

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7210585号
(P7210585)

(45)発行日 令和5年1月23日(2023.1.23)

(24)登録日 令和5年1月13日(2023.1.13)

(51)国際特許分類	F I			
G 1 1 C 11/56 (2006.01)	G 1 1 C 11/56	2 1 0		
G 1 1 C 11/54 (2006.01)	G 1 1 C 11/54			
G 1 1 C 16/10 (2006.01)	G 1 1 C 16/10	1 4 0		

請求項の数 18 (全35頁)

(21)出願番号	特願2020-529370(P2020-529370)	(73)特許権者	500147506
(86)(22)出願日	平成30年10月30日(2018.10.30)		シリコン ストリージ テクノロジー
(65)公表番号	特表2021-504868(P2021-504868 A)		インコーポレイテッド
(43)公表日	令和3年2月15日(2021.2.15)		S I L I C O N S T O R A G E T E C
(86)国際出願番号	PCT/US2018/058297		H N O L O G Y , I N C .
(87)国際公開番号	WO2019/108334		アメリカ合衆国 9 5 1 3 4 カリフォル
(87)国際公開日	令和1年6月6日(2019.6.6)	(74)代理人	ニア州 サンノゼ ホルガー ウェイ 4 5 0
審査請求日	令和3年10月13日(2021.10.13)		110000626
(31)優先権主張番号	15/826,345	(72)発明者	弁理士法人英知国際特許商標事務所
(32)優先日	平成29年11月29日(2017.11.29)		トラン、ヒュー バン
(33)優先権主張国・地域又は機関	米国(US)	(72)発明者	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
			ティワリ、ビビン
			アメリカ合衆国 9 4 5 6 8 カリフォル
			最終頁に続く

(54)【発明の名称】 人工ニューラルネットワーク内のアナログニューロモーフィックメモリのための高精度かつ高効率な調整機構及びアルゴリズム

(57)【特許請求の範囲】

【請求項 1】

不揮発性メモリセルのアレイ内で、迅速にプログラムすることができる前記不揮発性メモリセルを識別する方法であって、該方法は、
不揮発性メモリセルを消去するステップと、
前記不揮発性メモリセルを第1の電圧でプログラミングするステップと、
読み出し動作中に前記不揮発性メモリセルを通る第1の電流を測定するステップと、
前記不揮発性メモリセルを第2の電圧でプログラミングするステップと、
読み出し動作中に前記不揮発性メモリセルを通る第2の電流を測定するステップと、
前記第2の電流と前記第1の電流との間の差が閾値を超える場合、前記不揮発性メモリセルが迅速にプログラムされ得ることを示すデータを格納するステップと、を含む、方法。

【請求項 2】

前記データは、前記不揮発性メモリセルのアドレスを含む、請求項1に記載の方法。

【請求項 3】

前記データは、単一ビットを構成する、請求項1に記載の方法。

【請求項 4】

前記メモリセルは、分割2ゲートフラッシュメモリセルである、請求項1に記載の方法。

【請求項 5】

前記メモリセルは、分割3ゲートフラッシュメモリセルである、請求項1に記載の方法。

【請求項 6】

前記メモリセルは、分割４ゲートフラッシュメモリセルである、請求項１に記載の方法。
【請求項７】

前記メモリセルは、サブ閾値領域で読み出される、請求項１に記載の方法。

【請求項８】

不揮発性メモリセルのアレイ内で、迅速にプログラムすることができる前記不揮発性メモリセルを識別する方法であって、該方法は、

不揮発性メモリセルをプログラミングするステップと、

前記不揮発性メモリセルを第１の電圧で消去するステップと、

読み出し動作中に前記不揮発性メモリセルを通る第１の電流を測定するステップと、

前記不揮発性メモリセルを第２の電圧で消去するステップと、

読み出し動作中に前記不揮発性メモリセルを通る第２の電流を測定するステップと、

前記第２の電流と前記第１の電流との間の差が閾値を超える場合、前記不揮発性メモリセルが迅速にプログラムされ得ることを示すデータを格納するステップと、を含む、方法

。

【請求項９】

前記データは、前記不揮発性メモリセルのアドレスを含む、請求項８に記載の方法。

【請求項１０】

前記データは、単一ビットを構成する、請求項８に記載のシステム。

【請求項１１】

前記メモリセルは、分割２ゲートフラッシュメモリセルである、請求項８に記載の方法。

【請求項１２】

前記メモリセルは、分割３ゲートフラッシュメモリセルである、請求項８に記載の方法。

【請求項１３】

前記メモリセルは、分割４ゲートフラッシュメモリセルである、請求項８に記載の方法。

【請求項１４】

前記メモリセルは、サブ閾値領域で読み出される、請求項８に記載の方法。

【請求項１５】

セルのプログラミング速度特性に基づいて、複数の不揮発性アナログニューロモーフィックメモリセルをプログラミングする方法であって、該方法は、

セルの前記プログラミング速度特性を決定するステップと、

前記プログラミング速度特性が第１の値を有する場合、前記セルの浮遊ゲートで所望の電荷レベルを達成するために、前記セルに第１の調整アルゴリズムを実行するステップと、

前記プログラミング速度特性が第２の値を有する場合、前記セルの前記浮遊ゲートで前記所望の電荷レベルを達成するために、前記セルに第２の調整アルゴリズムを実行するステップと、を含む、

前記メモリセルはアレイに配置され、水平方向に配置されたワード線及びソース線、並びに垂直方向に配置されたビット線を有し、前記ビット線は電流出力ニューロンとして動作し、

電圧入力は、制御ゲートに提供され、

ダイオード接続された調整された基準セルに流れる入力電流は、前記電圧入力を提供するために使用される、方法。

【請求項１６】

前記ダイオード接続された調整された基準セルは、前記アレイ内のメモリセルの行に使用される、請求項１５に記載の方法。

【請求項１７】

セルのプログラミング速度特性に基づいて、複数の不揮発性アナログニューロモーフィックメモリセルをプログラミングする方法であって、該方法は、

セルの前記プログラミング速度特性を決定するステップと、

前記プログラミング速度特性が第１の値を有する場合、前記セルの浮遊ゲートで所望の電荷レベルを達成するために、前記セルに第１の調整アルゴリズムを実行するステップと、

10

20

30

40

50

前記プログラミング速度特性が第2の値を有する場合、前記セルの前記浮遊ゲートで前記所望の電荷レベルを達成するために、前記セルに第2の調整アルゴリズムを実行するステップと、を含み、

前記メモリセルはアレイに配置され、水平方向に配置されたワード線及びソース線、並びに垂直方向に配置されたビット線を有し、前記ビット線は電流出力ニューロンとして動作し、

電圧入力は、前記ワード線に対するものであり、

ダイオード接続された調整された基準セルに流れる入力電流は、前記電圧入力を提供するために使用される、方法。

【請求項18】

前記ダイオード接続された調整された基準セルは、行に使用される、請求項17に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願)

本特許出願は、2017年11月29日に出願された米国特許出願第15/826,345号の利益を主張する。

【0002】

(発明の分野)

人工ニューラルネットワークにおいて使用されるアナログニューロモフィックメモリ内のセルを調整するための多数の実施形態が開示されている。

【背景技術】

【0003】

人工ニューラルネットワークは、多数の入力によって異なり得る機能を推定する又は近似するために使用され、概ね未知であ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）によく似ている。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。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。

【0004】

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。CMOS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CMOS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。

【0005】

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15/594,439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フィック

10

20

30

40

50

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ている第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、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の部分の上方に配置され、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の部分の上方に配置され、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、をそれぞれが含む複数のメモリセルを含む。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格納するように構成されている。複数のメモリセルは、第1の複数の入力に、格納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ている。

10

【0006】

アナログニューロモーフィック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、各浮動ゲートは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が挙げられる。先行技術は、セルが所望の量の電荷を含むことを確実にするために、各セルを調整するための高速かつ正確な機構を欠いている。

【0007】

人工ニューラルネットワークで使用されるアナログニューロモーフィックメモリを調整するための改良された機構及びアルゴリズムが必要とされている。

20

【0008】

【0009】

【0010】

【0011】

【0012】

【0013】

【0014】

【0015】

【0016】

【0017】

【0018】

【0019】

【0020】

【0021】

【0022】

【0023】

【0024】

【0025】

【0026】

【0027】

【0028】

【0029】

【0030】

【0031】

【0032】

【0033】

【0034】

【0035】

【0036】

30

40

50

【0037】

【0038】

【0039】

【0040】

【0041】

【0042】

【0043】

【0044】

【0045】

【0046】

10

【0047】

【0048】

【0049】

【0050】

【0051】

【0052】

【0053】

【0054】

【0055】

【図面の簡単な説明】

20

【0056】

【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ゲートの不揮発性メモリセルの側面断面図である。

【図3】図2のメモリセルに対する従来のアレイアーキテクチャを示す図である。

【図4】従来の2ゲートの不揮発性メモリセルの側面断面図である。

【図5】図4のメモリセルに対する従来のアレイアーキテクチャを示す図である。

【図6】従来の4ゲートの不揮発性メモリセルの側面断面図である。

【図7】図6のメモリセルに対する従来のアレイアーキテクチャを示す図である。

【図8A】均等な間隔であるニューラルネットワークの重みレベルの割当てを示す図である。

30

【図8B】不均等な間隔であるニューラルネットワークの重みレベルの割当てを示す図である。

【図9】双方向調整アルゴリズムを示すフロー図である。

【図10】電流比較を使用した重みマッピングを示すブロック図である。

【図11】電圧比較を使用した重みマッピングを示すブロック図である。

【図12】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ニューラルネットワークを示す図である。

【図13】ベクトル乗算器マトリクスを示すブロック図である。

【図14】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15】4ゲートのメモリセルのアレイの第1のアーキテクチャを示す概略図である。

40

【図16】4ゲートのメモリセルのアレイの第1のアーキテクチャを示す概略図である。

【図17】4ゲートのメモリセルのアレイの第2のアーキテクチャを示す概略図である。

【図18】4ゲートのメモリセルのアレイの第2のアーキテクチャを示す概略図である。

【図19】4ゲートのメモリセルのアレイの第3のアーキテクチャを示す概略図である。

【図20】4ゲートのメモリセルのアレイの第4のアーキテクチャを示す概略図である。

【図21】4ゲートのメモリセルのアレイの第5のアーキテクチャを示す概略図である。

【図22】4ゲートのメモリセルのアレイの第6のアーキテクチャを示す概略図である。

【図23】2ゲートのメモリセルのアレイの第1のアーキテクチャを示す概略図である。

【図24】2ゲートのメモリセルのアレイの第2のアーキテクチャを示す概略図である。

【図25】電流／電圧ログ変換器を示す図である。

50

【図 2 6】電圧／電流ログ変換器を示す図である。

【図 2 7】接地基準（Gnd-referred）電流加算器を示す図である。

【図 2 8】V d d 基準（Vdd-referred）電流加算器を示す図である。

【図 2 9】不揮発性メモリアレイの N^2 ニューラルネット入力の利用を示す図である。

【図 3 0】不揮発性メモリアレイの N^2 ニューラルネット入力の利用を示す図である。

【図 3 1】周期的にシフトする入力線を有する不揮発性メモリアレイのニューラルネット入力の利用を示す図である。

【図 3 2】図 1 5 のメモリアレイアーキテクチャであるが、周期的にシフトする入力線を有するメモリアレイアーキテクチャを示す概略図である。

【図 3 3】図 2 0 のメモリアレイアーキテクチャであるが、周期的にシフトする入力線を有するメモリアレイアーキテクチャを示す概略図である。

【図 3 4】人工ニューラルネットワークで利用することができる先行技術の不揮発性メモリセルの断面図である。

【図 3 5】人工ニューラルネットワークで使用するための不揮発性メモリセルのアレイを示す。

【図 3 6】図 3 5 の不揮発性メモリセルで動作を実行するための動作電圧を示す。

【図 3 7】人工ニューラルネットワークで使用するための不揮発性メモリセルのアレイを示す。

【図 3 8】図 3 7 の不揮発性メモリセルで動作を実行するための動作電圧を示す。

【図 3 9】人工ニューラルネットワークで使用するための不揮発性メモリセルのアレイを示す。

【図 4 0】図 3 9 の不揮発性メモリセルで動作を実行するための動作電圧を示す。

【図 4 1】人工ニューラルネットワークで使用するための不揮発性メモリセルのアレイを示す。

【図 4 2】図 4 1 の不揮発性メモリセルで動作を実行するための動作電圧を示す。

【図 4 3】不揮発性メモリセルのアレイ内の異なるプログラミング速度特性を有するセルを調整するためのアルゴリズムを示す。

【図 4 4】不揮発性メモリセルのアレイ内で迅速にプログラムすることができるセルを識別するためのアルゴリズムを示す。

【図 4 5】粗いプログラミングのシーケンス及び微細なプログラミングのシーケンスを含むセルを調整するためのアルゴリズムを示す。

【図 4 6】不揮発性メモリセルを調整するための均一ステップアルゴリズムを示す。

【図 4 7】不揮発性メモリセルを調整するための粗いステップ及び微細なステップを含む、均一なログステップアルゴリズムを示す。

【図 4 8】不揮発性メモリセルを調整するための粗いステップ及び微細なステップを含む、バイナリサーチステップアルゴリズムを示す。

【図 4 9】不揮発性メモリセルのための調整シーケンスの波形を示す。

【図 5 0】不揮発性メモリセルのための調整シーケンスの波形を示す。

【発明を実施するための形態】

【0 0 5 7】

本発明の人工ニューラルネットワークは、C M O S 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。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, 0 2 9, 1 3 0 号（「' 1 3 0 号特許」）は、分割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。メモリセルを図 2 に示す。各メモリセル 1 0 は、半導体基板 1 2 内に形成され、チャネル領域 1 8 をその間に有するソース領域及びドレイン領域 1 4 / 1 6 を含む。浮遊ゲート 2 0 は、チャネル領域 1 8 の第 1 の部分の上方に形成され、チャネル領域 1 8 の第 1 の部分から絶縁され（かつその伝導率を制御する）、及びドレイン領域 1 6 の一部の上方に形成される。制御ゲート 2 2 は、チャネル領域 1 8 の第 2 の部分の上方に配設され、チャネル領域 1 8 の第 2 の部分から絶縁される（かつその伝導率を制御する）第 1 の部分 2 2 a と、浮遊ゲート 2 0 の上

10

20

30

40

50

方に延在する第２の部分２２ｂと、を有する。浮遊ゲート２０及び制御ゲート２２は、ゲート酸化物２６によって基板１２から絶縁される。

【００５８】

メモリセルは、制御ゲート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から制御ゲート２２まで中間絶縁体２４を通して浮遊ゲート２０の電子をトンネリングさせる。

【００５９】

メモリセルは、制御ゲート２２に正電圧、及びドレイン１６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１４からドレイン１６に向かって流れることになる。電子は、制御ゲート２２と浮遊ゲート２０との間の隙間に達すると、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２０からの静電引力に起因してゲート酸化物２６を通して浮遊ゲート２０のに注入される。

【００６０】

メモリセルは、ドレイン１６及び制御ゲート２２に正の読み出し電圧を印加することによって読み出される（制御ゲート下のチャネル領域をオンにする）。浮遊ゲート２０が正に帯電する（すなわち、電子を消去し、ドレイン１６に正に結合する）場合、浮遊ゲート２０下のチャネル領域の部分は、次に同様にオンになり、電流は、チャネル領域１８を流れ、これは、消去された状態つまり「１」の状態として検知される。浮遊ゲート２０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２０下のチャネル領域の部分は、ほとんど又は完全にオフになり、電流は、チャネル領域１８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。

【００６１】

先行技術のメモリアレイのアーキテクチャを図３に示す。メモリセル１０は、行及び列に配置される。各列において、メモリセルは、ミラー方式にエンドツーエンドで配置され、その結果、それらのメモリセルは、それぞれが共通ソース領域１４（source、S）を共有するメモリセルのペアとして形成され、メモリセルペアの隣接するセットはそれぞれ共通ドレイン領域１６（drain、D）を共有する。任意の所与のメモリセルの行のソース領域１４は全て、ソース線１４ａによって電氣的に接続される。任意の所与のメモリセルの列のドレイン領域１６は全て、ビット線１６ａによって電氣的に接続される。任意の所与のメモリセルの行の制御ゲート２２は全て、制御ゲート線２２ａによって電氣的に接続される。したがって、メモリセルを個々にプログラム及び読み出しすることができる一方、メモリセル消去は１行ずつ実行される（メモリセルの各行は、制御ゲート線２２ａへの高電圧の適用によって一緒に消去される）。特定のメモリセルを消去する場合は、同じ行にある全てのメモリセルもまた消去される。

【００６２】

当業者は、ソース及びドレインは互換可能であり得ることを理解しており、ここで浮遊ゲートは、図４（２ゲートメモリセル）に示されるように、ドレインの代わりにソースの上方に部分的に延在し得る。図５は、メモリセル１０、ソース線１４ａ、ビット線１６ａ、及び制御ゲート線２２ａを含む、対応するメモリセルアーキテクチャを最も良く示している。図から明らかなように、同じ行のメモリセル１０は、同じソース線１４ａ及び同じ制御ゲート線２２ａを共有し、一方同じ列の全てのセルのドレイン領域は、同じビット線１６ａに電氣的に接続されている。レイの設計は、デジタルアプリケーション用に最適化されており、例えば、１．６Ｖ及び７．６Ｖを選択した制御ゲート線２２ａ及びソース線１４ａにそれぞれ印加し、選択したビット線１６ａを接地することによるなど、選択したセルの個々のプログラミングを可能にする。同じペア内の選択されていないメモリセルを妨害することは、選択されていないビット線１６ａに２ボルト超の電圧を印加し、残りの線を接地することによって回避される。消去に関与しているプロセス（浮遊ゲート２０から制御ゲート２２までの電子のファウラーノルドハイムトンネリング）は、ドレイン電

10

20

30

40

50

圧（すなわち、同じソース線 1 4 a を共有する行方向で 2 つの隣接するセルにより異なり得る電圧のみ）に弱く影響を受けるだけであるので、メモリセル 1 0 を個々に消去することはできない。

【0063】

3 つ以上のゲートを有する分割ゲートメモリセルもまた既知である。例えば、図 6 に示されるように、4 ゲートのメモリセルは、ソース領域 1 4、ドレイン領域 1 6、チャネル領域 1 8 の第 1 の部分の上方の浮遊ゲート 2 0、チャネル領域 1 8 の第 2 の部分の上方の選択ゲート 2 8、浮遊ゲート 2 0 の上方の制御ゲート 2 2、及びソース領域 1 4 の上方の消去ゲート 3 0 を有し、それらは既知である（例えば、米国特許第 6, 747, 310 号を参照されたく、これは参照によりあらゆる目的で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てい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。

【0064】

別のタイプの先行技術の分割ゲート 3 ゲートメモリセルを図 3 4 に示す。図 3 4 の分割ゲートメモリセルは、別個の制御ゲートを有さないことを除いて、図 6 の分割ゲートメモリセルと同一である。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 6 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。

（0001）

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に 4 つの端子に印加され得る典型的な電圧範囲を示す。

表 1：図 3 4 のフラッシュメモリセルの動作

【表 1】

	WL	WL－ 非選択	BL	BL－ 非選択	EG	EG－ 非選択	SL	SL－ 非選択
読み出し	0.7～ 2.2V	－0.5V／ 0V	0.6～ 2V	0V／ FLT	0～ 2.6V	0～ 2.6V	0V	0V／ FLT／VB
消去	－0.5V／ 0V	－.5V／ 0V	0V	0V	11.5V	0～ 2.6V	0V	0V
プログラム	1～ 1.5V	－.5V／ 0V	1～ 3μA	Vinh (約1.8V)	4.5V	0～ 2.6V	7～ 9V	0～1V／ FLT

【0065】

4 ゲートのメモリセルアレイのアーキテクチャは、図 7 に示されるように構成され得る。この実施形態では、それぞれの水平の選択ゲート線 2 8 a は、メモリセルのその行の選択ゲート 2 8 全てを電氣的に一緒に接続する。それぞれの水平の制御ゲート線 2 2 a は、メモリセルのその行の制御ゲート 2 2 全てを電氣的に一緒に接続する。それぞれの水平のソース線 1 4 a は、ソース領域 1 4 を共有するメモリセルの 2 つの行に対するソース領域 1 4 全てを電氣的に一緒に接続する。それぞれのビット線 1 6 a は、メモリセルのその列のドレイン領域 1 6 全てを電氣的に一緒に接続する。それぞれの消去ゲート線 3 0 a は、消去ゲート 3 0 を共有するメモリセルの 2 つの行に対する消去ゲート 3 0 全てを電氣的に一緒に接続する。以前のアーキテクチャと同様に、個々のメモリセルは、独立してプログラム及び読み出しを行うことができる。しかしながら、セルを個々に消去する方法はない。消去は、消去ゲート線 3 0 a に高圧正電圧を印加することによって実行され、このことは同じ消去ゲート線 3 0 a を共有するメモリセルの両方の行の同時消去をもたらす。例示的な動作電圧は、下の表 2 内のものを含み得る（この実施形態では、選択ゲート線 2 8 a

はワード線（word line、WL）と呼ばれることがある）：

表 2

【表 2】

	WL		BL		SL		CG		EG	
	選択	非選択	選択	非選択	選択	非選択	選択	非選択	選択	非選択
消去	0V	0V	0V	0V	0V	0V	0V	0V	11.5V	0V
読み出し	2.5V	0V	0.8V	0V	0V	0V	2.5V	2.5V	0V	0V
プログラム	1V	0V	1 μ A	2.5V	4.5V	0.5V	10.5V	0/2.5V	4.5V	0.5V

【0066】

上記の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に線を再構成する。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。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、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、セル格納がアナログであるか、又は多数の別個の値（16又は32の異なる値など）のうちの1つを最低限格納することができ、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格納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。

メモリセルのプログラミング及び格納

【0067】

メモリセル内に格納されたニューラルネットワークの重みレベルの割り当ては、図8Aに示されるような均等な間隔、又は図8Bに示されるような不均等な間隔であり得る。不揮発性メモリセルのプログラミングは、図9に示されるものなどの双方向調整アルゴリズムを使用して実装され得る。I_{cell}は、プログラムされる標的セルの読み出し電流であり、I_{target}は、セルが理想的にプログラムされる際の所望の読み出し電流である。標的セル読み出し電流I_{cell}を読み出し（工程1）、標的読み出し電流I_{target}と比較する（工程2）。標的セル読み出し電流I_{cell}が、標的読み出し電流I_{target}より大きい場合は、プログラミング調整プロセスを実行して（工程3）浮遊ゲート（ルックアップテーブルを使用して制御ゲートの所望のプログラミング電圧V_{CG}を決定する）の電子数を増大させ（工程3a～3b）、必要に応じてこれを繰り返すことができる（工程3c）。標的セル読み出し電流I_{cell}が、標的読み出し電流I_{target}より小さい場合は、消去調整プロセスを実行して（工程4）浮遊ゲート（ルックアップテーブルを使用して消去ゲートの所望の消去電圧V_{EG}を決定する）の電子数を減少させ（工程4a～4b）、必要に応じてこれを繰り返すことができる（工程4c）。プログラミング調整プロセスが標的読み出し電流を通り過ぎると、（許容できるデルタ値内で）標的読み出し電流を達成するまで次に消去調整プロセスを実行し（工程3d及び工程4aで始まる）、逆もまた同様である（工程4d及び工程3aで始まる）。

【0068】

不揮発性メモリセルのプログラミングは、プログラミング調整を使用して一方向の調整アルゴリズムを使用して代わりに実装され得る。このアルゴリズムを使用して、メモリセルが最初に完全に消去され、次に標的セルの読み出し電流が標的閾値に達するまで図9のプログラミング調整工程3a～3cが実行される。あるいは、不揮発性メモリセルの調整は、消去調整を使用して一方向の調整アルゴリズムを使用して実装され得る。このアプローチでは、メモリセルが最初に完全にプログラムされ、次に標的セルの読み出し電流が標的閾値に達するまで図9の消去調整工程4a～4cが実行される。

【0069】

図10は、電流比較を使用した重みマッピングを示す図である。重みデジタルビット（例えば、各シナプシスに対して5ビットの重み、メモリセルの標的デジタル重みを表す）は、デジタルアナログ変換器（digital-to-analog converter、DAC）40に入力され、DAC40はビットを電圧 V_{out} （例えば、64電圧レベル-5ビット）に変換する。 V_{out} は、電圧電流変換器 V/I Conv42によって電流 I_{out} （例えば、64電流レベル-5ビット）に変換される。電流は、電流比較器IComp44に供給される。プログラム又は消去アルゴリズムの有効化は、メモリセル10に入力される（例えば、消去：EG電圧をインクリメント、又はプログラム：CG電圧をインクリメント）。外へ出るメモリセル電流 $I_{cellout}$ （すなわち、読み出し動作から）は、電流比較器IComp44に供給される。電流比較器IComp44は、メモリセル電流 $I_{cellout}$ を重みデジタルビット由来の電流 I_{out} と比較してメモリセル10に格納された重みを示す信号を生成する。

【0070】

図11は、電圧比較を使用した重みマッピングを示す図である。重みデジタルビット（例えば、各シナプシスに対して5ビットの重み）は、デジタルアナログ変換器（DAC）40に入力され、DAC40はビットを電圧 V_{out} （例えば、64電圧レベル-5ビット）に変換する。 V_{out} は、電圧比較器VComp46に供給される。プログラム又は消去アルゴリズムの有効化は、メモリセル10に入力される（例えば、消去：EG電圧をインクリメント、又はプログラム：CG電圧をインクリメント）。外へ出るメモリセル電流 $I_{cellout}$ は、電圧 V_{2out} への変換（例えば、64電圧レベル-5ビット）のために電流電圧変換器 I/V Conv48に供給される。電圧 V_{2out} は、電圧比較器VComp46に供給される。電圧比較器VComp46は、電圧 V_{out} 及び V_{2out} を比較してメモリセル10に格納された重みを示す信号を生成する。

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

【0071】

図12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例を概念的に図示する。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。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CB1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、層C1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、C1の異なる特徴マップを生成する。

【0072】

C1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

10

20

30

40

50

プスC B 1は、2次元アレの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、アレは必ずしも物理的な2次元アレ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。

【0073】

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ルの活性化関数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）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。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ルの活性化関数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。

【0074】

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の一部を使用して実行される。図13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とマトリクスとの乗算（vector-by-matrix multiplication、VMM）アレのブロック図である。具体的には、VMM 32は、不揮発性メモリセルのアレイ33、消去ゲート及びワード線ゲートデコーダ34、制御ゲートデコーダ35、ビット線デコーダ36、並びにソース線デコーダ37を含み、それらのデコーダはメモリアレイ33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37はまた、メモリセルアレの出力をデコードする。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格納する。第2に、メモリアレイは、効果的に、入力に、メモリアレイに格納された重みを乗じて出力を生成し、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、その場でのメモリ計算のために電力効率も良い。

【0075】

メモリアレイの出力は、メモリセルアレの出力を合計してその量み込み用の単一の値を生成する、差動加算器（加算オペアンプなど）38に供給される。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（例えば上記の説明ではC 1）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

10

20

30

40

50

成する。

【0076】

図14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VMM
32aに提供される。入力VMM32aによって生成された出力は、次に、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次のVMM（隠しレベル
1）32bへの入力として提供される、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。

10

【0077】

図15は、ドレイン加算マトリクス乗算器として配置された4ゲートのメモリセル（す
なわち、図6に示されるものなど）のアレイを図示する。図15のアレイに関する様々な
ゲート線及び領域線は、消去ゲート線30aが、水平ではなく垂直に延びる（すなわち、
各消去ゲート線30aは、メモリセルのその列に関して全ての消去ゲート30と一緒に接
続する）ことを除いて、（対応する構造に同じ要素番号を用いて）図7中のものと同じで
あり、その結果、各メモリセル10は、独立してプログラム、消去、及び読み出しを行う
ことができる。そのセルに適切な重み値を使用してメモリセルのそれぞれがプログラムさ
れた後で、アレイは、ドレイン加算マトリクス乗算器として作動する。マトリクス入力は
、 $V_{in0} \dots V_{in7}$ であり、選択ゲート線28aに印加される。図15のアレイに
関して出力のマトリクス $I_{out0} \dots I_{outN}$ は、ビット線16aで生成される。
各出力 I_{out} は、列内の全てのセルについてセル電流 I に、セルに格納された重み W を
乗じた値の合計である。

20

$$I_{out} = \sum (I_{ij} * W_{ij})$$

【0078】

各メモリセル（又はメモリセルのペア）は、その列内のメモリセル（又はメモリセルの
ペア）に格納された重み値の合計によって決まる出力電流 I_{out} として表される重み値
を有する単一のシナプスとして作動する。任意の所与のシナプスの出力は、電流の形態で
ある。したがって、第1の段階後のそれぞれの後続のVMM段階は、好ましくは、前のV
MM段階からの入力電流を入力電圧 V_{in} として使用される電圧に変換するための回路機
構を含む。図16は、そのような電流電圧変換回路機構の例を図示しているが、これは入
力電流 $I_{in0} \dots I_{inN}$ を入力電圧 $V_{in0} \dots V_{inN}$ にログ変換する改変さ
れたメモリセルの行である。

30

【0079】

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。

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

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

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

40

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

ベクターマトリクス乗算器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：

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

$$I_{out} = (w_a / w_p) * I_{in} = W * I_{in}$$

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

【0080】

図17及び図18は、ドレイン加算マトリクス乗算器として配置された4ゲートのメモ
リセル（すなわち、図6に示されるものなど）のアレイの別の構成を図示する。図17及
び図18のアレイに関する線は、ソース線14aが水平ではなく垂直に延びる（すなわち

50

、各ソース線14aは、メモリセルのその列に関して全てのソース領域14と一緒に接続する)こと、及び消去ゲート線30aが垂直ではなく水平に延びる(すなわち、各消去ゲート線30aは、メモリセルペアのその行に関して全ての消去ゲート30aと一緒に接続する)ことを除いて図15及び図16のアレイ内のものと同じであり、その結果各メモリセルは、独立してプログラム、消去、及び読み出しを行うことができる。マトリクス入力 $V_{in0} \dots V_{inN}$ は、選択ゲート線28aのままであり、マトリクス出力 $I_{out0} \dots I_{outN}$ は、ビット線16aのままである。

【0081】

図19は、ゲートカップリング／ソース加算マトリクス乗算器として配置された4ゲートのメモリセル(すなわち、図6に示されるものなど)のアレイの別の構成を図示する。図19のアレイの線は、選択ゲート線28aが垂直に延び、メモリセルの各列についてそれらのうちの2つが存在することを除いて、図15及び図16中のものと同じである。具体的には、メモリセルの各列は、2つの選択ゲート線、すなわち奇数行メモリセルの全ての選択ゲート28と一緒に接続する第1の選択ゲート線28a1、及び偶数行メモリセルの全ての選択ゲート28と一緒に接続する第2の選択ゲート線28a2を含む。

【0082】

図19の頂部と底部において回路は、入力電流 $I_{in0} \dots I_{inN}$ を入力電圧 $V_{in0} \dots V_{inN}$ にログ変換するように働く。この図で示されるマトリクス入力は、 $V_{in0} \dots V_{in5}$ であり、選択ゲート線28a1及び28a2に印加される。具体的には、入力 V_{in0} は、列1内の奇数セルについて選択線28a1に印加される。 V_{in1} は、列1内の偶数セルについて選択ゲート線28a2に印加される。 V_{in2} は、列2内の奇数セルについて選択ゲート線28a1に印加される。 V_{in3} は、列2内の偶数セルについて選択ゲート線28a2に印加されるなど。マトリクス出力 $I_{out0} \dots I_{out3}$ は、ソース線14aに提供される。ビット線16aは、固定バイアス電圧 V_{BLRd} でバイアスされる。各出力 I_{out} は、メモリセルのその行内の全てのセルについてセル電流 I に、セルに格納された重み W を乗じた値の合計である。したがって、このアーキテクチャについて、メモリセルの各行は、その行内のメモリセルに格納された重み値の合計によって決まる出力電流 I_{out} として表される重み値を有する単一のシナプスとして作動する。

【0083】

図20は、ゲートカップリング／ソース加算マトリクス乗算器として配置された4ゲートのメモリセル(すなわち、図6に示されるものなど)のアレイの別の構成を図示する。図20のアレイの線は、ビット線16が垂直に延び、メモリセルの各列についてそれらのうちの2つが存在することを除いて、図19中のものと同じである。具体的には、メモリセルの各列は、2本のビット線、すなわち隣接する対のメモリセルの全てのドレイン領域と一緒に接続する第1のビット線16a1(同じビット線コンタクトを共有する2つのメモリセル)、及び次の隣接する対のメモリセルの全てのドレイン領域と一緒に接続する第2のビット線16a2を含む。マトリクス入力 $V_{in0} \dots V_{inN}$ は、選択ゲート線28a1及び28a2のままであり、マトリクス出力 $I_{out0} \dots I_{outN}$ は、ソース線14aのままである。全ての第1のビット線16a1のセットは、バイアスレベル(例えば、1.2V)でバイアスされており、全ての第2のビット線16a2のセットは、別のバイアスレベル(例えば、0V)でバイアスされている。ソース線14aは、実質上のバイアスレベル(例えば、0.6V)でバイアスされる。共通ソース線14aを共有するメモリセルの各ペアについて、出力電流は、上部セルから下部セルを差し引いた差動出力になる。したがって、各出力 I_{out} は、これらの差動出力の合計である。

$$I_{out} = \sum (I_{iju} * W_{iju} - I_{ijd} * W_{ijd})$$

$$SL \text{ 電圧 } \sim 1/2 \cdot V_{dd}, \sim 0.6V$$

したがって、このアーキテクチャについて、ペアのメモリセルの各行は、そのペアのメモリセルの行内のメモリセルに格納された重み値(例えば、1つの正の重みと1つの負の重み)によって決まる差動出力の合計である出力電流 I_{out} として表される重み値を有

する単一のシナプスとして作動する。

【0084】

図21は、ゲートカップリング／ソース加算マトリクス乗算器として配置された4ゲートのメモリセル（すなわち、図6に示されるものなど）のアレイの別の構成を図示する。図21のアレイの線は、消去ゲート30aが水平に延び、制御ゲート線22aが垂直に延び、メモリセルの各列についてそれらのうちの2つが存在することを除いて、図20中のものと同じである。具体的には、メモリセルの各列は、2つの制御ゲート線、すなわち奇数行メモリセルの全ての制御ゲート22aと一緒に接続する第1の制御ゲート線22a1、及び偶数行メモリセルの全ての制御ゲート22aと一緒に接続する第2の制御ゲート線22a2を含む。マトリクス入力 $V_{in0} \dots V_{inN}$ は、選択ゲート線28a1及び28a2のままであり、マトリクス出力 $I_{out0} \dots I_{outN}$ は、ソース線14aのままである。

10

【0085】

図22は、ソース加算マトリクス乗算器として配置された4ゲートのメモリセル（すなわち、図6に示されるものなど）のアレイの別の構成を図示する。図22のアレイの線及び入力、図17中のものと同じである。しかしながら、出力がビット線16aに提供される代わりに、それらはソース線14aに提供される。マトリクス入力 $V_{in0} \dots V_{inN}$ は、選択ゲート線28aのままである。

【0086】

図23は、ドレイン加算マトリクス乗算器として配置された2ゲートのメモリセル（すなわち、図1に示されるものなど）のアレイの構成を図示する。図23のアレイの線は、水平のソース線14aが垂直のソース線14aに置き換えられたことを除いて、図5中のものと同じである。具体的には、各ソース線14aは、メモリセルのその列のソース領域全てに接続されている。マトリクス入力 $V_{in0} \dots V_{inN}$ は、制御ゲート線22aに印加される。マトリクス出力 $I_{out0} \dots I_{outN}$ は、ビット線16aに生成される。各出力 I_{out} は、列内の全てのセルについてセル電流 I に、セルに格納された重み W を乗じた値の合計である。メモリセルの各列は、その列内のメモリセルに格納された重み値の合計によって決まる出力電流 I_{out} として表される重み値を有する単一のシナプスとして作動する。

20

【0087】

図24は、ソース加算マトリクス乗算器として配置された2ゲートのメモリセル（すなわち、図1に示されるものなど）のアレイの構成を図示する。図24のアレイの線は、制御ゲート線22aが垂直に延び、メモリセルの各列についてそれらのうちの2つが存在することを除いて、図5中のものと同じである。具体的には、メモリセルの各列は、2つの制御ゲート線、すなわち奇数行メモリセルの全ての制御ゲート22aと一緒に接続する第1の制御ゲート線22a1、及び偶数行メモリセルの全ての制御ゲート22aと一緒に接続する第2の制御ゲート線22a2を含む。

30

【0088】

この構成用のマトリクス入力は、 $V_{in0} \dots V_{inN}$ であり、制御ゲート線22a1及び22a2に印加される。具体的には、入力 V_{in0} は、列1内の奇数行セルについて制御ゲート線22a1に印加される。 V_{in1} は、列1内の偶数行セルについて制御ゲート線22a2に印加される。 V_{in2} は、列2内の奇数行セルについて制御ゲート線22a1に印加される。 V_{in3} は、列2内の偶数行セルについて選択ゲート線22a2に印加されるなど。マトリクス出力 $I_{out0} \dots I_{outN}$ は、ソース線14aに生成される。共通ソース線14aを共有するメモリセルの各ペアについて、出力電流は、上部セルから下部セルを差し引いた差動出力になる。したがって、このアーキテクチャについて、ペアのメモリセルの各行は、そのペアのメモリセルの行内のメモリセルに格納された重み値によって決まる差動出力の合計である出力電流 I_{out} として表される重み値を有する単一のシナプスとして作動する。

40

【0089】

50

図15～図16、図19及び図20の実施形態に関する例示の動作電圧は以下を含む。
【表3】

	EG		WL		CG		BL		SL	
	選択	非選択	選択	非選択	選択	非選択	選択	非選択	選択	非選択
消去	VEGerase	gnd	gnd	gnd	gnd	VCGerinhibit	gnd	gnd	gnd	gnd
プログラム	VEGprg/ gnd	gnd	VWLprg	gnd	VCGprg	gnd	lprog	VBLprginh	VSprg	gnd
読み出し	gnd	gnd	VWLrd	gnd	VCGrd	gnd	VBLrd	gnd	VSrd	float/ gnd

おおよその数値として以下が挙げられる。

【表4】

VEGerase	8～11.5v
VCGerinhibit	3.5～8v
VEGprg	4～6v
VWLprg	0.8～1.2v
VCGprg	6～10v
VBLprginh	1～2.5v
lprog	0.2～1μa
VSprg	3～5V
VWLrd	0.4～2.0V
VCGrd	0～2.5V
VBLrd	1～2V
VSrd	0～0.6V

【0090】

図17～図18及び図22の実施形態に関する例示の動作電圧は以下を含む。

【表5】

	EG		WL		CG			BL		SL	
	選択	非選択	選択	非選択	選択	非選択－ 共有EG	非選択	選択	非選択	選択	非選択
消去	VEGerase	gnd	gnd	gnd	gnd	VCGerinh	gnd	gnd	VBLerinh/ gnd	VSlerinh	gnd
プログラミング	VEGprg/ gnd	gnd	VWLprg	gnd	VCGprg/ gnd	gnd	gnd	lprog	VBLprginh	VSprg	gnd
読み出し	gnd	gnd	VWLrd	gnd	VCGrd	gnd	gnd	VBLrd	gnd	VSrd	float/gnd

おおよその数値として以下が挙げられる。

10

20

30

40

50

【表 6】

VEGerase	7~10v
VSLeinh	3.5~6v
VCGerinh	3.5~7v
VBLeinh	1~2.5v
VEGprg	4~6v
VWLprg	0.8~1.2v
VCGprg	6~10v
VBLeprinh	1~2.5v
Iprog	0.2~1 μ a
VSprg	3~5V
VWLrd	0.4~2.0V
VCGrd	1~2.5V
VBLeird	1~2V
VSrd	0~0.6V

10

【0091】

20

図25は、本発明で使用するための例示の電流電圧ログ変換器50を図示する（WL＝選択ゲート線、CG＝制御ゲート線、EG＝消去ゲート線）。メモリは、弱反転領域にバイアスされる、 $I_{ds} = I_0 \cdot e^{(V_g - V_{th})/kV_t}$ 。図26は、本発明で使用するための例示の電圧電流ログ変換器52を図示する。メモリは、弱反転領域にバイアスされる。図27は、本発明で使用するための接地基準電流加算器54を図示する。図28は、本発明で使用するためのV_{dd}基準電流加算器56を下に図示する。負荷の例として、ダイオード、不揮発性メモリセル、及びレジスタが挙げられる。

【0092】

上述のメモリアレイ構成は、フィードフォワード分類エンジンを実装する。トレーニングは、メモリセル内に「重み」値を格納することによって完了し（シナプスアレイの生成）、このことは、個々のセルのサブスレッショルドスロープファクタが変更されたことを意味する。ニューロンは、シナプスの出力を加算すること及びニューロン閾値に応じて始動させる又は始動させないことによって実行される（すなわち、決定を下す）。

30

【0093】

以下の工程を使用して、入力電流I_Eを処理することができる（例えば、入力電流は、画像認識のための特徴計算の出力から直接来ている）：

工程1－不揮発性メモリで、より簡単に処理するためにログスケールに変換する。

・バイポーラトランジスタを使用した入力電流電圧変換。バイポーラトランジスタのバイアス電圧V_{BE}は、エミッタ電流と対数関係を有する。

$$V_{BE} = a \cdot \ln I_E - b \rightarrow V_{BE} \propto \ln I_E$$

40

－ 式中、a（比）及びb（バイアス又はオフセット）は定数である

・V_{BE}電圧は、メモリセルがサブスレッショルド領域で操作されるように発生する。

工程2－発生したバイアス電圧V_{BE}をワード線（サブスレッショルド領域内の）に印加する。

・CMOSトランジスタの出力電流I_{DRAIN}は、入力電圧（V_{GS}）、熱電圧（U_T）及びカップ（ $k = C_{ox} / (C_{ox} + C_{dep})$ ）と指数関数的関係を有し、式中、C_{ox}及びC_{dep}は、浮遊ゲートの電荷に線形従属である。

$$I_{DRAIN} \propto \exp(k V_{BE} / U_T), \text{又は}$$

$$\ln I_{DRAIN} \propto k V_{BE} / U_T$$

・I_{DRAIN}の対数は、複数のV_{BE}及び浮遊ゲートの電荷（カップに関連した）と線形関

50

係を有し、式中、 U_T は、所与の温度で一定である。

・出力＝入力×重みの関係がシナプスに対して存在する。

【0094】

セルのそれぞれの出力 (I_{DRAIN}) は、読み出しモードにおいて結び付けられて、アレイ又はアレイのセクタ内の各シナプスの値を合計し得る。 I_{DRAIN} は、ひとたび合計されると、電流比較器に送り込まれ、単一知覚ニューラルネットワークの比較に応じて「論理」0又は1を出力することができる。1つの知覚(1つのセクタ)は上述されている。各知覚からの出力は、複数の知覚のセクタの次のセットに送り込まれ得る。

【0095】

メモリベースの畳み込みニューラルネットワークにおいて、入力の設定は、特定の重みを乗じて、隠し層又は出力層用の所望の結果を生成する必要がある。上述のように、1つの技術は、先行する画像をスキャンすることである(例えば、水平及び垂直の方向の両方で画像全体にXピクセルずつシフトされる $M \times M$ フィルタ(カーネル))を使用した $N \times N$ マトリクス。ピクセルのスキャンは、メモリアレイに対する十分な入力が存在する限り、同時に少なくとも部分的に実行され得る。例えば、図29に示されるように、 $M=6$ のフィルタサイズ(すなわち、36ピクセルの 6×6 アレイ)を使用して、 $N \times N$ 画像アレイを $X=2$ のシフトを用いてスキャンすることができる。その例において、フィルタ内の6ピクセルの第1の行は、 N^2 個の入力のメモリアレイへの入力のうち最初の6個に提供される。次に、フィルタ内の6ピクセルの第2の行は、 N^2 個の入力の第2のN個の入力のうち入力の最初の6個に提供されるなど。これは、図29の図の第1の行において表現されており、ここで、点は、上に示したように入力による乗算のためのメモリアレイ内に格納された重みを表す。次に、フィルタは、2ピクセルずつ右側にシフトされ、シフトされたフィルタ内の6ピクセルの第1の行は、第1のN個の入力の第3から第8の入力に提供され、6ピクセルの第2の行は、第2のN個の入力の第3から第8の入力に提供されるなど。ひとたびフィルタが画像の右側までシフトされると、フィルタは、左側に戻って再配置されるが、2ピクセルずつ下にシフトされ、ここでプロセスは、 $N \times N$ 画像全体がスキャンされるまで再び繰り返す。水平にシフトされたスキャンの各セットは、どの N^2 個のメモリアレイの入力が、乗算用のデータを有するかを示す台形により表現され得る。

【0096】

したがって、スキャン間の2ピクセルのシフト及び 6×6 のフィルタサイズを使用した $N \times N$ 画像アレイのスキャンは、 N^2 個の入力及び $((N-4)/2)^2$ 個の行を必要とする。図30は、メモリアレイ内の重みがどのようにフィルタスキャンのために格納されるかを示す台形をグラフィカルに示す。網掛け領域の各行は、水平のスキャンのうち1セットの間に入力に適用される重みを表現する。矢印は、メモリアレイの直線の入力線を示す(例えば、入力データを受信する図15中の入力線28aは、直線的にメモリアレイにわたってずっと延在し、各1つが常に同じ行のメモリセルにアクセスし、図19のアレイの場合、各入力線は、常に同じ列のメモリセルにアクセスする)。白い領域は、入力に供給されるデータがない場所を示す。したがって、白い領域は、メモリセルアレイの非効率的な使用を示す。

【0097】

図31に示されるようにメモリアレイを再構成することによって、効率を高め、入力の総数を低減することができる。具体的には、メモリアレイの入力線は、別の行又は列に周期的にシフトされ、したがってアレイの未使用の部分を低減し、したがってスキャンを実行するのに必要なアレイに繰り返される入力線の数低減する。具体的には、シフト $X=2$ である本例の場合では、矢印は、2行又は2列ずつ各入力線が周期的にシフトし、広く離間したメモリセル利用の台形を密集したメモリセル利用の矩形に変形することを示す。メモリセル部分の間の余分な空間は、このシフトを実行する電線束に必要とされるが、メモリセルアレイに必要な入力数は、大きく低減される($5n+6$ のみ)。

【0098】

図32は、図15のアレイであるが、入力線として使用される線28aに対する2行の

周期的なシフトを伴うアレイを図示する。入力線に対する行の周期的なシフトは、図 1 7、図 2 2、及び図 2 3 のアレイで同様に実行され得る。図 3 3 は、図 2 0 のアレイであるが、入力線として使用される線 2 8 a 1 及び 2 8 a 2 に対する 2 列の周期的なシフトを伴うアレイを図示する。入力線に対する列の周期的なシフトは、図 1 9、図 2 1、及び図 2 4 のアレイで同様に実行され得る。

改善された調整機構及びアルゴリズム

【0099】

改善された調整機構及びアルゴリズムのための実施形態についてここで説明する。調整は、所望の量の電荷が不揮発性メモリセルの浮遊ゲートに格納されていることを検証するプロセスであり、すなわち、不揮発性メモリセルが所望の値を格納していることを確実にするためのプロセスである。

【0100】

図 3 5 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 VMM 3 5 0 0 を示し、入力層と次の層との間のシナプスとして利用される。VMM 3 5 0 0 は、不揮発性メモリセルのメモリアレイ 3 5 0 3、基準アレイ 3 5 0 1、及び基準アレイ 3 5 0 2 を備える。基準アレイ 3 5 0 1 及び 3 5 0 2 は、端子 B L R 0 - 3 内に流れる電流入力を電圧入力 W L 0 - 3 に変換する役割を果たす。実際には、基準メモリセルは、マルチプレクサを通してそれらの中に流れる電流入力にダイオード接続されている。基準セルは、標的基準レベルに調整される。メモリアレイ 3 5 0 3 は、2 つの目的を果たす。第 1 に、VMM 3 5 0 0 により使用される重みを格納する。第 2 に、メモリアレイ 3 5 0 3 は、メモリアレイに格納された重みによって、入力（端子 B L R 0 - 3 に提供される電流入力であって、基準アレイ 3 5 0 1 及び 3 5 0 2 がこれらの電流入力を入力電圧に変換してワード線 W L 0 - 3 に供給する）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。

【0101】

図 3 6 は、VMM 3 5 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る電圧を示す。行は、読み出し、消去、及びプログラムの動作を示す。

【0102】

図 3 7 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 VMM 3 7 0 0 を示し、入力層と次の層との間のシナプスとして利用される。VMM 3 7 0 0 は、不揮発性メモリセルのメモリアレイ 3 7 0 3、基準アレイ 3 7 0 1、及び基準アレイ 3 7 0 2 を備える。VMM 3 7 0 0 は、VMM 3 7 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、VMM 3 5 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。

【0103】

図 3 8 は、VMM 3 7 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る電圧を示す。行は、読み出し、消去、及びプログラムの動作を示す。

【0104】

図 3 9 は、図 6 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 VMM 3 9 0 0 を示し、入力層と次の層との間のシナプスとして利用される。VMM 3 9 0 0 は、不揮発性メモリセルのメモリアレイ 3 9 0 3、基準アレイ 3 9 0 1、及び基準アレイ 3 9 0 2 を備える。基準アレイ 3 9 0 1 及び 3 9 0 2 は、端子 B L R 0 - 3 内に流れる電流入力を電

圧入力CGO-3に変換する役割を果たす。実際には、基準メモリセルは、マルチプレクサを通してそれらの中に流れる電流入力にダイオード接続されている。基準セルは、標的基準レベルに調整される。メモリアレイ3903は、2つの目的を果たす。第1に、VMM3900により使用される重みを格納する。第2に、メモリアレイ3903は、メモリアレイに格納された重みによって、入力（端子BLRO-3に提供される電流入力であって、基準アレイ3901及び3902がこれらの電流入力を入力電圧に変換して制御ゲートCGO-3に供給する）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、入力はワード線に提供され、出力は読み出し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。

10

【0105】

VMM3900は、メモリアレイ3903内のメモリセルの一方向調整を実装する。すなわち、各セルは消去され、次いで、浮遊ゲートの所望の電荷に達するまで部分的にプログラムされる。過度に多くの電荷が浮遊ゲートにある場合（誤った値がセル内に格納されるなど）、セルは消去されなければならない、一連の部分的なプログラミング動作をやり直さなければならない。示されるように、同じ消去ゲートを共有する2つの行は、一緒に消去される必要がある（ページ消去として知られる）、その後、各セルは、浮遊ゲートの所望の電荷に達するまで部分的にプログラムされる。

【0106】

図40は、VMM3900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に印加される電圧を示す。行は、読み出し、消去、及びプログラムの動作を示す。

20

【0107】

図41は、図6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VMM4100を示し、入力層と次の層との間のシナプスとして利用される。VMM4100は、不揮発性メモリセルのメモリアレイ4103、基準アレイ4101、及び基準アレイ4102を備える。VMM4100は、VMM4100が双方向調整を実装することを除いてVMM3900と同様であり、各個々のセルは、浮動ゲートの所望の電荷量に達するために、完全に消去され、部分的にプログラムされ、必要に応じて部分的に消去され得る。示されるように、基準アレイ4101及び4102は、端子BLRO-3内の入力電流を制御ゲート電圧CGO-3に変換し（マルチプレクサを介したダイオード接続された基準セルの作用を通じて）、行方向でメモリセルに印加される。電流出力（ニューロン）は、ビット線に接続されたメモリセルからの全ての電流を合計するビット線にある。

30

【0108】

図42は、VMM4100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に印加される電圧を示す。行は、読み出し、消去、及びプログラムの動作を示す。

40

【0109】

図43は、調整アルゴリズム4300を示す。調整アルゴリズム4300は、個々のメモリセルが高速又は低速であり得ることを認識する。高速セルは迅速にプログラムされ得るものであり、低速セル又はノーマルセルは、高速セルと比較してプログラム動作中に同じ状態を達成するために、より大量の時間を要するものである。この差は、個々のセル特性及び分散の物理的性質によるものである。

50

【0110】

調整アルゴリズム4300は、以下の一連の工程を含む。最初に、メモリセルのページが消去される（工程4301）。次いで、システムは、構成シーケンス中に事前設定されたルックアップテーブル4320に基づいて、メモリアレイ内のどのセルが高速セルであるかを判定し、セルの特性が変化する場合（高速から低速、又は低速から高速に）、動作中に更新され得る（工程4302）。ルックアップテーブル4320は、例えば、全ての高速不揮発性メモリセルのアドレスのリストを含み得る。あるいは、アレイ内の各セルに対するエントリを含んでもよく、エントリは、「1」が高速セルを示し、「0」が低速ノーマルセルを示す単一のビットであってもよい。

【0111】

セルが高速セルである場合、高速調整アルゴリズムが実装され、この場合、比較的大きな電荷が、部分的プログラミング動作を通じて高速セルの浮遊ゲートに追加される（工程4303）。各部分的プログラミング動作の後、検証シーケンスが実行され、読み出し動作におけるセルを通る I_{cell} が、 $I_{target1}$ よりも大きいかどうかを決定する（工程4304）。いいえの場合、部分的プログラミング動作が再び実行される。はいの場合、 $I_{cell} < I_{margin_0V}$ であるかどうか判定される（工程4305）。はいの場合、所望の状態がメモリセル内で達成され、調整シーケンスが完了する（工程4306）。そうでない場合、セルは、意図したよりも速くプログラムされており、ルックアップテーブル4320において高速セルとしてマークされる（工程4307）。あまりにも多くの電荷が浮遊ゲートにあるため、セルは使用されず、再度消去されなければならない（工程4302）。

【0112】

工程4302の結論が、セルが低速セルであるということである場合、低調整アルゴリズムが実装され、部分的プログラミング動作を通じて、より小さい電荷が低速セルの浮遊ゲートに追加される（工程4308）。各部分的プログラミング動作の後、検証シーケンスが実行され、読み出し動作におけるセルを通る I_{cell} が、 $I_{target1}$ よりも大きいかどうかを決定する（工程4309）。いいえの場合、部分的プログラミング動作が再び実行される。はいの場合、 $I_{cell} < I_{margin_0V}$ であるかどうか判定される（工程4310）。はいの場合、所望の状態がメモリセル内で達成され、調整シーケンスが完了する（工程4311）。そうでない場合、セルは、意図されるよりも速くプログラムされており、高速セルとしてルックアップテーブル4320にマークされる（工程4307）。あまりにも多くの電荷が浮遊ゲートにあるため、セルは使用されず、再度消去されなければならない（工程4302）。高速調整アルゴリズムは、大きな書き込み（例えば、プログラム）電圧増分又は広い書き込みパルス幅で実装することができ、低調整アルゴリズムは、小さい書き込み電圧増分又は狭い書き込みパルス幅で実装することができる。

【0113】

図44は、調整アルゴリズム4400を示す。調整アルゴリズム4400は、高速セルであるメモリアレイ内のセルを識別するために、構成シーケンス中に使用することができる。調整アルゴリズム4400は、以下の一連の工程を含む。ページが消去される（工程4401）。電圧 $V_{CG_diagnosis}$ でセルをプログラムする（工程4402）。電流、 I_{cell1} 、が、読み出し動作中にセルを通して測定される（工程4403）。セルは、電圧 $V_{CG_diagnosis} + dV$ でプログラムされる（工程4404）。電流、 I_{cell2} 、が、読み出し動作中にセルを通して測定される（工程4405）。 I_{cell2} と I_{cell1} との間の差が I_{cell_0V} を超えるかどうかを決定する（工程4406）。いいえの場合、セルは、ノーマル又は低速セルである（工程4407）。はいの場合、セルは高速セルであり、ルックアップテーブル4320において高速セルとして識別される（工程4408）。図44に含まれるグラフに示されるように、高速セルは、工程4404に回答して、その浮遊ゲートに電荷を蓄積する。増分プログラミング電圧 d は、小さい電圧であり、ノーマル又は低速のセルの蓄積電荷に影響を与えない

10

20

30

40

50

。すなわち、高速セルは、比較的小さいプログラミング電圧に応答して、それらの浮遊ゲートに電荷を蓄積する。あるいは、2つの消去パルスを使用して、高速ビット領域に入るメモリセルのデルタ電流を抽出することができる。この場合、全てのセルが、最初に深くプログラムされる。

【0114】

図45は、調整アルゴリズム4500を示す。調整アルゴリズム4500は、粗いアルゴリズム4510（セルが比較的大きな増分でプログラムされる）と、微細なアルゴリズム4511（セルが比較的小さい増分でプログラムされる）とを組み合わせる。これにより、所望の電圧のオーバーシュートの発生が最小限に抑えられ、システムの全体的な速度が増加する。

【0115】

調整アルゴリズム4500は、以下の一連の工程を含む。ページが消去される（工程4501）。次いで、粗いアルゴリズム4510が実行され、これは工程4502及び工程4503を含む。セルが、 $V_{CG-C_init} + dV - C$ 、 $N_pulseC = NC_i + 1$ でプログラムされる（工程4502）。電流、 I_{cell} 、が、読み出し動作中にセルを通して測定され、 $I_{cell} > I_{cell_offset}$ であるかどうかについて判定がなされる（工程4503）。はいの場合、工程4504及び工程4505を含む微細アルゴリズム4511が始まる。セルが、 $V_{CGFinit} (= V_{CG-C_last} - V_{step}) + dV - F$ 、 $N_pulseF = NF_i + 1$ でプログラムされる（工程4504）。電流、 I_{cell} 、が、読み出し動作中にセルを通して測定され、 $I_{cell} > I_{cell_target}$ かどうかについて判定がなされる（工程4505）。はいの場合、所望の電荷が達成され、調整プロセスが完了する（工程4506）。いいえの場合、 $f \cdot N_pulseF = N_{max}$ であるかどうか判定される（工程4507）。はいの場合、セルは不良セルであると判定され、そのようなものとしてルックアップテーブルにマークされる（工程4508）。いいえの場合、工程4504が繰り返される。工程4503の結果がいいえの場合、 $N_pulseC = NC_{max}$ であるかどうか判定される（工程4509）。はいの場合、セルは不良セルであると判定され、そのようなものとしてルックアップテーブルにマークされる（工程4508）。いいえの場合、工程4502が繰り返される。あるいは、粗い電圧レベルの代わりに広い書き込みパルス幅を使用することができ、微細な電圧レベルの代わりに狭い書き込みパルス幅を使用することができる。

【0116】

図43及び図45の調整動作に関する追加の詳細が、図46～図50を参照してここで説明される。

【0117】

図46は、均一ステップアルゴリズム4600の例示的な値を示す。この実施例では、おおよそのプログラム標的電圧は8Vである。この標的電圧は、例えば、ルックアップテーブル又は標的プログラミング電圧に対する電流標的 I_{target} のモデルから抽出される。セルは、最初に4Vでプログラムされる。その後、微細なプログラミングは、最大ステップ数が400の、0.01Vのステップサイズでのプログラミングを含む。

【0118】

図47は、均一ログステップ（10で除算）、粗／微細アルゴリズム4700の例示的な値を示す。この実施例では、おおよそのプログラム標的電圧は8Vである。セルは、最初に4Vでプログラムされる。その後、粗いプログラミングは、0.4V（ $= 4V / 10$ ）のログステップサイズでのプログラミングを含み、最大のステップ数は10である。その後、第1の微細なプログラミングのシーケンスが行われ、浮遊ゲートの電圧を0.2Vだけ低減させるための部分消去動作が行われ、その後、0.1Vのステップサイズでのプログラミングが行われ、最大ステップ数は6である。次いで、第2の微細なプログラミングのシーケンスが行われ、0.1Vの部分消去動作が行われ、その後、0.01V（ $= 0.1V / 10$ ）のログステップサイズでのプログラミングが行われ、最大のステップ数は

10

20

30

40

50

20である。したがって、パルスの合計数は46である。

【0119】

図48は、バイナリサーチステップ、粗／微細アルゴリズム4800を示す。この実施例では、おおよそのプログラム標的電圧は、再び8Vである。セルは、最初に4Vでプログラムされる。その後、粗いプログラミングが、デルタプログラミング電圧が2で割られるステップで行われ、すなわち2V、1V、0.5V、0.25V、0.125V、及び0.625Vのステップで行われる。その後、浮動ゲートの電圧を0.0625Vだけ低減する消去ステップが行われ、その後、0.01Vの固定増分での微細なプログラミングが行われ、最大のステップ数は19である。パルスの合計数は25である。

【0120】

図49は、アルゴリズム4800の例示的な波形を示す。見て分かるように、粗いプログラミングは、Icellが閾値を超えるまで電圧を増加させた。次いで、電圧を0.0625Vだけ低減させるために電圧ステップダウンが実行され、その後、微細なプログラミングが行われる。ステップダウンは、次のプログラミングパルスの潜在的なプログラムオーバーシュートを回避するためある。

【0121】

図50は、本明細書に記載される粗／微細アルゴリズムの別の例示的な波形を示す。最初に粗いプログラミングが行われる。第2に、微細なプログラミングが行われる。第3に、適切な電圧が達成されると、プログラミング電荷の総量が記録され、及び／又は工程が記録され、その結果、その後、適切なプログラミングレベルは、所望の正確な電荷で単一の定電圧プログラミング動作で達成され得る。あるいは、調整アルゴリズムは、広いパルス幅及び／又は大きな電圧増分から構成される粗いパルス、並びに一定パルス幅及び／又は定電圧パルスから構成される微細パルスから構成することができる。

【0122】

代替の実施形態では、図15、図16、図17、図18、図19、図20、図21、図22、図25、図26、図32、及び図33の実施形態は、入力、選択ゲート（ワード線）の代わりに制御ゲートであるように改変され得る。同様に、図39及び図41の実施形態は、入力、制御ゲートの代わりに選択ゲート（ワード線）であるように改変され得る。

【0123】

本明細書で使用される場合、「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、「電気的に結合された」は、「に直接電気的に結合された」（要素と一緒に電気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「間接的に電気的に結合された」（要素と一緒に電気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、並びに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。

10

20

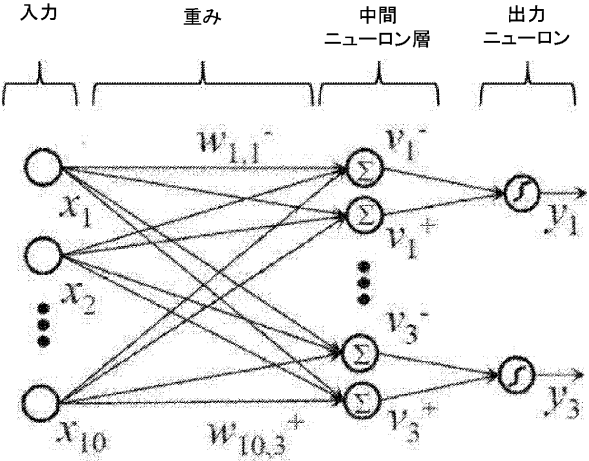
30

40

50

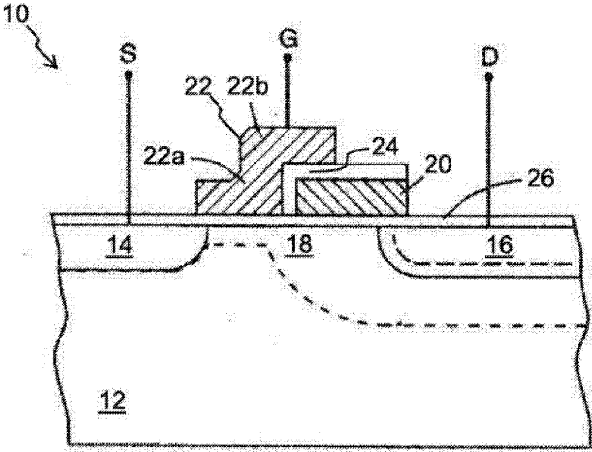
【図面】

【図 1】



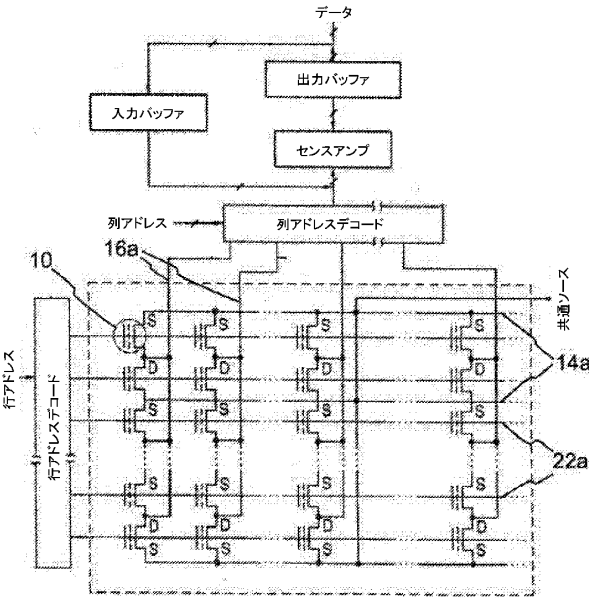
(先行技術)

【図 2】



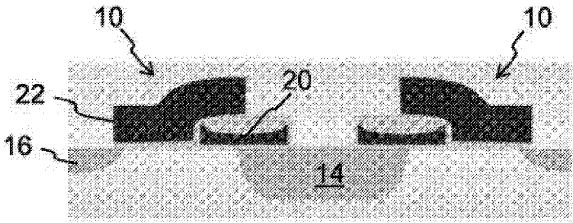
(先行技術)

【図 3】



(先行技術)

【図 4】



(先行技術)

10

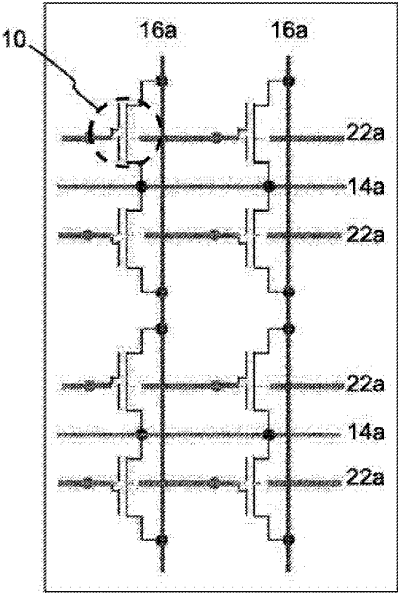
20

30

40

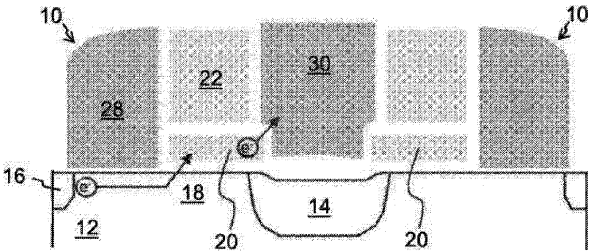
50

【図 5】



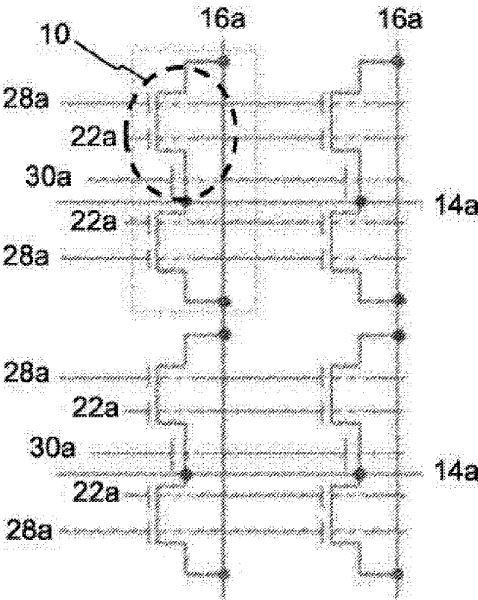
(先行技術)

【図 6】



(先行技術)

【図 7】



(先行技術)

【図 8 A】

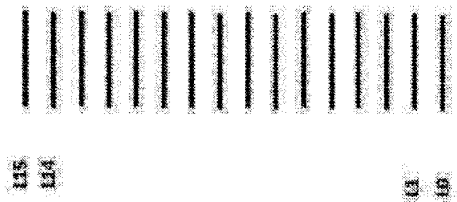


FIG. 8A

10

20

30

40

50

【図 8 B】

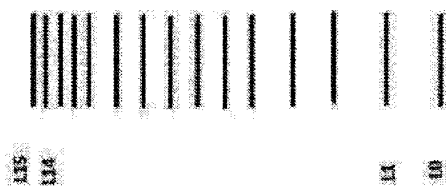
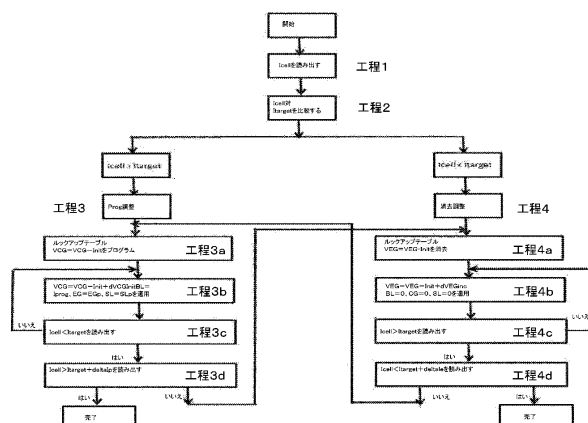
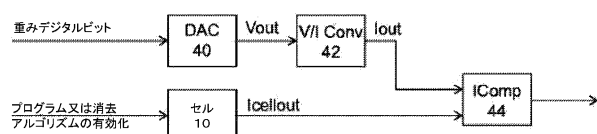


FIG. 8B

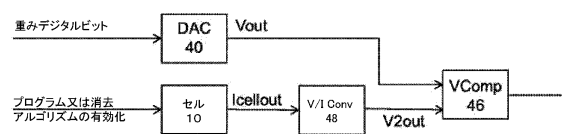
【圖 9】



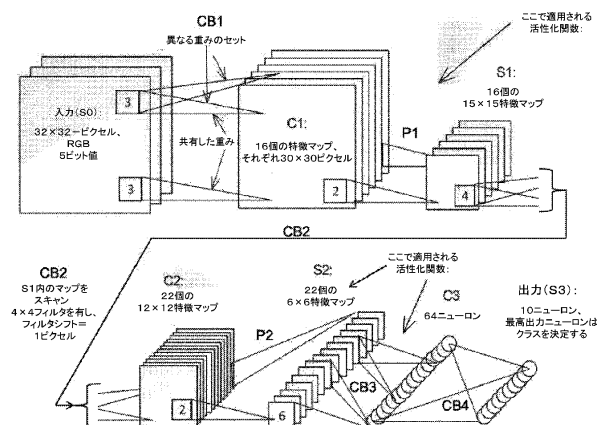
【図 10】



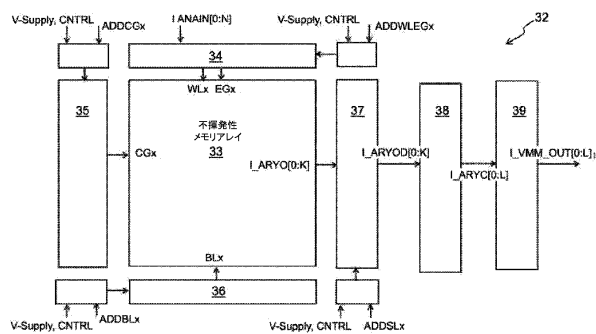
【図 11】



【図 12】



【図 13】



【図 1 4】

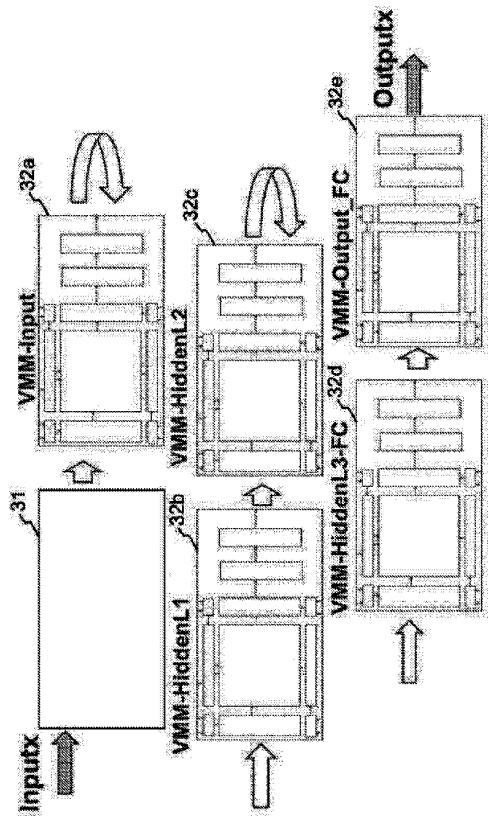


FIG. 14

【図 1 5】

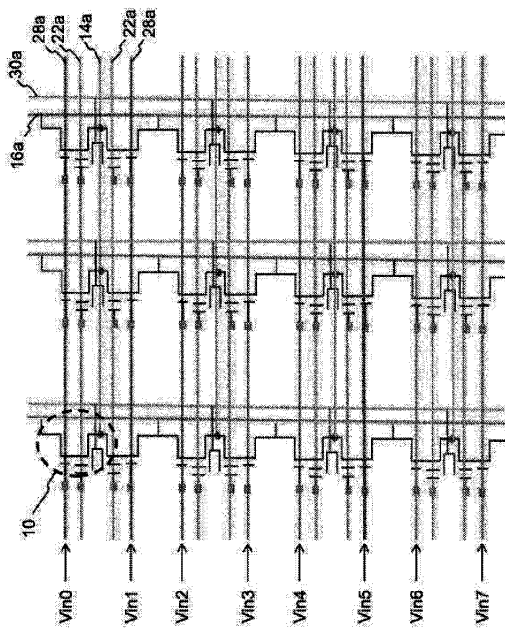


FIG. 15

【図 1 6】

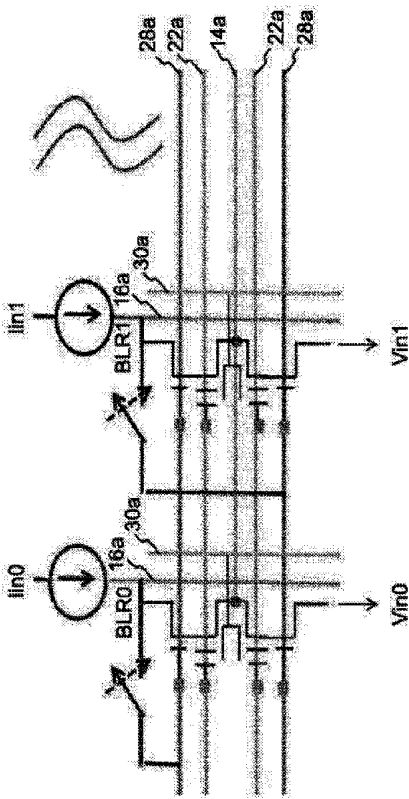


FIG. 16

【図 1 7】

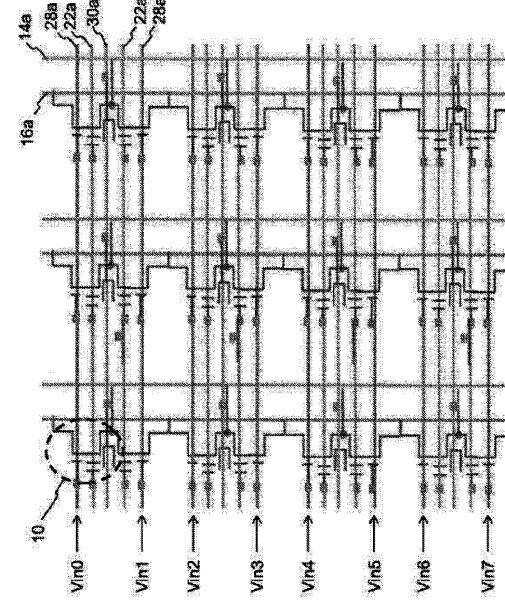


FIG. 17

10

20

30

40

50

【図 18】

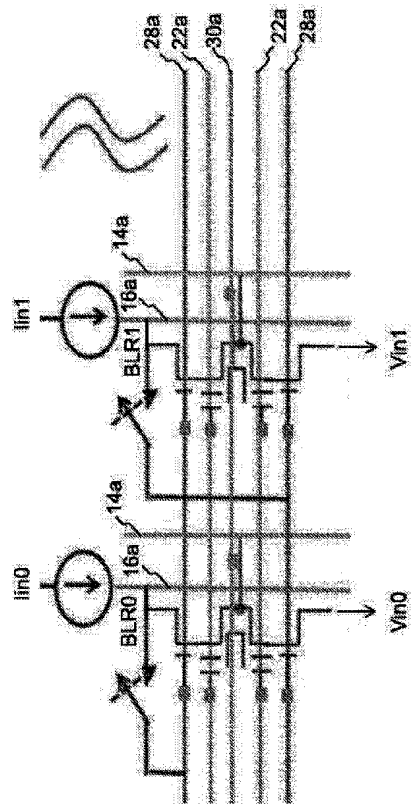


FIG. 18

【図 19】

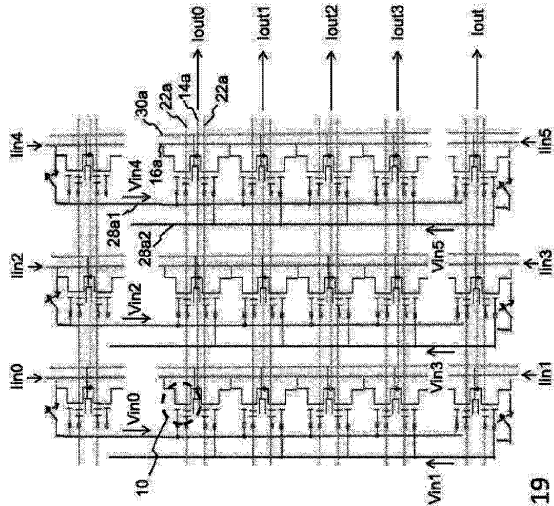


FIG. 19

【図 20】

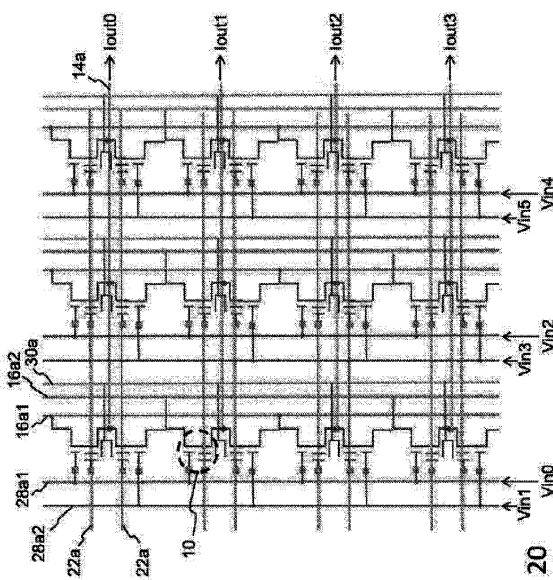


FIG. 20

【図 21】

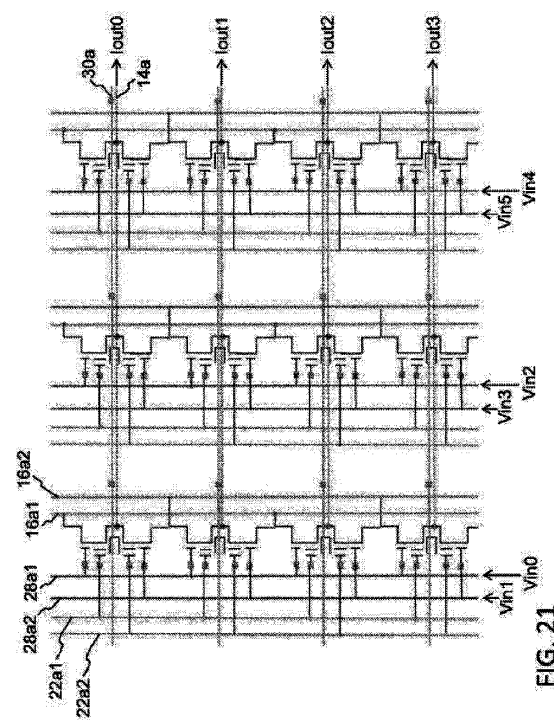


FIG. 21

10

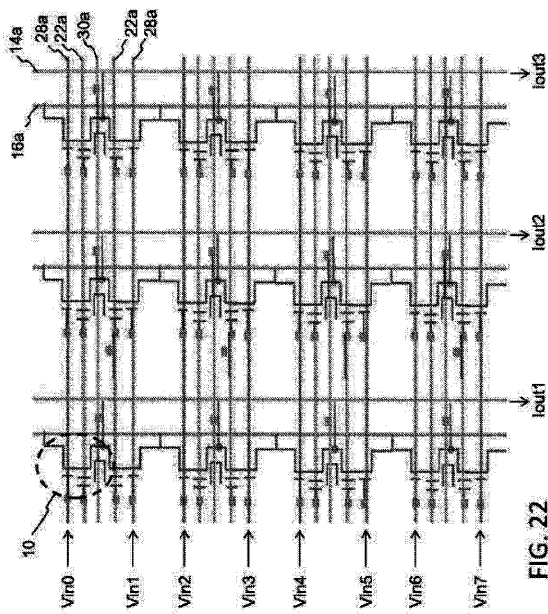
20

30

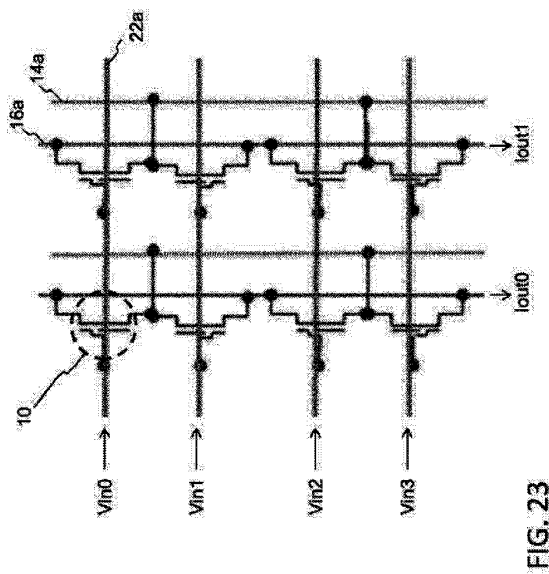
40

50

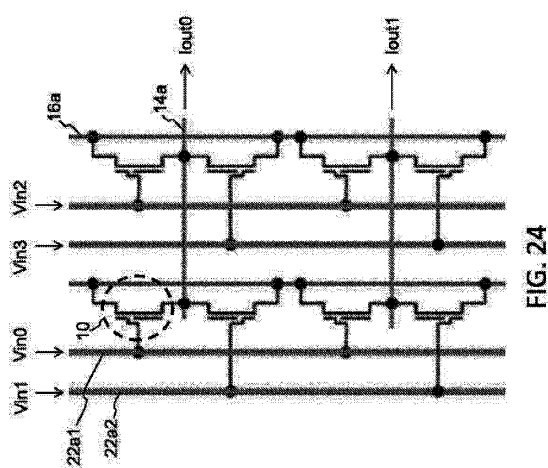
【図 2 2】



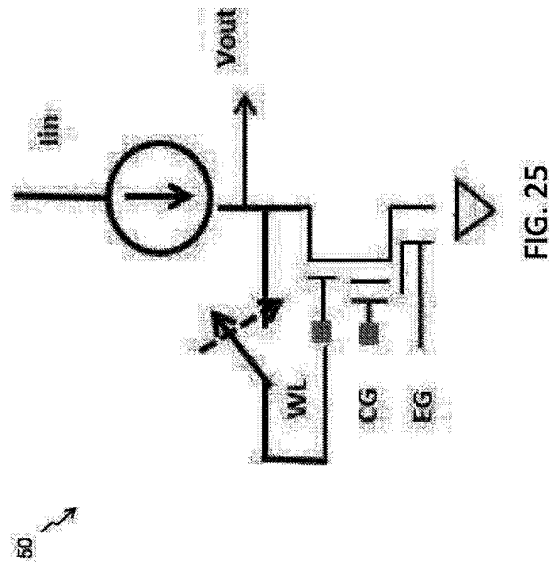
【図 2 3】



【図 2 4】



【図 2 5】



10

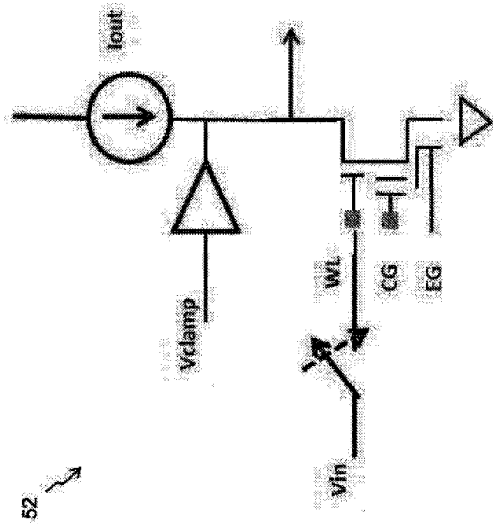
20

30

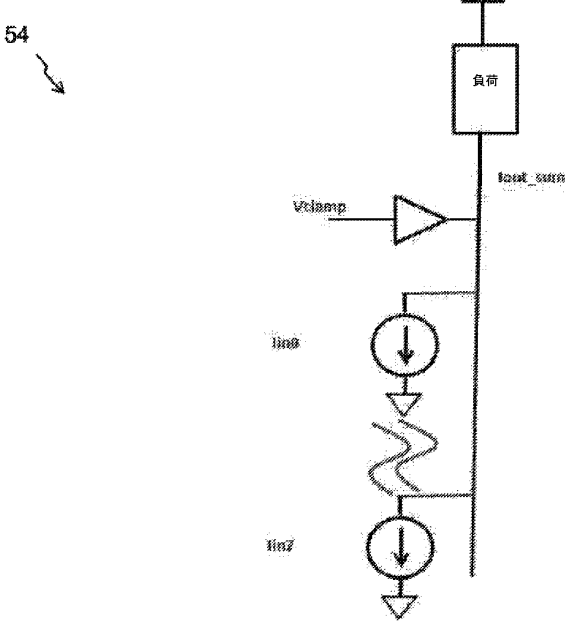
40

50

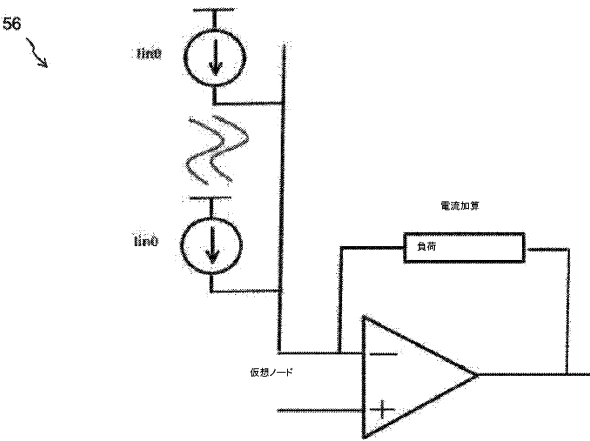
【図 2 6】



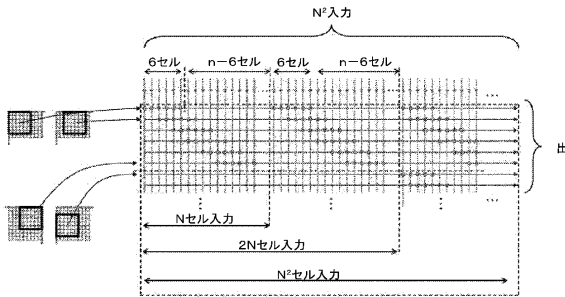
【図 2 7】



【図 2 8】



【図 2 9】



10

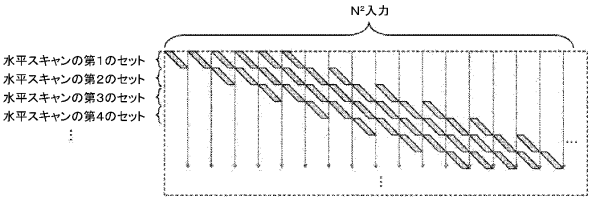
20

30

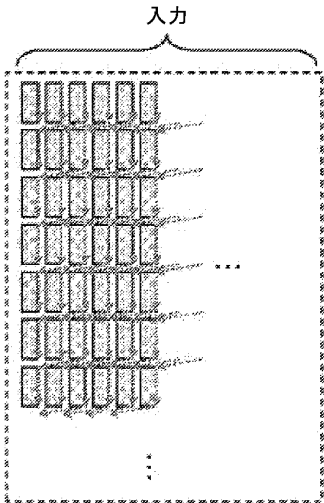
40

50

【図 3 0】



【図 3 1】



【図 3 2】

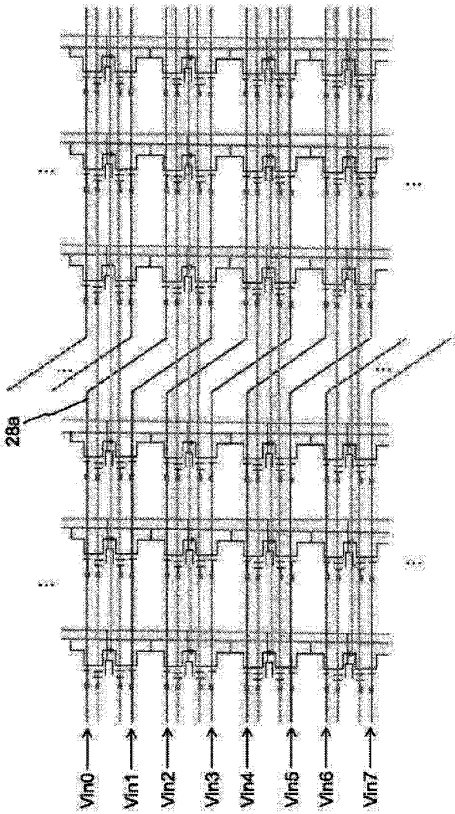


FIG. 32

【図 3 3】

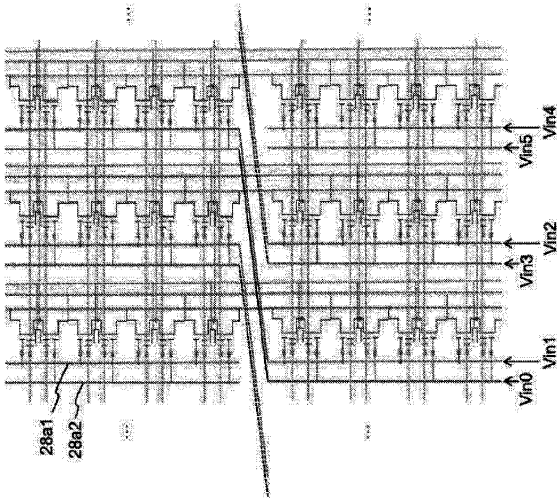


FIG. 33

10

20

30

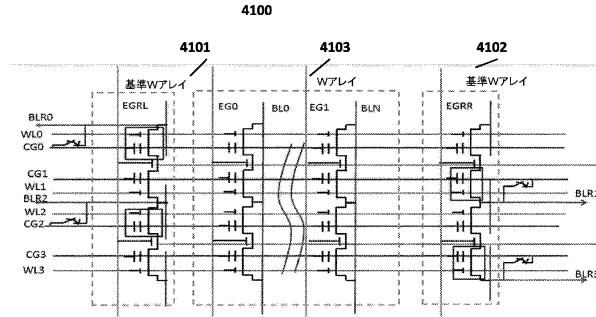
40

50

【図 40】

	WL	WL-非選択	BL	BL-非選択	CG	CG-同じセクタを非選択	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 (ビームアップ)	0V	0-2.6V	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	0V
プログラム	0.7-1V	-0.5V/0V	0.1-1uA	V _{th} (1-2V)	4-11V	0-2.6V	0-2.6V	4.5-5V	0-2.6V	4.5-5V	0-1V

【図 41】

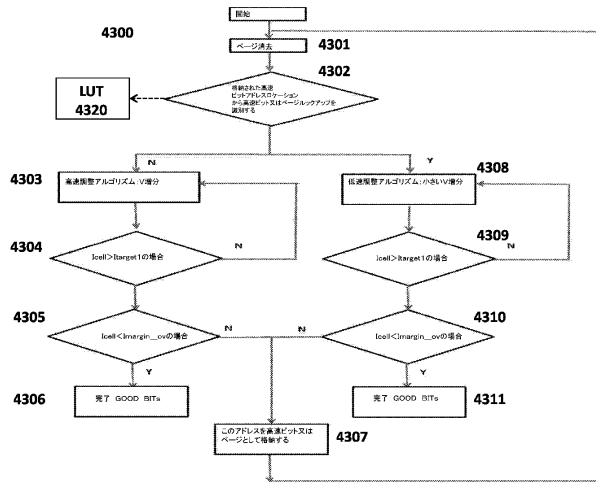


10

【図 42】

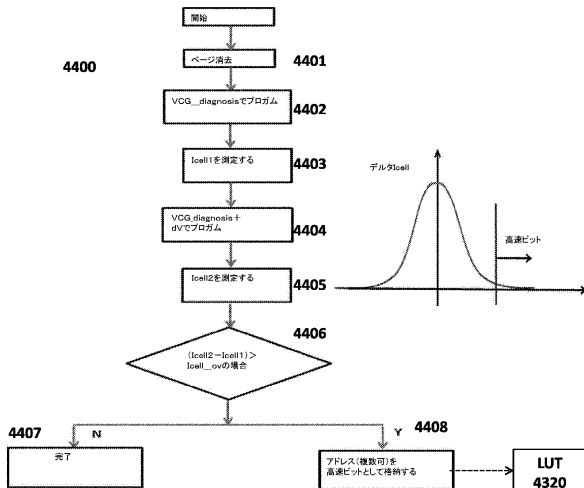
	WL	WL-非選択	BL	BL-非選択	CG	CG-同じセクタを非選択	CG-非選択	EG	EG-非選択	SL	SL-非選択
読み出し	1.0-2V	-0.5V/0V	0.6-2V (ビームアップ)	0V	0-2.6V	0-2.6V	0-2.6V	0-2.6V	0-2.6V	0V	0V
消去	0V	0V	0V	0V	0V	CGINH(4-9V)	0-2.6V	5-12V	0-2.6V	0V	0V
プログラム	0.7-1V	-0.5V/0V	0.1-1uA	V _{th} (1-2V)	4-11V	0-2.6V	0-2.6V	4.5-5V	0-2.6V	4.5-5V	0-1V

【図 43】

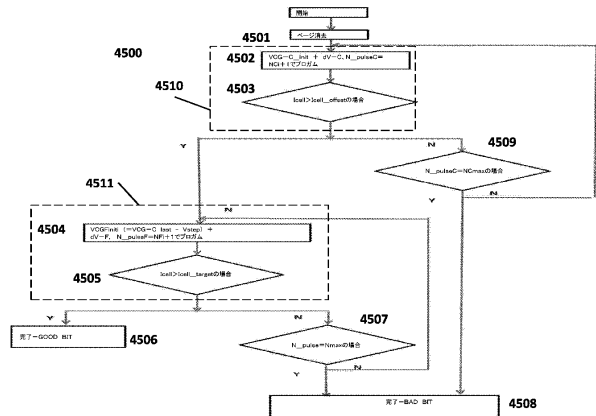


20

【図 44】



【図 45】



30

40

50

【図 46】

均一ステップアルゴリズム
4600

V標的	8V
V開始	4V
均一ステップ	
ステップサイズ	0.01V
#ステップ	400
全 #ステップ:	400

【圖 47】

均一ログステップ
粗／微細アルゴリズム
4700

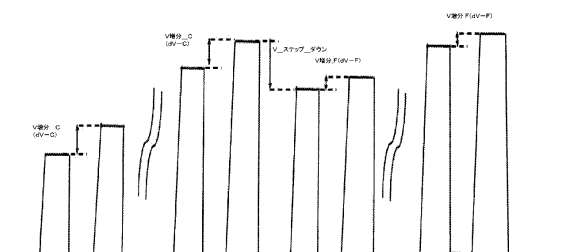
V標的		8V
V開始		4V
粗一均一ステップ		
	ステップサイズ	0.4V
	#ステップ	10
微細1		
	1ステップダウン	0.2V
	ステップサイズ	0.1V
	#ステップ	6
微細2		
	1ステップダウン	0.1V
	ステップサイズ	0.01V
	#ステップ	20
全#ステップ:		46

【図 48】

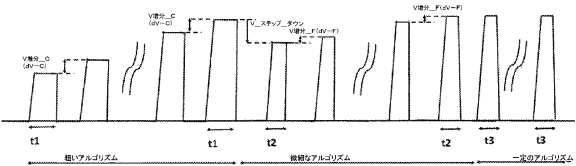
バイナリサーチステップ
粗／微細アルゴリズム
4800

V標的	8V
V開始	4V
粗—バイナリサーチステップ	
ステップ1	2V
ステップ2	1V
ステップ3	0.5V
ステップ4	0.25V
ステップ5	0.125V
ステップ6	0.0625V
微細1	
1ステップダウン	0.0625V
ステップサイズ	0.01V
#ステップ	19
	25

【図 49】



【図 50】



10

20

30

40

50

フロントページの続き

- ニア州、ダブリン、アスターウッド ドライブ 5 5 9 9
(72)発明者 ドー、ナン
アメリカ合衆国 9 5 0 7 0 カリフォルニア州、サラトガ、ウォールナット アベニュー 2 0 4 5 1
(72)発明者 レムケ、スティーブン
アメリカ合衆国 9 5 0 0 6 カリフォルニア州、ボルダー クリーク、エイコーン ドライブ 5 5 4
(72)発明者 ハリハラン、サントシュ
アメリカ合衆国 9 5 1 2 8 カリフォルニア州、サンノゼ、アパート 7、トピーカ アベニュー 6 8
(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、プリストル ベイ コモン 1 8 4 8
審査官 後藤 彰
(56)参考文献 特表 2 0 1 0 - 5 3 0 5 9 4 (J P , A)
国際公開第 2 0 1 7 / 2 0 0 8 8 3 (W O , A 1)
特開 2 0 0 7 - 3 2 3 7 1 6 (J P , A)
特開 2 0 1 5 - 0 5 3 0 9 8 (J P , A)
特開 2 0 0 8 - 1 3 0 1 8 2 (J P , A)
(58)調査した分野 (Int.Cl., D B 名)
G 1 1 C 1 6 / 1 0
G 1 1 C 1 1 / 5 4
G 1 1 C 1 1 / 5 6