートデコーダ回路を示す。

【図 3 4】ベクトル乗算器マトリクスにおけるワード線を制御するための電流電圧回路を示す。

【図 3 5】ベクトル乗算器マトリクスにおけるワード線を制御するための別の電流電圧回路を示す。

10

【図 3 6】ベクトル乗算器マトリクスにおける制御ゲート線を制御するための電流電圧回路を示す。

【図 3 7】ベクトル乗算器マトリクスにおける制御ゲートラインを制御するための別の電流電圧回路を示す。

【図 3 8】ベクトル乗算器マトリクスにおける制御ゲートラインを制御するための別の電流電圧回路を示す。

【図 3 9】ベクトル乗算器マトリクスにおけるワード線を制御するための別の電流電圧回路を示す。

【図 4 0】ベクトル乗算器マトリクスにおけるワード線を制御するための別の電流電圧回路を示す。

20

【図 4 1】ベクトル乗算器マトリクスにおけるワード線を制御するための別の電流電圧回路を示す。

【図 4 2】図 9 のベクトル乗算器マトリクスの動作電圧を示す。

【図 4 3】図 1 0 のベクトル乗算器マトリクスの動作電圧を示す。

【図 4 4】図 1 1 のベクトル乗算器マトリクスの動作電圧を示す。

【図 4 5】図 1 2 のベクトル乗算器マトリクスの動作電圧を示す。

【発明を実施するための形態】

【0056】

本発明の人工ニューラルネットワークは、CMOS 技術及び不揮発性メモリアレイの組み合わせを利用する。

30

不揮発性メモリセル

【0057】

デジタル不揮発性メモリは、周知である。例えば、米国特許第 5,029,130 号（「130 号特許」）は、スプリットゲート不揮発性メモリセルのアレイを開示し、あらゆる目的のために参照により本明細書に組み込まれる。このようなメモリセルを図 2 に示す。各メモリセル 210 は、半導体基板 12 内に形成され、チャンネル領域 18 をその間に有するソース領域 14 とドレイン領域 16 と、を含む。浮遊ゲート 20 は、チャンネル領域 18 の第 1 の部分の上方に形成され、チャンネル領域 18 の第 1 の部分から絶縁されて（かつチャンネル領域 18 の第 1 の部分導電性を制御して）おり、またソース領域 16 の一部分の上方に形成される。ワード線端子 22（典型的には、ワード線に結合される）は、チャンネル領域 18 の第 2 の部分の上方に配設され、チャンネル領域 18 の第 2 の部分から絶縁され、（かつチャンネル領域 18 の第 2 の部分導電性を制御する）第 1 の部分と、浮遊ゲート 20 の上方に、かつ上方で延在する第 2 の部分と、を有する。浮遊ゲート 20 及びワード線端子 22 は、ゲート酸化物によって基板 12 から絶縁される。ビット線 24 はドレイン領域 16 に結合される。

40

【0058】

メモリセル 210 は、ワード線端子 22 に高圧正電圧を印加することによって消去され（ここで電子は、浮遊ゲートから除去される）、それにより、ファウラーノルドハイムトンネリングによって浮遊ゲート 20 からワード線端子 22 まで中間絶縁体を通して浮遊ゲート 20 の電子をトンネリングさせる。

50

【 0 0 5 9 】

メモリセル 2 1 0 は、ワード線端子 2 2 に正電圧、及びソース 1 6 に正電圧を印加することによってプログラムされる（ここで電子は、浮遊ゲートに印加される）。電子電流は、ソース 1 6 からドレイン 1 4 に向かって流れることになる。電子は、ワード線端子 2 2 と浮遊ゲート 2 0 との間隙に達すると、加速し発熱する。熱せられた電子のいくらかは、浮遊ゲート 2 0 からの静電引力に起因してゲート酸化物 2 6 を通って浮遊ゲート 2 0 に注入される。

【 0 0 6 0 】

ドレイン 1 4 及びワード線端子 2 2 に正の読み出し電圧を印加する（ワード線端子下のチャンネル領域をターンオンする）ことによって、メモリセル 2 1 0 が読み出される。浮遊ゲート 2 0 が正に帯電する（すなわち、電子を消去し、ドレイン 1 6 に正に結合する）場合、浮遊ゲート 2 0 下のチャンネル領域の部分は、次に同様にオンになり、電流は、チャンネル領域 1 8 を流れ、これは、消去された状態つまり「 1 」の状態として感知される。浮遊ゲート 2 0 が負に帯電する（すなわち、電子でプログラムされる）場合、次に浮遊ゲート 2 0 下のチャンネル領域の部分は、ほとんど又は完全にオフになり、電流は、チャンネル領域 1 8 を流れず（又はわずかに流れる）、プログラムされた状態又は「 0 」の状態として検出される。

【 0 0 6 1 】

表 1 は、読み出し、消去及びプログラム動作を実行するためのメモリセル 2 1 0 の端子に印加され得る典型的な電圧範囲を示す。

表 1：図 2 のフラッシュメモリセル 2 1 0 の動作

【表 1】

	WL	BL	SL
読み出し	2~3V	0.6~2V	0V
消去	約11~13V	0V	0V
プログラム	1~2V	1~3 μ A	9~10V

【 0 0 6 2 】

他のスプリットゲートメモリセル構成が知られている。例えば、図 3 は、ソース領域 1 4 と、ドレイン領域 1 6 と、チャンネル領域 1 8 の第 1 の部分の上方にある浮遊ゲート 2 0 と、チャンネル領域 1 8 の第 2 の部分の上方にある選択ゲート 2 8（典型的には、ワード線に結合される）と、浮遊ゲート 2 0 の上方にある制御ゲート 2 2 と、及びソース領域 1 4 の上方にある消去ゲート 3 0 と、を備える 4 ゲートメモリセル 3 1 0 を示す。この構成は、米国特許第 6, 7 4 7, 3 1 0 号に記載され、この米国特許は、あらゆる目的のため参照により本明細書に組み込まれる。ここで、全てのゲートは、浮遊ゲート 2 0 を除いて、非浮遊ゲートであり、それらは電圧源に電氣的に接続される又は接続可能であることを意味する。プログラミングは、浮遊ゲート 2 0 にそれ自体を注入するチャンネル領域 1 8 からの熱せられた電子によって示される。消去は、浮遊ゲート 2 0 から消去ゲート 3 0 へトンネリングする電子によって示される。

【 0 0 6 3 】

表 2 は、読み出し、消去及びプログラム動作を実行するためのメモリセル 3 1 0 の端子に印加され得る典型的な電圧範囲を示す。

表 2：図 3 のフラッシュメモリセル 3 1 0 の動作

10

20

30

40

50

【表 2】

	WL/SG	BL	CG	EG	SL
読み出し	1.0~2V	0.6~2V	0~2.6V	0~2.6V	0V
消去	-0.5V/0V	0V	0V/-8V	8~12V	0V
プログラム	1V	1 μ A	8~11V	4.5~9V	4.5~5V

10

【0064】

図4は、スプリットゲート3ゲートメモリセル410を示す。メモリセル410は、メモリセル410が別個の制御ゲートを有しないことを除いて、図3のメモリセル310と同一である。消去動作（消去ゲートを通しての消去）及び読み出し動作は、制御ゲートバイアスがないことを除いて、図3のものと同様である。プログラミング動作もまた、制御ゲートバイアスなしで行われるため、ソース線のプログラム電圧は、制御ゲートバイアスの不足を補償するためにより高い。

【0065】

表3は、読み出し、消去及びプログラム動作を実行するためのメモリセル410の端子に印加され得る典型的な電圧範囲を示す。

20

表3：図4のフラッシュメモリセル410の動作

【表3】

	WL/SG	BL	EG	SL
読み出し	0.7~2.2V	0.6~2V	0~2.6V	0V
消去	-0.5V/0V	0V	11.5V	0V
プログラム	1V	2~3 μ A	4.5V	7~9V

30

【0066】

図5は、積層ゲートメモリセル510を示す。メモリセル510は、浮遊ゲート20がチャネル領域18の上方に延在し、制御ゲート22が絶縁層によって分離された浮遊ゲート20の上方に延在することを除いて、図2のメモリセル210と同様である。消去、プログラミング、及び読み出し動作は、メモリセル210について前述したものと同様の方法で動作する。

40

【0067】

表4は、読み出し、消去及びプログラム動作を実行するためのメモリセル510の端子に印加され得る典型的な電圧範囲を示す。

表4：図5のフラッシュメモリセル510の動作

50

【表 4】

	CG	BL	SL	Pサブ
読み出し	2~5V	0.6~2V	0V	0V
消去	-8~-10V/0V	FLT	FLT	8~10V/15~20V
プログラム	8~12V	3~5V	0V	0V

10

【0068】

上記の人工ニューラルネットワークにおける不揮発性メモリセルのタイプのうちの1つを含むメモリアレイを利用するために、2つの修正が行われる。第1に、以下に更に説明されるように、アレイ内の他のメモリセルのメモリ状態に悪影響を与えずに各メモリセルを個々にプログラム、消去、及び読み出しすることができるよう線を構成する。第2に、メモリセルの連続（アナログ）プログラミングを提供する。

【0069】

具体的には、アレイ内の各メモリセルのメモリ状態（すなわち、浮遊ゲートの電荷）を、完全に消去された状態から完全にプログラムされた状態へ、独立して、かつ他のメモリセルの異常が最小で連続的に変更することができる。別の実施形態において、アレイ内の各メモリセルのメモリ状態（すなわち、浮遊ゲートの電荷）を、完全にプログラムされた状態から完全に消去された状態へ、及び逆もまた同様に、独立して、かつ他のメモリセルの異常が最小で連続的に変更することができる。このことは、セル記憶がアナログであるか、又は多数の別個の値（16又は64の異なる値など）のうちの1つを最低限記憶することができ、それは、メモリアレイ内の全てのセルの非常に正確、かつ個々の調整を可能にし、またメモリアレイを記憶に対して理想的にし、ニューラルネットワークのシナプシスの重みに微調整を加えることを意味する。

20

不揮発性メモリセルアレイを使用するニューラルネットワーク

【0070】

図6は、不揮発性メモリアレイを利用するニューラルネットワークの非限定実施例を概念的に示す。この例は、顔認識アプリケーションのために不揮発性メモリアレイニューラルネットを使用するが、不揮発性メモリアレイベースのニューラルネットワークを使用して他の任意の適切なアプリケーションを実施することができる。

30

【0071】

S0は入力であり、この例では、5ビット精度の32×32ピクセルRGB画像である（すなわち、各色R、G、及びBにつき1つずつの3つの32×32ピクセルアレイ、各ピクセルが5ビット精度である）。S0からC1に行くシナプスCB1は、異なる重みのセット及び共有される重みの両方を有し、入力画像を3×3ピクセルの重なり合うフィルタでスキャンし（カーネル）、1ピクセル（又はモデルによって決まるように2ピクセル以上）ずつフィルタをシフトする。具体的には、画像の3×3部分における9ピクセルの値（すなわち、フィルタ又はカーネルと呼ばれる）は、シナプスCB1に提供され、それによってこれらの9個の入力値に適切な重みを乗じ、その乗算の出力の合計後、単一の出力値が決定され、特徴マップC1の層の1つのピクセルを生成するためにC1の第1のニューロンによって与えられる。3×3フィルタは次に右側に1ピクセルだけシフトされ（すなわち、3ピクセルの列を右側に追加し、左側で3ピクセルの列をドロップする）、そのため、この新しく位置づけられたフィルタの9ピクセル値が、シナプスCB1に提供されるため、それらに同じ重みを乗じ、関連するニューロンによって第2の単一の出力値を決定する。このプロセスを、3×3フィルタが32×32ピクセル画像全体にわたって3色全て及び全てのビットについてスキャンするまで続ける（精度値）。プロセスは次に

40

50

、層 C 1 の特徴マップ全てが計算されるまで、異なる重みのセットを使用して繰り返されて、C 1 の異なる特徴マップを生成する。

【 0 0 7 2 】

C 1 において、本例では、それぞれ 30×30 ピクセルを有する 16 個の特徴マップが存在する。各ピクセルは、入力とカーネルとの乗算から抽出された新しい特徴ピクセルであり、したがって、各特徴マップは、2次元アレイであり、したがってこの例では、シナプス C B 1 は、2次元アレイの 16 層を構成する（本明細書で言及されるニューロン層及びアレイは、必ずしも物理的関係ではなく論理的な関係であり、すなわち、アレイは必ずしも物理的な 2次元アレイに配向されないことに留意する）。16 個の特徴マップのそれぞれは、フィルタスキャンに適用される 16 個の異なるシナプス重みのセットの 1 つによって生成される。C 1 特徴マップは全て、境界同定など同じ画像特徴の異なる態様を対象とすることができる。例えば、第 1 のマップ（この第 1 のマップを生成するのに使用される全てのスキャンに共有される第 1 の重みセットを使用して生成される）は、円形エッジを識別することができ、第 2 のマップ（第 1 の重みセットと異なる第 2 の重みセットを使用して生成される）は、方形エッジ又は特定の特徴のアスペクト比などを識別することができる。

10

【 0 0 7 3 】

活性化関数 P 1（プーリング）は、C 1 から S 1 へ行く前に適用され、各特徴マップ中の連続する、重なり合わない 2×2 領域からの値をプーリングする。プーリング段階の目的は、平均して近隣の位置にすること（又は \max 関数もまた使用され得る）、例えばエッジ位置の依存を低減すること、及び次の段階に行く前にデータサイズを低減することである。S 1 において、16 個の 15×15 特徴マップ（すなわち、16 個のそれぞれ 15×15 ピクセルの異なるアレイ）が存在する。S 1 から C 2 へ行く C B 2 内のシナプス及び関連するニューロンは、S 1 内のマップを 1 ピクセルのフィルタシフトを使用して 4×4 フィルタでスキャンする。C 2 において、22 個の 12×12 特徴マップが存在する。活性化関数 P 2（プーリング）は、C 2 から S 2 へ行く前に適用され、各特徴マップ中の連続する、重なり合わない 2×2 領域からの値をプーリングする。S 2 において、22 個の 6×6 特徴マップが存在する。活性化関数は、S 2 から C 3 へ行くシナプス C B 3 で適用され、ここで C 3 内の全てのニューロンは S 2 内の全てのマップに接続する。C 3 において、64 個のニューロンが存在する。C 3 から出力 S 3 へ行くシナプス C B 4 は、S 3 を C 3 と完全に接続する。S 3 における出力は、10 個のニューロンを含み、ここで最も高い出力ニューロンは、クラスを決定する。この出力は、例えば、元の画像の内容の同定又は分類を示すことができる。

20

30

【 0 0 7 4 】

シナプスの各レベルは、不揮発性メモリセルのアレイ又はアレイの一部を使用して実行される。図 7 は、不揮発性メモリセルを含み、入力層と次の層との間のシナプスとして利用されるベクトルマトリクス乗算（VMM）アレイのブロック図である。具体的には、VMM 3 2 は、不揮発性メモリセルのアレイ 3 3、消去ゲート及びワード線ゲートデコーダ 3 4、制御ゲートデコーダ 3 5、ビット線デコーダ 3 6、並びにソース線デコーダ 3 7 を含み、それらのデコーダはメモリアレイ 3 3 に対する入力をデコードする。この例におけるソース線デコーダ 3 7 はまた、メモリアレイの出力をデコードする。あるいは、ビット線デコーダ 3 6 は、メモリアレイの出力をデコードすることができる。メモリアレイは、2つの目的を果たす。第 1 に、VMM により使用される重みを記憶する。第 2 に、メモリアレイは、効果的に、入力に、メモリアレイに記憶された重みを乗じ、それを出力線（ソース線またはビット線）ごとに加算して出力を生成し、これは、次の層への入力又は最後の層への入力になる。乗算及び加算関数を実行することによって、メモリアレイは、別個の乗算及び加算論理回路の必要性をなくし、また、その場でのメモリ計算のために電力効率も良い。

40

【 0 0 7 5 】

メモリアレイの出力は差動加算器（加算オペアンプなど）3 8 に供給され、差動加算器

50

はメモリセルアレイの出力を合計してその畳み込み用の単一の値を生成する。差動加算器は、正の入力で正の重みと負の重みとの合計を実現するようなものである。合計された出力値は、次に出力を整流する活性化関数回路 39 に供給される。活性化関数には、シグモイド、 \tanh 、又はReLU関数が含まれ得る。整流された出力値は、次の層（例えば上の説明ではC1）として特徴マップの要素になり、次いで、次のシナプスに適用されて次の特徴マップ層又は最後の層を生成する。したがって、この例では、メモリアレイは、複数のシナプスを構成し（ニューロンの前の層から、又は画像データベースなどの入力層から入力を受信する）、加算オペアンプ 38 及び活性化関数回路 39 は、複数のニューロンを構成する。

【0076】

図8は、様々なレベルのVMMのブロック図である。図14に示されるように、入力は、デジタルアナログ変換器31によってデジタルからアナログに変換され、入力VMM32aに提供される。入力VMM32aによって生成された出力は、次のVMM（隠しレベル1）32bへの入力として提供され、VMM（隠しレベル1）32bは、同様にして、次のVMM（隠しレベル2）32bへの入力として提供される出力を生成する、出力を生成する、などとなる。VMMの32の様々な層は、畳み込みニューラルネットワーク（convolutional neural network、CNN）のシナプス及びニューロンの異なる層として機能する。各VMMは、スタンドアローンの不揮発性メモリアレイであり得、又は複数のVMMは、同じ不揮発性メモリアレイの異なる部分を利用することができ、又は複数のVMMは、同じ不揮発性メモリアレイの重なり合う部分を利用することができる。図8に示される例は、5つの層（32a、32b、32c、32d、32e）：1つの入力層（32a）、2つの隠れ層（32b、32c）、及び2つの完全に接続された層（32d、32e）を含む。当業者であれば、これは単なる例示であり、代わりにシステムが2つを超える隠れ層及び2つを超える完全に接続された層を備え得ることを理解するであろう。

ベクトルマトリクス乗算（VMM）アレイ

【0077】

図9は、図2に示されるタイプのメモリセルに特に適したニューロンVMM900を示し、入力層と次の層との間のシナプス及びニューロンの一部として利用される。VMM900は、不揮発性メモリセルのメモリアレイ903、基準アレイ901、及び基準アレイ902を備える。基準アレイ901及び902は、端子BLR0~3内に流れ込む電流入力を電圧入力WL0~3に変換する役割を果たす。図示の基準アレイ901及び902は、列方向にある。一般に、基準アレイ方向は入力線に直交する。実際には、基準メモリセルは、マルチプレクサ（1つのマルチプレクサ及び基準ビット線をバイアスするための1つのカスコードトランジスタVBLRを含むマルチプレクサ914）を介して接続されたダイオードであり、電流入力はそれらに流れ込む。基準セルは、標的基準レベルに調整される。

【0078】

メモリアレイ903は、2つの目的を果たす。第1に、VMM900により使用される重みを記憶する。第2に、メモリアレイ903は、入力（端子BLR0~3に提供される電流入力であって、基準アレイ901及び902がこれらの電流入力を入力電圧に変換してワード線WL0~3に供給する）にメモリアレイに記憶された重みを有効に乗算して、出力を生成し、これは次の層への入力又は最終の層への入力となる。乗算関数を実行することによって、メモリアレイは、別個の乗算論理回路の必要性をなくし、また電力効率も良い。ここで、電圧入力はワード線に提供され、出力は読み出し（推論）動作中にビット線に現れる。ビット線の電流は、ビット線に接続されたメモリセルからの全ての電流の合計関数を実行する。

【0079】

図42は、VMM900の動作電圧を示す。表中の列は、選択セルのワード線、非選択セルのワード線、選択セルのビット線、非選択セルのビット線、選択セルのソース線、及び非選択セルのソース線に印加された電圧を示す。行は、読み出し、消去、及びプログラ

10

20

30

40

50

ムの動作を示す。

【 0 0 8 0 】

図 1 0 は、図 2 に示されるタイプのメモリセルに特に適したニューロン V M M 1 0 0 0 を示し、入力層と次の層との間のシナプス及びニューロンの一部として利用される。V M M 1 0 0 0 は、不揮発性メモリセルのメモリアレイ 1 0 0 3、基準アレイ 1 0 0 1、及び基準アレイ 1 0 0 2 を備える。V M M 1 0 0 0 は、V M M 1 0 0 0 においてワード線が垂直方向に延びることを除いて、V M M 9 0 0 と同様である。2つの基準アレイ 1 0 0 1 (頂部で、偶数行の電圧への基準変換入力電流を提供する) 及び 1 0 0 2 (底部で、奇数行の電圧への基準変換入力電流を提供する) が存在する。ここで、入力はワード線に提供され、出力は読み出し動作中にソース線に現れる。ソース線の電流は、ソース線に接続されたメモリセルからの全ての電流の合計関数を実行する。

10

【 0 0 8 1 】

図 4 3 は、V M M 1 0 0 0 の動作電圧を示す。表中の列は、選択セルのワード線、非選択セルのワード線、選択セルのビット線、非選択セルのビット線、選択セルのソース線、及び非選択セルのソース線に印加された電圧を示す。行は、読み出し、消去、及びプログラムの動作を示す。

【 0 0 8 2 】

図 1 1 は、図 3 に示されるタイプのメモリセルに特に適したニューロン V M M 1 1 0 0 を示し、入力層と次の層との間のシナプス及びニューロンの一部として利用される。V M M 1 1 0 0 は、不揮発性メモリセルのメモリアレイ 1 1 0 1、基準アレイ 1 1 0 2 (偶数行の入力電圧への基準変換入力電流を提供する)、及び基準アレイ 1 1 0 3 (奇数行の入力電圧への基準変換入力電流を提供する) を備える。V M M 1 1 0 0 は、V M M 1 1 0 0 を除き V M M 9 0 0 と同様である。V M M 1 1 0 0 は、メモリセルの行の制御ゲートに結合された制御線 1 1 0 6、及びメモリセルの隣接する行の消去ゲートに結合された制御線 1 1 0 7 を更に備える。ここで、ワード線、制御ゲート線、消去ゲート線は同じ方向である。V M M は、選択された基準メモリセル及びスイッチ 1 1 0 5 (m u x 1 1 1 4 の一部) のビット線コンタクトに選択的に結合する基準ビット線選択トランジスタ 1 1 0 4 (m u x 1 1 1 4 の一部) を更に備え、このビット線選択トランジスタ 1 1 0 4 は、特定の選択された基準メモリセルのために、基準ビット線を制御線 1 1 0 6 に選択的に結合する。ここで、入力は (メモリアレイ 1 1 0 1 の) ワード線に提供され、出力は読み出し動作中にビット線 1 1 0 9 などのビット線に現れる。ビット線の電流は、ビット線に接続されたメモリセルからの全ての電流の合計関数を実行する。

20

30

【 0 0 8 3 】

図 4 4 は、V M M 1 1 0 0 の動作電圧を示す。表中の列は、選択セルのワード線、非選択セルのワード線、選択セルのビット線、非選択セルのビット線、選択セルの制御ゲート、選択セルと同じセクタ内の非選択セルの制御ゲート、選択セルとは異なるセクタ内の非選択セルの制御ゲート、選択セルの消去ゲート、非選択セルの消去ゲート、選択セルのソース線、及び非選択セルのソース線の電圧を示す。行は、読み出し、消去、及びプログラムの動作を示す。

【 0 0 8 4 】

図 1 2 は、図 3 に示されるタイプのメモリセルに特に適したニューロン V M M 1 2 0 0 を示し、入力層と次の層との間のシナプス及びニューロンの一部として利用される。V M M 1 2 0 0 は、V M M 1 2 0 0 で、消去ゲート線 1 2 0 1 などの消去ゲート線が垂直方向に延びることを除いて、V M M 1 1 0 0 と同様である。ここで、入力はワード線に提供され、出力はソース線に現れる。ビット線の電流は、ビット線に接続されたメモリセルからの全ての電流の合計関数を実行する。

40

【 0 0 8 5 】

図 4 5 は、V M M 1 2 0 0 の動作電圧を示す。表中の列は、選択セルのワード線、非選択セルのワード線、選択セルのビット線、非選択セルのビット線、選択セルの制御ゲート、選択セルと同じセクタ内の非選択セルの制御ゲート、選択セルとは異なるセクタ内の非

50

選択セルの制御ゲート、選択セルの消去ゲート、非選択セルの消去ゲート、選択セルのソース線、及び非選択セルのソース線の電圧を示す。行は、読み出し、消去、及びプログラムの動作を示す。

【0086】

図13は、図3に示されるタイプのメモリセルに特に適したニューロンVMM1300を示し、入力層と次の層との間のシナプス及びニューロンの一部として利用される。VMM1300は、不揮発性メモリセルのメモリアレイ1301、及び基準アレイ1302（アレイの頂部で）を備える。あるいは、別の基準アレイが、図10のものと同様に、底部に位置することができる。他の点では、VMM1300は、VMM1300で、制御ゲート線1303などの制御ゲート線が垂直方向に伸び（したがって、行方向の基準アレイ1302が入力制御ゲート線に直交する）、消去ゲート線1304などの消去ゲート線が水平方向に伸びることを除いて、VMM1200と同様である。ここで、入力は制御ゲート線に提供され、出力はソース線に現れる。一実施形態では、偶数行のみが使用され、別の実施形態では、奇数行のみが使用される。ソース線の電流は、ソース線に接続されたメモリセルからの全ての電流の合計関数を実行する。

10

【0087】

ニューラルネットワークについて本明細書に記載されるように、フラッシュセルは、サブ閾値領域で動作するように構成されることが好ましい。

【0088】

本明細書に記載されたメモリセルは、弱反転にバイアスされる。

20

$$I_{ds} = I_o * e^{(V_g - V_{th})/kV_t} = w * I_o * e^{(V_g)/kV_t}$$

$$w = e^{(-V_{th})/kV_t}$$

【0089】

入力電流を入力電圧に変換するためにメモリセルを使用するI-Vログ変換器について：

$$V_g = k * V_t * \log [I_{ds} / w p * I_o]$$

【0090】

ベクトルマトリクス乗算器VMMとして使用されるメモリアレイについて、出力電流は以下である：

$$I_{out} = w a * I_o * e^{(V_g)/kV_t}、すなわち$$

$$I_{out} = (w a / w p) * I_{in} = W * I_{in}$$

$$W = e^{(V_{thp} - V_{tha})/kV_t}$$

30

【0091】

ワード線又は制御ゲートは、入力電圧に対するメモリセルの入力として使用することができる。

【0092】

あるいは、フラッシュメモリセルは、線形領域で動作するように構成することができる。

$$I_{ds} = \frac{1}{2} * (V_{gs} - V_{th}) * V_{ds} ; \quad C_{ox} = u * C_{ox} * W / L$$

$$W = \frac{2 * I_{ds}}{(V_{gs} - V_{th}) * V_{ds}}$$

【0093】

I-V線形変換器では、線形領域で動作するメモリセルを使用して、入力/出力電流を入力/出力電圧に線形に変換することができる。

40

【0094】

ESFベクトルマトリクス乗算器の他の実施形態は、参照により本明細書に組み込まれる米国特許出願第15/826,345号に記載されているとおりである。ソース線又はビット線は、ニューロン出力（電流和出力）として使用することができる。

【0095】

図14は、ビット線デコーダ回路1400の一実施形態を示す。ビット線デコーダ回路1400は、列デコーダ1402及びアナログニューロモフィックニューロン（「ANN」）列デコーダ1403を備え、これらの各々はVMMアレイ1401に結合される。VMMアレイは、先に論じたVMM設計（VMM900、1000、1100、1200

50

、及び1300など)又は他のVMM設計のうちのいずれかに基づくことができる。

【0096】

アナログニューロモーフックシステムによる1つの課題は、システムが(読み出し動作を伴う)プログラム及び検証を行うことができなければならない、アレイ内の全てのセルが選択され読み出されるANN読み出しを実行することができなければならないことである。換言すれば、ビット線デコーダは、場合によっては1つのビット線のみを選択しなければならない、他の例では全てのビット線を選択しなければならない。

【0097】

ビット線デコーダ回路1400は、この目的を達成する。列デコーダ1402は、従来の列デコーダ(プログラム及び消去、又はPE、デコード経路)であり、プログラム及びプログラム検証(感知動作)などの個々のビット線を選択するために使用することができる。列デコーダ1402の出力は、プログラム、PE検証、及び消去を制御するためのプログラム/消去(PE)列ドライバ回路に結合される(図14には示されていない)。ANN列デコーダ1403は、一度に全てのビット線での読み出し動作を可能にするように具体的に設計された列デコーダである。ANN列デコーダ1403は、ビット線(ここではBL0)に結合された例示的な選択トランジスタ1405及び出力回路(例えば、tanh、シグモイド、ReLUなどの電流加算器及び活性化関数)1406を備える。同じデバイスのセットが、他のビット線の各々に取り付けられる。選択トランジスタ1405などの選択トランジスタの全ては、選択線1404に結合される。ANN読み出し動作中、選択線1404が有効にされ、選択トランジスタ1405などの選択トランジスタの各々がオンになり、次いで、各ビット線からの電流が回路1406及び出力などの出力回路によって受信される。

【0098】

図15は、ビット線デコーダ回路1500の一実施形態を示す。ビット線デコーダ回路1500は、VMMアレイ1501に結合される。VMMアレイは、先に論じたVMM設計(VMM900、1000、1100、1200、及び1300など)又は他のVMM設計のうちのいずれかに基づくことができる。

【0099】

選択トランジスタ1502及び1503は、一対の相補的制御信号(V0及びVB_0)によって制御され、ビット線(BL0)に結合される。選択トランジスタ1504及び1505は、別の対の相補的制御信号(V1及びVB_1)によって制御され、別のビット線(BL1)に結合される。選択トランジスタ1502及び1504は、プログラミングを有効にするために同じ出力に結合され、選択トランジスタ1503及び1505は、禁止プログラミングのためなどの同じ出力に結合される。トランジスタ1502/1503/1504/1505の出力線(プログラム及び消去PEデコード経路)は、プログラム、PE検証、及び消去(図示せず)を制御するためのPE列ドライバ回路に結合されるようなものである。

【0100】

選択トランジスタ1506は、ビット線(BL0)に結合され、出力及び活性化関数回路1507(例えば、tanh、シグモイド、ReLUなどの電流加算及び活性化関数)に結合される。選択トランジスタ1506は制御線1508によって制御される。

【0101】

BL0のみが活性化されると、制御線1508はデアサートされ、信号V0はアサートされ、したがってBL0のみを読み取る。ANN読み出し動作中に、制御線1508がアサートされ、選択トランジスタ1506及び同様のトランジスタがオンにされ、全てのビット線は、全てのニューロン処理のために読み出される。

【0102】

図16は、ビット線デコーダ回路1600の一実施形態を示す。ビット線デコーダ回路1600は、VMMアレイ1601に結合される。VMMアレイは、先に論じたVMM設計(VMM900、1000、1100、1200、及び1300など)又は他のVMM

10

20

30

40

50

設計のうちのいずれかに基づくことができる。

【0103】

選択トランジスタ1601は、ビット線(BL0)に結合され、出力及び活性化関数回路1603に結合される。選択トランジスタ1602は、ビット線(BL0)及び共通出力(PEデコード経路)に結合される。

【0104】

BL0のみが活性化されると、選択トランジスタ1602が活性化され、BL0が共通出力に取り付けられる。ANN読み出し動作中、選択トランジスタ1601及び同様のトランジスタがオンにされ、全てのビット線が読み出される。

【0105】

図14、15及び16のデコードに関して、非選択トランジスタについては、トランジスタ漏れがメモリセル性能に影響を及ぼすことを低減するために、負バイアスを与えることができる。あるいは、アレイがANN動作中である間に、PEデコード経路に負バイアスを与えることができる。負バイアスは、 $-0.1V \sim -0.5V$ 以上であり得る。

【0106】

図17は、VMN1700を示す。VMMシステム1700は、VMMアレイ1701及び基準アレイ1720(これは、VMM900、1000、1100、1200、及び1300、又は他のVMM設計など、前述のVMM設計のいずれかに基づくことができる)、低電圧行デコーダ1702、高電圧行デコーダ1703、参照セル低電圧列デコーダ1704(列方向の基準アレイのために示され、行方向に変換を出力するための入力を提供することを意味する)、ビット線PEDライバ1712、ビット線マルチプレクサ1706、活性化関数回路及び加算器1707、制御論理1705、及びアナログバイアス回路1708を備える。

【0107】

図示のように、基準セル低電圧列デコーダ1704は、列方向の基準アレイ1720のためにあり、行方向に変換を出力するための入力を提供することを意味する。基準アレイが行方向にある場合、基準デコーダは、アレイの頂部及び/又は底部で行われ、列方向に変換を出力するための入力を提供する必要がある。

【0108】

低電圧行デコーダ1702は、読み出し及びプログラム動作のためのバイアス電圧を提供し、高電圧行デコーダ1703のデコード信号を提供する。高電圧行デコーダ1703は、プログラム及び消去動作のための高電圧バイアス信号を提供する。参照セル低電圧列デコーダ1704は、参照セルのデコード関数を提供する。ビット線PEDライバ1712は、プログラム、検証、及び消去におけるビット線の制御関数を提供する。バイアス回路1705は、様々なプログラム、消去、プログラム検証、及び読み出し動作に必要な複数の電圧を提供する共有バイアスブロックである。

【0109】

図18は、VMMシステム1800を示す。VMMシステム1800は、VMMシステム1800が、赤色アレイ1801、ビット線PEDライバBLDRV1802、高電圧列デコーダ1803、NVRセクタ1804、及び基準アレイ1820を更に備えることを除いて、VMMシステム1700と同様である。高電圧列デコーダ1803は、垂直デコード線の高電圧バイアスを提供する。赤色アレイ1802は、欠陥のあるアレイ部分を置き換えるためのアレイ冗長性を提供する。NVR(不揮発性レジスタ別情報セクタ)セクタ1804は、ユーザ情報、デバイスID、パスワード、セキュリティキー、トリムビット、構成ビット、製造情報などを記憶するために使用されるアレイセクタであるセクタである。

【0110】

図19は、VMMシステム1900を示す。VMMシステム1900は、VMMシステム1900が参照システム1999を更に備えることを除いて、VMMシステム1800と同様である。基準システム1999は、基準アレイ1901、基準アレイ低電圧行デコ

10

20

30

40

50

ーダ 1902、基準アレイ高電圧行デコーダ 1903、及び基準アレイ低電圧列デコーダ 1904 を備える。基準システムは、複数の VMM システムにわたって共有することができる。VMM システムは、NVR セクタ 1905 を更に備える。

【0111】

基準アレイ低電圧行デコーダ 1902 は、基準アレイ 1901 を伴う読み出し及びプログラミング動作のためのバイアス電圧を提供し、基準アレイ高電圧行デコーダ 1903 のデコード信号も提供する。基準アレイ高電圧行デコーダ 1903 は、基準アレイ 1901 を伴うプログラム及び動作のための高電圧バイアスを提供する。基準アレイ低電圧列デコーダ 1904 は、基準アレイ 1901 のデコード関数を提供する。基準アレイ 1901 は、プログラム検証のための基準ターゲット又はセルマージンを提供する（辺縁セルを探索する）ようなものである。

10

【0112】

図 20 は、ワード線ドライバ 2000 を示す。ワード線ドライバ 2000 は、ワード線（本明細書に示される例示的なワード線 WL0、WL1、WL2、及び WL3 など）を選択し、そのワード線にバイアス電圧を提供する。各ワード線は、制御線 2001 によって制御される選択 *iso*（絶縁）トランジスタ 2002 などの選択トランジスタに取り付けられる。*iso* トランジスタ 2002 は、ワード線デコードトランジスタから消去（例えば、8 ~ 12 V）などの高電圧を絶縁するために使用され、これは I/O トランジスタ（例えば、1.8 V、3.3 V）で実装することができる。ここで、任意の動作中に、制御線 2001 が活性化され、*iso* トランジスタ 2002 を選択するのに類似した全ての選択トランジスタがオンにされる。例示的なバイアストランジスタ 2003（ワード線デコード回路の一部）は、ワード線を第 1 のバイアス電圧（3 V など）に選択的に結合させ、例示的なバイアストランジスタ 2004（ワード線デコード回路の一部）は、ワード線を第 2 のバイアス電圧（第 1 のバイアス電圧より低く、接地を含み、間のバイアスであり、非使用メモリ行からの漏れを低減するための負電圧バイアスである）に選択的に結合させた。ANN 読み出し動作中、全ての使用ワード線が選択され、第 1 のバイアス電圧に結び付けられる。全ての非使用ワード線は、第 2 のバイアス電圧に結び付けられる。プログラム動作などの他の動作中に、一方のワード線のみが選択され、他方のワード線が第 2 のバイアス電圧に結び付けられるまで選択され、これは、アレイ漏れを低減するために負バイアス（例えば、-0.3 ~ -0.5 V 以上）であり得る。

20

30

【0113】

図 21 は、ワード線ドライバ 2100 を示す。ワード線ドライバ 2100 は、バイアストランジスタ 2103 などの頂部トランジスタがバイアス電圧に個別に結合され得ることを除いて、ワード線ドライバ 2000 と同様であり、このようなトランジスタは全てワード線ドライバ 2000 内のように共に結び付けられない。これにより、全てのワード線は、同時に異なる独立した電圧を同時に有することが可能になる。

【0114】

図 22 は、ワード線ドライバ 2200 を示す。ワード線ドライバ 2200 は、バイアストランジスタ 2103 及び 2104 がデコーダ回路 2201 及びインバータ 2202 に結合されることを除いて、ワード線ドライバ 2100 と同様である。したがって、図 22 は、ワード線ドライバ 2200 内のデコードサブ回路 2203 を示す。

40

【0115】

図 23 は、ワード線ドライバ 2300 を示す。ワード線ドライバ 2300 は、バイアストランジスタ 2103 及び 2104 がシフトレジスタ 2301 のステージ 2302 の出力に結合されることを除いて、ワード線ドライバ 2100 と同様である。シフトレジスタ 1301 は、データのシリアルシフト（レジスタをシリアル計時する）によって、データパターンのシフトに応じて 1 つ以上の行が同時に有効になることを可能にするなど、各列を独立して制御することを可能にする。

【0116】

図 24 は、ワード線ドライバ 2400 を示す。ワード線ドライバ 2400 は、各選択ト

50

ランジスタがコンデンサ 2403 などのコンデンサに更に結合されることを除いて、ワード線ドライバ 2000 と同様である。コンデンサ 2403 は、トランジスタ 2401 によって線 2440 の電圧をサンプリングするためにトランジスタ 2401 によって有効にされた、動作の開始時のワード線にプリチャージ又はバイアスを提供することができる。コンデンサ 2403 は、各ワード線に対する入力電圧をサンプリング及び保持するように作用する (S/H)。トランジスタ 2401 は、VMM アレイの ANN 動作 (アレイ電流加算器及び活性化関数) 中にオフであり、S/H コンデンサの電圧がワード線の (浮遊) 電圧源として機能することを意味する。あるいは、コンデンサ 2403 は、メモリアレイからのワード線容量によって提供され得る。

【0117】

図 25 は、ワード線ドライバ 2500 を示す。ワード線ドライバ 2500 は、バイアストランジスタ 2501 及び 2502 がそれぞれスイッチ 2503 及び 2504 に接続されることを除いて、前述のワード線ドライバと同様である。スイッチ 2503 は、opa (演算増幅器) 2505 及びスイッチ 2504 の出力を受信し、スイッチ 2504 は、opa 2505 の負入力に基準入力を提供し、これにより、opa 2505、トランジスタ 2501、スイッチ 2503 及び 2504 によって提供された閉ループの作用により、コンデンサ 2403 によって記憶された電圧を本質的に提供する。このようにして、スイッチ 2503 及び 2504 が閉じられると、入力 2506 の電圧がトランジスタ 2501 によってコンデンサ 2403 に重畳される。あるいは、コンデンサ 2403 は、メモリアレイからのワード線容量によって提供され得る。

【0118】

図 26 は、ワード線ドライバを示す。ワード線ドライバ 2600 は、電圧をワード線 WL0 に駆動するためにコンデンサ 2604 の電圧の電圧バッファとして作用する増幅器 2601 の追加を除いて、前述したワード線ドライバと同様であり、S/H コンデンサの電圧がワード線のための (浮遊) 電圧源として機能することを意味する。これは、例えば、ワード線とワード線の結合がコンデンサの電圧に影響を及ぼすことを回避するためである。

【0119】

図 27 は、高電圧ソース線デコーダ回路 2700 を表す。高電圧ソース線デコーダ回路は、図示のように構成されたトランジスタ 2701、2702 及び 2703 を備える。トランジスタ 2703 は、ソース線を低電圧に選択解除するために使用される。トランジスタ 2702 は、高電圧をアレイのソース線に駆動するために使用され、トランジスタ 2701 はソース線の電圧を監視するために使用される。トランジスタ 2702、2701 及びドライバ回路 (例えば、opa) は、PVT (プロセス、電圧、温度) 及び変動電流負荷条件を維持するために、閉ループ方式 (力/感知) で構成される。SLE (駆動ソース線ノード) 及び SLB (監視されたソース線ノード) は、ソース線の一方の端部にあり得る。あるいは、SLE は、ソース線の一方の端部にあり、SLN が他方の端部にあり得る。

【0120】

図 28 は、図 2 に示されるタイプのメモリセルと共に使用するのに適切である、ワード線デコーダ回路 2801、ソース線デコーダ回路 2804 及び高電圧レベルシフタ 2808 を備える VMM 高電圧デコード回路を示す。

【0121】

ワード線デコーダ回路 2801 は、図示のように構成された、PMOS 選択トランジスタ 2802 (信号 HVO_B によって制御される) 及び NMOS 選択トランジスタ 2803 (信号 HVO_B によって制御される) を備える。

【0122】

ソース線デコーダ回路 2804 は、NMOS 監視トランジスタ 2805 (信号 HVO によって制御される) と、駆動トランジスタ 2806 (信号 HVO によって制御される) と、図示のように構成された (信号 HVO_B によって制御される) トランジスタ 2807 を備える。

【0123】

10

20

30

40

50

高電圧レベルシフタ 2808 は、有効な信号 EN を受信し、高電圧信号 HV 及びその補体 HVO_B を出力する。

【0124】

図 29 は、図 3 に示されるタイプのメモリセルと共に使用するのに適切である、消去ゲートデコーダ回路 2901、制御ゲートデコーダ回路 2904、ソース線デコーダ回路 2907 及び高電圧レベルシフタ 2911 を備える VMM 高電圧デコード回路を示す。

【0125】

消去ゲートデコーダ回路 2901 及び制御ゲートデコーダ回路 2904 は、図 28 のワード線デコーダ回路 2801 と同じ設計を使用する。

【0126】

ソース線デコーダ回路 2907 は、図 28 のソース線デコーダ回路 2804 と同じ設計を使用する。

【0127】

高電圧レベルシフタ 2911 は、図 28 の高電圧レベルシフタ 2808 と同じ設計を使用する。

【0128】

図 30 は、例示的なワード線 WL0、WL1、WL2、及び WL3 のワード線デコーダ 300 を示す。例示的なワード線 WL0 は、プルアップトランジスタ 3001 及びプルダウントランジスタ 3002 に結合される。プルアップトランジスタ 3001 が活性化されると、WL0 が有効になる。プルダウントランジスタ 3002 が活性化されると、WL0 は無効にされる。図 30 の関数は、絶縁トランジスタを有さない図 21 と同様である。

【0129】

図 31 は、例示的な制御ゲート線 CG0、CG1、CG2、及び CG3 の制御ゲートデコーダ 3100 を示す。例示的な制御ゲート線 CG0 は、プルアップトランジスタ 3101 及びプルダウントランジスタ 3102 に結合される。プルアップトランジスタ 3101 が活性化されると、CG0 が有効にされる。プルダウントランジスタ 3102 が活性化されると、CG0 は無効にされる。図 31 の選択及び選択解除関数は、制御ゲートの図 30 の選択及び選択関数と同様である。

【0130】

図 32 は、例示的な制御ゲート線 CG0、CG1、CG2、及び CG3 の制御ゲートデコーダ 3200 を示す。制御ゲートデコーダ 3200 は、制御ゲートデコーダ 3200 が各制御ゲート線に結合されたコンデンサ 3203 などのコンデンサを含むことを除いて、制御ゲートデコーダ 3100 と同様である。これらのサンプルアンドホールド (S/H) コンデンサは、動作前に各制御ゲート線にプリチャージバイアスを提供することができ、S/H コンデンサの電圧が制御ゲート線の (浮遊) 電圧源として機能することを意味する。S/H コンデンサは、メモリセルからの制御ゲート容量によって提供され得る。

【0131】

図 33 は、例示的な制御ゲート線 CG0、CG1、CG2、及び CG3 の制御ゲートデコーダ 3300 を示す。制御ゲートデコーダ 3300 は、制御ゲートデコーダ 3300 がバッファ 3301 (opa など) を更に備えることを除いて、制御ゲートデコーダ 3200 と同様である。

【0132】

図 34 は、電流電圧回路 3400 を示す。回路は、構成されたダイオード接続基準セル回路 3450 及びサンプルアンドホールド回路 3460 を備える。回路 3450 は、入力電流源 3401、NMOS トランジスタ 3402、カスコーディングバイアストラジスタ 3403、及び基準メモリセル 3404 を備える。サンプルアンドホールド回路は、スイッチ 3405 及び S/H コンデンサ 3406 からなる。メモリ 3404 は、ワード線を供給するためなど、入力電流を電圧に変換するために、ビット線のバイアスを有するダイオード接続構成でバイアスされる。

【0133】

10

20

30

40

50

図35は、電流電圧回路3500を示しており、電流電圧回路3500は、電流電圧回路3400と同様であるが、それにS/Hコンデンサの後に増幅器3501を追加している。電流電圧回路3500は、構成されたダイオード接続基準セル回路3550、サンプルアンドホールド回路3470、及び増幅器ステージ3562を備える。

【0134】

図36は、ダイオード接続構成における制御ゲートの電流電圧回路3400と同じ設計である電流電圧回路3600を示す。電流電圧回路3600は、構成されたダイオード接続基準セル回路3650及びサンプルアンドホールド回路3660を備える。

【0135】

図37は、基準回路3750とS/H回路3760との間にバッファ3790が配置される電流電圧回路3700を示す。

10

【0136】

図38は、ダイオード接続構成で接続された制御ゲートを有する図35と同様の電流電圧回路3800を示す。電流電圧回路3800は、構成されたダイオード接続基準セル回路3550、サンプルアンドホールド回路3870、及び増幅器ステージ3862を備える。

【0137】

図39は、図2のメモリセルに適用された図34と同様の電流電圧回路3900を示す。電流電圧回路3900は、構成されたダイオード接続基準セル回路3950及びサンプルアンドホールド回路3960を備える。

20

【0138】

図40は、図2のメモリセルに適用される図37と同様の電流電圧回路4000を示し、バッファ4090は、基準回路4050とS/H回路4060との間に配置される。

【0139】

図41は、図2のメモリセルに適用された図38と同様の電流電圧回路4100を示す。電流電圧回路4100は、構成されたダイオード接続基準セル回路4150、サンプルアンドホールド回路4170、及び増幅器ステージ4162を備える。

【0140】

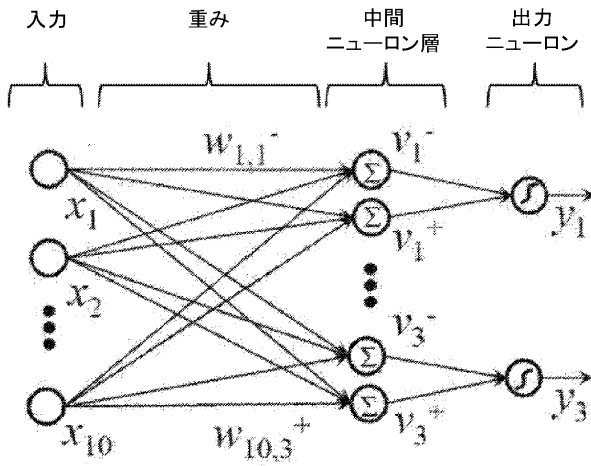
本明細書で使用される場合、「の上方に(over)」及び「に(on)」という用語は両方とも、「に直接」(中間材料、要素、又は空間がそれらの間に何ら配設されない)、及び「の上に間接的に」(中間材料、要素、又は空間がそれらの間に配設される)を包括的に含むことに留意するべきである。同様に、「隣接した」という用語は、「直接隣接した」(中間材料、要素、又は空間がそれらの間に何ら配設されない)、及び「間接的に隣接した」(中間材料、要素、又は空間がそれらの間に配設される)を含み、「に取り付けられた」は、「に直接取り付けられた」(中間材料、要素、又は空間がそれらの間に何ら配設されない)、及び「に間接的に取り付けられた」(中間材料、要素、又は空間がそれらの間に配設される)を含み、「電氣的に結合された」は、「に直接電氣的に結合された」(要素と一緒に電氣的に連結する中間材料又は要素がそれらの間にない)、及び「に間接的に電氣的に結合された」(要素と一緒に電氣的に連結する中間材料又は要素がそれらの間にある)を含む。例えば、要素を「基板の上方に」形成することは、その要素を基板に直接、中間材料/要素をそれらの間に何ら伴わずに、形成すること、及びその要素を基板の上に間接的に、1つ以上の中間材料/要素をそれらの間に伴って、形成することを含み得る。

30

40

【図面】

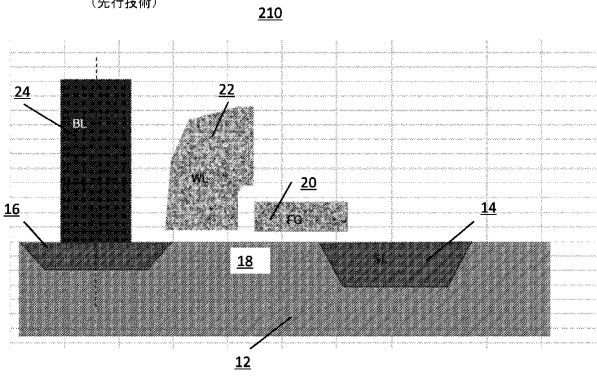
【図 1】



(先行技術)

【図 2】

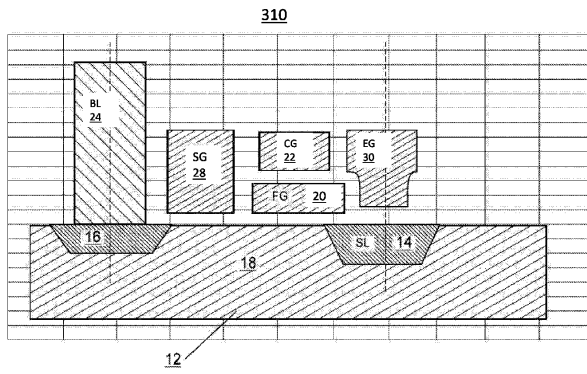
(先行技術)



10

【図 3】

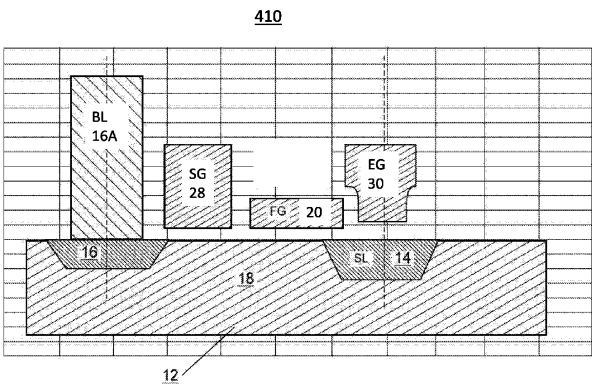
(先行技術)



12

【図 4】

(先行技術)



12

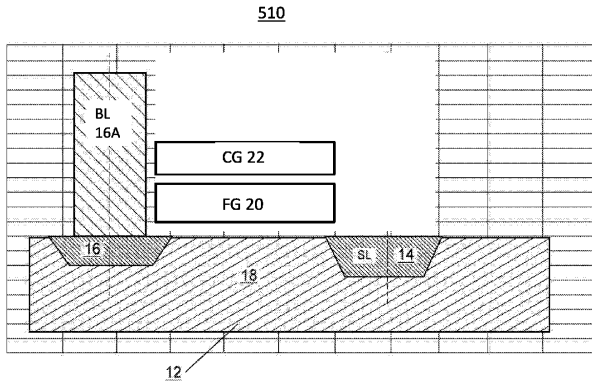
30

40

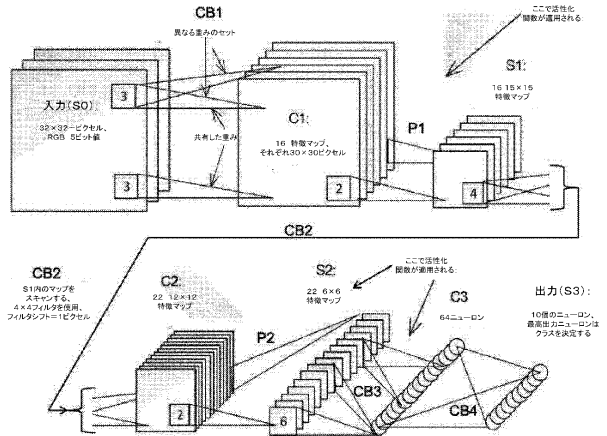
50

【図5】

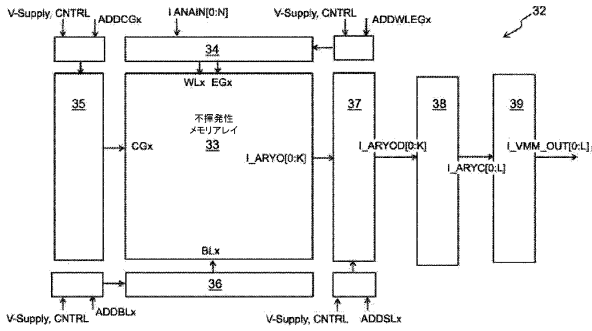
(先行技術)



【図6】



【図7】



【図8】

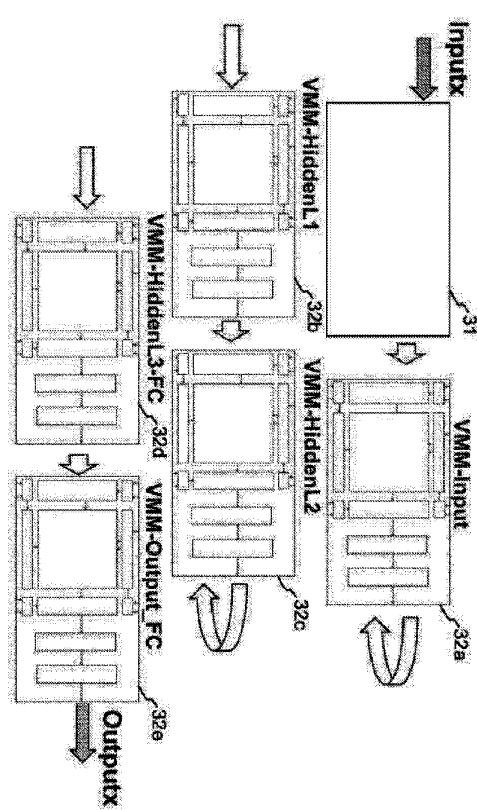


FIGURE 8

10

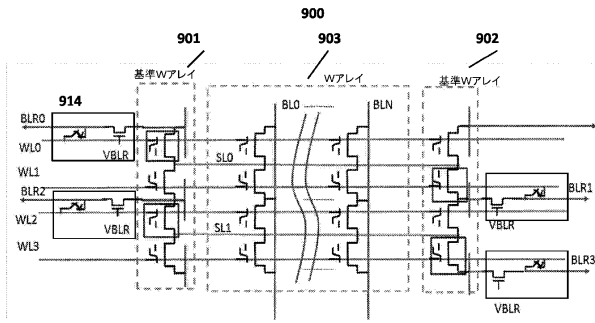
20

30

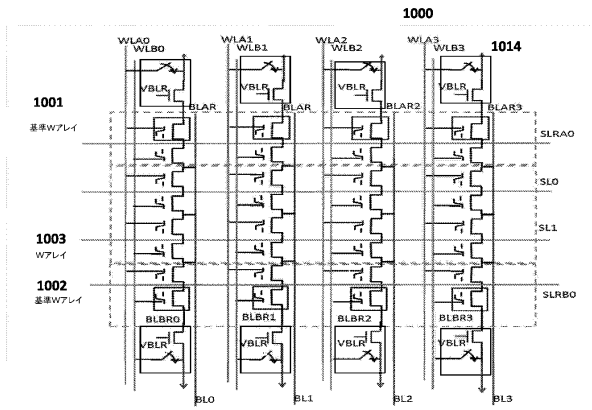
40

50

【 図 9 】

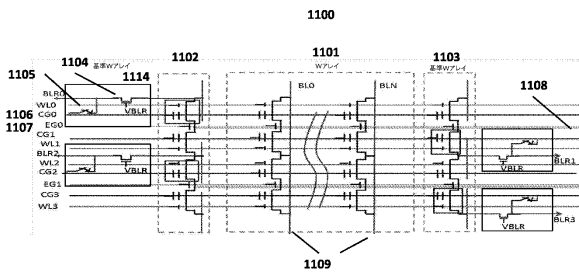


【 図 10 】

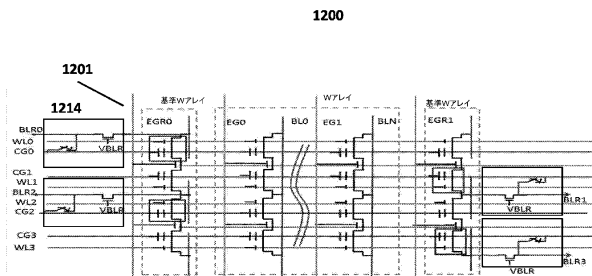


10

【 図 11 】

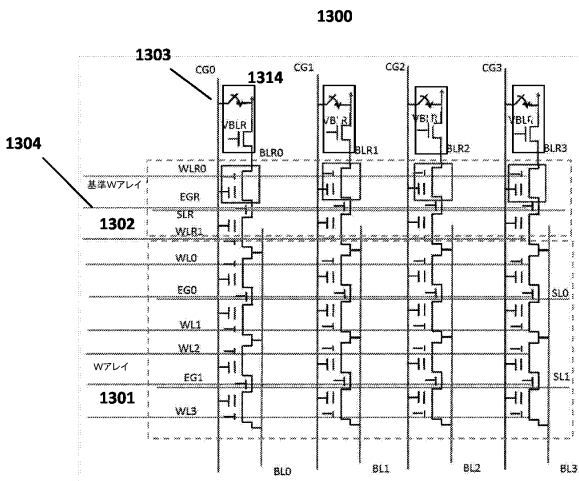


【 図 12 】

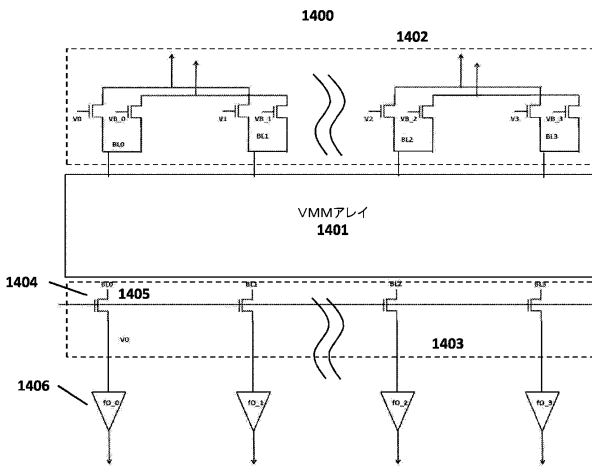


20

【 図 13 】



【 図 14 】

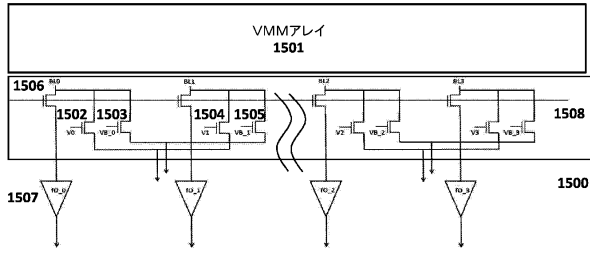


30

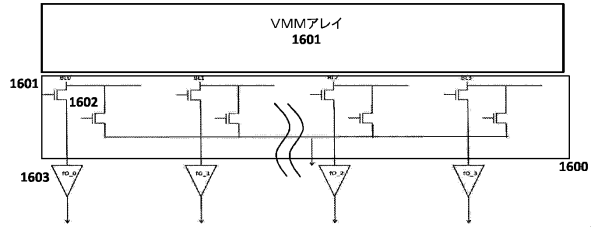
40

50

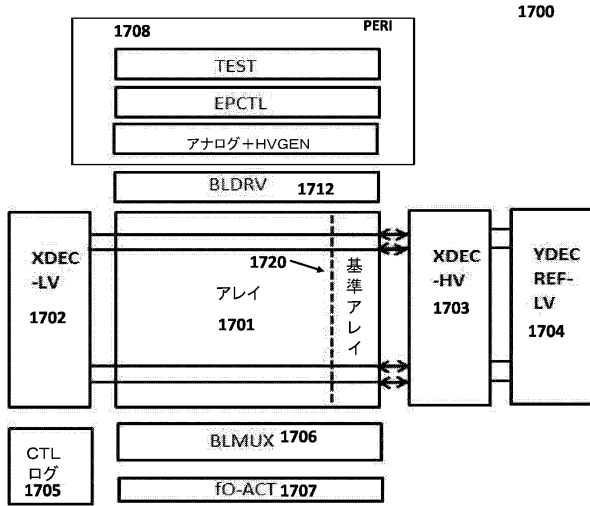
【 図 15 】



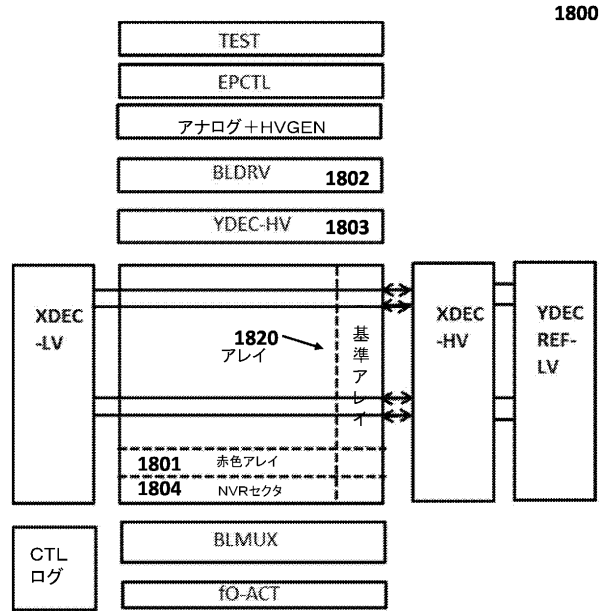
【 図 16 】



【 図 17 】



【 図 18 】



10

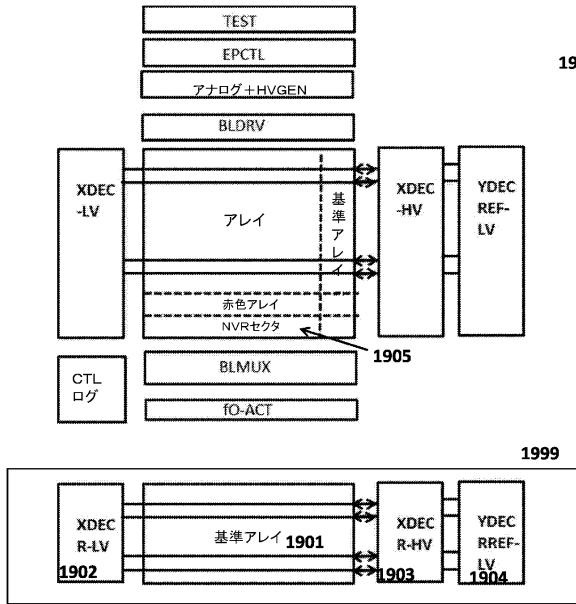
20

30

40

50

【図 19】



【図 20】

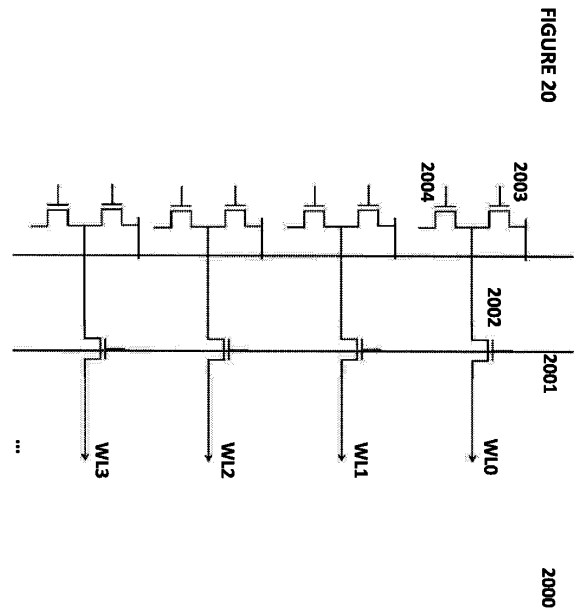


FIGURE 20

【図 21】

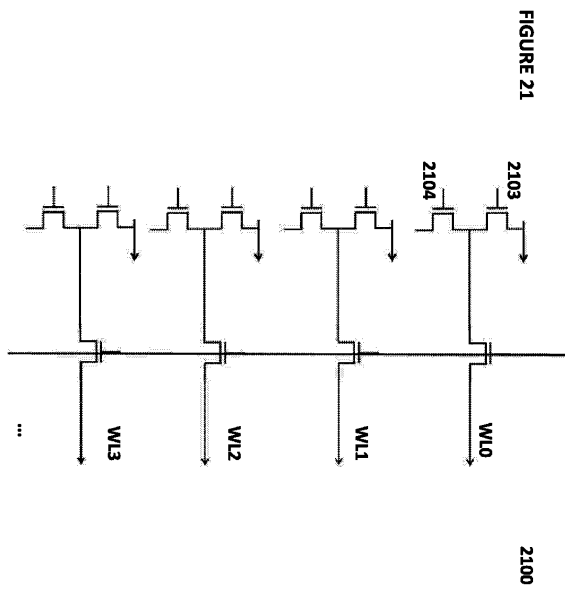


FIGURE 21

【図 22】

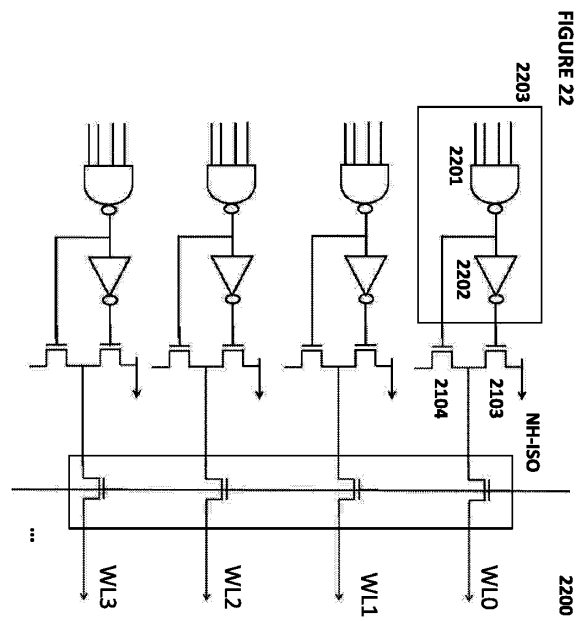


FIGURE 22

10

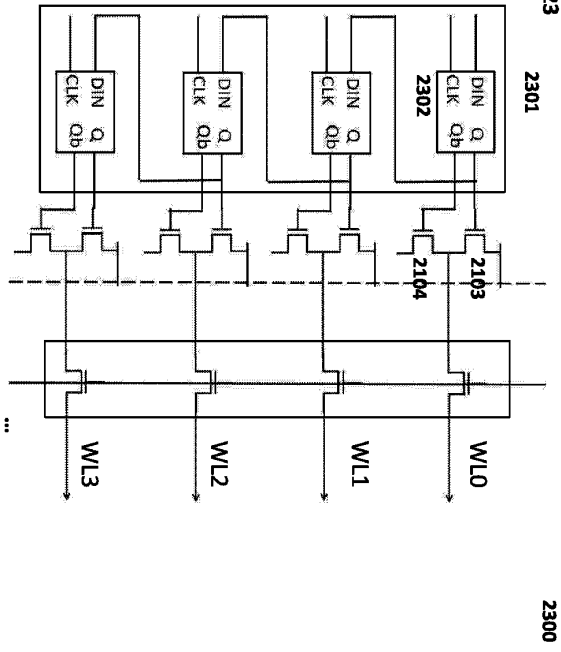
20

30

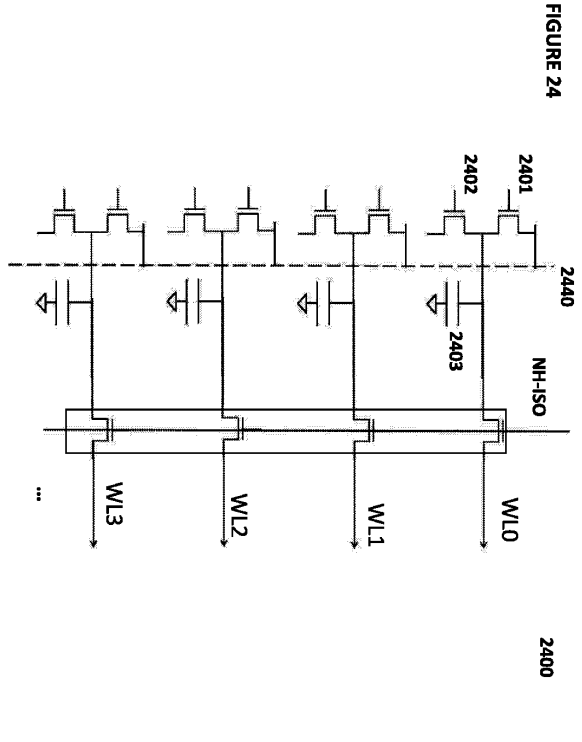
40

50

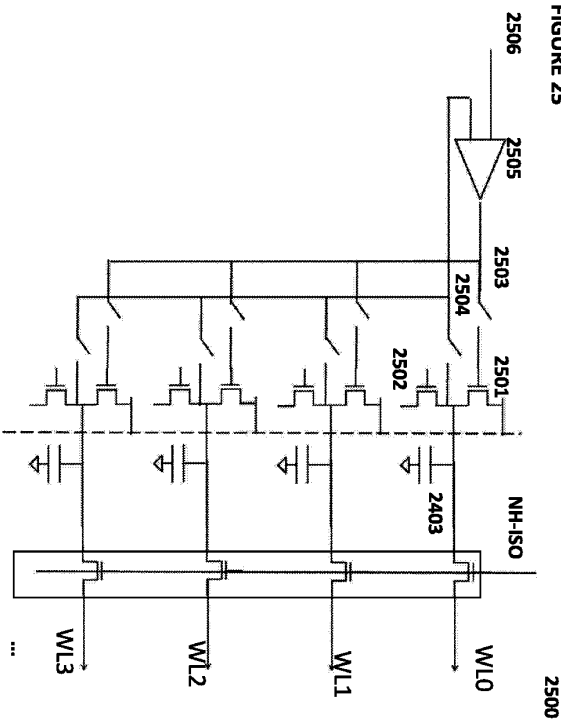
【 2 3 】



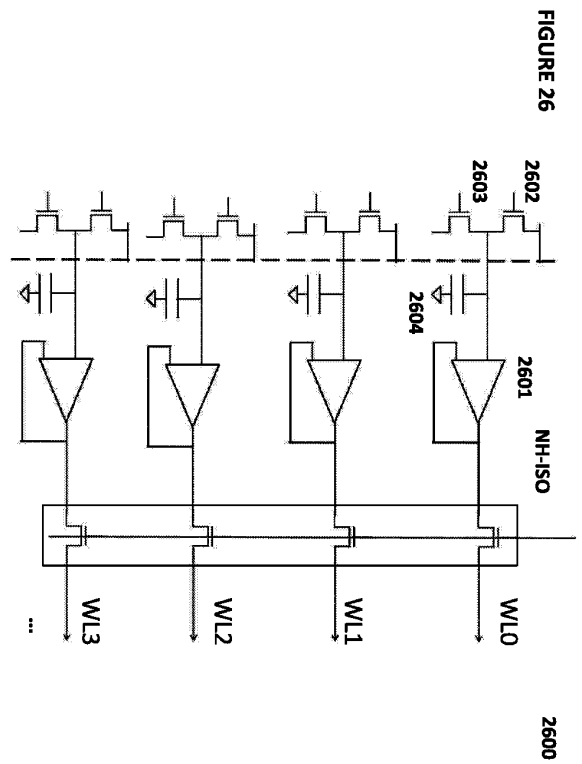
【 2 4 】



【 2 5 】



【 2 6 】



10

20

30

40

50

【 3 1 】

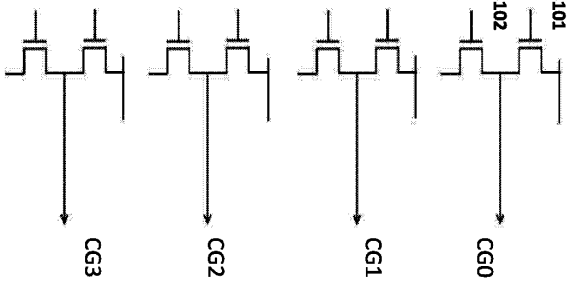


FIGURE 31

3100

【 3 2 】

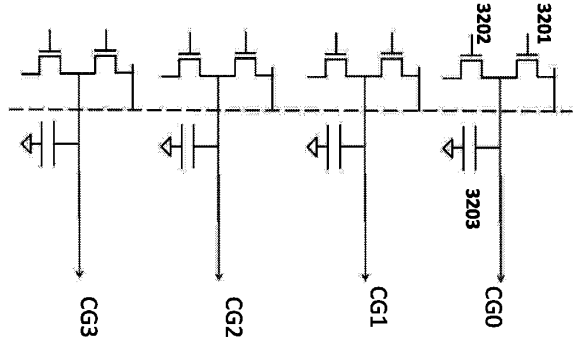


FIGURE 32

3200

【 3 3 】

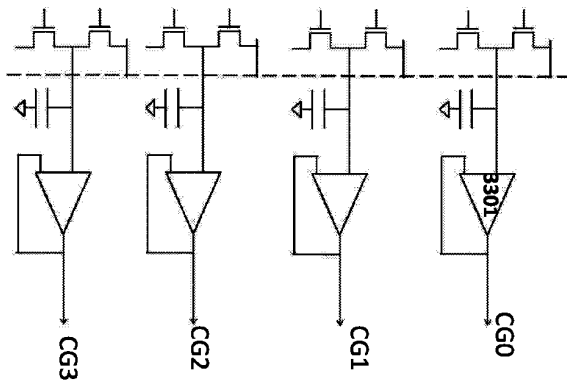


FIGURE 33

3300

【 3 4 】

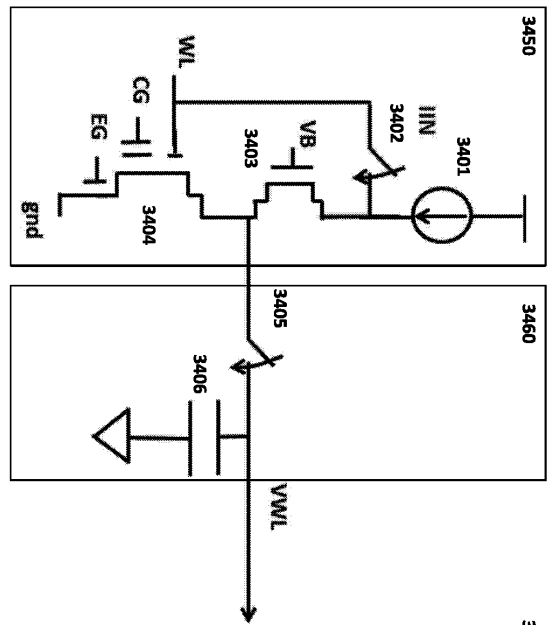


FIGURE 34

3400

10

20

30

40

50

【 3 5 】

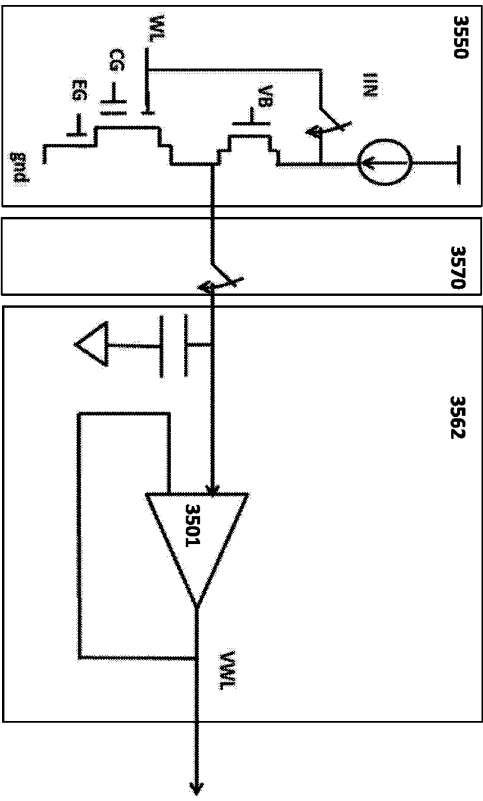


FIGURE 35

【 3 6 】

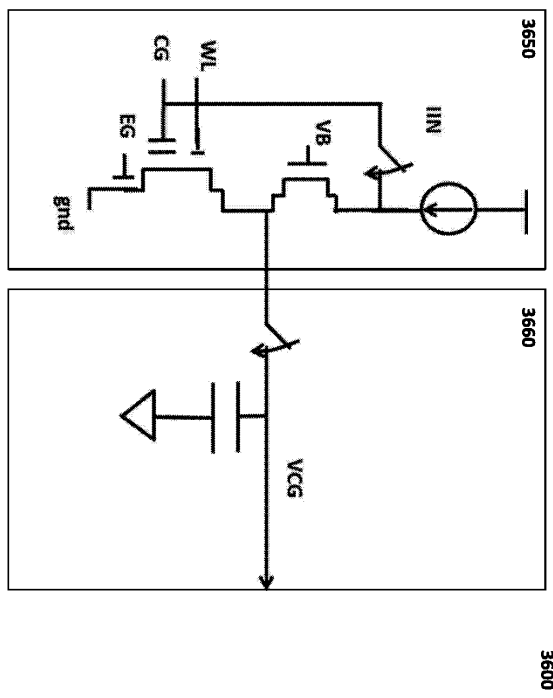


FIGURE 36

10

20

【 3 7 】

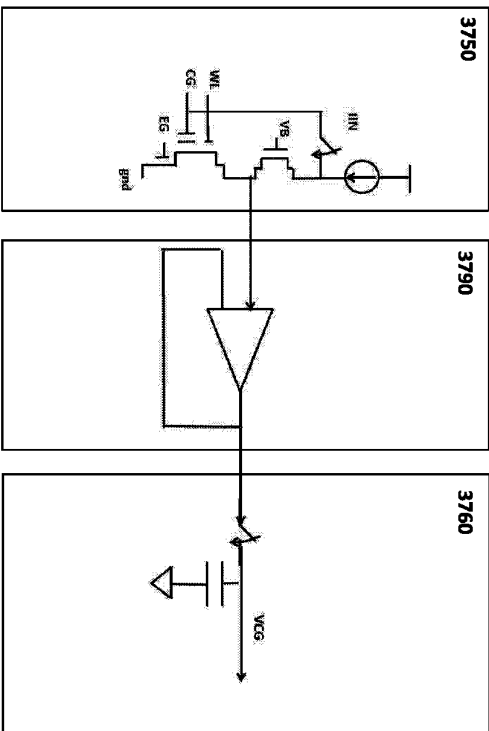


FIGURE 37

【 3 8 】

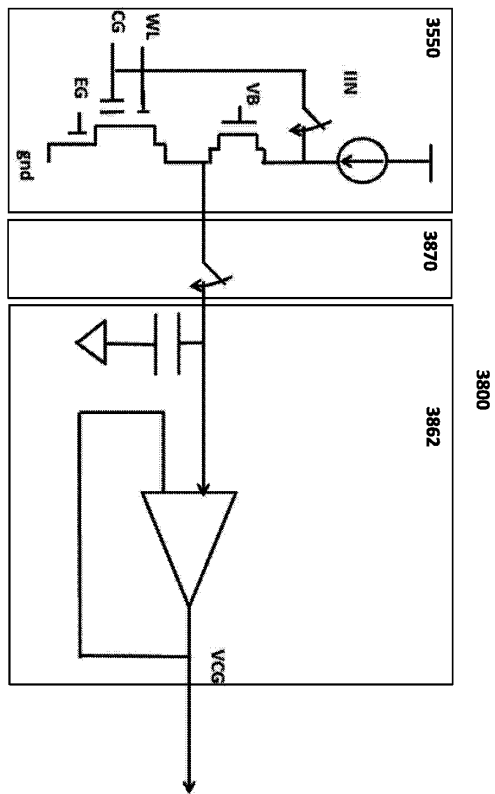


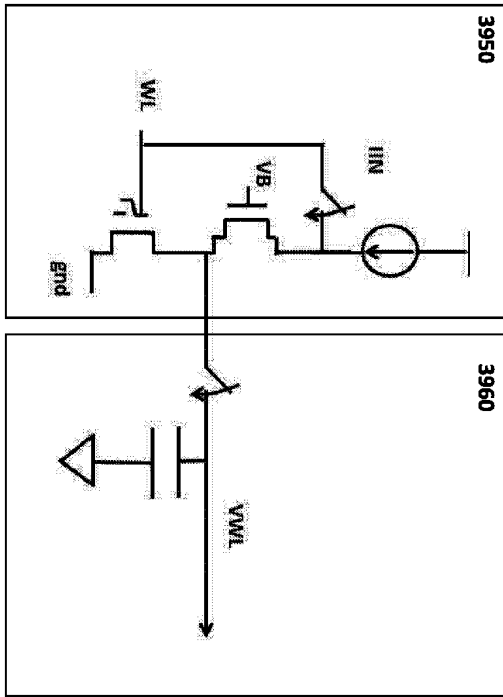
FIGURE 38

30

40

50

【図 39】



【図 40】

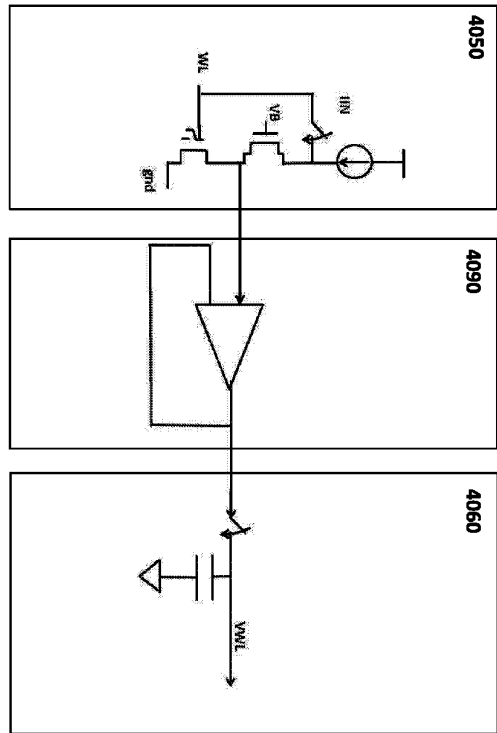


FIGURE 40

FIGURE 39

10

20

【図 41】

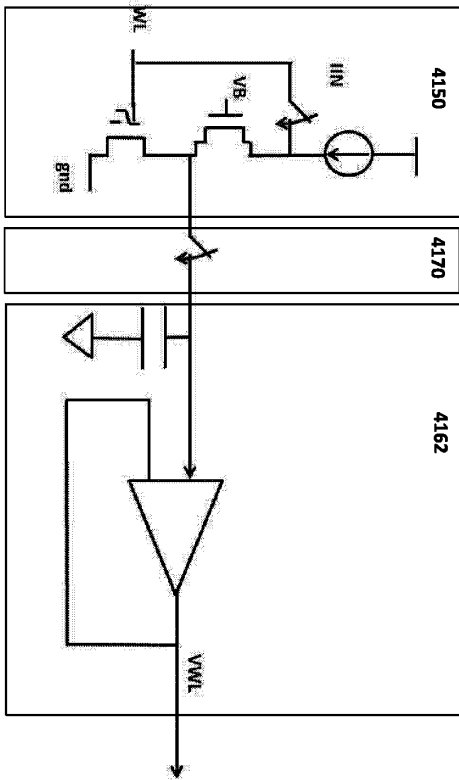


FIGURE 41

【図 42】

	WL	WL-非選択	BL	BL-非選択	SL	SL-非選択
読み出し	1-3.5V	-0.5V/0V	0.6~2V (0ニューロン)	0.6V-2V/0V	0V	0V
消去	-5-13V	0V	0V	0V	0V	0V
プログラム	1-2V	-0.5V/0V	0.1-3 uA	Vinh 約 2.5V	4-10V	0~1V/FLT

30

40

50

【 図 4 3 】

	WL	WL-非選択	BL	BL-非選択	SL	SL-非選択
読み出し	1-3.5V	-0.5V/0V	0.6-2V	0.6V-2V/0V	約1V (1ニューロン)	0V
消去	-5-13V	0V	0V	0V	0V	SL-禁止 (約4-8V)
プログラミング	1-2V	-0.5V/0V	0.1-3 μA	V _{inh} 約2.5V	4-10V	0-1V/FILT

【 図 4 4 】

	WL	WL-非選択	BL	BL-非選択	CG	CG-同じセクター 非選択	CG-非選択	EG	EG-非選択	SL	SL-非選択
読み出し	1.0-2V	-0.5V/0V	0.6-2V (1ニューロン)	0V	0-2.6V		0-2.6V	0-2.6V	0-2.6V	0V	0V
消去	0V	0V	0V	0V	0V		0-2.6V	0-2.6V	5-12V	0-2.6V	0V
プログラム	0.7-1V	-0.5V/0V	0.1-1μA	V _{inh} (1-2V)	4-11V		0-2.6V	0-2.6V	4.5-5V	0-2.6V	4.5-5V 0-1V

【 図 4 5 】

	WL	WL-非選択	BL	BL-非選択	CG	CG-同じセクター 非選択	CG-非選択	EG	EG-非選択	SL	SL-非選択
読み出し	1.0-2V	-0.5V/0V	0.6-2V (1ニューロン)	0V	0-2.6V		0-2.6V	0-2.6V	0-2.6V	0V	0V
消去	0V	0V	0V	0V	0V	CGINH(4-8V)	0-2.6V	5-12V	0-2.6V	0V	0V
プログラム	0.7-1V	-0.5V/0V	0.1-1μA	V _{inh} (1-2V)	4-11V		0-2.6V	0-2.6V	4.5-5V	0-2.6V	4.5-5V 0-1V

10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

ニア州、サンノゼ、プリストル ベイ コモン 1 8 4 8

(72)発明者 リ、アン

アメリカ合衆国 9 5 1 2 1 カリフォルニア州、サンノゼ、リンドミュア ドライブ 3 3 8 5

(72)発明者 ブー、サン

アメリカ合衆国 9 5 1 3 8 カリフォルニア州、サンノゼ、ダンナ コート 4 3 1

(72)発明者 パム、ヒエン

ベトナム国 ホーチミン、ゴ バップ ディストリクト、ワード 1 5、レ ダック トー ストリート、7 3 6 / 1 6 3 / 1 8

(72)発明者 ニュエン、カ

ベトナム国 ホーチミン、ゴ バップ ディストリクト、ワード 1 6、ニュエン パン ルオン ストリート、3 5 0 / 1 6 0

(72)発明者 トラン、ハン

ベトナム国 ホーチミン、ゴ バップ ディストリクト、ワード 1 4、カンギア アパ - トメント、ロット エー、ルーム 5 . 0 2

審査官 後藤 彰

(56)参考文献 国際公開第 2 0 1 7 / 2 0 0 8 8 3 (W O , A 1)

特表 2 0 1 9 - 5 1 7 1 3 8 (J P , A)

特開平 0 3 - 1 7 4 6 7 9 (J P , A)

(58)調査した分野 (Int.Cl. , D B 名)

G 1 1 C 1 6 / 0 8

G 1 1 C 1 1 / 5 4

G 1 1 C 1 6 / 0 4

G 0 6 G 7 / 6 0

G 0 6 N 3 / 0 6 3