



(10) **DE 10 2020 121 214 B4** 2024.10.31

(12) **Patentschrift**

(21) Aktenzeichen: **10 2020 121 214.3**
(22) Anmeldetag: **12.08.2020**
(43) Offenlegungstag: **20.05.2021**
(45) Veröffentlichungstag
der Patenterteilung: **31.10.2024**

(51) Int Cl.: **G11C 11/54 (2006.01)**
G06N 3/04 (2023.01)

Innerhalb von neun Monaten nach Veröffentlichung der Patenterteilung kann nach § 59 Patentgesetz gegen das Patent Einspruch erhoben werden. Der Einspruch ist schriftlich zu erklären und zu begründen. Innerhalb der Einspruchsfrist ist eine Einspruchsgebühr in Höhe von 200 Euro zu entrichten (§ 6 Patentkostengesetz in Verbindung mit der Anlage zu § 2 Abs. 1 Patentkostengesetz).

(30) Unionspriorität:
10-2019-0146341 **15.11.2019** **KR**
16/892,637 **04.06.2020** **US**

(73) Patentinhaber:
**SAMSUNG ELECTRONICS CO., LTD., Suwon-si,
Gyeonggi-do, KR**

(74) Vertreter:
**KUHLEN & WACKER Patent- und
Rechtsanwaltsbüro PartG mbB, 85354 Freising,
DE**

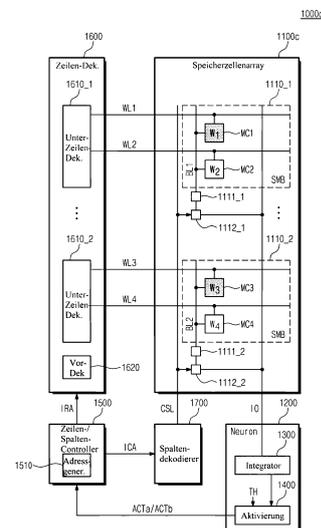
(72) Erfinder:
**Yu, Haksoo, Suwon-si, Gyeonggi-do, KR; Kim,
Nam Sung, Suwon-si, Gyeonggi-do, KR; Sohn,
Kyomin, Suwon-si, Gyeonggi-do, KR; Youn,
Jaeyoun, Suwon-si, Gyeonggi-do, KR**

(56) Ermittelte Stand der Technik:
siehe Folgeseiten

(54) Bezeichnung: **Neuromorphe Vorrichtung basierend auf einem Speicher**

(57) Hauptanspruch: Neuromorphe Vorrichtung (1000a), umfassend:
ein erstes Speicherzellenarray (1100), das erste Speicherzellen (MC1, MC3) enthält, die einer ersten Adresse entsprechen, wobei die ersten Speicherzellen (MC1, MC3) eingerichtet sind, um erste Gewichte (W_1, W_3) zu speichern, und zweite Speicherzellen (MC2, MC4), die einer zweiten Adresse entsprechen, wobei die zweiten Speicherzellen eingerichtet sind, um zweite Gewichte (W_2, W_4) zu speichern; und
eine erste Neuronenschaltung (1200, 1210), die einen ersten Integrator (1300) und eine erste Aktivierungsschaltung (1400) enthält, wobei der erste Integrator (1300) eingerichtet ist, um erste Lesesignale aus den ersten Speicherzellen (MC1, MC3) zu summieren und ein erstes Summensignal an die erste Aktivierungsschaltung (1400) auszugeben und zweite Lesesignale aus den zweiten Speicherzellen (MC2, MC4) zu summieren und zweite Summensignale an die erste Aktivierungsschaltung (1400) auszugeben, und die erste Aktivierungsschaltung (1400) eingerichtet ist, um ein erstes Aktivierungssignal (ACTa) auf der Grundlage des ersten Summensignals auszugeben und ein zweites Aktivierungssignal (ACTb) auf der Grundlage der zweiten Summensignale auszugeben, die von dem Integrator (1300) ausgegeben werden; und
einen Adressengenerator (1510), der eingerichtet ist, um die zweite Adresse auf der Grundlage des ersten Aktivierungssignals (ACTa) zu erzeugen und eine dritte Adresse auf der Grundlage des zweiten Aktivierungssignals (ACTb) zu

erzeugen; und
einen Zeilendekodierer (1600), der eingerichtet ist, um eine erste interne Zeilenadresse der ersten Adresse bzw. eine zweite interne Zeilenadresse der zweiten Adresse zu dekodieren; und
einen Spaltendekodierer (1700), der eingerichtet ist, um eine erste interne Spaltenadresse der ersten Adresse bzw. eine zweite interne Spaltenadresse der zweiten ...



(56) Ermittelter Stand der Technik:

DE	690 28 617	T2
US	6 654 730	B1
US	10 360 971	B1
US	2019 / 0 102 170	A1

Yu, Shimeng. Neuro-inspired computing with emerging nonvolatile memorys. Proceedings of the IEEE, 2018, 106. Jg., Nr. 2, S. 260-285.

BeschreibungQUERVERWEISE AUF VERWANDTE
ANMELDUNGEN

[0001] Diese Anmeldung beansprucht die Priorität der koreanischen Patentanmeldung Nr. 10-2019-0146341, die am 15. November 2019 beim koreanischen Amt für geistiges Eigentum eingereicht wurde

HINTERGRUND

[0002] Die erfinderischen Konzepte beziehen sich auf eine neuromorphe Vorrichtung und insbesondere auf eine speicherbasierte neuromorphe Vorrichtung.

[0003] Künstliche neuronale Netzwerke, die biologischen neuronalen Netzwerken ähnlich sind, stehen im Rampenlicht. Die künstlichen neuronalen Netzwerke werden in verschiedenen Bereichen wie maschinelles Lernen, Auswahl, Schlussfolgerung, Vorhersage, Erkennung, Analyse, Übersetzung und Diagnose eingesetzt. Die künstlichen neuronalen Netzwerke können künstliche Neuronen enthalten, die biologischen Neuronen ähnlich sind (z. B. neuronale Zellen) und eine Vielzahl von Schichten bilden. Ein Synapsengewicht kann eine Verbindungsstärke zwischen den künstlichen Neuronen anzeigen und kann durch maschinelles Lernen gelernt und verändert werden.

[0004] Mit der Zunahme der Anzahl der Schichten eines künstlichen neuronalen Netzwerks und der Anzahl der künstlichen Neuronen nehmen auch die Synapsengewichte und Biases zu, die die Verbindungsstärken zwischen künstlichen Neuronen anzeigen. Im Falle der Implementierung eines künstlichen neuronalen Netzwerks auf einem Halbleiterchip sind Technologien für hohe Integration und niedrige Leistung vorteilhaft, um die zunehmenden Synapsengewichte und Biases zu speichern und eine Vielzahl von künstlichen Neuronen zu implementieren.

[0005] US 2019102170A1 offenbart eine Compute-in-Memory (CIM)-Schaltung, die einen Multiply-Accumulate (MAC)-Vorgang basierend auf einer stromerfassenden Auslesetechnik ermöglicht und einen Operationsverstärker, der mit einer Bitleitung einer Spalte von Bitzellen gekoppelt ist, die in einem Speicherarray der CIM-Schaltung enthalten sind, um zu bewirken, dass die Bitzellen wie ideale Stromquellen fungieren, um einen analogen Spannungswert zu bestimmen, der vom Operationsverstärker für gegebene, in der gespeicherte Zustände ausgegeben wird Bitzellen und für gegebene Eingabeaktivierungen für die Bitzellen. Der analoge Spannungswert, wird von der Verarbeitungsschaltung der CIM-Schaltung erfasst und in einen digitalen Wert

umgewandelt, um einen MAC-Wert (Multiply-Accumulate) zu berechnen.

[0006] US 6654730B 1 offenbart eine Rechenvorrichtung für ein neuronales Netzwerk die Neuronenoperationen parallel durch mehrere Recheneinheiten ausführt und ein Betriebsverfahren für ein neuronales Netzwerk.

[0007] DE69028617T2 offenbart ein Datenverarbeitungssystem mit einem darin verpackten Speicher zur Realisierung einer groß angelegten und schnellen parallelen verteilten Verarbeitung und insbesondere ein Datenverarbeitungssystem für die neuronale Netzwerkverarbeitung.

[0008] US10360971B1 offenbart Techniken für die Funktionalität künstlicher neuronaler Netzwerke innerhalb eines dynamischen Direktzugriffsspeichers. Innerhalb eines Speicherblocks wird auf eine Vielzahl dynamischer Direktzugriffszellen zugegriffen. Daten innerhalb der mehreren dynamischen Direktzugriffszellen werden mithilfe mehrerer Leseverstärker erfasst, die den mehreren dynamischen Direktzugriffszellen zugeordnet sind. Mehrere Auswahlleitungen, die mit den mehreren Leseverstärkern gekoppelt sind, werden aktiviert, um das Lesen der Daten innerhalb der mehreren dynamischen Direktzugriffszellen zu erleichtern, wobei die Aktivierung eine Funktion von Eingaben in eine Schicht innerhalb eines neuronalen Netzwerks ist.

[0009] Die Veröffentlichung Yu, Shimeng: „Neuro-Inspired Computing With Emerging Nonvolatile Memory“ offenbart eine zeitgemäße Übersicht über den Stand der Technik, die Herausforderungen und die Perspektiven des neuroinspirierten Computings mit neuen nichtflüchtigen Speichergeräten.

ZUSAMMENFASSUNG

[0010] Ausführungsformen der erfinderischen Konzepte sehen eine speicherbasierte neuromorphe Vorrichtung vor.

[0011] Nach einer Beispielausführungsform kann eine neuromorphe Vorrichtung ein Speicherzellenarray enthalten, das erste Speicherzellen enthält, die einer ersten Adresse entsprechen, wobei die ersten Speicherzellen eingerichtet sind, um erste Gewichte (Wichtungen) zu speichern, und zweite Speicherzellen, die einer zweiten Adresse entsprechen, wobei die zweiten Speicherzellen eingerichtet sind, um zweite Gewichte zu speichern, und eine erste Neuronenschaltung, die einen ersten Integrator und eine erste Aktivierungsschaltung enthält, wobei der erste Integrator eingerichtet sind, um erste Lesesignale aus den ersten Speicherzellen zu summieren und ein erstes Summensignal an die erste Aktivierungsschaltung auszugeben und zweite Lesesig-

nale aus den zweiten Speicherzellen zu summieren und zweite Summensignale an die erste Aktivierungsschaltung auszugeben, und die erste Aktivierungsschaltung eingerichtet ist, um ein erstes Aktivierungssignal auf der Grundlage des ersten Summensignals auszugeben und ein zweites Aktivierungssignal auf der Grundlage der zweiten Summensignale der ersten Lesesignale auszugeben, die von dem Integrator ausgegeben werden. Die Erfindung ist in den beigefügten unabhängigen Ansprüchen definiert. Weiterentwicklungen der Erfindung sind in den abhängigen Ansprüchen angegeben.

KURZE BESCHREIBUNG DER ZEICHNUNGEN

[0012] Die vorstehenden und andere Objekte und Merkmale der erfinderischen Konzepte werden durch eine detaillierte Beschreibung von Beispielausführungsformen davon unter Bezugnahme auf die beigefügten Zeichnungen deutlich.

Fig. 1 veranschaulicht ein künstliches neuronales Netzwerk nach einer Ausführungsform der erfinderischen Konzepte;

Fig. 2 veranschaulicht ein Blockdiagramm einer neuromorphen Vorrichtung nach einer Ausführungsform der erfinderischen Konzepte;

Fig. 3 veranschaulicht ein Blockdiagramm einer neuromorphen Vorrichtung nach einer anderen Ausführungsform der erfinderischen Konzepte;

Fig. 4 veranschaulicht ein Blockdiagramm einer neuromorphen Vorrichtung nach einer anderen Ausführungsform der erfinderischen Konzepte;

Fig. 5 veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung nach einer anderen Ausführungsform der erfinderischen Konzepte;

Fig. 6 veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung nach einer anderen Ausführungsform der erfinderischen Konzepte;

Fig. 7 veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung aus **Fig. 5**;

Fig. 8 veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung aus **Fig. 6**;

Fig. 9 veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung aus **Fig. 7**;

Fig. 10 veranschaulicht ein Timing-Diagramm einer Operation einer neuromorphen Vorrichtung von **Fig. 9**;

Fig. 11 veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung aus **Fig. 7**;

Fig. 12 veranschaulicht ein Timing-Diagramm einer Operation einer neuromorphen Vorrichtung von **Fig. 11**;

Fig. 13 veranschaulicht einen Schaltplan eines Integrators nach einer Ausführungsform der erfinderischen Konzepte;

Fig. 14 veranschaulicht einen Schaltplan einer Summenschaltung nach einer Ausführungsform der erfinderischen Konzepte;

Fig. 15 veranschaulicht einen Schaltplan einer Aktivierungsschaltung nach einer Ausführungsform der erfinderischen Konzepte;

Fig. 16 veranschaulicht ein Blockdiagramm einer neuromorphen Vorrichtung nach einer Ausführungsform der erfinderischen Konzepte;

Fig. 17 veranschaulicht ein künstliches neuronales Netzwerk, das von einer neuromorphen Vorrichtung aus **Fig. 16** ausgeführt werden kann;

Fig. 18 ist ein Ablaufdiagramm, das ein Betriebsverfahren einer neuromorphen Vorrichtung von **Fig. 16** veranschaulicht;

Fig. 19 veranschaulicht ein Blockdiagramm einer neuromorphen Vorrichtung nach einer Ausführungsform der erfinderischen Konzepte; und

Fig. 20 veranschaulicht ein Blockdiagramm eines System-on-Chips nach einer Ausführungsform der erfinderischen Konzepte.

DETAILLIERTE BESCHREIBUNG

[0013] Wenn die Begriffe „ungefähr“ oder „im Wesentlichen“ in dieser Spezifikation in Verbindung mit einem Zahlenwert verwendet werden, ist beabsichtigt, dass der zugehörige Zahlenwert eine Fertigungstoleranz (z. B. $\pm 10\%$) um den angegebenen Zahlenwert herum enthält. Wenn darüber hinaus die Wörter „im Allgemeinen“ und „im Wesentlichen“ in Verbindung mit geometrischen Formen verwendet werden, ist beabsichtigt, dass die Genauigkeit der geometrischen Form nicht erforderlich ist, sondern dass der Spielraum für die Form im Rahmen der Offenbarung liegt. Unabhängig davon, ob numerische Werte oder Formen als „ungefähr“ oder „wesentlich“ modifiziert werden, ist es ferner so zu verstehen, dass diese Werte und Formen so zu verstehen sind, dass sie eine Fertigungs- oder Betriebstoleranz (z. B. $\pm 10\%$) um die angegebenen numerischen Werte oder Formen einschließen.

[0014] Fig. 1 veranschaulicht ein künstliches neuronales Netzwerk nach einer Ausführungsform der erfinderischen Konzepte.

[0015] Das künstliche neuronale Netzwerk ANN kann eine Eingabeschicht IL, zumindest eine verborgene Schicht HL und eine Ausgabeschicht OL enthalten. Das künstliche neuronale Netzwerk ANN kann ein zum Beispiel tiefgehendes neuronales Netzwerk (DNN), ein rekurrentes neuronales Netzwerk (RNN), ein faltendes neuronales Netzwerk (CNN), ein gepulstes neuronales Netzwerk (SNN, Spiking Neural Network) usw. sein. Die Eingabeschicht IL kann eine Vielzahl von Eingabeknoten IN1 bis IN3 enthalten und Eingabedaten IN empfangen. Die Ausgabeschicht OL kann eine Vielzahl von Ausgabeknoten ON1 bis ON3 enthalten und kann Ausgabedaten OUT auf der Grundlage von Daten ausgeben, die durch zumindest eine verborgene Schicht HL hindurchgehen. Die zumindest eine verborgene Schicht HL kann eine Vielzahl von verborgenen Knoten HN1 bis HN8 enthalten und kann die Verbindungen zwischen der Eingabeschicht IL und der Ausgabeschicht OL einrichten. Obwohl die in Fig. 1 dargestellte Beispielanzahl der Knoten IN1 bis IN3, HN1 bis HN8 und ON1 bis ON3, aus denen die jeweiligen Schichten IL, HL und OL bestehen, die Anzahl der verborgenen Schichten HL und die Verbindungsbeziehungen zwischen den Knoten IN1 bis IN3, HN1 bis HN8 und ON1 bis ON3 in den vorliegenden Ausführungsformen dargestellt sind, kann die Anzahl der Knoten, Schichten und Verbindungen mehr oder weniger groß sein. Beispielsweise können ein oder mehrere Modusknoten IN1 bis IN3, HN1 bis HN8 und ON1 bis ON3, eine oder mehrere verborgene Schichten HL und eine oder mehrere Verbindungsbeziehungen zwischen den Knoten IN1 bis IN3, HN1 bis HN8 und ON1 bis ON3 ferner in das künstliche neuronale Netzwerk ANN einbezogen werden. Ein Knoten kann ein künstliches Neuron sein, das einem Neuron eines biologischen neuronalen Netzwerks entspricht und als „Neuronenschaltung“ oder „Perzeptron“ bezeichnet werden kann. Die Verbindungsstärke zwischen Neuronenschaltungen kann einer Verbindungsstärke einer Synapse des biologischen neuronalen Netzwerks entsprechen, und die Verbindungsstärke zwischen Neuronenschaltungen kann durch ein Gewicht (z. B. ein Synapsengewicht) ausgedrückt werden. Wie im Neuron des biologischen neuronalen Netzwerks kann eine Neuronenschaltung auf der Grundlage von Signalen von anderen Neuronenschaltungen, Gewichten für die anderen Neuronenschaltungen, Biases und einem Schwellenwert der Neuronenschaltung ein Aktivierungssignal an andere Neuronenschaltungen ausgeben. Eine Hardware-Vorrichtung, in der das künstliche neuronale Netzwerk ANN implementiert ist und die das künstliche neuronale Netzwerk ANN ausführt, wird im Folgenden beschrieben.

[0016] Fig. 2 veranschaulicht ein Blockdiagramm einer neuromorphen Vorrichtung nach einer Ausführungsform der erfinderischen Konzepte.

[0017] Eine neuromorphe Vorrichtung 1000a kann eine „neuromorphe Rechenvorrichtung“, ein „neuromorpher Chip“, ein „neuromorpher Prozessor“, ein „neuromorphes System“, ein „Beschleuniger für neuronale Netzwerke“, ein „Beschleunigerchip für neuronale Netzwerke“, ein „Prozessor für künstliche Intelligenz (KI)“, eine „KI-Speichervorrichtung“, ein „Prozessor“, eine „Speichervorrichtung“ oder eine Kombination davon sein. Die neuromorphe Vorrichtung 1000a kann eine speicherbasierte neuromorphe Vorrichtung sein, die eine Struktur oder eine Architektur einer Daten speichernden Speichervorrichtung verwendet. In einigen Beispielausführungsformen kann die neuromorphe Vorrichtung 1000a eine Vielzahl von Verarbeitungselementen enthalten, die gleichzeitig und/oder simultan die Verarbeitung des neuronalen Netzwerks durchführen, wie z. B. einen Satz von Verarbeitungselementen, die gleichzeitig und/oder simultan Gewichte auf mehreren Kanälen summieren. In einigen Beispielausführungsformen kann die neuromorphe Vorrichtung 1000a eingerichtet sein, um das neuronale Netzwerk nacheinander zu verarbeiten, wie z. B. eine Sequenz von Operationen für jeden von mehreren Kanälen, wie im Folgenden beschrieben wird. Die neuromorphe Vorrichtung 1000a kann ein Speicherzellenarray 1100 und eine Neuronenschaltung 1200 enthalten.

[0018] Das Speicherzellenarray 1100 kann die Speicherzellen MC1 bis MC4 enthalten. Zum Beispiel kann jede der Speicherzellen MC1 bis MC4 eine statische Direktzugriffsspeicher (SRAM)-Zelle, eine dynamische Direktzugriffsspeicher (DRAM)-Zelle, eine Thyristor-Direktzugriffsspeicher (TRAM)-Zelle, eine NAND-Flash-Speicherzelle, eine NOR-Flash-Speicherzelle, eine resistive Direktzugriffsspeicher (RRAM)-Zelle, eine ferroelektrische Direktzugriffsspeicher (FRAM)-Zelle, eine Phasenänderungs-Direktzugriffsspeicher (PRAM)-Zelle und eine magnetische Direktzugriffsspeicher (MRAM)-Zelle sein. Jede der Speicherzellen MC1 bis MC4 kann eine DRAM-Zelle mit einem Zelltransistor „T“ und einem Zellkondensator „C“ sein, und die neuromorphe Vorrichtung 1000a kann eine DRAM-basierte neuromorphe Vorrichtung sein. Die erfinderischen Konzepte sind jedoch nicht darauf beschränkt. Beispielsweise kann jede der Speicherzellen MC1 bis MC4 ein einzelnes Bit speichern. Ein Zellspannungspegel Vcell (z. B. eine Speicherspannung) jeder der Speicherzellen MC1 bis MC4 kann einem der Spannungspegel VS0 und VS1 entsprechen, die logische Werte (z. B. „0“ und „1“) angeben, die durch ein einzelnes Bit ausgedrückt werden und sich voneinander unterscheiden. Beispielsweise kann jede der Speicherzellen MC1 bis MC4 einer Einzelpegelzelle (SLC) entsprechen. Die Gewichte W_1 bis W_4 können jeweils über

eine Ein-/Ausgabelitung IO (z. B. eine Ein- und Ausgabelitung IO) in die Speicherzellen MC1 bis MC4 geschrieben werden. Die Speicherzellen MC1 bis MC4 können jeweils die Gewichte W_1 bis W_4 speichern. Jedes der Gewichte W_1 bis W_4 kann eine Verbindungsstärke der Neuronenschaltung 1200 anzeigen. Obwohl in **Fig. 2** nicht dargestellt, kann zumindest eine der Speicherzellen MC1 bis MC4 anstelle eines Gewichts einen Bias speichern, oder eine andere Speicherzelle als die Speicherzellen MC1 bis MC4 kann einen Bias speichern. Beispielsweise kann die Anzahl der Bits eines Gewichts und die Anzahl der Bits eines Bias der Anzahl der Bits entsprechen, die jede der Speicherzellen MC1 bis MC4 speichern kann. Auf die Speicherzellen MC1 bis MC4 kann zum Beispiel jeweils über eine Ein-/Ausgabelitung IO zugegriffen werden. Obwohl die Anzahl der Speicherzellen MC1 bis MC4 in der vorliegenden Ausführungsform dargestellt ist, kann die Anzahl der Speicherzellen mehr oder weniger als dargestellt sein, z. B. können eine oder mehrere Speicherzellen weiter im Speicherzellenarray 1100 enthalten sein. Entsprechende Gewichte und/oder Biases können jeweils in die eine oder mehrere weiter eingeschlossene Speicherzellen geschrieben und/oder gespeichert werden.

[0019] Die Neuronenschaltung 1200 kann einen Integrator 1300 und eine Aktivierungsschaltung 1400 enthalten. Der Integrator 1300 kann die Gewichte W_1 bis W_4 empfangen, die von den Speicherzellen MC1 bis MC4 über die Ein-/Ausgabelitung IO übertragen werden. Der Integrator 1300 kann zumindest einen Teil der Gewichte W_1 bis W_4 integrieren, summieren und/oder akkumulieren. In dem Fall, dass ein Teil der Speicherzellen MC1 bis MC4 einen Bias anstelle eines Gewichts speichert, kann der Integrator 1300 das/die Gewicht(e) und die Bias(en) empfangen und summieren. Der Integrator 1300 kann als „Akkumulator“ oder „Addierer“ bezeichnet werden. Zum Beispiel kann der Integrator 1300 zumindest ein oder mehrere Gewichte der Gewichte W_1 bis W_4 an der Ein-/Ausgabelitung IO summieren. Der Integrator 1300 kann mehrere Bits der Gewichte an der einen Ein-/Ausgabelitung IO summieren. Der Integrator 1300 kann ein Summensignal, das ein Summenergebnis anzeigt, der Aktivierungsschaltung 1400 zuführen.

[0020] Die Aktivierungsschaltung 1400 kann auf der Grundlage des Summensignals des Integrators 1300 ein Aktivierungssignal ACTa/ACTb ausgeben, erzeugen, aktivieren und/oder auslösen. Die Aktivierungsschaltung 1400 kann zum Beispiel das Aktivierungssignal ACTa/ACTb ausgeben, indem sie ein Schwellwertsignal TH (oder als „Schwellwert“, „Referenzsignal“ oder „Referenzwert“ bezeichnet) mit dem Summensignal des Integrators 1300 vergleicht. Wenn das Summensignal des Integrators 1300 das Schwellwertsignal TH überschreitet, kann die Aktivi-

erungsschaltung 1400 das Aktivierungssignal ACTa/ACTb ausgeben und/oder aktivieren, oder, zum Beispiel, wenn das Summensignal des Integrators 1300 das Schwellwertsignal TH nicht überschreitet, kann die Aktivierungsschaltung 1400 das Aktivierungssignal ACTa/ACTb deaktivieren und/oder nicht ausgeben. Das Aktivierungssignal ACTa/ACTb kann als „Ausgabesignal“ bezeichnet werden. Die Aktivierungsschaltung 1400 kann das Aktivierungssignal ACTa/ACTb unter Verwendung verschiedener Aktivierungsfunktionen, wie z. B. einer Schrittfunktion, einer Sigmoidfunktion, einer Funktion der gleichgerichteten Lineareinheit (ReLU), einer Leaky-ReLU-Funktion und einer Tanh-Funktion, ausgeben.

[0021] Da zum Beispiel zumindest ein oder mehrere Gewichte wiederholt über die Ein-/Ausgabelitung IO übertragen werden, kann die Neuronenschaltung 1200 wiederverwendet werden. Der Integrator 1300 kann Lesesignale aus den Speicherzellen MC1 und MC3, die die Gewichte W_1 und W_3 an der einen Ein-/Ausgabelitung IO (①) speichern, summieren. Die Aktivierungsschaltung 1400 kann das Aktivierungssignal ACTa ausgeben, indem sie ein Summensignal, das der Integrator 1300 als Summe der Gewichte W_1 und W_3 ausgibt, mit dem Schwellwertsignal TH vergleicht. Als nächstes kann der Integrator 1300 Lesesignale aus den Speicherzellen MC2 und MC4, die die Gewichte W_2 und W_4 an der einen Ein-/Ausgabelitung IO (②) speichern, weiter summieren. Die Aktivierungsschaltung 1400 kann ferner das Aktivierungssignal ACTb ausgeben, indem sie ein Summensignal, das der Integrator 1300 als Summe der Gewichte W_2 und W_4 ausgibt, mit dem Schwellwertsignal TH vergleicht. Da die eine Neuronenschaltung 1200 wiederholt und unterschiedlich zumindest einen Teil der unterschiedlichen Gewichte W_1 bis W_4 über die Ein-/Ausgabelitung IO empfängt und wiederholt und unterschiedlich die Aktivierungssignale ACTa und ACTb ausgibt, kann die Neuronenschaltung 1200 wiederverwendet werden. Mindestens eines der Gewichte W_2 und W_4 kann die Verbindungsstärke zwischen der Neuronenschaltung 1200, die die Lesesignale von den Speicherzellen MC1 und MC3 empfängt, und der Neuronenschaltung 1200, die die Lesesignale von den Speicherzellen MC2 und MC4 empfängt, anzeigen. Die beiden vorstehend beschriebenen Neuronenschaltungen 1200 sind Beispiele für eine physikalisch implementierte Neuronenschaltung 1200, aber die Neuronenschaltung 1200 kann als zwei oder mehr Knoten des künstlichen neuronalen Netzwerks ANN arbeiten und kann wiederverwendet werden.

[0022] In einer Beispielausführungsform können die Speicherzellen MC1 bis MC4 ausgewählt werden, um die Lesesignale von den Speicherzellen MC1 bis MC4 auszugeben. Im Fall von „①“ können die Speicherzellen MC1 und MC3 ausgewählt werden,

und im Fall von „Ⓜ“ können die Speicherzellen MC2 und MC4 ausgewählt werden. Die Lesesignale können beispielsweise aus den Speicherzellen MC1 bis MC4 den Produkten von Adressen zur Auswahl der Speicherzellen MC1 bis MC4 und den in den Speicherzellen MC1 bis MC4 gespeicherten Gewichten W_1 bis W_4 entsprechen. In einem anderen Beispiel können die Lesesignale aus den Speicherzellen MC1 bis MC4 Gewichte und/oder Biases enthalten, die in den Speicherzellen MC1 bis MC4 gespeichert sind. Die Anzahl der Speicherzellen des Speicherzellenarrays 1100, eine Stelle einer ausgewählten Speicherzelle, die Anzahl der ausgewählten Speicherzellen usw. sind nur Beispiele und können entsprechend angepasst werden.

[0023] Fig. 3 veranschaulicht ein Blockdiagramm einer neuromorphen Vorrichtung nach einer anderen Ausführungsform der erfinderschen Konzepte. Eine zusätzliche Beschreibung in Verbindung mit Komponenten, die die gleichen Referenznummern aufweisen, wird weggelassen, um Redundanz zu vermeiden, und es werden hauptsächlich die Unterschiede zwischen den neuromorphen Vorrichtungen 1000a und 1000b beschrieben.

[0024] Die neuromorphe Vorrichtung 1000b kann das Speicherzellenarray 1100b und die Neuronenschaltung 1200 enthalten. Das Speicherzellenarray 1100b kann die Speicherzellen MC1b bis MC4b enthalten. Wie bei der neuromorphen Vorrichtung 1000a kann jede der Speicherzellen MC1b bis MC4b eine DRAM-Zelle sein, die den Zelltransistor „T“ und den Zellkondensator „C“ enthält, aber jede der Speicherzellen MC1b bis MC4b kann zwei Bits speichern. Der Zellspannungspegel V_{cell} jeder der Speicherzellen MC1b bis MC4b kann einem der Spannungspegel VS00, VS01, VS10 und VS11 entsprechen, die logische Werte (z. B. 00, 01, 10 und 11) angeben, die durch zwei Bits ausgedrückt werden und sich voneinander unterscheiden. In diesem Fall kann ein Gewicht und/oder Bias, die in jeder der Speicherzellen MC1b bis MC4b gespeichert sind, aus zwei Bits bestehen. In einer Ausführungsform kann jede der Speicherzellen MC1b bis MC4b zwei oder mehr Bits speichern, der Zellspannungspegel V_{cell} kann weiter unterteilt sein, um logische Werte anzuzeigen, die durch zwei oder mehr Bits ausgedrückt werden, und ein Gewicht und/oder Bias kann zwei oder mehr Bits aufweisen. Jede der Speicherzellen MC1b bis MC4b kann einer mehrstufigen Zelle (MLC), einer dreistufigen Zelle (TLC), einer vierstufigen Zelle (QLC) usw. entsprechen.

[0025] Die Neuronenschaltung 1200 kann unabhängig von der Anzahl der Bits, die jede der Speicherzellen MC1b bis MC4b speichern kann, wiederverwendet werden. Der Integrator 1300 kann Lesesignale aus den Speicherzellen MC1b und MC3b summieren, wobei die Gewichte W_1b und W_3b jeweils ein

Multi-Bit an der einen Ein-/Ausgabeleitung IO (Ⓜ) speichern. Die Aktivierungsschaltung 1400 kann das Aktivierungssignal ACTa ausgeben, indem sie ein Summensignal, das der Integrator 1300 als Summe der Gewichte W_1b und W_3b ausgibt, mit dem Schwellwertsignal TH vergleicht. Als nächstes kann der Integrator 1300 Lesesignale aus den Speicherzellen MC2 und MC4, in denen die Gewichte W_2b und W_4b gespeichert sind, die jeweils ein Multi-Bit an der einen Ein-/Ausgabeleitung IO (Ⓜ) sind, weiter summieren. Die Aktivierungsschaltung 1400 kann ferner das Aktivierungssignal ACTb ausgeben, indem sie ein Summensignal, das der Integrator 1300 als Summe der Gewichte W_2b und W_4b ausgibt, mit dem Schwellwertsignal TH vergleicht. In der vorstehenden Beschreibung kann eine Speicherzelle, sofern nicht anders definiert, ein einzelnes Bit oder zwei oder mehr Bits speichern.

[0026] Fig. 4 veranschaulicht ein Blockdiagramm einer neuromorphen Vorrichtung 1000c nach einer anderen Ausführungsform der erfinderschen Konzepte. Zusätzliche Beschreibungen in Verbindung mit Komponenten, die die gleichen Referenznummern aufweisen, werden ausgelassen, um Redundanz zu vermeiden, und es werden hauptsächlich die Unterschiede zwischen den neuromorphen Vorrichtungen 1000a, 1000b und 1000c beschrieben.

[0027] Das Speicherzellenarray 1100c kann die Unter-Speicherblöcke 1110_1 bis 1110_2 enthalten. Alle Speicherzellen des Speicherzellenarrays 1100 können in die Unter-Speicherblöcke 1110_1 bis 1110_2 unterteilt werden. Der Unter-Speicherblock 1110_1 kann die Speicherzellen MC1 und MC2 enthalten. Die Speicherzellen MC1 und MC2 können jeweils mit Wortleitungen WL1 und WL2 verbunden (z. B. gekoppelt) und gemeinsam mit einer Bitleitung BL1 verbunden sein. Der Unter-Speicherblock 1110_2 kann die Speicherzellen MC3 und MC4 enthalten. Die Speicherzellen MC3 und MC4 können jeweils mit den Wortleitungen WL3 und WL4 verbunden sein und gemeinsam mit einer Bitleitung BL2 verbunden werden.

[0028] Das Speicherzellenarray 1100c kann einen Bitleitungs-Leseverstärker 1111_1 enthalten, der mit der Bitleitung BL1 verbunden ist. Der Bitleitungs-Leseverstärker 1111_1 kann einen Spannungspegel der Bitleitung BL1 lesen und/oder verstärken, der sich ändern kann, wenn eine der Speicherzellen MC1 und MC2 ausgewählt wird. Beispielsweise kann nur eine der Speicherzellen MC1 und MC2, die mit dem Bitleitungs-Leseverstärker 1111_1 verbunden sind, ausgewählt werden. Das Speicherzellenarray 1100 kann ferner einen Bitleitungs-Leseverstärker 1111_2 enthalten, der mit der Bitleitung BL2 verbunden ist. Der Bitleitungs-Leseverstärker 1111_2 kann einen Spannungspegel der Bitleitung BL2 lesen und verstärken, der sich ändern kann,

wenn eine der Speicherzellen MC3 und MC4 ausgewählt wird. Beispielsweise kann nur eine der Speicherzellen MC3 und MC4, die mit dem Bitleitungs-Leseverstärker 1111_2 verbunden sind, ausgewählt werden. Obwohl das Speicherzellenarray 1100c als die Speicherzellen MC1 bis MC4 enthaltend veranschaulicht wird, kann das Speicherzellenarray z. B. die Speicherzellen MC1b bis MC4b enthalten.

[0029] Der Bitleitungs-Leseverstärker 1111_1 kann ein Gewicht und/oder einen Bias wiederherstellen, indem er erneut einen Spannungspegel speichert. Der gespeicherte Spannungspegel kann durch Abtasten und/oder Verstärken des Spannungspegels der Bitleitung BL1 in der aus den Speicherzellen MC1 und MC2 ausgewählten Speicherzelle erhalten werden. In ähnlicher Weise kann der Bitleitungs-Leseverstärker 1111_2 ein Gewicht und/oder einen Bias wiederherstellen, indem er einen Spannungspegel, der durch Abtasten und/oder Verstärken des Spannungspegels der Bitleitung BL2 erhalten werden kann, in der aus den Speicherzellen MC3 und MC4 ausgewählten Speicherzelle speichert. Eine Zugriffsoperation (z. B. Lesen) auf die ausgewählten Speicherzellen kann zerstörungsfrei sein.

[0030] Das Speicherzellenarray 1100c kann ferner einen Spaltenauswahlschalter 1112_1 enthalten, der die Bitleitung BL1 elektrisch mit der Ein-/Ausgabeleitung IO verbindet. Der Spaltenauswahlschalter 1112_1 kann die Bitleitung BL1 basierend auf einem Spaltenauswahlsignal, das durch eine Spaltenauswahlleitung CSL übertragen wird, elektrisch mit der Ein-/Ausgabeleitung IO verbinden. Das Speicherzellenarray 1100c kann ferner einen Spaltenauswahlschalter 1112_2 enthalten, der die Bitleitung BL2 elektrisch mit der Ein-/Ausgabeleitung IO verbindet. Der Spaltenauswahlschalter 1112_2 kann die Bitleitung BL2 basierend auf dem durch die Spaltenauswahlleitung CSL übertragenen Spaltenauswahlsignal elektrisch mit der Ein-/Ausgabeleitung IO verbinden. Beispielsweise kann jeder der Spaltenauswahlschalter 1112_1 und 1112_2 mit derselben Spaltenauswahlleitung CSL verbunden werden und dasselbe Spaltenauswahlsignal über dieselbe Spaltenauswahlleitung CSL empfangen. Das Spaltenauswahlsignal kann es ermöglichen, dass jede der Bitleitungen BL1 und BL2 elektrisch mit derselben Ein-/Ausgabeleitung IO verbunden wird.

[0031] Die Anzahl der Unter-Speicherblöcke des Speicherzellenarrays 1100, die Anzahl der Speicherzellen pro Unter-Speicherblock, die Anzahl der Wortleitungen pro Unter-Speicherblock und die Anzahl der Bitleitungen pro Unter-Speicherblock sind jedoch nur eine beispielhafte Ausführungsform, und ihre Anzahl kann mehr oder weniger als dargestellt sein. Außerdem können z. B. die Bitleitungs-Leseverstärker 1111_1 neben dem Unter-Speicherblock 1110_1 so viel wie die Anzahl der zum Unter-Speicherblock

1110_1 gehörenden Bitleitungen BL1 angeordnet sein, und die Bitleitungs-Leseverstärker 1111_2 können neben dem Unter-Speicherblock 1110_2 so viel wie die Anzahl der zum Unter-Speicherblock 1110_2 gehörenden Bitleitungen BL2 angeordnet sein. Die Bitleitungs-Leseverstärker 1111_1 und 1111_2 und die Spaltenauswahlschalter 1112_1 und 1112_2 können zwischen den Unter-Speicherblöcken 1110_1 bis 1110_2 angeordnet werden. Der Unter-Speicherblock 1110_1 kann zwischen dem Bitleitungs-Leseverstärker 1111_1 und einem weiteren Bitleitungs-Leseverstärker neben dem Unter-Speicherblock 1110_1 zwischengeschaltet werden. Die Speicherzellen MC1 bis MC4 können identisch zueinander implementiert werden, die Unter-Speicherblöcke 1110_1 bis 1110_2 können identisch zueinander implementiert werden, die Bitleitungs-Leseverstärker 1111_1 bis 1111_2 können identisch zueinander implementiert werden, und die Spaltenauswahlschalter 1112_1 bis 1112_2 können identisch zueinander implementiert werden.

[0032] Die neuromorphe Vorrichtung 1000c kann ferner einen Zeilen-/Spalten-Controller 1500, der einen Adressgenerator 1510 enthält, einen Zeilendekodierer 1600 und einen Spaltendekodierer 1700 enthalten. Der Zeilen-/Spalten-Controller 1500 kann den Zeilendekodierer 1600 und den Spaltendekodierer 1700 auf der Grundlage der Aktivierungssignale ACTa und ACTb steuern. Der Adressgenerator 1510 kann auf der Grundlage der Aktivierungssignale ACTa und ACTb eine Adresse zur Auswahl zumindest eines Teils der Speicherzellen MC1 bis MC4 des Speicherzellenarrays 1100 erzeugen. Der Adressgenerator 1510 kann immer dann eine interne Zeilenadresse IRA und eine interne Spaltenadresse ICA erzeugen, wenn jedes der Aktivierungssignale ACTa und ACTb empfangen wird. Der Adressgenerator 1510 kann zum Beispiel einen Zähler enthalten, der als Antwort auf die Aktivierungssignale ACTa und ACTb einen Zählvorgang durchführt und einen Zählwert (oder eine Adresse) erzeugt. Außerdem kann der Zeilen-/Spalten-Controller 1500 auf der Grundlage der Aktivierungssignale ACTa und ACTb dem Zeilendekodierer 1600 ein Steuersignal zur Steuerung des Zeilendekodierers 1600 zuführen, so dass Wortleitungen entsprechend der internen Zeilenadresse IRA ausgewählt werden. Auf der Grundlage der Aktivierungssignale ACTa und ACTb kann der Zeilen-/Spalten-Controller 1500 dem Spaltendekodierer 1700 ein Steuersignal zur Steuerung des Spaltendekodierers 1700 zuführen, so dass Spaltenauswahlleitungen entsprechend der internen Spaltenadresse ICA ausgewählt werden.

[0033] Der Zeilendekodierer 1600 kann die interne Zeilenadresse IRA dekodieren und zumindest eine der Wortleitungen WL1 bis WL4 auswählen, die mit den Speicherzellen MC1 bis MC4 des Speicherzellenarrays 1100c verbunden sind. Zum Beispiel kann

der Zeilendekodierer 1600 im Fall von „①“ aus **Fig. 2** und/oder **Fig. 3** die interne Zeilenadresse IRA dekodieren und kann die Wortleitungen WL1 und WL3 auswählen. Zum Beispiel kann der Zeilendekodierer 1600 im Fall von „②“ aus **Fig. 2** und/oder **Fig. 3** die interne Zeilenadresse IRA dekodieren und die Wortleitungen WL2 und WL4 auswählen. Die dem Zeilendekodierer 1600 zugeführte interne Zeilenadresse IRA im Fall von „①“ und die dem Zeilendekodierer 1600 zugeführte interne Zeilenadresse IRA im Fall von „②“ können sich voneinander unterscheiden.

[0034] Der Zeilendekodierer 1600 kann die Unter-Zeilendekodierer 1610_1 bis 1610_2 und einen Vor-Zeilendekodierer 1620 enthalten. Beispielsweise kann der Unter-Zeilendekodierer 1610_1 die interne Zeilenadresse IRA dekodieren und eine der zum Unter-Speicherblock 1110_1 gehörenden Wortleitungen WL1 und WL2 auswählen. Der Unter-Zeilendekodierer 1610_2 kann die interne Zeilenadresse IRA dekodieren und eine der zum Unter-Speicherblock 1110_2 gehörenden Wortleitungen WL3 und WL4 auswählen. Die Unter-Zeilendekodierer 1610_1 bis 1610_2 können so implementiert werden, dass sie zueinander identisch sind. Der Zeilendekodierer 1600 kann so viele Unter-Zeilendekodierer enthalten wie die Anzahl der Unter-Speicherblöcke des Speicherzellenarrays 1100.

[0035] Der Vor-Zeilendekodierer 1620 kann die interne Zeilenadresse IRA vor den Unter-Zeilendekodierern 1610_1 bis 1610_2 dekodieren und/oder die Unter-Zeilendekodierer 1610_1 bis 1610_2 auswählen. Beispielsweise kann der Vor-Zeilendekodierer 1620 ein oberes Bit eines Zeilenadressenbits, das zur Auswahl von Wortleitungen jedes der Unter-Speicherblöcke 1110_1 und 1110_2 verwendet wird, auf ein Don't-Care-Bit setzen. In diesem Fall kann der Vor-Zeilendekodierer 1620 gleichzeitig zwei oder mehr Unter-Zeilendekodierer 1610_1 und 1610_2 auswählen. Im Falle von „①“ in **Fig. 2** und/oder **Fig. 3** kann der Vor-Zeilendekodierer 1620 zur Auswahl der Speicherzellen MC1 und MC3 gleichzeitig die Unter-Zeilendekodierer 1610_1 und 1610_2 auswählen und die Unter-Zeilendekodierer 1610_1 und 1610_2 können jeweils die Wortleitungen WL1 bzw. WL3 auswählen. Im Falle von „②“ in **Fig. 2** und/oder **Fig. 3** kann der Vor-Zeilendekodierer 1620 zur Auswahl der Speicherzellen MC2 und MC4 gleichzeitig die Unter-Zeilendekodierer 1610_1 und 1610_2 auswählen und die Unter-Zeilendekodierer 1610_1 und 1610_2 können die Wortleitungen WL2 bzw. WL4 auswählen. In einem anderen Beispiel kann der Vor-Zeilendekodierer 1620 einen der Unter-Zeilendekodierer 1610_1 und 1610_2 auswählen. In diesem Fall kann eine Wortleitung aus dem Speicherzellenarray 1100 ausgewählt werden.

[0036] Der Spaltendekodierer 1700 kann die interne Spaltenadresse ICA dekodieren und das Spalte-

nauswahlsignal über die Spaltenauswahlleitung CSL übertragen (z. B. ausgeben). Beispielsweise kann der Spaltendekodierer 1700 einen Auswahlvorgang starten, nachdem ein Auswahlvorgang des Zeilendekodierers 1600 eingeleitet wurde. Wie vorstehend beschrieben, kann der Spaltendekodierer 1700 das Spaltenauswahlsignal an jeden der Spaltenauswahlschalter 1112_1 bis 1112_2 übertragen. Der Spaltendekodierer 1700 kann die Bitleitungen BL1 bis BL2 auswählen, indem er das Spaltenauswahlsignal überträgt. Wenn der Zeilendekodierer 1600 also gleichzeitig zwei oder mehr Wortleitungen WL1 und WL3 oder WL2 und WL4 auswählt, dann erfassen und verstärken die Bitleitungs-Leseverstärker 1111_1 und 1111_2 die Gewichte W_1 und W_3 oder W_2 und W_4 der Speicherzellen MC1 und MC3 oder MC2 und MC4, und dann wird das Spaltenauswahlsignal an die Spaltenauswahlschalter 1112_1 und 1112_2 angelegt, die Gewichte W_1 und W_3 oder W_2 und W_4 können gleichzeitig an die Ein-/Ausgabeleitung IO übertragen (oder diesen zugeführt) werden. Der Integrator 1300 kann die Gewichte W_1 und W_3 an der Ein-/Ausgabeleitung IO oder die Gewichte W_2 und W_4 an der Ein-/Ausgabeleitung IO summieren. Da hier die Speicherzellen MC1 und MC2 mit der Bitleitung BL1, die Bitleitung BL1 mit dem Spaltenwahlschalter 1112_1, die Speicherzellen MC3 und MC4 mit der Bitleitung BL2 verbunden werden können, kann die Bitleitung BL1 mit dem Spaltenwahlschalter 1112_1 verbunden werden, die Bitleitung BL2 kann mit dem Spaltenauswahlschalter 1112_2 verbunden werden und die Spaltenauswahlschalter 1112_1 und 1112_2 können mit derselben Spaltenauswahlleitung CSL verbunden werden, wobei die internen Spaltenadressen ICA, die den Speicherzellen MC1 bis MC4 entsprechen, identisch sein können.

[0037] In einer Ausführungsform kann der Adressgenerator 1510 die interne Zeilenadresse IRA und die interne Spaltenadresse ICA nacheinander erhöhen oder verringern. Beispielsweise kann der Zeilendekodierer 1600 die interne Zeilenadresse IRA dekodieren und die Wortleitungen WL1 und WL3 auswählen, und der Spaltendekodierer 1700 kann die interne Spaltenadresse ICA dekodieren und die Spaltenauswahlleitung CSL auswählen. Als nächstes kann der Zeilendekodierer 1600 die interne Zeilenadresse IRA, die je nach Aktivierungssignal ACTa erhöht oder verringert wird, dekodieren und die Wortleitungen WL2 und WL4 auswählen, und der Spaltendekodierer 1700 kann die interne Spaltenadresse ICA dekodieren und die Spaltenauswahlleitung CSL auswählen. Beispielsweise können der Zeilendekodierer 1600 und der Spaltendekodierer 1700 nacheinander die Speicherzellen MC1 und MC3 und die Speicherzellen MC2 und MC4 auswählen.

[0038] Der Integrator 1300, die Aktivierungsschaltung 1400, der Zeilen-/Spaltendekodierer 1500, der Zeilendekodierer 1600, der Spaltendekodierer

1700, der Adressgenerator 1510 und der Vor-Zeilendekodierer 1620 können Verarbeitungsschaltungen, wie z. B. Hardware, die Logikschaltungen enthält, eine Hardware/Software-Kombination, wie z. B. ein Prozessor, der Software ausführt, oder eine Kombination davon umfassen. Zum Beispiel kann die Verarbeitungsschaltung insbesondere eine zentrale Verarbeitungseinheit (CPU), eine arithmetische Logikeinheit (ALU), einen digitalen Signalprozessor, einen Mikrocomputer, ein Field-Programmable-Gate-Array (FPGA) und eine programmierbare Logikeinheit, einen Mikroprozessor, eine anwendungsspezifische integrierte Schaltung (ASIC) usw. umfassen, ist aber nicht darauf beschränkt.

[0039] Fig. 5 veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung nach einer anderen Ausführungsform der erfinderischen Konzepte. Zusätzliche Beschreibung, die mit Komponenten mit den gleichen Referenznummern zusammenhängen, werden ausgelassen, um Redundanz zu vermeiden, und es werden hauptsächlich Unterschiede zwischen den neuromorphen Vorrichtungen 1000c und 1000d beschrieben.

[0040] Das Speicherzellenarray 1100d kann ferner die Unter-Speicherblöcke 1120_1 bis 1130_1 enthalten. Die Unter-Speicherblöcke 1120_1 bis 1130_1 können Speicherzellen MC enthalten, die mit den Wortleitungen WL1 und WL2 verbunden sind, die vom Unter-Zeilendekodierer 1610_1 ausgewählt werden, und können so implementiert werden, dass sie zueinander identisch sind. Das Speicherzellenarray 1100d kann ferner die Unter-Speicherblöcke 1120_2 bis 1130_2 enthalten. Die Unter-Speicherblöcke 1120_2 bis 1130_2 können Speicherzellen MC enthalten, die mit den Wortleitungen WL3 und WL4 verbunden sind, die vom Unter-Zeilendekodierer 1610_2 ausgewählt werden, und können so implementiert werden, dass sie zueinander identisch sind. Das Speicherzellenarray 1100d kann ferner Bitleitungs-Leseverstärker 1121_1 bis 1121_2 und 1131_1 bis 1131_2 enthalten, die im Wesentlichen identisch mit den Bitleitungs-Leseverstärkern 1111_1 bis 1111_2 sind, sowie Spaltenauswahlschalter 1122_1 bis 1122_2 und 1132_1 bis 1132_2, die im Wesentlichen identisch mit den Spaltenauswahlschaltern 1112_1 bis 1112_2 sind. Die Bitleitungs-Leseverstärker 1111_1 und 1131_1 und die Spaltenauswahlschalter 1112_1 und 1132_1 können neben den Unter-Speicherblöcken 1110_1 bis 1130_1 angeordnet werden. Die Bitleitungs-Leseverstärker 1111_2 und 1131_2 und die Spaltenauswahlschalter 1112_2 und 1132_2 können angrenzend an die Unter-Speicherblöcke 1110_2 bis 1130_2 angeordnet werden.

[0041] Ein Unter-Wortleitungstreiber-Array, das die Wortleitungen WL1 und WL2 und einen Unter-Speicherblock ansteuert, kann ferner zwischen den

Unter-Speicherblöcken 1110_1 bis 1130_1 angeordnet sein. Wie in der vorstehenden Beschreibung können ein Unter-Wortleitungstreiber-Array und ein Unter-Speicherblock ferner zwischen den Unter-Speicherblöcken 1110_2 bis 1130_2 angeordnet sein. Ein Unter-Speicherblock, ein Bitleitungs-Leseverstärker und ein Spaltenauswahlschalter können ferner zwischen den Unter-Speicherblöcken 1110_1 und 1110_2, zwischen den Unter-Speicherblöcken 1120_1 und 1120_2 und zwischen den Unter-Speicherblöcken 1130_1 und 1130_2 angeordnet sein.

[0042] Die neuromorphe Vorrichtung 1000d kann ferner die Neuronenschaltungen 1220 bis 1230 enthalten, wobei die Neuronenschaltung 1210 der Neuronenschaltung 1200 von Fig. 4 entspricht. Die Anzahl der zusätzlichen Neuronenschaltungen 1220 bis 1230 kann der Anzahl der Unter-Speicherblöcke entsprechen, die entlang einer Richtung angeordnet sind, in der sich die Wortleitungen WL1 bis WL4 erstrecken. Die Neuronenschaltungen 1210 bis 1230 können jeweils die Integratoren 1310 bis 1330 und jeweils die Aktivierungsschaltungen 1410 bis 1430 enthalten. Die Neuronenschaltungen 1210 bis 1230 können im Wesentlichen identisch mit der Neuronenschaltung 1200 arbeiten und können jeweils die Aktivierungssignale ACT1 bis ACT3 ausgeben. Die Neuronenschaltung 1210 kann zumindest zwei oder mehr der in den Speicherzellen MC der Unter-Speicherblöcke 1110_1 bis 1110_2 an einer Ein-/Ausgabeleitung IO gespeicherten Gewichte summieren. Wie in der Neuronenschaltung 1210 kann jede der Neuronenschaltungen 1220 und 1230 zwei oder mehr Gewichte an einer Ein-/Ausgabeleitung IO summieren.

[0043] Der Spaltendekodierer 1700 kann zusätzliche Unter-Spaltendekodierer 1720 bis 1730 enthalten, wobei der Unter-Spaltendekodierer 1710 dem Unter-Spaltendekodierer 1700 aus Fig. 4 entspricht. Jeder der Unter-Spaltendekodierer 1710 bis 1730 kann die interne Spaltenadresse ICA dekodieren und kann das Spaltenauswahlsignal über die Spaltenauswahlleitung CSL übertragen. Beispielsweise kann der Unter-Spaltendekodierer 1710 die Speicherzellen MC der Unter-Speicherblöcke 1110_1 bis 1110_2 auswählen, der Unter-Spaltendekodierer 1720 kann die Speicherzellen MC der Unter-Speicherblöcke 1120_1 bis 1120_2 auswählen und der Unter-Spaltendekodierer 1730 kann die Speicherzellen MC der Unter-Speicherblöcke 1130_1 bis 1130_2 auswählen.

[0044] Im Falle von „①“ aus Fig. 2 und/oder Fig. 3 kann der Zeilendekodierer 1600 gleichzeitig die Wortleitungen WL1 und WL3 und die Speicherzellen MC der mit den Wortleitungen WL1 und WL3 verbundenen Unter-Speicherblöcke 1110_1 bis 1130_2 auswählen. Im Falle von „②“ aus Fig. 2 oder Fig. 3 kann der Zeilendekodierer 1600 gleichzeitig die Wortlei-

tungen WL2 und WL4 und die Speicherzellen MC der Unter-Speicherblöcke 1110_1 bis 1130_2, die mit den Wortleitungen WL2 und WL4 verbunden sind, auswählen. Die Unter-Spaltendekodierer 1710 bis 1730 können die Spaltenauswahlsignale über die Spaltenauswahlleitungen CSL an die Spaltenauswahlschalter 1112_1 bis 1132_2 anlegen. Die Neuronenschaltung 1210 kann gleichzeitig Lesesignale von Speicherzellen, die durch den Zeilendekodierer 1600 und den Spaltendekodierer 1700 ausgewählt wurden, an der Ein-/Ausgabeleitung IO summieren, die Neuronenschaltung 1220 kann gleichzeitig Lesesignale von Speicherzellen, die durch den Zeilendekodierer 1600 und den Spaltendekodierer 1700 ausgewählt wurden, an der Ein-/Ausgabeleitung IO summieren, die Neuronenschaltung 1230 kann gleichzeitig Lesesignale von Speicherzellen, die vom Zeilendekodierer 1600 und vom Spaltendekodierer 1700 ausgewählt wurden, an der Ein-/Ausgabeleitung IO summieren, und die Neuronenschaltungen 1210 bis 1230 können die Aktivierungssignale ACT1 bis ACT3 ausgeben. Die Neuronenschaltung 1210 kann ein Summensignal des Integrators 1310 mit dem Schwellwertsignal TH vergleichen, um das Aktivierungssignal ACT1 auszugeben, die Neuronenschaltung 1220 kann ein Summensignal des Integrators 1320 mit dem Schwellwertsignal TH vergleichen, um das Aktivierungssignal ACT2 auszugeben, und die Neuronenschaltung 1230 kann ein Summensignal des Integrators 1330 mit dem Schwellwertsignal TH vergleichen, um das Aktivierungssignal ACT3 auszugeben. Daher können alle, ein Teil oder keines der Aktivierungssignale ACT1 bis ACT3 aktiviert werden, oder alle, ein Teil oder keines der Aktivierungssignale ACT1 bis ACT3 können deaktiviert werden.

[0045] In einer Ausführungsform kann die Neuronenschaltung 1210 den Zeilendekodierer 1600, den Unter-Spaltendekodierer 1710 und/oder den Integrator 1310 durch Verwendung des aktivierten Aktivierungssignals ACT1 zurücksetzen. Da der Unter-Spaltendekodierer 1710 die Speicherzellen MC auswählt, auf die über die mit der Neuronenschaltung 1210 verbundene Ein-/Ausgabeleitung IO zugegriffen wird, und der Integrator 1300 in der Neuronenschaltung 1210 enthalten ist, kann die vorstehend beschriebene Rücksetzoperation eine Selbstrücksetzoperation sein. Die Neuronenschaltung 1210 kann den Unter-Spaltendekodierer 1720 oder den Unter-Spaltendekodierer 1730, der an die Neuronenschaltung 1210 angrenzt, zurücksetzen. Jeder von den Unter-Spaltendekodierern 1710 und 1720 kann das Spaltenauswahlsignal als Reaktion auf das aktivierte Aktivierungssignal ACT1 deaktivieren. Der Unter-Spaltendekodierer 1720 kann das Spaltenauswahlsignal als Reaktion auf das aktivierte Aktivierungssignal ACT2 deaktivieren. Der Integrator 1300 kann als Reaktion auf das aktivierte Aktivierungssignal ACT1 einen Spannungspegel der Ein-

/Ausgabeleitung IO auf einen Vorladepegel oder einen Rücksetzpegel ansteuern. Wie in der Neuronenschaltung 1210 können die Neuronenschaltungen 1220 und 1230 jeweils Rücksetzvorgänge durchführen. Beispielsweise kann der Zeilendekodierer 1600 Zeilenauswahlsignale deaktivieren, die als Reaktion auf das aktivierte Aktivierungssignal ACT1/ACT2/ACT3 über ausgewählte Wortleitungen übertragen werden.

[0046] Fig. 6 veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung nach einer anderen Ausführungsform der erfindatorischen Konzepte. Auf eine zusätzliche Beschreibung in Verbindung mit Komponenten, die die gleichen Referenznummern aufweisen, wird verzichtet, um Redundanz zu vermeiden, und es werden hauptsächlich Unterschiede zwischen den neuromorphen Vorrichtungen 1000d und 1000e beschrieben.

[0047] Verglichen mit der neuromorphen Vorrichtung 1000d kann die neuromorphe Vorrichtung 1000e zusätzlich die Schreibtreiber 1810 bis 1830 enthalten. Nach der Auswahl von Speicherzellen durch den Zeilendekodierer 1600 und den Spaltendekodierer 1700 (siehe Fig. 5) kann der Schreibtreiber 1810 über die Ein-/Ausgabeleitung IO Gewichte in Speicherzellen schreiben und/oder aktualisieren, die aus Speicherzellen der Unter-Speicherblöcke 1110_1 bis 1110_2 ausgewählt wurden. Wie im Schreibtreiber 1810 können die Schreibtreiber 1820 und 1830 Schreiboperationen durchführen.

[0048] Der Zeilendekodierer 1600 kann eine externe Zeilenadresse ERA dekodieren und kann eine der Wortleitungen WL1 bis WL4 auswählen. Der Vor-Zeilendekodierer 1620 kann die externe Zeilenadresse ERA dekodieren und kann einen der Unter-Zeilendekodierer 1610_1 bis 1610_2 auswählen. Im Falle der Dekodierung der internen Zeilenadresse IRA kann der Vor-Zeilendekodierer 1620 ein oberes Bit eines Zeilenadressenbits, das zur Auswahl von Wortleitungen jedes der Unter-Speicherblöcke 1110_1 und 1110_2 verwendet wird, auf ein Don't-Care-Bit setzen. Im Falle der Dekodierung der externen Zeilenadresse ERA kann der Vor-Zeilendekodierer 1620 das obere Bit gültig dekodieren. Der Zeilendekodierer 1600 kann gleichzeitig zwei oder mehr Wortleitungen auswählen, wenn die neuromorphe Vorrichtung 1000e eine neuronale Netzwerk-Operation (z. B. eine „neuromorphe Rechenoperation“ oder „neuromorphe Verarbeitung“) unter Verwendung der im Speicherzellenarray 1100d gespeicherten Gewichte und Biases durchführt. Der Zeilendekodierer 1600 kann in einem Fall, in dem die neuromorphe Vorrichtung 1000e die im Speicherzellenarray 1100 gespeicherten Gewichte und Biases aktualisiert, eine Wortleitung auswählen. Jeder der Unter-Spaltendekodierer 1710 bis 1730 kann eine externe Spaltenadresse ECA dekodieren, um das Spalte-

nauswahlsignal über die Spaltenauswahlleitung CSL zu übertragen, oder er kann die interne Spaltenadresse ICA dekodieren, um das Spaltenauswahlsignal über die Spaltenauswahlleitung CSL zu übertragen.

[0049] Beispielsweise können die externe Zeilenadresse ERA und die externe Spaltenadresse ECA an die neuromorphe Vorrichtung 1000e übertragen und/oder dieser zugeführt werden, um die im Speicherzellenarray 1100 gespeicherten Gewichte und Biases zu aktualisieren. Da die neuromorphe Vorrichtung 1000e die neuronale Netzwerk-Operation durchführt, können die interne Zeilenadresse IRA und die interne Spaltenadresse ICA intern vom Adressgenerator 1510 auf der Grundlage der Aktivierungssignale ACT1 bis ACT3, die von den Neuronschaltungen 1210 bis 1230 ausgegeben werden, erzeugt werden.

[0050] Fig. 7 veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung aus Fig. 5. Zur Vermeidung von Redundanz wird auf eine zusätzliche Beschreibung von Komponenten mit den gleichen Referenznummern verzichtet und die neuromorphe Vorrichtung 1000d wird unter Bezugnahme auf Fig. 5 und 7 beschrieben.

[0051] Das Beispiel in Fig. 5 veranschaulicht den Unter-Zeilendekodierer 1610_1, wie er mit den Wortleitungen WL1 und WL2 verbunden ist, aber ein Beispiel ist in Fig. 7 dargestellt, wobei der Unter-Zeilendekodierer 1610_1 mit den Wortleitungen WL1[X:0] verbunden ist (wobei X eine ganze Zahl von 1 oder mehr ist), die interne Zeilenadresse IRA dekodiert, um eine der Wortleitungen WL1 [X:0] auszuwählen (d. h. WL1 und WL2, die zu WL1[X:0] gehören), und gibt über die ausgewählte Wortleitung ein Zeilenauswahlsignal aus. Die übrigen Unter-Zeilendekodierer 1610_2 und 1610_3 können so implementiert werden, dass sie im Wesentlichen identisch mit dem Unter-Zeilendekodierer 1610_1 sind. Das Beispiel in Fig. 5 veranschaulicht den Unter-Spaltendekodierer 1710, wie er mit der einen Spaltenauswahlleitung CSL verbunden ist, aber ein Beispiel ist in Fig. 7 dargestellt, wobei der Unter-Spaltendekodierer 1710 mit den Spaltenauswahlleitungen CSL[Y:0] (wobei Y eine ganze Zahl von 1 oder mehr ist) (CSL gehört zu CSL[Y:0]) verbunden ist, die interne Spaltenadresse ICA dekodiert, um eine der Spaltenauswahlleitungen CSL[Y:0] auszuwählen, und ein Spaltenauswahlsignal über die ausgewählte Spaltenauswahlleitung ausgibt. Die übrigen Unter-Spaltendekodierer 1720 bis 1730 (von Fig. 5) können so implementiert werden, dass sie im Wesentlichen identisch mit dem Unter-Spaltendekodierer 1710 sind. Der Unter-Speicherblock 1110_1 kann die Speicherzellen MC enthalten, die den Wortleitungen WL1[X:0] und Bitleitungen BL[K:0] (wobei K eine ganze Zahl von 1 oder mehr ist) entsprechen, die gleichzeitig durch Spaltenauswahlsignale ausge-

wählt werden, die über die Spaltenauswahlleitungen CSL[Y:0] übertragen werden. Außerdem können die Bitleitungen-Leseverstärker 1111_1, die jeweils mit den Bitleitungen BL[K:0] verbunden sind, neben dem Unter-Speicherblock 1110_1 angeordnet werden. Die Unter-Speicherblöcke 1110_2 bis 1110_3 können so implementiert werden, dass sie im Wesentlichen identisch mit dem Unter-Speicherblock 1110_1 sind, und die Bitleitungen-Leseverstärker 1111_2 bis 1111_3 können jeweils angrenzend an die Unter-Speicherblöcke 1110_2 bis 1110_3 angeordnet werden.

[0052] Das Speicherzellenarray 1100 kann die Spaltenauswahlschalter 1112_1 bis 1112_3 enthalten. Die Spaltenauswahlschalter 1112_1 können die Bitleitungen BL[K:0] des Unter-Speicherblocks 1110_1 elektrisch mit den lokalen Ein-/Ausgabeleitungen LIO1[K:0] als Reaktion auf das Spaltenauswahlsignal verbinden, das durch eine Spaltenauswahlleitung übertragen wird, die jeweils aus den Spaltenauswahlleitungen CSL[Y:0] ausgewählt wird. Die Spaltenauswahlschalter 1112_2 können die Bitleitungen BL[K:0] des Unter-Speicherblocks 1110_2 als Reaktion auf das Spaltenauswahlsignal, das durch eine Spaltenauswahlleitung übertragen wird, die jeweils aus den Spaltenauswahlleitungen CSL[Y:0] ausgewählt wird, elektrisch mit lokalen Ein-/Ausgabeleitungen LIO2[K:0] verbinden. Die Spaltenauswahlschalter 1112_3 können die Bitleitungen BL[K:0] des Unter-Speicherblocks 1110_3 als Reaktion auf das Spaltenauswahlsignal, das durch eine aus den Spaltenauswahlleitungen CSL[Y:0] ausgewählte Spaltenauswahlleitung übertragen wird, elektrisch mit den lokalen Ein-/Ausgabeleitungen LIO3[K:0] verbinden.

[0053] Beispielsweise können die Unter-Zeilendekodierer 1610_1, 1610_2 und 1610_3 jeweils die Wortleitungen WL1[0], WL2[0] und WL3[0] auswählen und der Unter-Spaltendekodierer 1710 kann die Spaltenauswahlleitung CSL[0] auswählen (oder natürlich kann nur einen Teil der Unter-Zeilendekodierer 1610_1, 1610_2 und 1610_3 Wortleitungen auswählen und eine der übrigen Spaltenauswahlleitung CSL[Y:1] kann ausgewählt werden). Die Bitleitungen-Leseverstärker 1111_1, 1111_2 und 1111_3 können Spannungspegel der Bitleitungen BL[K:0] abtasten, die durch Gewichte verändert werden, die in Speicherzellen gespeichert sind, die mit den Wortleitungen WL1[0], WL2[0] und WL3[0] verbunden sind. Die Spaltenauswahlschalter 1112_1, 1112_2 und 1112_3, die mit der Spaltenauswahlleitung CSL [0] verbunden sind, können die Bitleitungen BL[K:0] elektrisch mit den lokalen Ein-/Ausgabeleitungen LIO1[K:0], LIO2[K:0] bzw. LIO3[K:0] verbinden. Das heißt, in den ausgewählten Speicherzellen gespeicherte Gewichte oder Lesesignale aus den ausgewählten Speicherzellen, die die Gewichte enthalten, können auf die lokalen Ein-/Ausgabeleitungen LIO1 [K:0], LIO2[K:0] und LIO3[K:0] übertragen werden.

Zusätzlich können die Bitleitungs-Leseverstärker 1111_1, 1111_2 und 1111_3 die abgetasteten Spannungspegel der Bitleitung BL[K:0] verstärken und die in den Speicherzellen, die mit den Wortleitungen WL1[0], WL2[0] und WL3[0] verbunden sind, gespeicherten Gewichte wiederherstellen.

[0054] Der Integrator 1300 kann lokale Leseverstärker 1311_1 enthalten, die jeweils mit den lokalen Ein-/Ausgabeleitungen LIO1[K:0] verbunden sind, lokale Leseverstärker 1311_2, die jeweils mit den lokalen Ein-/Ausgabeleitungen LIO2[K:0] verbunden sind, und lokale Leseverstärker 1311_3, die jeweils mit den lokalen Ein-/Ausgabeleitungen LIO3[K:0] verbunden sind. Die lokalen Leseverstärker 1311_1 können globale Ein-/Ausgabeleitungen GIO[K:0] auf der Grundlage von Lesesignalen ansteuern, die an die lokalen Ein-/Ausgabeleitungen LIO1[K:0] übertragen werden, d. h. auf der Grundlage der Spannungspegel der lokalen Ein-/Ausgabeleitungen LIO1[K:0]. Die lokalen Leseverstärker 1311_2 können die globalen Ein-/Ausgabeleitungen GIO[K:0] auf der Grundlage der Spannungspegel der lokalen Ein-/Ausgabeleitungen LIO2[K:0] ansteuern. Die lokalen Leseverstärker 1311_3 können die globalen Ein-/Ausgabeleitungen GIO[K:0] auf der Grundlage der Spannungspegel der lokalen Ein-/Ausgabeleitungen LIO3[K:0] ansteuern. Wie vorstehend beschrieben, können bei der Durchführung einer neuronalen Netzwerk-Operation zumindest zwei oder mehr der Unter-Zeilendekodierer 1610_1 bis 1610_3 gleichzeitig Wortleitungen auswählen. Dementsprechend können zumindest zwei oder mehr der lokalen Leseverstärker 1311_1 bis 1311_3 gleichzeitig die globalen Ein-/Ausgabeleitungen GIO[K:0] ansteuern. Beispielsweise können die lokalen Leseverstärker 1311_1, 1311_2 und 1311_3, die jeweils mit den lokalen Ein-/Ausgabeleitungen LIO1[0], LIO2[0] und LIO3[0] verbunden sind, die globalen Ein-/Ausgabeleitungen GIO[0] gleichzeitig und jeweils ansteuern. Wie in der vorstehenden Beschreibung können die lokalen Leseverstärker 1311_1, 1311_2 und 1311_3, die jeweils mit den lokalen Ein-/Ausgabeleitungen LIO1[0], LIO2[0] und LIO3[0] verbunden sind, die globalen Ein-/Ausgabeleitungen GIO[1] gleichzeitig und jeweils ansteuern. Lokale Leseverstärker und lokale Ein-/Ausgabeleitungen können für jeden der Unter-Speicherblöcke, die in Wortleitungsrichtung angeordnet sind, vorgesehen werden. Die unter Bezugnahme auf Fig. 2 bis 6 beschriebenen Ein-/Ausgabeleitungen IO können die lokalen Ein-/Ausgabeleitungen LIO1[K:0] bis LIO3[K:0] und die globalen Ein-/Ausgabeleitungen GIO[K:0] enthalten.

[0055] Der Integrator 1300 kann globale Leseverstärker 1312 enthalten, die jeweils mit den globalen Ein-/Ausgabeleitungen GIO[K:0] und einer Summenschaltung 1313 verbunden sind. Die globalen Leseverstärker 1312 können die Spannungspegel der globalen Ein-/Ausgabeleitungen GIO[K:0] erfassen

und verstärken. Die globalen Leseverstärker 1312 können die Gewichte oder Biases ausgewählter Speicherzellen, die jeweils an den globalen Ein-/Ausgabeleitungen GIO[K:0] summiert werden, erfassen und verstärken. Die globalen Leseverstärker 1312 können der Summenschaltung 1313 die erfassten Spannungspegel der globalen Ein-/Ausgabeleitungen GIO[K:0] zuführen. Der Integrator 1300 kann Lesesignale, die die Gewichte oder Biases der Speicherzellen MC enthalten, an den globalen Ein-/Ausgabeleitungen GIO[K:0] unter Verwendung der lokalen Leseverstärker 1311_1 bis 1311_3 und der globalen Leseverstärker 1312 summieren. Die Summenschaltung 1313 kann die Spannungspegel der globalen Ein-/Ausgabeleitungen GIO[K:0], die von den globalen Leseverstärkern 1312 zugeführt werden, summieren, ein Summensignal SOUT erzeugen und/oder das Summensignal SOUT der Aktivierungsschaltung 1400 zuführen. Nur der Unter-Spaltendekodierer 1710 und die Neuronenschaltung 1200 werden unter Bezugnahme auf Fig. 7 beschrieben, aber die übrigen Unter-Spaltendekodierer 1720 bis 1730 und die übrigen Neuronenschaltungen 1220 bis 1230 von Fig. 5 können so implementiert werden, dass sie identisch mit dem Unter-Spaltendekodierer 1710 und der Neuronenschaltung 1200 sind.

[0056] Fig. 8 zeigt im Detail ein Blockdiagramm einer neuromorphen Vorrichtung aus Fig. 6. Zur Vermeidung von Redundanz wird auf eine zusätzliche Beschreibung in Verbindung mit Komponenten, die die gleichen Referenznummern aufweisen, verzichtet, eine neuromorphe Vorrichtung 1000e wird unter Bezugnahme auf Fig. 6 und 8 beschrieben, und es wird hauptsächlich ein Unterschied zwischen der neuromorphen Vorrichtung 1000e und der neuromorphen Vorrichtung 1000d beschrieben.

[0057] Das Speicherzellenarray 1100 kann Eingabe-/Ausgabeschalter 1910_1 enthalten, die die lokalen Ein-/Ausgabeleitungen LIO1[K:0] elektrisch mit den globalen Ein-/Ausgabeleitungen GIO[K:0] verbinden, bzw. Eingabe-/Ausgabeschalter 1910_2, die die lokalen Ein-/Ausgabeleitungen LIO2[K:0] mit den globalen Ein-/Ausgabeleitungen GIO[K:0] bzw. Eingabe-/Ausgabeschaltern 1910_3 elektrisch verbinden, die die lokalen Ein-/Ausgabeleitungen LIO3[K:0] mit den globalen Ein-/Ausgabeleitungen GIO[K:0] elektrisch verbinden. Um beispielsweise im Speicherzellenarray 1100 gespeicherte Gewichte und/oder Biases zu aktualisieren, kann der Unter-Zeilendekodierer 1610_1 die Wortleitung WL1[0] auswählen und der Unter-Spaltendekodierer 1710 kann die Spaltenauswahlleitung CSL[0] auswählen. Die Ein-/Ausgabeschalter 1910_1 können die lokalen Ein-/Ausgabeleitungen LIO1[K:0] elektrisch mit den jeweiligen globalen Ein-/Ausgabeleitungen GIO[K:0] verbinden. Danach kann der Schreiber 1810 über die globalen Ein-/Ausgabeleitungen GIO[K:0], die Ein-/Ausgabeschalter 1910_1, die lokalen

Ein-/Ausgabeleitungen LIO1[K:0], die Spaltenauswahlschalter 1112_1 und die Bitleitungen BL[K:0] neue Gewichte und/oder neue Biases in die vom Unter-Zeilendekodierer 1610_1 und dem Unter-Spaltendekodierer 1710 ausgewählten Speicherzellen schreiben. Die Ein-/Ausgabeschalter 1910_1 bis 1910_3 können die neuen Gewichte oder die neuen Biases an die Speicherzellen MC übertragen. Wie in der vorstehenden Beschreibung kann eine Schreib- (oder Aktualisierungs-) Operation auf den verbleibenden Speicherzellen durchgeführt werden.

[0058] Fig. 9 zeigt im Detail ein Blockdiagramm einer neuromorphen Vorrichtung aus Fig. 7. Zusätzliche Beschreibungen von Komponenten, die die gleichen Referenznummern aufweisen, werden ausgelassen, um Redundanz zu vermeiden. Komplementäre Bitleitungen BLB[K:0] der Bitleitungen BL[K:0], komplementäre lokale Ein-/Ausgabeleitungen LIOB1[K:0], LIOB2[K:0] und LIOB3[K:0] der lokalen Ein-/Ausgabeleitungen LIO1[K:0], LIO2[K:0] und LIO3[K:0], und die komplementären globalen Ein-/Ausgabeleitungen GIOB[K:0] der globalen Ein-/Ausgabeleitungen GIO[K:0] können weiter im Speicherzellenarray 1100 und im Integrator 1300 der neuromorphen Vorrichtung 1000d/1000e angeordnet werden. Eine Leitung und eine komplementäre Leitung können ein Paar sein, die eine Leitung kann eine der Leitungen BL[K:0], LIO1[K:0], LIO2[K:0], LIO3[K:0] und GIO[K:0] sein, und die komplementäre Leitung kann eine der Leitungen BLB[K:0], LIOB1[K:0], LIOB2[K:0], LIOB3[K:0] und GIOB[K:0] sein. Die komplementäre Leitung kann einen Spannungspegel aufweisen, der mit einem Spannungspegel der einen Leitung in einem Vorlademodus identisch ist, und kann einen Spannungspegel aufweisen, der komplementär zu einem Spannungspegel der einen Leitung in den übrigen Modi (z. B. einem Aktivierungsmodus, einem Lesemodus und einem Schreibmodus) außer dem Vorlademodus ist. Beispielsweise kann die komplementäre Leitung in einer Richtung geändert oder angetrieben werden, die einer Richtung entgegengesetzt ist, in der ein Spannungspegel der einen Leitung geändert wird.

[0059] Unter Bezugnahme auf Fig. 9 können die Speicherzellen MC mit den Wortleitungen WL1[0] und WL2[0] verbunden werden. Beispielsweise kann jede der Speicherzellen MC ein einzelnes Bit speichern. Zum Beispiel kann der Zellenspannungspegel Vcell der Speicherzelle MC, die mit der Wortleitung WL1[0] und der Bitleitung BL[0] verbunden ist, ein Spannungspegel VC1 sein, der einem ersten logischen Wert entspricht, der Zellenspannungspegel Vcell der Speicherzelle MC, die mit der Wortleitung WL1[0] und der Bitleitung BL[1] verbunden ist, ein Spannungspegel VC0 sein, der einem zweiten logischen Wert entspricht, der Zellenspannungspegel Vcell der Speicherzelle MC, die mit der Wortleitung WL2[0] und der Bitleitung BL[0] verbunden ist,

kann der Spannungspegel VC1 sein, der dem ersten logischen Wert entspricht, und der Zellenspannungspegel Vcell der Speicherzelle MC, die mit der Wortleitung WL2[0] und der Bitleitung BL[1] verbunden ist, kann der Spannungspegel VC1 sein, der dem ersten logischen Wert entspricht. Die vorstehend beschriebenen Spannungspegel sind jedoch nur eine beispielhafte Ausführungsform. Wenn die Wortleitungen WL1[0] und WL2[0] von den Unter-Zeilendekodierern 1610_1 bzw. 1610_2 ausgewählt werden, kann die Ladungsteilung zwischen der Speicherzelle MC und der Bitleitung BL[0] sowie zwischen der Speicherzelle MC und der Bitleitung BL[1] erfolgen.

[0060] Die Bitleitungs-Leseverstärker 1111_1 und 1111_2 können jeweils eine Spannungspegeldifferenz der Bitleitung BL[0] und der komplementären Bitleitung BLB[0] sowie eine Spannungspegeldifferenz der Bitleitung BL[1] und der komplementären Bitleitung BLB[1] erfassen und verstärken. Auf der Grundlage der erfassten Ergebnisse können die Bitleitungs-Leseverstärker 1111_1 die Bitleitung BL[0] und die komplementäre Bitleitung BLB[0] mit den Spannungspegeln VBL1 und VBLO und die Bitleitung BL[1] und die komplementäre Bitleitung BLB[1] mit den Spannungspegeln VBLO und VBL1 ansteuern. Basierend auf den erfassten Ergebnissen können die Bitleitungs-Leseverstärker 1111_2 die Bitleitung BL[0] und die komplementäre Bitleitung BLB[0] mit den Spannungspegeln VBL1 und VBLO ansteuern und die Bitleitung BL[1] und die komplementäre Bitleitung BLB[1] mit den Spannungspegeln VBL1 und VBLO ansteuern. Die vorstehend beschriebenen Spannungspegel VBL1 und VBLO sind jedoch nur Beispiele und können je nach dem Zellspannungspegel Vcell geändert werden.

[0061] Beispielsweise können die Bitleitung BL[0]BL[1] und die komplementäre Bitleitung BLB[0]BLB[1] zusammen im selben Unter-Speicherblock (eine gefaltete Bitleitungsarchitektur) angeordnet werden. In einem anderen Beispiel können die Bitleitung BL[0]BL[1] und die komplementäre Bitleitung BLB[0]BLB[1] jeweils in verschiedenen Unter-Speicherblöcken (eine offene Bitleitungsarchitektur) angeordnet sein.

[0062] Die Spaltenauswahlschalter 1112_1 und 1112_2 können jeweils die Bitleitung BL[0] mit der lokalen Ein-/Ausgabelitung LIO[0], die komplementäre Bitleitung BLB[0] mit der komplementären lokalen Ein-/Ausgabelitung LIOB[0], die Bitleitung BL[1] mit der lokalen Ein-/Ausgabelitung LIO[1] und die komplementäre Bitleitung BLB[1] mit der komplementären lokalen Ein-/Ausgabelitung LIOB[1] elektrisch verbinden. Wenn die Spaltenauswahlschalter 1112_1 durch das über die Spaltenauswahlleitung CSL[0] übertragene Spaltenauswahlsignal eingeschaltet werden, können die Bitleitungs-Leseverstärker 1111_1 die lokale Ein-/Ausgabelitung LIO[0]

und die komplementäre lokale Ein-/Ausgabeleitung LIOB0[0] mit den Spannungspegeln VBL1 und VBLO und die lokale Ein-/Ausgabeleitung LIO0[1] und die komplementäre lokale Ein-/Ausgabeleitung LIOB0[1] mit den Spannungspegeln VBLO und VBL1 ansteuern. Wenn die Spaltenauswahlschalter 1112_2 durch das über die Spaltenauswahlleitung CSL[0] übertragene Spaltenauswahlsignal eingeschaltet werden, können die Bitleitungs-Leseverstärker 1111_2 die lokale Ein-/Ausgabeleitung LIO1[0] und die komplementäre lokale Ein-/Ausgabeleitung LIOB1[0] mit den Spannungspegeln VBL1 und VBLO und die lokale Ein-/Ausgabeleitung LIO1[1] und die komplementäre lokale Ein-/Ausgabeleitung LIOB1[1] mit den Spannungspegeln VBL1 und VBLO ansteuern.

[0063] Der lokale Leseverstärker 1311_1, der mit der lokalen Ein-/Ausgabeleitung LIO0[0] und der komplementären lokalen Ein-/Ausgabeleitung LIOB0[0] verbunden ist, kann die globale Ein-/Ausgabeleitung GIO[0] und die komplementäre globale Ein-/Ausgabeleitung GIOB[0] auf der Grundlage der Spannungspegel VBL1 und VBLO ansteuern. Der lokale Leseverstärker 1311_2, der mit der lokalen Ein-/Ausgabeleitung LIO1[0] und der komplementären lokalen Ein-/Ausgabeleitung LIOB1[0] verbunden ist, kann die globale Ein-/Ausgabeleitung GIO[0] und die komplementäre globale Ein-/Ausgabeleitung GIOB[0] auf der Grundlage der Spannungspegel VBL1 und VBLO ansteuern.

[0064] Der lokale Leseverstärker 1311_1, der mit der lokalen Ein-/Ausgabeleitung LIO0[1] und der komplementären lokalen Ein-/Ausgabeleitung LIOB0[1] verbunden ist, kann die globale Ein-/Ausgabeleitung GIO[1] und die komplementäre globale Ein-/Ausgabeleitung GIOB[1] auf der Grundlage der Spannungspegel VBLO und VBL1 ansteuern. Der lokale Leseverstärker 1311_2, der mit der lokalen Ein-/Ausgabeleitung LIO1[1] und der komplementären lokalen Ein-/Ausgabeleitung LIOB1[1] verbunden ist, kann die globale Ein-/Ausgabeleitung GIO[1] und die komplementäre globale Ein-/Ausgabeleitung GIOB[1] auf der Grundlage der Spannungspegel VBL1 und VBLO ansteuern.

[0065] Beispielsweise kann jeder der Spannungspegel der globalen Ein-/Ausgabeleitungen GIO[0] und GIO[1] und der komplementären globalen Ein-/Ausgabeleitungen GIOB[0] und GIOB[1] einer der Spannungspegel VGIO0 bis VGIOR (wobei R eine ganze Zahl von 1 oder mehr ist) sein, die den logischen Werten einer Summe von Gewichten (oder Biases) entsprechen, die in zwei Speicherzellen MC gespeichert sind. Gewichte und/oder Biases der Speicherzellen MC, die jeweils den Zellenspannungspegel Vcell (= VC1) aufweisen, können an der globalen Ein-/Ausgabeleitung GIO[0] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0] summiert

werden. Gewichte oder Biases der Speicherzelle MC mit dem Zellenspannungspegel Vcell (= VC0) und der Speicherzelle MC mit dem Zellenspannungspegel Vcell (= VC1) können an der globalen Ein-/Ausgabeleitung GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[1] summiert werden. Dementsprechend können die Spannungspegel der globalen Ein-/Ausgabeleitungen GIO[0] und GIO[1] voneinander verschieden sein, und die Spannungspegel der komplementären globalen Ein-/Ausgabeleitungen GIOB[0] und GIOB[1] können voneinander verschieden sein.

[0066] Die Beschreibung wird mit Bezug auf **Fig. 9** gegeben, da die 1-Bit-Gewichte (oder Biases) zweier Speicherzellen MC an der globalen Ein-/Ausgabeleitung GIO[0]/GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0]/GIOB[1] summiert werden. Allerdings kann sich die Anzahl der logischen Werte und die Anzahl der Spannungspegel VGIO0 bis VGIOR, abhängig von der Anzahl der an der globalen Ein-/Ausgabeleitung GIO[0]/GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0]/GIOB[1] summierten Gewichte (oder Biases) oder der Anzahl der Bits der Biases (oder Gewichte), die an der globalen Ein-/Ausgabeleitung GIO[0]/GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0]/GIOB[1] summiert werden, weiter erhöhen.

[0067] Der globale Leseverstärker 1312 kann mit der globalen Ein-/Ausgabeleitung GIO[0] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0] verbunden werden und kann die Spannungspegel der globalen Ein-/Ausgabeleitung GIO[0] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0] erfassen und verstärken. Der globale Leseverstärker 1312 kann mit der globalen Ein-/Ausgabeleitung GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[1] verbunden werden und kann die Spannungspegel der globalen Ein-/Ausgabeleitung GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[1] erfassen und verstärken. Die Schemenschaltung 1313 kann die Spannungspegel der globalen Ein-/Ausgabeleitungen GIO[1:0] und die Spannungspegel der komplementären globalen Ein-/Ausgabeleitungen GIOB[1:0], die von den globalen Leseverstärkern 1312 zugeführt werden, summieren, kann das Summensignal SOUT und ein komplementäres Summensignal SOUTB erzeugen und kann das Summensignal SOUT und das komplementäre Summensignal SOUTB der Aktivierungsschaltung 1400 zuführen.

[0068] **Fig. 10** veranschaulicht ein Timing-Diagramm einer Operation einer neuromorphen Vorrichtung von **Fig. 9**. Zu Beginn der Operation können die Spannungspegel der in **Fig. 9** dargestellten Leitungen WL1[0], WL2[0], BL[1:0], BLB[1:0], LIO0[1:0],

LIOB0[1:0], LIO1[1:0], LIOB1[1:0], GIO[1:0] und GIOB[1:0] Vorladepegel eines Vorlademodus sein.

[0069] Unter der Steuerung des Zeilen/Spalten-Controllers 1500 können die Unter-Zeilendekodierer 1610_1 und 1610_2 im Aktivierungsmodus die Wortleitungen WL1[0] und WL2[0] entsprechend der internen Zeilenadresse IRA auswählen. Wenn die Wortleitungen WL1[0] und WL2[0] ausgewählt sind, kann die Ladungsteilung zwischen den Bitleitungen BL[1:0] und den Speicherzellen MC erfolgen. Ein Beispiel ist in **Fig. 9** dargestellt, da die Bitleitungen BL[1:0] direkt mit den Speicherzellen MC verbunden sind. Da die komplementären Bitleitungen BLB[1:0] jedoch auch mit beliebigen anderen Speicherzellen verbunden sein können, kann die Ladungsteilung zwischen den komplementären Bitleitungen BLB[1:0] und den anderen Speicherzellen erfolgen.

[0070] Wenn die Ladungsteilung zwischen den Bitleitungen BL[1:0] und den Speicherzellen MC abgeschlossen ist, kann die neuronale Netzwerk-Operation der neuromorphen Vorrichtung 1000d eingeleitet werden. Nach der Ladungsteilung können die Bitleitungs-Leseverstärker 1111_1 und 1111_2 jeweils eine Spannungspegeldifferenz der Bitleitung BL[0] und der komplementären Bitleitung BLB[0] sowie eine Spannungspegeldifferenz der Bitleitung BL[1] und der komplementären Bitleitung BLB[1] erfassen und verstärken. Die Bitleitungen BL[1:0] und die komplementären Bitleitungen BLB[1:0] können auf Spannungspegel entwickelt oder verstärkt werden, die den logischen Werten entsprechen, die die Zellspannungspegel Vcell der Speicherzellen MC anzeigen. Wenn die neuronale Netzwerk-Operation der neuromorphen Vorrichtung 1000d eingeleitet wird, kann der Unter-Spaltendekodierer 1710 auch die Spaltenauswahlleitung CSL[0], die der internen Spaltenadresse ICA entspricht, unter der Steuerung des Zeilen/Spalten-Controllers 1500 aktivieren oder auswählen. Wenn der Unter-Spaltendekodierer 1710 die Spaltenauswahlschalter 1112_1 und 1112_2 unter Verwendung des über die Spaltenauswahlleitung CSL[0] übertragenen Spaltenauswahlsignale einschaltet, können die lokalen Ein-/Ausgabeleitungen LIO0[1:0] und LIO1[1:0] und die komplementären lokalen Ein-/Ausgabeleitungen LIOB0[1:0] und LIOB1[1:0] Spannungspegel der Bitleitungen BL[1:0] und der komplementären Bitleitungen BLB[1:0] aufweisen, und die lokalen Leseverstärker 1311_1 und 1311_2 können die globalen Ein-/Ausgabeleitungen GIO[1:0] und die komplementären globalen Ein-/Ausgabeleitungen GIOB[1:0] basierend auf den Spannungspegeln der lokalen Ein-/Ausgabeleitungen LIO0[1:0] und LIO1[1:0] und der komplementären lokalen Ein-/Ausgabeleitungen LIOB0[1:0] und LIOB1[1:0] ansteuern (i.e., erste Summierung). Nach der ersten Summierung kann die Summenschaltung 1313 die Spannungspegel der globalen Ein-/Ausgabeleitungen GIO[K:0] und die Span-

nungspegel der komplementären globalen Ein-/Ausgabeleitungen GIOB[K:0] summieren (d. h. zweite Summierung). Nach der zweiten Summierung können die Unter-Zeilendekodierer 1610_1 und 1610_2 sowie die Unter-Spaltendekodierer 1710_1 unter der Steuerung des Zeilen/Spalten-Controllers 1500 und/oder auf der Grundlage eines Aktivierungssignals der Aktivierungsschaltung 1400 die Spannungspegel der Leitungen WL1[0], WL2[0], BL[1] treiben:0], BLB[1:0], LIO0[1:0], LIOB0[1:0], LIO1[1:0], LIOB1[1:0], GIO[1:0] und GIOB[1:0] auf die Vorladepegel des Vorlademodus ansteuern. Wenn das Timing-Diagramm von **Fig. 10** wiederholt wird, kann die neuronale Netzwerk-Operation der neuromorphen Vorrichtung 1000d ausgeführt werden.

[0071] **Fig. 11** veranschaulicht im Detail ein Blockdiagramm einer neuromorphen Vorrichtung aus **Fig. 7**. Eine Beschreibung wird sich auf die Unterschiede zwischen der neuromorphen Vorrichtung 1000d von **Fig. 11** und der neuromorphen Vorrichtung 1000d von **Fig. 9** konzentrieren. Im Gegensatz zu den Speicherzellen MC aus **Fig. 9** kann jede der Speicherzellen MC aus **Fig. 11** zwei Bits speichern. Zum Beispiel kann der Zellenspannungspegel Vcell der Speicherzelle MC, die mit der Wortleitung WL1[0] und der Bitleitung BL[0] verbunden ist, ein Spannungspegel VC11 sein, der einem ersten logischen Wert entspricht, der Zellenspannungspegel Vcell der Speicherzelle MC, die mit der Wortleitung WL1[0] und der Bitleitung BL[1] verbunden ist, kann ein Spannungspegel VC00 sein, der einem zweiten logischen Wert entspricht, der Zellenspannungspegel Vcell der Speicherzelle MC, die mit der Wortleitung WL2[0] und der Bitleitung BL[0] verbunden ist, kann der Spannungspegel VC01 sein, der einem dritten logischen Wert entspricht, und der Zellenspannungspegel Vcell der Speicherzelle MC, die mit der Wortleitung WL2[0] und der Bitleitung BL[1] verbunden ist, kann der Spannungspegel VC10 sein, der einem vierten logischen Wert entspricht. Die vorstehend beschriebenen Spannungspegel sind jedoch nur Beispiele. Wie vorstehend beschrieben, können die Speicherzellen MC zwei oder mehr Bits speichern, und die Anzahl der Logikwerte und die Anzahl der Zellenspannungspegel können sich in Abhängigkeit von zwei oder mehr Bits weiter erhöhen. Aus Gründen der Veranschaulichung und Beschreibung wird jedoch davon ausgegangen, dass jede der Speicherzellen MC zwei Bits speichert.

[0072] Wenn die Wortleitungen WL1[0] und WL2[0] von den Unter-Zeilendekodierern 1610_1 bzw. 1610_2 jeweils ausgewählt werden, kann die Ladungsteilung zwischen der Speicherzelle MC und den Bitleitungen BL[1:0] erfolgen. Die Bitleitungen BL[1:0], die die Wortleitung WL1[0] schneiden, können durch die Ladungsteilung die Spannungspegel VBL00 und VBL11 aufweisen. Die Bitleitungen BL[1:0], die die Wortleitung WL2[0] schneiden, können

durch die Ladungsteilung die Spannungspegel VBL10 und VBL01 aufweisen. Wenn die Spaltenauswahlschalter 1112_1 und 1112_2 durch das Spaltenauswahlsignal, das über die Spaltenauswahlleitung CSL[0] übertragen wird, eingeschaltet werden, können die lokalen Ein-/Ausgabeleitungen LIO0[1:0] die gleichen Spannungspegel VBL00 und VBL11 wie die Bitleitungen BL[1:0] aufweisen, die die Wortleitung WL1[0] schneiden, und die lokalen Ein-/Ausgabeleitungen LIO1[1:0] können die gleichen Spannungspegel VBL10 und VBL01 wie die Bitleitungen BL[1:0] aufweisen, die die Wortleitung WL2[0] schneiden. Jede der komplementären Bitleitungen BLB[1:0] und die komplementären lokalen Ein-/Ausgabeleitungen LIOB0[1:0] und LIOB1[1:0] können einen Vorladepegel VPBL aufweisen. In Bezug auf die Speicherzellen MC, die die Zellenspannungspegeln Vcell (= VC11, VC00, VC01 und VC10) aufweisen, können die entsprechenden Spannungspegeldifferenzen der Bitleitungen BL[0], BL[1], BL[0] und BL[1] und der komplementären Bitleitungen BLB[0], BLB[1], BLB[0], und BLB[1] und die entsprechenden Spannungspegeldifferenzen zwischen den lokalen Ein-/Ausgabeleitungen LIO0[0], LIO0[1], LIO1[0] und LIO1[1] und den komplementären lokalen Ein-/Ausgabeleitungen LIOB0[0], LIOB0[1], LIOB1[0] und LIOB1[1] logische Werte anzeigen, die jeweils den Zellenspannungspegeln Vcell (VC11, VC00, VC01 und/oder VC10) entsprechen. Die vorstehend beschriebenen Spannungspegeldifferenzen können durch die Ladungsteilung zwischen den Speicherzellen MC und den Bitleitungen BL[0], BL[1], BL[0] und BL[1] oder durch die Bitleitungs-Leseverstärker 1111_1 und 1111_2 gebildet werden. Beispielsweise können die Bitleitungs-Leseverstärker 1111_1 und 1111_2 entsprechende Spannungspegeldifferenzen zwischen den Bitleitungen BL[1:0] und den komplementären Bitleitungen BLB[1:0] aufrechterhalten, während die neuronale Netzwerk-Operation durchgeführt wird.

[0073] Die lokalen Leseverstärker 1311_1 und 1311_2 können die globale Ein-/Ausgabeleitung GIO[0] und die komplementäre globale Ein-/Ausgabeleitung GIOB[0] basierend auf den Spannungspegeln VBL11, VPBL, VBL01 und VPBL der lokalen Ein-/Ausgabeleitungen LIO0[0], der komplementären lokalen Ein-/Ausgabeleitungen LIOB0[0], der lokalen Ein-/Ausgabeleitung LIO1[0] und der komplementären lokalen Ein-/Ausgabeleitung LIOB1[0] ansteuern. Die lokalen Leseverstärker 1311_1 und 1311_2 können die globale Ein-/Ausgabeleitung GIO[1] und die komplementäre globale Ein-/Ausgabeleitung GIOB[1] basierend auf den Spannungspegeln VBL00, VPBL, VBL10 und VPBL der lokalen Ein-/Ausgabeleitungen LIO0[1], der komplementären lokalen Ein-/Ausgabeleitungen LIOB0[1], der lokalen Ein-/Ausgabeleitung LIO1[1] und der komplementären lokalen Ein-/Ausgabeleitung LIOB1[1] ansteuern.

[0074] Beispielsweise kann jeder der Spannungspegel der globalen Ein-/Ausgabeleitungen GIO[0] und GIO[1] und der komplementären globalen Ein-/Ausgabeleitungen GIOB[0] und GIOB[1] einer der Spannungspegel VGIO0 bis VGIOS (wobei S eine ganze Zahl von 1 oder mehr ist) sein, die den logischen Werten einer Summe von Gewichten (oder Biases) entsprechen, die in zwei Speicherzellen MC gespeichert sind. Gewichte und/oder Biases der Speicherzellen MC mit dem Zellenspannungspegel Vcell (= VC11) und dem Zellenspannungspegel Vcell (= VC01) können an der globalen Ein-/Ausgabeleitung GIO[0] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0] summiert werden. Gewichte und/oder Biases der Speicherzellen MC mit dem Zellenspannungspegel Vcell (= VC00) und dem Zellenspannungspegel Vcell (= VC10) können an der globalen Ein-/Ausgabeleitung GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[1] summiert werden. Eine Spannungspegeldifferenz der globalen Ein-/Ausgabeleitung GIO[0] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0] kann eine Summe von Gewichten oder Biases der Speicherzellen MC mit dem Zellenspannungspegel Vcell (= VC11) und dem Zellenspannungspegel Vcell (= VC01) anzeigen. Eine Spannungspegeldifferenz der globalen Ein-/Ausgabeleitung GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[1] kann eine Summe von Gewichten oder Biases der Speicherzellen MC, die den Zellenspannungspegel Vcell (= VC00) aufweisen, und dem Zellenspannungspegel Vcell (= VC10) anzeigen.

[0075] Die Beschreibung wird mit Bezug auf **Fig. 11** gegeben, da 2-Bit-Gewichte (und/oder Biases) zweier Speicherzellen MC an der globalen Ein-/Ausgabeleitung GIO[0]/GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0]/GIOB[1] summiert werden. Allerdings kann sich die Anzahl der logischen Werte und die Anzahl der Spannungspegel VGIO0 bis VGIOS abhängig von der Anzahl der Gewichte (und/oder Biases), die an der globalen Ein-/Ausgabeleitung GIO[0]/GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0]/GIOB[1] summiert werden, oder der Anzahl der Biases (und/oder Gewichte), die an der globalen Ein-/Ausgabeleitung GIO[0]/GIO[1] und der komplementären globalen Ein-/Ausgabeleitung GIOB[0]/GIOB[1] summiert werden, weiter erhöhen. Zum Beispiel kann „S“ größer als „R“ sein.

[0076] **Fig. 12** veranschaulicht ein Timing-Diagramm einer Operation einer neuromorphen Vorrichtung von **Fig. 11**. Es wird hauptsächlich ein Unterschied zwischen dem Timing-Diagramm von **Fig. 12** und dem Timing-Diagramm von **Fig. 10** beschrieben.

[0077] Nachdem die Ladungsteilung zwischen den Bitleitungen BL[1:0] und den Speicherzellen MC eingeleitet wurde, kann die neuronale Netzwerk-Opera-

tion der neuromorphen Vorrichtung 1000d eingeleitet werden. Entsprechende Spannungspegeldifferenzen zwischen den Bitleitungen BL[1:0] und den komplementären Bitleitungen BLB[1:0], die durch die Ladungsteilung gebildet werden, können logischen Werten entsprechen, die die Zellspannungspegel Vcell der Speicherzellen MC anzeigen. Wenn der Unter-Spaltendekodierer 1710 die Spaltenauswahlschalter 1112_1 und 1112_2 unter Verwendung des über die Spaltenauswahlleitung CSL[0] übertragenen Spaltenauswahlsignals einschaltet, können die lokalen Ein-/Ausgabeleitungen LIO0[1:0] und LIO1[1:0] und die komplementären lokalen Ein-/Ausgabeleitungen LIOB0[1:0] und LIOB1[1:0] Spannungspegel der Bitleitungen BL[1:0] und der komplementären Bitleitungen BLB[1:0] aufweisen, und die lokalen Leseverstärker 1311_1 und 1311_2 können die globalen Ein-/Ausgabeleitungen GIO[1:0] und die komplementären globalen Ein-/Ausgabeleitungen GIOB[1:0] basierend auf den Spannungspegeln der lokalen Ein-/Ausgabeleitungen LIO0[1:0] und LIO1[1:0] und der komplementären lokalen Ein-/Ausgabeleitungen LIOB0[1:0] und LIOB1[1:0] ansteuern (z. B. erste Summierung). Nach der ersten Summierung kann die Summenschaltung 1313 die Spannungspegel der globalen Ein-/Ausgabeleitungen GIO[K:0] und die Spannungspegel der komplementären globalen Ein-/Ausgabeleitungen GIOB[K:0] summieren (z. B. zweite Summierung).

[0078] Nach der zweiten Summierung und vor dem Vorladungsmodus können die Bitleitungs-Leseverstärker 1111_1 und 1111_2 entsprechende Spannungspegeldifferenzen zwischen den Bitleitungen BL[1:0] und den komplementären Bitleitungen BLB[1:0] verstärken und/oder entwickeln. Die Zellspannungspegel Vcell, die in den Speicherzellen MC gespeichert wurden, können durch die Entwicklungsoperationen der Bitleitungs-Leseverstärker 1111_1 und 1111_2 wiederhergestellt werden. Nach der Wiederherstellungsoperation können die Unter-Zeilendekodierer 1610_1 und 1610_2 und die Unter-Spaltendekodierer 1710_1 unter der Kontrolle des Zeilen-/Spalten-Controllers 1500 die Spannungspegel der Leitungen WL1[0], WL2[0], BL[1:0], BLB[1:0], LIO0[1:0], LIOB0[1:0], LIO1[1:0], LIOB1[1:0], GIO[1:0] und GIOB[1:0] auf die Vorladepegel des Vorlademodus ansteuern. Wenn das Timing-Diagramm von **Fig. 12** wiederholt wird, kann der Neuronale Netzwerk-Operation der neuromorphen Vorrichtung 1000d ausgeführt werden.

[0079] **Fig. 13** veranschaulicht einen Schaltplan eines Integrators nach einer Ausführungsform der erfinderischen Konzepte. **Fig. 13** wird unter Bezugnahme auf **Fig. 8** beschrieben. Der Integrator 1310 kann die Transistoren M0 bis M5 enthalten. Die Transistoren M0 bis M2 können jeweils und gleichzeitig die globale Ein-/Ausgabeleitung GIO[0] in Abhängigkeit von den Spannungspegeln der lokalen Ein-/Aus-

gabeleitungen LIO0[0] bis LIO2[0] ansteuern. Beispielsweise können die Transistoren M0 bis M2 jeweils an den lokalen Leseverstärkern 1311_1 bis 1311_3 platziert werden. Der Transistor M3 kann die globale Ein-/Ausgabeleitung GIO[0] in Abhängigkeit von einem Rücksetzsignal RST1 zurücksetzen und/oder vorladen. Das Rücksetzsignal RST1 kann das Aktivierungssignal ACTa/ACTb der Aktivierungsschaltung 1400 oder ein Signal sein, das auf der Grundlage des Aktivierungssignals ACTa/ACTb erzeugt wird. Basierend auf einem Auswahlsignal SE1 und einem komplementären Auswahlsignal SEB1 können die an die Versorgungsspannungen VDD und GND angeschlossenen Transistoren M4 und M5 die globale Ein-/Ausgabeleitung GIO[0] auswählen oder nicht. Zum Beispiel kann das Auswahlsignal SE1 und das komplementäre Auswahlsignal SEB1 durch den Unter-Spaltendekodierer 1710 erzeugt werden, der die interne Spaltenadresse ICA oder die externe Spaltenadresse ECA dekodiert. Zum Beispiel können die Transistoren M3 bis M5 am globalen Leseverstärker 1312 platziert werden. Der Integrator 1310 kann ferner Transistoren enthalten, die mit den globalen Ein-/Ausgabeleitungen GIO[K:1], den komplementären globalen Ein-/Ausgabeleitungen GIOB[K:0], den lokalen Ein-/Ausgabeleitungen LIO1[K:1] bis LIO3[K:1] und den komplementären lokalen Ein-/Ausgabeleitungen LIOB1[K:0] bis LIOB3[K:0] verbunden sind, Transistoren, die mit dem Rücksetzsignal RST1 verbunden sind, Transistoren, die mit dem Auswahlsignal SE1 und dem komplementären Auswahlsignal SEB1 verbunden sind, die so verbunden sind, dass sie der in **Fig. 13** dargestellten Verbindung ähnlich sind. In **Fig. 13** sind ein Typ von Transistoren, eine Verbindungsbeziehung der Transistoren und die Versorgungsspannungen nur Beispiele und können entsprechend angepasst werden.

[0080] **Fig. 14** veranschaulicht einen Schaltplan einer Summenschaltung nach einer Ausführungsform der erfinderischen Konzepte. **Fig. 14** wird unter Bezugnahme auf **Fig. 8** beschrieben. Die Summenschaltung 1313 kann die Transistoren M6 bis M11 enthalten. Die Transistoren M6 bis M8 können jeweils und gleichzeitig eine Ausgabeleitung in Abhängigkeit von den Spannungspegeln der globalen Ein-/Ausgabeleitungen GIO[K:0], die von den globalen Leseverstärkern 1312 zugeführt werden, so ansteuern, dass das Summensignal SOUT erzeugt wird. Der Transistor M9 kann die Ausgabeleitung in Abhängigkeit von einem Rücksetzsignal RST2 zurücksetzen oder vorladen. Beispielsweise kann das Rücksetzsignal RST2 im Wesentlichen identisch mit dem Rücksetzsignal RST1 sein. Basierend auf einem Auswahlsignal SE2 und einem komplementären Auswahlsignal SEB2 können die an die Versorgungsspannungen VDD und GND angeschlossenen Transistoren M10 und M11 die Ausgabeleitung auswählen oder nicht. Beispielsweise können das Aus-

wahlsignal SE2 und das komplementäre Auswahl-signal SEB2 im Wesentlichen identisch mit dem Auswahl-signal SE1 und dem komplementären Auswahl-signal SEB1 sein. Die Summenschaltung 1313 kann ferner Transistoren, die jeweils und gleichzeitig abhängig von den Spannungspegeln der komplementären globalen Ein-/Ausgabeleitungen GIOB [K:0] eine komplementäre Ausgabeleitung ansteuern, einen Transistor, der mit dem Rücksetz-signal RST2 verbunden ist, und Transistoren, die mit dem Auswahl-signal SE2 und dem komplementären Auswahl-signal SEB2 verbunden sind, die so verbunden sind, dass sie dem in **Fig. 14** dargestellten Anschluss ähnlich sind, enthalten. In **Fig. 14** sind die Art der Transistoren, die Verbindungsbeziehung der Transistoren und die Versorgungsspannungen nur Beispiele und können entsprechend angepasst werden.

[0081] Fig. 15 veranschaulicht einen Schaltplan einer Aktivierungsschaltung nach einer Ausführungsform des erfinderischen Konzepts. Die Aktivierungsschaltung 1400 kann einen Komparator 1411 enthalten, der die Signale SOUT und SOUTB mit dem Schwellwertsignal TH vergleicht. Wenn z. B. das Summensignal SOUT das Schwellwertsignal TH überschreitet und das komplementäre Summensignal SOUTB das Schwellwertsignal TH nicht überschreitet, kann der Komparator 1411 das Aktivierungssignal ACT aktivieren und dann das Aktivierungssignal ACT deaktivieren. Ein Zeitraum, in dem das Aktivierungssignal ACT aktiviert wird, kann im Voraus festgelegt werden. Alternativ kann der Komparator 1411 das Aktivierungssignal ACT deaktivieren, wenn das Summensignal SOUTB das Schwellwertsignal TH nicht überschreitet und das komplementäre Summensignal SOUTB das Schwellwertsignal TH überschreitet. Das Aktivierungssignal ACT kann z. B. ein Impulssignal sein, das in Abhängigkeit von einem Vergleichsergebnis selektiv ausgegeben wird.

[0082] Fig. 16 zeigt ein Blockdiagramm einer neuromorphen Vorrichtung nach einer Ausführungsform des erfinderischen Konzepts. Eine neuromorphe Vorrichtung 2000 kann die Bänke 2100, einen Bus 2200, einen programmierbaren Router 2300 und eine Peripherieschaltung 2400 enthalten.

[0083] Die Bank 2100 kann ein Speicherzellenarray 2110, ein Neuronenschaltungs-Array 2120, einen Zeilen-/Spalten-Controller 2150, einen Zeilendekodierer 2160, einen Spaltendekodierer 2170 und ein Schreiber-Array 2180 enthalten. Die Komponenten 2110, 2150, 2160 und 2170 können im Wesentlichen identisch mit den Komponenten 1100, 1500, 1600 und 1700 sein, die unter Bezugnahme auf **Fig. 2 bis 15** beschrieben sind. Das Neuronenschaltungs-Array 2120 kann die Neuronen-Schaltungen 1200 und 1210 bis 1230 enthalten, die mit Bezug

auf **Fig. 2 bis 15** beschrieben sind. Das Schreiber-Array 2180 kann die Schreiber 1810 bis 1830 enthalten, die in Bezug auf **Fig. 2 bis 15** beschrieben sind.

[0084] Der Bus 2200 kann einen Verbindungspfad zwischen den Banken 2100, einen Verbindungspfad zwischen programmierbaren Routern 2300, einen Verbindungspfad zwischen der Peripherieschaltung 2400 und den Banken 2100 oder einen Verbindungspfad zwischen der Peripherieschaltung 2400 und den programmierbaren Routern 2300 vorsehen. Beispielsweise in dem Fall, in dem im Speicherzellenarray 2110 gespeicherte Gewichte oder Biases aktualisiert werden, können eine Adresse (einschließlich einer externen Bankadresse, einer externen Zeilenadresse und/oder einer externen Spaltenadresse), die Speicherzellen angibt, in denen frühere Gewichte oder Biases gespeichert sind, neue Gewichte und neue Biases über den Bus 2200 an die Bänke übertragen werden. In dem Fall, in dem im anderen Beispiel die neuromorphe Vorrichtung 2000 eine neuronale Netzwerk-Operation durchführt, können Aktivierungssignale des Neuronenschaltungs-Arrays 2120 einer ersten Bank 2100 an eine zweite Bank 2100 (z. B. eine Zielbank) übertragen werden. In diesem Fall kann der Zeilen-/Spalten-Controller 2150 der zweiten Bank 2100 eine interne Adresse (z. B. eine interne Zeilenadresse und eine interne Spaltenadresse) zur Auswahl von Speicherzellen des Speicherzellenarrays 2110 auf der Grundlage der übertragenen Aktivierungssignale erzeugen.

[0085] Der programmierbare Router 2300 kann die Aktivierungssignale des Neuronenschaltungs-Arrays 2120 der ersten Bank 2100 über den Bus 2200 empfangen, kann auf der Grundlage der empfangenen Aktivierungssignale eine interne Bankadresse erzeugen, die der zweiten Bank 2100 entspricht, und kann die empfangenen Aktivierungssignale an die zweite Bank 2100 übertragen. Dabei kann die interne Bankadresse das Speicherzellenarray 2110 der entsprechenden Bank 2100 angeben. Der programmierbare Router 2300 kann einen Bankadressengenerator 2310 enthalten, der auf der Grundlage der empfangenen Aktivierungssignale eine Vielzahl von internen Bankadressen erzeugt, die einer Vielzahl von verschiedenen Bänken 2100 entsprechen, und einen Vermittler 2320, der die empfangenen Aktivierungssignale jeweils auf der Grundlage der Vielzahl von internen Bankadressen an die Vielzahl von verschiedenen Bänken 2100 überträgt.

[0086] Im Einzelnen kann der Bankadressengenerator 2310 die Aktivierungssignale ACT1 des Neuronenschaltungs-Arrays 2120 einer ersten Bank BK1 empfangen und auf der Grundlage der Aktivierungssignale ACT1 interne Bankadressen erzeugen, die eine zweite Bank BK2 anzeigen. Der Vermittler 2320 kann die Aktivierungssignale ACT1 auf der

Grundlage der internen Bankadressen des Bankadressengenerators 2310 an den Zeilen-/Spalten-Controller 2150 der zweiten Bank BK2 übermitteln. Der Bankadressengenerator 2310 kann die Aktivierungssignale ACT2 des Neuronenschaltungs-Arrays 2120 der zweiten Bank BK2 empfangen und auf der Grundlage der Aktivierungssignale ACT2 interne Bankadressen erzeugen, die die erste Bank BK1 anzeigen. Der Vermittler 2320 kann die Aktivierungssignale ACT2 auf der Grundlage der internen Bankadressen des Bankadressengenerators 2310 dem Zeilen-/Spalten-Controller 2150 der ersten Bank BK1 übermitteln. Wie vorstehend beschrieben, können die Aktivierungssignale ACT1 des Neuronenschaltungs-Arrays 2120 der ersten Bank BK1 dem Zeilen-/Spalten-Controller 2150 der ersten Bank BK1 und die Aktivierungssignale ACT2 des Neuronenschaltungs-Arrays 2120 der zweiten Bank BK2 dem Zeilen-/Spalten-Controller 2150 der ersten Bank BK1 zugeführt werden. Der Zeilen-/Spalten-Controller 2150 der ersten Bank BK1 kann die interne Zeilenadresse IRA und die interne Spaltenadresse ICA auf der Grundlage der Aktivierungssignale ACT1 oder der Aktivierungssignale ACT2 erzeugen. Der Zeilen-/Spalten-Controller 2150 der ersten Bank BK1 kann die interne Zeilenadresse IRA und die interne Spaltenadresse ICA auf der Grundlage der Aktivierungssignale ACT1 oder der Aktivierungssignale ACT2 erzeugen. Ein Beispiel wird veranschaulicht, da die erste Bank BK1 und die zweite Bank BK2 nebeneinander liegen, aber die erste Bank BK1 und die zweite Bank BK2 dürfen nicht nebeneinander liegen.

[0087] Der programmierbare Router 2300 kann die Anzahl der Bänke 2100 der neuromorphen Vorrichtung 2000 auf der Grundlage der empfangenen Aktivierungssignale dynamisch anpassen. In dem Fall, in dem die neuromorphe Vorrichtung 2000 den neuronalen Netzwerkbetrieb durchführt, kann das Neuronenschaltungs-Array 2120 die Schichten IL, HL und OL von **Fig. 1** bilden, da das Neuronenschaltungs-Array 2120 in einer Bank 2100 wiederverwendet wird, und die Neuronenschaltungs-Arrays 2120 können die Schichten IL, HL und OL von **Fig. 1** bilden, da die Neuronenschaltungs-Arrays 2120 zwischen mehreren Bänken 2100 wiederverwendet werden. Der Vermittler 2320 des programmierbaren Routers 2300 kann die Kommunikation zwischen den mit dem Bus 2200 verbundenen Bänken 2100, allen anderen programmierbaren Routern 2300 und der Peripherieschaltung 2400 vermitteln. Beispielsweise können Gewichte, die im Speicherzellenarray 2110 einer ersten Bank 2100 gespeichert sind, die Verbindungsstärken des Neuronenschaltungs-Arrays 2120 anzeigen, die in der ersten Bank 2100 wiederverwendet werden. Als ein anderes Beispiel können Gewichte, die im Speicherzellenarray 2110 der ersten Bank 2100 gespeichert sind, die Verbindungsstärke des Neuronenschaltungs-Arrays 2120 der ersten Bank

2100 und des Neuronenschaltungs-Arrays 2120 der zweiten Bank 2100 anzeigen.

[0088] Die Peripherieschaltung 2400 kann einen Taktpuffer 2410, einen Befehls- und Adresspuffer 2420, einen Befehlsdekodierer 2430, einen Adressdemultiplexer 2440, einen Auffrischungs-Controller 2450, eine Daten-Ein-/Ausgabeschaltung 2460, einen DQ-Puffer 2470 und einen DQS-Puffer 2480 enthalten. Der Taktpuffer 2410 kann ein Taktsignal CK von außen empfangen und das Taktsignal ICK allen anderen Komponenten der Peripherieschaltung 2400 zuführen. Der Befehls- und Adresspuffer 2420 kann ein Befehls- und Adresssignal CA empfangen, das auf dem vom Taktpuffer 2410 empfangenen Taktsignal CK basiert. Das Befehls- und Adress-Signal CA kann die externe Bankadresse EBA, die externe Zeilenadresse ERA und die externe Spaltenadresse ECA enthalten. Das Befehls- und Adress-Signal CA kann außerdem verschiedene Befehle für das Speicherzellenarray 2110 enthalten, wie z. B. einen Aktivierungsbefehl, einen Lesebefehl, einen Schreibbefehl und einen Vorladebefehl. Die neuromorphe Vorrichtung 2000 kann in Synchronisation mit dem Taktsignal CK arbeiten. Der Befehls- und Adresspuffer 2420 kann den Befehl dem Befehlsdekodierer 2430 und die Adresse dem Adressdemultiplexer 2440 zuführen. Der Befehlsdekodierer 2430 kann den Befehl dekodieren und die Komponenten 2100, 2200, 2300, 2300, 2410, 2420 und 2440 bis 2480 der neuromorphen Vorrichtung 2000 auf der Grundlage eines Decodierungsergebnisses steuern. Der Adressdemultiplexer 2440 kann eine Adresse empfangen und die empfangene Adresse den Bänken 2100 über den Bus 2200 oder den programmierbaren Router 2300 als externe Zeilenadresse ERA oder externe Spaltenadresse ECA zuführen. Der Adressdemultiplexer 2440 kann der Bank 2100 die externe Zeilenadresse ERA oder die externe Spaltenadresse ECA als externe Zeilenadresse ERA oder als externe Spaltenadresse ECA zuführen, die die externe Bankadresse EBA angibt.

[0089] Der Auffrischungs-Controller 2450 kann einen Auffrischungs-Vorgang zur Wiederherstellung von Gewichten oder Biases durchführen, die im Speicherzellenarray 2110 gespeichert sind. Beispielsweise kann der Auffrischungs-Controller 2450 eine Auffrischungszeilenadresse von Wortleitungen, die mit Speicherzellen verbunden sind, in denen Gewichte oder Biases gespeichert sind, die für das Auffrischen vorgesehen sind, sowie eine Auffrischungs-Bankadresse der Bank 2100, in der die Speicherzellen platziert sind, erzeugen. Der Auffrischungs-Controller 2450 kann die Bänke 2100 in Abhängigkeit von der Auffrischungs-Bankadresse auswählen und die Auffrischungszeilenadresse über den Bus 2200 oder den programmierbaren Router 2300 an die Zeilendekodierer 2160 der Bänke 2100 übertragen. Der Zeilendekodierer 2160 kann

Speicherzellen auswählen, die der Auffrischungszeilenadresse entsprechen, und Bitleitungs-Leseverstärker, die mit den ausgewählten Speicherzellen verbunden sind, können in den ausgewählten Speicherzellen gespeicherte Gewichte oder Biases auffrischen oder wiederherstellen.

[0090] Unter der Steuerung des Befehlsdekodierers 2430 kann der Daten-Ein-/Ausgabeschaltkreis 2460 Schreibdaten (z. B. Gewichte oder Biases) verarbeiten, die in die Bänke 2100 geschrieben werden sollen, und kann die Schreibdaten über den Bus 2200 oder den programmierbaren Router 2300 an das Schreibtreiber-Array 2180 der Bänke 2100 übertragen (z. B. diesem zuführen). Die Daten-Eingabe-/Ausgabeschaltung 2460 kann den DQ-Puffer 2470 und den DQS-Puffer 2480 zum Empfang der Schreibdaten von einer externen Vorrichtung steuern. Unter Steuerung des Befehlsdekodierers 2430 kann die Daten-Ein-/Ausgabeschaltung 2460 Lesedaten einschließlich Aktivierungssignalen der Neuronenschaltungs-Arrays 2120 der Bänke 2100 über den Bus 2200 oder den programmierbaren Router 2300 empfangen und die Lesedaten verarbeiten. Die Daten-Eingabe/Ausgabeschaltung 2460 kann den DQ-Puffer 2470 und den DQS-Puffer 2480 zur Ausgabe der gelesenen Daten an die externe Vorrichtung steuern. Der DQ-Puffer 2470 und der DQS-Puffer 2480 können jeweils ein Schreib-DQ-Signal einschließlich der Schreibdaten und ein Schreib-DQS-Signal zur Erfassung (und/oder Abtastung) des Schreib-DQ-Signals empfangen bzw. ein Lese-DQ-Signal einschließlich der Lesedaten und ein Lese-DQS-Signal nach außen übertragen. Dabei kann das DQ-Signal ein bidirektionales Daten-Eingabe-/Ausgabesignal sein, das das Schreib-DQ-Signal oder das Lese-DQ-Signal ist, und das DQS-Signal kann ein bidirektionales Datenabtastsignal sein, das das Schreib-DQS-Signal oder das Lese-DQS-Signal ist.

[0091] Ferner können der Taktpuffer 2410, der Befehls- und Adresspuffer 2420, der Befehlsdekodierer 2430, der Adressdemultiplexer 2440, der Auffrischungs-Controller 2450, die Daten-Ein-/Ausgabeschaltung 2460, der DQ-Puffer 2470 und der DQS-Puffer 2480 Verarbeitungsschaltungen, wie z. B. Hardware, die Logikschaltungen enthält, eine Hardware/Software-Kombination, wie z. B. einen Prozessor, der Software ausführt, oder eine Kombination davon umfassen. Die Verarbeitungsschaltung kann zum Beispiel insbesondere eine zentrale Verarbeitungseinheit (CPU), eine arithmetisch-logische Einheit (ALU), einen digitalen Signalprozessor, einen Mikrocomputer, ein Field-Programmable Gate-Array (FPGA) und eine programmierbare Logikeinheit, einen Mikroprozessor, eine anwendungsspezifische integrierte Schaltung (ASIC) usw. umfassen, ist aber nicht darauf beschränkt.

[0092] Fig. 17 veranschaulicht ein künstliches neuronales Netzwerk, das von einer neuromorphen Vorrichtung aus Fig. 16 ausgeführt werden kann. Ein künstliches neuronales Netzwerk 2000 kann im Wesentlichen identisch mit dem künstlichen neuronalen Netzwerk ANN von Fig. 1 sein. Die externe Bankadresse EBA, die externe Zeilenadresse ERA und die externe Spaltenadresse ECA, die in dem Befehls- und Adressensignal CA enthalten sind, das der neuromorphen Vorrichtung 2000 zugeführt wird, können den Eingabedaten (z. B. den Eingabedaten IN von Fig. 1) des künstlichen neuronalen Netzwerks 2000 entsprechen. Verschiedene Befehle, die in dem der neuromorphen Vorrichtung 2000 zugeführten Befehls- und Adress-Signal CA enthalten sind, können den Eingabedaten des künstlichen neuronalen Netzwerks 2000 entsprechen. Lesedaten, die im Lese-DQ-Signal enthalten sind, das von der neuromorphen Vorrichtung 2000 ausgegeben wird, können den Ausgabedaten (z. B. den Ausgabedaten OUT von Fig. 1) des künstlichen neuronalen Netzwerks 2000 entsprechen. Die vorstehend beschriebene interne Bankadresse IBA, die interne Zeilenadresse IRA und die interne Spaltenadresse ICA können Verbindungen (z. B. Synapsen) zwischen den Knoten IN1 bis IN3 der Eingabeschicht IL und den Knoten HN1 bis HN4 der verborgenen Schicht HL, Verbindungen zwischen den Knoten HN1 bis HN8 der verborgenen Schichten HL und Verbindungen zwischen den Knoten HN5 bis HN8 der verborgenen Schicht HL und den Knoten ON1 bis ON3 der Ausgabeschicht OL anzeigen.

[0093] Fig. 18 ist ein Ablaufdiagramm, das ein Betriebsverfahren einer neuromorphen Vorrichtung von Fig. 16 veranschaulicht. In der Operation S110 kann der Befehls- und Adresspuffer 2420 der neuromorphen Vorrichtung 2000 das Befehls- und Adressensignal CA von einer externen Vorrichtung empfangen. In der Operation S120 kann der Befehlsdekodierer 2430 einen Befehl dekodieren und bestimmen, ob ein in der Operation S110 empfangener Befehl ein Schreibbefehl ist.

[0094] Wenn es sich bei dem Befehl um den Schreibbefehl (Y) handelt, kann der Befehlsdekodierer 2430 in der Operation S130 über den Bus 2200 die Bank 2100 auswählen, die einer Bankadresse entspricht, die in einer in der Operation S110 empfangenen Adresse enthalten ist. Der Adressdemultiplexer 2440 kann die externe Zeilenadresse ERA und die externe Spaltenadresse ECA, die in der in der Operation S110 empfangenen Adresse enthalten sind, über den Bus 2200 an die Bank 2100 übertragen, die der Bankadresse entspricht. Der DQ-Puffer 2470 und der DQS-Puffer 2480 können das Schreib-DQ-Signal einschließlich der Schreibdaten und das Schreib-DQS-Signals empfangen, und die Daten-Eingabe-/Ausgabeschaltung 2460 kann die Schreibdaten verarbeiten und die Schreibdaten über den

Bus 2200 oder den programmierbaren Router 2300 an die Bank 2100 übertragen, die der Bankadresse entspricht. Hier können die Schreibdaten gewichtet oder verzerrt werden. Der Zeilendekodierer 2160 und der Spaltendekodierer 2170 der Bank 2100, die der Bankadresse entsprechen, können Speicherzellen auswählen, die der externen Zeilenadresse ERA und der externen Spaltenadresse ECA entsprechen, und das Schreibtreiber-Array 2180 kann die Schreibdaten in die ausgewählten Speicherzellen schreiben.

[0095] Wenn es sich bei dem Befehl um einen Lesebefehl (N) handelt, kann die neuromorphe Vorrichtung 2000 in der Operation S140 die neuronale Netzwerk-Operation durchführen. Der in der Operation S110 empfangene Befehl kann z. B. der Lesebefehl sein. In diesem Fall können der in der Operation S110 empfangene Befehl und die in der Operation S110 empfangene Adresse als Eingabedaten angesehen werden, die in die Eingabeschicht IL des neuronalen Netzwerks ANN eingegeben werden müssen, die unter Bezugnahme auf **Fig. 1** beschrieben wird. Der Befehlsdekodierer 2430 kann über den Bus 2200 die Bank 2100 auswählen, die einer Bankadresse entspricht, die in der in der Operation S110 empfangenen Adresse enthalten ist. Der Adressdemultiplexer 2440 kann die externe Zeilenadresse ERA und die externe Spaltenadresse ECA, die in der in der Operation S110 empfangenen Adresse enthalten sind, über den Bus 2200 an die Bank 2100 übertragen, die der Bankadresse entspricht. Das Neuronenschaltungs-Array 2120 kann die in den Speicherzellen gespeicherten Gewichte und/oder Biases summieren, die der externen Zeilenadresse ERA und der externen Spaltenadresse ECA entsprechen, und kann Aktivierungssignale ausgeben. Das heißt, Wortleitungen und Spaltenauswahlleitungen, die der externen Zeilenadresse ERA und der externen Spaltenadresse ECA entsprechen, können eine Rolle von Eingabedaten spielen, die in die Eingabeschicht IL des neuronalen Netzwerks ANN eingegeben und mit Gewichten für die Eingabeschicht IL multipliziert werden. Der Zeilen-/Spalten-Controller 2150 kann die interne Zeilenadresse IRA und die interne Spaltenadresse ICA auf der Grundlage von Aktivierungssignalen erzeugen. Das Neuronenschaltungs-Array 2120 kann in den Speicherzellen gespeicherte Gewichte und/oder Biases entsprechend der internen Zeilenadresse IRA und der internen Spaltenadresse ICA summieren und kann darüber hinaus Aktivierungssignale ausgeben. Wenn das vorstehend beschriebene Verfahren wiederholt wird, kann das Neuronenschaltungs-Array 2120 wiederverwendet werden. Die interne Zeilenadresse IRA und die interne Spaltenadresse ICA können als Eingabedaten betrachtet werden, die in die verborgene Schicht HL oder die Ausgabeschicht OL des neuronalen Netzwerks ANN eingegeben werden, die unter Bezugnahme auf **Fig. 1** beschrieben sind, und mit Gewichten für die verborgene Schicht

HL oder die Ausgabeschicht OL multipliziert werden. Wortleitungen und Spaltenauswahlleitungen, die der internen Zeilenadresse IRA und der internen Spaltenadresse ICA entsprechen, können eine Rolle als Eingabedaten spielen, die in die verborgene Schicht HL des neuronalen Netzwerks ANN eingegeben werden. Außerdem kann der programmierbare Router 2300 auf der Grundlage von Aktivierungssignalen eine Bankadresse erzeugen, die der anderen Bank 2100 entspricht. Der Zeilen-/Spalten-Controller 2150 der anderen Bank 2100 kann die interne Zeilenadresse IRA und die interne Spaltenadresse ICA auf der Grundlage der vom programmierbaren Router 2300 zugeführten Aktivierungssignale erzeugen. Das Neuronenschaltungs-Array 2120 der anderen Bank 2100 kann die in den Speicherzellen gespeicherten Gewichte oder Biases entsprechend der internen Zeilenadresse IRA und der internen Spaltenadresse ICA summieren und Aktivierungssignale ausgeben. Der Zeilen-/Spalten-Controller 2150 kann ferner die interne Zeilenadresse IRA und die interne Spaltenadresse ICA auf der Grundlage der Aktivierungssignale erzeugen. Das Neuronenschaltungs-Array 2120 kann in den Speicherzellen gespeicherte Gewichte und/oder Biases summieren, die der internen Zeilenadresse IRA und der internen Spaltenadresse ICA entsprechen, und kann ferner Aktivierungssignale ausgeben. Wenn das vorstehend beschriebene Verfahren wiederholt wird, kann die neuromorphe Vorrichtung 2000 die neuronale Netzwerk-Operation durchführen.

[0096] In der Operation S150 kann die Daten-Ein-/Ausgabeschaltung 2460 die Aktivierungssignale von den Banken 2100 über den Bus 2200 oder den programmierbaren Router 2300 empfangen und Lesedaten erzeugen. Unter der Steuerung der Daten-Eingabe-/Ausgabeschaltung 2460 können der DQ-Puffer 2470 und der DQS-Puffer 2480 ein Lese-DQ-Signal, das die Lesedaten enthält, und ein DQS-Signal nach außen ausgeben.

[0097] In einer Ausführungsform kann das unter Bezugnahme auf **Fig. 2** bis 16 beschriebene Schwellenwertsignal TH im Voraus festgelegt werden. In einer anderen Ausführungsform kann das Schwellenwertsignal TH durch den in der Operation S110 erhaltenen Befehl aktualisiert werden. (Eine) Information (en) über einen Spannungspegel und/oder einen Strompegel des Schwellenwertsignals TH kann/können in einem Teilbereich des Speicherzellenarrays 1100/2110 oder in Registern gespeichert werden. Der in der Operation S110 empfangene Befehl kann ein Schreibebeleg zur Anforderung einer Schreiboperation des Teilbereichs des Speicherzellenarrays 1100/2110, ein Modusregister-Setzbefehl, ein Modusregister-Schreibebeleg, ein Mehrzweckbefehl (MPC) usw. sein. Obwohl in **Fig. 16** nicht dargestellt, kann die Peripherieschaltung 2400 weiterhin einen Spannungs- und/oder Stromgenerator zur Erzeu-

gung des Schwellwertsignals TH enthalten, der Befehlsdekodierer 2430 kann die im Teilbereich des Speicherzellenarrays 1100/2100 oder in Registern gespeicherte(n) Information(en) über das Schwellwertsignal TH auf der Grundlage des in der Operation S110 empfangenen Befehls aktualisieren, und der Spannungs- oder Stromgenerator kann das Schwellwertsignal TH in Bezug auf die aktualisierte (n) Information(en) auf verschiedene Weise ändern. Beispielsweise kann der Spannungs- oder der Stromgenerator ein Schwellwertsignal TH oder zwei oder mehr Schwellwertsignale TH erzeugen. Daher kann das eine Schwellwertsignal TH den Neuronenschaltungs-Arrays 2120 aller Bänke 2100 gemeinsam zugeführt werden, oder die zwei oder mehr Schwellwertsignale TH können den Neuronenschaltungs-Arrays 2120 der Bänke 2100 zugeführt werden. Das heißt, die Schwellwertsignale TH der Neuronenschaltungs-Arrays 2120 der Banken 2100 können identisch oder voneinander verschieden sein, und die Schwellwertsignale TH der Neuronenschaltungen 1210 bis 1230 können identisch oder voneinander verschieden sein.

[0098] Fig. 19 veranschaulicht ein Blockdiagramm einer neuromorphen Vorrichtung nach einer Ausführungsform der erfinderischen Konzepte. Eine neuromorphe Vorrichtung 3000 kann Speicher- (Verarbeitungs-) Plättchen 3100 und 3200 und ein Pufferplättchen 3300 enthalten. Die neuromorphe Vorrichtung 3000 kann eine Vielzahl von Kanälen CH1 bis CH3 enthalten. Dabei kann jeder der Kanäle CH1 bis CH3 das Taktsignal CK, das Befehls- und Adress-Signal CA, das DQ-Signal und das DQS-Signal aus Fig. 16 enthalten. Durch Silizium-Durchkontaktierungen TSV können physikalische und/oder elektrische Pfade zwischen den Speicherbausteinen 3100 und 3200 und dem Pufferplättchen 3300 vorgehen werden.

[0099] Das Speicherplättchen 3100 kann aktive Bereiche 3110 und einen Durchisolierungsbereich 3120 enthalten. Alle Komponenten der neuromorphen Vorrichtungen 1000a bis 1000e oder die Komponenten 2100, 2200 und 2300 der neuromorphen Vorrichtung 2000, die unter Bezugnahme auf Fig. 16 beschrieben sind, können in den aktiven Bereichen 3110 angeordnet werden. Die Durchkontaktierungen TSV können im Durchisolierungsbereich 3120 angeordnet werden, oder es können darin Schaltungen zum Senden oder Empfangen von Signalen durch die Durchkontaktierungen TSV angeordnet werden. Der Speicherplättchen 3200 kann so implementiert werden, dass es im Wesentlichen identisch mit dem Speicherplättchen 3100 ist.

[0100] Das Pufferplättchen 3300 (z. B. ein „Kernplättchen“ oder ein „Logikplättchen“) kann die aktiven Bereiche 3310 und den Durchisolierungsbereich 3320 enthalten. Schaltungen zum Empfang oder

zur Ausgabe von Signalen, die von außen über die Kanäle CH1 bis CH3 übertragen werden, alle Komponenten der neuromorphen Vorrichtungen 1000a bis 1000e und/oder alle Komponenten der neuromorphen Vorrichtung 2000, die unter Bezugnahme auf Fig. 16 beschrieben sind, können im aktiven Bereich 3310 angeordnet werden. Die Durchkontaktierungen TSV und/oder Schaltungen zum Senden oder Empfangen von Signalen durch die Durchkontaktierungen TSV können im Durchisolierungsbereich 3320 angeordnet werden.

[0101] Fig. 20 veranschaulicht ein Blockdiagramm eines System-on-Chips nach einer Ausführungsform der erfinderischen Konzepte. Ein System-on-Chip (SoC) 4000 kann ein Anwendungsprozessor (AP) sein. Das System-on-Chip 4000 kann eine neuronale Verarbeitungseinheit (NPU) 4100 enthalten.

[0102] Die NPU 4100 kann Kerne 4110, einen Bus 4120, programmierbare Router 4130 und eine Schnittstellen-Logikschaltung 4140 enthalten. Der Kern 4110 kann ein Speicherzellenarray 4111, Neuronenschaltungs-Arrays 4112, einen Zeilen-/Spalten-Controller 4115, einen Zeilendekodierer 4116, einen Spaltendekodierer 4117 und ein Schreiber-Array 4118 enthalten. Der Kern 4110 und die Komponenten 4111 bis 4118 können im Wesentlichen identisch mit der Bank 2100 und den Komponenten 2110 bis 2180 sein. Wie im Kern der CPU 4200 oder der GPU 4300 kann der Kern 4110 einen Befehl abrufen, ausgeben und/oder senden und eine arithmetische Operation, eine logische Operation und/oder eine Verschiebeoperation einer arithmetischen Logikeinheit (ALU) ausführen. Der Kern 4110 kann die neuronale Netzwerk-Operation unter Verwendung der Komponenten 4111 bis 4118 durchführen. Der Bus 4120 und die programmierbaren Router 4130 können im Wesentlichen identisch mit dem Bus 2200 und den vorstehend beschriebenen programmierbaren Routern 2300 sein.

[0103] Die Schnittstellen-Logikschaltung 4140 kann Schaltungen enthalten, die eingerichtet sind, um die Kerne 4110 und einen von den Kernen 4110 gemeinsam genutzten Cache-Speicher miteinander zu verbinden und zu koppeln. Die Schnittstellen-Logikschaltung 4140 kann zumindest einen Teil oder alle Komponenten 2410 bis 2480 der Peripherieschaltung 2400 der neuromorphen Vorrichtung 2000 enthalten. Die Schnittstellen-Logikschaltung 4140 kann mit anderen Komponenten 4200 bis 4700 im SoC 4100 über den Bus 4800 als Schnittstellenschaltung des NPU 4100 kommunizieren und kann die Kerne 4310 und einen Systembus 4800 verbinden. Das System-on-Chip 4000 kann eine zentrale Verarbeitungseinheit (CPU) 4200, eine Graphikverarbeitungseinheit (GPU) 4300, eine Schnittstellenschaltung 4400, die die Kommunikation mit einer externen Vorrichtung oder einer Benutzervorrichtung

durchführt, ein externer Speicher-Controller 4500, die eine (nicht abgebildete) externe Speichervorrichtung steuert, ein Multimedia-Controller 4600, die eine (nicht abgebildete) Multimediavorrichtung steuert, einen Anzeige-Controller 4700, die eine (nicht abgebildete) Anzeigevorrichtung steuert, und den Systembus 4800 umfassen, der eine Schnittstelle und Kommunikations- (oder Übertragungs-) Wege zwischen den Komponenten 4100 bis 4700 bereitstellt.

[0104] In einer Beispielausführungsform können einige oder alle der Systeme nach einer der Beispielausführungsformen, einschließlich einiger oder aller von der Peripherieschaltung 2400 und der Neuronenschaltung 1200, (einschließlich einiger oder aller von einem Integrator 1300, einer Aktivierungsschaltung 1400, einem Zeilen-/Spaltendekodierer 1500, einem Zeilendekodierer 1600, einem Spaltendekodierer 1700, einem Adressgenerator 1510, einem Vor-Zeilendekodierer 1620, dem Taktpuffer 2410, dem Befehls- und Adresspuffer 2420, dem Befehlsdekodierer 2430, dem Adressdemultiplexer 2440, dem Auffrischungs-Controller 2450, der Daten-Ein-/Ausgabeschaltung 2460, dem DQ-Puffer 2470 und dem DQS-Puffer 2480) eine oder mehrere Instanzen enthalten, darin enthalten sein und/oder durch eine oder mehrere Instanzen (z. B. Artikel, Stücke, Einheiten usw.) von Verarbeitungsschaltungen, wie z. B. Hardware, die Logikschaltungen enthält; eine Hardware/Software-Kombination, wie z. B. ein Prozessor, der Software ausführt; oder eine Kombination davon implementiert werden. Zum Beispiel kann die Verarbeitungsschaltung insbesondere eine zentrale Verarbeitungseinheit (CPU), eine arithmetische Logikeinheit (ALU), einen digitalen Signalprozessor, einen Mikrocomputer, ein Field-Programmable-Gate-Array (FPGA), ein System-on-Chip (SoC), eine programmierbare Logikeinheit, einen Mikroprozessor, einen anwendungsspezifischen integrierten Schaltung (ASIC) usw. enthalten, ist aber nicht darauf beschränkt. In einigen Beispielausführungsformen kann die Verarbeitungsschaltung eine nicht flüchtige computerlesbare Speichervorrichtung oder einen Speicher, z. B. ein Festkörperlaufwerk (SSD), das ein Programm von Befehlen speichert, und einen Prozessor umfassen, der eingerichtet ist, um das Programm von Befehlen auszuführen, um die Funktionalität einiger oder aller Systeme entsprechend einer der Beispielausführungsformen zu implementieren. Es wird davon ausgegangen, dass ein Element (z. B. eine Verarbeitungsschaltung, digitale Schaltungen usw.), das als „Implementierung“ eines Elements (z. B. ein Integrator 1300, eine Aktivierungsschaltung 1400, ein Zeilen-/Spaltendekodierer 1500, ein Zeilendekodierer 1600, ein Spaltendekodierer 1700, ein Adressgenerator 1510 und ein Vor-Zeilendekodierer 1620 usw.) beschrieben wird, so verstanden wird, dass es die Funktionalität des implementierten Elements (z. B. die Funktionalität

der Neuronenschaltung 1200, der Peripherieschaltung 2400 usw.) implementiert.

[0105] Eine neuromorphe Vorrichtung nach einer Ausführungsform der erfinderischen Konzepte, kann auf der Grundlage eines hochintegrierten und leistungsarmen Speichers arbeiten. Die neuromorphe Vorrichtung kann Gewichte und Biases aktualisieren, indem sie eine Schreiboperation des Speichers verwendet. Die neuromorphe Vorrichtung kann eine neuronale Netzwerk-Operation durchführen, indem sie eine zerstörungsfreie Leseoperation des Speichers verwendet. Die neuromorphe Vorrichtung kann die Anzahl der Neuronenschaltungen durch Wiederverwendung der Neuronenschaltungen erhöhen.

[0106] Während die erfinderischen Konzepte unter Bezugnahme auf beispielhafte Ausführungsformen derselben beschrieben wurden, wird es für Fachleute offensichtlich sein, dass verschiedene Änderungen und Modifikationen daran vorgenommen werden können, ohne vom Geist und Umfang der erfinderischen Konzepte, wie sie in den folgenden Ansprüchen dargelegt werden, abzuweichen.

Patentansprüche

1. Neuromorphe Vorrichtung (1000a), umfassend:
 ein erstes Speicherzellenarray (1100), das erste Speicherzellen (MC1, MC3) enthält, die einer ersten Adresse entsprechen, wobei die ersten Speicherzellen (MC1, MC3) eingerichtet sind, um erste Gewichte (W_1 , W_3) zu speichern, und zweite Speicherzellen (MC2, MC4), die einer zweiten Adresse entsprechen, wobei die zweiten Speicherzellen eingerichtet sind, um zweite Gewichte (W_2 , W_4) zu speichern; und
 eine erste Neuronenschaltung (1200, 1210), die einen ersten Integrator (1300) und eine erste Aktivierungsschaltung (1400) enthält, wobei der erste Integrator (1300) eingerichtet ist, um erste Lesesignale aus den ersten Speicherzellen (MC1, MC3) zu summieren und ein erstes Summensignal an die erste Aktivierungsschaltung (1400) auszugeben und zweite Lesesignale aus den zweiten Speicherzellen (MC2, MC4) zu summieren und zweite Summensignale an die erste Aktivierungsschaltung (1400) auszugeben, und die erste Aktivierungsschaltung (1400) eingerichtet ist, um ein erstes Aktivierungssignal (ACTa) auf der Grundlage des ersten Summensignals auszugeben und ein zweites Aktivierungssignal (ACTb) auf der Grundlage der zweiten Summensignale auszugeben, die von dem Integrator (1300) ausgegeben werden; und
 einen Adressgenerator (1510), der eingerichtet ist, um die zweite Adresse auf der Grundlage des ersten Aktivierungssignals (ACTa) zu erzeugen und eine dritte Adresse auf der Grundlage des zweiten

Aktivierungssignals (ACTb) zu erzeugen; und einen Zeilendekodierer (1600), der eingerichtet ist, um eine erste interne Zeilenadresse der ersten Adresse bzw. eine zweite interne Zeilenadresse der zweiten Adresse zu dekodieren; und einen Spaltendekodierer (1700), der eingerichtet ist, um eine erste interne Spaltenadresse der ersten Adresse bzw. eine zweite interne Spaltenadresse der zweiten Adresse zu dekodieren, wobei der Zeilendekodierer (1600) und der Spaltendekodierer (1700) nacheinander die ersten Speicherzellen und die zweiten Speicherzellen auswählen, wobei der Zeilendekodierer (1600) eingerichtet ist, um die erste interne Zeilenadresse (IRA) zu dekodieren, um gleichzeitig erste Wortleitungen (WL1, WL2) auszuwählen, die jeweils mit den ersten Speicherzellen (MC1, MC3) verbunden sind, und die zweite interne Zeilenadresse zu dekodieren, um gleichzeitig zweite Wortleitungen (WL3, WL4) auszuwählen, die jeweils mit den zweiten Speicherzellen (MC2, MC4) verbunden sind.

2. Neuomorphe Vorrichtung nach Anspruch 1, wobei zumindest eines der zweiten Gewichte (W_2 , W_4) eine Verbindungsstärke zwischen der ersten Neuronenschaltung (1200, 1210), die die ersten Lesesignale empfängt, und der Neuronenschaltung (1200, 1210), die die zweiten Lesesignale empfängt, anzeigt.

3. Neuomorphe Vorrichtung nach Anspruch 1, wobei der Zeilendekodierer (1600) eingerichtet ist, um ein oberes Bit eines Zeilenadressenbits, das zur Auswahl von Wortleitungen (WL1, WL2) von Unter-Speicherblöcken (1110_1, 1110_2) verwendet wird, auf ein Don't-Care-Bit zu setzen, und wobei die ersten Wortleitungen (WL1, WL2) und die zweiten Wortleitungen (WL3, WL4) jeweils in den Unter-Speicherblöcken (1110_1, 1110_2) angeordnet sind.

4. Neuomorphe Vorrichtung nach Anspruch 1, wobei die erste Aktivierungsschaltung (1400) eingerichtet ist, um den Zeilendekodierer (1600), den Spaltendekodierer (1700) und/oder den Integrator (1300) unter Verwendung zumindest eines von dem ersten Aktivierungssignal (ACTa) und dem zweiten Aktivierungssignal (ACTb) zurückzusetzen.

5. Neuomorphe Vorrichtung nach Anspruch 4, wobei das erste Speicherzellenarray (1100) ferner enthält: dritte Speicherzellen, die der dritten Adresse entsprechen, wobei die dritten Speicherzellen eingerichtet sind, um dritte Gewichte zu speichern; und vierte Speicherzellen, die einer vierten Adresse entsprechen, wobei die vierten Speicherzellen eingerichtet sind, um vierte Gewichte zu speichern, und wobei der Spaltendekodierer (1700) als Reaktion auf zumindest eines von dem ersten Aktivierungs-

signal (ACTa) und dem zweiten Aktivierungssignal (ACTb) eingerichtet ist, um ein erstes Spaltenauswahlsignal zu deaktivieren, das erste Bitleitungen auswählt, die mit den ersten Speicherzellen und mit den zweiten Speicherzellen verbunden sind, und eingerichtet ist, um ein zweites Spaltenauswahlsignal zu deaktivieren, das zweite Bitleitungen auswählt, die mit den dritten Speicherzellen und mit den vierten Speicherzellen verbunden sind.

6. Neuomorphe Vorrichtung (1000d) nach Anspruch 1, ferner umfassend: einen programmierbaren Router (2300), der eingerichtet ist, um eine Bankadresse auf der Grundlage von zumindest einem von dem ersten Aktivierungssignal oder dem zweiten Aktivierungssignal zu erzeugen; ein zweites Speicherzellenarray, das dritte Speicherzellen enthält, die der Bankadresse entsprechen und eingerichtet sind, um dritte Gewichte zu speichern; und eine zweite Neuronenschaltung (1220), die einen zweiten Integrator (1320) und eine zweite Aktivierungsschaltung (1420) enthält, wobei der zweite Integrator (1320) eingerichtet ist, um dritte Lesesignale aus den dritten Speicherzellen zu summieren und dritte Summensignale auszugeben, und die zweite Aktivierungsschaltung (1420) eingerichtet ist, um ein drittes Aktivierungssignal auf der Grundlage des dritten Summensignals der dritten Lesesignale auszugeben, und wobei zumindest eines der dritten Gewichte eine Verbindungsstärke zwischen der ersten Neuronenschaltung (1210) und der zweiten Neuronenschaltung (1220) anzeigt.

7. Neuomorphe Vorrichtung nach Anspruch 1, wobei die ersten Speicherzellen (MC1, MC3) und die zweiten Speicherzellen (MC2, MC4) eingerichtet sind, um jeweils ein oder mehrere Bits zu speichern.

8. Neuomorphe Vorrichtung (1000d), umfassend: eine Neuronenschaltung (1210), wobei die Neuronenschaltung (1210) einen Integrator (1310) und eine Aktivierungsschaltung (1410) enthält, die ein Summensignal des Integrators (1310) mit einem Schwellwertsignal vergleicht, um ein Impulssignal auszugeben, wobei der Integrator (1310) enthält einen ersten lokalen Leseverstärker (1311_1), der eingerichtet ist, um an eine erste globale Ein-/Ausgabeleitung ein erstes Gewicht auszugeben, das in einer ersten Speicherzelle gespeichert und von einem ersten Bitleitungs-Leseverstärker (1111_1) erfasst wird; einen zweiten lokalen Leseverstärker (1311_1), der eingerichtet ist, um an die erste globale Ein-/Ausgabeleitung ein zweites Gewicht auszugeben, das in einer zweiten Speicherzelle gespeichert und von einem zweiten Bitleitungs-Leseverstärker (1111_1)

erfasst wird;
 einen dritten lokalen Leseverstärker (1311_2), der eingerichtet ist, um an eine zweite globale Ein-/Ausgabeleitung ein drittes Gewicht auszugeben, das in einer dritten Speicherzelle gespeichert und von einem dritten Bitleitungs-Leseverstärker (1111_2) erfasst wird;
 einen vierten lokalen Leseverstärker (1311_2), der eingerichtet ist, um an die zweite globale Ein-/Ausgabeleitung ein viertes Gewicht auszugeben, das in einer vierten Speicherzelle gespeichert und von einem vierten Bitleitungs-Leseverstärker (1111_2) erfasst wird;
 einen ersten globalen Leseverstärker (1312), der eingerichtet ist, um die ersten und zweiten Gewichte zu erfassen, die an der ersten globalen Ein-/Ausgabeleitung summiert werden;
 einen zweiten globalen Leseverstärker (1312), der eingerichtet ist, um das dritte und vierte Gewicht, das an der zweiten globalen Ein-/Ausgabeleitung summiert wird, zu erfassen; und
 eine Summenschaltung (1313), die eingerichtet ist, um das erste bis vierte von dem ersten und zweiten globalen Leseverstärker erfasste Gewicht zu summieren, um das Summensignal auszugeben.

9. Neuromorphe Vorrichtung (1000e) nach Anspruch 8, ferner umfassend:
 ein Speicherzellenarray, wobei das Speicherzellenarray enthält erste bis vierte Speicherzelle;
 ersten bis vierten Bitleitungs-Leseverstärker, die eingerichtet sind, um das erste bis vierte Gewicht an erste bis vierte lokale Ein-/Ausgabeleitung auszugeben;
 einen ersten Ein-/Ausgabeschalter (1910_1), der die erste lokale Ein-/Ausgabeleitung (LIO1) elektrisch mit der ersten globalen Ein-/Ausgabeleitung (GIO) verbindet;
 einen zweiten Ein-/Ausgabeschalter (1910_2), der die zweite lokale Ein-/Ausgabeleitung (LIO2) elektrisch mit der ersten globalen Ein-/Ausgabeleitung (GIO) verbindet;
 einen dritten Ein-/Ausgabeschalter (1910_3), der die dritte lokale Ein-/Ausgabeleitung (LIO3) elektrisch mit der zweiten globalen Ein-/Ausgabeleitung verbindet; und
 einen vierten Ein-/Ausgabeschalter, der die vierte lokale Ein-/Ausgabeleitung elektrisch mit der zweiten globalen Ein-/Ausgabeleitung verbindet.

10. Neuromorphe Vorrichtung (1000e) nach Anspruch 9, ferner umfassend:
 einen Zeilendekodierer (1610_1), der eingerichtet ist, um eine erste Wortleitung (WL1), die mit der ersten und dritten Speicherzelle verbunden ist, und eine zweite Wortleitung (WL2), die mit der zweiten und vierten Speicherzelle verbunden ist, zu steuern; und
 einen Spaltendekodierer (1710_1), der eingerichtet ist, um eine Spaltenauswahlleitung (CSL) zu

steuern, über die ein Spaltenauswahlsignal übertragen wird, wobei die Spaltenauswahlleitung (CSL) zumindest eine mit der ersten Speicherzelle verbundene erste Bitleitung elektrisch mit der ersten lokalen Ein-/Ausgabeleitung verbindet und eine mit der zweiten Speicherzelle verbundene zweite Bitleitung elektrisch mit der zweiten lokalen Ein-/Ausgabeleitung verbindet oder eine mit der dritten Speicherzelle verbundene dritte Bitleitung elektrisch mit der dritten lokalen Ein-/Ausgabeleitung verbindet und eine mit der vierten Speicherzelle verbundene vierte Bitleitung elektrisch mit der vierten lokalen Ein-/Ausgabeleitung verbindet.

11. Neuromorphe Vorrichtung nach Anspruch 10, ferner umfassend:
 einen Adressgenerator (1510), der eingerichtet ist, um auf der Grundlage des Impulssignals eine interne Zeilenadresse (IRA) und eine interne Spaltenadresse (ICA) zu erzeugen.

Es folgen 19 Seiten Zeichnungen

FIG. 2

1000a

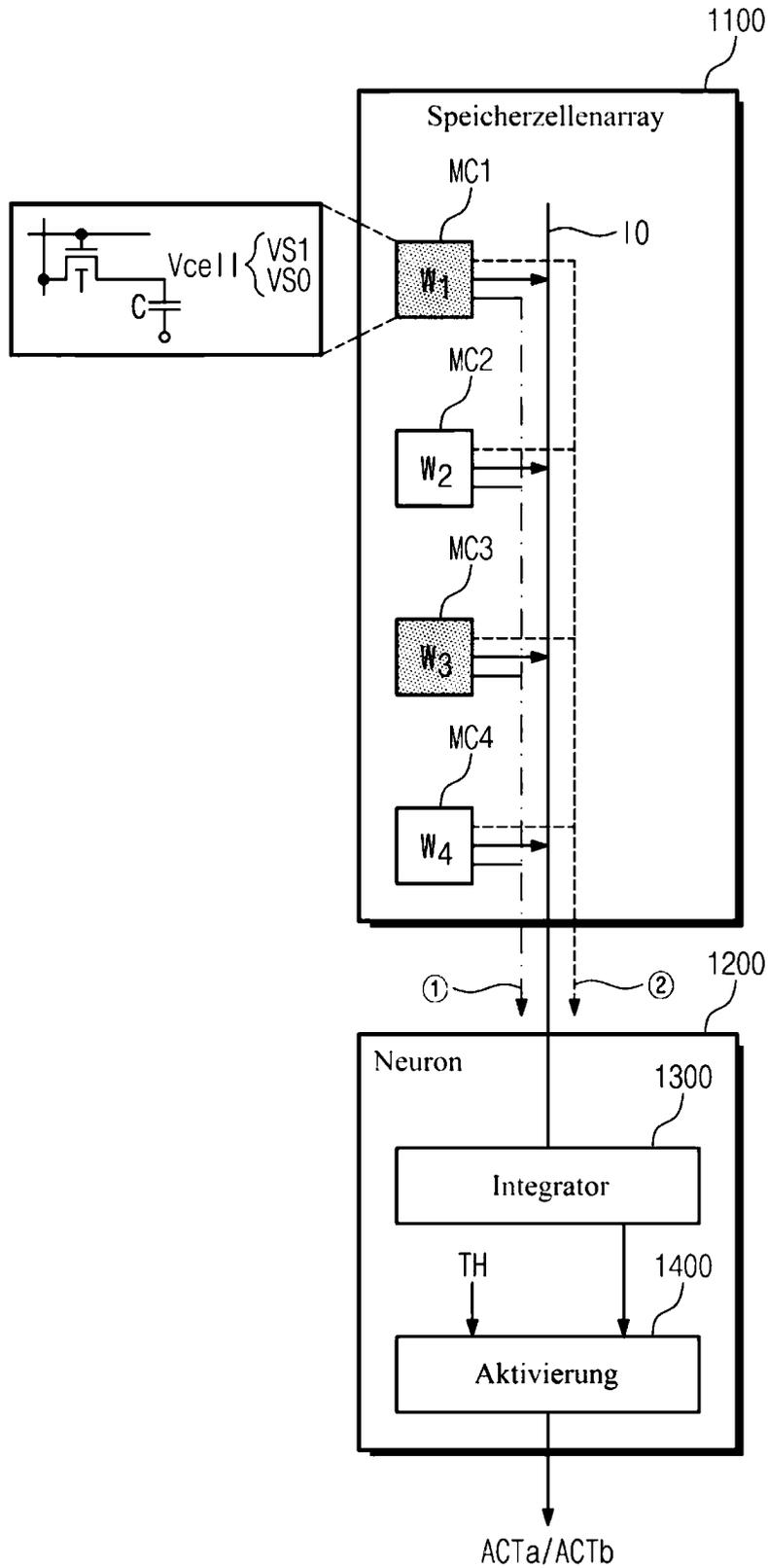


FIG. 3

1000b

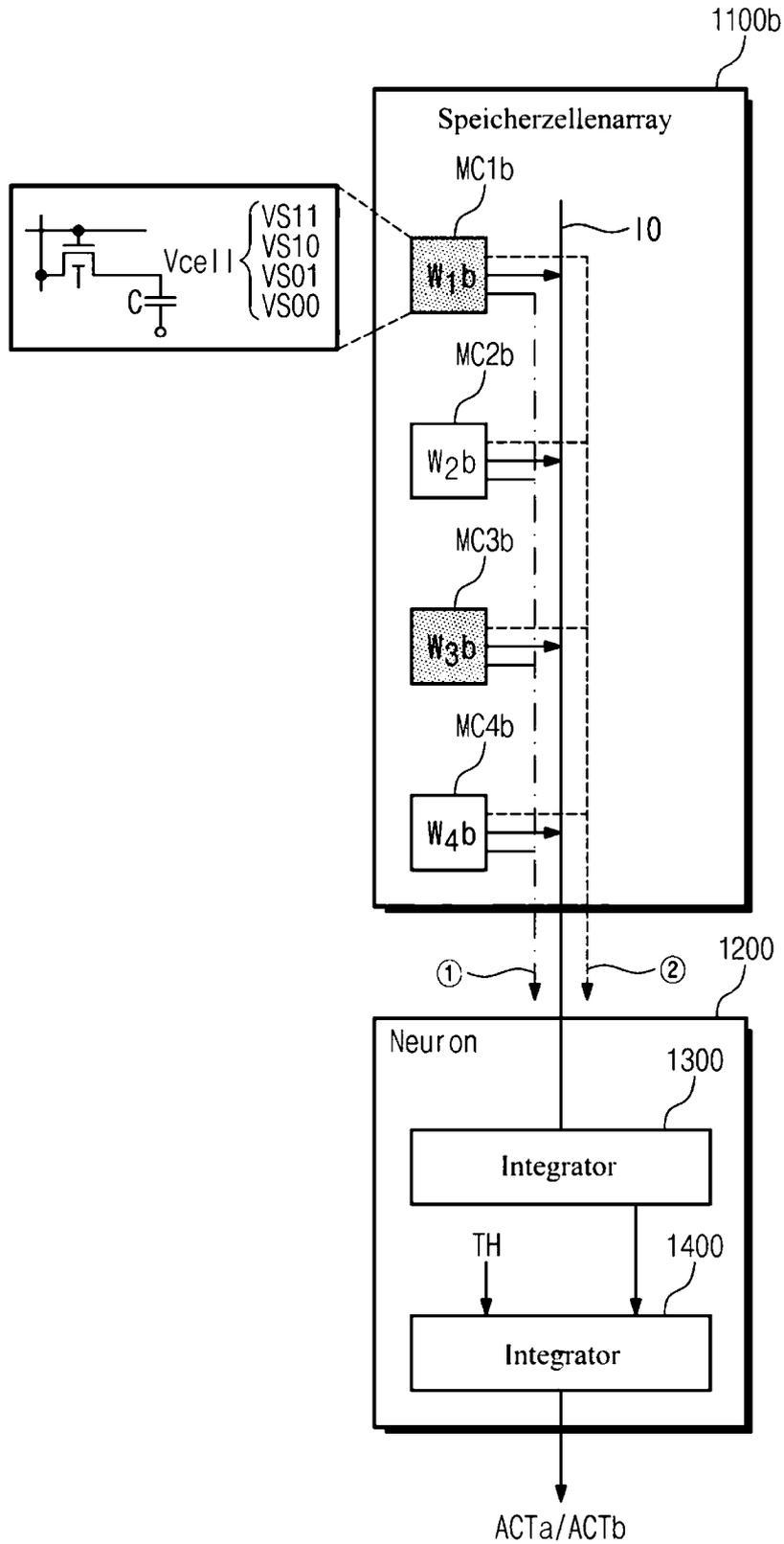


FIG. 4

1000c

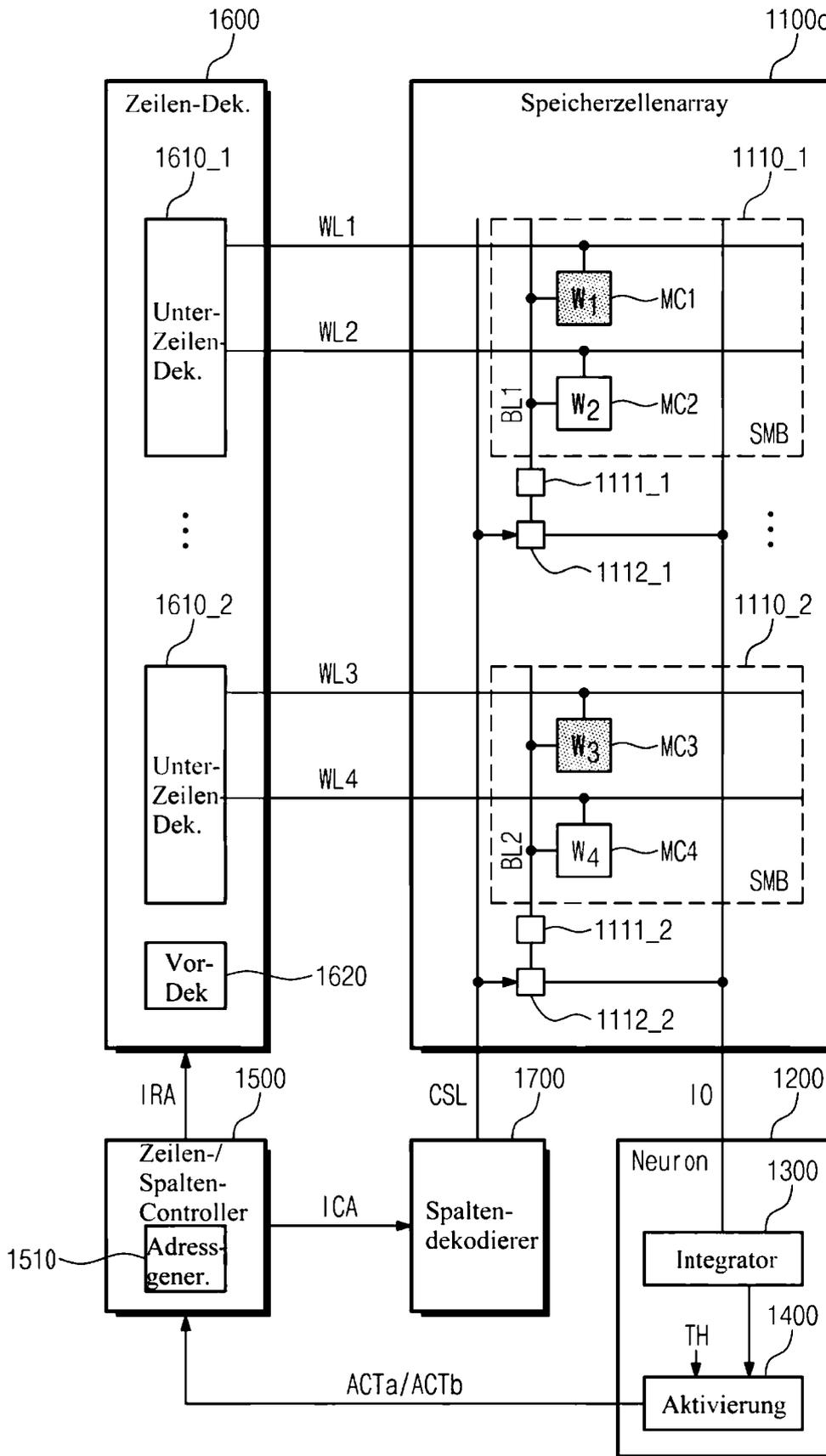


FIG. 5

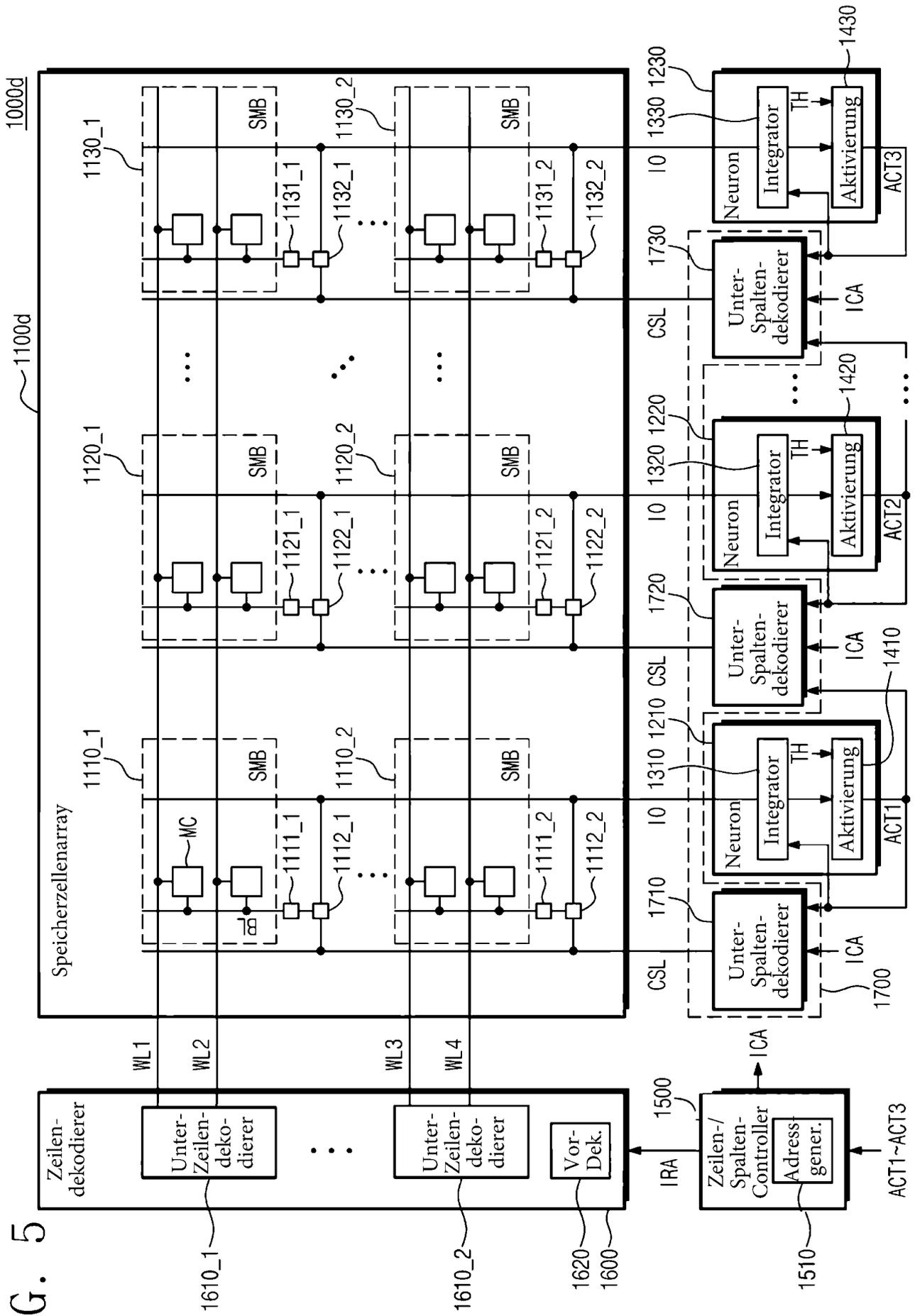


FIG. 6

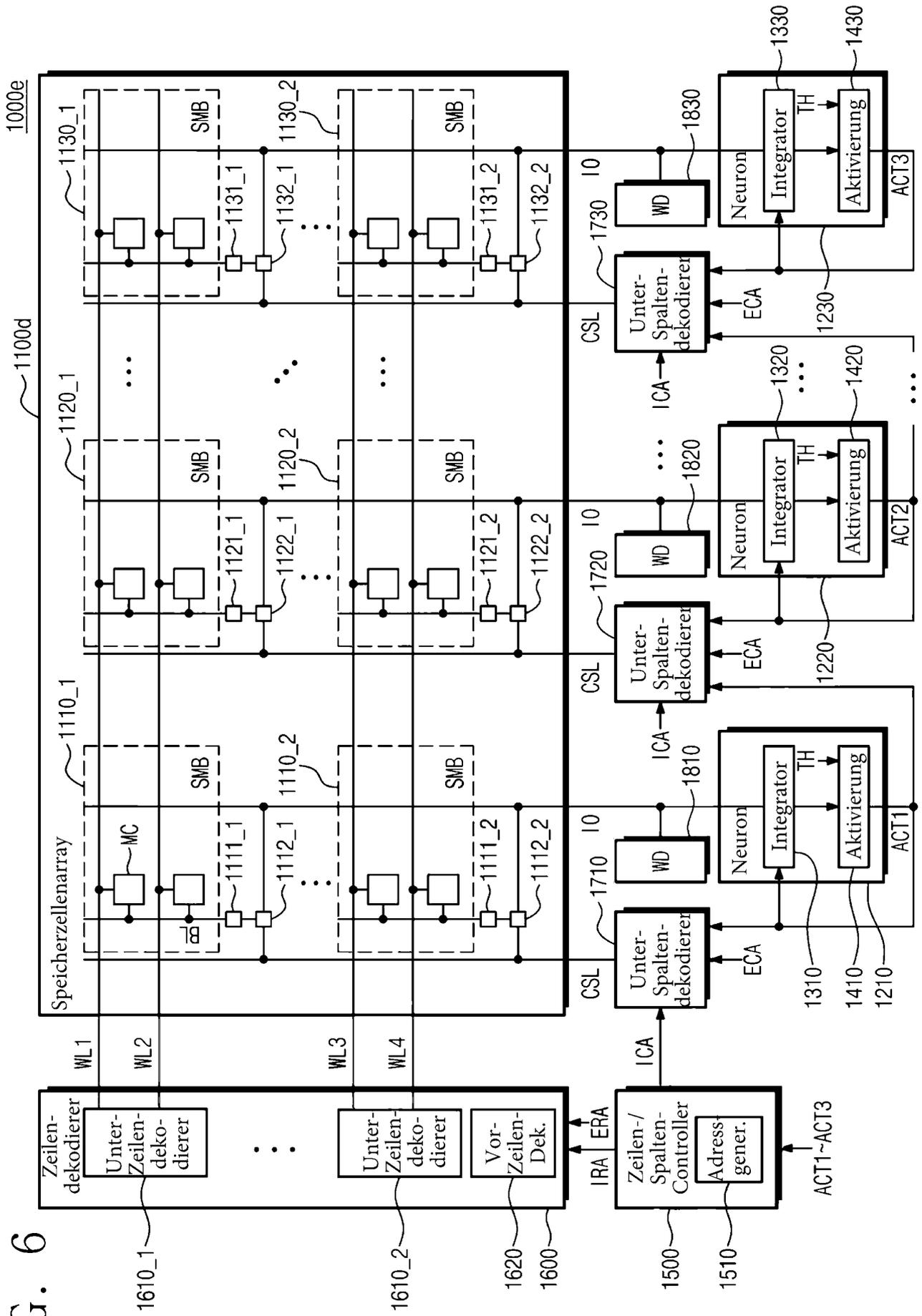


FIG. 7

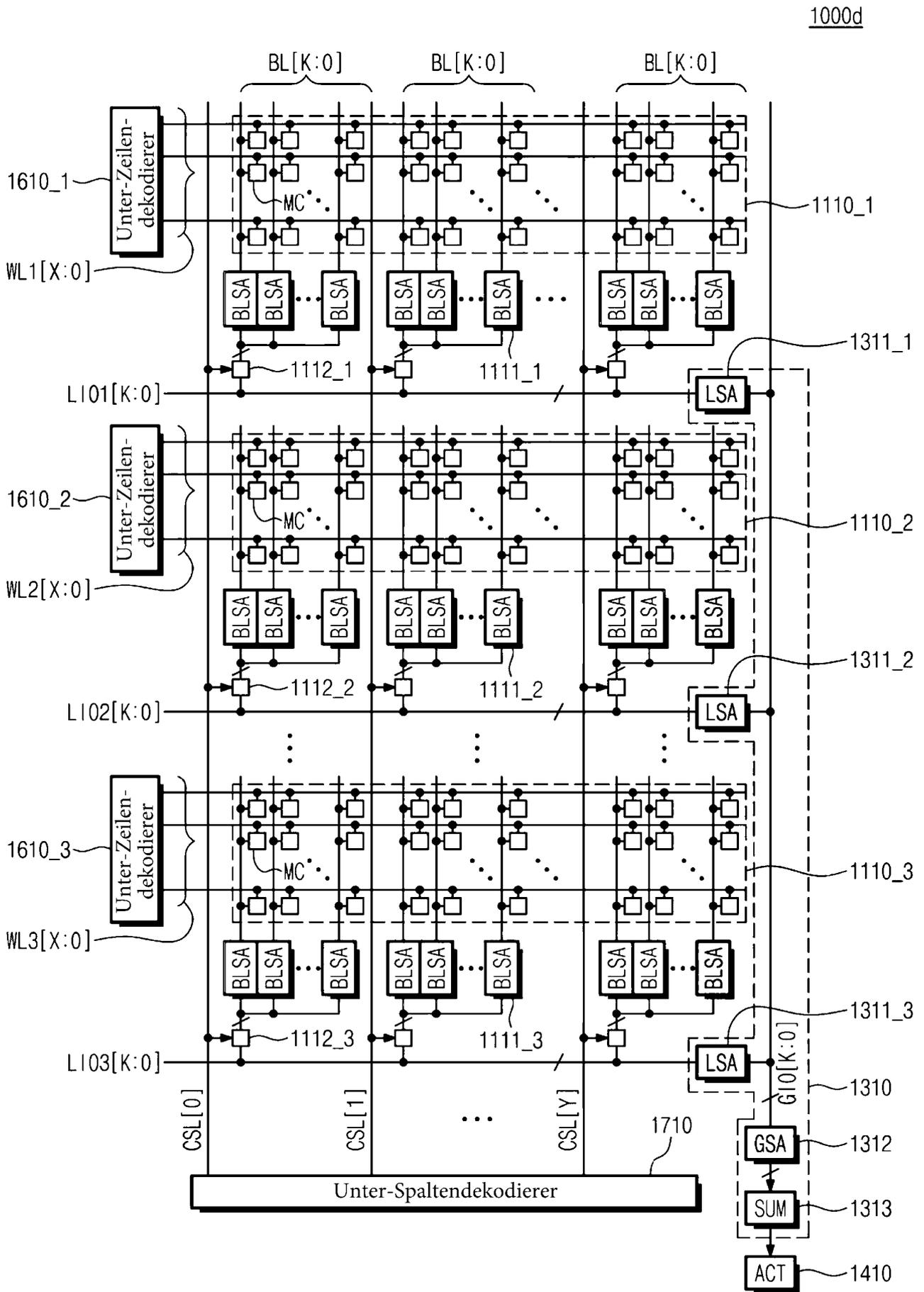


FIG. 8

1000e

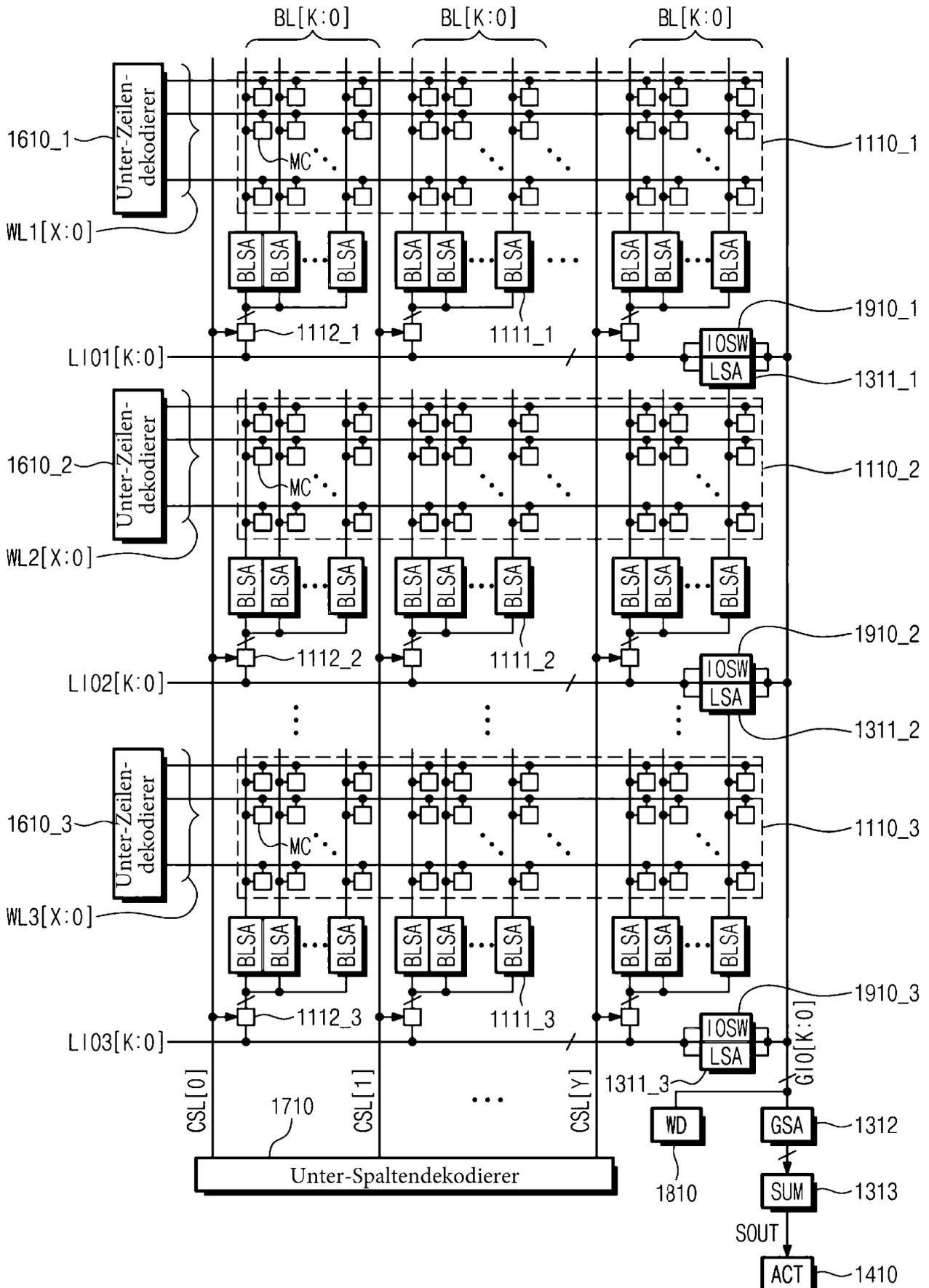


FIG. 9

1000d

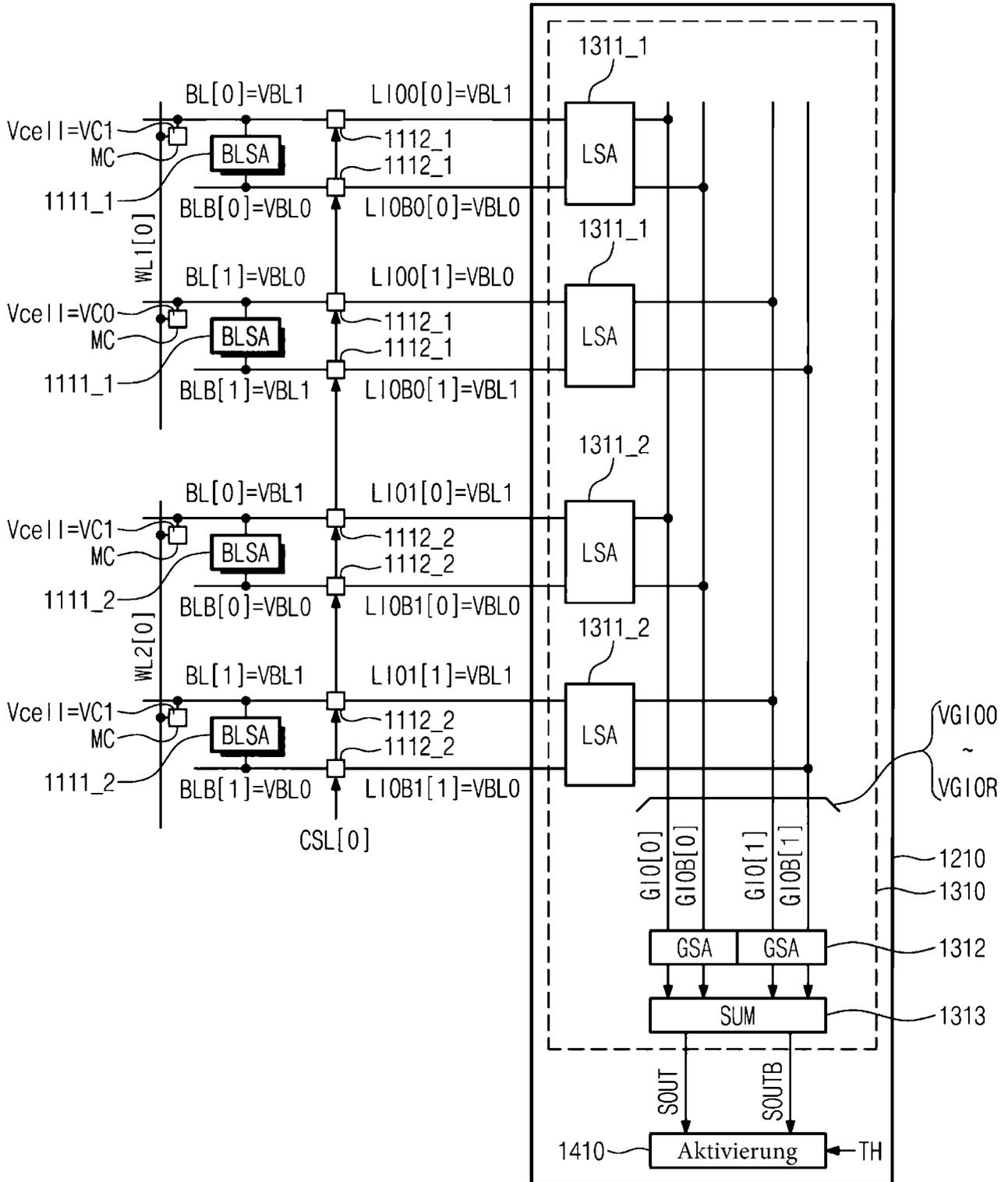


FIG. 10

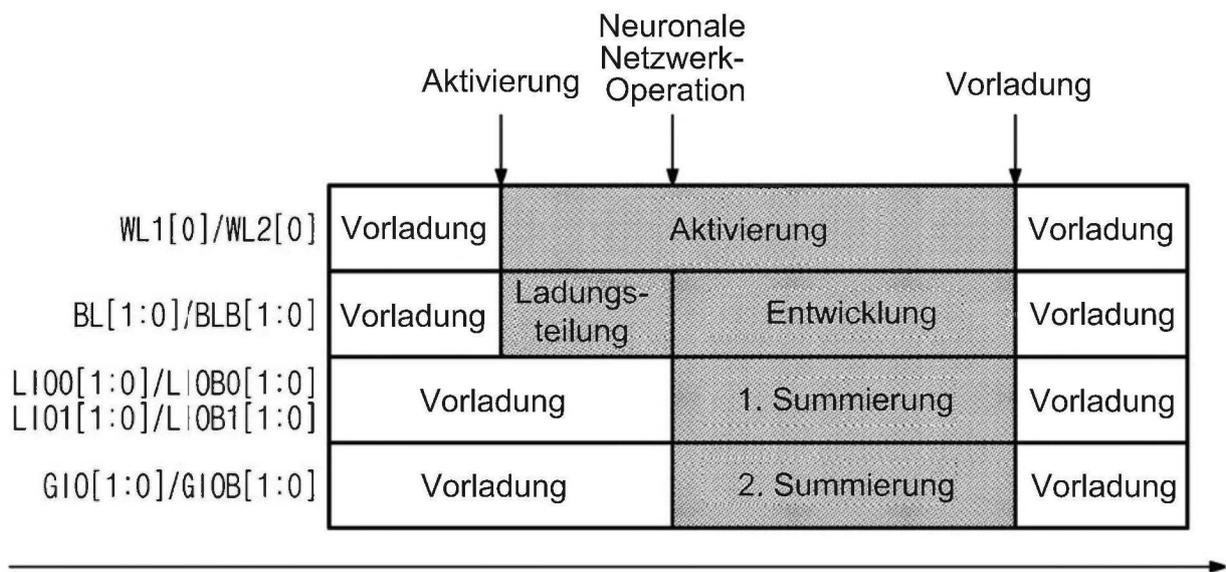


FIG. 11

1000d

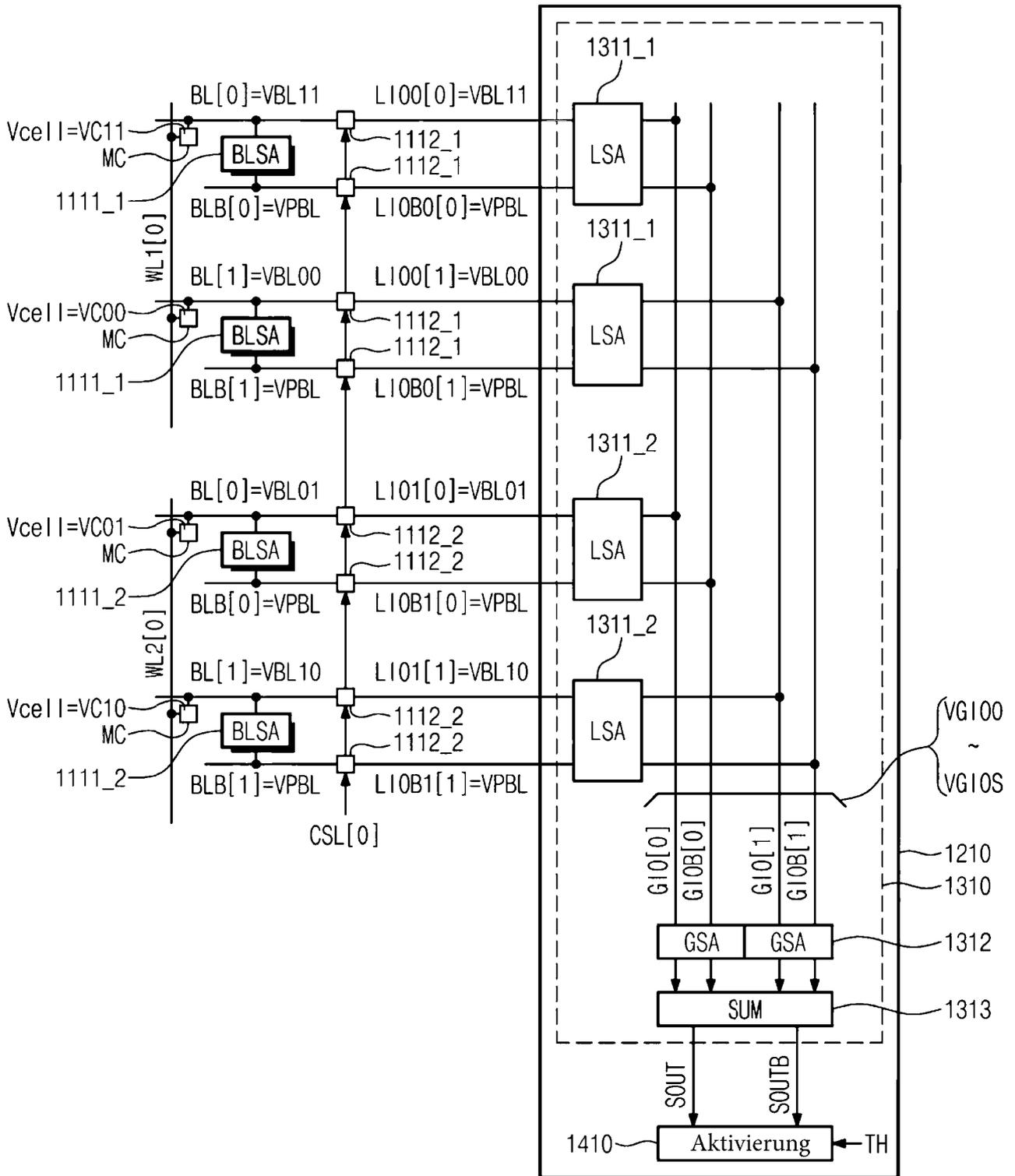


FIG. 12

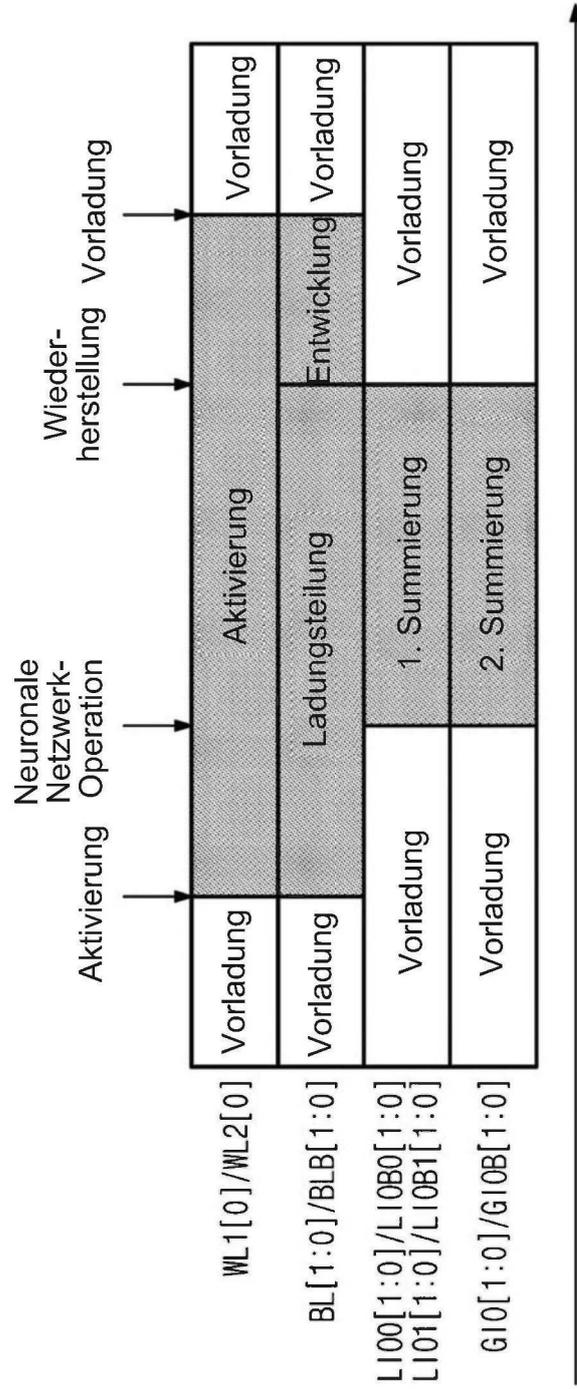


FIG. 13

1310

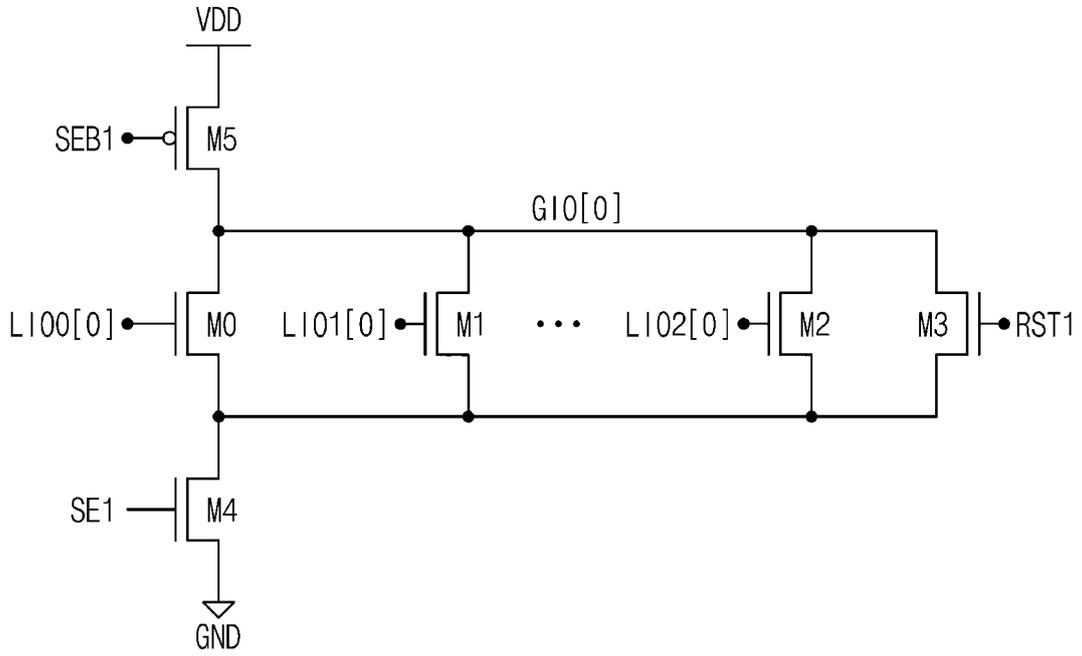


FIG. 14

1313

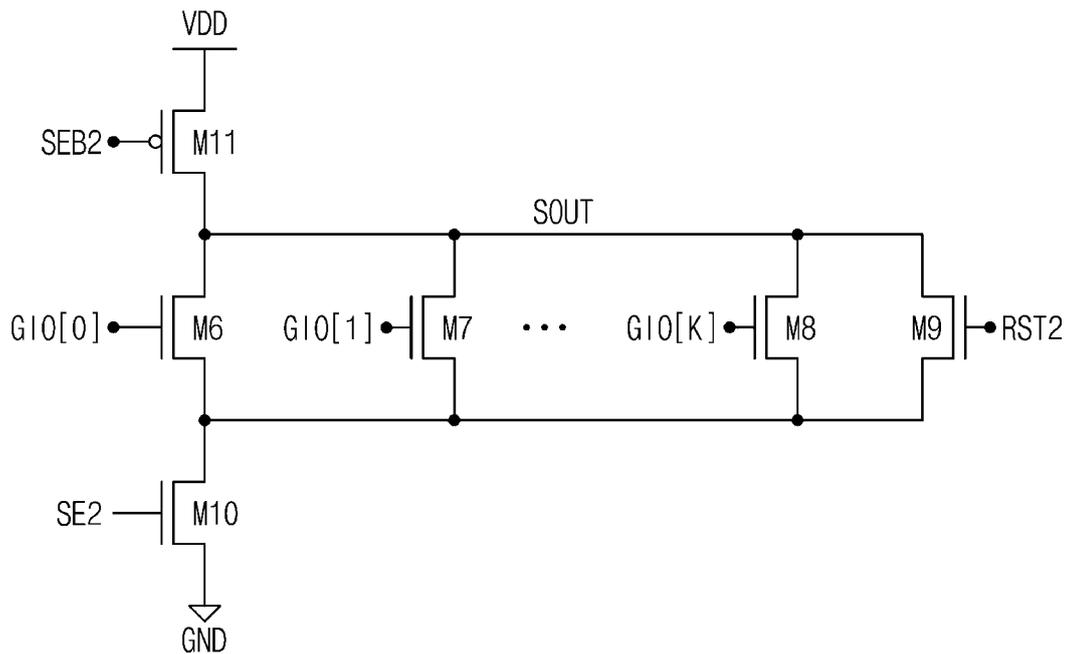


FIG. 15

1410

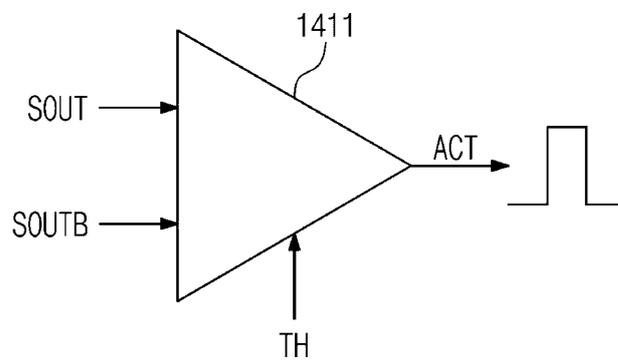


FIG. 16

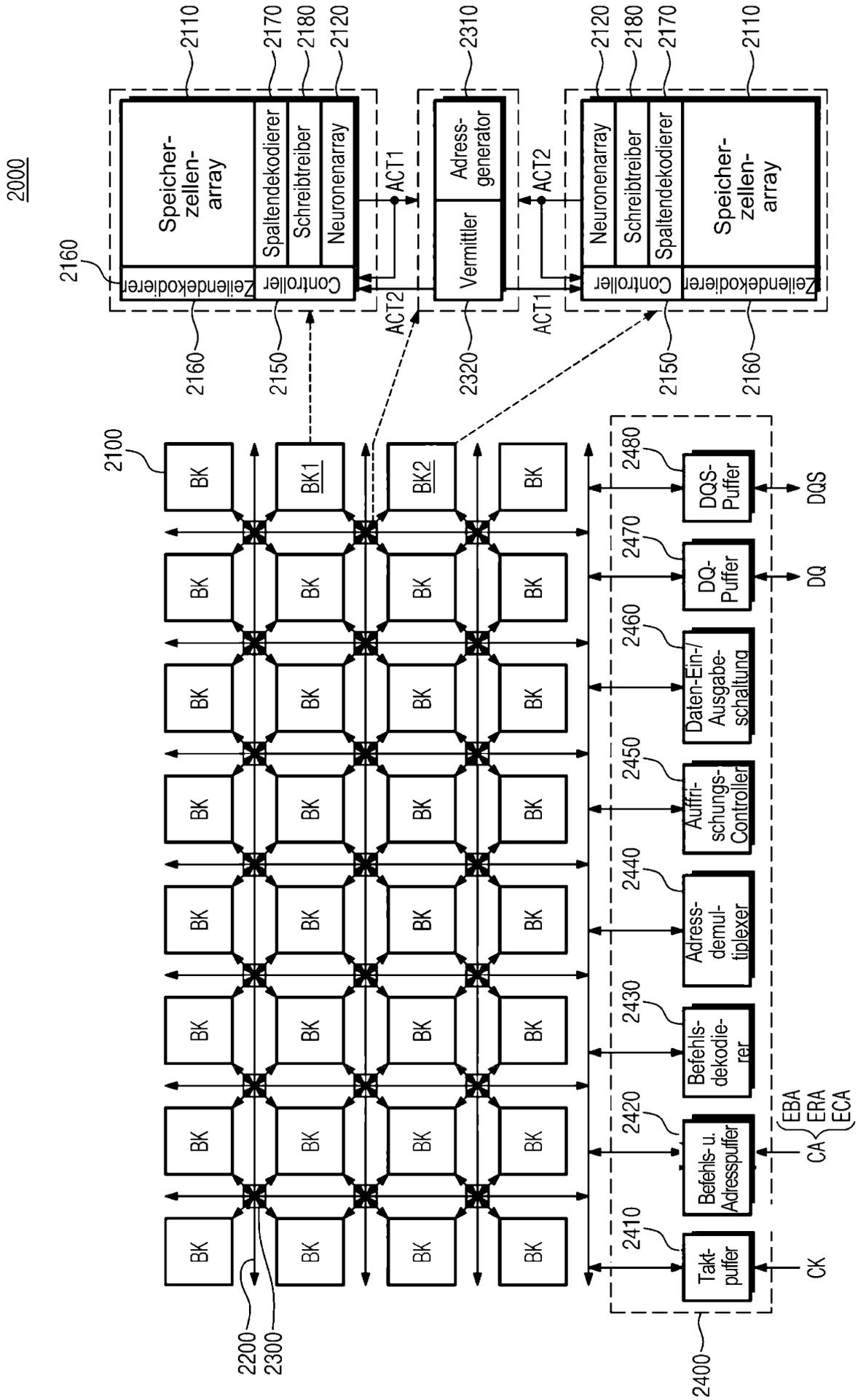


FIG. 18

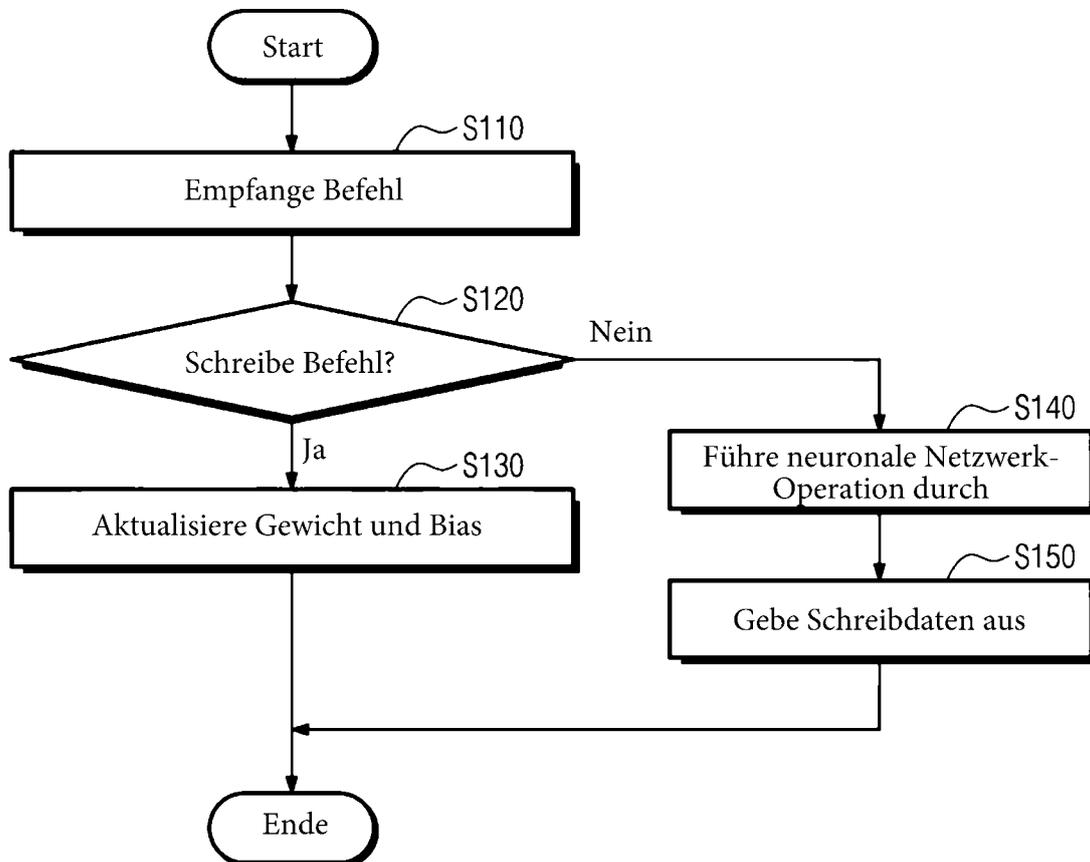


FIG. 19

3000

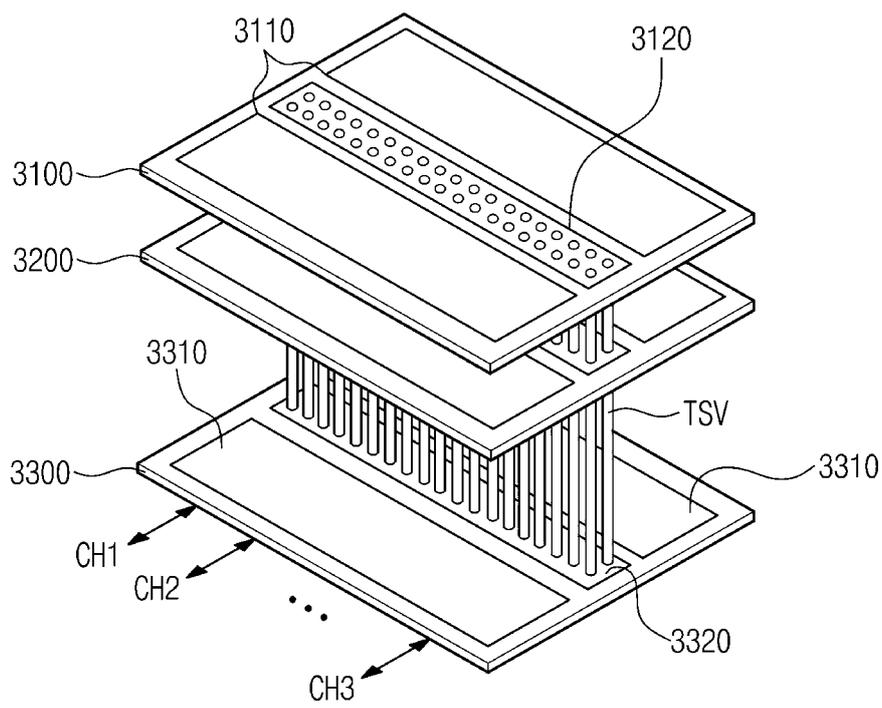


FIG. 20

