



(12) 发明专利

(10) 授权公告号 CN 103282919 B

(45) 授权公告日 2016. 02. 17

(21) 申请号 201180063280. 5

G11C 11/54(2006. 01)

(22) 申请日 2011. 10. 18

(56) 对比文件

(30) 优先权数据

12/982, 505 2010. 12. 30 US

W0 2009113993 A1, 2009. 09. 17, 权利要求 1-15, 说明书第 15-40 页, 及图 1-29.

(85) PCT国际申请进入国家阶段日

2013. 06. 28

US 2008/0162391 A1, 2008. 07. 03, 说明书第 34-37, 44-48 段.

(86) PCT国际申请的申请数据

PCT/EP2011/068183 2011. 10. 18

US 5212766 A, 1993. 05. 18, 全文.

US 6865434 B1, 2005. 03. 08, 全文.

(87) PCT国际申请的公布数据

W02012/089360 EN 2012. 07. 05

审查员 周正

(73) 专利权人 国际商业机器公司

地址 美国纽约

(72) 发明人 D·S·莫德哈 张立伦

R·K·蒙托耶

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 于静 张亚非

(51) Int. Cl.

G06N 3/063(2006. 01)

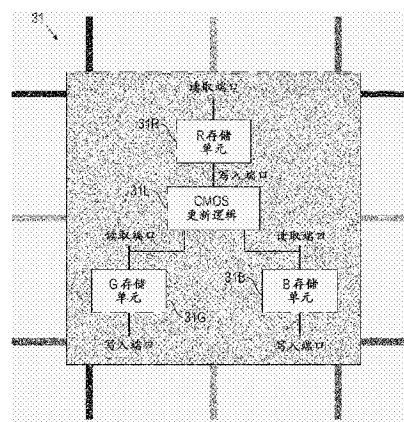
权利要求书3页 说明书7页 附图15页

(54) 发明名称

强化学习的电子突触

(57) 摘要

本发明具体实施例提供用以强化学习的电子突触装置。电子突触配置为互连前突触电子神经元及后突触电子神经元。电子突触包含配置为存储电子突触的状态以及存储用以更新电子突触的状态的元信息的存储器组件。电子突触还包含更新模块,其配置为基于元信息更新电子突触的状态,以响应针对强化学习的更新信号。更新模块配置为基于元信息更新电子突触的状态,以响应基于学习规则的针对强化学习的延迟更新信号。



1. 一种用于强化学习的电子突触系统,包含:
多个电子神经元;
交叉式数组,配置为互连所述多个电子神经元,该交叉式数组包含:
多个轴突及多个树突,使得该轴突与该树突彼此横贯;以及
多个电子突触,其中每一电子突触位于耦合于树突及轴突间的该交叉式数组的交叉点,每一电子突触配置为互连前突触电子神经元及后突触电子神经元;
其中每一电子突触包含:
三个存储器组件,配置为维持信息,其中一个存储器组件存储该电子突触的状态,并且另外两个存储器组件存储用以更新该电子突触的该状态的元信息;
三个轴突路径横贯于三个树突路径而形成 9 个交叉点的 3×3 数组,使得所述三个存储器组件位于该数组的对角线的所述交叉点上,提供 R 位、G 位及 B 位;以及
更新模块,配置为基于该元信息更新该电子突触的该状态,以响应针对强化学习的更新信号。
2. 如权利要求 1 所述的系统,其中该电子突触能够异步地并行操作。
3. 如权利要求 1 或 2 所述的系统,其中电子突触状态基于延迟更新信号根据强化学习而程序化。
4. 如权利要求 1 或 2 所述的系统,其中:
每一电子突触的该更新模块配置为更新该电子突触,使得:
响应该前突触神经元脉冲信号,该电子突触的该状态通过读取信号而更新,该读取信号经由该电子突触而从该前突触神经元传送至该后突触神经元;
响应该前突触神经元脉冲信号以及接着响应该后突触神经元脉冲信号,更新该电子突触的该状态使得该电子突触根据学习规则而设定;以及
响应该后突触神经元脉冲信号及接着响应该前突触神经元脉冲信号,更新该电子突触的该状态,其中该电子突触根据学习规则而重设。
5. 如权利要求 1 或 2 所述的系统,其中:
每一电子突触的该更新模块配置为基于强化学习而更新该电子突触,使得该电子突触基于脉冲时间相依可塑性而更新以响应更新信号。
6. 如权利要求 1 或 2 所述的系统,其中每一电子突触还包含多个轴突路径横贯于多个树突路径而形成交叉点的数组,使得该存储器组件位于该数组的对角线的该交叉点上。
7. 如权利要求 1 所述的系统,其中
该交叉式数组包含具有 N^2 电子突触的 $N \times N$ 交叉点;
每一电子突触针对电子强化学习而配置为:
并行地读取 R 行;
并行地读取且设定 G 列;
并行地重设 G 行;
并行地读取且设定 B 行;
并行地设定 B 列;
估计在 R 行及列上的一些设定位;以及
当报酬信号抵达时,提供通用值信号,且设定及重设该交叉式数组中的所有 $N^2 R$ 位。

8. 如权利要求 1、2 或 7 所述的系统, 其中每一电子突触包含六端装置, 其中两端用于读取、两端用于设定、以及两端用于重设。

9. 如权利要求 1、2 或 7 所述的系统, 其中每一电子突触包含基于静态随机存取存储器 (SRAM) 的存储器组件。

10. 如权利要求 1、2 或 7 所述的系统, 其中每一电子突触包含基于动态随机存取存储器 (DRAM) 的存储器组件。

11. 如权利要求 1、2 或 7 所述的系统, 其中每一电子突触配置为互连前突触电子神经元的轴突以及后突触电子神经元的树突。

12. 一种用于强化学习的电子突触装置, 包含:

电子突触, 配置为互连前突触电子神经元及后突触电子神经元, 该电子突触包含:

三个存储器组件, 配置为维持信息, 其中一个存储器组件存储该电子突触的状态, 并且另外两个存储器组件存储用以更新该电子突触的该状态的元信息;

三个轴突路径横贯于三个树突路径而形成 9 个交叉点的 3x3 数组, 使得所述三个存储器组件位于该数组的对角线的所述交叉点上, 提供 R 位、G 位及 B 位; 以及

更新模块, 配置为基于该元信息更新该电子突触的该状态, 以响应针对强化学习的更新信号。

13. 如权利要求 12 所述的装置, 其中该更新模块配置为基于该元信息更新该电子突触的该状态, 以响应基于学习规则的针对强化学习的延迟更新信号。

14. 如权利要求 12 或 13 所述的装置, 其中:

该更新模块配置为基于电子强化学习更新该电子突触的该状态, 使得该电子突触基于脉冲时间相依可塑性学习规则而更新。

15. 如权利要求 13 所述的装置, 其中:

该更新模块配置为更新该电子突触, 使得:

响应该前突触神经元脉冲信号, 该电子突触的该状态通过读取信号而更新, 该读取信号经由该电子突触而从该前突触神经元传送至该后突触神经元;

响应该前突触神经元脉冲信号以及接着响应该后突触神经元脉冲信号, 更新该电子突触的该状态使得该电子突触根据学习规则而设定; 以及

响应该后突触神经元脉冲信号及接着响应该前突触神经元脉冲信号, 更新该电子突触的该状态, 其中该电子突触根据学习规则而重设。

16. 如权利要求 12、13 或 15 所述的装置, 其中:

该更新模块配置为更新该电子突触, 使得若在时窗内接收更新信号:

响应该前突触神经元脉冲信号, 该电子突触的该状态通过读取信号而更新, 该读取信号经由该电子突触而从该前突触神经元传送至该后突触神经元;

响应该前突触神经元脉冲信号以及接着响应该后突触神经元脉冲信号, 更新该电子突触的该状态, 其中该电子突触根据学习规则而设定; 以及

响应该后突触神经元脉冲信号及接着响应该前突触神经元脉冲信号, 更新该电子突触的该状态, 其中该电子突触根据学习规则而重设。

17. 如权利要求 16 所述的装置, 其中该电子突触包含六端装置, 其中两端用于读取、两端用于设定、以及两端用于重设。

18. 如权利要求 12、13、15 或 17 所述的装置, 其中

该电子突触配置为互连前突触电子神经元的轴突以及后突触电子神经元的树突。

19. 一种用于强化学习的电子突触方法, 包括:

基于元信息而更新电子突触的状态, 以响应基于学习规则的针对强化学习的延迟更新信号;

其中该电子突触配置为互连前突触电子神经元及后突触电子神经元, 该电子突触包含:

三个存储器组件, 配置为维持信息, 其中一个存储器组件存储该电子突触的状态, 并且另外两个存储器组件存储用以更新该电子突触的该状态的元信息; 以及

三个轴突路径横贯于三个树突路径而形成 9 个交叉点的 3x3 数组, 使得所述三个存储器组件位于该数组的对角线的所述交叉点上, 提供 R 位、G 位及 B 位。

20. 如权利要求 19 所述的方法, 其进一步包括:

基于电子强化学习而更新该电子突触的该状态, 使得该电子突触基于脉冲时间相依可塑性学习规则而更新。

21. 如权利要求 19 所述的方法, 其中该方法更新该电子突触, 使得:

响应该前突触神经元脉冲信号, 该电子突触的该状态通过读取信号而更新, 该读取信号经由该电子突触而从该前突触神经元传送至该后突触神经元;

响应该前突触神经元脉冲信号以及接着响应该后突触神经元脉冲信号, 更新该电子突触的该状态使得该电子突触根据学习规则而设定; 以及

响应该后突触神经元脉冲信号及接着响应该前突触神经元脉冲信号, 更新该电子突触的该状态, 其中该电子突触根据学习规则而重设。

22. 如权利要求 19 所述的方法, 其中该方法更新该电子突触, 使得若在时窗内接收更新信号:

响应该前突触神经元脉冲信号, 该电子突触的该状态通过读取信号而更新, 该读取信号经由该电子突触而从该前突触神经元传送至该后突触神经元;

响应该前突触神经元脉冲信号以及接着响应该后突触神经元脉冲信号, 更新该电子突触的该状态, 其中该电子突触根据学习规则而设定; 以及

响应该后突触神经元脉冲信号及接着响应该前突触神经元脉冲信号, 更新该电子突触的该状态, 其中该电子突触根据学习规则而重设。

23. 如权利要求 19 所述的方法, 其中

该电子突触包含六端装置, 其中两端用于读取、两端用于设定、以及两端用于重设; 以及

该电子突触配置为互连前突触电子神经元的轴突以及后突触电子神经元的树突。

强化学习的电子突触

背景技术

[0001] 本发明一般涉及神经形态 (neuromorphic) 及突触器 (synapatron) 系统,特别是涉及在突触交叉式数组中产生脉冲时间相依可塑性。

[0002] 神经形态及突触器系统,还称作人造神经网络,是允许电子系统以实质类似于生物脑部的方式运作的运算系统。神经形态及突触器系统一般不利用控制 0 与 1 的传统数字模型。反之,神经形态及突触器系统产生功能上大致等效于生物脑部神经元的处理组件间的连接。神经形态及突触器系统可包含以生物神经元作模型的许多电子电路。

[0003] 在生物系统中,突触是指一神经元的轴突与另一神经元的树突间的连接点,对突触而言,二个神经元分别称为前突触 (pre-synaptic) 及后突触 (post-synaptic)。我们个人经验的本质存储于突触的传导中。突触传导随着时间以每个脉冲时间相依可塑性 (spike-timing dependent plasticity, STDP) 的前突触及后突触神经元的相对脉冲时间的函数而改变。若后突触神经元在前突触神经元发射后才发射,则 STDP 规则增加此突触的传导,若二个神经元发射的顺序颠倒,则 STDP 规则降低突触的传导。

发明内容

[0004] 本发明具体实施例提供配置为强化学习的电子突触。在一具体实施例中,电子突触配置为互连前突触电子神经元及后突触电子神经元。电子突触包含配置为存储电子突触的状态以及存储用以更新电子突触的状态的元信息的存储器组件。电子突触还包含更新模块,其配置为基于元信息更新电子突触的状态,以响应针对强化学习的更新信号。更新模块配置为基于元信息更新电子突触的状态,以响应基于学习规则的针对强化学习的延迟更新信号。

[0005] 在另一具体实施例中,本发明提供一种系统,其包含多个电子神经元以及配置为互连多个电子神经元的交叉式数组。交叉式数组包含多个轴突及多个树突,使得轴突与树突彼此横贯。交叉式数组还包含多个电子突触,其中每一电子突触位于耦合于树突及轴突间的交叉式数组的交叉点,每一电子突触配置为互连前突触电子神经元及后突触电子神经元。

[0006] 本发明的优选实施例提供空分多重存取电子突触,包含六端装置,其中两端用于读取、两端用于设定、以及两端用于重设。

[0007] 本发明的这些及其它特征、方面及优点可参考以下的详细说明、后附的权利要求书、以及所附随的附图而了解。

附图说明

[0008] 图 1A 根据本发明一具体实施例显示具有电子突触的交叉式数组的神经形态及突触器系统的示意图;

[0009] 图 1B 根据本发明一具体实施例显示在前突触路径及后突触路径的交叉点的电子突触的示意图;

- [0010] 图 2 根据本发明一具体实施例显示在涉及读取操作的交叉点的电子突触的示意图；
- [0011] 图 3 根据本发明一具体实施例显示在涉及 STDP- 设定操作的交叉点的电子突触的示意图；
- [0012] 图 4 根据本发明一具体实施例显示在涉及 STDP- 重设操作的交叉点的电子突触的示意图；
- [0013] 图 5 根据本发明一具体实施例显示在涉及 STDP- 设定操作的交叉点的电子突触的示意图；
- [0014] 图 6 根据本发明一具体实施例显示包括交叉点数组的电子突触的示意图；
- [0015] 图 7 根据本发明一具体实施例显示涉及 R 位的 STDP 操作的电子突触的示意图；
- [0016] 图 8 根据本发明一具体实施例显示涉及 G 位的 STDP 操作的电子突触的示意图；
- [0017] 图 9 根据本发明一具体实施例显示涉及 B 位的 STDP 操作的电子突触的示意图；
- [0018] 图 10 根据本发明一具体实施例显示电子突触的交叉式数组的示意图；
- [0019] 图 11 根据本发明一具体实施例显示电子突触的示意图；
- [0020] 图 12 根据本发明一具体实施例显示基于静态随机存取存储器 (static random access memory, SRAM) 的电子突触的示意图；
- [0021] 图 13 根据本发明一具体实施例显示基于动态随机存取存储器 (dynamic random access memory, DRAM) 的电子突触的示意图；以及
- [0022] 图 14 显示适用于实施本发明一具体实施例的信息处理系统的高阶方块示意图。

具体实施方式

[0023] 本发明具体实施例提供配置为供强化学习 (reinforcement learning, RL) 的电子突触。本发明具体实施例还提供利用此用于 RL 的电子突触的神经形态及突触器系统,其包含实施脉冲时间相依可塑性 (spike-timing dependent plasticity, STDP) 的交叉式数组。

[0024] 现在参考图 1A,其根据本发明一具体实施例显示具有交叉式数组的神经形态及突触器系统 10 的示意图。在一范例中,交叉式数组可包含一“超高密度交叉式数组”,其节距在约 0.1 奈米 (nm) 至 10 微米 (μm) 的一范围内。神经形态及突触器系统 10 包括具有多个神经元 14、16、18、及 20 的交叉式数组 12。这些神经元在此也称作“电子神经元”。神经元 14 及 16 为轴突神经元,而神经元 18 及 20 为树突神经元。轴突神经元 14 及 16 显示为分别具有连接至轴突路径 (axons) 26 及 28 的输出 22 及 24。树突神经元 18 及 20 显示为分别具有连接至树突路径 (dendrites) 34 及 36 的输入 30 及 32。轴突神经元 14 及 16 也包含输入且接收沿树突的信号,然而为了简化说明,并未显示这些输入及树突。因此,当轴突神经元 14 及 16 接收沿树突连接的输入时,其将作用为树突神经元。同样地,当树突神经元 18 及 20 将信号沿其轴突连接送出时,其将作用为轴突神经元。当任何神经元 14、16、18 及 20 发射时,其将送出一脉冲至其轴突及其树突连接。

[0025] 轴突 26、28 及树突 34、36 间的每一连接可由一突触装置 31 形成。突触装置所在的交点在此可称作“交叉点 (cross-point junctions)”。每一神经元 14、16、18 及 20 包括一对 RC 电路 48。一般而言,根据本发明一具体实施例,当接收自树突输入连接 (图未示) 的输入超过一临界时,轴突神经元 14 及 16 将“发射”(传输一脉冲)。当轴突神经元 14 及 16

发射时,其维持一 A-STDP 变量,此变量随由在其中一 RC 电路 48 中的电阻及电容值所确定的相对长且预定的时间常数衰减。举例来说,在一具体实施例中,此时间常数可为约 50ms。A-STDP 变量可通过使用电流镜或等效电路确定横跨电容的电压而取样。此变数用以达成轴突 STDP,其通过对自相关神经元最后一次发射起的时间进行编码,更细节的部分将描述于后。轴突 STDP 用以控制“增效 (potentiation)”,其在本文定义为增加的突触传导率。

[0026] 当树突神经元 18、20 发射时,其也维持 D-STDP 变量,此变量随由在其中一 RC 电路 48 中的电阻及电容值所确定的相对长且预定的时间常数衰减。举例来说,在一具体实施例中,此时间常数可为约 50ms。在其它具体实施例中,此变量可根据指数曲线以外的其它函数而以时间为函数衰减。举例来说,变量可根据线性、多项式、或二次函数而衰减。在本发明其它具体实施例中,变量可随时间增加而非减少。在任何情况下,此变量可用以达成树突 STDP,其通过对自相关神经元最后一次发射起的时间进行编码,将于下文中做进一步的讨论。树突 STDP 用以控制“抑制 (depression)”,其在本文定义为减少的突触传导率。

[0027] 电子突触 31 的功能包括:读取状态、以及根据 STDP 及基于 RL 的 STDP 而程序化状态。电子突触 31 可节能,使其适用于异步实施。此外,电子突触可省空间,使其适用于交叉式实施。图 1B 根据本发明一具体实施例显示在一前突触路径 26 及一后突触路径 36 的交叉点的电子突触 31 的透视图。

[0028] 参考图 2,关于在神经元 14 的轴突 26 与另一神经元 20 上的树突 36 间接触的交叉点的突触 31,二个神经元分别称为前突触 (pre-synaptic) 及后突触 (post-synaptic)。当前突触神经元 14 发射时,一“读取”信号从前突触神经元 14 传送至后突触神经元 20。参考图 3,当前突触神经元 14 发射且接着后突触神经元 20 发射时,则突触 31 为 STDP- 设定。参考图 4,当后突触神经元 20 发射且接着前突触神经元 14 发射时,则突触 31 为 STDP- 重设。

[0029] 强化学习 (RL) 一般包含基于行为后果的学习,其中 RL 模块基于过去事件选择行为。由 RL 模块所接收的一强化信号 (如一报酬信号) 为指示行为成功的一报酬 (一数值)。RL 模块接着学习选择可随时间增加报酬的行为。在根据本发明的强化学习的一实施中,STDP- 设定与 STDP- 重设操作并无立即发生。反而,若在一时窗内出现一报酬 (“值”) 信号,则施以 STDP- 设定或 STDP- 重设操作。

[0030] 根据本发明一具体实施例,突触 31 执行多个信息位。在一范例中,根据 RGB 架构,突触 31 维持三个位,包括位 R、位 G、及位 B。位 R 用于读取,位 G 用于 STDP- 设定,而位 B 用于 STDP- 重设。一开始,位 G 及 B 设定为 0 作为其自然状态。若前突触神经元发射且接着后突触神经元发射时,则针对 STDP- 设定而设定位 G (例如设定为 1)。若后突触神经元发射且接着前突触神经元发射时,则针对 STDP- 设定而设定位 B (例如设定为 1)。

[0031] 在一具体实施例中,后突触神经元发射且接着后突触神经元发射,则 STDP- 重设实施于位 B 及 G。举例来说,基于时间常数衰减 (例如 1 秒) 将位 B 及 G 重设为 0。在另一具体实施例中,重设位 B 及 G 包含一随机程序来重设 B 及 G,而与神经元发射无关。

[0032] 在一具体实施例中,当出现一报酬如下,则设定及重设位 R:

[0033] 当一报酬出现时:

[0034] 若 $G=1$ 且 $B=0$,则设定 R,

[0035] 若 $B=1$ 且 $G=0$,则重设 R,

[0036] 若 $G=1$ 且 $B=1$,或 $G=0$ 且 $B=0$,则不对 R 采取行动。

[0037] (when a reward occurs:

[0038] if $G=1$ and $B=0$, then set R ,

[0039] if $B=1$ and $G=0$, then reset R ,

[0040] if $G=1$ and $B=1$, or $G=0$ and $B=0$, take no action on R .)

[0041] 参考图 5, 在一实施中, 突触 31 包含 $n \times n$ 交叉点数组。图 6 显示包含九个交叉点的 3×3 数组的一范例突触 31 ($n=3$), 其中使用三个对角线交叉点。

[0042] 用以读取位 R 的逻辑位于突触 31 的周边, 如图 7 中的范例所示, 其进一步描述读取突触 31 的位 R , 其中当前突触神经元发射时, 其传送一读取脉冲至后突触神经元。接着, 当脉冲从前突触神经元经由突触 31 的 R 交叉点抵达时, 后突触神经元异步地读取脉冲。位 R 的设定 / 重设的逻辑包含于突触 31 内。在一实施中, 位 R 可使用 DRAM 装置而实施。

[0043] 用以设定位 G 的逻辑位于突触 31 的周边, 如图 8 中的范例所示, 其进一步描述设定突触 31 的位 G , 其中当后突触神经元发射时, 其传送一警示脉冲至前突触神经元。根据其最后的发射的时间, 前突触神经元可能设定一前突触设定脉冲。后突触神经元总是传送一后突触设定脉冲。若前突触设定及后突触设定脉冲两者一起抵达位 G 的交叉点, 则设定位 G 。

[0044] 此外, 用以重设位 G 的逻辑设置于突触 31 的周边。在一具体实施例中, 位 G 较佳的设定值为零, 且在一特定时间常数 (例如 1 秒) 后重设。在其它具体实施例中, 重设 G 包含重设位 G 的一随机 (random stochastic) 程序, 其以完全异步的方式, 而与神经元发射无关。在一范例中, 程序具有约 1 秒的平均重设时间且具有一重尾分布 (heavy tail distribution)。在一范例中, G 的重设由前突触神经元所启动。在一实施中, 位 G 可使用 DRAM 装置而实施。

[0045] 用以重设位 B 的逻辑设置于突触 31 的周边。参考图 9, 在一具体实施例中, 若后突触神经元发射且接着前突触神经元发射, 则设定位 B 。若前突触神经元发射且接着后突触神经元发射, 则设定位 G 。前突触神经元发射时将警示一后突触神经元。取决于其最后的发射, 后突触神经元可能地设定一后突触设定脉冲。前突触神经元总是传送一前突触设定脉冲。若前突触脉冲及后突触设定脉冲两者一起抵达 B 位, 则设定位 B 。此外, 用以重设 B 的逻辑位于方块 31 中。位 B 较佳设定零, 且在一特定时间常数 (例如 1 秒) 后简单地重设。在另一具体实施例中, 一随机程序以完全异步的方式重设位 B , 而与神经元的发射无关。在一范例中, 程序具有约 1 秒的平均重设时间, 且具有一重尾分布。在一范例中, 重设由后突触神经元所启动。在一实施中, 位 B 可使用 DRAM 装置而实施。

[0046] 参考图 10, 在一具体实施例中, 本发明提供用以实施根据所述本发明具体实施例的电子强化学习突触的系统 70。系统 70 包含 RGB 突触方块 31 的 $N \times N$ 交叉式数组, 其可异步地并行操作 (N 行及 N 列)。系统 70 还包含 N 个前突触神经元 (如前 1、前 2、...、前 N) 以及 N 个后突触神经元 (如后 1、后 2、...、后 N), 其经由突触 31 的交叉式数组而互连。在一实施中, 每一后突触神经元 31 包含一电子混合模式 (模拟 - 数字) 的异步神经元。

[0047] 状态将根据异步实施的 STDP 及基于 RL 的 STDP 而程序化。当前突触神经元发射时, 一读取信号从前突触神经元传送至后突触神经元, 其在脉冲到达时异步地读取脉冲, 且可能地设定一前设定 (pre-set) 脉冲以及总是设定一后设定 (post-set) 脉冲。

[0048] 在本发明一具体实施例中, 每一电子突触 31 配置为互连一前突触电子神经元及

一后突触电子神经元。电子突触 31 包含存储器组件（例如图 11 中的存储器装置 31R、31G、31B），其配置为存储电子突触的状态以及存储用以更新电子突触状态的元信息。每一电子突触单元 31 还包含一更新模块（例如图 11 中的模块 31L），其配置为基于元信息而更新电子突触的状态，以响应针对强化学习的一更新信号。更新模块配置为基于元信息而更新电子突触的状态，以响应基于学习规则的针对强化学习的一延迟更新信号。

[0049] 图 11 描述 R、G、B 突触数组 31 为一突触单元（块）的范例实施，其可与其它突触单元 31 并行地操作，而不需相位（且不需对读取、设定及重设的分时多重存取）。每一突触单元 31 可与其它突触单元 31 完全异步地操作，因而不需要频率。

[0050] 在本发明一具体实施例中，每一 RGB 突触单元 31 包含一数字互补金属氧化物半导体（complementary metal-oxide-semiconductor, CMOS）更新逻辑 31L 于局部突触单元级，可用以写入 R 单元。在一实施中，单元 31 包含分别针对位 R、B、及 G 的存储器组件 31R、31B、及 31G。存储器组件可包含静态随机存取存储器（static random access memory, SRAM）、动态随机存取存储器（dynamic random access memory, DRAM）、相变存储器（Phase-change memory, PCM）、磁性隧道接合（magnetic tunnel junction, MTJ）等。在此具体实施例中，突触单元 31 包含空分多重存取电子突触，其中电子突触表示为六端装置，其中两端用于读取、两端用于设定、以及两端用于重设。

[0051] 在本发明另一具体实施例中，更新模块 31L 包含软件模块，其包括计算机可读取程序代码以执行于一处理器（如图 14 中的信息处理系统 100）上，其中软件模块包括配置为更新电子突触状态的计算机可读取程序代码，如本文中根据本发明具体实施例所述。

[0052] R 存储器单元维持突触的状态。G 及 B 存储器单元维持用于突触的后续更新的元信息。神经元确定存储器单元的读取 / 写入信息。在图 11 中，突触 31 提供自前突触神经元至后突触神经元的连接，其合作以启动适当的字符线及位线来读取 / 写入 R、G、B 存储器单元。神经元使用写入端口外部地仅写入 G 及 B 存储器单元。G 及 B 存储器单元由更新逻辑 31L 使用读取端口而内部地读取，以相应地更新 R 存储器单元。R 存储器单元系使用读取端口而外部地读取，且由更新逻辑 31L 使用写入端口而内部地写入（即更新）。

[0053] 突触的状态可具有一个或多个位，其存储指示突触传导率级的多个值。在一具体实施例中，R 存储器单元存储突触状态，其中突触状态为 1- 位突触（0 为指示一连接的传导状态，1 为指示无连接的非传导状态）。神经元可通过读取 R 存储器单元确定经过突触的一连接。为更新 R 单元用于学习操作，耦合至突触的前突触神经元及后突触神经元实施一程序，以写入 B 及 G 存储器单元，供强化学习。神经元使用写入端口将更新值存储至 B 及 / 或 G 存储器单元。之后，来自 B 或 G 存储器单元的一更新值用来更新 R 存储器单元为突触的状态，以响应一进入的报酬信号。R 存储器单元用 B 存储器单元的值或以 G 存储器单元的值更新，其取决于后来进入的报酬信号为一强化信号（延迟的更新），如前所述。在一范例中，STDP 值存储于 G 或 B 存储器单元，且之后突触的状态通过以来自 G 或 B 单元的值更新 R 存储器单元而更新。

[0054] 在一实施中，平行的字符线（水平）及位线（垂直）用以存取存储器单元。每一存储器单元具有一读取字符线、读取位线、写入字符线及写入位线。在一范例中，更新逻辑 31L 实施 B 及 G 存储器单元元信息的逻辑互斥或（XOR）组合，以更新突触的 R 存储器单元状态。突触单元 31 以 SRAM 及 DRAM 实施提供强化学习。参考图 12，在基于 SARM 的 RGB 单元

实施中,每一 SRAM 单元 31 为可转换的(可由周边电路以行或列存取)。参考图 13,在基于 SRAM 及 DRAM 的实施中,在 DRAM 存储器组件中的数据随时间衰减至一基态。频率信号用以计时交叉式数组中存储器单元的操作。存储器单元可同步或异步地存取。

[0055] 当前突触神经元发射且接着后突触神经元发射时,设定突触。当后突触神经元发射且接着前突触神经元发射时,重设突触,且若一报酬(值)信号于一时窗内出现,则施以 STDP- 设定或重设。

[0056] 学习突触的电子强化还包含:并行地读取 R 行、并行地读取且设定 G 列、并行地重设 G 行、并行地读取且设定 B 行、并行地设定 B 列、估计在 R 行及列上的一些设定位、以及当一报酬信号抵达时,实施 / 提供一通用值信号,且并行地设定及重设交叉式数组中的所有 N^2R 位。

[0057] 图 14 为适用于实施本发明的一具体实施例的信息处理系统 100 的高阶方块图。计算机系统包括一个或多个处理器,例如处理器 102。处理器 102 连接至通信架构 104(例如通信总线、交叉条、或网络)。

[0058] 计算机系统可包括显示接口 106,其自通信架构 104(或帧缓冲器,图未示)转送图像、文字、及其它数据,供显示于显示单元 108。计算机系统也包括主要存储器 110,较佳为随机存取存储器(random access memory, RAM),且也可包括次要存储器 112。次要存储器 112 可包括例如硬盘驱动器 114 及 / 或可移式存储驱动器 116,代表例如软盘驱动器、磁带驱动器、或光学驱动器。可移式存储驱动器 116 以本领域技术人员公知的方式,自可移式存储单元 118 读取及 / 或对其写入。举例而言,可移式存储单元 118 代表软盘、光盘、磁带、或光盘等,其由可移式存储驱动器 116 作读取及写入。将了解到,可移式存储单元 118 包括具有计算机软件及 / 或数据存储于其中的计算机可读取介质。

[0059] 在替代具体实施例中,次要存储器 112 可包括其它类似手段,允许计算机程序或其它指令加载计算机系统。这样的手段可包括例如可移式存储单元 120 与接口 122。这样的手段的范例可包括程序包及封包接口(例如在视频游戏装置中可找到的)、可移式存储器芯片(例如 EPROM 或 PROM)及相关插槽、及允许软件及数据自可移式存储单元 120 移至计算机系统的其它可移式存储单元 120 与接口 122。

[0060] 计算机系统也可包括通信接口 124。通信接口 124 允许软件及数据在计算机系统及外部装置之间转移。通信接口 124 的范例可包括调制解调器、网络接口(例如以太网络卡)、通信端口、或 PCMCIA 插槽及插卡等。通过通信接口 124 进行转移的软件与数据可为信号的形式,此信号可例如为电子、电磁、光学、或可由通信接口 124 接收的其它信号。这样的信号通过通信路径(即信道)126 提供给通信接口 124。此通信路径 126 承载信号,且可利用电线、缆线、光纤、电话线、蜂窝电话链路、射频(RF)链接、及 / 或其它通信信道而实施。

[0061] 在此文件中,“计算机程序介质”及“计算机可用介质”及“计算机可读取介质”等词一般是指如主要存储器 110 与次要存储器 112、可移式存储驱动器 116、及安装于硬盘驱动器 114 的硬盘等介质。

[0062] 计算机程序(也称为计算机控制逻辑)存储于主要存储器 110 及 / 或次要存储器 112。计算机程序也可经由通信接口 124 而接收。当执行这样的计算机程序时,将使得计算机系统能执行在此所述的本发明各特征。特别是,当执行计算机程序时,将使得处理器 102 能执行计算机系统的各特征。因此,这样的计算机程序代表计算机系统的控制器。

[0063] 根据以上描述,本发明提供实施本发明的具体实施例的系统、计算机程序产品、及方法。权利要求内所主张的单数组件并不代表“唯一”,而是指“一个或多个”,除非有特别说明。本领域技术人员所公知或随后可得知的与上述范例具体实施例的组件均等的所有结构与功能,均包含在本发明的主张内。

[0064] 在此所使用的术语仅用于描述特定具体实施例的目的且并不意欲为本发明的限制。如本文中所使用,除非上下文清楚地另外指示,否则单数形式“一”(a、an) 及“该”(the) 也有包括复数形式的意思。应进一步理解,术语“包含 (comprises 及 / 或 comprising)”于本说明书中使用时表示所陈述的特征、实体、步骤、操作、单元及 / 或部件的存在,但并不排除一个或多个其它特征、实体、步骤、操作、单元、部件及 / 或其群组的存在或添加。

[0065] 后附的权利要求中的所有手段或步骤功能组件的对应结构、材料、行为及等效物意欲包括用于结合其它所主张的组件以执行功能的任何结构、材料或行为。本发明的描述以出于说明及描述的目的呈现,其并不意欲为详尽的或限制本发明于所揭示的形式。对本领域技术人员而言,在不偏离本发明的范围及精神的情况下,许多修改及变化是显而易见的。具体实施例选择并描述为可最佳地解释本发明的原理及实际应用,且使得其它本领域技术人员能够理解本发明各种具体实施例具有各种修改以适用于所预期的特定用途。

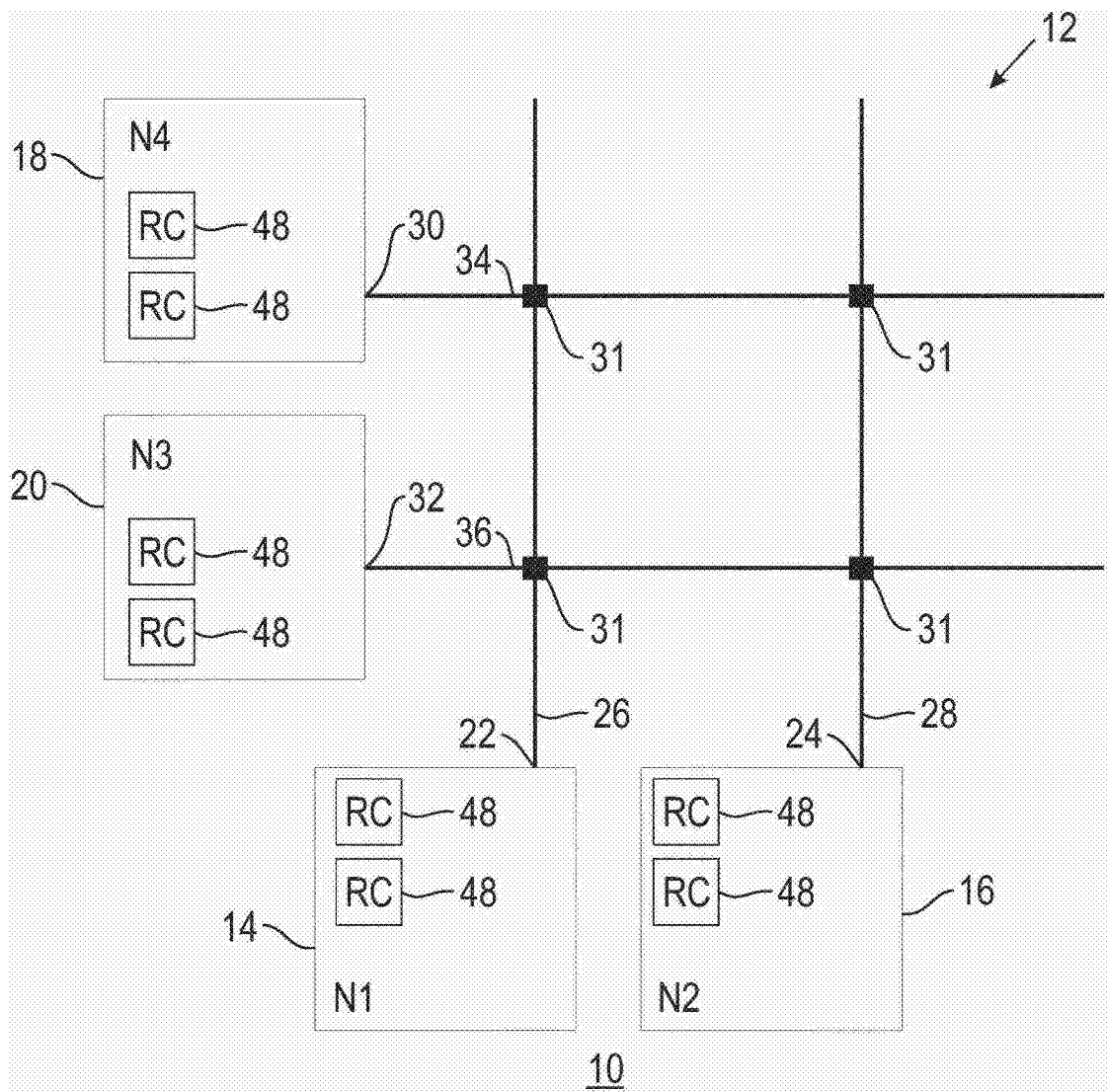


图 1A

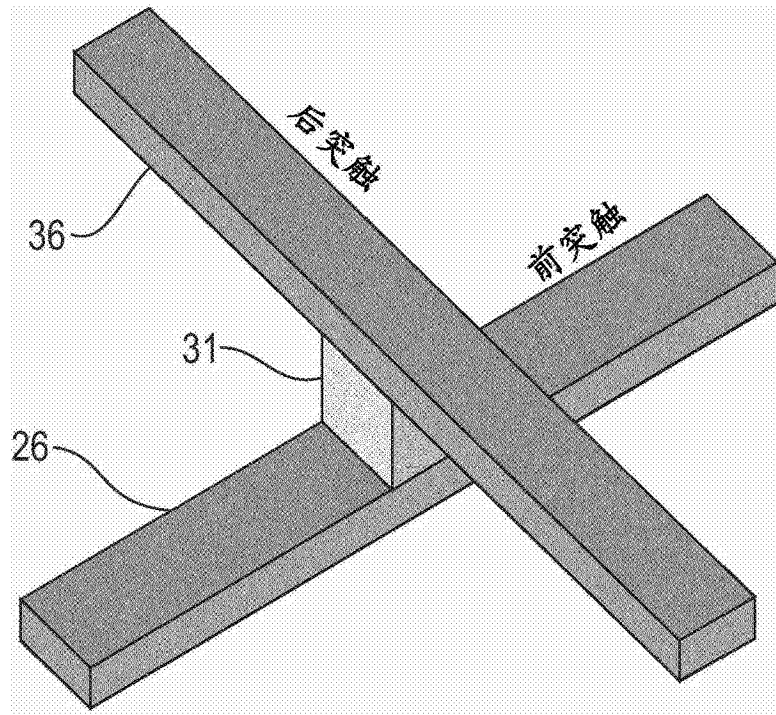


图 1B

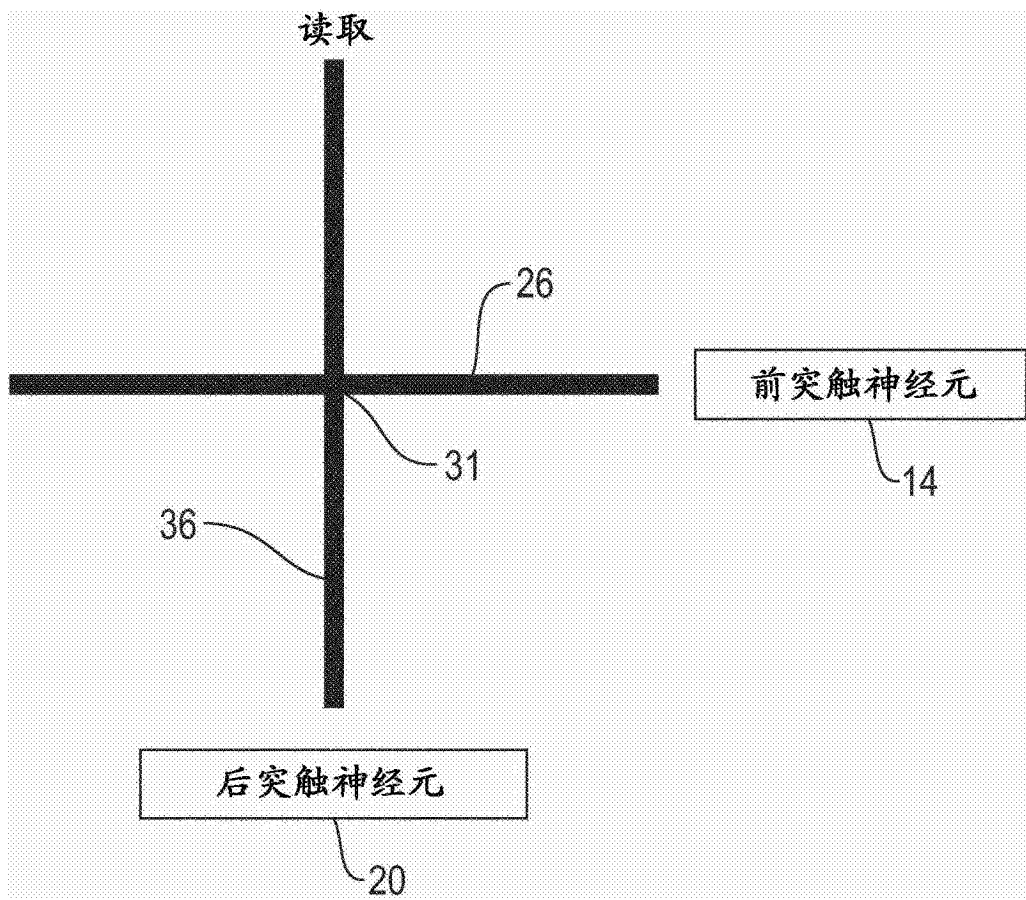


图 2

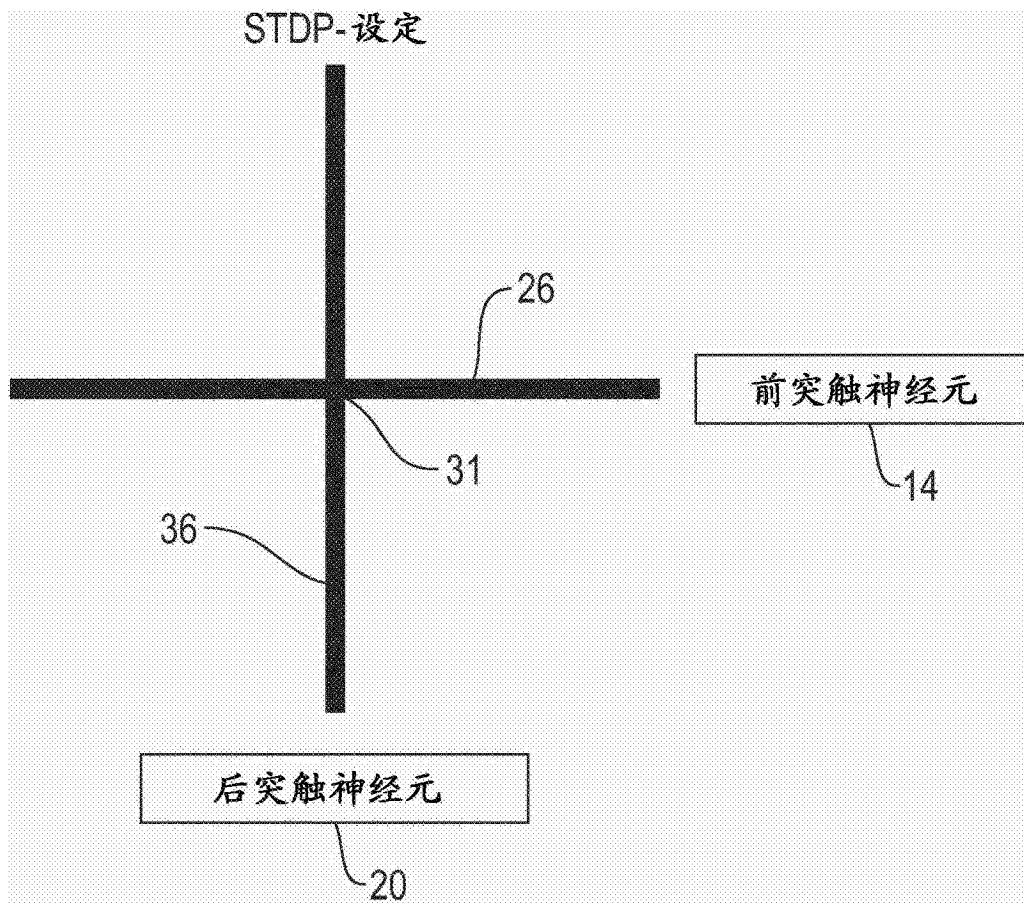


图 3

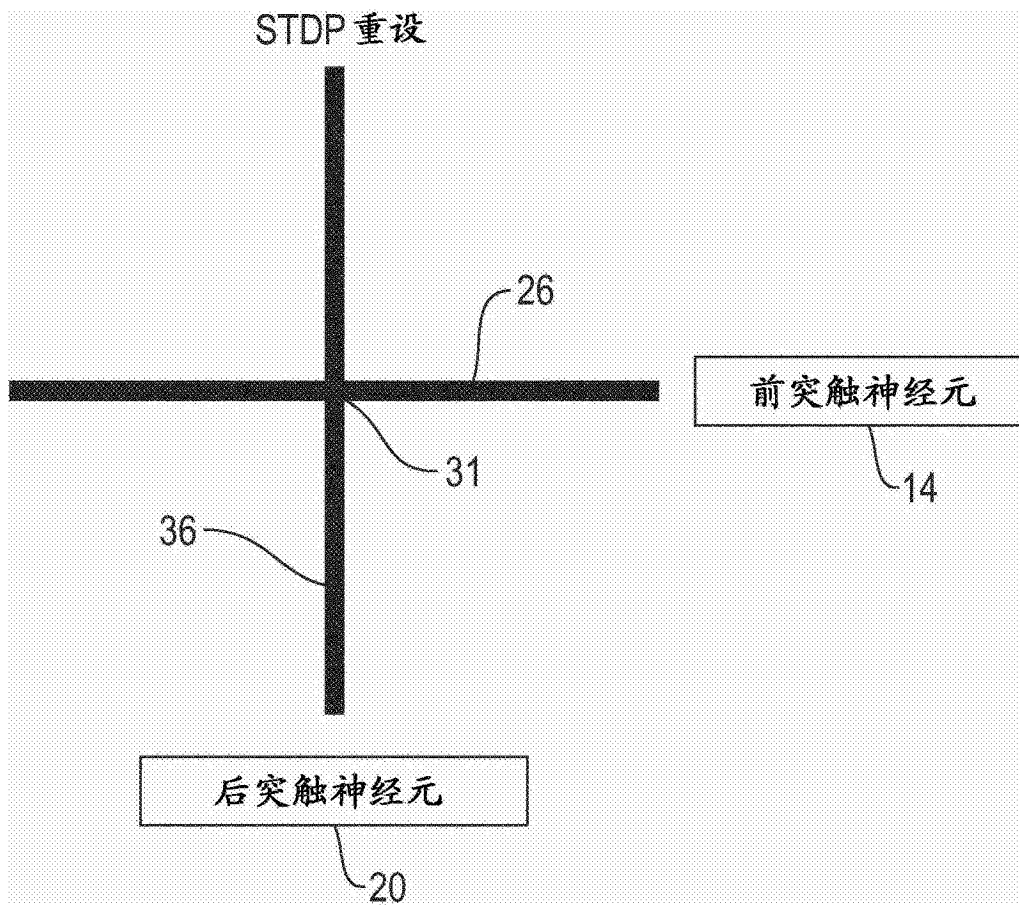


图 4

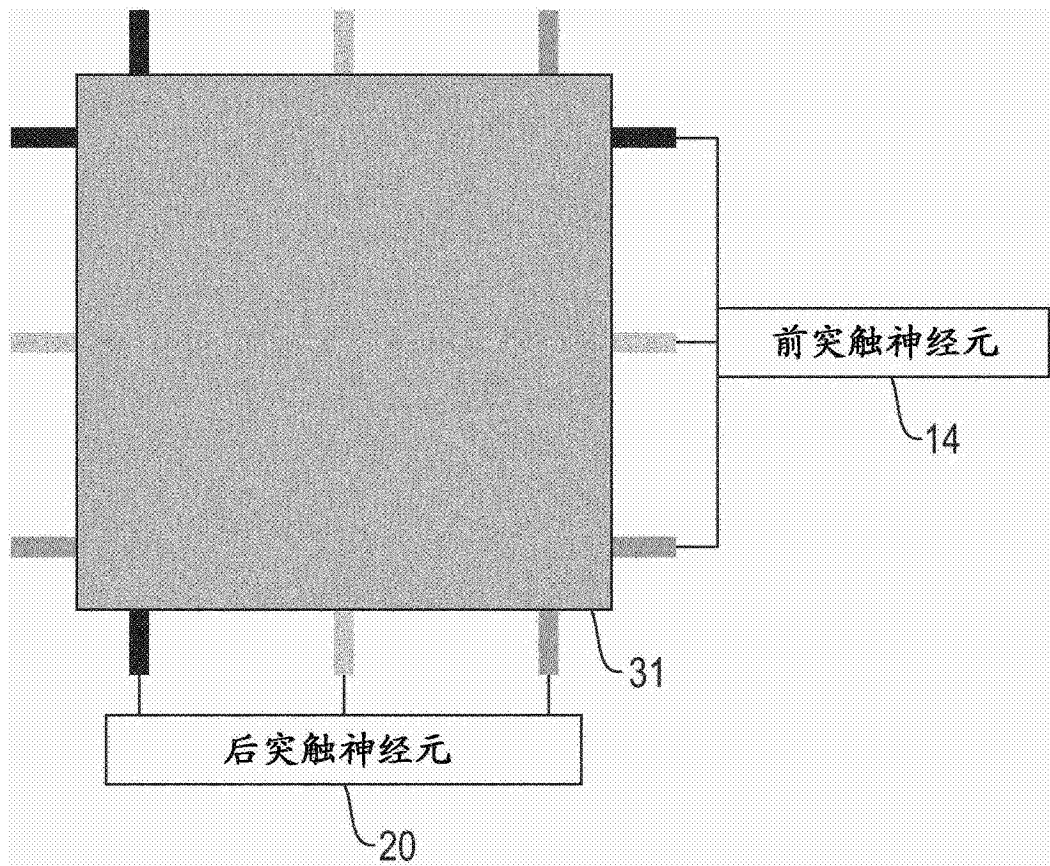


图 5

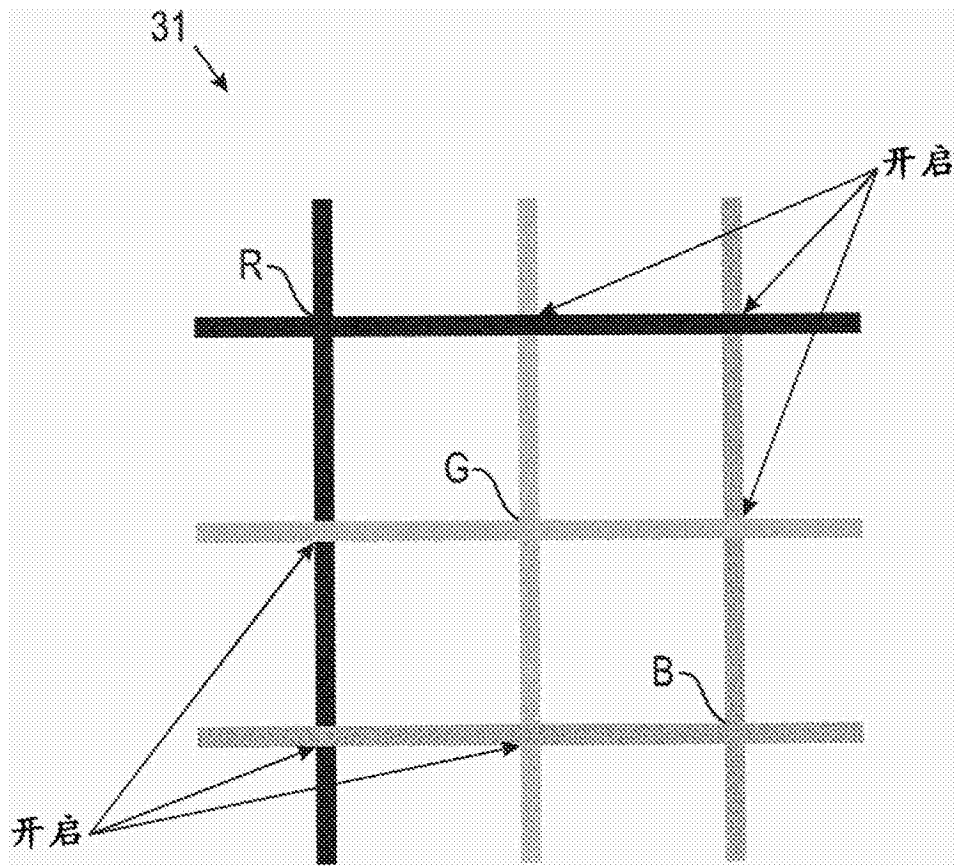


图 6

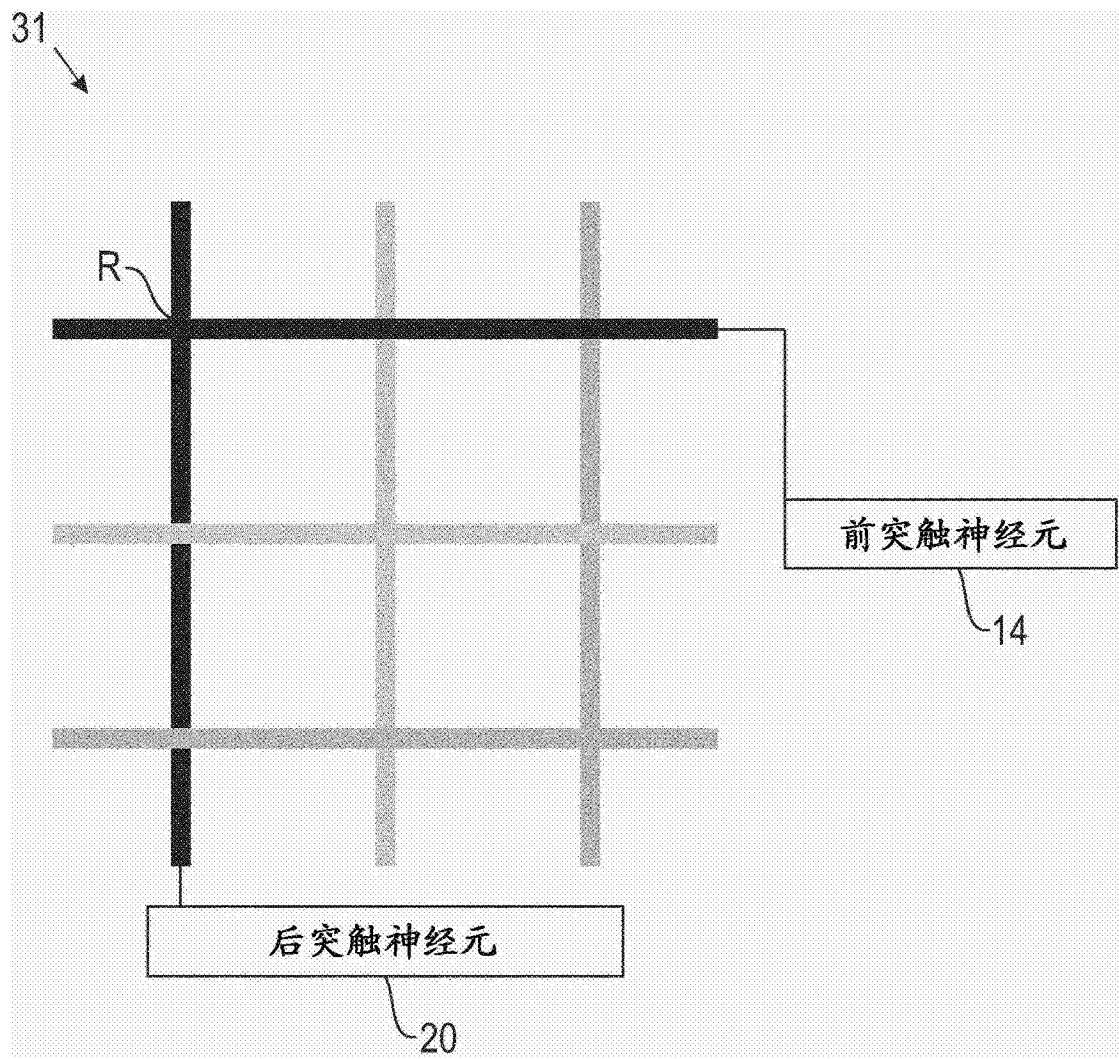


图 7

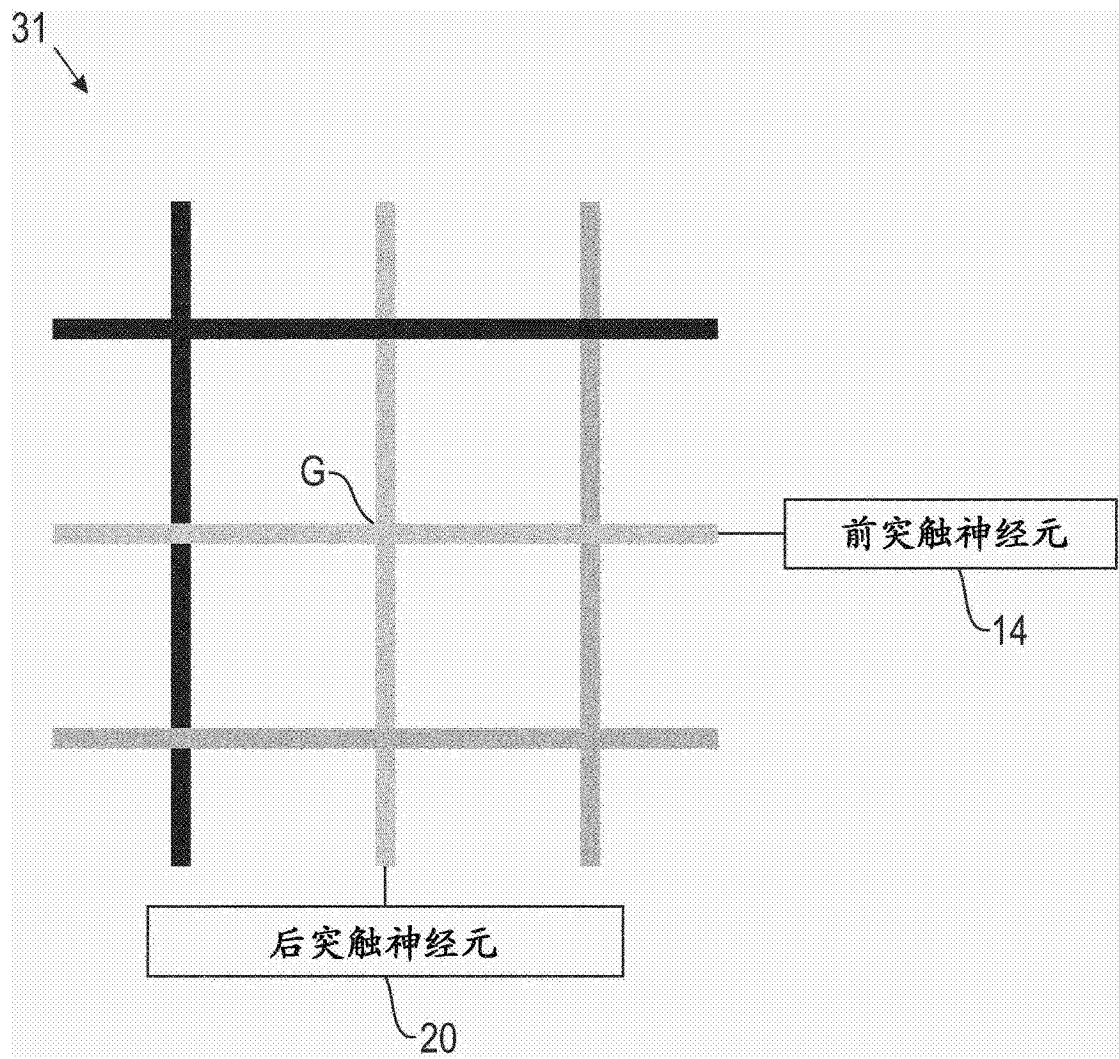


图 8

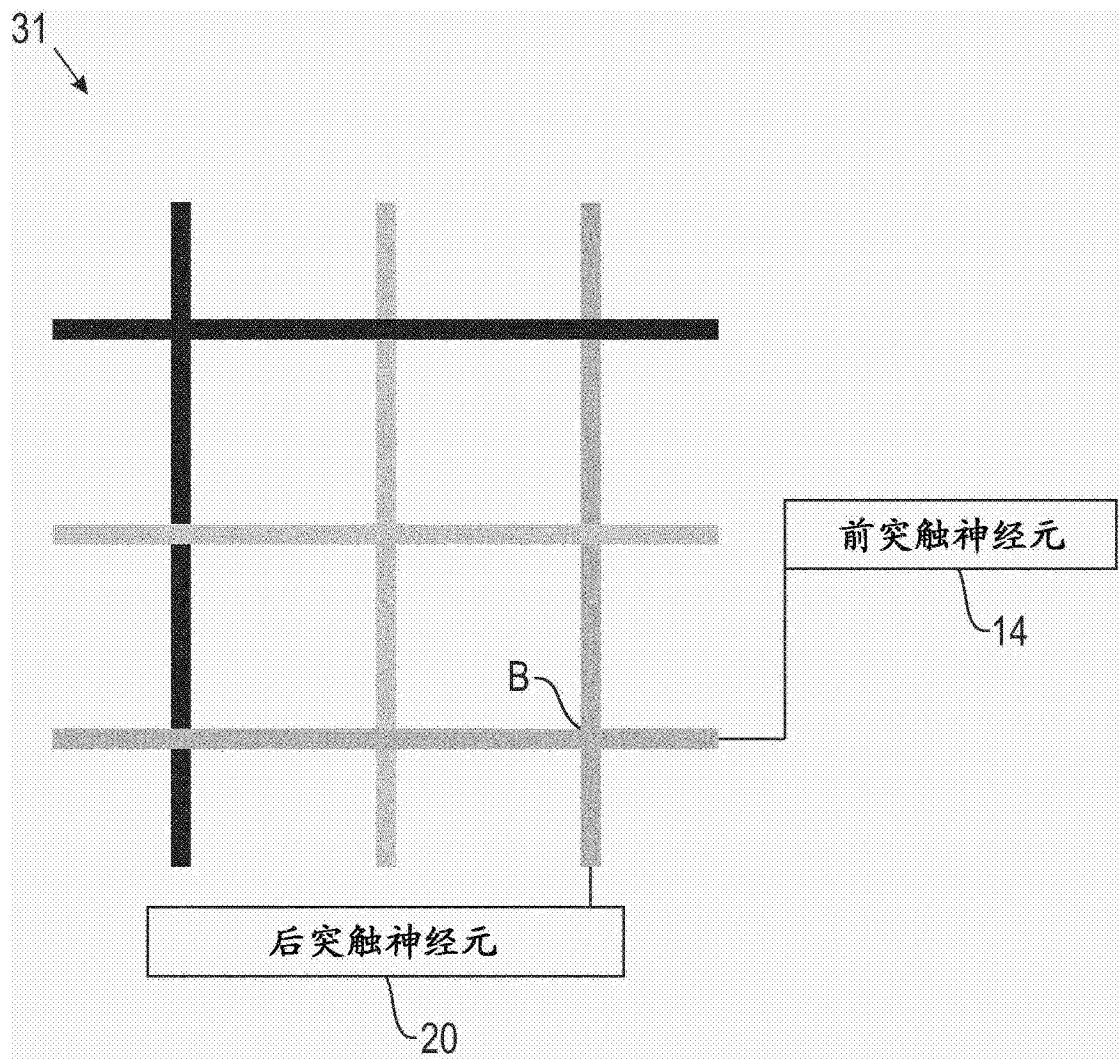


图 9

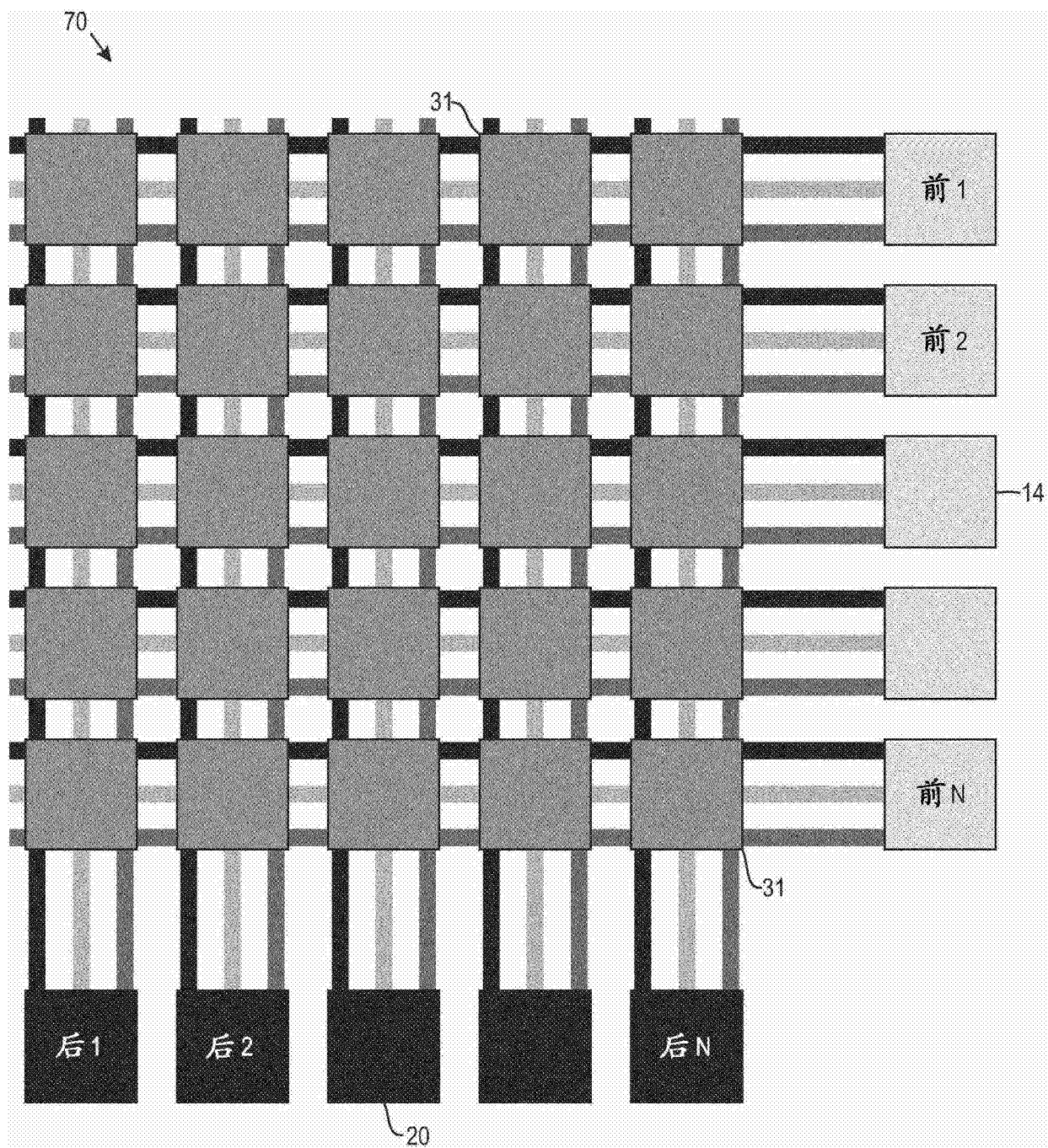


图 10

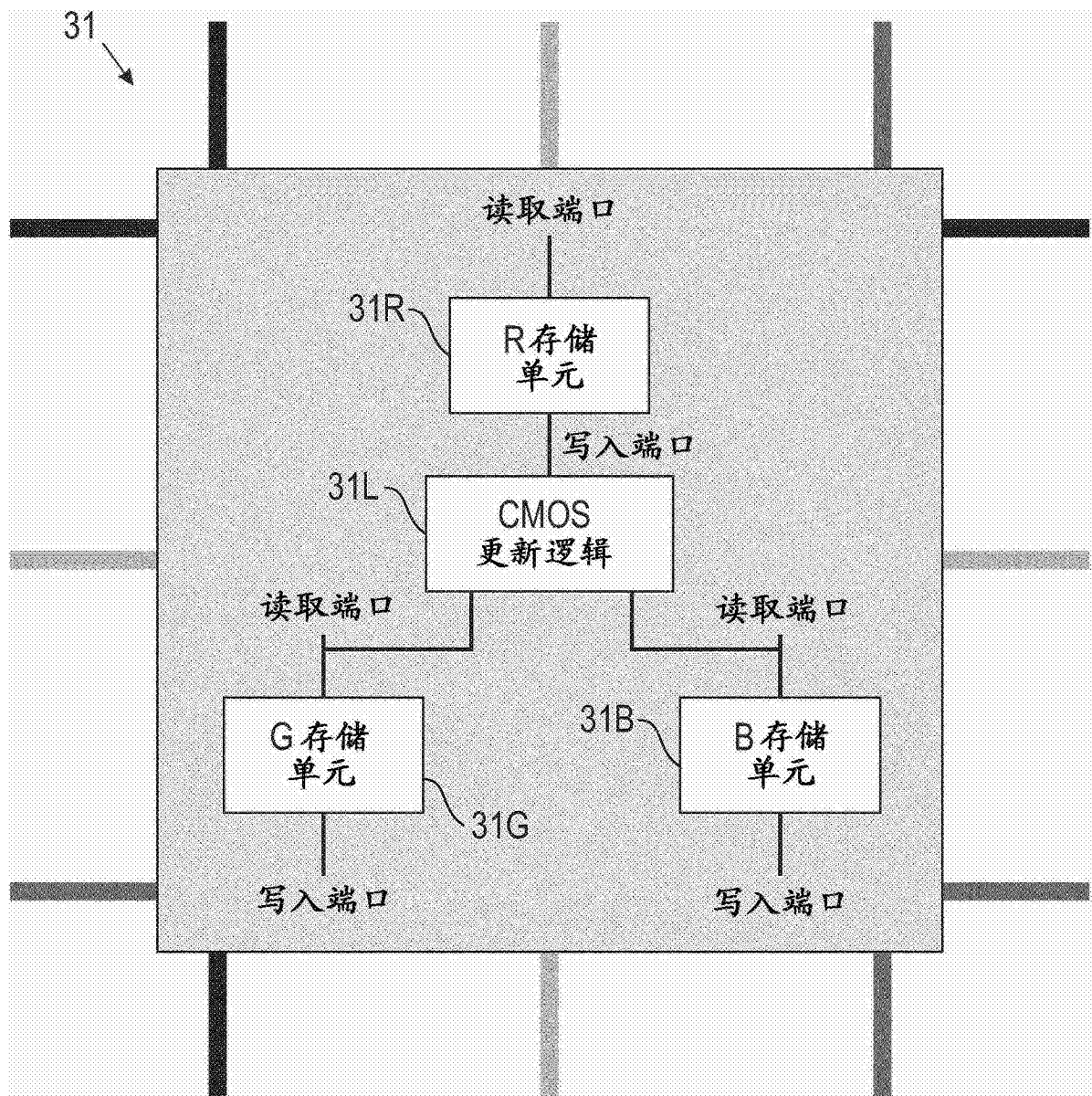


图 11

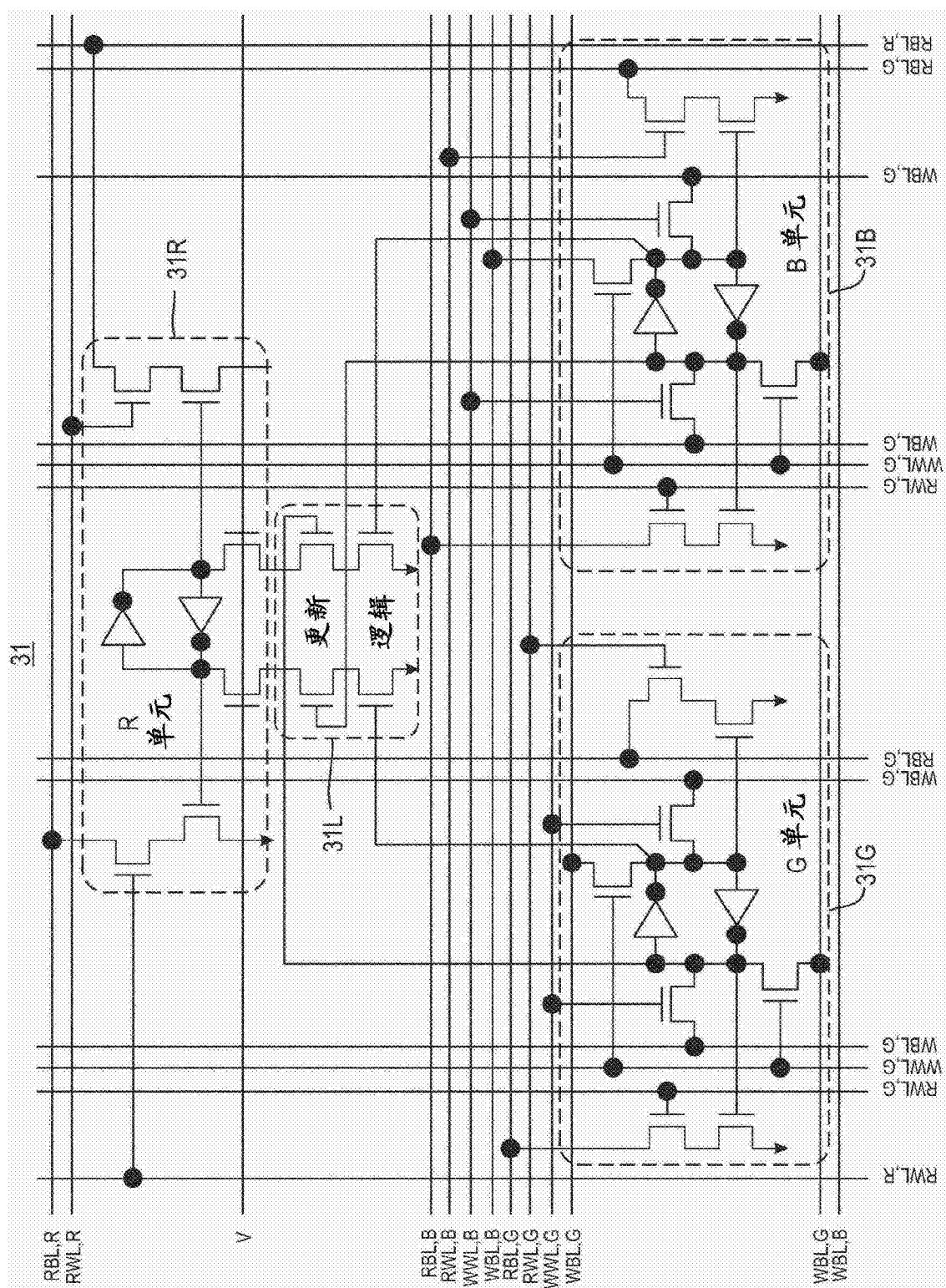


图 12

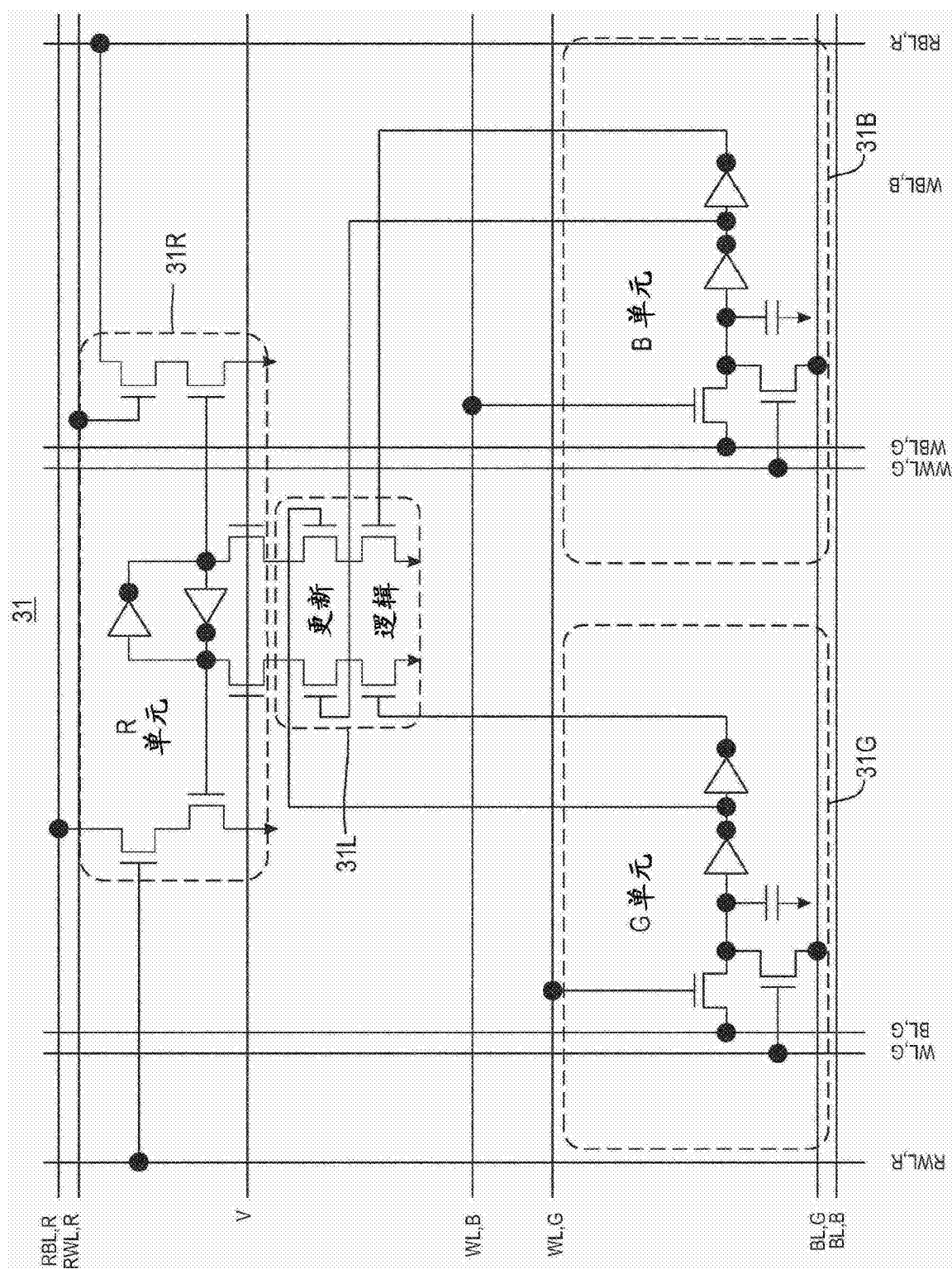


图 13

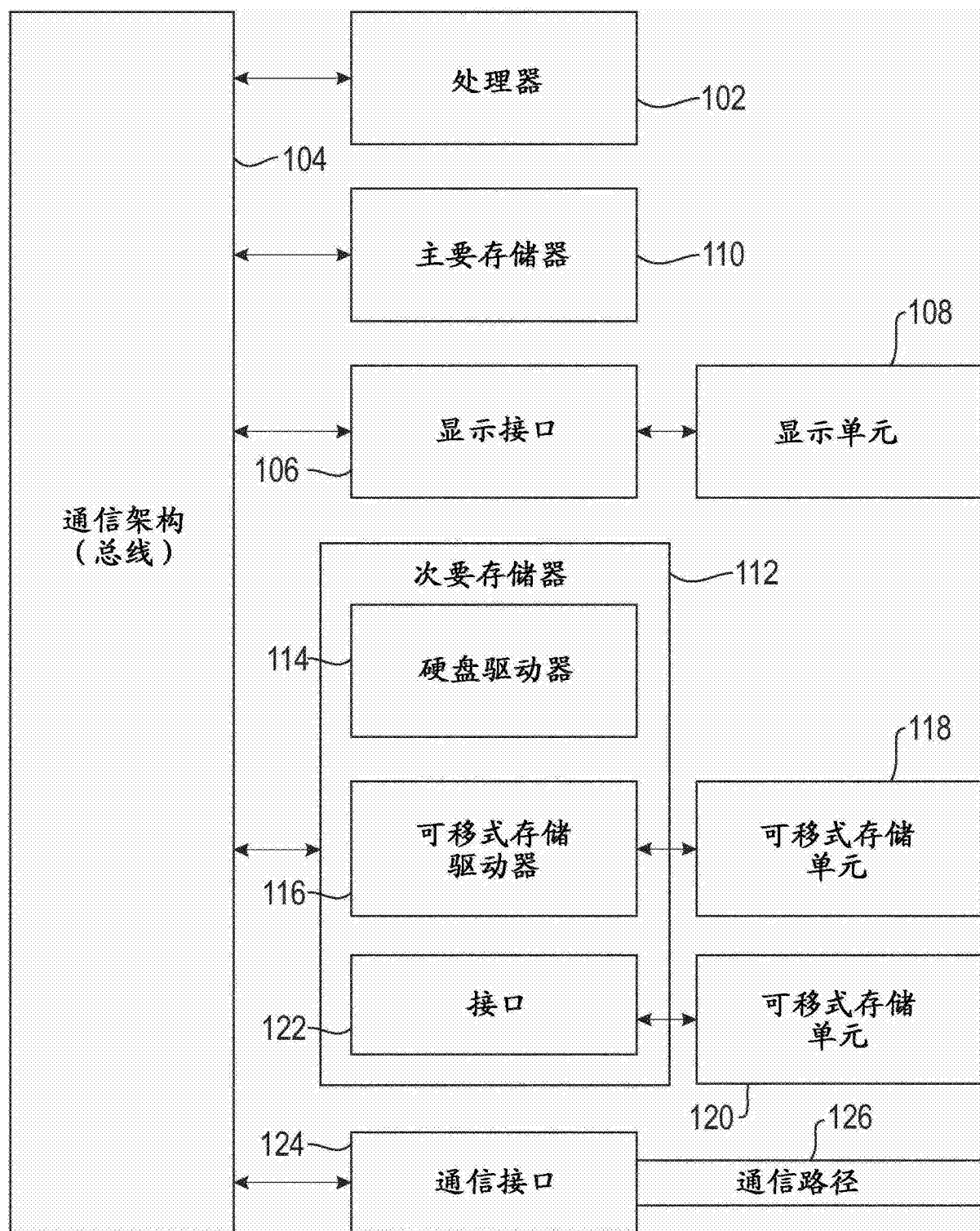


图 14