(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2023年6月22日(22.06.2023)

WIPO | PCT

(10) 国際公開番号
WO 2023/112674 A1

(51) 国際特許分類:
G06N 3/063 (2023.01) *G06G 7/60* (2006.01)
G06G 7/14 (2006.01) *G11C 11/54* (2006.01)
G06G 7/16 (2006.01) *G11C 13/00* (2006.01)
G06G 7/19 (2006.01) *H10B 63/00* (2023.01)

(21) 国際出願番号： PCT/JP2022/044141

(22) 国際出願日： 2022年11月30日(30.11.2022)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(30) 優先権データ：
特願 2021-201499 2021年12月13日(13.12.2021) JP

(71) 出願人: ヌヴォトンテクノロジージャパン株式会社 (**NUVOTON TECHNOLOGY CORPORATION JAPAN**) [JP/JP]; 〒6178520 京都府長岡京市神足焼町１番地 Kyoto (JP).

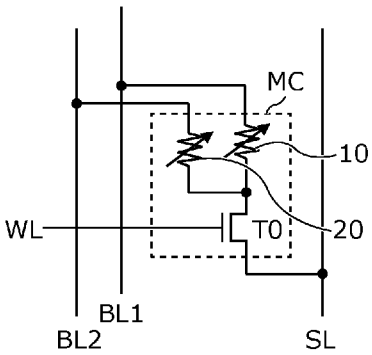
(72) 発明者: 安原 隆太郎 (**YASUHARA, Ryutaro**). 藤井 覚 (**FUJII, Satoru**). 村岡 俊作 (**MURAOKA, Shunsaku**).

(74) 代理人: 新居 広守, 外 (**NII, Hiromori** et al.); 〒5320011 大阪府大阪市淀川区西中島５丁目３番１０号タナカ・イトーピア新大阪ビル６階新居国際特許事務所内 Osaka (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG,

(54) **Title:** ARTIFICIAL INTELLIGENCE PROCESSING DEVICE, AND LEARNING AND INFERENCE METHOD FOR ARTIFICIAL INTELLIGENCE PROCESSING DEVICE

(54) 発明の名称： 人工知能処理装置および人工知能処理装置の学習推論方法



(57) **Abstract:** Provided is an artificial intelligence processing device (200) that uses variable resistance non-volatile storage elements, wherein: a first variable resistance non-volatile storage element (10) and a second variable resistance non-volatile storage element (20) which have differing properties are mounted on a single substrate (70); and, when voltage pulses of identical polarity and identical voltage are continuously applied a plurality of times, the ratio of the amount of change in conductance at the second voltage pulse to the amount of change of conductance at the first voltage pulse in the first variable resistance non-volatile storage element (10) is smaller than the ratio of the amount of change in conductance at the second voltage pulse to the amount of change in conductance at the first voltage pulse in the second variable resistance non-volatile storage element (20).

(57) 要約：抵抗変化型不揮発性記憶素子を用いた人工知能処理装置（２００）であって、特性の異なる第１の抵抗変化型不揮発性記憶素子（１０）と第２の抵抗変化型不揮発性記憶素子（２０）とが１つの基板（７０）上に搭載され、同一極性かつ同一電圧の電圧パルスを複数回連続して印加すると、第１の抵抗変化型不揮発性記憶素子（１０）における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率が、第２の抵抗変化型不揮発性記憶素子（２０）における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率よりも小さい。

WO 2023/112674 A1

［続葉有］

SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

明　細　書

発明の名称 ：

人工知能処理装置および人工知能処理装置の学習推論方法

技術分野

[0001] 本開示は、人工知能処理装置およびその学習推論方法に関し、特に、与えられる電気的信号に応じてその抵抗値が変化する抵抗変化型不揮発性記憶素子を用いた人工知能処理装置に関する。

背景技術

[0002] 情報通信技術の進展に伴い、あらゆるものがインターネットに繋がるＩｏＴ（Internet of Things）技術の到来が注目されている。ＩｏＴ技術において、様々な電子機器がインターネットに接続されることで、機器の高性能化が期待されているが、更なる高性能化を実現する技術として、電子機器自らが学習と判断を行う人工知能（AI:Artificial Intelligence）技術の研究開発が近年活発に行われている。

[0003] 人工知能技術において、人間の脳型情報処理を工学的に模倣したニューラルネットワーク技術が用いられており、ニューラルネットワーク演算を高速、低消費電力で実行する半導体集積回路の研究開発が盛んに行われている。

[0004] 特許文献１に、従来のニューラルネットワーク演算回路が開示されている。ニューラルネットワークは複数の入力が各々異なる結合重み係数を有するシナプスと呼ばれる結合で接続されたニューロンと呼ばれる（パーセプトロンと呼ばれる場合もある）基本素子から構成され、複数のニューロンが互いに接続されることで、画像認識や音声認識といった高度な演算処理を行うことができる。ニューロンでは各入力と各結合重み係数を乗算したものを全て加算した積和演算動作が行われる。積和演算回路は入力や結合重み係数を格納するメモリ回路とレジスタ回路、入力と結合重み係数を乗算する乗算回路、乗算結果を累積加算するアキュムレータ回路、及びこれらの回路ブロックの動作制御を行う制御回路で構成される。これからの回路ブロックは全てデ

ジタル回路で構成される。

[0005] 特許文献２に、従来のニューラルネットワーク演算回路の別の例が開示されている。ニューラルネットワーク演算回路を多階調のアナログ抵抗値、あるいは、その逆数であるコンダクタンス（以下、単に「コンダクタンス」という）が設定可能な抵抗変化型不揮発性メモリを用いて構成するものであり、不揮発性メモリ素子に結合重み係数に相当するコンダクタンスを格納し、入力に相当する電圧値を不揮発性メモリ素子に印加し、このとき不揮発性メモリ素子に流れるアナログ電流値を利用する。ニューロンで行われる積和演算動作は、複数の結合重み係数を複数の不揮発性メモリ素子にコンダクタンスとして格納し、入力に相当する電圧値を複数の不揮発性メモリ素子に印加し、複数の不揮発性メモリ素子に流れる電流値を合算したアナログ電流値を積和演算の結果として得ることで行われる。ここで、不揮発性メモリにコンダクタンスを書き込む際の書き込み方法としては、事前に導出した結合重み係数から各不揮発性メモリに書き込むコンダクタンスを算出し、各不揮発性メモリにコンダクタンスを書き込む。

[0006] 非特許文献１に、従来のニューラルネットワーク演算回路のさらに別の例が開示されている。こちらにおいてもニューラルネットワーク演算回路をコンダクタンスが設定可能な抵抗変化型不揮発性メモリを用いて構成するものであり、不揮発性メモリ素子に結合重み係数に相当するコンダクタンスを格納し、入力に相当するアナログ電圧値を不揮発性メモリ素子に印加し、このとき不揮発性メモリ素子に流れるアナログ電流値を利用する点では同じである。ここで、不揮発性メモリにコンダクタンスを書き込む際の書き込み方法としては、書き込む前の時点でのコンダクタンスから書き込み後に設定されるコンダクタンスの間の変化量が初めに導出され、そのコンダクタンスの変化量に応じた書き込みが不揮発性メモリ素子に対して行われる。

[0007] すなわち、特許文献２のようなニューラルネットワーク演算回路の例においては、一般的に不揮発性メモリに書き込まれるコンダクタンスそのものを基に書き込み動作が行われるのに対し、非特許文献１のようなニューラルネ

ットワーク演算回路においては、一般的に不揮発性メモリに書き込まれるコンダクタンスの書き込み前後の変化量を基に書き込み動作が行われる点が異なる。特許文献２、非特許文献１に開示されている不揮発性メモリ素子を用いたニューラルネットワーク演算回路は、いずれも前述したデジタル回路で構成されるニューラルネットワーク演算回路と比べて低消費電力化が実現可能であり、コンダクタンスが設定可能な抵抗変化型不揮発性メモリのプロセス開発、デバイス開発、及び回路開発が近年盛んに行われている。

先行技術文献

特許文献

[0008] 特許文献1：特開２００１－１８８７６７号公報

特許文献2：国際公開第２０１９／０４９７４１号

非特許文献

[0009] 非特許文献1：M. Prezioso、 et al.、 “Training and operation of an integrated neuromorphic network based on metal-oxide memristors、” Nature、 no. 521、pp. 61-64、 2015.

発明の概要

発明が解決しようとする課題

[0010] しかしながら、前述した従来のニューラルネットワーク演算回路は以下に示す課題があった。

[0011] すなわち、不揮発性メモリに書き込まれるコンダクタンスそのものを基に書き込み動作が行われるニューラルネットワーク演算回路は、事前に導出したコンダクタンスそのものを用いて書き込みを行うため、精度よくコンダクタンスを不揮発性メモリ素子に書き込むことが出来ることから、製品出荷後（つまり、ニューラルネットワーク演算回路等の人工知能処理装置の出荷後）には主にニューラルネットワークでの積和演算のみを処理する「推論用人工知能処理装置」に適している。しかしながら、一般的にソフトウェアでの処理を前提としたニューラルネットワークにおいては、結合重み係数の更新

、すなわち学習は結合重み係数の変化量を基に行われるため、製品出荷後にコンダクタンスの更新を頻繁に行う「学習用人工知能処理装置」においては、コンダクタンスそのものを基に書き込み動作が行われるニューラルネットワーク演算回路を用いると学習が効率的に行われないという点が課題である。

[0012] 図１Ａは、コンダクタンスそのものを基に書き込み動作が行われるニューラルネットワーク演算回路における学習時のプロセスを示す。このプロセスでは、各不揮発性メモリ素子の結合重み係数を用いて推論を実施し（Ｓ１０）、その推論結果と教師ラベルとの差を確認し（Ｓ１１）、各不揮発性メモリ素子について結合重み係数の更新時の変化量を算出し（Ｓ１２）、各不揮発性メモリ素子について現在の結合重み係数を読み出し（Ｓ１３）、読み出した結合重み係数に基づいて各不揮発性メモリ素子について更新後の結合重み係数を算出し（Ｓ１４）、算出した結合重み係数を各不揮発性メモリ素子に書き込む（Ｓ１５）。つまり、結合重み係数自体を書き込む場合には、６つのステップが必要となる。

[0013] 図１Ｂは、コンダクタンスの書き込み前後の変化量を基に書き込み動作が行われるニューラルネットワーク演算回路における学習時のプロセスを示す。このプロセスでは、各不揮発性メモリ素子の結合重み係数を用いて推論を実施し（Ｓ１０）、その推論結果と教師ラベルとの差を確認し（Ｓ１１）、各不揮発性メモリ素子について結合重み係数の更新時の変化量を算出し（Ｓ１２）、各不揮発性メモリ素子について、算出した結合重み係数の変化量だけ現在の結合重み係数が変化するように結合重み係数を更新する（Ｓ２０）。つまり、結合重み係数の変化量を書き込む場合には、４つのステップで済む。図１Ｂにおける学習プロセスは、ソフトウェアベースでニューラルネットワーク演算を行う場合と同じプロセスで実施することができ、図１Ａに示す学習プロセスよりも手順が少ないことが分かる。

[0014] また、コンダクタンスの書き込み前後の変化量を基に書き込み動作が行われるニューラルネットワーク演算回路は、変化量を基に書き込み動作を行う

ことから、図１Ｂにおける学習プロセスでコンダクタンスを更新でき、製品出荷後にコンダクタンスの更新を頻繁に行う「学習用人工知能処理装置」に適している。しかしながら、図１Ｂに示される学習プロセスが実装されたニューラルネットワーク演算回路では、製品出荷前の「初期設定」時においては、事前に導出したコンダクタンスそのものを用いて書き込みを行うことが出来ず、電圧パルスを複数回連続で打つことにより事前に導出した値にまでコンダクタンスを変化させることが必要となる。その際、書き込み特性にばらつきが生じ、初期設定を高効率・高精度に行えないという課題がある。

[0015] 本開示は上記課題を鑑みてなされたものであり、製品出荷時等における高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が可能な抵抗変化型不揮発性記憶素子を用いた人工知能処理装置およびその学習推論方法を提供することにある。

課題を解決するための手段

[0016] 本開示の一形態に係る人工知能処理装置は、１つの基板と、前記１つの基板に搭載され、積和演算を行う積和演算回路とを備え、前記積和演算回路は、前記積和演算における積の演算に用いられる結合重み係数をコンダクタンスとして保持する、特性の異なる第１の抵抗変化型不揮発性記憶素子と第２の抵抗変化型不揮発性記憶素子とを含み、同一極性かつ同一電圧の電圧パルスを複数回連続して印加されると、前記第１の抵抗変化型不揮発性記憶素子における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率が、前記第２の抵抗変化型不揮発性記憶素子における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率よりも小さい。

[0017] 本開示の一形態に係る人工知能処理装置の学習推論方法は、上記人工知能処理装置の学習推論方法であって、結合重み係数を変更する場合に、当該場合が、初期設定のために結合重み係数を変更する第１のケースであるか、学習において結合重み係数を変更する第２のケースであるかを判断する判断ス

テップと、前記判断ステップでの判断の結果、前記第１のケースであると判断した場合に、前記第１の抵抗変化型不揮発性記憶素子のコンダクタンスを変更する第１の書き換えステップと、前記判断ステップでの判断の結果、前記第２のケースであると判断した場合に、前記第２の抵抗変化型不揮発性記憶素子のコンダクタンスを変更する第２の書き換えステップと、推論する場合に、前記積和演算回路から出力される、前記第１の抵抗変化型不揮発性記憶素子に流れる電流および前記第２の抵抗変化型不揮発性記憶素子に流れる電流の合計を、積和演算における１つの積として、用いる推論ステップとを含む。

発明の効果

[0018] 本開示の抵抗変化型不揮発性記憶素子を用いた人工知能処理装置およびその学習推論方法によれば、製品出荷時等における高効率・高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が可能になる。

図面の簡単な説明

[0019] [図1A]図１Ａは、コンダクタンスそのものを基に書き込み動作が行われるニューラルネットワーク演算回路における学習時のプロセスを示す図である。

[図1B]図１Ｂは、コンダクタンスの書き込み前後の変化量を基に書き込み動作が行われるニューラルネットワーク演算回路における学習時のプロセスを示す図である。

[図2]図２は、実施形態に係る単一の電圧パルスにより設定コンダクタンスに概ね到達する抵抗変化型不揮発性記憶素子の断面模式図である。

[図3]図３は、実施形態に係る同一極性、同一電圧の複数電圧パルスの連続印加により徐々にコンダクタンスが変化する抵抗変化型不揮発性記憶素子の断面模式図である。

[図4]図４は、実施形態に係る単一の電圧パルスにより設定コンダクタンスに概ね到達する抵抗変化型不揮発性記憶素子を、同一極性、同一電圧の複数電圧パルスの連続印加により抵抗変化させた際のコンダクタンスの変化を示す

図である。

[図5]図５は、実施形態に係る同一極性、同一電圧の複数電圧パルスの連続印加により徐々にコンダクタンスが変化する抵抗変化型不揮発性記憶素子を、同一極性、同一電圧の複数電圧パルスの連続印加により抵抗変化させた際のコンダクタンスの変化を示す図である。

[図6A]図６Ａは、従来技術におけるメモリセルの回路図である。

[図6B]図６Ｂは、従来技術におけるメモリセルの断面図である。

[図7A]図７Ａは、実施形態に係るメモリセルの回路図である。

[図7B]図７Ｂは、実施形態に係るメモリセルの一例を示す断面図である。

[図7C]図７Ｃは、実施形態に係るメモリセルの、図７Ｂとは異なる一例を示す断面図である。

[図8A]図８Ａは、別の実施形態に係るメモリセルの回路図である。

[図8B]図８Ｂは、図８Ａに記載された第１の抵抗変化型不揮発性記憶素子が搭載された人工知能処理装置のメモリセルの断面図の一例を示した図である。

[図8C]図８Ｃは、図８Ａに記載された、第２の抵抗変化型不揮発性記憶素子が搭載された人工知能処理装置のメモリセルの断面図の一例を示した図である。

[図9A]図９Ａは、実施例に係る人工知能処理装置のモデルを示すブロック図である。

[図9B]図９Ｂは、図９Ａに示されるニューロンの機能を示す図である。

[図10A]図１０Ａは、図９Ｂに示されるニューロンを実現する回路例を示す図である。

[図10B]図１０Ｂは、図９Ｂに示されるニューロンを実現する別の回路例を示す図である。

[図11]図１１は、図１０Ａに示されるニューロンからなる人工知能処理装置の全体の構成を示すブロック図である。

[図12A]図１２Ａは、図１０Ａに示されるメモリセルに対する書き込みおよび

読み出し時におけるワード線ＷＬ、ビット線ＢＬ１、ＢＬ２、および、ソース線ＳＬを介した印加電圧の例を示す図である。

[図12B]図１２Ｂは、図１０Ｂに示されるメモリセルに対する書き込みおよび読み出し時におけるワード線ＷＬ１、ＷＬ２、ビット線ＢＬ、および、ソース線ＳＬを介した印加電圧の例を示す図である。

[図13]図１３は、図１１に示される制御回路の動作例を示すフローチャートである。

発明を実施するための形態

[0020] 以下、本開示に係る実施形態について図面を参照しながら説明する。なお、以下で説明する実施形態はいずれも実施の一具体例を示すものである。したがって、以下の実施形態で示される数値、形状、材料、構成要素、構成要素の配置位置及び接続形態、ステップ、ステップの順序などはあくまで一例であり、本開示を限定する主旨ではない。本開示は請求の範囲だけによって限定される。

[0021] 以下の実施形態における構成要素のうち、本開示の最上位概念を示す独立請求項に記載されていない構成要素については、本開示の課題を達成するのに必ずしも必要ではないが採用し得る形態を構成するものとして説明される。

[0022] 本発明者は、抵抗変化型不揮発性記憶素子を用いた人工知能処理装置における製品出荷時等における高精度の結合重み係数の設定（初期設定）および製品出荷後等における高効率な結合重み係数の更新（学習）の動作効率を高めるべく鋭意検討を行った結果、以下のような知見を得た。

[0023] 本発明者らは、抵抗変化型不揮発性記憶素子を用いた人工知能処理装置において、特性の異なる２種類の抵抗変化型不揮発性記憶素子の両方を１つの基板上に搭載し、同一極性かつ同一電圧の連続した電圧パルス印加により、片方の不揮発性記憶素子は少しずつコンダクタンスが変化し、もう一方の不揮発性記憶素子は１発目の電圧パルス印加によりコンダクタンスが大きく変化し２発目以降の電圧パルス印加によるコンダクタンス変化量が小さい構成

とすることにより、従来課題であった抵抗変化型不揮発性記憶素子を用いた人工知能処理装置における製品出荷時等における高精度の結合重み係数の設定（初期設定）および製品出荷後等における高効率な結合重み係数の更新（学習）の動作効率化の両立が可能となることを見出した。

[0024] 当該知見の詳細は、以下において、実施形態とともに適宜説明される。

[0025] （実施形態）

［抵抗変化型不揮発性記憶素子の構成］

まず、実施形態に係る２種類の抵抗変化型不揮発性記憶素子の構成の一例について説明する。

[0026] 図２は、実施形態に係る第１の抵抗変化型不揮発性記憶素子１０の構成例を示す模式図である。第１の抵抗変化型不揮発性記憶素子１０は、製品出荷時等における高精度の結合重み係数の設定（初期設定）に好適な抵抗変化型不揮発性記憶素子である。

[0027] 図２に示すように、第１の抵抗変化型不揮発性記憶素子１０は、基板１と、基板１の上に形成された第１電極２と、第１電極２の上に金属酸化物層として形成された抵抗変化層３と、抵抗変化層３の上に形成された第２電極４とを備えている。第１電極２及び第２電極４は、抵抗変化層３と電気的に接続されている。すなわち、第１の抵抗変化型不揮発性記憶素子１０は、第１電極２と、第２電極４と、第１電極２と第２電極４との間に介在する抵抗変化層３とを備える。

[0028] なお、第１電極２は、第２電極４と同等のサイズでもよく、また第１電極２、第２電極４及び抵抗変化層３の配置は、上下逆に配置してもよいし、横向けに配置してもよい。

[0029] 基板１は、例えばトランジスタ等の回路素子が形成されたシリコン基板により構成される。また、第１電極２あるいは第２電極４の少なくとも一方は、貴金属である金属、例えば、Ａｕ（金）、Ｐｔ（白金）、Ｉｒ（イリジウム）、Ｐｄ（パラジウム）、Ｒｕ（ルテニウム）などのうちの１つの材料を用いて構成される。例えば、第２タンタル酸化物層３ｂに接触する第２電極

４が貴金属で構成され、第１電極２が貴金属または非貴金属で構成される。このような電極の特徴により、第１の抵抗変化型不揮発性記憶素子１０は、後述する図４に示される抵抗変化特性を有する。

[0030] 抵抗変化層３は、第１電極２及び第２電極４間に印加される電圧パルスに応じて抵抗（言い換えると、コンダクタンス）が変化する。抵抗変化層３は、金属酸化物で構成されており、第１タンタル酸化物層３ａと第２タンタル酸化物層３ｂとが積層されて構成されている。ここで、第２タンタル酸化物層３ｂの酸素含有率は、第１タンタル酸化物層３ａの酸素含有率よりも高くなっている。

[0031] 第１タンタル酸化物層３ａの組成をＴａＯｘとした場合に、０＜ｘ＜２．５であり、且つ、第２タンタル酸化物層３ｂの組成をＴａＯｙとした場合に、ｘ＜ｙであればよい。

[0032] 図３は、実施形態に係る第２の抵抗変化型不揮発性記憶素子２０の構成例を示す模式図である。第２の抵抗変化型不揮発性記憶素子２０は、製品出荷後等における高効率な結合重み係数の更新（学習）に好適な抵抗変化型不揮発性記憶素子である。

[0033] 図３に示すように、第２の抵抗変化型不揮発性記憶素子２０は、第１の抵抗変化型不揮発性記憶素子１０と同様に、基板１１と、基板１１の上に形成された第１電極１２と、第１電極１２の上に金属酸化物層として形成された抵抗変化層１３と、抵抗変化層１３の上に形成された第２電極１４とを備えている。第１電極１２及び第２電極１４は、抵抗変化層１３と電気的に接続されている。すなわち、第２の抵抗変化型不揮発性記憶素子２０は、第１電極１２と、第２電極１４と、第１電極１２と第２電極１４との間に介在する抵抗変化層１３とを備える。

[0034] なお、第１電極１２は第２電極１４と同等のサイズでもよく、また第１電極１２、第２電極１４及び抵抗変化層１３の配置は、上下逆に配置してもよいし、横向けに配置してもよい。

[0035] 基板１１は、例えばトランジスタ等の回路素子が形成されたシリコン基板

により構成される。また、第１電極１２あるいは第２電極１４の少なくとも一方は、非貴金属である金属、例えば、ＴｉＮ（窒化チタン）、ＴａＮ（窒化タンタル）などのうちの１つの材料を用いて構成される。例えば、第２タンタル酸化物層３ｂに接触する第２電極１４が非貴金属で構成され、第１電極１２が貴金属または非貴金属で構成される。このような電極の特徴により、第２の抵抗変化型不揮発性記憶素子２０は、後述する図５に示される抵抗変化特性を有する。

[0036] 抵抗変化層１３は、第１電極１２及び第２電極１４間に印加される電圧パルスに応じて抵抗（言い換えると、コンダクタンス）が変化する。抵抗変化層１３は、金属酸化物で構成されており、第１タンタル酸化物層１３ａと第２タンタル酸化物層１３ｂとが積層されて構成されている。ここで、第２タンタル酸化物層１３ｂの酸素含有率は、第１タンタル酸化物層１３ａの酸素含有率よりも高くなっている。

[0037] 第１タンタル酸化物層１３ａの組成をＴａＯｘとした場合に、０＜ｘ＜２．５であり、且つ、第２タンタル酸化物層１３ｂの組成をＴａＯｙとした場合に、ｘ＜ｙであればよい。

[0038] ［抵抗変化型不揮発性記憶素子の製造方法］

次に、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の製造方法の一例について、第１の抵抗変化型不揮発性記憶素子１０の製造方法の場合を用いて説明する。

[0039] まず、基板１上に、スパッタリング法により、第１電極２を形成する。その後、Ｔａターゲットをアルゴンガス及び酸素ガス中でスパッタリングする所謂反応性スパッタリング法によって、第１電極２の上にタンタル酸化物層を形成する。ここで、タンタル酸化物層における酸素含有率は、アルゴンガスに対する酸素ガスの流量比を変えることにより容易に調整することができる。なお、基板温度は特に加熱することなく室温とすることができる。

[0040] 次に、上記のようにして形成されたタンタル酸化物層の最表面を酸化することによりその表面を改質する。あるいは、高濃度の酸素含有率を有するタ

ンタル酸化物（例えばＴａ２Ｏ５）ターゲットを用いて、より酸素含有率の高い層をスパッタ法で形成する。これにより、先に形成されたタンタル酸化物層の表面に、当該タンタル酸化物層の酸化されなかった領域（第１領域）よりも酸素含有率の高い領域（第２領域）が形成される。

[0041] これらの第１領域及び第２領域が第１タンタル酸化物層３ａ及び第２タンタル酸化物層３ｂにそれぞれ相当し、このようにして形成された第１タンタル酸化物層３ａ及び第２タンタル酸化物層３ｂによって抵抗変化層３が構成されることになる。

[0042] 次に、上記のようにして形成された抵抗変化層３の上に、スパッタリング法により、第２電極４を形成する。

[0043] 最後に、第１の抵抗変化型不揮発性記憶素子１０を形成するために、所望のマスクを用いて、第１電極２、酸素不足型の第１タンタル酸化物層３ａ、第２タンタル酸化物層３ｂおよび第２電極４をパターニングして、抵抗変化層３を第１電極２、第２電極４で挟持した第１の抵抗変化型不揮発性記憶素子１０を形成する。

[0044] なお、第１の抵抗変化型不揮発性記憶素子１０を形成するに当たり、本工程では同じマスクを用いて一括してパターニングを行ったが、工程ごとに個別にパターニングを行ってもかまわない。

[0045] なお、第１電極２及び第２電極４並びに抵抗変化層３の大きさ及び形状は、フォトマスク及びフォトリソグラフィによって調整することができる。本実施形態では、第２電極４及び抵抗変化層３の大きさを０．１μｍ×０．１μｍ（面積０．０１μｍ２）とし、第１電極２と抵抗変化層３とが接する部分の大きさも０．１μｍ×０．１μｍ（面積０．０１μｍ２）としたが、このような大きさ及び形状に限定されるものではなく、レイアウト設計によって適宜変更可能である。

[0046] なお、第２の抵抗変化型不揮発性記憶素子２０についても、第１電極１２あるいは第２電極１４の具体的な材料の相違を除いて、製造方法の基本的な流れは、第１の抵抗変化型不揮発性記憶素子１０と同様である。

[0047] ［抵抗変化型不揮発性記憶素子の動作方法と抵抗変化特性］

次に、上述した製造方法により得られた第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の動作について説明する。

[0048] 以下では、抵抗変化層３の抵抗値が所定の高い値（例えば、３００ｋΩ）にある場合を高抵抗状態、同じく所定の低い値（例えば、１２ｋΩ）にある場合を低抵抗状態といい、同一極性、同一電圧の複数電圧パルスの連続印加により、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスはその間を変化するものとする。

[0049] また、以下では、負極性の書き込み電圧パルスを第１の抵抗変化型不揮発性記憶素子１０または第２の抵抗変化型不揮発性記憶素子２０の第１電極２及び第２電極４間に印加することにより、抵抗変化層３のコンダクタンスが増加し、抵抗変化層３が高抵抗状態から低抵抗状態へ変化する過程を低抵抗化（あるいは、「セット」ともいう）と称し、正極性の書き込み電圧パルスを第１電極２及び第２電極４間に印加することにより、抵抗変化層３のコンダクタンスが減少し、抵抗変化層３が低抵抗状態から高抵抗状態へ変化する過程を高抵抗化（あるいは、「リセット」ともいう）と称する。また、抵抗変化型不揮発性記憶素子の低抵抗化および高抵抗化を、コンダクタンスあるいは抵抗値の「書き込み」ともいう。

[0050] このような書き込みを繰り返すことにより、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０が不揮発性記憶素子として動作する。

[0051] ここで、初期過程について説明する。本実施形態では、第１回目の上記書き込みの前に、通常、一度だけ、初期過程を実行する。初期過程とは、その後の低抵抗化および高抵抗化において安定した抵抗変化動作を実現するための準備過程であり、「ブレイク」あるいは「フォーミング」とも呼ばれる。

[0052] 一般に、製造直後の第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０は通常抵抗変化時の高抵抗状態よりさらに

高い初期抵抗値を有しており、その状態で通常動作時の低抵抗化電圧パルスあるいは高抵抗化電圧パルスを印加しても抵抗変化は起こらない。

[0053] そこで、初期過程において、初期電圧パルスを第１電極２及び第２電極４間に印加する。

[0054] 以降は、高抵抗化あるいは低抵抗化における同一極性、同一電圧の複数電圧パルスの連続印加により、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスは高抵抗状態および低抵抗状態の間を変化することになる。

[0055] つまり、初期過程は、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０を製造した後にまだ電圧が印加されたことのない初期状態の第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０に対して行なわれる過程である。

[0056] 上述した初期過程を経ることにより、フィラメントと呼ばれる周囲の酸素不足度よりも大きい酸素不足度を持つ局所領域が抵抗変化層３内に形成される。

[0057] なお、本実施形態では初期過程を経ることによりフィラメントを形成したが、必ずしも初期過程を経てフィラメントを形成する必要は無く、抵抗変化型不揮発性記憶素子の形成時に酸素不足度が０％よりも十分に大きい酸化物層を設けることで代用してもよい。

[0058] 次に、本実施形態に係る第１の抵抗変化型不揮発性記憶素子１０に固有の特性について、説明する。図４は、本実施形態に係る第１の抵抗変化型不揮発性記憶素子１０において、同一極性、同一電圧の複数電圧パルスの連続印加により、第１の抵抗変化型不揮発性記憶素子１０のコンダクタンスを高抵抗状態あるいは低抵抗状態に変化させた際の抵抗変化特性である。横軸は、電圧パルスの印加回数を示し、縦軸は、コンダクタンスを示す。 図４に示すように、第１の抵抗変化型不揮発性記憶素子１０に同一極性、同一電圧の複数電圧パルスの連続印加を行い低抵抗化した際には、図４における矩形波形の上辺における黒丸のプロットに示されるように、１発目の電圧パルスで

のコンダクタンスの変化量が大きく、１発目の電圧パルスにより高抵抗状態から低抵抗状態に近い状態にまでコンダクタンスが増加する。その後、連続して２発目、３発目とパルス印加を行っても、１発目の電圧パルスでのコンダクタンスの変化量と比べると非常に小さい変化量に留まる。同様に、第１の抵抗変化型不揮発性記憶素子１０に同一極性、同一電圧の複数電圧パルスの連続印加を行い高抵抗化した際には、図４における矩形波形の下辺における黒丸のプロットに示されるように、１発目の電圧パルスでのコンダクタンスの変化量が大きく、１発目の電圧パルスにより低抵抗状態から高抵抗状態に近い状態にまでコンダクタンスが減少する。その後、連続して２発目、３発目とパルス印加を行っても、１発目の電圧パルスでのコンダクタンスの変化量と比べると非常に小さい変化量に留まる。

[0059] このような抵抗変化特性を示す第１の抵抗変化型不揮発性記憶素子１０においては、素子と接続する電流制限回路により低抵抗状態におけるコンダクタンス自体（言い換えると、多階調のアナログ抵抗値）を調整することが可能である。つまり、第１の抵抗変化型不揮発性記憶素子１０は、連続する電圧パルスの印加における初回の電圧パルスによってコンダクタンスが大きく変化し、その後の電圧パルスでは変化量が小さいという特性を有する。そのため、第１の抵抗変化型不揮発性記憶素子１０は、製品出荷時等における高精度の結合重み係数の設定（初期設定）に好適な抵抗変化型不揮発性記憶素子であるといえる。つまり、第１の抵抗変化型不揮発性記憶素子１０と電流制限回路を用いることにより、抵抗変化型不揮発性記憶素子を用いた人工知能処理装置において、製品出荷前および製品出荷後のファームウェア更新、学習モデルの更新、定期メンテナンス、および第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスの変更では結合重み係数更新が不十分な時に、事前に各素子のコンダクタンス設定値を取得して各素子に書き込むことが可能となる。

[0060] 次に、本実施形態に係る第２の抵抗変化型不揮発性記憶素子２０に固有の特性について、説明する。図５は、本実施形態に係る第２の抵抗変化型不揮

発性記憶素子２０において、同一極性、同一電圧の複数電圧パルスの連続印加により、第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスを高抵抗状態あるいは低抵抗状態に変化させた際の抵抗変化特性である。横軸は、電圧パルスの印加回数を示し、縦軸は、コンダクタンスを示す。

[0061] 図５に示すように、第２の抵抗変化型不揮発性記憶素子２０に同一極性、同一電圧の複数電圧パルスの連続印加を行い低抵抗化した際には、図５における立ち上がり波形における黒丸のプロットに示されるように、１発目の電圧パルスでのコンダクタンスの変化の割合が第１の抵抗変化型不揮発性記憶素子１０における１発目の電圧パルスでのコンダクタンスの変化の割合よりも小さい。その後、連続して２発目、３発目とパルス印加を行った際には、引き続き低抵抗化方向へのコンダクタンスの変化が生じる。同様に、第２の抵抗変化型不揮発性記憶素子２０に同一極性、同一電圧の複数電圧パルスの連続印加を行い高抵抗化した際には、図５における立ち下がり波形における黒丸のプロットに示されるように、１発目の電圧パルスでのコンダクタンスの変化の割合が第１の抵抗変化型不揮発性記憶素子１０における１発目の電圧パルスでのコンダクタンスの変化の割合よりも小さい。その後、連続して２発目、３発目とパルス印加を行った際には、引き続き高抵抗化方向へのコンダクタンスの変化が生じる。

[0062] このような抵抗変化特性を示す第２の抵抗変化型不揮発性記憶素子２０においては、パルス印加前のコンダクタンスによらず、低抵抗化方向の極性を持つ電圧パルスあるいは高抵抗化方向の極性を持つ電圧パルスを印加することにより一定量のコンダクタンスの増加（低抵抗化）あるいはコンダクタンスの減少（高抵抗化）を生じさせることが可能である。つまり、第２の抵抗変化型不揮発性記憶素子２０は、連続する電圧パルスによって徐々にコンダクタンスが変化する特性を有する。そのため、第２の抵抗変化型不揮発性記憶素子２０は、製品出荷後等における高効率な結合重み係数の更新（学習）に好適な抵抗変化型不揮発性記憶素子であるといえる。つまり、第２の抵抗変化型不揮発性記憶素子２０を用いることにより、抵抗変化型不揮発性記憶

素子を用いた人工知能処理装置において、製品出荷後の学習のための結合重み係数の更新時に、各素子に対し直接的に一定量のコンダクタンスの増加あるいは減少させる書き込みが可能となる。同一極性かつ同一電圧の電圧パルスを複数回連続して印加されると、第２の抵抗変化型不揮発性記憶素子２０における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率は、第１の抵抗変化型不揮発性記憶素子１０における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率よりも大きい。

[0063] 以上のように、単一の電圧パルスにより設定コンダクタンスに概ね到達する第１の抵抗変化型不揮発性記憶素子１０と、同一極性、同一電圧の複数電圧パルスの連続印加により徐々にコンダクタンスが変化する第２の抵抗変化型不揮発性記憶素子２０とを比較すると、単一の電圧パルスによりコンダクタンスが設定可能な第１の抵抗変化型不揮発性記憶素子１０はより強固に特定の抵抗状態を保っていると考えられることから、書き込み後のリテンション特性は高いと考えられる。一方、同一極性、同一電圧の複数電圧パルスの連続印加により徐々にコンダクタンスが変化する第２の抵抗変化型不揮発性記憶素子２０は、緩やかにコンダクタンスを変化させることから、書き込み自体による素子劣化度合いが小さく、エンデュランス特性は高いと考えられる。

[0064] ［抵抗変化型不揮発性記憶素子の配置］

図６Ａおよび図６Ｂは、従来技術における抵抗変化型不揮発性記憶素子の回路図および断面図である。

[0065] 図６Ａは、従来の抵抗変化型不揮発性記憶素子が搭載された人工知能処理装置のメモリセルの回路図を示したものである。メモリセルＭＣは抵抗変化型不揮発性記憶素子ＲＰとセルトランジスタＴ０の直列接続されたもので構成され、１つのセルトランジスタＴ０と１つの抵抗変化型不揮発性記憶素子ＲＰから構成される『１Ｔ１Ｒ』型のメモリセルである。メモリセルＭＣの

ワード線ＷＬはセルトランジスタＴ０のゲート端子に接続され、ビット線ＢＬは抵抗変化型不揮発性記憶素子ＲＰに接続され、ソース線ＳＬはセルトランジスタＴ０のソース端子に接続される。

[0066] 図６Ｂは、従来の抵抗変化型不揮発性記憶素子ＲＰが搭載された人工知能処理装置のメモリセルの断面図を示したものである。基板６０上に拡散領域６１ａ、６１ｂが形成されており、拡散領域６１ａがセルトランジスタＴ０のソース端子として、拡散領域６１ｂがセルトランジスタのドレイン端子として作用する。拡散領域６１ａ、６１ｂ間がセルトランジスタＴ０のチャネル領域として作用し、このチャネル領域上に酸化膜６２、ポリシリコンで形成されるゲート電極６３が形成され、セルトランジスタＴ０として動作する。セルトランジスタＴ０のソース端子である拡散領域６１ａはビア６４ａを介して第１配線層６５ａであるソース線ＳＬに接続される。セルトランジスタＴ０のドレイン端子である拡散領域６１ｂはビア６４ｂを介して第１配線層６５ｂに接続される。さらに、第１配線層６５ｂはビア６６を介して第２配線層６７に接続され、第２配線層６７はビア６８ａを介して抵抗変化型不揮発性記憶素子ＲＰに接続される。抵抗変化型不揮発性記憶素子ＲＰは第１電極２、抵抗変化層３、第２電極４から構成される。抵抗変化型不揮発性記憶素子ＲＰはビア６８ｂを介して第３配線層６９であるビット線ＢＬに接続される。

[0067] この構成においては、抵抗変化型不揮発性記憶素子ＲＰとしては、単一の電圧パルスにより設定コンダクタンスに概ね到達する実施形態に係る第１の抵抗変化型不揮発性記憶素子１０もしくは同一極性、同一電圧の複数電圧パルスの連続印加により徐々にコンダクタンスが変化する実施形態に係る第２の抵抗変化型不揮発性記憶素子２０のいずれか一方のみが人工知能処理装置に搭載される。そのため、製品出荷時等における高効率・高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）を両立することが困難である。

[0068] 図７Ａ～図７Ｃは、実施形態に係る抵抗変化型不揮発性記憶素子の回路図

、断面図の例を示す図である。

[0069] 図７Ａは、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の両方が搭載された人工知能処理装置のメモリセルの回路図を示したものである。メモリセルＭＣは単一の電圧パルスによりコンダクタンスが設定可能な第１の抵抗変化型不揮発性記憶素子１０、同一極性、同一電圧の複数電圧パルスの連続印加により徐々にコンダクタンスが変化する第２の抵抗変化型不揮発性記憶素子２０とセルトランジスタＴ０が接続されたもので構成され、１つのセルトランジスタＴ０と２つの抵抗変化型不揮発性記憶素子から構成される『１Ｔ２Ｒ』型のメモリセルである。メモリセルＭＣのワード線ＷＬはセルトランジスタＴ０のゲート端子に接続され、ビット線ＢＬ１は第１の抵抗変化型不揮発性記憶素子１０に接続され、ビット線ＢＬ２は第２の抵抗変化型不揮発性記憶素子２０に接続され、ソース線ＳＬはセルトランジスタＴ０のソース端子に接続される。

[0070] 図７Ｂは、図７Ａに示される第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の両方が搭載された人工知能処理装置のメモリセルの断面図の一例を示したものである。半導体の基板７０上に拡散領域７１ａ、７１ｂが形成されており、拡散領域７１ａがセルトランジスタＴ０のソース端子として、拡散領域７１ｂがセルトランジスタのドレイン端子として作用する。拡散領域７１ａ、７１ｂ間がセルトランジスタＴ０のチャネル領域として作用し、このチャネル領域上に酸化膜７２、ポリシリコンで形成されるゲート電極７３が形成され、セルトランジスタＴ０として動作する。セルトランジスタＴ０のソース端子である拡散領域７１ａはビア７４ａを介して第１配線層７５ａであるソース線ＳＬに接続される。セルトランジスタＴ０のドレイン端子である拡散領域７１ｂはビア７４ｂを介して第１配線層７５ｂに接続される。さらに、第１配線層７５ｂはビア７６ａを介して第２配線層７７に接続され、第２配線層７７はビア７８ａを介して第１の抵抗変化型不揮発性記憶素子１０に接続される。第１の抵抗変化型不揮発性記憶素子１０は第１電極２、抵抗変化層３、第２電極４から構成され

る。第１の抵抗変化型不揮発性記憶素子１０はビア７８ｂを介して第３配線層７９であるビット線ＢＬ１に接続される。また、同時に、第２配線層７７はビア７６ｃを介して第２の抵抗変化型不揮発性記憶素子２０に接続される。第２の抵抗変化型不揮発性記憶素子２０は第１電極１２、抵抗変化層１３、第２電極１４から構成される。第２の抵抗変化型不揮発性記憶素子２０はビア７６ｂを介して第１配線層７５ｃであるビット線ＢＬ２に接続される。

[0071] 図７Ｃは、図７Ａに示される第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の両方が搭載された人工知能処理装置のメモリセルの断面図の、先ほど（つまり、図７Ｂに示される例）と異なる例を示したものである。半導体の基板８０上に拡散領域８１ａ、８１ｂが形成されており、拡散領域８１ａがセルトランジスタＴ０のソース端子として、拡散領域８１ｂがセルトランジスタのドレイン端子として作用する。拡散領域８１ａ、８１ｂ間がセルトランジスタＴ０のチャネル領域として作用し、このチャネル領域上に酸化膜８２、ポリシリコンで形成されるゲート電極８３が形成され、セルトランジスタＴ０として動作する。セルトランジスタＴ０のソース端子である拡散領域８１ａはビア８４ａを介して第１配線層８５ａであるソース線ＳＬに接続される。セルトランジスタＴ０のドレイン端子である拡散領域８１ｂはビア８４ｂを介して第１配線層８５ｂに接続される。さらに、第１配線層８５ｂはビア８６ｃを介して第２配線層８７ｂに接続され、第２配線層８７ｂはビア８８ａを介して第１の抵抗変化型不揮発性記憶素子１０に接続される。第１の抵抗変化型不揮発性記憶素子１０は第１電極２、抵抗変化層３、第２電極４から構成される。第１の抵抗変化型不揮発性記憶素子１０はビア８８ｂを介して第３配線層８９であるビット線ＢＬ１に接続される。また、同時に、第１配線層８５ｂはビア８６ａを介して第２の抵抗変化型不揮発性記憶素子２０に接続される。第２の抵抗変化型不揮発性記憶素子２０は第１電極１２、抵抗変化層１３、第２電極１４から構成される。第２の抵抗変化型不揮発性記憶素子２０はビア８６ｂを介して第２配線層８７ａであるビット線ＢＬ２に接続される。この配置においては

、基板平面と鉛直の方向から見て、第１の抵抗変化型不揮発性記憶素子１０の中心と第２の抵抗変化型不揮発性記憶素子２０の中心とが一致しない配置となる。

[0072] なお、図７Ｂおよび図７Ｃにおいては、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０を第１配線層と第３配線層の間の異なる配線層間に配置したが、他の配線層間、例えば第２配線層と第４配線層の間の異なる配線層間に配置しても同様の効果が得られる。

[0073] 従来例として記載した図６Ａにおいては、ビット線ＢＬあるいはソース線ＳＬを流れる電流、すなわち抵抗変化型不揮発性記憶素子ＲＰを流れる電流がメモリセルＭＣを流れる電流として定義されるが、本実施形態にかかる図７Ａにおいては、メモリセルＭＣを流れる電流は、ビット線ＢＬ１とビット線ＢＬ２を流れる電流の合計、あるいはソース線ＳＬを流れる電流、すなわち、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０を流れる電流の合計として定義される。

[0074] 図７Ａ～図７Ｃに示されるメモリセルＭＣでは、ワード線ＷＬにおける信号が１つのニューロンに入力される入力信号に相当し、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスの合計が、その入力信号に対応する１つの結合重み係数に相当し、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０を流れる電流の合計（つまり、ビット線ＢＬ１とビット線ＢＬ２を流れる電流の合計、あるいはソース線ＳＬを流れる電流）が入力信号と結合重み係数との積に相当する。

[0075] 図８Ａ～図８Ｃは、別の実施形態に係る抵抗変化型不揮発性記憶素子の回路図、断面図の例を示す図である。

[0076] 図８Ａは、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の両方が搭載された人工知能処理装置のメモリセルの回路図を示したものである。メモリセルＭＣ１は単一の電圧パルスによりコンダクタンスが設定可能な第１の抵抗変化型不揮発性記憶素子１０とセル

トランジスタＴ１が接続されたもので構成され、メモリセルＭＣ２は同一極性、同一電圧の複数電圧パルスの連続印加により徐々にコンダクタンスが変化する第２の抵抗変化型不揮発性記憶素子２０とセルトランジスタＴ２が接続されたもので構成され、２つのセルトランジスタＴ１およびＴ２と２つの抵抗変化型不揮発性記憶素子から構成される『２Ｔ２Ｒ』型のメモリセルである。メモリセルＭＣ１のワード線ＷＬ１はセルトランジスタＴ１のゲート端子に接続され、メモリセルＭＣ２のワード線ＷＬ２はセルトランジスタＴ２のゲート端子に接続され、ビット線ＢＬは第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０に接続され、ソース線ＳＬはセルトランジスタＴ１およびＴ２のソース端子に接続される。

[0077] 図８Ｂは、図８Ａに記載された、第１の抵抗変化型不揮発性記憶素子１０が搭載された人工知能処理装置のメモリセルＭＣ１の断面図の一例を示したものである。半導体の基板９０上に拡散領域９１ａ、９１ｂが形成されており、拡散領域９１ａがセルトランジスタＴ１のソース端子として、拡散領域９１ｂがセルトランジスタＴ１のドレイン端子として作用する。拡散領域９１ａ、９１ｂ間がセルトランジスタＴ１のチャネル領域として作用し、このチャネル領域上に酸化膜９２、ポリシリコンで形成されるゲート電極９３が形成され、セルトランジスタＴ１として動作する。セルトランジスタＴ１のソース端子である拡散領域９１ａはビア９４ａを介して第１配線層９５ａであるソース線ＳＬに接続される。セルトランジスタＴ１のドレイン端子である拡散領域９１ｂはビア９４ｂを介して第１配線層９５ｂに接続される。さらに、第１配線層９５ｂはビア９６を介して第２配線層９７に接続され、第２配線層９７はビア９８ａを介して第１の抵抗変化型不揮発性記憶素子１０に接続される。第１の抵抗変化型不揮発性記憶素子１０は第１電極２、抵抗変化層３、第２電極４から構成される。第１の抵抗変化型不揮発性記憶素子１０はビア９８ｂを介して第３配線層９９であるビット線ＢＬに接続される。

[0078] 図８Ｃは、図８Ａに記載された、第２の抵抗変化型不揮発性記憶素子２０

が搭載された人工知能処理装置のメモリセルＭＣ２の断面図の一例を示したものである。図８Ｂと共通の基板９０上に拡散領域１０１ａ、１０１ｂが形成されており、拡散領域１０１ａがセルトランジスタＴ２のソース端子として、拡散領域１０１ｂがセルトランジスタＴ２のドレイン端子として作用する。拡散領域１０１ａ、１０１ｂ間がセルトランジスタＴ２のチャネル領域として作用し、このチャネル領域上に酸化膜１０２、ポリシリコンで形成されるゲート電極１０３が形成され、セルトランジスタＴ２として動作する。セルトランジスタＴ２のソース端子である拡散領域１０１ａはビア１０４ａを介して図８Ｂと共通の第１配線層９５ａであるソース線ＳＬに接続される。セルトランジスタＴ２のドレイン端子である拡散領域１０１ｂはビア１０４ｂを介して第１配線層１０５に接続される。第１配線層１０５はビア１０６ａを介して第２の抵抗変化型不揮発性記憶素子２０に接続される。第２の抵抗変化型不揮発性記憶素子２０は第１電極１２、抵抗変化層１３、第２電極１４から構成される。第２の抵抗変化型不揮発性記憶素子２０はビア１０６ｂを介して第２配線層１０７に接続される。第２配線層１０７はビア１０８を介して図８Ｂと共通の第３配線層９９であるビット線ＢＬに接続される。

[0079] なお、図８Ｂおよび図８Ｃにおいては、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０を第１配線層と第３配線層の間の異なる配線層間に配置したが、他の配線層間、例えば第２配線層と第４配線層の間の異なる配線層間に配置しても同様の効果が得られる。

[0080] 従来例として記載した図６Ａにおいては、ビット線ＢＬあるいはソース線ＳＬを流れる電流、すなわち抵抗変化型不揮発性記憶素子ＲＰを流れる電流がメモリセルＭＣを流れる電流として定義されるが、本実施形態にかかる図８Ａにおいては、ビット線ＢＬあるいはソース線ＳＬを流れる電流は、メモリセルＭＣ１を流れる電流とメモリセルＭＣ２を流れる電流の合計、すなわち、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０を流れる電流の合計として定義される。

[0081] 図８Ａ～図８Ｃに示されるメモリセルＭＣ１とメモリセルＭＣ２とを合わせたメモリセルに着目すると、ワード線ＷＬ１およびＷＬ２に共通に流れる信号が１つのニューロンに入力される入力信号に相当し、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスの合計が、その入力信号に対応する１つの結合重み係数に相当し、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０を流れる電流の合計（つまり、ビット線ＢＬを流れる電流、あるいはソース線ＳＬを流れる電流）が入力信号と結合重み係数との積に相当する。

[0082] 図７Ａ～Ｃおよび図８Ａ～Ｃで示されるように、単一の電圧パルスにより設定コンダクタンスに概ね到達する第１の抵抗変化型不揮発性記憶素子１０と、同一極性、同一電圧の複数電圧パルスの連続印加により徐々にコンダクタンスが変化する第２の抵抗変化型不揮発性記憶素子２０とを両方搭載した実施形態に係る人工知能処理装置においては、両方の素子が使用できるため、第１の抵抗変化型不揮発性記憶素子１０は、製品出荷前および製品出荷後のファームウェア更新、学習モデルの更新、定期メンテナンス、および第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスの変更では結合重み係数更新が不十分な時にコンダクタンスを変更し、第２の抵抗変化型不揮発性記憶素子２０は、製品出荷後の学習のための結合重み係数の更新時にコンダクタンスを変更するという使用方法を取ることにより、製品出荷時等における高効率・高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が可能になる。つまり、結合重み係数を変更する（つまり、書き込む）際には、その際の用途に応じて、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の一方だけに対してコンダクタンスを変更し、人工知能処理装置を使用して推論する（つまり、結合重み係数を読み出す）際には、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の両方に流れる電流の合計値を用いる。

[0083] また、人工知能処理装置において、既存のニューラルネットワークの結合重み係数の設定を用いた転移学習あるいは強化学習に適用して既存のニューラルネットワークの結合重み係数設定をそのまま用いる第１のニューラルネットワーク領域と、新たに学習を行う第２のニューラルネットワーク領域とがある場合には、第１のニューラルネットワーク領域の結合重み係数の設定には第１の抵抗変化型不揮発性記憶素子１０に対してコンダクタンスを変更し、第２のニューラルネットワーク領域の結合重み係数の設定には第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスを変更するという使用方法により、効率的に転移学習あるいは強化学習を行うことが出来る。

[0084] ［実施例］

次に、本開示に係る人工知能処理装置の実施例について、説明する。

[0085] 図９Ａは、実施例に係る人工知能処理装置２００のモデルを示すブロック図である。人工知能処理装置２００は、入力層２０１、複数の隠れ層２０２および出力層２０３で構成されるニューラルネットワークである。各層（入力層２０１、隠れ層２０２、出力層２０３）は、複数のニューロン２１０で構成される。各ニューロン２１０には、前段の層を構成するニューロン２１０からの出力データがシナプス２１１を経て入力される。

[0086] 図９Ｂは、図９Ａに示されるニューロン２１０の機能を示す図である。ニューロン２１０は、前段の層を構成するニューロン２１０からの出力データを、シナプス２１１を経て入力データx_iとして受け取り、受け取った入力データx_iに、そのシナプス２１１に対応する結合重み係数w_iを乗じた積（$w_i \cdot x_i$）を、全ての入力データx_iについて加算する積和演算（$\Sigma w_i \cdot x_i$）を行う。そして、ニューロン２１０は、積和演算の結果に対して、内部に有するバイアスｂを加算し、得られた結果（$\Sigma w_i \cdot x_i + b$）を、内部に有するステップ関数等の活性化関数ｆに入力することで、出力データｙを生成し、次の段の層を構成する各ニューロン２１０に出力する。

[0087] 図１０Ａは、図９Ｂに示されるニューロン２１０を実現する回路例を示す図である。本図では、２次元上に配置されたメモリセルＭＣｉｊ（ｉ＝０～

ｍ、ｊ＝０～ｎ）として、図７Ａに示される１Ｔ２Ｒ型のメモリセルＭＣが用いられた回路例が示されている。ニューロン２１０は、積和演算回路２１５、ワード線選択回路２３０、判定回路２５０、および、カラムゲート（トランジスタＹＴｉ１、ＹＴｉ２、トランジスタＤＴｉ）で構成される。

[0088] 積和演算回路２１５は、ニューロン２１０における積和演算を行う回路であり、図７Ａに示される１Ｔ２Ｒ型のメモリセルが、複数、列方向に並べられ、かつ、ビット線ＢＬｉ１、ＢＬｉ２、ソース線ＳＬｉのそれぞれを共通にするように、接続されている。各メモリセルから出力される電流が、それぞれ、入力データx_iと結合重み係数w_iとの積（$w_i \cdot x_i$）に相当し、それらの電流が合流し、ソース線ＳＬｉに流れる電流（あるいは、ビット線ＢＬｉ１およびＢＬｉ２に流れる電流の合計）が積和演算の結果$\Sigma w_i \cdot x_i$に相当する。この積和演算回路２１５では、入力データx_iは、「１」または「０」であり、結合重み係数w_iは、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスの合計に対応する。

[0089] なお、本図には、図９Ａに示される全てのニューロン２１０に相当する複数の積和演算回路２１５が行方向に並べられて配置されており、それらのうち、１つの積和演算回路２１５が、１つのニューロン２１０における積和演算を行う。

[0090] ワード線選択回路２３０は、積和演算回路２１５を構成するメモリセルＭＣｉ０～ＭＣｉｎに含まれるトランジスタＴｉのゲート端子に、ワード線ＷＬ０～ＷＬｎを介して、行単位で、メモリセルを選択または非選択にするための入力データｘ０～ｘｎを供給する回路である。

[0091] 判定回路２５０は、ニューロン２１０が有する活性化関数ｆを実行する回路であり、積和演算回路２１５から出力される積和演算の結果を示すソース線ＳＬｉに流れる電流、あるいは、ビット線ＢＬｉ１およびＢＬｉ２に流れる電流の合計（$\Sigma w_i \cdot x_i$）に、内部に有するバイアスｂを加算した値（$\Sigma w_i \cdot x_i + b$）と、所定の閾値とを比較して出力する。判定回路２５０は、

行方向に並ぶ複数の積和演算回路２１５に対して、並列に処理を行うことができる。

[0092] カラムゲートを構成するトランジスタＹＴｉ１、ＹＴｉ２は、メモリセルＭＣｉ０～ＭＣｉｎに対する書き込みおよび読み出し時において、ゲート端子に入力される信号に応じて、それぞれ、メモリセルＭＣｉ０～ＭＣｉｎのビット線ＢＬｉ１、ＢＬｉ２と、所定の電源電圧とを接続したり非接続にしたりする。また、トランジスタＤＴｉは、メモリセルＭＣｉ０～ＭＣｉｎに対する書き込みおよび読み出し時においてゲート端子に入力される信号に応じて、ソース線ＳＬｉと所定の電源電圧とを接続したり非接続にしたりする。

[0093] 図１０Ｂは、図９Ｂに示されるニューロン２１０を実現する別の回路例（ニューロン２１０ａ）を示す図である。本図では、２次元上に配置されたメモリセルＭＣｉｊ（ｉ＝０～ｍ、ｊ＝０～ｎ）として、図８Ａに示される２Ｔ２Ｒ型のメモリセルＭＣが複数、列方向に並べられ、かつ、ビット線ＢＬｉ、ソース線ＳＬｉのそれぞれを共通にするように、接続されている。ニューロン２１０ａは、積和演算回路２１５ａ、ワード線選択回路２３０ａ、判定回路２５０ａ、および、カラムゲート（トランジスタＹＴｉ、トランジスタＤＴｉ）で構成される。

[0094] このニューロン２１０ａは、図１０Ａに示されるニューロン２１０と、基本的な機能は同じであるが、１Ｔ２Ｒ型のメモリセルＭＣｉｊで構成される図１０Ａのニューロン２１０と異なり、２Ｔ２Ｒ型のメモリセルＭＣｉｊに対応した接続構成となっている。ワード線選択回路２３０ａは、メモリセルＭＣｉ０～ＭＣｉｎのそれぞれに対して、２つのワード線ＷＬｊ１およびＷＬｊ２を出力する。判定回路２５０ａは、積和演算回路２１５ａから出力される積和演算の結果を示すビット線ＢＬｉ、あるいは、ソース線ＳＬｉに流れる電流（$\Sigma w_i \cdot x_i$）に、内部に有するバイアスｂを加算した値（$\Sigma w_i \cdot x_i + b$）と、所定の閾値とを比較して出力する。カラムゲートについては、各ビット線ＢＬｉついて、所定の電源電圧との接続および非接続を切り替え

る１個のトランジスタＹＴｉが設けられる。

[0095] 図１１は、図１０Ａに示されるニューロン２１０からなる人工知能処理装置２００の全体の構成を示すブロック図である。人工知能処理装置２００は、メモリセルアレイ２２０、ワード線選択回路２３０、カラムゲート２４０、判定回路２５０、書き込み回路２６０および制御回路２７０で構成される。

[0096] メモリセルアレイ２２０は、図１０Ａに示されるメモリセルＭＣｉｊ（ｉ＝０～ｍ、ｊ＝０～ｎ）に相当する。ワード線選択回路２３０、カラムゲート２４０、判定回路２５０は、図１０Ａで説明されたものと同じである。書き込み回路２６０は、図１０Ａで説明された所定の電源電圧を供給する回路であり、メモリセルＭＣｉｊに対して所望のコンダクタンス（言い換えれば、多階調のアナログ抵抗値）を書き込むための電流制限回路を有する。

[0097] 制御回路２７０は、人工知能処理装置２００の全体を制御することで、メモリセルＭＣｉｊに対する書き込みおよび読み出しを制御する回路であり、例えば、プログラムを格納したメモリおよびプロセッサ等で構成される。より詳しくは、制御回路２７０は、人工知能処理装置２００の結合重み係数を変更する（つまり、書き込む）際には、製品出荷時等における高精度の結合重み係数の設定（初期設定）か製品出荷後等における高効率な結合重み係数の更新（学習）かの用途に応じて、各メモリセルＭＣｉｊを構成する第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の一方だけに対してコンダクタンスを変更するように、人工知能処理装置２００を制御する。また、制御回路２７０は、人工知能処理装置２００を使用して（つまり、結合重み係数を用いて）推論を行う際には、各メモリセルＭＣｉｊを構成する第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０に流れる電流の合計値を用いるように、人工知能処理装置２００を制御する。

[0098] なお、図１１は、図１０Ａに示されるニューロン２１０からなる人工知能処理装置２００の全体の構成を示すブロック図であるが、図１０Ｂに示され

るニューロン２１０ａからなる人工知能処理装置２００の全体の構成を示すブロック図については、上述した接続配線等を除いて、図１１と同様であるので、図示及びその説明を省略する。

[0099] 図１２Ａは、図１０Ａに示されるメモリセルＭＣｉｊに対する書き込みおよび読み出し時におけるワード線ＷＬ、ビット線ＢＬ１、ＢＬ２、および、ソース線ＳＬを介した印加電圧の例を示す図である。ここでは、第ｉカラムのメモリセルＭＣｉｊを構成する第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０をリセット（高抵抗化）する場合、セット（低抵抗化）する場合、および、メモリセルＭＣｉｊから読み出す場合における印加電圧の例が示されている。図１２Ａにおいて、「モード１」は、第１の抵抗変化型不揮発性記憶素子１０へのコンダクタンスの書き込み（第１の書き換えステップ）を示し、「モード２」は、第２の抵抗変化型不揮発性記憶素子２０へのコンダクタンスの書き込み（第２の書き換えステップ）を示す。なお、「メモリセルＭＣｉｊに対する書き込み」とは、メモリセルＭＣｉｊに結合重み係数を設定または変更することを意味し、「メモリセルＭＣｉｊに対する読み出し」とは、メモリセルＭＣｉｊを流れる電流を計測することを意味する。

[0100] （１）メモリセルＭＣｉｊに書き込む（結合重み係数を設定または変更する）場合

第１の抵抗変化型不揮発性記憶素子１０をリセット（「高抵抗化」）する場合には（「モード１」）、ワード線ＷＬを介してトランジスタＴｉをオンにするパルス電圧V_{g_on}（例えば、２Ｖ）がゲート端子に供給され、ビット線ＢＬ１にリセット電圧V_{reset}（例えば、２Ｖ）が印加され、ビット線ＢＬ２に基準電圧V_{ss}（例えば、０Ｖ）が印加され、ソース線ＳＬに基準電圧V_{ss}（例えば、０Ｖ）が印加される。これにより、第１の抵抗変化型不揮発性記憶素子１０だけに対して、第１の抵抗変化型不揮発性記憶素子１０の下方の端子を基準として上方の端子に正電圧が印加され、書き込み回路２６０に内蔵される電流制限回路による制限電流に応じた抵抗値に高抵抗化される。

[0101] 第２の抵抗変化型不揮発性記憶素子２０をリセット（「高抵抗化」）する場合には（「モード２」）、ワード線ＷＬを介してトランジスタＴｉをオンにするパルス電圧V_{g_on}（例えば、２Ｖ）がゲート端子に供給され、ビット線ＢＬ１に基準電圧V_{ss}（例えば、０Ｖ）が印加され、ビット線ＢＬ２にリセット電圧V_{reset}（例えば、２Ｖ）が印加され、ソース線ＳＬに基準電圧V_{ss}（例えば、０Ｖ）が印加される。これにより、第２の抵抗変化型不揮発性記憶素子２０だけに対して、第２の抵抗変化型不揮発性記憶素子２０の下方の端子を基準として上方の端子に正電圧が印加され、書き込み回路２６０に内蔵される電流制限回路による制限電流に応じた抵抗値に高抵抗化される。

[0102] 第１の抵抗変化型不揮発性記憶素子１０をセット（「低抵抗化」）する場合には（「モード１」）、ワード線ＷＬを介してトランジスタＴｉをオンにするパルス電圧V_{g_on}（例えば、２Ｖ）がゲート端子に供給され、ビット線ＢＬ１に基準電圧V_{ss}（例えば、０Ｖ）が印加され、ビット線ＢＬ２にセット電圧V_{set}（例えば、２Ｖ）が印加され、ソース線ＳＬにセット電圧V_{set}（例えば、２Ｖ）が印加される。これにより、第１の抵抗変化型不揮発性記憶素子１０だけに対して、第１の抵抗変化型不揮発性記憶素子１０の下方の端子を基準として上方の端子に負電圧が印加され、書き込み回路２６０に内蔵される電流制限回路による制限電流に応じた抵抗値に低抵抗化される。

[0103] 第２の抵抗変化型不揮発性記憶素子２０をセット（「低抵抗化」）する場合には（「モード２」）、ワード線ＷＬを介してトランジスタＴｉをオンにするパルス電圧V_{g_on}（例えば、２Ｖ）がゲート端子に供給され、ビット線ＢＬ１にセット電圧V_{set}（例えば、２Ｖ）が印加され、ビット線ＢＬ２に基準電圧V_{ss}（例えば、０Ｖ）が印加され、ソース線ＳＬにセット電圧V_{set}（例えば、２Ｖ）が印加される。これにより、第２の抵抗変化型不揮発性記憶素子２０だけに対して、第２の抵抗変化型不揮発性記憶素子２０の下方の端子を基準として上方の端子に負電圧が印加され、書き込み回路２６０に内蔵される電流制限回路による制限電流に応じた抵抗値に低抵抗化される。

[0104] （２）メモリセルＭＣｉｊから読み出す場合（推論ステップ）

第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０を流れる電流の合計を計測する場合には（「読み出し」）、ワード線ＷＬを介してトランジスタＴｉをオンにする読み出し電圧$Ｖ_{g_read}$（例えば、１Ｖ）がゲート端子に供給され、ビット線ＢＬ１およびＢＬ２に読み出し電圧$Ｖ_{read}$（例えば、０．４Ｖ）が印加され、ソース線ＳＬに基準電圧$Ｖ_{Vss}$（例えば、０Ｖ）が印加される。これにより、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０に対して、読み出し電圧$Ｖ_{read}$が印加され、メモリセルＭＣｉｊから、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０を流れる電流の合計（つまり、一つの積（$ｗ_{i}$・$ｘ_{i}$）が出力される。よって、積和演算回路２１５を構成する全てのメモリセルＭＣｉｊから出力される電流がソース線ＳＬに流れ（ビット線ＢＬ１およびＢＬ２に流れる電流の合計でもある）、積和演算の結果（Σ$ｗ_{i}$・$ｘ_{i}$）として、判定回路２５０で計測される。

[0105] 図１２Ｂは、図１０Ｂに示されるメモリセルＭＣｉｊに対する書き込みおよび読み出し時におけるワード線ＷＬ１、ＷＬ２、ビット線ＢＬ、および、ソース線ＳＬを介した印加電圧の例を示す図である。ここでは、第ｉカラムのメモリセルＭＣｉｊを構成する第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０をリセット（高抵抗化）する場合、セット（低抵抗化）する場合、および、メモリセルＭＣｉｊから読み出す場合における印加電圧の例が示されている。

[0106] （１）メモリセルＭＣｉｊに書き込む（結合重み係数を設定または変更する）場合

第１の抵抗変化型不揮発性記憶素子１０をリセット（「高抵抗化」）する場合には（「モード１」）、ワード線ＷＬ１を介してトランジスタＴｉ１をオンにするパルス電圧$Ｖ_{g_on}$（例えば、２Ｖ）がゲート端子に供給され、ワード線ＷＬ２を介してトランジスタＴｉ２をオフにするパルス電圧$Ｖ_{g_off}$（例えば、０Ｖ）がゲート端子に供給され、ビット線ＢＬにリセット電圧$Ｖ_{res}$

$_{et}$（例えば、２Ｖ）が印加され、ソース線ＳＬに基準電圧$_{Vss}$（例えば、０Ｖ）が印加される。これにより、第１の抵抗変化型不揮発性記憶素子１０だけに対して、第１の抵抗変化型不揮発性記憶素子１０の下方の端子を基準として上方の端子に正電圧が印加され、書き込み回路２６０に内蔵される電流制限回路による制限電流に応じた抵抗値に高抵抗化される。

[0107] 第２の抵抗変化型不揮発性記憶素子２０をリセット（「高抵抗化」）する場合には（「モード２」）、ワード線ＷＬ１を介してトランジスタＴｉ１をオフにするパルス電圧V_{g_off}（例えば、０Ｖ）がゲート端子に供給され、ワード線ＷＬ２を介してトランジスタＴｉ２をオンにするパルス電圧V_{g_on}（例えば、２Ｖ）がゲート端子に供給され、ビット線ＢＬにリセット電圧V_{reset}（例えば、２Ｖ）が印加され、ソース線ＳＬに基準電圧$_{Vss}$（例えば、０Ｖ）が印加される。これにより、第２の抵抗変化型不揮発性記憶素子２０だけに対して、第２の抵抗変化型不揮発性記憶素子２０の下方の端子を基準として上方の端子に正電圧が印加され、書き込み回路２６０に内蔵される電流制限回路による制限電流に応じた抵抗値に高抵抗化される。

[0108] 第１の抵抗変化型不揮発性記憶素子１０をセット（「低抵抗化」）する場合には（「モード１」）、ワード線ＷＬ１を介してトランジスタＴｉ１をオンにするパルス電圧V_{g_on}（例えば、２Ｖ）がゲート端子に供給され、ワード線ＷＬ２を介してトランジスタＴｉ２をオフにするパルス電圧V_{g_off}（例えば、０Ｖ）がゲート端子に供給され、ビット線ＢＬに基準電圧$_{Vss}$（例えば、０Ｖ）が印加され、ソース線ＳＬにセット電圧V_{set}（例えば、２Ｖ）が印加される。これにより、第１の抵抗変化型不揮発性記憶素子１０だけに対して、第１の抵抗変化型不揮発性記憶素子１０の下方の端子を基準として上方の端子に負電圧が印加され、書き込み回路２６０に内蔵される電流制限回路による制限電流に応じた抵抗値に低抵抗化される。

[0109] 第２の抵抗変化型不揮発性記憶素子２０をセット（「低抵抗化」）する場合には（「モード２」）、ワード線ＷＬ１を介してトランジスタＴｉ１をオフにするパルス電圧V_{g_off}（例えば、０Ｖ）がゲート端子に供給され、ワー

ド線ＷＬ２を介してトランジスタＴｉ２をオンにするパルス電圧V_{g_on}（例えば、２Ｖ）がゲート端子に供給され、ビット線ＢＬに基準電圧V_{ss}（例えば、０Ｖ）が印加され、ソース線ＳＬにセット電圧V_{set}（例えば、２Ｖ）が印加される。これにより、第２の抵抗変化型不揮発性記憶素子２０だけに対して、第２の抵抗変化型不揮発性記憶素子２０の下方の端子を基準として上方の端子に負電圧が印加され、書き込み回路２６０に内蔵される電流制限回路による制限電流に応じた抵抗値に低抵抗化される。

[0110] （２）メモリセルＭＣｉｊから読み出す場合（推論ステップ）

第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０を流れる電流の合計を計測する場合には（「読み出し」）、ワード線ＷＬ１およびＷＬ２を介してトランジスタＴｉ１およびＴｉ２をオンにする読み出し電圧V_{g_read}（例えば、１Ｖ）がゲート端子に供給され、ビット線ＢＬに読み出し電圧V_{read}（例えば、０．４Ｖ）が印加され、ソース線ＳＬに基準電圧V_{ss}（例えば、０Ｖ）が印加される。これにより、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０に対して、読み出し電圧V_{read}が印加され、メモリセルＭＣｉｊから、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０を流れる電流（その合計が一つの積（$w_i \cdot x_i$）に相当する）が出力される。よって、積和演算回路２１５を構成する全てのメモリセルＭＣｉｊから出力される電流がソース線ＳＬおよびビット線ＢＬに流れ、積和演算の結果（$\Sigma w_i \cdot x_i$）として、判定回路２５０で計測される。

[0111] 図１３は、図１１に示される制御回路２７０の動作例を示すフローチャートである。

[0112] 制御回路２７０は、今から行う結合重み係数の設定処理が、初期設定のために結合重み係数を変更する第１のケースであるか、学習において結合重み係数を変更する第２のケースであるかを判断する（Ｓ３０）。なお、第１のケースには、人工知能処理装置２００の出荷前および出荷後のファームウェア更新、学習モデルの更新、定期メンテナンス、および第２の抵抗変化型不

揮発性記憶素子２０のコンダクタンスの変更では結合重み係数の更新が不十分な時のうちの少なくとも一つが含まれ、第２のケースには、人工知能処理装置２００の出荷後の学習のための結合重み係数の更新時が含まれる。

[0113] その結果、結合重み係数の設定処理が初期設定（つまり、第１のケース）であると判断した場合には（Ｓ３０で「初期設定」）、制御回路２７０は、カラムゲート２４０またはワード線選択回路２３０を制御することで、メモリセルＭＣｉｊのそれぞれについて、第１の抵抗変化型不揮発性記憶素子１０を選択し（Ｓ３１）、事前に導出した結合重み係数を設定する（Ｓ３２）。つまり、制御回路２７０は、メモリセルＭＣｉｊのそれぞれについて、第１の抵抗変化型不揮発性記憶素子１０に対して（つまり、「モード１」で）、事前に導出したコンダクタンスを書き込む（第１の書き換えステップ）。

[0114] 一方、ステップＳ３０において、結合重み係数の設定処理が学習による結合重み係数の更新（つまり、第２のケース）であると判断した場合には（Ｓ３０で「学習による結合重み係数の更新」）、制御回路２７０は、現在のメモリセルアレイ２２０に保持された結合重み係数を用いて推論を実施し（Ｓ３５）、その推論結果と教師ラベルとの差を確認した後に（Ｓ３６）、メモリセルＭＣｉｊのそれぞれについて、更新時の結合重み係数の変化量を算出する（Ｓ３７）。そして、制御回路２７０は、カラムゲート２４０またはワード線選択回路２３０を制御することで、メモリセルＭＣｉｊのそれぞれについて、第２の抵抗変化型不揮発性記憶素子２０を選択し（Ｓ３８）、算出した結合重み係数の変化量だけ現在の結合重み係数が変化するように結合重み係数を更新する（Ｓ３９）。つまり、制御回路２７０は、メモリセルＭＣｉｊのそれぞれについて、第２の抵抗変化型不揮発性記憶素子２０に対して（つまり、「モード２」で）、変化量の分だけコンダクタンスを更新する（第２の書き換えステップ）。

[0115] これにより、初期設定か学習かの用途に応じて、その用途に適した第１の抵抗変化型不揮発性記憶素子１０または第２の抵抗変化型不揮発性記憶素子２０だけに対して書き込みが行われ、製品出荷時等における高精度の結合重

み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が実現される。

[0116] なお、図１３では、各ステップは、制御回路２７０によって実行されたが、それらの一部または全部のステップが、人工知能処理装置２００の外部に置かれた別のプロセッサ等の制御回路によって実行されてもよい。

[0117] 以上のように、本開示に係る人工知能処理装置２００は、１つの基板１等と、１つの基板１等に搭載され、積和演算を行う積和演算回路２１５とを備え、積和演算回路２１５は、積和演算における積の演算に用いられる結合重み係数をコンダクタンスとして保持する、特性の異なる第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０とを含み、同一極性かつ同一電圧の電圧パルスを複数回連続して印加されると、第１の抵抗変化型不揮発性記憶素子１０における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率が、第２の抵抗変化型不揮発性記憶素子２０における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率よりも小さい。

[0118] これにより、連続する電圧パルスの印加における初回の電圧パルスによってコンダクタンスが大きく変化し、その後の電圧パルスでは変化量が小さいという特性を有する第１の抵抗変化型不揮発性記憶素子１０を、結合重み係数の設定（初期設定）に用い、一方、連続する電圧パルスによって徐々にコンダクタンスが変化する特性を有する第２の抵抗変化型不揮発性記憶素子２０を、結合重み係数の更新（学習）に用いることができる。よって、製品出荷時等における高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が可能な抵抗変化型不揮発性記憶素子を用いた人工知能処理装置が実現される。

[0119] また、本開示に係る人工知能処理装置２００は、１つの基板１等と、１つの基板１等に搭載され、積和演算を行う積和演算回路２１５とを備え、積和演算回路２１５は、第１電極２等、第２電極４等、および第１電極２等と第

２電極４等との間に挟まれた抵抗変化層３等を含み、積和演算における積の演算に用いられる結合重み係数をコンダクタンスとして保持する複数の抵抗変化型不揮発性記憶素子を有し、複数の抵抗変化型不揮発性記憶素子は、特性の異なる第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０とを含み、第１の抵抗変化型不揮発性記憶素子１０において第１電極２等と第２電極４等のうち少なくとも片方の電極が貴金属電極であり、第２の抵抗変化型不揮発性記憶素子２０において第１電極２等と第２電極４等のうち少なくとも片方の電極が非貴金属電極である。

[0120] これにより、同一極性かつ同一電圧の電圧パルスを複数回連続して印加されると、第１の抵抗変化型不揮発性記憶素子１０における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率が、第２の抵抗変化型不揮発性記憶素子２０における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率よりも小さいという特性が得られる。よって、第１の抵抗変化型不揮発性記憶素子１０を結合重み係数の設定（初期設定）に用い、一方、第２の抵抗変化型不揮発性記憶素子２０を結合重み係数の更新（学習）に用いることで、製品出荷時等における高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が可能な抵抗変化型不揮発性記憶素子を用いた人工知能処理装置が実現される。

[0121] なお、具体的には、上記貴金属電極はＩｒ、Ｐｔのうちいずれか１つを含み、上記非貴金属電極はＴｉＮ、ＴａＮのうちいずれか１つを含む。

[0122] また、本開示に係る人工知能処理装置２００は、１つの基板１等と、１つの基板１等に搭載され、積和演算を行う積和演算回路２１５とを備え、積和演算回路２１５は、積和演算における積の演算に用いられる結合重み係数をコンダクタンスとして保持する、特性の異なる第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０とを含み、第１の抵抗変化型不揮発性記憶素子１０は、第２の抵抗変化型不揮発性記憶素子２０よ

りも高いリテンション特性を有し、第２の抵抗変化型不揮発性記憶素子２０は、第１の抵抗変化型不揮発性記憶素子１０よりも高いエンデュランス特性を有する。

[0123] これにより、書き込み後のリテンション特性が高い第１の抵抗変化型不揮発性記憶素子１０を、結合重み係数の設定（初期設定）に用い、一方、書き込みによる素子劣化度合いが小さい、エンデュランス特性が高い第２の抵抗変化型不揮発性記憶素子２０を、結合重み係数の更新（学習）に用いることができる。よって、製品出荷時等における高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が可能な抵抗変化型不揮発性記憶素子を用いた人工知能処理装置が実現される。

[0124] ここで、積和演算回路２１５は、第１の抵抗変化型不揮発性記憶素子１０に流れる電流および第２の抵抗変化型不揮発性記憶素子２０に流れる電流を合計し、得られた電流の合計を、積和演算における１つの積として、出力する。これにより、第１の抵抗変化型不揮発性記憶素子１０のコンダクタンスと第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスとによって一つの結合重み係数が構成され、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０とを含むメモリセルを一つのニューロンに対応させることが可能になる。

[0125] また、基板１等の上方には、複数の配線層が形成され、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０とは、互いに異なる配線層の間に配置されてもよい。これにより、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０とで、異なる製造プロセスを適用できる。

[0126] また、基板１等の上方には、配線層が形成され、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０とは、配線層を介して互いに接続されてもよい。これにより、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０とは、共通の配線層で接

続され、配線層の共通化によって構造が簡素化される。

[0127] また、基板１等の平面と鉛直の方向から見て、第１の抵抗変化型不揮発性記憶素子１０の中心と第２の抵抗変化型不揮発性記憶素子２０の中心とは、一致しなくてもよい。これにより、第１の抵抗変化型不揮発性記憶素子１０と第２の抵抗変化型不揮発性記憶素子２０とで、平面視における領域を区別したうえで、形成することができる。

[0128] また、第１の抵抗変化型不揮発性記憶素子１０は、人工知能処理装置２００の出荷前および出荷後のファームウェア更新、学習モデルの更新、定期メンテナンス、および第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスの変更では結合重み係数の更新が不十分な時のうちの少なくとも一つの時に、コンダクタンスが変更され、第２の抵抗変化型不揮発性記憶素子２０は、人工知能処理装置２００の出荷後の学習のための結合重み係数の更新時に、コンダクタンスが変更されてもよい。これにより、製品出荷時等における高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が可能になる。

[0129] また、人工知能処理装置２００は、既存のニューラルネットワークの結合重み係数の設定を用いた転移学習あるいは強化学習に適用され、既存のニューラルネットワークの結合重み係数設定をそのまま用いた第１のニューラルネットワーク領域と、新たに学習を行う第２のニューラルネットワーク領域とを有し、第１のニューラルネットワーク領域の結合重み係数の設定には、第１の抵抗変化型不揮発性記憶素子１０のコンダクタンスが更新され、第２のニューラルネットワーク領域の結合重み係数の設定には、第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスが更新される。これにより、書き込み後のリテンション特性が高い第１の抵抗変化型不揮発性記憶素子１０が結合重み係数の設定（初期設定）に用いられ、一方、書き込みによる素子劣化度合いが小さい、エンデュランス特性が高い第２の抵抗変化型不揮発性記憶素子２０が結合重み係数の更新（学習）に用いられるので、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の

それぞれの特性を生かした人工知能処理装置２００が実現される。

[0130] また、本開示に係る人工知能処理装置２００の学習推論方法は、結合重み係数を変更する場合に、当該場合が、初期設定のために結合重み係数を変更する第１のケースであるか、学習において結合重み係数を変更する第２のケースであるかを判断する判断ステップ（Ｓ３０）と、判断ステップでの判断の結果、第１のケースであると判断した場合に、第１の抵抗変化型不揮発性記憶素子１０のコンダクタンスを変更する第１の書き換えステップ（Ｓ３１～Ｓ３２）と、判断ステップでの判断の結果、第２のケースであると判断した場合に、第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスを変更する第２の書き換えステップ（Ｓ３５～Ｓ３９）と、推論する場合に、積和演算回路２１５から出力される、第１の抵抗変化型不揮発性記憶素子１０に流れる電流および第２の抵抗変化型不揮発性記憶素子２０に流れる電流の合計を、積和演算における１つの積として、用いる推論ステップ（読み出し）とを含む。

[0131] これにより、用途を判定したうえで、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の一方だけに対してコンダクタンスの更新が行われ、製品出荷時等における高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が実現される。

[0132] ここで、第１のケースには、人工知能処理装置２００の出荷前および出荷後のファームウェア更新、学習モデルの更新、定期メンテナンス、および第２の抵抗変化型不揮発性記憶素子２０のコンダクタンスの変更では結合重み係数の更新が不十分な時のうちの少なくとも一つが含まれ、第２のケースには、人工知能処理装置２００の出荷後の学習のための結合重み係数の更新時が含まれる。これにより、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０のそれぞれの特性を生かした人工知能処理装置２００の学習推論方法が実現される。

[0133] 以上、本開示に係る人工知能処理装置およびその学習推論方法について、

実施形態に基づいて説明したが、本開示は、この実施形態に限定されるものではない。本開示の主旨を逸脱しない限り、当業者が思いつく各種変形を本実施形態に施したものや、実施形態における一部の構成要素を組み合わせて構築される別の形態も、本開示の範囲内に含まれる。

[0134] 例えば、上記実施形態では、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０は、タンタル酸化物で構成される抵抗変化層を有したが、抵抗変化層は、このような材料に限られず、ハフニウム等の遷移金属またはアルミニウムの酸化物で構成されてもよい。

[0135] また、上記実施形態では、第１の抵抗変化型不揮発性記憶素子１０および第２の抵抗変化型不揮発性記憶素子２０の抵抗変化層は、第１タンタル酸化物層と第２タンタル酸化物層の積層構造で構成されたが、このような積層構造に限られず、タンタル酸化物層等の単一層で構成されてもよい。

[0136] また、上記実施形態では、一つのニューロンは、１個の第１の抵抗変化型不揮発性記憶素子１０および１個の第２の抵抗変化型不揮発性記憶素子２０で構成されたが、少なくとも１個の第１の抵抗変化型不揮発性記憶素子１０および少なくとも１個の第２の抵抗変化型不揮発性記憶素子２０を備えるのであれば、２個以上の第１の抵抗変化型不揮発性記憶素子１０または２個以上の第２の抵抗変化型不揮発性記憶素子２０を備えてもよい。

[0137] また、上記実施形態では、人工知能処理装置２００は、図９Ａに示される構造をもつニューラルネットワークであったが、この構造に限られず、所望の数の層で構成され、各層が所望の数のニューロンで構成されるニューラルネットワークであってもよい。

産業上の利用可能性

[0138] 本開示の抵抗変化型不揮発性記憶素子を用いた人工知能処理装置は、製品出荷時等における高効率・高精度の結合重み係数の設定（初期設定）と製品出荷後等における高効率な結合重み係数の更新（学習）の両立が可能になるものであり、特に、ＩｏＴ向けエッジＡＩ処理装置などとして有用である。

符号の説明

[0139] １、１１、６０、７０、８０、９０　基板

２、１２　第１電極

３、１３　抵抗変化層

３ａ、１３ａ　第１タンタル酸化物層

３ｂ、１３ｂ　第２タンタル酸化物層

４、１４　第２電極

１０　第１の抵抗変化型不揮発性記憶素子

２０　第２の抵抗変化型不揮発性記憶素子

６１ａ、６１ｂ、７１ａ、７１ｂ、８１ａ、８１ｂ、９１ａ、９１ｂ、１０１ａ、１０１ｂ、　拡散領域

６２、７２、８２、９２、１０２　酸化膜

６３、７３、８３、９３、１０３　ゲート電極（ワード線）

６４ａ、６４ｂ、６６、６８ａ、６８ｂ、７４ａ、７４ｂ、７６ａ、７６ｂ、７６ｃ、７８ａ、７８ｂ、８４ａ、８４ｂ、８６ａ、８６ｂ、８６ｃ、８８ａ、８８ｂ、９４ａ、９４ｂ、９６、９８ａ、９８ｂ、１０４ａ、１０４ｂ、１０６ａ、１０６ｂ、１０８　ビア

６５ａ、６５ｂ、７５ａ、７５ｂ、７５ｃ、８５ａ、８５ｂ、９５ａ、９５ｂ、１０５　第１配線層

６７、７７、８７ａ、８７ｂ、９７、１０７　第２配線層

６９、７９、８９、９９　第３配線層

２００　人工知能処理装置

２０１　入力層

２０２　隠れ層

２０３　出力層

２１０、２１０ａ　ニューロン

２１１　シナプス

２１５、２１５ａ　積和演算回路

２２０　メモリセルアレイ

２３０、２３０ａ　　ワード線選択回路

２４０　　カラムゲート

２５０、２５０ａ　　判定回路

２６０　　書き込み回路

２７０　　制御回路

ＭＣ、ＭＣ１、ＭＣ２、ＭＣｉｊ　メモリセル

Ｔ０、Ｔ１、Ｔ２　セルトランジスタ

ＷＬ、ＷＬ１、ＷＬ２　ワード線

ＢＬ、ＢＬ１、ＢＬ２、ＢＬｉ１、ＢＬｉ２　ビット線

ＳＬ、ＳＬｉ　ソース線

請求の範囲

[請求項1] １つの基板と、

前記１つの基板に搭載され、積和演算を行う積和演算回路とを備え、

前記積和演算回路は、前記積和演算における積の演算に用いられる結合重み係数をコンダクタンスとして保持する、特性の異なる第１の抵抗変化型不揮発性記憶素子と第２の抵抗変化型不揮発性記憶素子とを含み、

同一極性かつ同一電圧の電圧パルスを複数回連続して印加されると、前記第１の抵抗変化型不揮発性記憶素子における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率が、前記第２の抵抗変化型不揮発性記憶素子における１発目の電圧パルスでのコンダクタンスの変化量に対する２発目の電圧パルスでのコンダクタンスの変化量の比率よりも小さい、

人工知能処理装置。

[請求項2] １つの基板と、

前記１つの基板に搭載され、積和演算を行う積和演算回路とを備え、

前記積和演算回路は、第１電極、第２電極、および前記第１電極と前記第２電極との間に挟まれた抵抗変化層を含み、前記積和演算における積の演算に用いられる結合重み係数をコンダクタンスとして保持する複数の抵抗変化型不揮発性記憶素子を有し、

前記複数の抵抗変化型不揮発性記憶素子は、特性の異なる第１の抵抗変化型不揮発性記憶素子と第２の抵抗変化型不揮発性記憶素子とを含み、

前記第１の抵抗変化型不揮発性記憶素子において第１電極と第２電極のうち少なくとも片方の電極が貴金属電極であり、

前記第２の抵抗変化型不揮発性記憶素子において第１電極と第２電極のうち少なくとも片方の電極が非貴金属電極である、

人工知能処理装置。

[請求項3] 前記貴金属電極はＩｒ、Ｐｔのうちいずれか１つを含み、

前記非貴金属電極はＴｉＮ、ＴａＮのうちいずれか１つを含む、

請求項２記載の人工知能処理装置。

[請求項4] １つの基板と、

前記１つの基板に搭載され、積和演算を行う積和演算回路とを備え、

前記積和演算回路は、前記積和演算における積の演算に用いられる結合重み係数をコンダクタンスとして保持する、特性の異なる第１の抵抗変化型不揮発性記憶素子と第２の抵抗変化型不揮発性記憶素子とを含み、

前記第１の抵抗変化型不揮発性記憶素子は、前記第２の抵抗変化型不揮発性記憶素子よりも高いリテンション特性を有し、

前記第２の抵抗変化型不揮発性記憶素子は、前記第１の抵抗変化型不揮発性記憶素子よりも高いエンデュランス特性を有する、

人工知能処理装置。

[請求項5] 前記積和演算回路は、前記第１の抵抗変化型不揮発性記憶素子に流れる電流および前記第２の抵抗変化型不揮発性記憶素子に流れる電流を合計し、得られた電流の合計を、前記積和演算における１つの積として、出力する、

請求項１～４のいずれか１項に記載の人工知能処理装置。

[請求項6] 前記基板の上方には、複数の配線層が形成され、

前記第１の抵抗変化型不揮発性記憶素子と前記第２の抵抗変化型不揮発性記憶素子とは、互いに異なる前記配線層の間に配置されている、

請求項１～５のいずれか１項に記載の人工知能処理装置。

[請求項7] 前記基板の上方には、配線層が形成され、

前記第１の抵抗変化型不揮発性記憶素子と前記第２の抵抗変化型不揮発性記憶素子とは、前記配線層を介して互いに接続されている、

請求項１～５のいずれか１項に記載の人工知能処理装置。

[請求項8] 前記基板の平面と鉛直の方向から見て、前記第１の抵抗変化型不揮発性記憶素子の中心と前記第２の抵抗変化型不揮発性記憶素子の中心とは、一致しない、

請求項１～４のいずれか１項に記載の人工知能処理装置。

[請求項9] 前記第１の抵抗変化型不揮発性記憶素子は、前記人工知能処理装置の出荷前および出荷後のファームウェア更新、学習モデルの更新、定期メンテナンス、および前記第２の抵抗変化型不揮発性記憶素子のコンダクタンスの変更では結合重み係数の更新が不十分な時のうちの少なくとも一つの時に、コンダクタンスが変更され、

前記第２の抵抗変化型不揮発性記憶素子は、前記人工知能処理装置の出荷後の学習のための結合重み係数の更新時に、コンダクタンスが変更される、

請求項１～８のいずれか１項に記載の人工知能処理装置。

[請求項10] 前記第１の抵抗変化型不揮発性記憶素子および前記第２の抵抗変化型不揮発性記憶素子のそれぞれは、第１電極、第２電極、および前記第１電極と前記第２電極との間に挟まれた抵抗変化層を含み、

前記第１の抵抗変化型不揮発性記憶素子において第１電極と第２電極のうち少なくとも片方の電極が貴金属電極であり、

前記第２の抵抗変化型不揮発性記憶素子において第１電極と第２電極のうち少なくとも片方の電極が非貴金属電極である、

請求項１、４～９のいずれか１項に記載の人工知能処理装置。

[請求項11] 前記第１の抵抗変化型不揮発性記憶素子は、前記第２の抵抗変化型不揮発性記憶素子よりも高いリテンション特性を有し、

前記第２の抵抗変化型不揮発性記憶素子は、前記第１の抵抗変化型

不揮発性記憶素子よりも高いエンデュランス特性を有する、

請求項１～３、５～１０のいずれか１項に記載の人工知能処理装置。

[請求項12] 前記人工知能処理装置は、

既存のニューラルネットワークの結合重み係数の設定を用いた転移学習あるいは強化学習に適用され、

前記既存のニューラルネットワークの結合重み係数設定をそのまま用いた第１のニューラルネットワーク領域と、

新たに学習を行う第２のニューラルネットワーク領域とを有し、

前記第１のニューラルネットワーク領域の結合重み係数の設定には、前記第１の抵抗変化型不揮発性記憶素子のコンダクタンスが更新され、

前記第２のニューラルネットワーク領域の結合重み係数の設定には、前記第２の抵抗変化型不揮発性記憶素子のコンダクタンスが更新される、

請求項１～１１のいずれか１項に記載の人工知能処理装置。

[請求項13] 請求項１～１２のいずれか１項に記載の人工知能処理装置の学習推論方法であって、

結合重み係数を変更する場合に、当該場合が、初期設定のために結合重み係数を変更する第１のケースであるか、学習において結合重み係数を変更する第２のケースであるかを判断する判断ステップと、

前記判断ステップでの判断の結果、前記第１のケースであると判断した場合に、前記第１の抵抗変化型不揮発性記憶素子のコンダクタンスを変更する第１の書き換えステップと、

前記判断ステップでの判断の結果、前記第２のケースであると判断した場合に、前記第２の抵抗変化型不揮発性記憶素子のコンダクタンスを変更する第２の書き換えステップと、

推論する場合に、前記積和演算回路から出力される、前記第１の抵

抗変化型不揮発性記憶素子に流れる電流および前記第２の抵抗変化型不揮発性記憶素子に流れる電流の合計を、積和演算における１つの積として、用いる推論ステップとを含む、

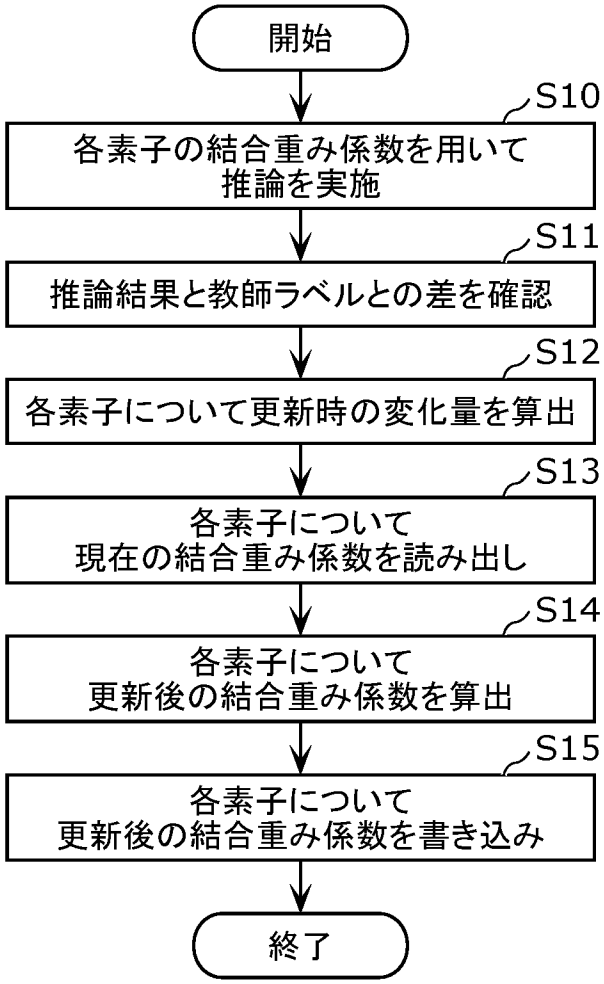
人工知能処理装置の学習推論方法。

[請求項14] 前記第１のケースには、前記人工知能処理装置の出荷前および出荷後のファームウェア更新、学習モデルの更新、定期メンテナンス、および前記第２の抵抗変化型不揮発性記憶素子のコンダクタンスの変更では結合重み係数の更新が不十分な時のうちの少なくとも一つが含まれ、

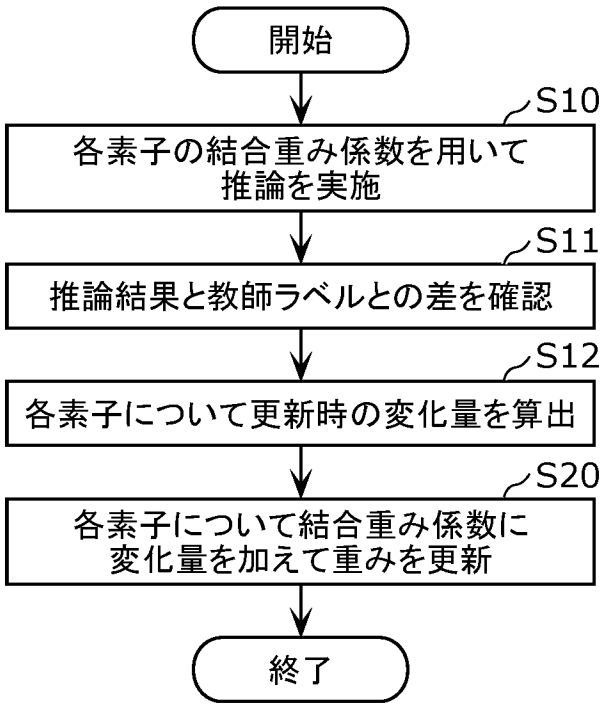
前記第２のケースには、前記人工知能処理装置の出荷後の学習のための結合重み係数の更新時が含まれる、

請求項１３記載の人工知能処理装置の学習推論方法。

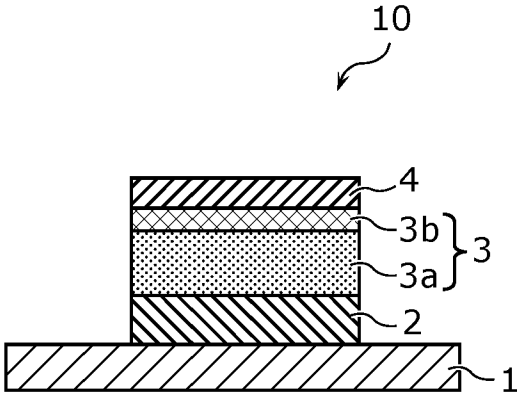
[図1A]


開始
S10
各素子の結合重み係数を用いて
推論を実施
S11
推論結果と教師ラベルとの差を確認
S12
各素子について更新時の変化量を算出
S13
各素子について
現在の結合重み係数を読み出し
S14
各素子について
更新後の結合重み係数を算出
S15
各素子について
更新後の結合重み係数を書き込み
終了

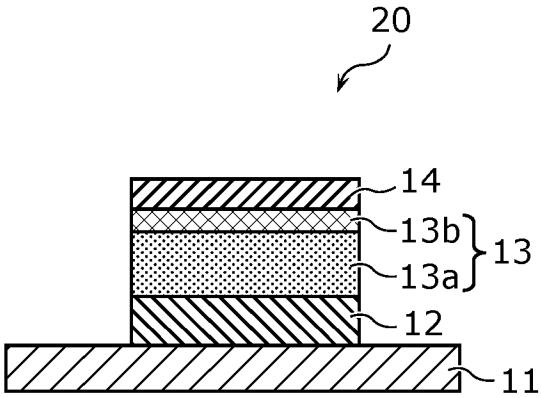
[図1B]


開始
S10
各素子の結合重み係数を用いて
推論を実施
S11
推論結果と教師ラベルとの差を確認
S12
各素子について更新時の変化量を算出
S20
各素子について結合重み係数に
変化量を加えて重みを更新
終了

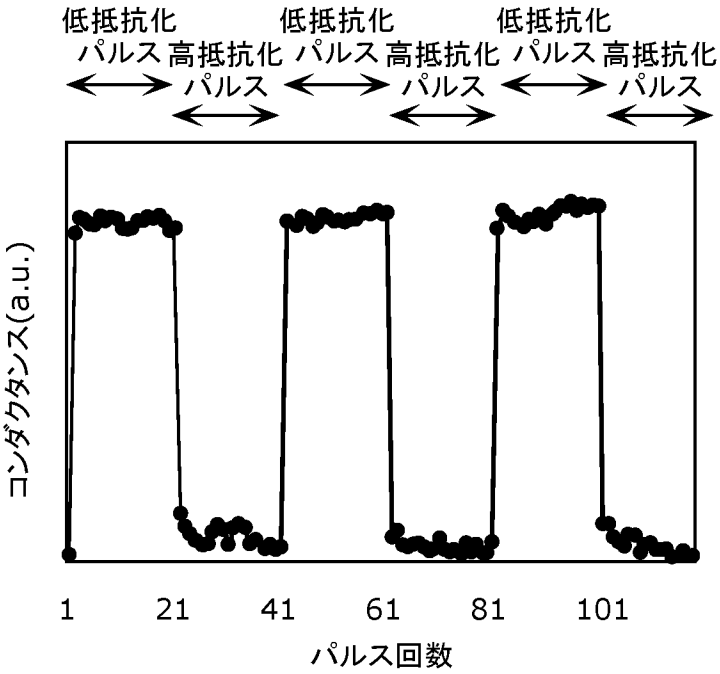
[図2]


10
4
3b
3
3a
2
1

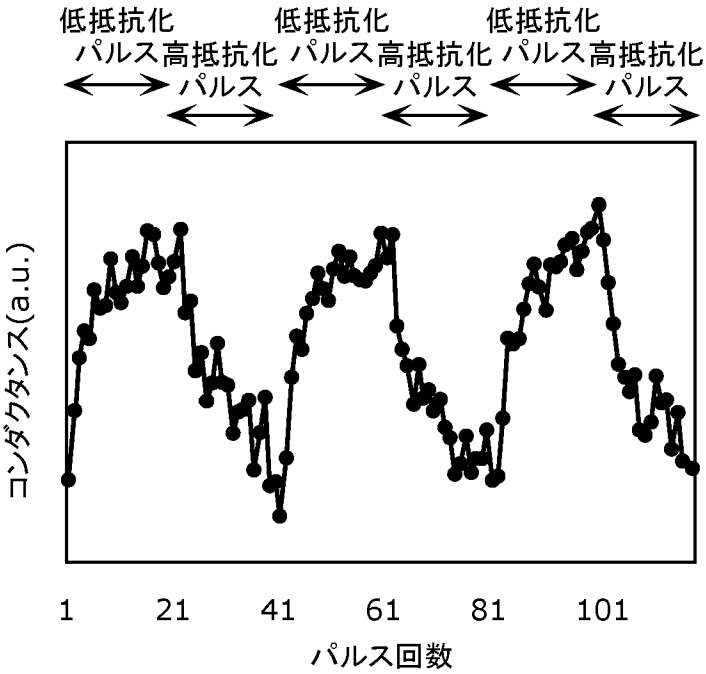
[図3]


20
14
13b
13
13a
12
11

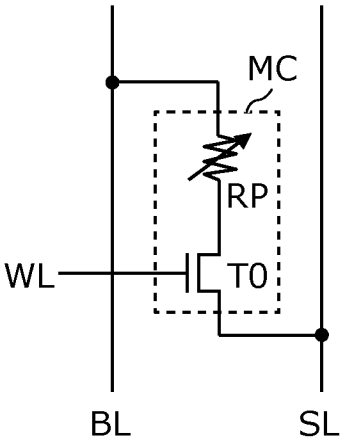
[図4]


低抵抗化パルス
高抵抗化パルス
低抵抗化パルス
高抵抗化パルス
低抵抗化パルス
高抵抗化パルス
コンダクタンス(a.u.)
1
21
41
61
81
101
パルス回数

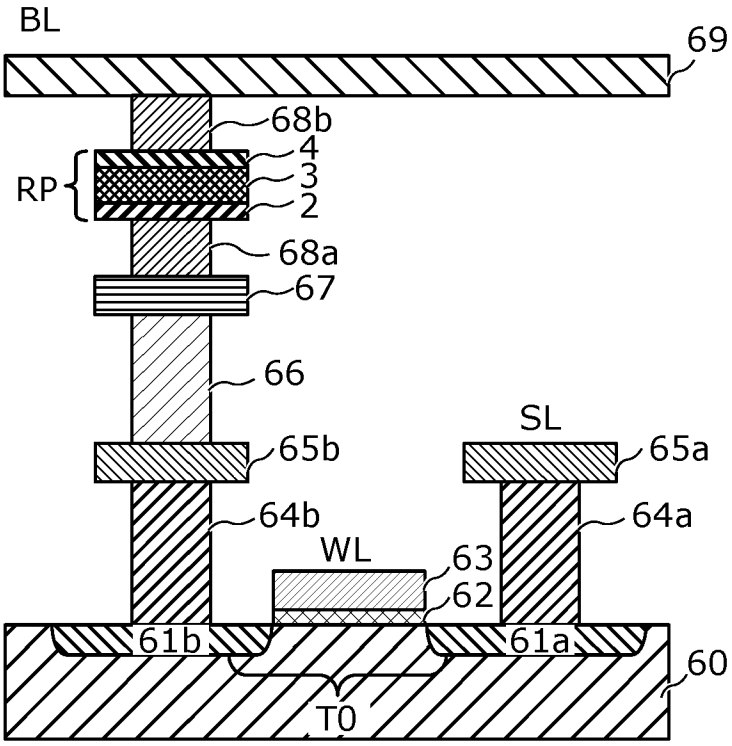
[図5]


低抵抗化パルス
高抵抗化パルス
低抵抗化パルス
高抵抗化パルス
低抵抗化パルス
高抵抗化パルス
コンダクタンス(a.u.)
1
21
41
61
81
101
パルス回数

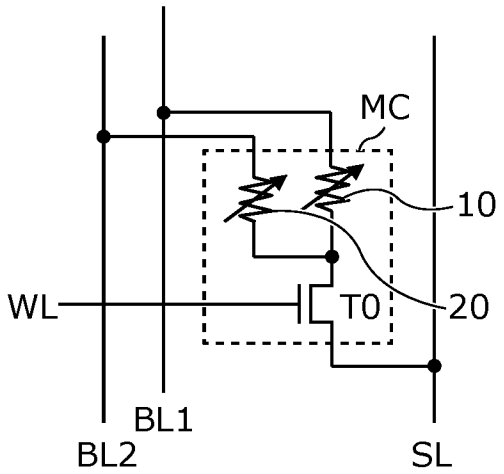
[図6A]


MC
RP
WL
T0
BL
SL

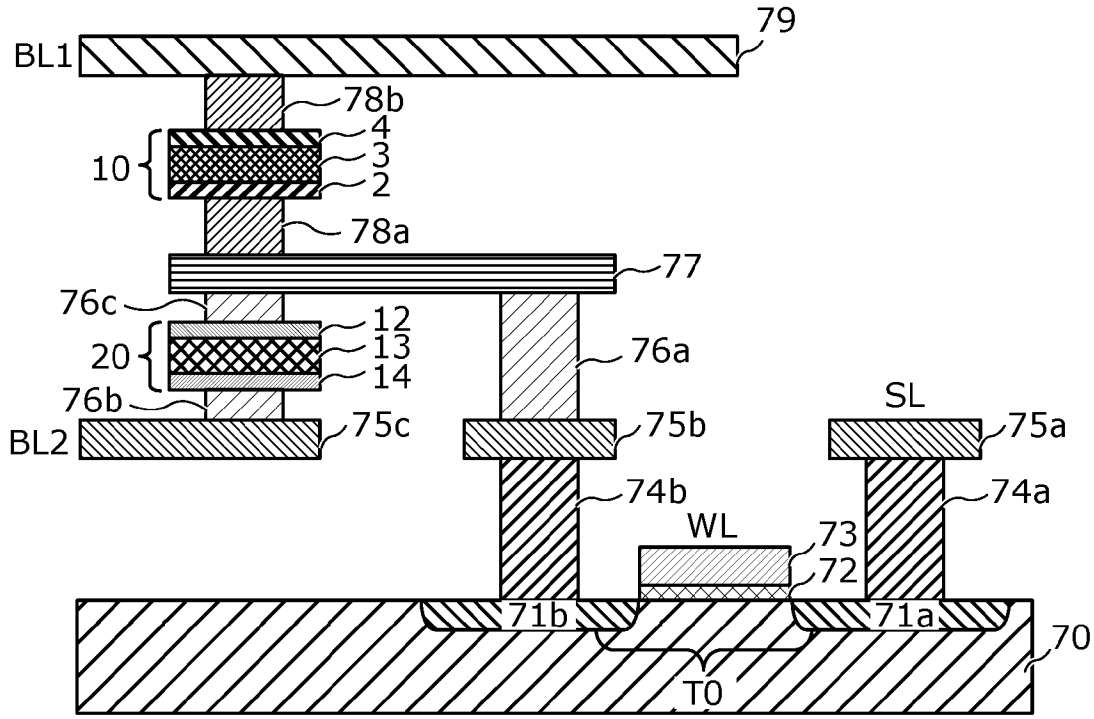
[図6B]


BL
69
68b
4
3
2
RP
68a
67
66
SL
65b
65a
64b
64a
WL
63
62
61b
61a
60
T0

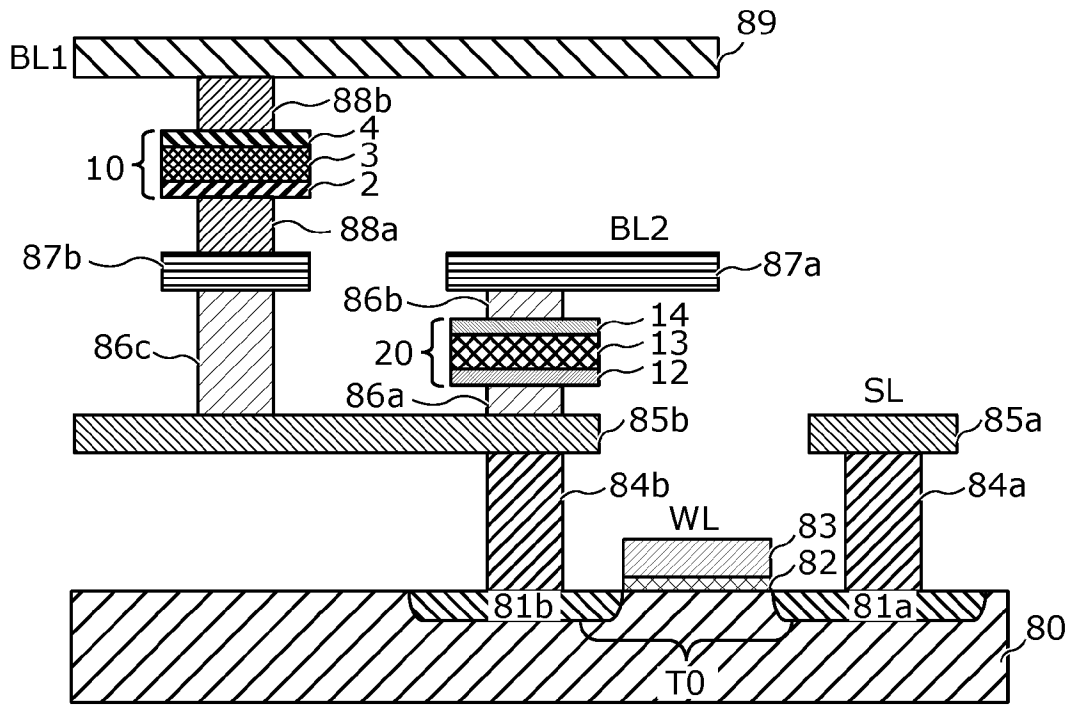
[図7A]


MC
10
WL
T0
20
BL1
BL2
SL

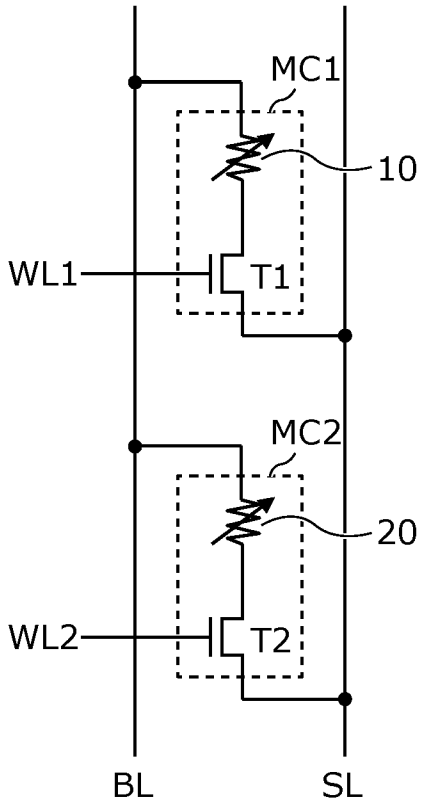
[図7B]


79
BL1
78b
4
10
3
2
78a
77
76c
12
20
13
14
76a
76b
SL
BL2
75c
75b
75a
74b
74a
WL
73
72
71b
71a
70
T0

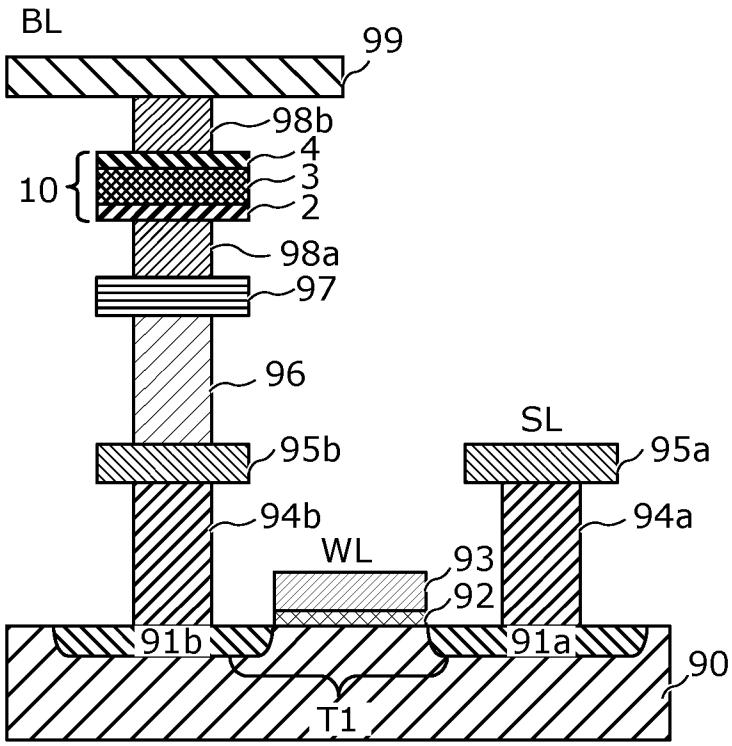
[図7C]


89
BL1
88b
4
10
3
2
88a
BL2
87b
87a
86b
14
86c
20
13
12
86a
SL
85b
85a
84b
84a
WL
83
82
81b
81a
80
T0

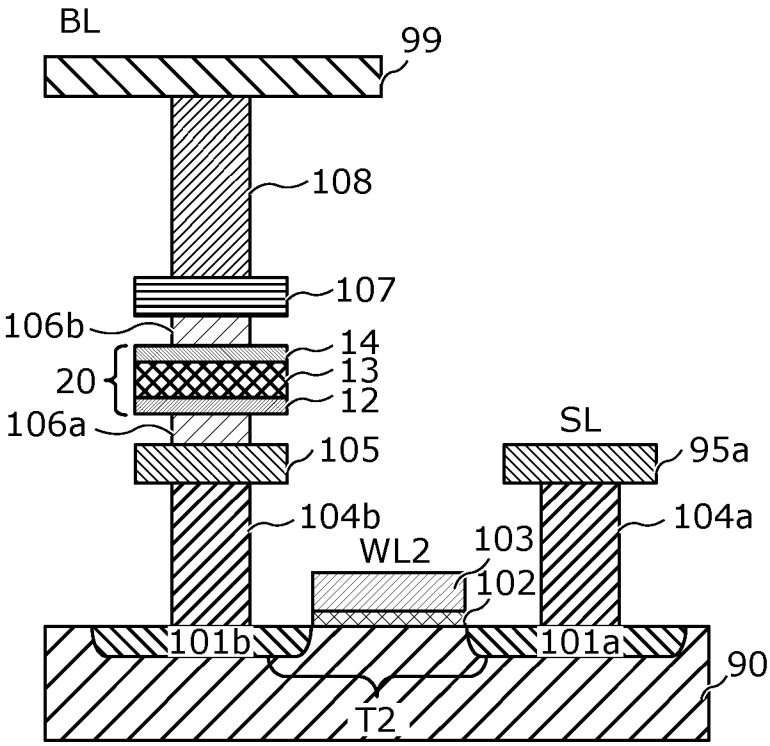
[図8A]


MC1
10
WL1
T1
MC2
20
WL2
T2
BL
SL

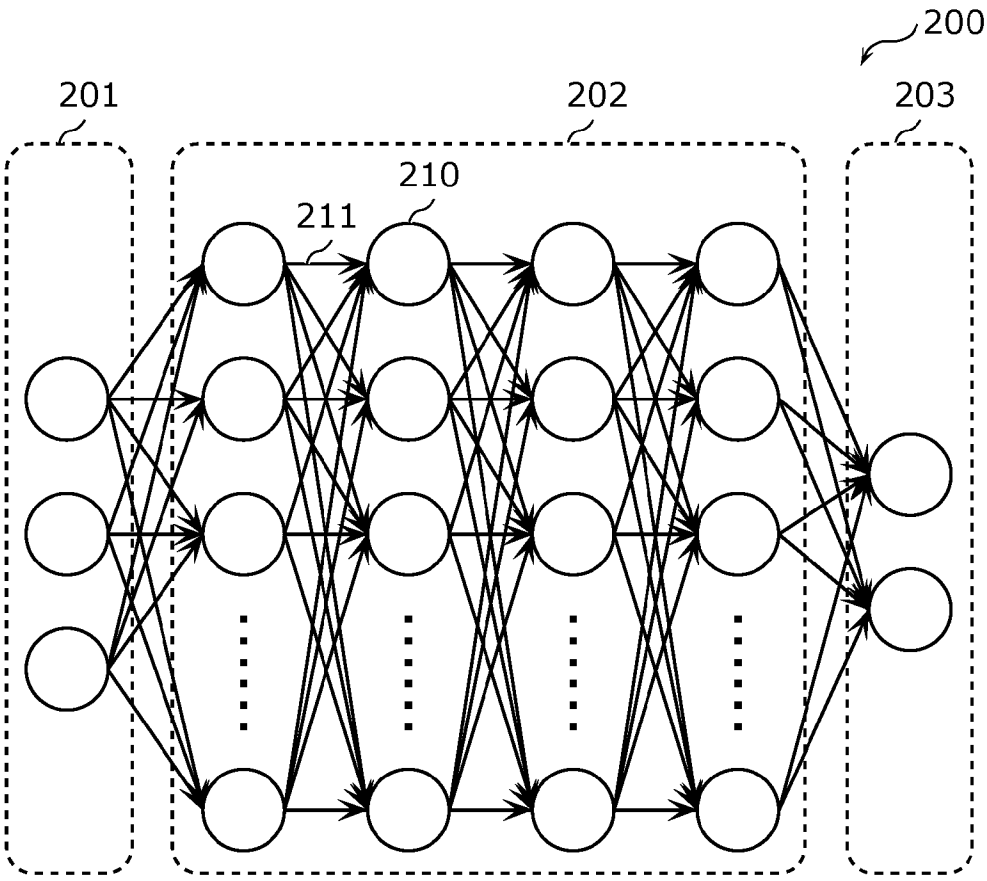
[図8B]


BL
99
98b
4
10
3
2
98a
97
96
SL
95b
95a
94b
94a
WL
93
92
91b
91a
90
T1

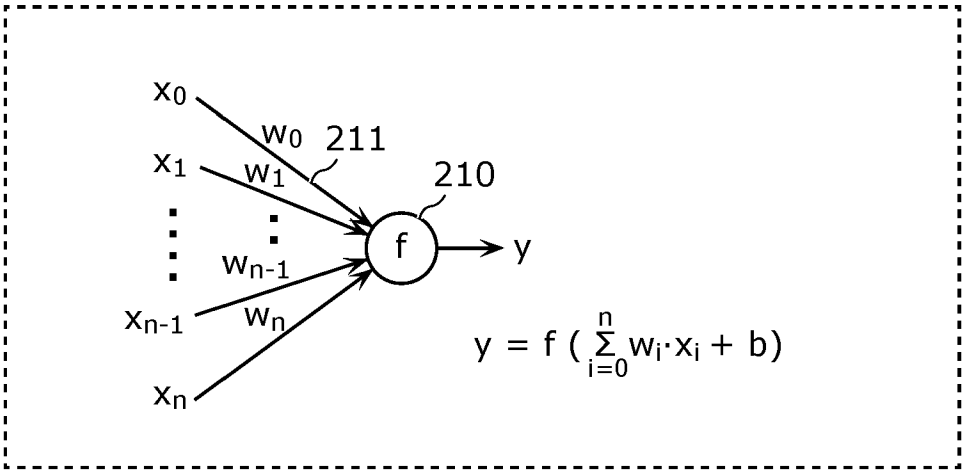
[図8C]


BL
99
108
107
106b
14
20
13
12
106a
105
SL
95a
104b
104a
WL2
103
102
101b
101a
90
T2

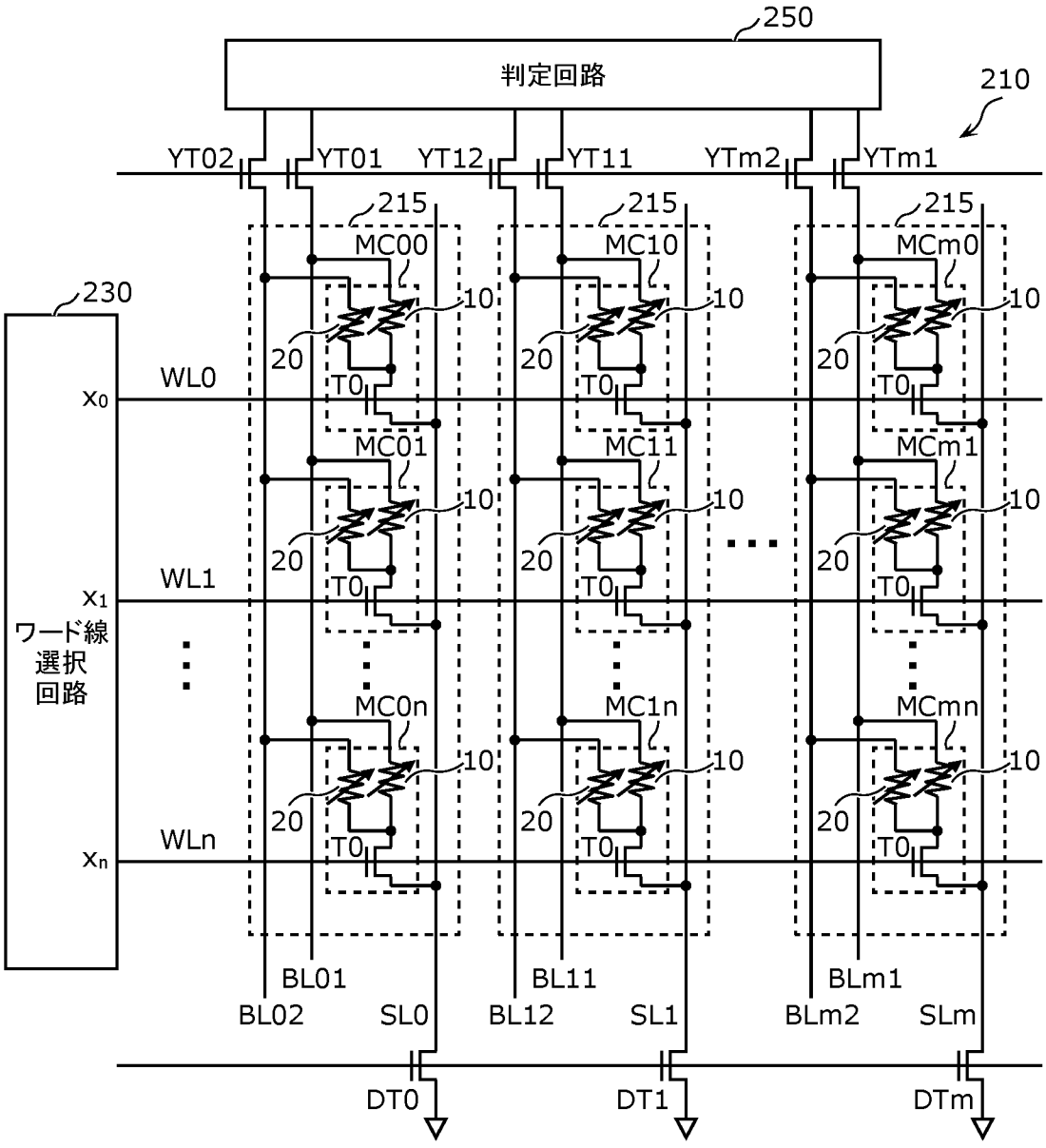
[図9A]


200
201
202
203
210
211

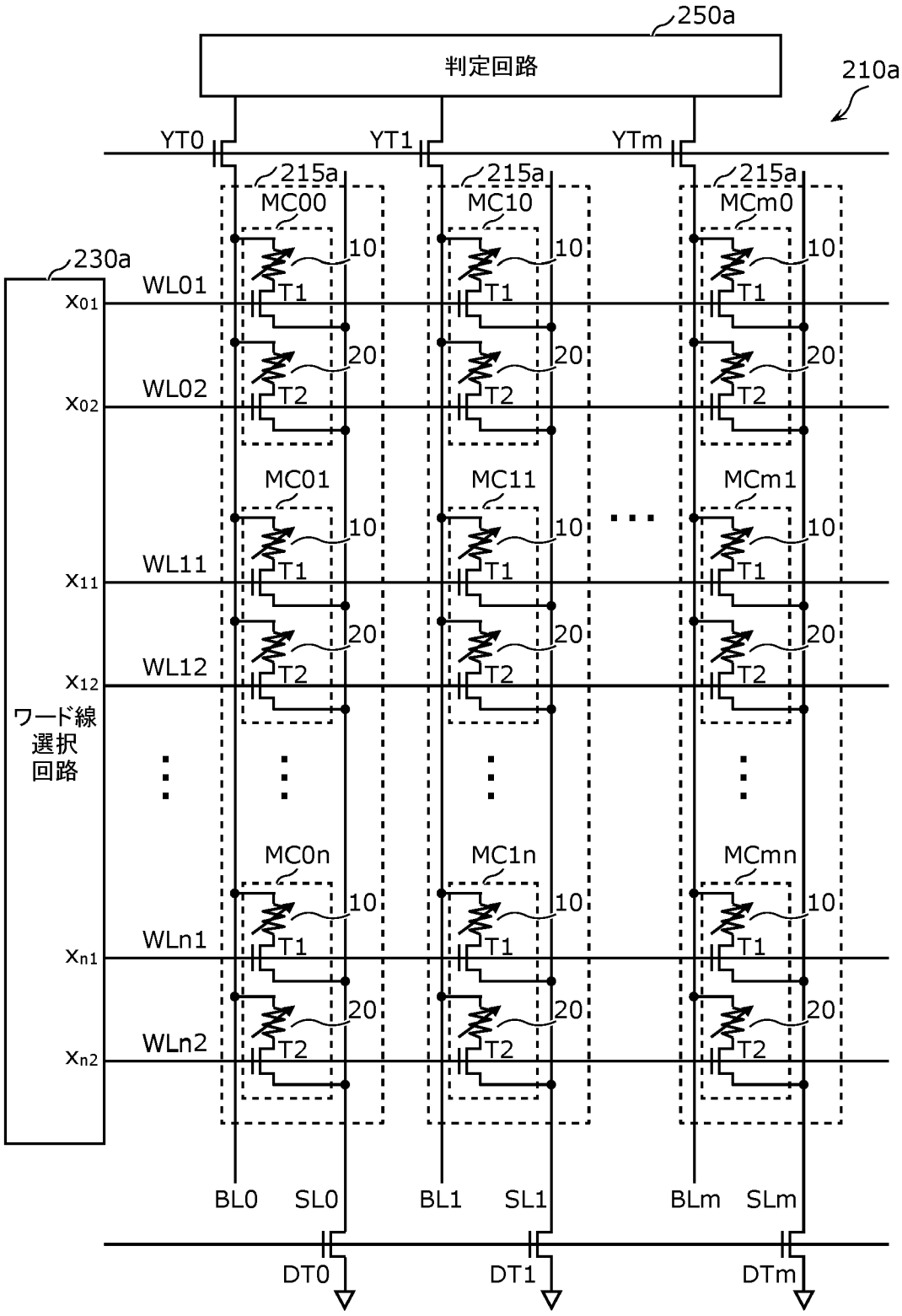
[図9B]


x0
x1
xn-1
xn
w0
w1
wn-1
wn
211
210
f
y
y = f (Σ wi·xi + b)
i=0
n

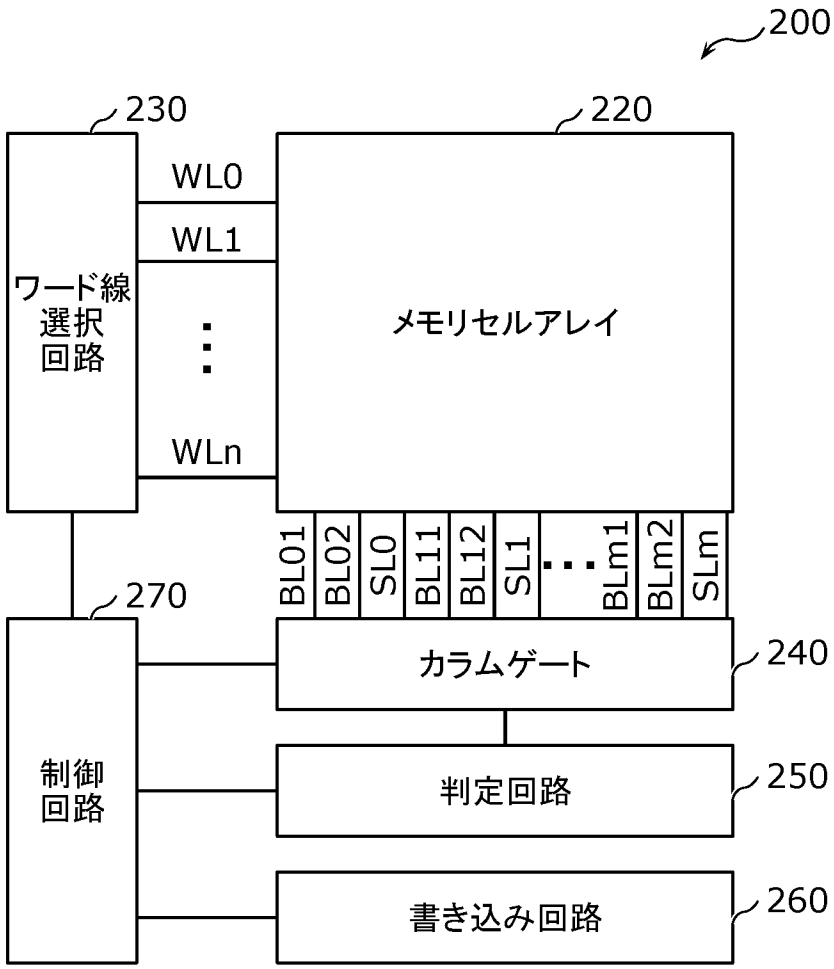
[図10A]


250
判定回路
210
YT02
YT01
YT12
YT11
YTm2
YTm1
215
MC00
MC10
MCm0
MC01
MC11
MCm1
MC0n
MC1n
MCmn
10
20
T0
230
ワード線
選択
回路
WL0
WL1
WLn
x0
x1
xn
BL01
BL02
BL11
BL12
BLm1
BLm2
SL0
SL1
SLm
DT0
DT1
DTm

[図10B]


250a
判定回路
210a
YT0
YT1
YTm
215a
MC00
MC10
MCm0
230a
WL01
WL02
WL11
WL12
WLn1
WLn2
x01
x02
x11
x12
xn1
xn2
10
20
T1
T2
MC01
MC11
MCm1
MC0n
MC1n
MCmn
ワード線
選択
回路
BL0
SL0
BL1
SL1
BLm
SLm
DT0
DT1
DTm

[図11]



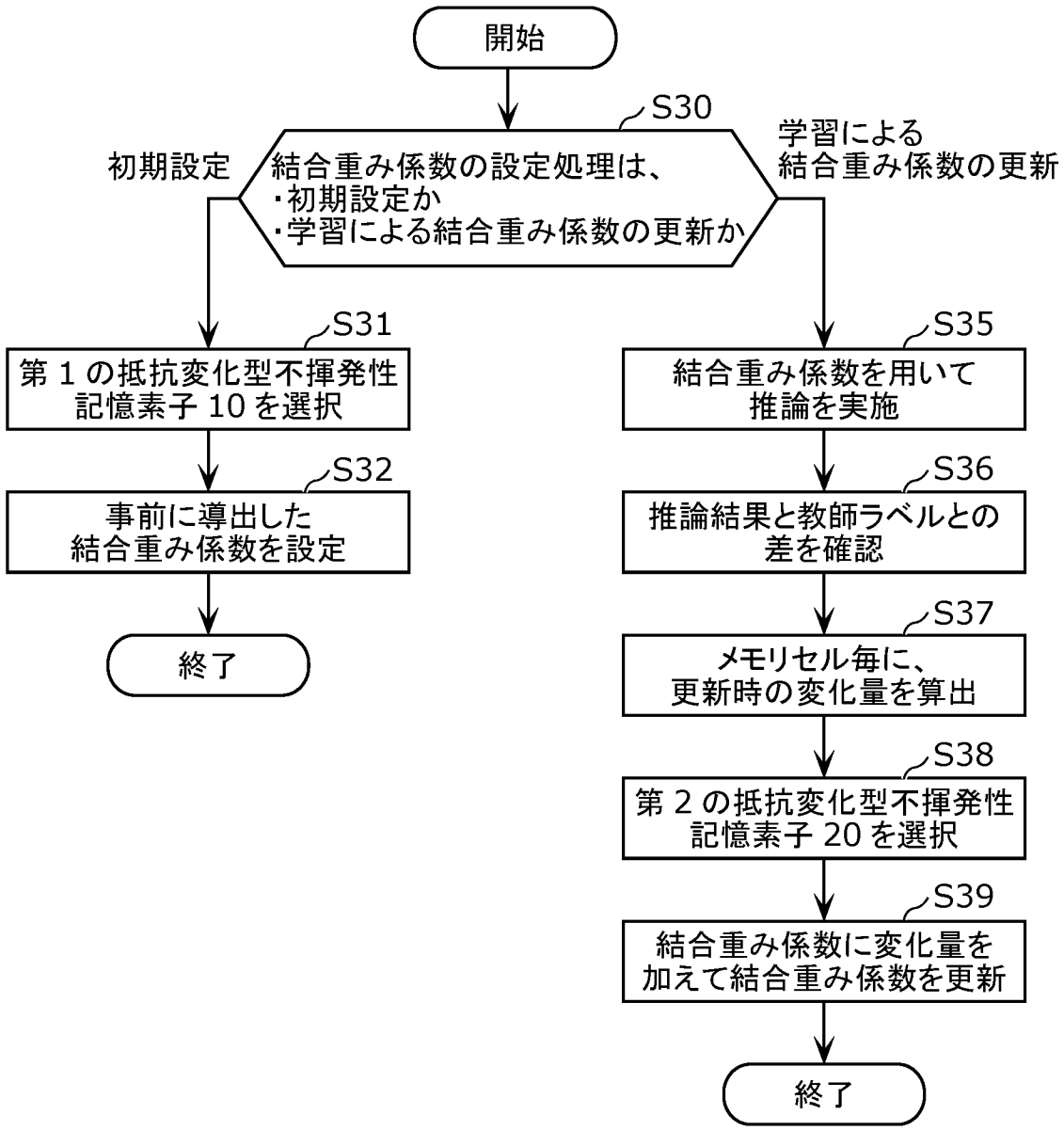
[図12A]

動作	リセット（高抵抗化）		セット（低抵抗化）		読み出し
	モード 1	モード 2	モード 1	モード 2	
WL	$V_{g\text{-}on}$	$V_{g\text{-}on}$	$V_{g\text{-}on}$	$V_{g\text{-}on}$	$V_{g\text{-}read}$
BL1	V_{reset}	V_{ss}	V_{ss}	V_{set}	V_{read}
BL2	V_{ss}	V_{reset}	V_{set}	V_{ss}	V_{read}
SL	V_{ss}	V_{ss}	V_{set}	V_{set}	V_{ss}

[図12B]

動作	リセット（高抵抗化）		セット（低抵抗化）		読み出し
	モード 1	モード 2	モード 1	モード 2	
WL1	$V_{g\text{-}on}$	$V_{g\text{-}off}$	$V_{g\text{-}on}$	$V_{g\text{-}off}$	$V_{g\text{-}read}$
WL2	$V_{g\text{-}off}$	$V_{g\text{-}on}$	$V_{g\text{-}off}$	$V_{g\text{-}on}$	$V_{g\text{-}read}$
BL	V_{reset}	V_{reset}	V_{ss}	V_{ss}	V_{read}
SL	V_{ss}	V_{ss}	V_{set}	V_{set}	V_{ss}

[図13]



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2022/044141

A. CLASSIFICATION OF SUBJECT MATTER

G06N 3/063(2023.01)i; ***G06G 7/14***(2006.01)i; ***G06G 7/16***(2006.01)i; ***G06G 7/19***(2006.01)i; ***G06G 7/60***(2006.01)i; ***G11C 11/54***(2006.01)i; ***G11C 13/00***(2006.01)i; ***H10B 63/00***(2023.01)i
FI: G06N3/063; G11C11/54; G11C13/00 215; G11C13/00 270J; G06G7/60; G06G7/16 510; G06G7/14; G06G7/19; G11C13/00 480B; G11C13/00 270H; H10B63/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06N3/063; G06G7/14; G06G7/16; G06G7/19; G06G7/60; G11C11/54; G11C13/00; H10B63/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996
Published unexamined utility model applications of Japan 1971-2023
Registered utility model specifications of Japan 1996-2023
Published registered utility model applications of Japan 1994-2023

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2020/0034686 A1 (SANDISK TECHNOLOGIES LLC) 30 January 2020 (2020-01-30) entire text, all drawings	1-14
A	NINOMIYA, T. et al. Data Retention of TaOx ReRAM in Low Current Operation and after Cycling. The 61st JSAP Spring Meeting 2014 [Extended Abstracts]. THE JAPAN SOCIETY OF APPLIED PHYSICS. 2014, p. 073, ISBN 978-4-86348-409-2 entire text, all drawings	1-14
A	WO 2010/004705 A1 (PANASONIC CORP.) 14 January 2010 (2010-01-14) entire text, all drawings	1-14

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 February 2023	**14 February 2023**
Name and mailing address of the ISA/JP **Japan Patent Office (ISA/JP)** **3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915** **Japan**	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2022/044141

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US 2020/0034686 A1	30 January 2020	DE 102019119744 A1	
		CN 110782027 A	
WO 2010/004705 A1	14 January 2010	US 2011/0002154 A1 entire text, all drawings	
		CN 101978496 A	

国際調査報告

国際出願番号
PCT/JP2022/044141

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06N 3/063(2023.01)i; G06G 7/14(2006.01)i; G06G 7/16(2006.01)i; G06G 7/19(2006.01)i; G06G 7/60(2006.01)i; G11C 11/54(2006.01)i; G11C 13/00(2006.01)i; H10B 63/00(2023.01)i
FI: G06N3/063; G11C11/54; G11C13/00 215; G11C13/00 270J; G06G7/60; G06G7/16 510; G06G7/14; G06G7/19; G11C13/00 480B; G11C13/00 270H; H10B63/00

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

G06N3/063; G06G7/14; G06G7/16; G06G7/19; G06G7/60; G11C11/54; G11C13/00; H10B63/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２３年
日本国実用新案登録公報	１９９６－２０２３年
日本国登録実用新案公報	１９９４－２０２３年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
A	US 2020/0034686 A1 (SANDISK TECHNOLOGIES LLC) 30.01.2020 (2020-01-30) 全文,全図	1-14
A	NINOMIYA, T et al., Data Retention of TaOx ReRAM in Low Current Operation and after Cycling, The 61st JSAP Spring Meeting 2014 [Extended Abstracts], The Japan Society of Applied Physics, 2014, p. 073, ISBN 978-4-86348-409-2 全文,全図	1-14
A	WO 2010/004705 A1 (パナソニック株式会社) 14.01.2010 (2010-01-14) 全文,全図	1-14

☐ Ｃ欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
“O” 口頭による開示、使用、展示等に言及する文献
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
“Y” 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
“&” 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
06.02.2023	14.02.2023

名称及びあて先
日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）
金田　孝之 5B 4808
電話番号 03-3581-1101 内線 3545

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2022/044141

引用文献	公表日	パテントファミリー文献	公表日
US 2020/0034686 A1	30.01.2020	DE 102019119744 A1 CN 110782027 A	
WO 2010/004705 A1	14.01.2010	US 2011/0002154 A1 全文,全図 CN 101978496 A	