

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 015 517**

51 Int. Cl.:

**G06N 3/049** (2013.01)

**G06N 3/065** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.04.2022 PCT/EP2022/060197**

87 Fecha y número de publicación internacional: **20.10.2022 WO22219195**

96 Fecha de presentación y número de la solicitud europea: **16.04.2022 E 22724651 (9)**

97 Fecha y número de publicación de la concesión europea: **12.02.2025 EP 4323923**

54 Título: **Red neuronal de impulsos de múltiples segmentos reconfigurable jerárquica**

30 Prioridad:

**16.04.2021 US 202163175570 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

**06.05.2025**

73 Titular/es:

**INNATERA NANOSYSTEMS B.V. (100.00%)  
Patrijsweg 20  
2289 EX Rijswijk, NL**

72 Inventor/es:

**ZJAJO, AMIR y  
KUMAR, SUMEET SUSHEEL**

74 Agente/Representante:

**MARTÍN DE LA CUESTA, Alicia María**

**ES 3 015 517 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Red neuronal de impulsos de múltiples segmentos reconfigurable jerárquica

5 **Campo técnico**

La presente invención se refiere a técnicas de reconocimiento automático de señales, y más particularmente, a un sistema y a un método para una red de neuronas de impulso de múltiples segmentos reconfigurable jerárquica.

10 **Antecedentes**

Las redes neuronales profundas (DNN) constituyen la base de un gran número de aplicaciones de aprendizaje automático; empezando por el reconocimiento de voz e imágenes, el número de aplicaciones que utilizan DNN ha aumentado de manera exponencial. Inicialmente, los aceleradores de hardware (redes profundas) se han implementado en lógica digital sincrónica convencional. El alto nivel de paralelismo de las redes neuronales no se reproduce en el procesamiento (normalmente) en serie y multiplexado en el tiempo en los sistemas digitales; por el contrario, las primitivas computacionales del emulador de DNN de hardware realizadas como nodos de computación analógicos, donde los elementos de memoria y procesamiento están ubicados conjuntamente, ofrecen mejoras significativas en lo que se refiere a la velocidad, el tamaño y el consumo de energía.

En los modelos de redes neuronales biológicas, cada neurona individual se comunica de manera asincrónica y a través de eventos dispersos, o impulsos. En una red neuronal de impulsos (SNN) basada en eventos de este tipo, sólo las neuronas que cambian de estado generan impulsos y pueden activar el procesamiento de señales en capas posteriores, con el consiguiente ahorro de recursos computacionales. En particular, las SNN incorporan arquitecturas distribuidas asincrónicas que procesan series temporales binarias dispersas por medio de cálculos locales activados por impulsos, retroalimentación local o global y aprendizaje en línea. En analogía con los sistemas de procesamiento neuronal biológico, los procesadores de SNN basados en señales mixtas conservan dos de sus características fundamentales: la representación explícita del tiempo y el uso explícito del espacio, instanciando circuitos físicos dedicados para cada elemento de neurona/sinapsis.

En esencia, tales implementaciones de SNN adoptan una representación híbrida de señales analógicas-digitales, es decir, los trenes de pulsos/impulsos transmiten información analógica en el momento de los eventos, que se convierten de nuevo en señales analógicas en las dendritas (entradas) de la neurona. La información se codifica mediante patrones de actividad que se producen en poblaciones de neuronas, y las sinapsis (una conexión con las neuronas subsiguientes) pueden adaptar su función dependiendo de los pulsos que reciben, lo que proporciona eficiencia energética en la transmisión de señales y flexibilidad para almacenar y recuperar información. Las SNN pueden aplicarse directamente al reconocimiento de patrones y la fusión de datos de sensores, basándose en el principio de que las características del dominio de amplitud, dominio de tiempo y dominio de frecuencia pueden codificarse en secuencias de impulsos únicas codificadas de manera espacial y temporal. La generación de estas secuencias se basa en el uso de uno o más segmentos de neuronas de impulso.

Los métodos conocidos actualmente para mapear SNN en un sustrato de hardware no permiten una gran flexibilidad en lo que respecta, por ejemplo, al número de neuronas usadas para una función particular dentro de la SNN, la conectividad a través de las sinapsis y la configuración de estas neuronas y sinapsis.

Además, aunque las operaciones elementales requeridas por una SNN se realizan de manera muy eficiente mediante circuitos electrónicos analógicos, el desajuste de fabricación induce distorsiones en sus propiedades funcionales. Véase M.J.M. Pelgrom, A.C.J. Duinmaijer, A.P.G. Welbers, "Matching properties of MOS transistors", IEEE Journal of Solid-State Circuits, vol. 24, n.º 5, págs. 1433-1439, 1989. Especialmente en geometrías de proceso más pequeñas y corrientes de funcionamiento más bajas, estos circuitos son cada vez más sensibles a los efectos cuánticos y al ruido externo, lo que reduce efectivamente la relación señal-ruido y limita el rendimiento del procesamiento. Véase K.L. Shepard, V. Narayanan, "Noise in deep submicron digital design", IEEE International Conference on Computer-Aided Design, págs. 524-531, 1997.

El impacto de estos hechos no ideales aumenta en el caso de las disposiciones grandes, donde el controlador, el sesgo, y las neuronas y las sinapsis se comparten por un mayor número de dispositivos, a lo largo de interconexiones más largas.

El documento WO 2020/099680 A2 da a conocer una red neuronal de impulsos para clasificar señales de entrada. La red neuronal de impulsos comprende una pluralidad de neuronas de impulso, y una pluralidad de elementos sinápticos que interconectan las neuronas de impulso para formar la red. Cada elemento sináptico está adaptado para recibir una señal de entrada sináptica y aplicar un peso a la señal de entrada sináptica para generar una señal de salida sináptica, pudiendo configurarse los elementos sinápticos para ajustar el peso aplicado por cada elemento sináptico. Además, cada una de las neuronas de impulso está adaptada para recibir una o más de las señales de salida sinápticas desde uno o más de los elementos sinápticos, y generar una señal de salida de tren de impulsos espacio-temporal en respuesta a la una o más señales de salida sinápticas recibidas. La red neuronal de impulsos

se divide en múltiples subredes, en la que cada subred comprende un subconjunto de las neuronas de impulso conectadas para recibir señales de salida sináptica desde un subconjunto de los elementos sinápticos. La subred está adaptada para generar una señal de patrón de salida de subred en respuesta a una señal de patrón de entrada de subred aplicada a la subred. Además, cada subred forma parte de uno o múltiples núcleos en una disposición de núcleos, consistiendo cada núcleo en una red programable de neuronas de impulso implementadas en hardware o una combinación de hardware y software. La comunicación entre núcleos en la disposición de núcleos se organiza a través de una estructura de interconexión programable.

El documento US 2012/317062 A1 da a conocer un circuito de red neuronal reconfigurable que comprende una disposición de sinapsis electrónicas que incluye múltiples sinapsis que interconectan una pluralidad de neuronas electrónicas digitales. Cada neurona comprende un integrador que integra impulsos de entrada y genera una señal cuando las entradas integradas superan un umbral. El circuito comprende además un módulo de control para reconfigurar la disposición de sinapsis. El módulo de control comprende una máquina de estado final global que controla la temporización para el funcionamiento del circuito, y un codificador de prioridad que permite que las neuronas de impulso accedan secuencialmente a la disposición de sinapsis.

El documento US 2019/303740 A1 da a conocer un núcleo neurosináptico que está adaptado para aplicar una pluralidad de pesos sinápticos a una pluralidad de activaciones de entrada para producir una pluralidad de activaciones de salida. Se leen los pesos sinápticos para uno de una pluralidad de núcleos lógicos. El núcleo neurosináptico está configurado para implementar dicho uno de la pluralidad de núcleos lógicos usando los pesos sinápticos. Se proporciona al menos un bloque de datos como activaciones de entrada contiguas al núcleo neurosináptico. Las activaciones de entrada se procesan por el núcleo neurosináptico para determinar al menos un bloque contiguo de activaciones de salida.

## Sumario

El segmento neurosináptico dado a conocer en la presente invención está organizado como disposiciones repetitivas de circuitos sinápticos y unidades neuronales. La presente invención abarca modalidades de cómo se dividen y se mapean estas disposiciones en segmentos, y se describen los métodos y mecanismos en los que pueden reconfigurarse y segmentarse la disposición neurosináptica y las conexiones que la acompañan, ofreciendo posteriormente una mayor flexibilidad en el mapeo de las SNN sobre un sustrato de hardware. La metodología de mapeo incorpora una división y un mapeo de segmentos activados por restricciones (en lo que se refiere al tamaño de cada segmento (el número de neuronas), su conectividad (topología y número de sinapsis) y su configuración (pesos y número de capas)), donde la definición elegida y la restricción limitada es una métrica de rendimiento vinculada con la función designada de la red de segmentos. Este enfoque, detallado en el resto de esta divulgación, da como resultado un medio eficaz para realizar grandes redes de neuronas de impulso que pueden reconocer patrones de señales sensoriales complejas.

La invención se define mediante el juego de reivindicaciones adjunto.

Según un primer aspecto de la divulgación, se da a conocer una disposición neurosináptica que comprende una pluralidad de neuronas de impulso, y una pluralidad de elementos sinápticos que interconectan las neuronas de impulso para formar una red neuronal de impulsos implementada al menos parcialmente en hardware. Cada elemento sináptico está dispuesto para recibir una señal de entrada sináptica desde al menos una de múltiples entradas y está adaptado para aplicar un peso a la señal de entrada sináptica para generar una señal de salida sináptica, pudiendo configurarse los elementos sinápticos para ajustar el peso aplicado por cada elemento sináptico. Además, cada una de las neuronas de impulso está dispuesta para recibir una o más de las señales de salida sinápticas desde uno o más de los elementos sinápticos, y está adaptada para generar una señal de salida de tren de impulsos espacio-temporal en respuesta a la una o más señales de salida sinápticas recibidas.

La disposición neurosináptica comprende bloques de peso y bloques de salida, comprendiendo cada bloque de peso uno o más de los elementos sinápticos, y comprendiendo cada bloque de salida una o más de las neuronas y un circuito de conmutación de neuronas. Cada bloque de salida puede conectarse eléctricamente a un subconjunto de bloques de peso a través del circuito de conmutación de neuronas y en el que el circuito de conmutación de neuronas está configurado para conectar eléctricamente de manera selectiva al menos un elemento sináptico comprendido dentro del subconjunto de bloques de peso al menos a una neurona comprendida dentro del bloque de salida respectivo, para obtener una división de la disposición neurosináptica en conjuntos de neuronas conectadas eléctricamente a elementos sinápticos seleccionados.

Según una realización del primer aspecto de la divulgación, el subconjunto de bloques de peso al que está conectado eléctricamente un bloque de salida forma una o múltiples columnas dentro de la disposición de elementos sinápticos y/o en el que los elementos sinápticos comprendidos en un bloque de peso particular se proporcionan dentro de la misma fila de la disposición neurosináptica y/o en el que cada bloque de salida está conectado a una columna de bloques de peso.

Según una realización del primer aspecto de la divulgación, cada uno de los circuitos de conmutación de neuronas

5 comprende trayectorias de señales de conmutación que comprenden cables conductores implementados en un circuito lógico del circuito de conmutación de neuronas, en los que las trayectorias de señales de conmutación están configuradas para poder conmutarse entre diferentes configuraciones, preferiblemente mediante el uso de puertas de transistor, en los que cada configuración determina qué al menos un elemento sináptico comprendido dentro del subconjunto de bloques de peso está conectado eléctricamente a qué al menos una neurona comprendida dentro del bloque de salida.

10 Según una realización del primer aspecto de la divulgación, el circuito de conmutación de neuronas está configurado para reconfigurar dinámicamente las trayectorias de señales de conmutación, preferiblemente en el que la reconfiguración dinámica se basa en una metodología de mapeo que incorpora una división y segmentación activadas por restricciones de la disposición neurosináptica, preferiblemente en el que la segmentación se basa en correspondencia del tamaño del bloque de peso y el bloque de salida con una relación señal-ruido de entrada.

15 Según una realización del primer aspecto de la divulgación, la segmentación de la disposición neurosináptica se realiza basándose en una o más reglas de aprendizaje, tasas de aprendizaje y/o mecanismos de (post-)plasticidad de manera que al menos dos de las subdisposiciones neurosinápticas son distintas en lo que se refiere a la una o más reglas de aprendizaje, tasas de aprendizaje y/o mecanismos de (post-)plasticidad.

20 Según una realización del primer aspecto de la divulgación, al menos uno de los bloques de peso está organizado como una estructura intercalada, tal como para facilitar la conmutación y/o combinación de elementos sinápticos dentro de la disposición neurosináptica, y/o en el que para cada bloque de salida, la conmutación puede controlarse de manera independiente, tal como para obtener una mayor flexibilidad de mapeo, y/o en el que al menos uno de los bloques de salida está organizado como una estructura intercalada y/o en el que la salida de neuronas de cualquiera de los bloques de salida difunde a una o múltiples subdisposiciones neurosinápticas.

25 Según una realización del primer aspecto de la divulgación, cada una de las neuronas dentro de uno de los bloques de salida conduce pasivamente como cambios analógicos graduales en el potencial eléctrico.

30 Según una realización del primer aspecto de la divulgación, la disposición neurosináptica está segmentada en subdisposiciones neurosinápticas que pueden separarse y tienen su propio circuito requerido, y/o están dispuestas para procesar, y optimizarse para, una sola modalidad.

35 Según una realización del primer aspecto de la divulgación, existen interconexiones tanto de largo alcance como de corto alcance entre neuronas de diferentes subdisposiciones neurosinápticas, y en las que existe una conectividad más densa entre subdisposiciones neurosinápticas proximales que entre subdisposiciones neurosinápticas más distantes.

40 Según una realización del primer aspecto de la divulgación, al menos uno de los bloques de salida puede controlarse en el sentido de que el periodo de acumulación y/o la constante de integración de una neurona particular dentro del al menos uno de los bloques de salida puede controlarse a través de señales de control.

45 Según una realización del primer aspecto de la divulgación, un potencial de membrana neuronal de una de las neuronas se implementa en el dominio analógico como una tensión a través de un condensador o como una variable multibit almacenada en capturadores digitales; o en el dominio digital usando circuitos lógicos de CMOS.

50 Según un segundo aspecto de la presente divulgación, se da a conocer un procesador neuronal de impulsos. El procesador neuronal de impulsos comprende un codificador de datos en impulsos que codifica datos digitales o analógicos en impulsos, una disposición neurosináptica según el primer aspecto de la divulgación, dispuesta para tomar los impulsos emitidos por el codificador de datos en impulsos como entrada y dispuesta para emitir un tren de impulsos espacio-temporal como resultado de la entrada, y un decodificador de impulsos dispuesto para decodificar los trenes de impulsos espacio-temporales que se originan a partir de la disposición neurosináptica.

55 Según un tercer aspecto de la presente divulgación, se da a conocer un método para configurar una disposición neurosináptica. La disposición neurosináptica comprende una pluralidad de neuronas de impulso, y una pluralidad de elementos sinápticos que interconectan las neuronas de impulso para formar la red implementada al menos parcialmente en hardware. Cada elemento sináptico está adaptado para recibir una señal de entrada sináptica desde al menos una de múltiples entradas y aplicar un peso a la señal de entrada sináptica para generar una señal de salida sináptica, pudiendo configurarse los elementos sinápticos para ajustar el peso aplicado por cada elemento sináptico. Además, cada una de las neuronas de impulso está adaptada para recibir una o más de las señales de salida sinápticas desde uno o más de los elementos sinápticos, y generar una señal de salida de tren de impulsos espacio-temporal en respuesta a la una o más señales de salida sinápticas recibidas.

65 El método comprende dividir la disposición neurosináptica en bloques de peso y bloques de salida, comprendiendo cada bloque de peso uno o más de los elementos sinápticos, y comprendiendo cada bloque de salida una o más de las neuronas y un circuito de conmutación de neuronas, hacer que cada bloque de salida pueda conectarse eléctricamente a un subconjunto de bloques de peso a través del circuito de conmutación de neuronas, y configurar

el circuito de conmutación de neuronas para conectar eléctricamente de manera selectiva al menos un elemento sináptico comprendido dentro del subconjunto de bloques de peso al menos a una neurona comprendida dentro del bloque de salida respectivo, para obtener una división de la disposición neurosináptica en conjuntos de neuronas conectadas eléctricamente a elementos sinápticos seleccionados.

Según una realización del tercer aspecto de la divulgación, el subconjunto de bloques de peso al que está conectado eléctricamente un bloque de salida forma una o más de múltiples columnas dentro de la disposición de elementos sinápticos y/o en el que los elementos sinápticos comprendidos en un bloque de peso particular se proporcionan dentro de la misma fila de la disposición neurosináptica.

Según una realización del tercer aspecto de la divulgación, cada uno de los circuitos de conmutación de neuronas comprende trayectorias de señales de conmutación que comprenden cables conductores implementados en un circuito lógico del circuito de conmutación de neuronas, en el que las trayectorias de señales de conmutación están configuradas para poder conmutarse entre diferentes configuraciones, preferiblemente mediante el uso de puertas de transistor, en el que cada configuración determina qué al menos un elemento sináptico comprendido dentro del subconjunto de bloques de peso está conectado eléctricamente a qué al menos una neurona comprendida dentro del bloque de salida.

### Breve descripción de los dibujos

Ahora se describirán realizaciones, únicamente a modo de ejemplo, con referencia a los dibujos esquemáticos adjuntos en los que los símbolos de referencia correspondientes indican partes correspondientes, y en los que:

la figura 1 muestra una representación esquemática de una disposición neurosináptica cuadrada con  $n^2$  sinapsis;

las figuras 2A-C muestran representaciones esquemáticas de organización de la disposición neurosináptica reconfigurable: en la figura 2A las unidades organizativas primarias son bloques de peso  $W_{ij}$  y bloques de salida  $O_j$  que reúnen  $m$  pesos y  $m$  neuronas, respectivamente, en la figura 2B se muestra la estructura de bloque de peso  $W_{ij}$ , en la figura 2C se muestra la estructura de un bloque de salida  $O_j$ ;

la figura 2D muestra una realización a modo de ejemplo de posibles configuraciones de un circuito de conmutación de neuronas en un decodificador de columnas/neuronas;

la figura 3 muestra esquemáticamente cómo dividir una disposición reconfigurable a modo de ejemplo (tal como se muestra por ejemplo en la figura 2) en múltiples segmentos/subdisposiciones;

las figuras 4A-D muestran esquemáticamente una organización a modo de ejemplo de la disposición neurosináptica reconfigurable: en la figura 4A,  $n^2$  sinapsis y  $n$  neuronas, en la figura 4B dividiendo la disposición en la figura 4A en dos segmentos, cada uno con  $n^2/2$  sinapsis y  $2 \times n$  neuronas, en la figura 4C dividiendo una disposición en cuatro segmentos, cada uno con  $n^2/4$  sinapsis y  $4 \times n$  neuronas, y en la figura 4D dividiendo una disposición en ocho segmentos, cada uno con  $n^2/8$  sinapsis y  $8 \times n$  neuronas;

la figura 5 muestra una representación a modo de ejemplo de una organización de la disposición neurosináptica reconfigurable: cada uno de los segmentos (subdisposiciones) puede tener un recuento sináptico diferente por neurona;

la figura 6A muestra una representación a modo de ejemplo de un procesador neuronal de impulsos que incorpora codificadores de datos en impulsos, una disposición neurosináptica multisegmentada de señal analógica/mixta, y decodificadores de impulsos;

la figura 6B muestra una representación a modo de ejemplo de una estructura de interconexión.

Las figuras se proporcionan únicamente para fines ilustrativos, y no sirven como restricción del alcance o la protección tal como se establece en las reivindicaciones.

### Descripción de realizaciones

A continuación en el presente documento, se describirán determinadas realizaciones en más detalle. Sin embargo, debe apreciarse que estas realizaciones no deben interpretarse como limitativas del alcance de protección para la presente divulgación.

La figura 1 muestra una representación esquemática de una disposición 100 neurosináptica cuadrada con  $n^2$  sinapsis y  $n$  neuronas. La disposición 100 neurosináptica comprende una matriz 150 de sinapsis con sinapsis 101 y una fila 160 de neuronas con neuronas 103. Las sinapsis y las neuronas pueden implementarse parcial o completamente en hardware.

Cada una de las sinapsis 101 en la matriz 150 de sinapsis tiene un determinado peso  $w_{ij}$  atribuido, donde  $i$  indica la  $i$ -ésima fila de la matriz 150 de sinapsis, y  $j$  indica la  $j$ -ésima columna de la matriz 150 de sinapsis. Cada fila de la matriz 150 de sinapsis se activa por una entrada 102 respectiva. La entrada 102 es una señal eléctrica y puede ser por ejemplo un impulso (tren) de tensión o corriente. Cada columna  $j$  de la matriz 150 de sinapsis activa una neurona 103 de salida respectiva que se indica como  $N_j$ .

Cada uno de los pesos  $w_{ij}$  indica la fuerza o amplitud de la conexión entre la entrada 102 respectiva a la  $i$ -ésima fila de la matriz 150 de sinapsis y la neurona  $N_j$  conectada a la  $j$ -ésima columna de la matriz 150 de sinapsis. Dicho de otro modo, si el peso  $w_{ij}$  aumenta, se refuerza la conexión entre la entrada 102 respectiva a la  $i$ -ésima fila y la neurona  $N_j$  conectada a la  $j$ -ésima columna de la matriz 150 de sinapsis.

Cuando el potencial de membrana de una neurona  $N_j$  particular, que está relacionado con la carga eléctrica de la membrana de la neurona, alcanza un valor específico (denominado el "umbral") la neurona se activa. Esta señal eléctrica se representa como un impulso o bien de tensión o bien de corriente y se envía a otra parte de la SNN a través de una línea 180 de salida de impulsos.

La fila 160 de neuronas puede controlarse en el sentido de que, por ejemplo, el periodo de acumulación y/o la constante de integración de una(s) neurona(s) 103 particular(es) puede controlarse a través de señales 170 de control (aunque sólo se muestra una señal 170 de control, puede haber una o más señales de control independientes para cada neurona). De este modo, puede controlarse la reacción de las neuronas 103 a las señales desde la matriz 150 de sinapsis.

Esta disposición neurosináptica cuadrada es un ejemplo de un hardware de red neuronal de impulsos como el que se describe en el documento US 63/107.498, que utiliza disposiciones configurables de neuronas de impulso, y sinapsis, conectadas usando una estructura de interconexión programable que facilita la implementación de cualquier topología de conexión arbitraria. En el mismo se describió una disposición neurosináptica eficiente que usa componentes distribuidos que se compone de  $n$  entradas que activan  $n^2$  sinapsis y  $n$  neuronas, es decir el sistema incorpora la presencia de sinapsis electrónicas en las uniones de la disposición, mientras que la periferia de la disposición incluye filas de los circuitos neuronales, que imitan la acción del soma y el cono axónico de las neuronas biológicas.

Una disposición de este tipo puede dividirse y mapearse en una disposición de segmentos, cada uno de los cuales contiene una red programable de neuronas de impulso.

Las figuras 2A-C muestran representaciones esquemáticas de organización de la disposición neurosináptica reconfigurable que comprende  $n^2$  elementos sinápticos y  $n$  neuronas. En la figura 2A, las unidades organizativas primarias son bloques 201 de peso y bloques 202 de salida que reúnen  $1 \times m$  elementos sinápticos y  $m$  neuronas, respectivamente.

La figura 2B muestra un solo bloque 201 de peso con líneas de entrada. Para soportar un mayor abanico de entrada, las salidas sinápticas de múltiples elementos sinápticos distribuidos en múltiples columnas de sinapsis pueden dirigirse a una sola neurona, aumentando de manera efectiva el abanico de entrada potencial de una neurona a  $n \times m$  entradas únicas. Esto se logra reuniendo las sinapsis en bloques de peso. La organización de la disposición neurosináptica reconfigurable describe un espacio de posibles arquitecturas controladas por el número  $m$  de elementos sinápticos dentro de un bloque de peso, con  $m \in [1, n]$ , donde el abanico de entrada máximo en una neurona es  $n \times m$ , el número de puertos de entrada es  $n \times m$ , y el número de neuronas difundidas a través de un solo impulso de entrada es  $n/m$ . Aumentar el número de  $W_{n,n/m}$  bloques de peso proporciona una mayor utilización de sinapsis cuando una entrada no se comparte entre todas las neuronas mapeadas.

El número de columnas de sinapsis (columnas de elementos sinápticos) en un bloque 201 de peso es por tanto igual a  $m$  tal como se muestra en la figura 2B, y  $n/m$  es el número total de bloques 201 de peso en cada fila de la disposición 150 neurosináptica, siendo  $n$  el número total de columnas en la matriz 150 de sinapsis (suponiendo una disposición cuadrada). Un bloque de peso puede comprender múltiples filas de elementos sinápticos, así como múltiples columnas de elementos sinápticos, sin embargo el número de filas en cada bloque 201 de peso es de 1 en este ejemplo. Cada bloque de peso se indica como  $W_{ij}$ , con  $i \in [1, n]$  y  $j \in [1, n/m]$ .

Obsérvese que aunque la figura 2A representa una sola línea de entrada para cada fila de bloques 201 de peso por motivos de simplicidad, cada fila de bloques 201 de peso recibe realmente múltiples entradas, por ejemplo un bloque 201 de peso que comprende  $m$  elementos sinápticos puede tener  $m$  entradas, tal como se muestra en la figura 2B, de modo que los bloques 201 de peso en cada fila se conectan a  $m$  líneas de entrada. De manera similar, la figura 2A representa una sola línea para transmitir las salidas de cada columna de bloques 201 de peso por motivos de simplicidad. Sin embargo, cada columna de bloques 201 de peso genera realmente múltiples salidas, por ejemplo un bloque 201 de peso que comprende  $m$  elementos sinápticos puede generar  $m$  salidas, tal como se muestra en la figura 2B, de modo que los bloques 201 de peso en cada columna se conectan a  $m$  líneas de salida.

La figura 2C muestra un solo bloque 202 de salida. El número de neuronas en cada bloque 202 de salida activadas

por una columna de los bloques 201 de peso es de  $m$ . Cada bloque de salida se indica como  $O_i$  con  $i \in [1, n/m]$ .

La disposición neurosináptica reconfigurable cuadrada puede dividirse por tanto en  $n/m$  bloques de salida de tamaño  $m$ , donde  $m \leq n$ . Por tanto, se forma una disposición neurosináptica que puede separarse que tiene  $n$  puertos de entrada únicos y  $n/m$  bloques de salida que comprenden  $m$  neuronas.

De este modo se aumenta el número total de entradas únicas y puede usarse o bien para aumentar el abanico de entrada a neuronas individuales y/o bien con el fin de proporcionar más granularidad en la distribución de impulsos. Las conexiones de retroalimentación se ajustan a escala fácilmente con la disposición para un mapeo eficiente de redes recurrentes y multicapa. El número de bloques de peso  $n/m$  por fila designa por tanto la división interna de la disposición neurosináptica  $n^2$ ; colectivamente, el número de entradas únicas se aumenta a  $n \times m$ . Si  $m = n$ , cada sinapsis tiene su propio puerto de entrada dedicado, mientras que para  $m = 1$ , se obtiene la disposición original  $n^2$ , por lo que una entrada se comparte a través de todas las neuronas.

En la figura 2B se representa la estructura de un solo bloque de peso  $w_{ij}$  para la organización de la disposición neurosináptica reconfigurable.

Cada uno de los  $W_{n,n/m}$  bloques de peso sigue la organización de la disposición, es decir un conjunto de  $m$  entradas, indicadas como  $in^i$  con  $i \in [1, n]$ , difunde a lo largo de una fila de bloques de peso, es decir, una sola entrada  $in_x^i$  con  $x \in [1, m]$  se distribuye a  $n/m$  neuronas directamente.

El  $x$ -ésimo elemento sináptico (con  $x \in [1, m]$ ) del bloque de peso  $w_{ij}$  (con  $i \in [1, n]$  y  $j \in [1, n/m]$ ) se indica como  $w_{ij}^x$ . Un conjunto de  $m$  entradas, indicadas  $in^i$  con  $i \in [1, n]$ , difunde a lo largo de la  $i$ -ésima fila de elementos sinápticos; es decir, una sola entrada  $in_x^i$  con  $x \in [1, m]$  en el conjunto de  $m$  entradas se distribuye a  $n/m$  neuronas.

En la figura 2C se muestra la estructura de un bloque de salida  $O_j$  para la organización de la disposición neurosináptica reconfigurable.

Una columna de bloques 201 de peso comprende múltiples columnas de sinapsis que pueden conectarse a un bloque 202 de salida y pasa las salidas 252 desde las columnas de sinapsis a un decodificador 250 de columnas/neuronas. El decodificador 250 de columnas/neuronas incluye un circuito 260 de conmutación para conectar las salidas 252 de columna de sinapsis a uno o más conjuntos de neuronas 254 del bloque 202 de salida, estando controlado el circuito 260 de conmutación por una o más señales 251 de selección. La señal de conmutación que procede del decodificador 250 de columnas/neuronas se ha conmutado a una o más neuronas 254

indicadas por  $N_x^j$  con  $x \in [1, m]$  para el bloque de salida  $O_j$  con  $j \in [1, n/m]$ , lo que permite la conexión de las salidas 252 de columna de sinapsis a diferentes combinaciones de neuronas o conjuntos de neuronas.

La señal 251 de selección puede usarse para seleccionar cómo el decodificador 250 de columnas/neuronas conmuta las salidas 252 de columna de sinapsis desde un conjunto de neuronas 254 a otro conjunto de neuronas 254 del bloque 202 de salida. Pueden usarse una o más señales 253 de control de neuronas para determinar las propiedades de funcionamiento de los conjuntos de neuronas 254, o de las neuronas individuales.

En la figura 2D se muestra una realización a modo de ejemplo de posibles configuraciones del circuito 260 de conmutación de neuronas en el decodificador 250 de columnas/neuronas cuando cuatro columnas 270 de sinapsis sirven como entrada al decodificador 250 de columnas/neuronas. En este ejemplo, el decodificador 250 de columnas/neuronas se dispone para dotar a cada grupo de cuatro columnas 270 de sinapsis adyacentes de cuatro posibles configuraciones 280 de neuronas de salida tal como se muestra, es decir las cuatro columnas de sinapsis pueden tener sus salidas dirigidas selectivamente a cuatro conjuntos diferentes de neuronas 254 del bloque 202 de salida. El decodificador 250 de columnas/neuronas puede conectarse con diferentes números de columnas 270 de sinapsis y neuronas 254 de salida, y el número de columnas de sinapsis y neuronas de salida puede ser diferente, y el número y la disposición de las configuraciones 280 de neuronas de salida pueden ser diferentes.

A continuación se comentan las cuatro configuraciones 280 del circuito 260 de conmutación de neuronas. En una primera configuración, las cuatro columnas 270 de sinapsis sirven como entrada respectiva a cuatro neuronas de salida a través de trayectorias 261 de señales de conmutación. En una segunda configuración, dos pares de las cuatro columnas 270 de sinapsis sirven como entrada a dos conjuntos de neuronas de salida a través de trayectorias 261 de señales de conmutación, es decir las columnas de sinapsis  $i_0$  e  $i_1$  sirven a la neurona de salida  $o_0$  y las columnas de sinapsis  $i_2$  e  $i_3$  sirven a la neurona de salida  $o_2$ . En una tercera configuración, tres columnas de sinapsis  $i_0$ ,  $i_1$  e  $i_2$  sirven a un conjunto de neuronas de salida  $o_0$ , y la columna de sinapsis restante  $i_3$  sirve a la neurona de salida  $o_3$  a través de las trayectorias 261 de señales de conmutación. En la cuarta configuración, las cuatro columnas de sinapsis sirven a un solo conjunto de neuronas de salida  $o_0$  a través de trayectorias 261 de señales de conmutación.

Las trayectorias 261 de señales de conmutación pueden comprender trayectorias de conducción implementadas en el circuito lógico del circuito 260 de conmutación, y pueden conmutarse entre diferentes configuraciones en una implementación particular, por ejemplo, mediante puertas de transistor configurables o fijas. La conexión particular de las trayectorias 261 de señales de conmutación puede establecerse en la fabricación inicial de la disposición neurosináptica, por ejemplo en la fábrica, y/o puede establecerse sobre la marcha a través de programación u otra configuración del decodificador 250 de columnas/neuronas y/o el circuito 260 de conmutación de neuronas. El circuito 260 de conmutación de neuronas, por ejemplo, puede reconfigurar dinámicamente las trayectorias 261 de señales de conmutación durante el funcionamiento de la disposición neurosináptica. Tal como se ha mencionado, pueden usarse una o más señales 251 de selección para seleccionar cómo conmuta el decodificador 250 de columnas/neuronas desde un conjunto de neuronas 254 a otro conjunto de neuronas 254. La reconfiguración dinámica puede basarse en una determinada metodología de mapeo que incorpora una división y un mapeo de segmentos activados por restricciones, de los cuales se describirán adicionalmente a continuación realizaciones a modo de ejemplo. Por ejemplo, la reconfiguración dinámica puede realizarse basándose en la relación señal-ruido  $S/N$  de un segmento particular en la disposición neurosináptica. Esto se explicará adicionalmente más adelante.

Por tanto, el circuito 260 de conmutación de neuronas determina qué columnas 270 de sinapsis envían salidas a qué conjuntos 280 de neuronas de salida. Dicho de otro modo, el circuito 260 de conmutación de neuronas implementa por tanto (como resultado de la configuración actualmente activa del circuito 260 de conmutación de neuronas) una segmentación particular de la disposición neurosináptica.

Para soportar un mayor abanico de entrada, los bloques de peso  $W_{n,n/m}$  pueden organizarse como una estructura intercalada para facilitar la conmutación/combinación de  $m$  columnas de sinapsis adyacentes. La configuración para cada conjunto de conmutación de  $n/m$  neuronas puede controlarse independientemente para una mayor flexibilidad de mapeo. Para el enrutamiento sensible a la latencia de impulsos recurrentes o conexiones entre capas, pueden considerarse varias opciones, por ejemplo ofrecer una trayectoria directa entre la neurona de salida  $i$  y las entradas seleccionadas, utilizar organización intercalada de grupos o conjuntos de neuronas, de manera que una entrada pueda enrutarse rápidamente a cualquier grupo, proporcionar una inyección flexible de impulsos en la disposición, donde la salida de neuronas puede difundir a grupos individuales o múltiples de manera automática.

La segmentación de la disposición permite la uniformidad en lo que se refiere a la robustez y la fiabilidad del rendimiento, es decir, la segmentación permite (i) una distribución de energía robusta (pueden colocarse buses metálicos anchos verticales (para una resistencia menor) en los límites de fila de sinapsis para aliviar los problemas de caída de IR), (ii) si una interfaz para el acceso a los circuitos de retención de peso requiere un reloj, la segmentación crea la jerarquía necesaria para una estrategia de rechazo de reloj, (iii) la segmentación mejora la integridad de la señal: en lugar de una sola presinapsis con una gran fuerza de activación, la segmentación permite un enfoque distribuido, ya que las señales pueden rechazarse/amplificarse después de cada segmento, (iv) la segmentación aumenta la capacidad de observación dentro de la disposición y simplifica los procesos de verificación.

La figura 3 muestra esquemáticamente cómo dividir una disposición reconfigurable a modo de ejemplo (tal como se muestra por ejemplo en la figura 2) en múltiples segmentos/subdisposiciones.

Una primera subdisposición 301 comprende bloques de peso  $W_{ij}$  con  $i \in [1, n]$  y  $j \in [1, x]$ . Una segunda subdisposición 302 comprende bloques de peso  $W_{ij}$  con  $i \in [1, n]$  y  $j \in [x + 1, y]$ . Una subdisposición 303 final comprende bloques de peso  $W_{ij}$  con  $i \in [1, n]$  y  $j \in [z + 1, n/m]$ . En este caso,  $x, y, z$  son números naturales y  $x \leq y \leq z \leq n/m$ . Pueden existir más subdisposiciones. Un número  $n$  de líneas 304 de entrada alimentan los diferentes bloques de peso. Cada uno de los bloques de peso  $w_{ij}$  está conectado a un bloque de salida  $O_j$ . Cada subdisposición puede comprender varios bloques de salida segmentados como bloques 305, 306, 307 de salida de las subdisposiciones 301, 302, 303 respectivamente. Cuando el potencial de membrana de una neurona particular  $N_j$  (que está relacionado con la carga eléctrica de la membrana de la neurona) alcanza un valor específico, la neurona se activa. Esta señal eléctrica se representa como un impulso o bien de tensión o bien de corriente y se envía a otra parte de la SNN a través de una línea 309 de salida de impulsos.

Cada uno de los bloques de salida puede controlarse en el sentido de que, por ejemplo, el periodo de acumulación y/o la constante de integración de una(s) neurona(s) particular(es) puede controlarse a través de señales 308 de control. De este modo, puede controlarse la reacción de las neuronas a las señales desde los bloques de peso en las subdisposiciones 301, 302, 303.

Tal como se mencionó anteriormente, las operaciones elementales requeridas por una SNN se realizan de manera muy eficiente mediante circuitos electrónicos analógicos, sin embargo, el desajuste de fabricación induce distorsiones en sus propiedades funcionales. Especialmente en geometrías de proceso más pequeñas y corrientes de funcionamiento más bajas, estos circuitos son cada vez más sensibles a los efectos cuánticos y al ruido externo, lo que reduce efectivamente la relación señal-ruido y limita el rendimiento del procesamiento. El impacto de estos hechos no ideales aumenta en el caso de las disposiciones grandes, donde el controlador, el sesgo, y las neuronas y las sinapsis se comparten por un mayor número de dispositivos, a lo largo de interconexiones más largas. La división de disposiciones grandes en segmentos más pequeños (subdisposiciones), cada uno con su propio circuito

requerido, permite la uniformidad del rendimiento y mitiga estos hechos no ideales si la metodología de mapeo incorpora una división y un mapeo de segmentos activados por restricciones.

5 En una realización, los segmentos pueden organizarse basándose en las características de la función de procesamiento de la señal objetivo (por ejemplo ganancia, filtrado, multiplicación, adición), o constantes de frecuencia/tiempo de funcionamiento.

10 En otra realización, pueden configurarse diferentes regiones/segmentos de la disposición neurosináptica basándose en el caso de aplicación objetivo, el número de canales en la cadena/trayectoria de señal de preprocesamiento, la banda de frecuencia de interés, la complejidad y las características del conjunto de rasgos, o la relación señal-ruido objetivo del sistema.

15 En otra realización, los segmentos pueden organizarse basándose en definiciones de inspiración biofísica/química (por ejemplo, evaluación espacio-temporal (a corto y largo plazo), artefactos de envejecimiento, costes de comunicación cerebral y regiones/capas).

20 En otra realización, los segmentos pueden definirse basándose en una regla de aprendizaje, una tasa de aprendizaje y un mecanismo basado en coste asociado o mecanismos de (post-)plasticidad, o basándose en definiciones predefinidas basadas en el rendimiento de factor de mérito.

25 En otra realización, los segmentos dentro de la disposición neurosináptica pueden configurarse basándose en los requisitos de regulación homeostática (local y global), o en las capacidades de propagación/calibración de errores y las limitaciones del mecanismo, o en el mecanismo limitador, como la razón de excitación/inhibición, la interferencia o el mecanismo de saturación de red.

30 Las figuras 4A-D muestran esquemáticamente una organización a modo de ejemplo de la disposición 400 neurosináptica reconfigurable: en la figura 4A,  $n^2$  sinapsis 401 y  $n$  neuronas 402, en la figura 4B dividiendo la disposición 400 en la figura 4A en dos segmentos, cada uno con  $n^2/2$  sinapsis y  $2 \times n$  neuronas, en la figura 4C dividiendo una disposición 400 en cuatro segmentos, cada uno con  $n^2/4$  sinapsis 401 y  $4 \times n$  neuronas 402, y en la figura 4D dividiendo una disposición 400 en ocho segmentos, cada uno con  $n^2/8$  sinapsis 401 y  $8 \times n$  neuronas 402.

35 Los segmentos neurosinápticos pueden componerse de  $n$  entradas que activan  $n^2$  sinapsis y  $n$  neuronas (como en las figuras 1-3, figura 4A). Al reducir las distancias dendríticas/de disposición, los complejos y compartimentos acortan los retardos y disminuyen el ruido, es decir para reducir el consumo de energía y aumentar el recuento de neuronas por conexión sináptica para la misma área, la disposición en la figura 4A puede subdividirse en dos segmentos (figura 4B), cada uno con  $n^2/2$  sinapsis y  $n$  neuronas; por tanto, el recuento sináptico total a través de dos segmentos permanece igual, mientras que el número de neuronas se duplica. De manera similar, al dividir la disposición en la figura 4A en cuatro segmentos (figura 4C), cada uno con  $n^2/4$  sinapsis puede cuadruplicarse el recuento total de neuronas, o en ocho segmentos (figura 4D), cada uno con  $n^2/8$  sinapsis puede aumentar el recuento de neuronas en ocho veces, respectivamente.

45 La capacidad de información de una disposición depende de la relación señal (S) – ruido (N)  $S/N$  y aumenta como  $\log_2(1 + S/N)$ . Sin embargo, el coste energético de pasar la señal a través de la disposición aumenta como  $\sqrt{S/N}$ . Por tanto, a medida que aumenta  $\sqrt{S/N}$ , la eficiencia de la disposición disminuye, es decir, todos los componentes intentan transmitir la misma señal, aunque la información por coste fijo unitario aumenta. El punto óptimo se produce cuando estas dos tendencias en competencia se equilibran. Como consecuencia, una mayor razón de coste fijo con respecto al coste de señalización da como resultado una mayor disposición óptima.

50 El tamaño óptimo de la disposición también depende de los costes en otras partes del circuito con vistas a distribuir la relación S/N asignada entre los componentes para maximizar el rendimiento en todo el sistema.

55 La figura 5 muestra una representación a modo de ejemplo de una organización de la disposición 500 neurosináptica reconfigurable: cada uno de los segmentos (subdisposiciones) 501 puede tener un recuento sináptico diferente por neurona. A nivel del segmento 501, para beneficiarse de una precisión aumentada de las salidas de sensor y/o de unas capacidades aumentadas de selección y extracción de características de discriminación, pueden definirse segmentos especializados, cada uno dedicado a procesar una sola modalidad.

60 Cada uno de los segmentos puede tener un recuento sináptico diferente por neurona, donde  $m = [1, n]$ , definido según la descripción de esta invención.

La relación  $S/N$  de una entrada afecta profundamente al tamaño óptimo de la disposición dado que el ruido de entrada impone un límite al que debe aproximarse la relación  $S/N$  de la disposición. Esto reduce la eficacia de una disposición grande con una relación  $S/N$  de entrada baja y el tamaño de la disposición más eficiente, es decir, dado que una entrada con una relación  $S/N$  baja contiene menos información, y una disposición más pequeña tiene una

menor capacidad de información, la disposición óptima adapta su capacidad a su entrada. La adaptación del tamaño de la disposición a la relación  $S/N$  de entrada sigue el principio de simorfosis, para adaptar las capacidades dentro de un sistema para obtener una solución óptima. Véase E.R. Weibel, "Symmorphosis: On form and function in shaping life", Cambridge, MA: Harvard University Press, 2000.

Para economizar en el número de neuronas (pero impedir que un componente realice múltiples tareas de manera subóptima) dentro de un segmento, deben potenciarse las capacidades de cálculo del/de los circuito(s) dentro de una neurona, es decir, dedicar más definiciones de neurorreceptores (AMPA, NMDA, GABA) o basadas en la conductancia neuronal (Na, Ca, K) a la modalidad particular de la neurona y, de ese modo, mejorar la sensibilidad y la relación señal-ruido de las neuronas. Tales neuronas pueden usar circuitos internos (basados en neurorreceptores o en conductancia) para realizar funciones que habitualmente requieren un circuito de varias neuronas.

Un neuromodulador puede difundirse ampliamente y aun así actuar de manera local y específica, afectando sólo a las neuronas con un neurorreceptor apropiado. El alcance de un neuromodulador se potencia adicionalmente porque su receptor se diversifica en múltiples subtipos que se acoplan a diferentes redes de señalización intrasegmento. Como consecuencia, un pequeño ajuste dirigido puede resintonizar y reconfigurar toda una red, lo que permite una modularidad de procesamiento eficiente y la reprogramación de tareas si se requiere otra funcionalidad de procesamiento de señales. Los neurorreceptores alteran la dependencia de la tensión de las corrientes y, posteriormente, afectan a la forma de las curvas de activación e inactivación.

Cada receptor representa un canal de (trans)conductancia múltiple que modela las características no lineales asociadas. Los receptores de NMDA pueden proporcionar modificaciones dependientes de la actividad del peso sináptico, mientras que los receptores de AMPA pueden permitir que una corriente sináptica rápida active el soma. Las modificaciones de receptor rápidas pueden actuar como un mecanismo de control para los cambios dependientes de la actividad en la eficacia sináptica. Si los grupos de impulsos dendríticos (ráfagas) son suficientes para superar el umbral de una neurona particular, el axón generará potenciales de acción; el impulso resultante se retroalimenta a la dendrita y, junto con las señales del soma, se multiplica y se añade a las señales del receptor de NMDA para, posteriormente, generar el control del peso.

En general, las señales de retroalimentación pueden ser un subconjunto de las disponibles y pueden actuar sobre un subconjunto de los parámetros del sistema (es decir, preservar las propiedades intrínsecas específicas cuando se compensan las perturbaciones). Véase T. O'Leary, A.H. Williams, A. France, E. Marder, "Cell types, network homeostasis, and pathological compensation from a biologically plausible ion channel expression model", *Neuron*, vol. 82, págs. 809-821, 2014.

Cuando se producen perturbaciones en la red, los elementos computacionales neuronales generan un impulso basándose en el signo, la intensidad y la fase de la perturbación. Si se distorsiona o altera una respuesta de resonancia de un grupo neuronal, cada neurona ajusta las fases de los impulsos. Posteriormente, los elementos computacionales neuronales pueden realizar cálculos dependientes del tiempo.

En niveles suficientemente altos de heterogeneidad (por ejemplo, heterogeneidades aumentadas en parámetros neurosinápticos como la capacitancia de integrador, el umbral de activación, el periodo refractario), la red demuestra un patrón espacialmente desordenado de actividad neuronal. Si los perfiles de actividad de red son estrechos, aumenta la necesidad de homogeneidad en la distribución de los estímulos espaciales codificados por la red.

La figura 6A muestra una representación a modo de ejemplo de un procesador 600 neuronal de impulsos que incorpora uno o más codificadores 606 de datos en impulsos, una disposición 607 neurosináptica multisegmentada de señal analógica/mixta, y decodificadores 611 de impulsos.

Se usan datos como entrada para módulos de entrada tales como una AER 601, interfaces tales como por ejemplo una SPI 602, I<sup>2</sup>C 603 y/o I/O (entrada/salida) 604 y/o un preprocesador DSP/CPU 605. La salida de estos módulos de entrada se envía a uno o más codificadores 606 de datos en impulsos. El uno o más codificadores 606 de datos en impulsos codifican los datos en impulsos que se usan como entrada para la disposición 607 neurosináptica de señal analógica/mixta.

Los impulsos que se emiten por la disposición 607 neurosináptica pueden decodificarse por un decodificador 611 de impulsos. Puede usarse un módulo 614 de búsqueda de patrones para interpretar patrones particulares dentro de la señal de salida de impulsos. Los impulsos decodificados pueden enviarse entonces a una o más interfaces de salida y/o lógica 614 de interrupción. Un módulo 614 de configuración y control permite la configuración y el control de las diferentes partes del procesador neuronal de impulsos, por ejemplo, estableciendo pesos y parámetros de configuración de los elementos sinápticos y neuronas dentro de la disposición 607 neurosináptica.

La disposición 607 neurosináptica puede segmentarse según cualquiera de las realizaciones mencionadas anteriormente. Las neuronas dentro de un segmento particular pueden conducir pasivamente, como cambios graduales (analógicos) en el potencial eléctrico, y por tanto dependen únicamente de cálculos analógicos, que son directos y energéticamente eficientes. Los potenciales de acción breves, intensos y de alto consumo de energía se

reservan para la señalización a mayor distancia, es decir, la interconexión dentro de la disposición neurosináptica multisegmentada de señal analógica/mixta sigue el paradigma del cerebro biológico, es decir, una conectividad densa en/entre segmentos proximales, y menos densa entre segmentos distantes. Por tanto, pueden existir tanto interconexiones 610 de largo alcance como interconexiones 609 de corto alcance entre las neuronas dentro de la disposición 607 neurosináptica segmentada.

La figura 6B muestra una representación a modo de ejemplo de una estructura de interconexión.

De nuevo, la interconexión 610 de largo alcance y la interconexión 609 de corto alcance pueden estar presentes entre los diferentes segmentos 608 dentro de la disposición 607 neurosináptica.

Una estructura de interconexión (doble) de este tipo se rige por objetivos de diseño característicos: para la comunicación 609 de corto alcance se prioriza la baja latencia entre segmentos proximales (obtenida con alta densidad de conexión), mientras que con la comunicación 610 de largo alcance se busca un alto rendimiento entre segmentos distantes. Aunque el tejido de interconexión de segmento a segmento puede implementarse con lógica CMOS digital, las propias disposiciones neurosinápticas pueden implementarse mediante el uso, por ejemplo, de tecnologías de señal analógica/mixta, digitales o de memoria no volátil.

En su forma más simple, el potencial de membrana neuronal puede implementarse en el dominio analógico como una tensión a través de un condensador o como una variable multibit almacenada en capturadores digitales. Alternativamente, las neuronas digitales pueden implementarse usando circuitos lógicos de CMOS, tales como sumadores, multiplicadores y contadores. Desde el punto de vista del diseño de circuitos, las sinapsis modulan los impulsos de entrada y los transforman en una carga que, como consecuencia, crea corrientes postsinápticas que se integran en la capacitancia de membrana de la neurona postsináptica. La implementación de sinapsis de silicio normalmente sigue la metodología de señales mixtas.

La coordinación espaciotemporal de los diferentes tipos de transmisión sináptica es fundamental para la inferencia de señales sensoriales y el rendimiento de aprendizaje asociado. Para una entrada fuerte, la red se organiza en un estado dinámico donde la activación neuronal se impulsa principalmente por las fluctuaciones de entrada, de manera similar a un estado asincrónico-irregular, es decir, las correlaciones cruzadas temporales y de recuento de impulsos por pares se acercan a cero. La depresión heterosináptica entre todas las sinapsis de entrada genera secuencias de actividad estables dentro de la red.

Las neuronas generan patrones temporizados; debido a la interacción entre el retardo de conductancia (y las reglas de plasticidad), la red utiliza un conjunto de grupos/segmentos neuronales, que forman secuencias de activación reproducibles y precisas. A medida que aumenta la conductancia (dando como resultado en consecuencia una excitación neta de la red), las tasas de activación pueden aumentar y volverse más uniformes con un coeficiente de variación menor. La red puede generar trenes de impulsos con relaciones de tiempo regulares (intervalos entre ráfagas y ciclos de trabajo) en diferentes segmentos.

La incorporación de un comportamiento sincronizado de baja frecuencia incluye una reorganización de la fase de modo que la fase óptima, es decir, la más excitable, se alinee con las características temporales de los eventos en el estímulo de entrada en curso. La secuencia de la corriente sináptica (es decir, hacia afuera, hacia adentro) disminuye la inestabilidad temporal en la generación de potenciales de acción en neuronas individuales y, en consecuencia, crea una red con mayor capacidad de control de la actividad.

La gran discrepancia en la eficiencia energética y el rendimiento cognitivo de los sistemas nerviosos biológicos y la computación convencional se ejemplifica profundamente con tareas relacionadas con interacciones en tiempo real con el entorno físico, en particular en presencia de una entrada sensorial no controlada o con ruido. Sin embargo, la red neuronal basada en eventos neuromórficos, debido a la capacidad de aprender mediante ejemplos, el paralelismo de la operación, la memoria asociativa, la optimización multifactorial, y la extensibilidad, es una opción inherente para sistemas cognitivos compactos y de baja energía que aprenden y se adaptan a los cambios en las estadísticas de las señales sensoriales complejas.

La presente arquitectura reconfigurable de múltiples segmentos para las redes neuromórficas permite diseños que ofrecen soluciones energéticamente eficientes para aplicaciones que van desde la detección de patrones de señales biomédicas (por ejemplo, clasificación de impulsos, detección de convulsiones, etc.), clasificación de imágenes (por ejemplo, dígitos escritos a mano), comandos de voz, y puede aplicarse en una amplia variedad de dispositivos, incluyendo sensores inteligentes o dispositivos ponibles en sistemas ciberfísicos e Internet de las cosas.

REIVINDICACIONES

1. Disposición (100) neurosináptica que comprende una pluralidad de neuronas de impulso, y una pluralidad de elementos sinápticos que interconectan las neuronas de impulso para formar una red neuronal de impulsos, en la que los elementos sinápticos y las neuronas de impulso están implementadas parcial o completamente en hardware;
- 5
- en la que cada elemento sináptico está dispuesto para recibir una señal de entrada sináptica desde al menos una de múltiples entradas y está adaptado para aplicar un peso a la señal de entrada sináptica para generar una señal de salida sináptica, pudiendo configurarse los elementos sinápticos para ajustar el peso aplicado por cada elemento sináptico;
- 10
- en la que cada una de las neuronas de impulso está dispuesta para recibir una o más de las señales de salida sinápticas desde uno o más de los elementos sinápticos, y está adaptada para generar una señal de salida de tren de impulsos espacio-temporal en respuesta a la una o más señales de salida sinápticas recibidas; y en la que
- 15
- la disposición (100) neurosináptica comprende bloques (201) de peso y bloques (202) de salida, comprendiendo cada bloque (201) de peso uno o más de los elementos sinápticos, y comprendiendo cada bloque (202) de salida una o más de las neuronas y un circuito (260) de conmutación de neuronas;
- 20
- en la que cada bloque de peso está comprendido en una columna de bloques de peso de la disposición neurosináptica;
- 25
- en la que cada bloque (202) de salida puede conectarse eléctricamente a un subconjunto de bloques (201) de peso, estando formado el subconjunto de bloques (201) de peso por una o más columnas de bloques de peso, a través del circuito (260) de conmutación de neuronas mediante la conexión de una salida de la una o más columnas de bloques de peso a uno o más conjuntos de neuronas del bloque de salida; y en la que el circuito (260) de conmutación de neuronas está configurado para conectar eléctricamente de manera selectiva al menos un elemento sináptico comprendido dentro del subconjunto de bloques (201) de peso al menos a una neurona comprendida dentro del bloque (202) de salida respectivo, para obtener una división de la disposición (100) neurosináptica en conjuntos de neuronas conectadas eléctricamente a elementos sinápticos seleccionados.
- 30
2. Disposición (100) neurosináptica según la reivindicación 1, en la que los elementos sinápticos comprendidos en un bloque (201) de peso particular se proporcionan dentro de la misma fila de la disposición (100) neurosináptica y/o en la que cada bloque (202) de salida está conectado a una columna de bloques (201) de peso.
- 35
3. Disposición (100) neurosináptica según la reivindicación 1 ó 2, en la que cada uno de los circuitos (260) de conmutación de neuronas comprende trayectorias (261) de señales de conmutación que comprenden cables conductores implementados en un circuito lógico del circuito (260) de conmutación de neuronas, en la que las trayectorias (261) de señales de conmutación están configuradas para poder conmutarse entre diferentes configuraciones, preferiblemente mediante el uso de puertas de transistor, en la que cada configuración determina qué al menos un elemento sináptico comprendido dentro del subconjunto de bloques (201) de peso está conectado eléctricamente a qué al menos una neurona comprendida dentro del bloque (202) de salida.
- 40
4. Disposición (100) neurosináptica según la reivindicación 3, en la que el circuito (260) de conmutación de neuronas está configurado para reconfigurar dinámicamente las trayectorias (261) de señales de conmutación, en la que la reconfiguración dinámica se basa en una metodología de mapeo que incorpora una división y segmentación activadas por restricciones de la disposición (100) neurosináptica, en la que la disposición (100) neurosináptica se divide y se segmenta en uno o más segmentos, teniendo cada segmento una relación señal (S) - ruido (N) S/N, una capacidad de información dada por  $\log_2(1 + S/N)$  y un
- 45
- coste energético de paso de una señal a través del segmento dado por  $\sqrt{S/N}$ , de tal manera que para cada segmento, la capacidad de información se equilibra con el coste energético de paso de una señal a través del segmento.
- 50
5. Disposición (100) neurosináptica según una cualquiera de las reivindicaciones anteriores, en la que la segmentación de la disposición (100) neurosináptica se realiza basándose en una o más reglas de aprendizaje, tasas de aprendizaje y/o mecanismos de (post-)plasticidad, de manera que al menos dos de las subdisposiciones neurosinápticas son distintas en lo que se refiere a la una o más reglas de aprendizaje, tasas de aprendizaje y/o mecanismos de (post-)plasticidad.
- 55
6. Disposición (100) neurosináptica según cualquiera de las reivindicaciones anteriores, en la que al menos
- 60
- 65

- 5 uno de los bloques (201) de peso está organizado como una estructura intercalada, tal como para facilitar la conmutación y/o combinación de elementos sinápticos dentro de la disposición (100) neurosináptica, y/o en la que para cada bloque (202) de salida, la conmutación puede controlarse de manera independiente, tal como para obtener una mayor flexibilidad de mapeo, y/o en la que al menos uno de los bloques (202) de salida está organizado como una estructura intercalada y/o en la que la salida de neuronas de cualquiera de los bloques (202) de salida difunde a una o múltiples subdisposiciones neurosinápticas.
7. Disposición (100) neurosináptica según una cualquiera de las reivindicaciones anteriores, en la que cada una de las neuronas dentro de uno de los bloques (202) de salida conduce pasivamente como cambios analógicos graduales en el potencial eléctrico.
- 10 8. Disposición (100) neurosináptica según una cualquiera de las reivindicaciones anteriores, en la que la disposición (100) neurosináptica está segmentada en subdisposiciones neurosinápticas que pueden separarse y tienen su propio circuito requerido, y/o están dispuestas para procesar, y optimizarse para, una sola modalidad.
- 15 9. Disposición (100) neurosináptica según la reivindicación 8, en la que existen interconexiones tanto de largo alcance como de corto alcance entre neuronas de diferentes subdisposiciones neurosinápticas, y en la que existe una conectividad más densa entre subdisposiciones neurosinápticas proximales que entre subdisposiciones neurosinápticas más distantes.
- 20 10. Disposición (100) neurosináptica según una cualquiera de las reivindicaciones anteriores, en la que al menos uno de los bloques (202) de salida puede controlarse en el sentido de que el periodo de acumulación y/o la constante de integración de una neurona particular dentro del al menos uno de los bloques (202) de salida puede controlarse a través de señales de control.
- 25 11. Disposición (100) neurosináptica según una cualquiera de las reivindicaciones anteriores, en la que un potencial de membrana neuronal de una de las neuronas se implementa en el dominio analógico como una tensión a través de un condensador o como una variable multibit almacenada en capturadores digitales; o en el dominio digital usando circuitos lógicos de CMOS.
- 30 12. Procesador neuronal de impulsos que comprende:
- 35 un codificador de datos en impulsos que codifica datos digitales o analógicos en impulsos;
- una disposición (100) neurosináptica según una cualquiera de las reivindicaciones anteriores, dispuesta para tomar los impulsos emitidos por el codificador de datos en impulsos como entrada y dispuesta para emitir un tren de impulsos espacio-temporal como resultado de la entrada; y
- 40 un decodificador de impulsos dispuesto para decodificar los trenes de impulsos espacio-temporales que se originan a partir de la disposición (100) neurosináptica.
13. Método para configurar una disposición (100) neurosináptica, en el que:
- 45 la disposición (100) neurosináptica comprende una pluralidad de neuronas de impulso, y una pluralidad de elementos sinápticos que interconectan las neuronas de impulso para formar una red neuronal de impulsos en el que los elementos sinápticos y las neuronas de impulso están implementadas parcial o completamente en hardware;
- 50 en el que cada elemento sináptico está adaptado para recibir una señal de entrada sináptica desde al menos una de múltiples entradas y aplicar un peso a la señal de entrada sináptica para generar una señal de salida sináptica, pudiendo configurarse los elementos sinápticos para ajustar el peso aplicado por cada elemento sináptico;
- 55 en el que cada una de las neuronas de impulso está adaptada para recibir una o más de las señales de salida sinápticas desde uno o más de los elementos sinápticos, y generar una señal de salida de tren de impulsos espacio-temporal en respuesta a la una o más señales de salida sinápticas recibidas; y
- 60 en el que el método comprende
- dividir la disposición (100) neurosináptica en bloques (201) de peso y bloques (202) de salida, comprendiendo cada bloque (201) de peso uno o más de los elementos sinápticos, y comprendiendo cada bloque (202) de salida una o más de las neuronas y un circuito (260) de conmutación de neuronas,
- 65 en el que cada bloque de peso está comprendido en una columna de bloques de peso de la disposición neurosináptica;

- 5 hacer que cada bloque (202) de salida pueda conectarse eléctricamente a un subconjunto de bloques (201) de peso, estando formado el subconjunto de bloques (201) de peso por una o más columnas de bloques de peso, a través del circuito (260) de conmutación de neuronas mediante la conexión de una salida de la una o más columnas de bloques de peso a uno o más conjuntos de neuronas del bloque de salida;
- 10 configurar el circuito (260) de conmutación de neuronas para conectar eléctricamente de manera selectiva al menos un elemento sináptico comprendido dentro del subconjunto de bloques (201) de peso al menos a una neurona comprendida dentro del bloque (202) de salida respectivo, para obtener una división de la disposición (100) neurosináptica en conjuntos de neuronas conectadas eléctricamente a elementos sinápticos seleccionados.
14. Método según la reivindicación 14, en el que los elementos sinápticos comprendidos en un bloque (201) de peso particular se proporcionan dentro de la misma fila de la disposición (100) neurosináptica.
15. Método según la reivindicación 13 ó 14, en el que cada uno de los circuitos (260) de conmutación de neuronas comprende trayectorias (261) de señales de conmutación que comprenden cables conductores implementados en un circuito lógico del circuito (260) de conmutación de neuronas, en el que las trayectorias (261) de señales de conmutación están configuradas para poder conmutarse entre diferentes configuraciones, preferiblemente mediante el uso de puertas de transistor, en el que cada configuración determina qué al menos un elemento sináptico comprendido dentro del subconjunto de bloques (201) de peso está conectado eléctricamente a qué al menos una neurona comprendida dentro del bloque (202) de salida.
- 20
- 25

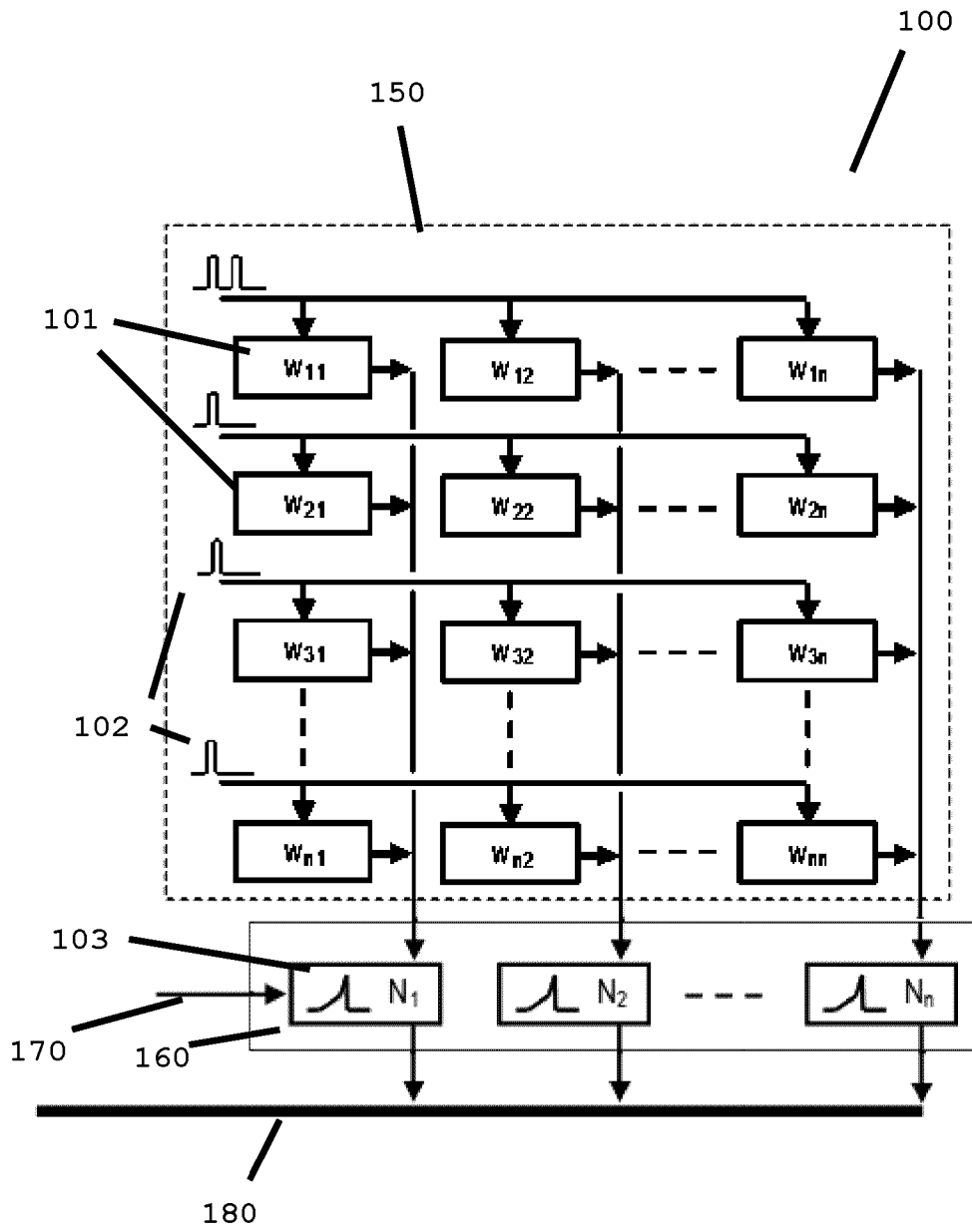


FIG. 1

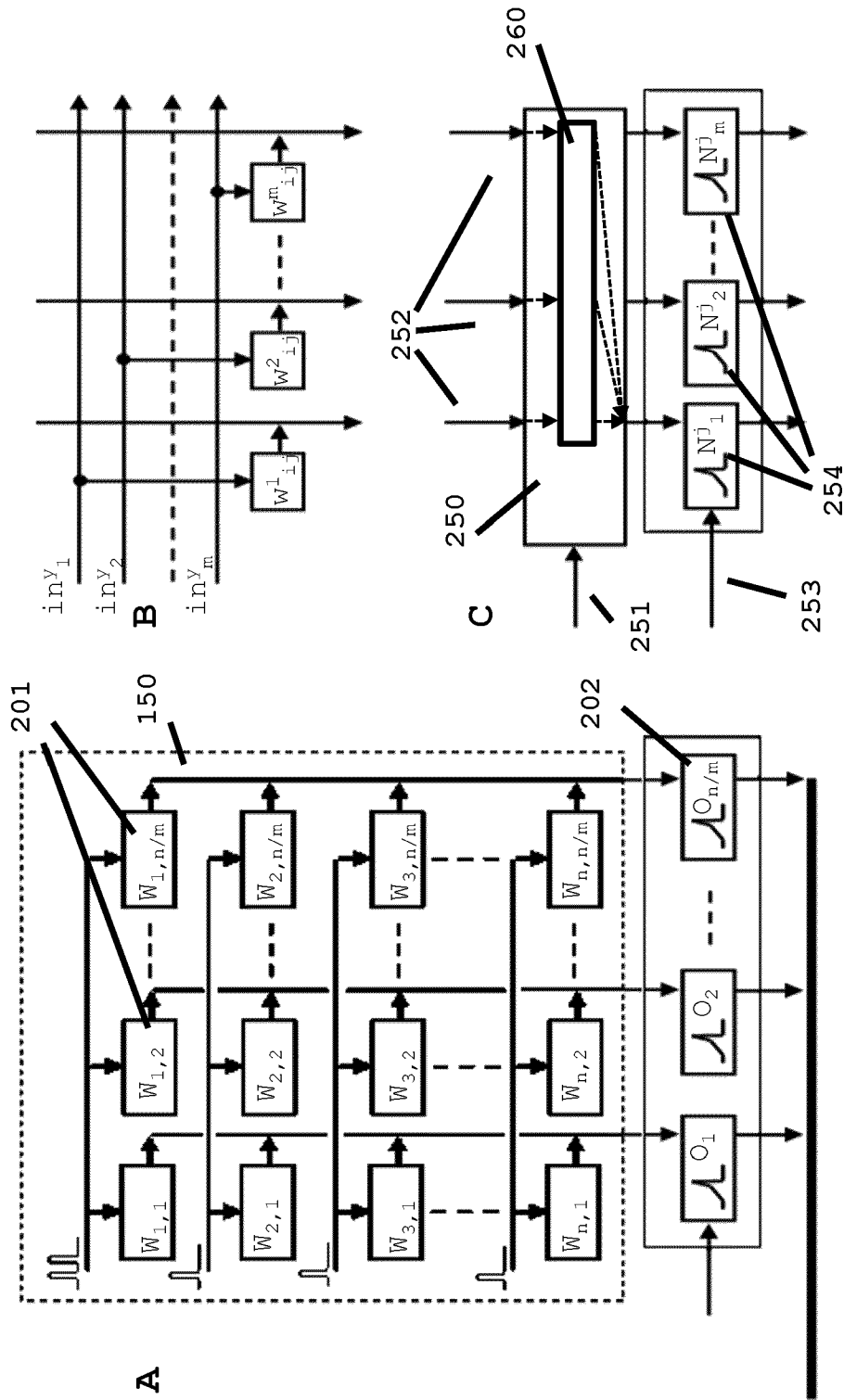


FIG. 2

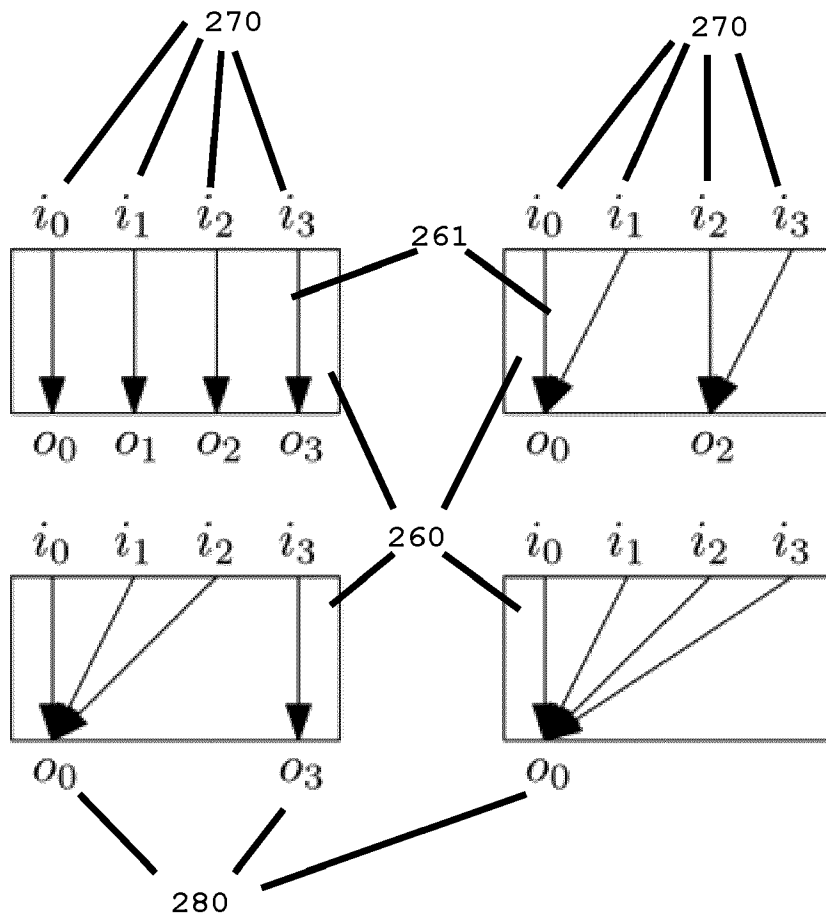


FIG. 2D

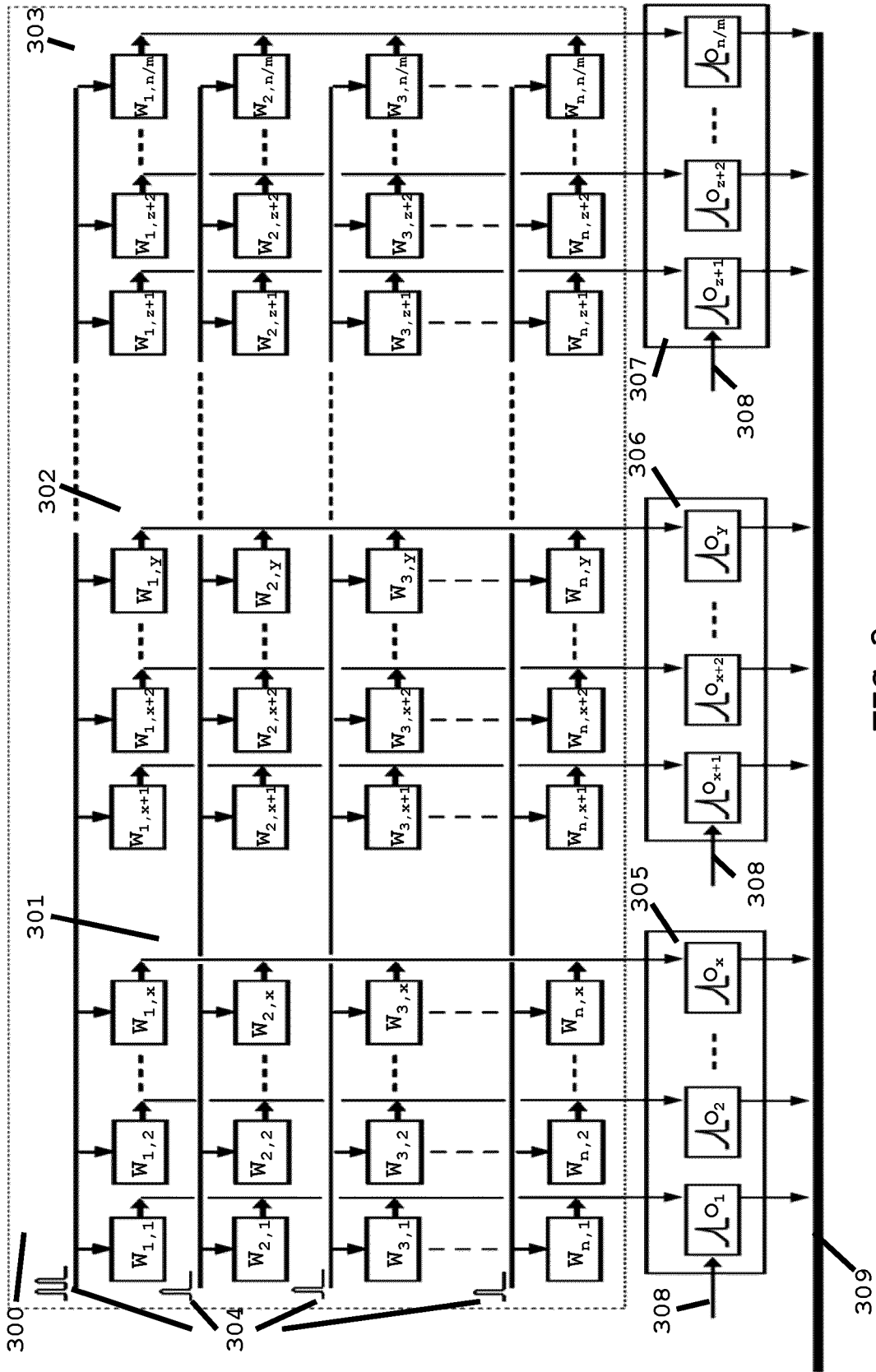


FIG. 3

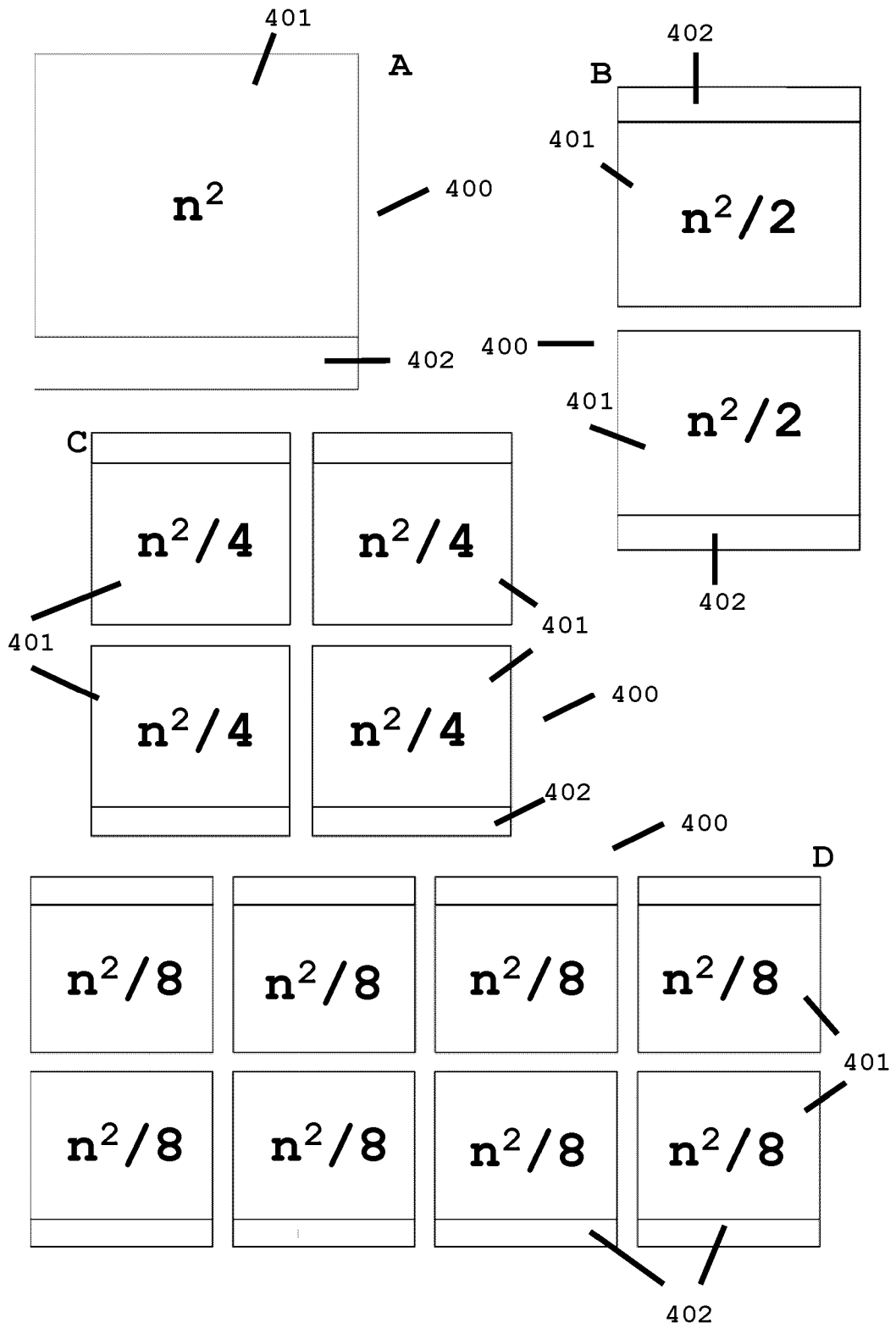


FIG. 4

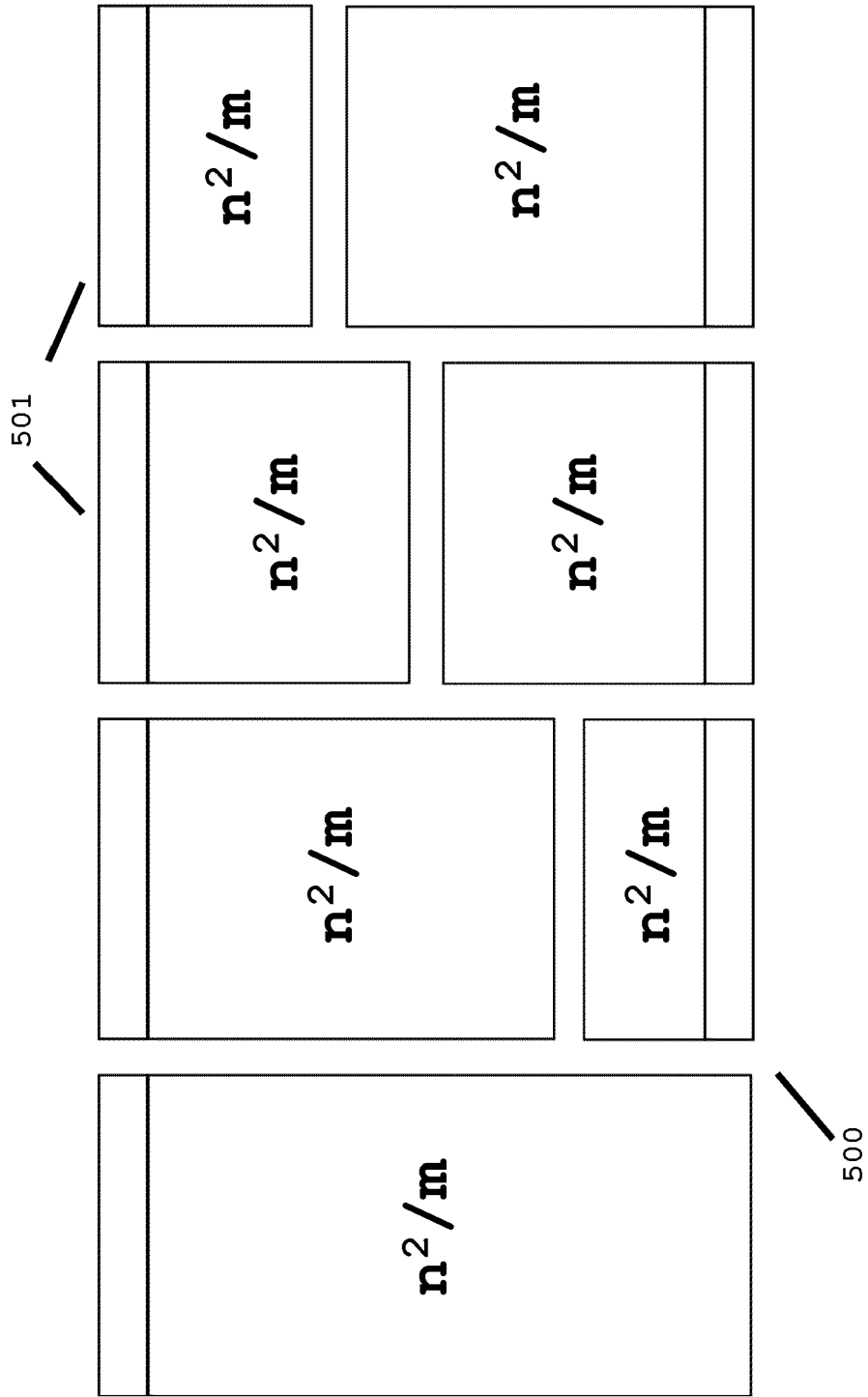


FIG. 5

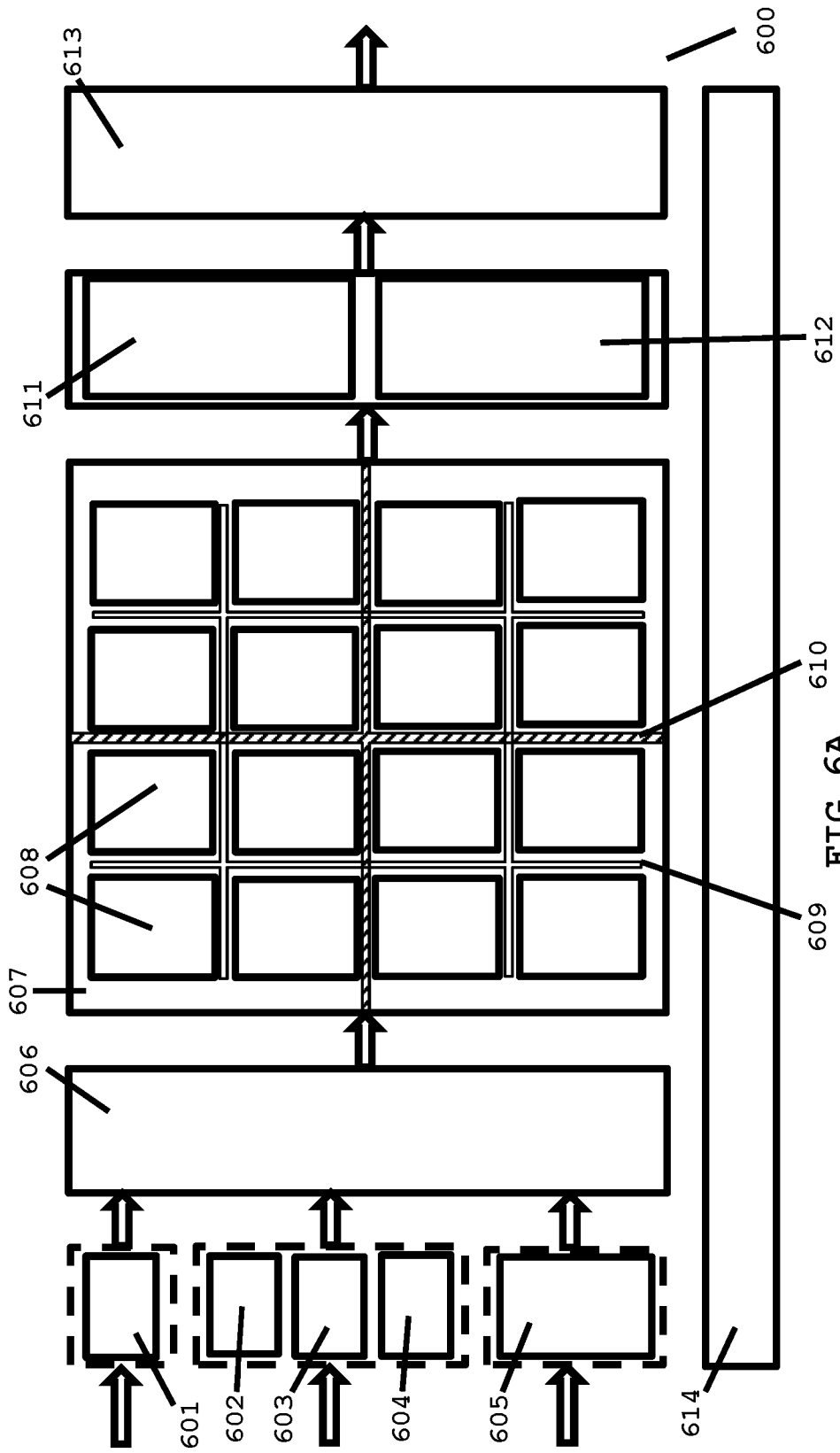


FIG. 6A

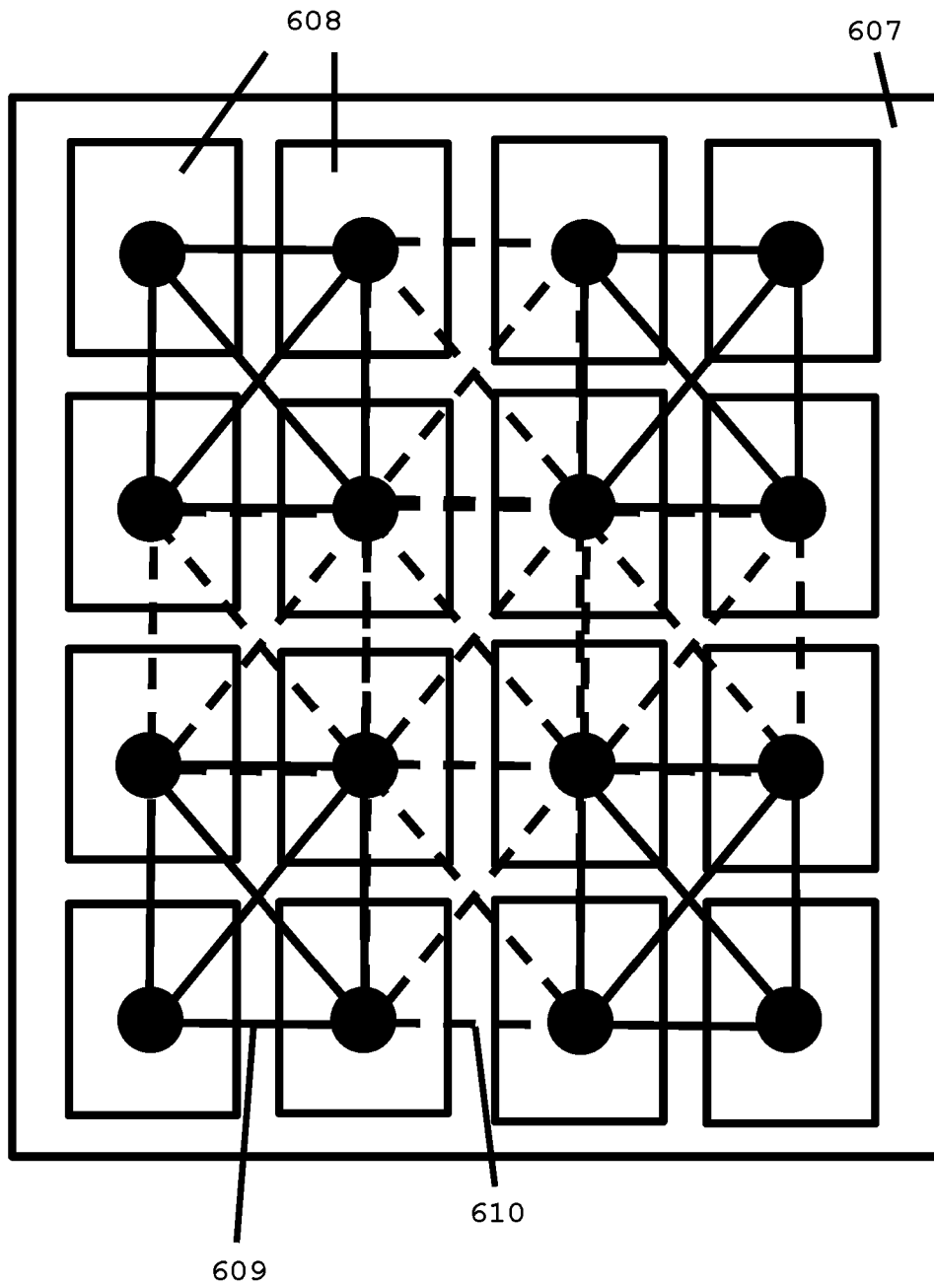


FIG. 6B