



(12) 发明专利

(10) 授权公告号 CN 111176552 B

(45) 授权公告日 2024.06.04

(21) 申请号 201910767783.9

(22) 申请日 2019.08.20

(65) 同一申请的已公布的文献号  
申请公布号 CN 111176552 A

(43) 申请公布日 2020.05.19

(30) 优先权数据  
10-2018-0138077 2018.11.12 KR

(73) 专利权人 三星电子株式会社  
地址 韩国京畿道

(72) 发明人 金炳喜 徐正旻 闵贤圭 梁承俊  
黄珠荣

(74) 专利代理机构 北京市立方律师事务所  
11330  
专利代理师 李娜

(51) Int. Cl.  
G06F 12/02 (2006.01)  
G06F 3/06 (2006.01)

G06N 3/0499 (2023.01)  
G06N 3/09 (2023.01)

(56) 对比文件  
CN 103714856 A, 2014.04.09  
CN 105700819 A, 2016.06.22  
CN 1853181 A, 2006.10.25  
WO 2013076763 A1, 2013.05.30  
US 2018143762 A1, 2018.05.24  
CN 106257399 A, 2016.12.28  
US 2013024609 A1, 2013.01.24  
US 2017090759 A1, 2017.03.30  
US 2017116520 A1, 2017.04.27  
CN 106201346 A, 2016.12.07  
KR 20060109042 A, 2006.10.19  
US 6105034 A, 2000.08.15  
金江军, 潘懋. 电子政务数据资源的开发利用. 地理与地理信息科学. 2003, (06), 全文.

审查员 刘茜茜

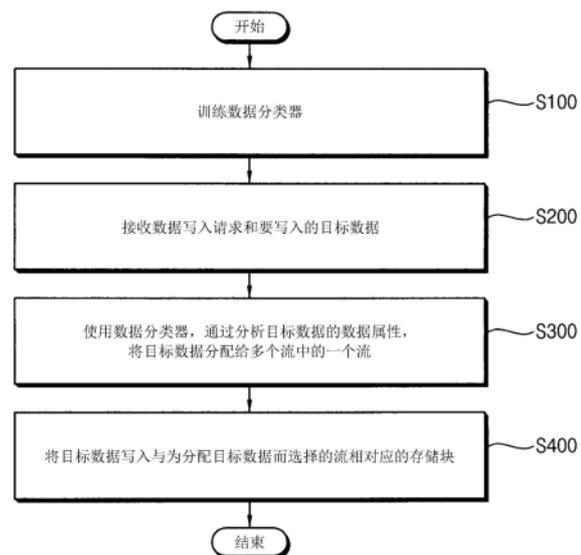
权利要求书2页 说明书16页 附图23页

(54) 发明名称

操作存储设备的方法、存储设备及包括其的存储系统

(57) 摘要

提供了操作存储设备的方法、存储设备及包括其的存储系统。所述存储系统包括多个存储块, 操作所述存储设备的方法包括: 接收数据写入请求和要写入的目标数据。使用诸如神经网络模型的数据分类器, 基于所述目标数据的数据属性, 将所述目标数据分配给从多个流中选择一个流。将所述目标数据写入被分配给了为分配所述目标数据而选择的流的存储块中, 从而将已经分配给流的所述目标数据写入被分配给了所述流的所述存储块中。



1. 一种操作存储设备的方法,所述存储设备包括存储器件的多个存储块,所述方法包括:

接收数据写入请求和要写入的目标数据;

使用数据分类器将所述目标数据分配给从多个流中选择一个流,所述数据分类器被配置为基于所述目标数据的数据属性从所述多个流中选择所述流,所述目标数据的数据属性与所述目标数据的规律性相关联,并且所述目标数据的规律性与所述目标数据中的数据位相关联;以及

将所述目标数据写入从所述多个存储块中选择一个存储块中,其中,所述存储块已经被分配给了所述目标数据已经被分配给的所述流,

其中,所述多个流中的各个流表示所述目标数据的规律性,并且

其中,将所述目标数据分配给从所述多个流中选择一个流包括:

在确定所述目标数据的多个数据位具有规律性的情况下,将所述目标数据分配给从所述多个流中选择的第一流,以及

在确定所述目标数据的所述多个数据位不具有规律性的情况下,将所述目标数据分配给从所述多个流中选择的第二流。

2. 根据权利要求1所述的方法,其中,将所述目标数据写入所述存储块中包括:在将所述目标数据分配给了从所述多个流中选择的第一流的情况下,如果从所述多个存储块中选择的第一存储块被分配给了所述第一流,则将所述目标数据写入所述第一存储块。

3. 根据权利要求2所述的方法,其中,将所述目标数据写入所述存储块中还包括:如果所述第一存储块尚未被分配给所述第一流,将从所述多个存储块中选择的空闲存储块分配给所述第一流,以写入所述目标数据。

4. 根据权利要求2所述的方法,其中,将所述目标数据写入所述存储块还包括:在确定已经被分配给所述第一流的所述第一存储块中的存储空间不足以存储所述目标数据的情况下,将从所述多个存储块中选择的空闲存储块分配给所述第一流。

5. 根据权利要求1所述的方法,还包括:

在将所述目标数据分配给从所述多个流中选择的所述流之前,训练所述数据分类器。

6. 根据权利要求1所述的方法,还包括:

执行对所述数据分类器的实时训练。

7. 根据权利要求6所述的方法,还包括:

基于所述数据分类器的所述实时训练来改变所述多个流的数目。

8. 一种存储设备,包括:

多个存储块;以及

存储控制器,所述存储控制器被配置为:

接收数据写入请求和要写入的目标数据;

使用数据分类器将所述目标数据分配给从多个流中选择一个流,所述数据分类器被配置为基于所述目标数据的数据属性从所述多个流中选择所述流,所述目标数据的数据属性与所述目标数据的规律性相关联,并且所述目标数据的规律性与所述目标数据中的数据位相关联,以及

将所述目标数据写入从所述多个存储块中选择一个存储块中,其中,所述存储块已

经被分配给了所述目标数据已经被分配给的所述流,其中,所述多个流中的各个流表示所述目标数据的规律性,并且

其中,将所述目标数据分配给从所述多个流中选择一个流包括:

在确定所述目标数据的多个数据位具有规律性的情况下,将所述目标数据分配给从所述多个流中选择的第一流,以及

在确定所述目标数据的所述多个数据位不具有规律性的情况下,将所述目标数据分配给从所述多个流中选择的第二流。

9. 根据权利要求8所述的存储设备,其中,  
所述存储控制器包括至少一个处理器,并且  
所述数据分类器由所述至少一个处理器驱动。

10. 根据权利要求8所述的存储设备,还包括:  
至少一个处理器,所述至少一个处理器位于所述存储控制器外部,并且  
所述数据分类器由所述至少一个处理器驱动。

11. 根据权利要求8所述的存储设备,其中,所述数据分类器由位于所述存储设备外部的至少一个处理器驱动。

12. 根据权利要求8所述的存储设备,还包括:

缓冲存储器,所述缓冲存储器被配置为在所述目标数据被写入所述存储块之前临时存储所述目标数据。

13. 一种存储系统,包括:

主机,所述主机被配置为提供数据写入请求和目标数据;以及  
存储设备,所述存储设备包括多个存储块和存储控制器,所述存储控制器被配置为基于所述数据写入请求将所述目标数据写入从所述多个存储块中选择一个存储块,

其中,所述存储控制器包括数据分类器,所述数据分类器被配置为基于所述目标数据的数据属性从多个流中选择一个流,所述目标数据的数据属性与所述目标数据的规律性相关联,并且所述目标数据的规律性与所述目标数据中的数据位相关联,

其中,所述存储控制器被配置为:

接收所述数据写入请求和所述目标数据;

使用所述数据分类器将所述目标数据分配给从所述多个流中选择的所述流;以及

将所述目标数据写入所述存储块,其中,所述存储块已经被分配给了所述目标数据已经被分配给的所述流,

其中,所述多个流中的各个流表示所述目标数据的规律性,并且

其中,将所述目标数据分配给从所述多个流中选择的所述流包括:

在确定所述目标数据的多个数据位具有规律性的情况下,将所述目标数据分配给从所述多个流中选择的第一流,以及

在确定所述目标数据的所述多个数据位不具有规律性的情况下,将所述目标数据分配给从所述多个流中选择的第二流。

14. 根据权利要求13所述的存储系统,其中,将所述目标数据分配给从所述多个流中选择的所述流并将所述目标数据写入所述存储块是独立于所述主机的控制来执行的。

## 操作存储设备的方法、存储设备及包括其的存储系统

[0001] 相关申请的交叉引用

[0002] 本申请要求2018年11月12日在韩国知识产权局(KIPO)提交的韩国专利申请No.10-2018-0138077的优先权,通过引用将该申请的全部内容并入本文。

### 技术领域

[0003] 本公开总体上涉及半导体集成电路,包括操作存储设备的方法、执行该方法的存储设备和/或包括该存储设备的存储系统。

### 背景技术

[0004] 某些类型的数据存储设备包括一个或更多个半导体存储器件。这些数据存储设备的示例包括固态硬盘(SSD)。与硬盘驱动器(HDD)相比,这些类型的数据存储设备可以有各种设计和/或性能优势。潜在优势的示例包括不存在移动的机械部件、更高的数据访问速度、稳定性、耐用性和/或低功耗。这些类型的数据存储设备可以包括非易失性存储器件(例如,闪存存储器)。在SSD的闪存存储器中,可以在写入操作之前执行擦除操作。随着数据被连续写入闪存存储器,数据可能分散在整个闪存存储器中。近来,研究人员正在进行关于获得可向其写入数据的空闲存储块或存储空间和/或增加存储设备的预期寿命的技术的各种研究项目。

### 发明内容

[0005] 一些示例实施例可以包括一种操作存储设备的方法,所述存储设备包括存储器件的多个存储块,其中,所述方法包括:接收数据写入请求和要写入的目标数据;使用数据分类器将所述目标数据分配给从多个流中选择一个流,所述数据分类器被配置为基于所述目标数据的数据属性从所述多个流中选择所述流;以及将所述目标数据写入从所述多个存储块中选择一个存储块中,其中,所述存储块已经被分配给了所述目标数据已经被分配给的所述流。

[0006] 一些示例实施例可以包括存储设备,所述存储设备包括多个存储块和存储控制器,所述存储控制器被配置为:接收数据写入请求和要写入的目标数据;使用数据分类器将所述目标数据分配给从多个流中选择一个流,所述数据分类器被配置为基于所述目标数据的数据属性从所述多个流中选择所述流,以及将所述目标数据写入从所述多个存储块中选择一个存储块中,其中,所述存储块已经被分配给了所述目标数据已经被分配给的所述流。

[0007] 一些示例实施例可以包括存储系统,所述存储系统包括:主机,所述主机被配置为提供数据写入请求和目标数据;以及存储设备,所述存储设备包括多个存储块和存储控制器,所述存储控制器被配置为基于所述数据写入请求将所述目标数据写入从所述多个存储块中选择一个存储块,其中,所述存储控制器包括数据分类器,所述数据分类器被配置为基于所述目标数据的数据属性从多个流中选择一个流,并且其中,所述存储控制器被配置

为:接收所述数据写入请求和所述目标数据;使用所述数据分类器将所述目标数据分配给从所述多个流中选择的所述流;以及将所述目标数据写入所述存储块,其中,所述存储块已经被分配给了所述目标数据已经被分配给的所述流。

[0008] 在一些示例实施例中,即使主机不支持多流方案,存储设备也可以控制多流方案。可以减少无用数据收集的开销,可以减少写入延迟和写入放大因子(WAF),可以增加写入吞吐量,和/或存储设备可以具有更高的性能和/或预期寿命。

#### 附图说明

[0009] 从以下结合附图的详细描述可以理解一些说明性的、非限制性的示例实施例。

[0010] 图1是示出根据一些示例实施例的操作存储设备的方法的流程图。

[0011] 图2是示出根据一些示例实施例的存储设备和包括该存储设备的存储系统的框图。

[0012] 图3是示出根据一些示例实施例的存储设备中所包括的存储控制器的示例的框图。

[0013] 图4是用于描述根据一些示例实施例的操作存储设备的方法的示图。

[0014] 图5A和图5B是用于描述根据一些示例实施例的存储设备中所包括的神经网络模型的示例的示图。

[0015] 图6是示出根据一些示例实施例的图1中的将目标数据分配给多个流中的一个流的示例的流程图。

[0016] 图7是用于描述根据一些示例实施例的图6的将目标数据分配给多个流中的一个流的操作的示图。

[0017] 图8是示出根据一些示例实施例的图1中的将目标数据分配给多个流中的一个流的另一示例的流程图。

[0018] 图9A和图9B是用于描述根据一些示例实施例的图8的将目标数据分配给多个流中的一个流的操作的示图。

[0019] 图10是示出根据一些示例实施例的图1中的将目标数据分配给多个流中的一个流的又一示例的流程图。

[0020] 图11是用于描述根据一些示例实施例的图10的将目标数据分配给多个流中的一个流的操作的示图。

[0021] 图12是示出根据一些示例实施例的图1中的将目标数据分配给多个流中的一个流的又一示例的流程图。

[0022] 图13是用于描述根据一些示例实施例的图12的将目标数据分配给多个流中的一个流的操作的示图。

[0023] 图14、图15和图16是示出根据一些示例实施例的图1中的将目标数据写入与为分配目标数据而选择的流相对应的存储块的示例的流程图。

[0024] 图17是示出根据一些示例实施例的操作存储设备的方法的性能的示图。

[0025] 图18是示出根据一些示例实施例的操作存储设备的方法的流程图。

[0026] 图19和图20是示出根据一些示例实施例的图18中的附加地实时训练数据分类器的示例的流程图。

[0027] 图21和图22是示出根据一些示例实施例的存储设备的框图和包括该存储设备的存储系统的框图。

[0028] 图23是示出根据一些示例实施例的存储设备中所包括的非易失性存储器件的示例的框图。

[0029] 图24是示出根据一些示例实施例的电子系统的框图。

### 具体实施方式

[0030] 可以参考示出了一些示例实施例的附图更全面地描述各种示例实施例。然而,本公开可以以许多不同的形式来实施,而限于本文阐述的示例实施例。在本申请中,相同的附图标记表示相同的元件。

[0031] 图1是示出根据一些示例实施例的操作存储设备的方法的流程图。

[0032] 在一些示例实施例中,这种存储设备可以包括多个存储块,并且独立于主机的控制以多流方案操作。将参考图2等来描述该存储设备的配置和包括该存储设备的存储系统的配置。

[0033] 在根据一些示例实施例的操作存储设备的方法中,可以训练和/或学习数据分类器(例如,神经网络模型(NNM))(操作S100)。可以预训练和/或预学习数据分类器,以基于多个流中的哪些流与多个数据相关联,将多个数据写入多个存储块中。例如,可以在存储设备外部预训练和/或预学习数据分类器,然后将数据分类器嵌入存储设备中。

[0034] (例如从外部设备)接收数据写入请求和要写入的目标数据(操作S200)。例如,可以从外部主机提供和/或由存储控制器接收数据写入请求和目标数据。

[0035] 使用数据分类器,通过分析目标数据的数据属性,选择多个流中的一个流来分配目标数据。也就是说,数据分类器通过分析目标数据的属性将目标数据分配给多个流中的一个流。(操作S300)。

[0036] 在一些示例实施例中,如参考图6和图7所描述的,用于选择流的数据属性(例如用于分配目标数据的数据属性)可以与表示数据的变化周期的有效期相关联,各个流可以表示目标数据的有效期或变化周期。在其他示例实施例中,如参考图8至图13所描述的,用于选择流的数据属性可以表示目标数据的规律性;和/或用于分配目标数据的数据属性可以与数据的规律性相关联。在其他示例实施例中,用于选择流的数据属性(例如,用于分配目标数据的数据属性)可以与目标数据的包括以下各种属性中的至少一个相关联:诸如可以由文件系统和/或存储器系统存储的元数据,和/或可以由操作系统确定的使用信息。

[0037] 基于为分配目标数据而选择的流,目标数据被写入与为分配目标数据而选择的流相对应的存储块中,从而将数据以单个流写入单个存储块中(操作S400)。在一些示例实施例中,具有相似有效期或变化周期的目标数据可以被分配给一个流和/或可以被存储在被分配给这个流的存储块中。在其他示例实施例中,具有相似规律性的目标数据可以被分配给一个流和/或可以被存储在被分配给这个流的存储块中。在其他示例实施例中,具有相似属性的目标数据可以被分配给一个流和/或可以被存储在被分配给这个流的存储块中。

[0038] 根据一些示例实施例的存储设备可以被配置为基于多流方案操作,在多流方案中,写入多个存储块的多个数据被分类到多个流中和/或由多个流来管理,使得存储设备将与相同的流相关联的数据写入同一个存储块中。在一些示例实施例中,包括在存储设备中

的一个存储块可以仅存储分配给相同的流的目标数据。在其他示例实施例中,分配给相同的流的目标数据可以存储在多个存储块中的被分配给相同流的不同存储块中。

[0039] 与由主机控制和/或驱动和/或取决于主机的多流方案不同,根据一些示例实施例的存储设备可以基于根据一些示例实施例的操作存储设备的方法,使用数据分类器为要写入的数据选择流,和/或可以使用自分配流以多流方案操作。也就是说,图1中的操作S300和/或S400可以独立于主机的控制来执行(例如,不管主机如何),即使主机不支持多流方案,存储设备也可以控制多流方案。因此,可以减少无用数据收集的开销,可以减少写入等待时间和/或写入放大因子(WAF)(例如,与由主机闪存控制器写入的数据量相比,可以减少由固态存储控制器写入的数据量),可以增加写入吞吐量,和/或存储设备可以具有更高的性能和/或预期寿命。

[0040] 图2是示出根据一些示例实施例的存储设备和包括该存储设备的存储系统的框图。

[0041] 参考图2,存储系统100包括主机200和存储设备300。

[0042] 主机200控制存储系统100的整体操作。尽管在图2中未示出,但是主机200可以包括主机处理器和/或主机存储器。主机处理器可以控制主机的操作。例如,主机处理器可以执行操作系统(OS)。主机存储器可以存储由主机处理器执行的指令和/或处理的数据。

[0043] 主机200向存储设备300提供数据写入请求WREQ和要写入的目标数据WDAT。例如,尽管在图2中未示出,但是由主机处理器执行的操作系统可以包括用于文件管理的文件系统和/或用于在操作系统级别控制包括存储设备300的外围设备的设备驱动程序。主机处理器执行的应用和/或应用程序的文件存储请求和/或文件系统的编程请求可以指示数据写入请求WREQ。

[0044] 存储设备300由主机200访问,和/或独立于主机200的控制以多流方案操作。存储设备300可以包括存储控制器310、多个非易失性存储器320a、320b和/或320c、和/或缓冲存储器330。

[0045] 存储控制器310可以基于从主机200接收到的命令和数据来控制存储设备300的操作(例如,数据写入操作)。

[0046] 存储控制器310可以基于参考图1描述的方法来操作。例如,存储控制器310包括神经网络模型NNM,该神经网络模型NNM被预训练,以基于多个流中的哪些流与多个数据相关联,将多个数据写入到包括在多个非易失性存储器320a、320b和/或320c中的多个存储块中。存储控制器310从主机200接收数据写入请求WREQ和要写入的目标数据WDAT,使用神经网络模型NNM通过分析目标数据WDAT来为分配目标数据WDAT选择多个流中的一个流,和/或将目标数据WDAT写入与为分配目标数据WDAT而选择的流相对应的存储块中,从而以单个流将数据写入单个存储块中。也就是说,存储控制器310可以控制存储设备300,使得存储设备300以多流方案操作。

[0047] 多个非易失性存储器320a、320b和/或320c可以针对每个流分别存储多个数据。例如,多个非易失性存储器320a、320b和/或320c可以存储目标数据WDAT、各种用户数据等。如参考图23所描述的,多个非易失性存储器320a、320b和/或320c均可以包括多个存储块。

[0048] 在一些示例实施例中,多个非易失性存储器320a、320b和/或320c均可以包括NAND闪存存储器。在其他示例实施例中,多个非易失性存储器320a、320b和/或320c均可以包括

以下各项之一:电可擦除可编程只读存储器 (EEPROM)、相变随机存取存储器 (PRAM)、电阻随机存取存储器 (RRAM)、纳米浮栅存储器 (NFGM)、聚合物随机存取存储器 (PoRAM)、磁随机存取存储器 (MRAM)、铁电随机存取存储器 (FRAM) 等。

[0049] 缓冲存储器330可以存储由存储控制器310执行的指令和/或处理的数据,和/或可以临时存储被存储在和/或将存储在多个非易失性存储器320a、320b和/或320c中的数据。例如,缓冲存储器330可以包括各种易失性存储器(例如动态随机存取存储器 (DRAM)、静态随机存取存储器 (SRAM) 等)中的至少一种。

[0050] 在一些示例实施例中,存储设备300可以是固态硬盘 (SSD)。在其他示例实施例中,存储设备300可以是以下各项之一:通用闪存 (UFS)、多媒体卡 (MMC)、嵌入式多媒体卡 (eMMC)、安全数字卡 (SD卡)、微型SD卡、记忆棒、芯片卡、通用串行总线 (USB) 卡、智能卡、紧凑型闪存 (CF) 卡等。

[0051] 在一些示例实施例中,存储设备300可以通过块可访问接口连接到主机200,该块可访问接口可以包括例如UFS、eMMC、串行高级技术附件 (SATA) 总线、高速非易失性存储器 (NVMe) 总线、串行连接SCSI (SAS) 总线等。存储设备300可以使用与多个非易失性存储器320a、320b和/或320c的存取容量相对应的块可访问地址空间来向主机200提供块可访问接口,以允许以存储块为单位对存储在多个非易失性存储器320a、320b和/或320c中的数据进行访问。

[0052] 在一些示例实施例中,存储系统100可以是诸如以下各项的任何移动系统:移动电话、智能电话、平板计算机、膝上型计算机、个人数字助理 (PDA)、便携式多媒体播放器 (PMP)、数码相机、便携式游戏机、音乐播放器、摄像机、视频播放器、导航设备、可穿戴设备、物联网 (IoT) 设备、万物互联网 (IoE) 设备、电子书阅读器、虚拟现实 (VR) 设备、增强现实 (AR) 设备和/或机器人设备。在其他示例实施例中,存储系统100可以是诸如个人计算机 (PC)、服务器计算机、工作站、数字电视、机顶盒和/或导航系统的任何计算系统。

[0053] 图3是示出根据一些示例实施例的存储设备中所包括的存储控制器的示例实施例的框图。

[0054] 参考图3,存储控制器310可以包括至少一个处理器311、数据分类器313、纠错码 (ECC) 块315、主机接口317和/或存储器接口319。

[0055] 处理器311可以响应于经由主机接口317从主机(例如,图2中的主机200)接收到的命令来控制存储控制器310的操作。在一些示例实施例中,处理器311可以通过采用用于操作存储设备(例如,图2中的存储设备300)的固件来控制各个组件。

[0056] 在一些示例实施例中,可以预训练数据分类器313(例如,神经网络模型NNM)。根据一些示例实施例,数据分类器313可以通过分析要写入的目标数据(例如,图2中的目标数据WDAT)和/或对目标数据进行分类,来将目标数据分配给流,以使存储设备300以多流方案操作。

[0057] 在一些示例实施例中,数据分类器313的至少一部分可以实现为硬件。例如,如图3所示,处理器311和数据分类器313可以实现为两个单独的硬件。又例如,尽管未在图3中示出,但是处理器311和数据分类器313可以实现为一个硬件。

[0058] 在其他示例实施例中,数据分类器313的至少一部分可以实现为指令代码和/或程序例程(例如软件程序)和/或可以存储在存储器(例如图2中的非易失性存储器320a、320b

和/或320c)中。例如,数据分类器313和/或包括在数据分类器313中的神经网络模型NNM可以由处理器311执行和/或驱动。

[0059] 在一些示例实施例中,用于纠错的ECC块315可以使用诸如以下纠错码来执行编码调制: Bose-Chaudhuri-Hocquenghem (BCH) 码、低密度奇偶校验 (LDPC) 码、turbo码、里所 (Reed-Solomon) 码、卷积码、递归系统卷积 (RSC) 码、网格编码调制 (TCM) 和/或块编码调制 (BCM); 和/或可以使用上述代码和/或其他纠错码来执行ECC编码和/或ECC解码。

[0060] 在一些示例实施例中,主机接口317可以提供主机200与存储设备300之间的物理连接。主机接口317可以提供与主机的总线格式相对应的接口,以用于主机200与存储设备300之间的通信。在一些示例实施例中,主机200的总线格式可以是小型计算机系统接口 (SCSI) 和/或串行连接SCSI (SAS) 接口。在其他示例实施例中,主机200可以利用诸如USB、高速外围组件互连 (PCI) (PCIe)、高级技术附件 (ATA)、并行ATA (PATA)、串行ATA (SATA) 和/或高速非易失性存储器 (NVM) (NVMe) 的总线格式。

[0061] 存储器接口319可以与非易失性存储器 (例如,图2中的非易失性存储器320a、320b和/或320c) 交换数据。存储器接口319可以向非易失性存储器320a、320b和/或320c传送数据,和/或可以接收从非易失性存储器320a、320b和/或320c读取的数据。在一些示例实施例中,存储器接口319可以经由一个信道连接到非易失性存储器320a、320b和/或320c。在其他示例实施例中,存储器接口319可以经由两个以上的信道连接到非易失性存储器320a、320b和/或320c。

[0062] 图4是用于描述根据一些示例实施例的操作存储设备的方法的示图。图4示出了图2中存储设备300的软件分层结构和存储空间结构。

[0063] 参考图4,存储设备300可以包括主机接口层 (HIL) 342、数据分类器344、闪存转换层 (FTL) 346、至少一个非易失性存储器 (NVM) 320和缓冲存储器330。

[0064] 主机接口层342可以接收从主机 (例如,图2中的主机200) 提供的各种请求、信号和/或数据。闪存转换层346可以将主机200提供的逻辑地址转换成非易失性存储器320中的物理地址,和/或可以管理存储在非易失性存储器320中的数据。

[0065] 图4中的非易失性存储器320和/或缓冲存储器330可以分别与图2中的非易失性存储器320a、320b和/或320c和/或缓冲存储器330基本相同。例如,非易失性存储器320可以包括多个存储块MB1、MB2、MB3……,和/或多个存储块MB1、MB2、MB3……均可以只存储相同的流的数据。例如,存储块MB1可以存储分配给第一流的数据D1、D3和/或D6,存储块MB2可以存储分配给第二流的数据D5、D7和/或D8,存储块MB3可以存储分配给第三流的数据D2、D4和/或D9。为了便于说明,分配给第一流的数据D1、D3和D6用矩形表示,分配给第二流的数据D5、D7和D8用圆形表示,分配给第三流的数据D2、D4和D9用三角形表示。在一些示例实施例中,缓冲存储器330可以临时存储被存储在和/或将存储在非易失性存储器320中的数据。

[0066] 在一些示例实施例中,图4中的数据分类器344可以与图3中的数据分类器313基本相同,和/或可以包括神经网络模型NNM。图4中的神经网络模型NNM可以与图2和/或图3中的神经网络模型NNM基本相同。在其他示例实施例中,图4中的数据分类器344可以不同于图3中的数据分类器313。在一些示例实施例中,可以预训练数据分类器344 (例如,神经网络模型NNM),以基于多个流中的哪个流与多个数据相关联,来将多个数据写入多个存储块MB1、MB2、MB3……中。例如,神经网络模型NNM可以是分类模型,神经网络模型NNM的输入可以是

原始数据(例如,目标数据WDAT),和/或神经网络模型NNM的输出可以是目标数据WDAT将被分配给的流的流ID SID(例如,数值)。原始数据(例如,目标数据WDAT)的至少一部分可以是神经网络模型NNM的输入。

[0067] 参照图4描述图1的操作存储设备的方法。

[0068] 首先,主机接口层342可以接收从主机200提供的数据写入请求WREQ和要写入的目标数据WDAT。目标数据WDAT可以在被提供给数据分类器344和/或闪存转换层346之前临时存储在缓冲存储器330中。

[0069] 在一些示例实施例中,数据分类器344可以从缓冲存储器330接收目标数据WDAT,和/或在没有任何其他信息的情况下,可以使用神经网络模型NNM分析作为原始数据的目标数据WDAT和/或对作为原始数据的目标数据WDAT进行分类,以预测、选择对应的流。

[0070] 在一些示例实施例中,数据分类器344可以包括一个或更多个神经网络模型NNM,该一个或更多个神经网络模型NNM可以被配置为例如单层感知器(SLP)、多层感知器(MLP)、卷积神经网络(CNN)、深层神经网络(DNN)、递归神经网络(RNN)等。在其他示例实施例中,数据分类器344可以利用其他分类模型、数据结构和/或自适应计算技术,包括(但不限于)贝叶斯统计分类器、遗传进化算法、模糊逻辑系统和诸如决策树的基于规则的和/或启发式分类器或它们的组合。

[0071] 闪存转换层346可以接收目标数据WDAT和/或由数据分类器344选择的流,和/或可以基于目标数据WDAT和/或流执行数据写入操作(例如,NAND操作)。例如,当数据分类器将目标数据分配给第一流时,可以将目标数据WDAT写入存储块MB1,当数据分类器将目标数据分配给第二流时,可以将目标数据WDAT写入存储块MB2,当数据分类器将目标数据分配给第三流时,可以将目标数据WDAT写入存储块MB3。

[0072] 在一些示例实施例中,当存储设备300以多流方案操作和/或被驱动时,主机200可以使用诸如文件信息、处理信息和/或文件系统信息的各种信息将流与对应信息之间的关系形成到流映射表中,和/或存储设备300可以通过使用神经网络模型NNM分析目标数据WDAT和/或对目标数据WDAT进行分类来选择流。在其他示例实施例中,主机200和/或存储设备300可以不包括流映射表。

[0073] 图5A和图5B是用于描述根据一些示例实施例的存储设备中所包括的神经网络模型的示例的示意图。

[0074] 参考图5A,通用神经网络可以包括输入层IL、多个隐藏层HL1、HL2、……、HLn和输出层OL。

[0075] 输入层IL可以包括i个输入节点 $x_1, x_2, \dots, x_i$ ,其中i是自然数。长度为i的输入数据(例如矢量输入数据)IDAT可以被输入到输入节点 $x_1, x_2, \dots, x_i$ ,使得输入数据IDAT的每个元素被输入到输入节点 $x_1, x_2, \dots, x_i$ 之一。

[0076] 多个隐藏层HL1、HL2、……、HLn可以包括n个隐藏层,其中n是自然数,并且可以包括多个隐藏节点 $h^1_1, h^1_2, h^1_3, \dots, h^1_m, h^2_1, h^2_2, h^2_3, \dots, h^2_m, \dots, h^n_1, h^n_2, h^n_3, \dots, h^n_m$ 。例如,隐藏层HL1可以包括m个隐藏节点 $h^1_1, h^1_2, h^1_3, \dots, h^1_m$ ,隐藏层HL2可以包括m个隐藏节点 $h^2_1, h^2_2, h^2_3, \dots, h^2_m$ ,隐藏层HLn可以包括m个隐藏节点 $h^n_1, h^n_2, h^n_3, \dots, h^n_m$ ,其中m是自然数。

[0077] 输出层OL可以包括j个输出节点 $y_1, y_2, \dots, y_j$ ,其中j是自然数。输出节点 $y_1,$

$y_2, \dots, y_j$ 中的每一个输出节点可以对应于待分类的类别中的一类。输出层OL可以为每个类输出与输入数据IDAT相关联的值(例如,类分数或简单的分数)。输出层OL可以被称作完全连接层,并且可以指示例如输入数据IDAT对应于汽车的概率。

[0078] 图5A中示出的神经网络的结构可以由关于被示出为的节点之间的分支(或连接)的信息以及分配给每个分支的加权值(未示出)来表示。一层内的节点彼此不连接,但不同层的节点可以全部或部分地彼此连接。

[0079] 每个节点(例如,节点 $h^1_j$ )可以接收前一个节点(例如,节点 $x_1$ )的输出,可以对接收的输出执行运算操作、估算和/或计算,和/或可以将运算操作、估算和/或计算的结果作为输出输出到下一个节点(例如,节点 $h^2_j$ )。每个节点可以通过将输入应用于特定函数(例如,非线性函数)来计算要输出的值。

[0080] 通常,预先设置神经网络的结构,和/或使用具有数据所属类的确定答案的数据适当地设置节点之间的连接的加权值。具有确定答案的数据称为“训练数据”,确定加权值的过程称为“训练”。神经网络在训练过程中“学习”。一组可独立训练的结构和加权值称为“模型”,通过具有所设置的加权值的模型预测输入数据所属类,然后输出预测值的过程称为“测试”过程。

[0081] 在一些示例实施例中,诸如SLP和/或MLP的神经网络可以具有如图5A所示的结构。

[0082] 诸如图5A所示的一些神经网络可能不适于处理输入图像数据(或输入声音数据),因为每个节点(例如,节点 $h^1_j$ )连接到前一层的所有节点(例如包括在层IL中的节点 $x_1, x_2, \dots, x_i$ ),然后加权值的数目随着输入图像数据的大小的增加而急剧增加。因此,已经研究出了通过将滤波技术与通用神经网络相结合而实现的卷积神经网络,使得卷积神经网络有效地训练二维图像(例如,输入图像数据)。

[0083] 参考图5B,卷积神经网络可以包括多个层CONV1、RELU1、CONV2、RELU2、POOL1、CONV3、RELU3、CONV4、RELU4、POOL2、CONV5、RELU5、CONV6、RELU6、POOL3和/或FC。

[0084] 与通用神经网络不同,卷积神经网络的每一层可以具有宽度、高度和/或深度的三个维度,因此输入到每一层的数据可以是具有宽度、高度和/或深度的三个维度的体数据。例如,如果图5B中的输入图像的大小为32个宽度(例如,32个像素)和32个高度以及三个颜色通道R、G和B,则对应于输入图像的输入数据IDAT的大小可以为 $32*32*3$ 。图5B中的输入数据IDAT可以被称作输入体数据(input volume data)和/或输入激活体(input activation volume)。

[0085] 卷积层CONV1、CONV2、CONV3、CONV4、CONV5和/或CONV6可以对输入体数据执行卷积运算。在图像处理中,卷积运算表示基于具有加权值的掩码来处理图像数据的运算,和/或通过输入值与加权值相乘和/或将相乘后的所有值进行累加来获得输出值的运算。掩码可以称为过滤器、窗口或内核。

[0086] 在一些示例实施例中,每个卷积层的参数可以包括一组可学习的滤波器。每个滤波器在空间上(沿宽度和/或高度)可以很小,但是可以贯穿输入体数据的整个深度。例如,在前向传递期间,每个滤波器可以在输入体数据的宽度和/或高度上滑动(更精确地说,卷积),并且可以在任何位置计算滤波器的条目与输入之间的点积。当过滤器在输入体数据的宽度和/或高度上滑动时,可以生成给出过滤器在每个空间位置的响应的二维激活图(activation map)。结果,可以通过沿着深度维度堆叠这些激活图来生成输出体数据。例

如,如果大小为 $32*32*3$ 的输入体数据通过具有四个补零滤波器的卷积层CONV1,卷积层CONV1的输出体数据的大小可以为 $32*32*12$ (例如,体数据的深度增加)。

[0087] RELU层RELU1、RELU2、RELU3、RELU4、RELU5和/或RELU6可以执行修正线性单元操作,这种操作对应于例如由函数 $f(x) = \max(0, x)$ (例如,对于所有负输入 $x$ ,输出为零)定义的激活函数。例如,如果大小为 $32*32*12$ 的输入体数据通过RELU层RELU1以执行修正线性单元操作,则RELU层RELU1的输出体数据的大小可以为 $32*32*12$ (例如,保持体数据的大小)。

[0088] 池化层POOL1、POOL2和/或POOL3可以沿着宽度和/或高度的空间维度对输入体数据执行下采样操作。例如,以 $2 \times 2$ 矩阵形式排列的四个输入值可以基于 $2 \times 2$ 滤波器被转换成一个输出值。例如,可以基于 $2 \times 2$ 最大池化来选择以 $2 \times 2$ 矩阵形式排列的四个输入值的最大值,或者可以基于 $2 \times 2$ 平均池化来获得以 $2 \times 2$ 矩阵形式排列的四个输入值的平均值。例如,如果大小为 $32*32*12$ 的输入体数据通过具有 $2*2$ 过滤器的池化层POOL1,汇集层POOL1的输出体数据的大小可以为 $16*16*12$ (例如,体数据的宽度和/或高度减小,和/或保持体数据的深度)。

[0089] 一些示例实施例可以包括这样的神经网络:一个卷积层(例如,CONV1)和/或一个RELU层(例如,RELU1)可以形成卷积神经网络中的一对CONV/RELU层,成对的CONV/RELU层可以重复地布置在卷积神经网络中,和/或池化层可以规律地(periodically)插入卷积神经网络中,以减小图像的空间大小和/或提取图像的特征。

[0090] 在一些示例实施例中,输出层和/或全连接层FC可以为每个类输出输入体数据IDAT的结果(例如,类分数)。例如,随着重复执行卷积操作和/或下采样操作,对应于二维图像的输入体数据IDAT可以被转换成一维矩阵或向量。例如,全连接层FC可以表示输入体数据IDAT对应于汽车的概率、对应于卡车的概率、对应于飞机的概率、对应于轮船的概率和对应于马的概率。

[0091] 在一些示例实施例中,包括在卷积神经网络中的层的类型和数目可以不限于参考图5B描述的示例,而是可以根据一些示例实施例进行改变。此外,尽管未在图5B中示出,但是卷积神经网络还可以包括诸如用于将对应于预测结果的分数值转换成概率值的softmax层、用于添加至少一个偏置的偏置添加层等的其他层。在一些示例实施例中,除了CNN之外,诸如DNN的神经网络可以具有如图5B所示的结构。

[0092] 鉴于对诸如神经网络之类的数据分类器的前述讨论,根据一些示例实施例,以下描述呈现了诸如神经网络之类的数据分类器的一些用途,以将目标数据分配给流和/或存储块。

[0093] 图6是示出图1中的将目标数据分配给多个流之一的示例的流程图。图7是用于描述图6的将目标数据分配给多个流之一的操作的示图。

[0094] 参考图1、图6和图7,在一些示例实施例中,当将目标数据分配多个流中的一个流时(操作S300),多个流可以表示有效期,该有效期是写入多个存储块中的多个数据的变化周期。也就是说,用于选择流的数据属性(例如用于分配目标数据的数据属性)可以与表示数据变化周期的有效期相关联。

[0095] 例如,当分析出目标数据的有效期或变化周期短于第一参考时间时(操作S311:是),可以从多个流中选择第一流来分配目标数据(操作S313)。当分析出目标数据的有效期或变化周期长于或等于第一参考时间时(操作S311:否),可以从多个流中选择第二流来分

配目标数据(操作S315)。

[0096] 在图7中,示出了具有三种不同类型的有效期的数据的变化周期或根据时间推移的变化周期。第一组数据DG1可以对应于第一有效期等级,和/或可以以时间间隔 $t_1$ 、 $t_2$ 、 $t_3$ 和 $t_4$ 被创建、修改和/或删除。第二组数据DG2可以对应于第二有效期等级,和/或可以以时间间隔 $t_5$ 、 $t_6$ 和 $t_7$ 被创建、修改和/或删除,时间间隔 $t_5$ 、 $t_6$ 和 $t_7$ 中的每一个时间间隔可以长于 $t_1$ 、 $t_2$ 、 $t_3$ 和 $t_4$ 中的每一个时间间隔。第三组数据DG3可以对应于第三有效期等级,和/或可以以时间间隔 $t_8$ 被创建、修改和/或删除,时间间隔 $t_8$ 可以长于时间间隔 $t_5$ 、 $t_6$ 和 $t_7$ 中的每一个时间间隔。第二组数据DG2可以具有比第一组数据DG1长的变化周期,第三组数据DG3可以具有比第二组数据DG2长的变化周期。

[0097] 在一些示例实施例中,指示数据的变化周期的有效期可以由有效期等级表示,该有效期等级是通过根据变化周期的长度将变化周期分类成 $N$ 个等级( $N$ 是自然数)而产生的。因此,具有相似有效期但没有完全相同的变化周期的数据段可以被分类成一组。例如,图7示出了这样的示例: $N=3$ ,每1小时改变的数据和每1小时10分钟改变的数据可以被分类为第一等级,每4小时改变的数据和每4小时40分钟改变的数据可以被分类为第二等级,每10小时改变的数据可以被分类为第三等级。作为分类结果,可以执行等级比较,使得第一等级的数据的有效期限比第二等级的数据的有效期限短和/或第三等级的数据的有效期限比第一等级的数据的有效期限长。

[0098] 图8是示出图1中的将目标数据分配给多个流中的一个流的另一示例的流程图。图9A和图9B是用于描述图8的将目标数据分配给多个流中的一个流的操作的示图。

[0099] 参考图1、图8、图9A和图9B,在一些示例实施例中,当将目标数据分配给多个流中的一个流时(操作S300),多个流可以表示写入多个存储块中的多个数据的规律性。也就是说,用于将目标数据分配给流的数据属性(例如,用于对目标数据进行分类的数据属性)可以与数据的规律性相关联。

[0100] 例如,当分析出包括在目标数据中的多个数据位具有至少一种规律性(例如,具有可识别的、期望的和/或预定的规律性)时(操作S321:是),可以从多个流中选择第一流来分配目标数据(操作S323)。例如,包括在目标数据中的多个数据位可以具有至少满足规律性阈值的规律性。当分析出包括在目标数据中的多个数据位具有随机性和/或不具有可识别的、期望的和/或预定的规律性(例如,不具有至少满足规律性阈值的规律性)时(操作S321:否),可以从多个流中选择第二流来分配目标数据(操作S325)。

[0101] 如图9A所示,数据DAT1可以包括多个数据位B10、B11、B12、B13、B14、B15、B16、B17、B18、B19、B1A、B1B、B1C、B1D、B1E和B1F。例如,每个数据位可以由十六进制代码表示和/或可以具有不同类型TA、TB、TC、TD和TE中的一种。图9A的数据DAT1可以无规律性地用数据位B10~B1F的不同类型TA~TE随机排列。

[0102] 如图9B所示,数据DAT2可以包括多个数据位B20、B21、B22、B23、B24、B25、B26、B27、B28、B29、B2A、B2B、B2C、B2D、B2E和B2F,每个数据位具有不同类型TA、TB、TC、TD和TE之一。图9B的数据DAT2可以具有规律性REG,该规律性具有以TA、TB、TA和/或TC的类型序列和/或类型顺序重复的模式。例如,在图9B的数据DAT2中,四个数据位B20~B23可以具有TA、TB、TA、TC的类型序列,接下来的四个数据位B24~B27可以具有不同的类型序列,接下来的四个数据位B28~B2B可以又具有TA、TB、TA、TC的类型序列。

[0103] 在一些示例实施例中,当目标数据是图9B的数据DAT2时,神经网络模型NNM可以确定目标数据具有可识别的、期望的和/或预定的规律性,和/或可以将目标数据分配给第一流。当目标数据是图9A的数据DAT1时,神经网络模型NNM可以确定目标数据具有随机性而没有可识别的、期望的和/或预定的规律性,和/或可以将目标数据分配给第二流。

[0104] 图10是示出图1中的将目标数据分配给多个流中的一个流的又一示例的流程图。图11是用于描述图10的将目标数据分配给多个流中的一个流的操作的示图。

[0105] 参考图1、图10和图11,在一些示例实施例中,当将目标数据分配给多个流中的一个流时(操作S300),与图8的示例类似,多个流可以表示写入多个存储块中的多个数据的规律性。

[0106] 例如,当分析出目标数据具有例如可识别的、期望的和/或预定的模式时(操作S331:是),可以从多个流中选择第一流来分配目标数据(操作S333)。当分析出目标数据不具有可识别的、期望的和/或预定的模式时(操作S331:否),可以从多个流中选择第二流来分配目标数据(操作S335)。

[0107] 如图11所示,数据DAT3可以包括多个数据位B30、B31、B32、B33、B34、B35、B36、B37、B38、B39、B3A、B3B、B3C、B3D、B3E和B3F,每个数据位可以是不同类型TA、TB、TC、TD和TE中的一种。图11的数据DAT3可以具有模式PT1,该模式PT1具有类型序列TA、TD、TA和TD。

[0108] 在一些示例实施例中,当目标数据是图11的数据DAT3时,神经网络模型NNM可以确定目标数据具有可识别的、期望的和/或预定的模式,和/或可以将目标数据分配给第一流。当目标数据是图9A的数据DAT1时,神经网络模型NNM可以确定目标数据不具有可识别的、期望的和/或预定的模式,和/或可以将目标数据分配给第二流。

[0109] 图12是示出图1中的将目标数据分配给多个流中的一个流的又一示例的流程图。图13是用于描述图12的将目标数据分配给多个流中的一个流的操作的示图。

[0110] 参考图1、图12和图13,在一些示例实施例中,当将目标数据分配给多个流中的一个流时(操作S300),与图8和图10的示例类似,多个流可以表示写入多个存储块中的多个数据的规律性。

[0111] 例如,当分析出目标数据具有可识别的、期望的和/或预定的第一模式时(操作S341:是),可以选择多个流中的第一流来分配目标数据(操作S343)。当分析出目标数据不具有第一模式时(操作S341:否),如果分析出目标数据具有可识别的、期望的和/或预定的第二模式(操作S345:是),则可以从多个流中选择第二流来分配目标数据(操作S347);如果分析出目标数据既不具有第一模式也不具有第二模式时(操作S341:否以及操作S345:否),则可以选择不同于第一流和第二流的另一个流(例如,第三流)来分配目标数据(操作S349)。

[0112] 如图13所示,数据DAT4可以包括多个数据位B40、B41、B42、B43、B44、B45、B46、B47、B48、B49、B4A、B4B、B4C、B4D、B4E和B4F,每个数据位可以具有不同类型TA、TB、TC、TD和TE中的一种。图13的数据DAT4可以具有模式PT2,该模式PT2具有类型序列TA、TB、TC、TD、TE。

[0113] 在一些示例实施例中,当目标数据是图11的数据DAT3时,神经网络模型NNM可以确定目标数据具有第一模式,和/或可以将目标数据分配给第一流。当目标数据是图13的数据DAT4时,神经网络模型NNM可以确定目标数据具有第二模式,和/或可以将目标数据分配给第二流。在其他情况下,神经网络模型NNM可以将目标数据分配给第三流。

[0114] 尽管图9A、图9B、图11和图13示出了可识别的、期望的和/或预定的规律性和/或可识别的、期望的和/或预定的模式与特定的类型序列和/或类型顺序相关联,但是一些示例实施例不限于此。例如,可识别的、期望的和/或预定的规律性和/或可识别的、期望的和/或预定的模式可以与是否包括特定的十六进制代码、是否包括特定的参数等相关联。

[0115] 尽管图6、图8、图10和图12示出了根据数据属性(例如,目标数据的可识别的、期望的和/或预定的标准)将目标数据分类成两个或三个组,和/或选择两个或三个流中的一个流来分配与其对应的写入数据,但是一些示例实施例不限于此。例如,根据一些示例实施例,可以改变可识别的、期望的和/或预定的数据属性的数目和/或与其对应的流的数目。

[0116] 在一些示例实施例中,目标数据的数据类型可以用作数据属性,通过该数据属性将目标数据分配给多个流中的一个流。例如,与存储设备的操作相关联的日志数据和/或提交数据以及与文件系统的操作相关联的元数据可以被分析为具有相对较短有效期的数据、具有可识别的、期望的和/或预定的规律性的数据和/或具有可识别的、期望的和/或预定的模式的数据。在一些示例实施例中,诸如视频数据、图像数据和/或语音数据的一般用户数据可以被分析为具有相对长有效期的数据、具有随机性的数据和/或没有可识别的、期望的和/或预定的模式的数据。

[0117] 图14、图15和图16是示出图1中的将目标数据写入与为分配目标数据而选择的流相对应的存储块的示例的流程图。

[0118] 参考图1和图14,在一些示例实施例中,当将目标数据写入与为分配目标数据而选择的流相对应的存储块中时(操作S400),如果目标数据被分配给了从多个流中选择的第一流(操作S411:是),则可以将目标数据写入多个存储块中的第一存储块(操作S413)。第一存储块可以对应于第一流。如果目标数据未被分配给第一流(操作S411:否),则目标数据可以被写入多个存储块中除了第一存储块之外的存储块中(操作S415)。也就是说,目标数据可以被写入已经用于第一流的存储块。

[0119] 参考图1和图15,当将目标数据写入与为分配目标数据而选择的流相对应的存储块时(操作S400),图15中的操作S411和S415可以分别与图14中的操作S411和S415基本相同。

[0120] 如果目标数据被分配给第一流(操作S411:是),并且第一存储块已经被分配给了第一流(操作S421:是),则可以将目标数据写入第一存储块(操作S413)。如果目标数据被分配给第一流(操作S411:是),并且第一存储块尚未被分配给第一流(操作S421:否),例如,如果第一流先被分配了和/或如果先写入了分配给第一流的数据,则可以将目标数据写入多个存储块中的空闲存储块(操作S423)。也就是说,存储块可以从空闲状态切换到开启状态,然后可以被使用。

[0121] 参考图1和图16,当将目标数据写入与为分配目标数据而选择的流相对应的存储块时(操作S400),图16中的操作S411和S415可以分别与图14中的操作S411和S415基本相同。

[0122] 如果目标数据被分配给了第一流(操作S411:是),并且对应于第一流的第一存储块中的存储空间足够(操作S431:否),则可以将目标数据写入第一存储块(操作S413)。如果目标数据被分配给了第一流(操作S411:是),并且第一存储块中的存储空间不足(操作S431:是),则可以将目标数据写入多个存储块中的空闲存储块(操作S433)。图16中的操作

S433可以类似于图15中的操作S423。

[0123] 尽管图14、图15和图16示出了基于目标数据被分配给第一流的示例来执行目标数据的写入操作,但是一些示例实施例不限于此。例如,可以以与图14、图15和图16所示方式相同的方式针对所有流中的每一个流执行写入操作。

[0124] 图17是示出根据一些示例实施例的操作存储设备的方法的性能的示图。

[0125] 参考图17,根据一些示例实施例,使用ext4文件系统进行一个实验,以检查包括在根据示例实施例的存储设备中的诸如神经网络模型NNM的数据分类器的数据分类的准确性。

[0126] 包括三个隐藏层的MLP模型被用作神经网络模型NNM。MLP模型中学习参数和/或训练参数的数目约为557827。模型的输入使用从文件的开头(例如,文件的一部分)开始的大约4K(例如,大约4096)大小。模型的输出根据数据类型(例如,日志数据为0,元数据为1,其他用户数据和/或一般数据为2)对目标数据进行分类。作为分类模型学习的结果,训练准确性约为97%,测试准确性约为97%。此外,当将大约373个日志文件、大约397个元文件和414个一般文件用作测试集时,如图17所示,使用学习模型对日志文件的推理准确性CASE1、对元文件的推理准确性CASE2和对通用文件的推理准确性CASE3分别为大约94%、93%和100%。

[0127] 图18是示出根据一些示例实施例的操作存储设备的方法的流程图。将省略与图1重复的描述。

[0128] 参考图18,在根据一些示例实施例的操作存储设备的方法中,图18中的操作S100、S200、S300和S400可以分别与图1中的操作S100、S200、S300和S400基本相同。

[0129] 在一些示例实施例中,可以附加地实时和/或在运行时间期间训练和/或学习诸如神经网络模型的数据分类器(操作S500)。如参考图19和图20所描述的,当附加地实时和/或在运行时间期间训练和/或学习数据分类器时,可以改变多个流的数目。在该示例中,在附加地训练和/或学习操作之后,可以选择更新后的多个流中的一个流来分配接收到的要写入的目标数据。

[0130] 图19和图20是示出在图18中附加地实时训练诸如神经网络模型的数据分类器的示例的流程图。

[0131] 参考图18和图19,当附加地实时训练数据分类器时(操作S500),可以分析存储设备的操作环境以检查是否满足第一条件(操作S511)。第一条件可以包括需要添加特定流的情况。例如,第一条件可以包括:分配给一个流的数据量太大(例如,第一流的大小超过大小阈值)和/或可以用两个以上的流分开管理的情况;有必要另外使用先前不存在的特定格式数据(例如,目标数据具有新的数据属性,数据分类器没有针对该数据属性将目标数据分配给多个流中的任何一个流)的情况等。

[0132] 当存储设备的操作环境满足第一条件时(操作S511:是),可以添加新的流(操作S513),因此可以增加存储设备中使用的流的数目。当存储设备的操作环境不满足第一条件时(操作S511:否),可以保持先前使用的流(操作S515)。

[0133] 参考图18和图20,在一些示例实施例中,当另外地实时训练诸如神经网络模型的数据分类器时(操作S500),可以分析存储设备的操作环境以检查是否满足第二条件(操作S521)。第二条件可以包括需要删除特定流的情况。例如,第二条件可以包括分配给特定流

的数据长时间不被使用(例如,所选择的流在阈值周期内未被选择用于目标数据的分配)的情况、分配给特定流的数据不再被使用(例如,已经被分配给特定流的目标数据在阈值周期内没有被访问)的情况等。

[0134] 当存储设备的操作环境满足第二条件时(操作S521:是),可以删除先前使用的流中的至少一个流(操作S523),因此可以减少存储设备中所使用的流的数目。当存储设备的操作环境不满足第二条件时(操作S521:否),可以保持先前使用的流(操作S525)。

[0135] 如本领域技术人员将理解的,一些发明构思的一些示例实施例可以体现为系统、方法、计算机程序产品和/或包含在具有计算机可读程序代码的一个或多个计算机可读介质中的计算机程序产品。计算机可读程序代码可以被提供给通用计算机的处理器、专用计算机的处理器或其他可编程数据处理设备的处理器。计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质。计算机可读存储介质可以是任何可以包含和/或存储供指令执行系统、装置或设备使用的程序和/或与指令执行系统、装置或设备结合使用的程序的有形介质。例如,计算机可读介质可以是非暂时性计算机可读介质。

[0136] 图21和图22是示出根据一些示例实施例的存储设备和包括该存储设备的存储系统的框图。将省略与图2重复的描述。

[0137] 参考图21,存储系统100a包括主机200和存储设备300a。

[0138] 除了图21中的存储设备300a还包括处理器311之外,图21的存储系统100a可以基本上与图2的存储系统100相同。

[0139] 处理器311可以包括在存储设备300a中,可以位于存储控制器310外部,和/或可以与包括在存储控制器310中的另一处理器311分开实现。在图21的示例中,诸如神经网络模型NNM的数据分类器可以由处理器311执行和/或驱动。

[0140] 参考图22,存储系统100b包括主机200、存储设备300b和处理器311。

[0141] 除了图22的存储系统100b还包括处理器311之外,图22的存储系统100b可以基本上与图2的存储系统100相同。

[0142] 在一些示例实施例中,处理器311可以位于存储设备300b的外部,和/或可以与包括在主机200中的处理器和/或包括在存储控制器310中的处理器(如图3所示)分开实现。在图22的示例中,诸如神经网络模型NNM的数据分类器可以由处理器311执行和/或驱动。

[0143] 在一些示例实施例中,存储设备300和/或存储控制器310可以包括处理器311。在包括图21和图22所示的示例实施例的一些示例实施例中,处理器311可以包括中央处理单元(CPU)、图形处理单元(GPU)、神经处理单元(NPU)、数字信号处理器(DSP)、图像信号处理器(ISP)和/或专用硬件中的至少一个。在一些示例实施例中,处理器311可以包括:逻辑电路;硬件/软件组合,例如,执行软件的处理器;和/或它们的组合。例如,处理器311可以包括但不限于中央处理单元(CPU)、算术逻辑单元(ALU)、数字信号处理器、微型计算机、现场可编程门阵列(FPGA)、片上系统(SoC)、可编程逻辑单元、微处理器、专用集成电路(ASIC)等。

[0144] 图23是示出根据一些示例实施例的存储设备中所包括的非易失性存储器件的示例的框图。

[0145] 参考图23,非易失性存储器件500包括存储单元阵列510、行译码器520、页面缓冲电路530、数据输入/输出(I/O)电路540、电压发生器550和/或控制电路560。

[0146] 存储单元阵列510经由多条串选择线SSL、多条字线WL和多条接地选择线GSL连接

到行译码器520。存储单元阵列510还经由多条位线BL连接到页面缓冲电路530。存储单元阵列510可以包括连接到多条字线WL和多条位线BL的多个存储单元(例如,多个非易失性存储单元)。存储单元阵列510可以被分成多个存储块BLK1、BLK2.....BLKz,每个存储块都包括存储单元。在一些示例实施例中,多个存储单元可以以二维(2D)阵列结构或三维(3D)垂直阵列结构排列。

[0147] 三维垂直阵列结构可以包括垂直定向的垂直单元串,使得至少一个存储单元位于另一个存储单元之上。该至少一个存储单元可以包括电荷俘获层。存储单元阵列的一些示例配置可以包括3D垂直阵列结构,在3D垂直阵列结构中三维存储阵列被配置为多个层级,字线和/或位线在层级之间共享:美国专利NO.7,679,133、No.8,553,466、No.8,654,587、No.8,559,235和美国专利公开号为2011/0233648的专利。

[0148] 控制电路560从存储控制器(例如,图2中的主机200和/或存储控制器310)接收命令CMD和地址ADDR,并基于命令CMD和地址ADDR控制非易失性存储器件500的擦除、编程和/或读取操作。擦除操作可以包括执行一系列擦除循环,和/或编程操作可以包括执行一系列编程循环。每个编程循环可以包括编程操作和/或编程验证操作。每个擦除循环可以包括擦除操作和/或擦除验证操作。读取操作可以包括正常读取操作和/或数据恢复读取操作。

[0149] 例如,控制电路560基于命令CMD可以生成用于控制电压发生器550的控制信号CON,和/或可以生成用于控制页面缓冲电路530的控制信号PBC,和/或可以基于地址ADDR生成行地址R\_ADDR和列地址C\_ADDR。控制电路560可以将行地址R\_ADDR提供给行译码器520和/或可以将列地址C\_ADDR提供给数据I/O电路540。

[0150] 行译码器520可以经由多条串选择线SSL、多条字线WL和/或多条接地选择线GSL连接到存储单元阵列510。

[0151] 电压发生器550可以基于功率PWR和/或控制信号CON产生用于非易失性存储器件500的操作的电压VS。电压VS可以经由行译码器520施加到多条串选择线SSL、多条字线WL和/或多条接地选择线GSL。此外,电压发生器550可以基于功率PWR和/或控制信号CON生成用于数据擦除操作的擦除电压VERS。擦除电压VERS可以直接或经由位线BL施加到存储单元阵列510。

[0152] 页面缓冲电路530可以经由多条位线BL连接到存储单元阵列510。页面缓冲电路530可以包括多个页面缓冲器。在一些示例实施例中,每个页面缓冲器可以连接到一条位线。在其他示例实施例中,每个页面缓冲器可以连接到两条以上的位线。

[0153] 页面缓冲电路530可以存储要编程到存储单元阵列510中的数据DAT和/或可以读取从存储单元阵列510感测到的数据DAT。也就是说,页面缓冲电路530可以根据非易失性存储器件500的操作模式用作写入驱动器和/或感测放大器。

[0154] 数据I/O电路540可以经由数据线DL连接到页面缓冲电路530。基于列地址C\_ADDR,数据输入/输出电路540可以经由页面缓冲电路530将来自非易失性存储器件500外部的数据DAT提供给存储单元阵列510,和/或可以将来自存储单元阵列510的数据DAT提供到非易失性存储器件500的外部。

[0155] 图24是示出根据一些示例实施例的电子系统的框图。

[0156] 参考图24,电子系统4000包括至少一个处理器4100、通信模块4200、显示/触摸模块4300、存储设备4400和存储器件4500。例如,电子系统4000可以是任何移动系统或任何计

算系统。处理器4100可以对应于和/或类似于图21和/或图22的处理器311。

[0157] 处理器4100控制电子系统4000的操作。处理器4100可以执行操作系统和/或至少一个应用来提供互联网浏览器、游戏、视频等。通信模块4200被实现为执行与外部设备的无线和/或有线通信。显示/触摸模块4300被实现为显示由处理器4100处理的数据和/或通过触摸面板接收数据。存储设备4400被实现为存储用户数据,和/或基于根据一些示例实施例的操作存储设备的方法来驱动存储设备4400。存储器件4500临时存储用于电子系统4000的处理操作的数据。图24中的处理器4100和/或存储设备4400可以分别对应于图2中的主机200和/或存储设备300。

[0158] 一些发明构思的一些示例实施例可以应用于包括存储设备和存储系统的各种电子设备和/或电子系统。例如,一些发明构思的一些示例实施例可以应用于诸如移动电话、智能电话、平板计算机、膝上型计算机、个人数字助理(PDA)、便携式多媒体播放器(PMP)、数码相机、便携式游戏机、音乐播放器、便携式摄像机、视频播放器、导航设备、可穿戴设备、物联网(IoT)设备、万物互联网(IoE)设备、电子书阅读器、虚拟现实(VR)设备、增强现实(AR)设备和/或机器人设备的系统。

[0159] 尽管已经描述了一些示例实施例,但是本领域技术人员将容易理解的是,在一些示例实施例中,在实质上不脱离本公开的主题的情况下,可以进行许多修改。因此,所有这些修改都旨在包括在如权利要求所限定的至少一些示例实施例的范围内。因此,应当理解的是,前述是对一些示例实施例的说明,并且不应当被解释为限于所公开的特定示例实施例,并且对所公开的示例实施例的修改以及其他示例实施例旨在包括在所附权利要求的范围内。

[0160] 如本文所使用的,诸如“当……时”和“在……时”之类的术语可以但不一定旨在暗示诸如事件序列之类的时间关系。例如,被描述为“当满足某个条件时”或“在满足某个条件时”发生的操作可以与满足条件同时执行,或者在满足条件期间执行,或者在满足条件之后执行,或者因为满足了条件或已经满足了条件所以执行,等等。类似地,诸如将操作与条件相关联的“在……情况下”之类的短语可以指示操作与满足条件同时执行,或者在满足条件期间执行,或者在满足条件之后执行,或者因为满足了条件或已经满足了条件所以执行,等等。所有这些合理的和逻辑上可能的并且与其他陈述不矛盾的解释都旨在包括在本公开中,本公开的范围被理解为仅由权利要求限定。

[0161] 本文所呈现和讨论的每个流程图描绘了以特定顺序排列的一组操作。然而,应当理解的是,每个流程图中的操作布置仅描绘了一个示例,并且根据本公开的主题,可以添加、移除、重新排列和/或同时执行操作。仅作为一个示例,在一些实施例中,如图1所示,训练数据分类器的操作S100可以发生在接收数据写入请求和要写入的目标数据的操作S200之前。然而,在其他实施例中,操作S100可以与操作S200同时发生和/或在操作S200之后发生。所有这些合理的和逻辑上可能的并且与其他陈述不矛盾的变化都旨在包括在本公开中,本公开的范围应当被理解为仅由权利要求限定。

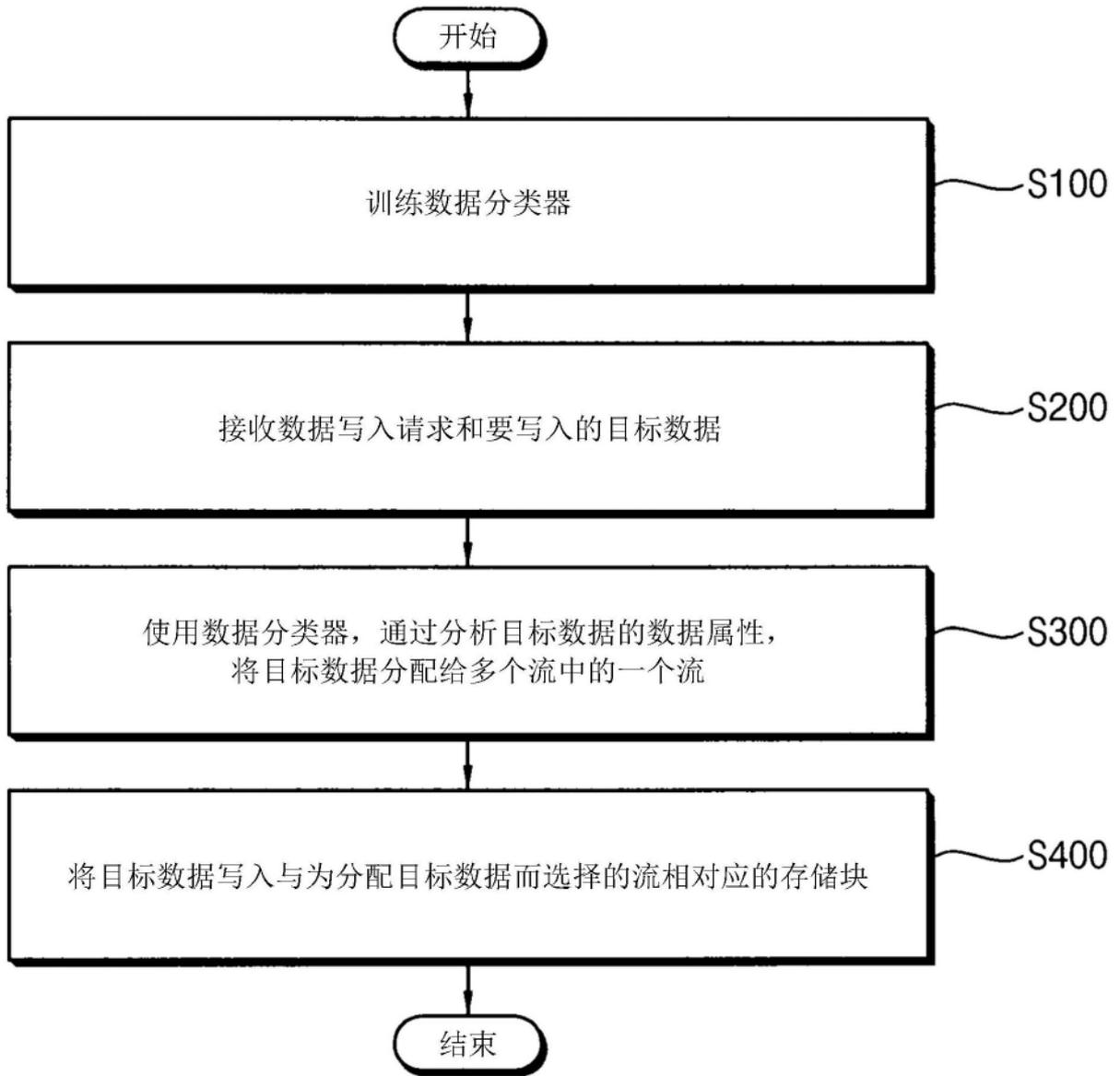


图1

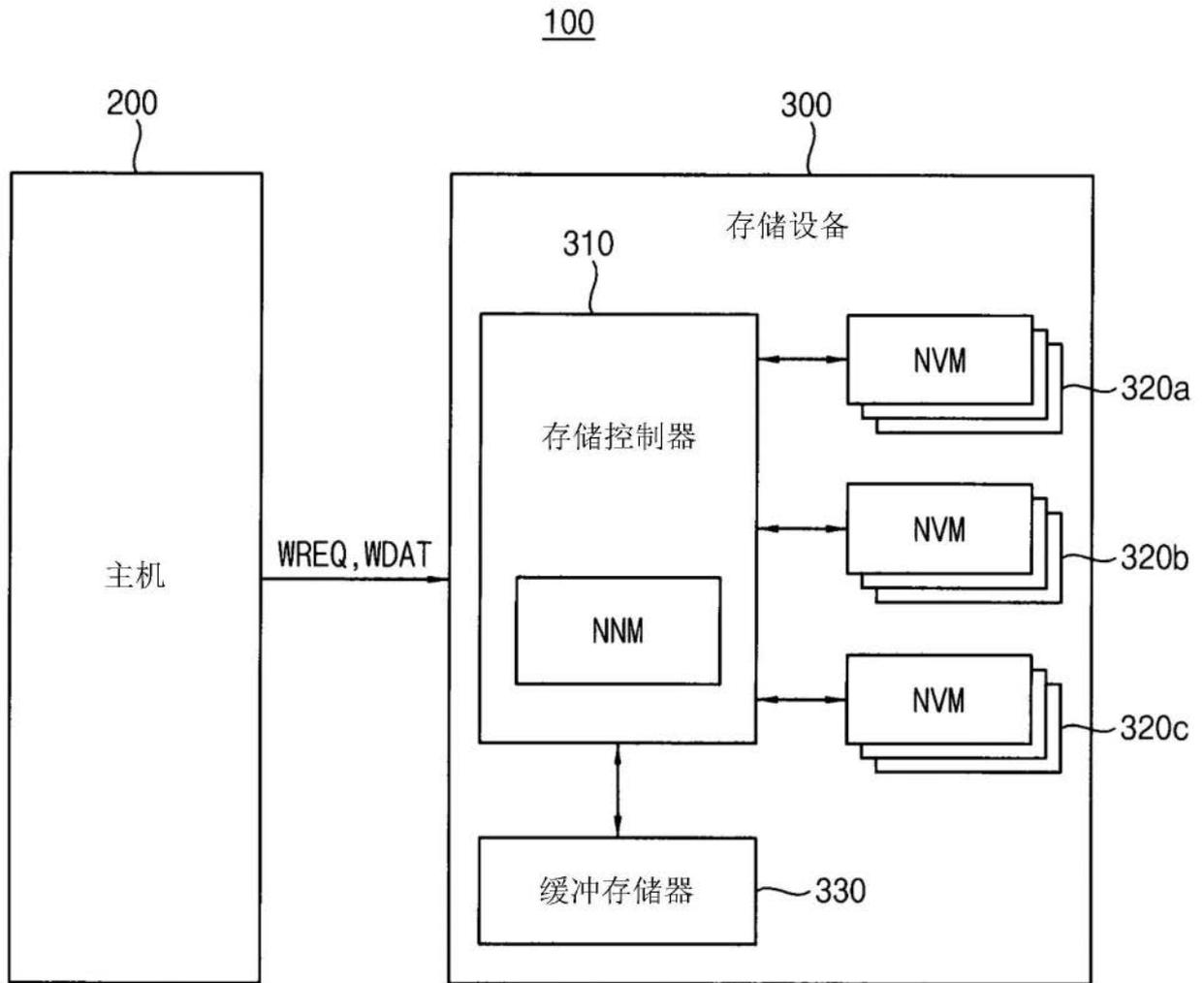


图2

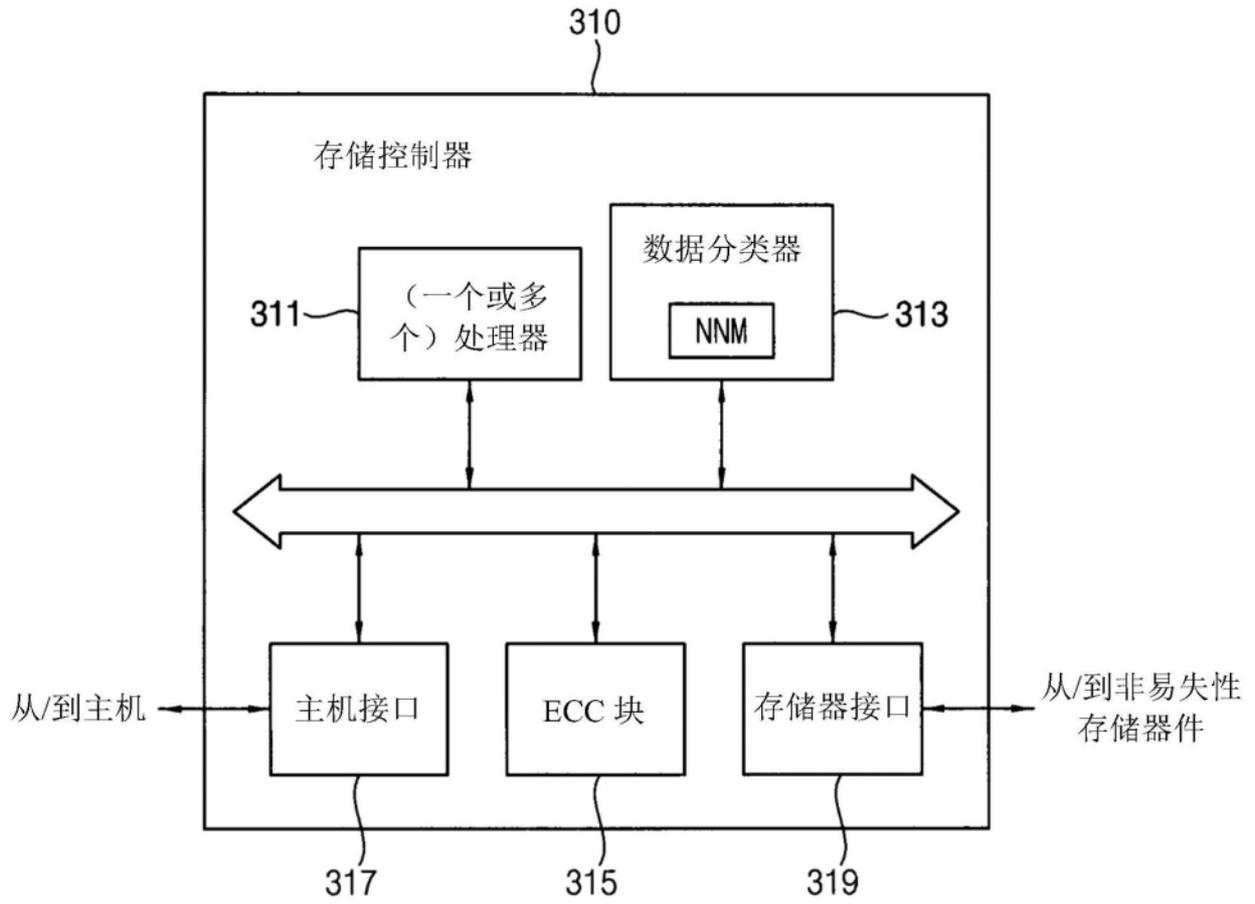


图3

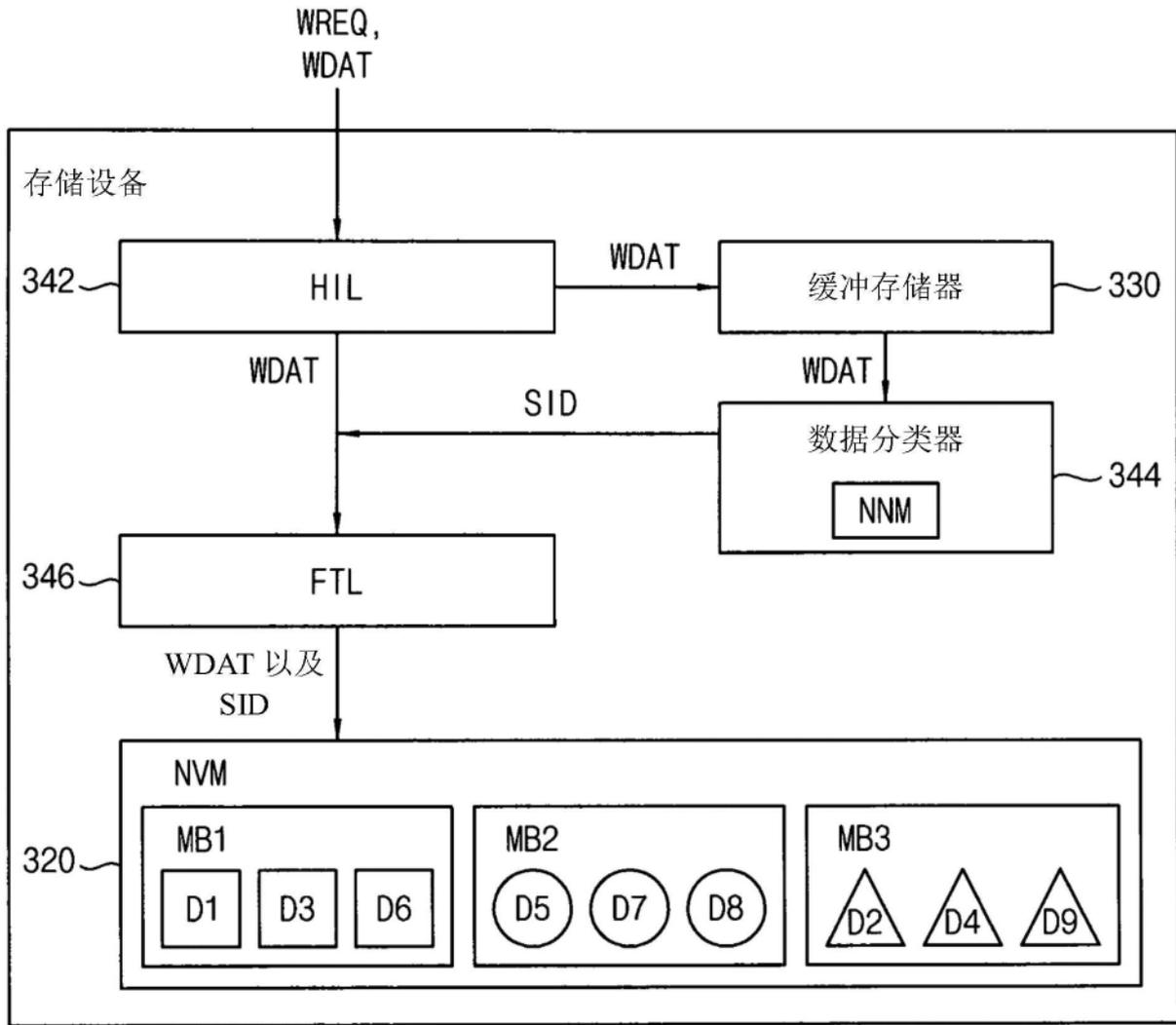


图4

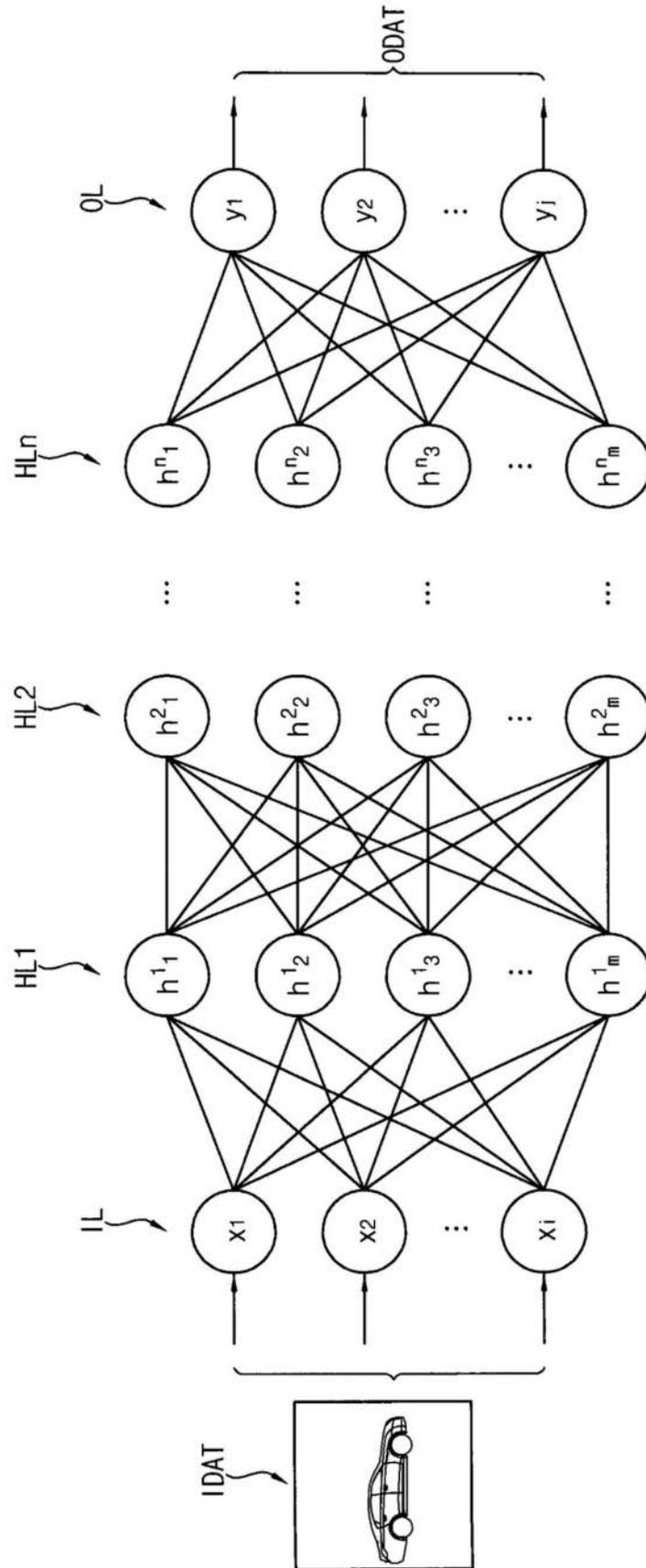


图5A

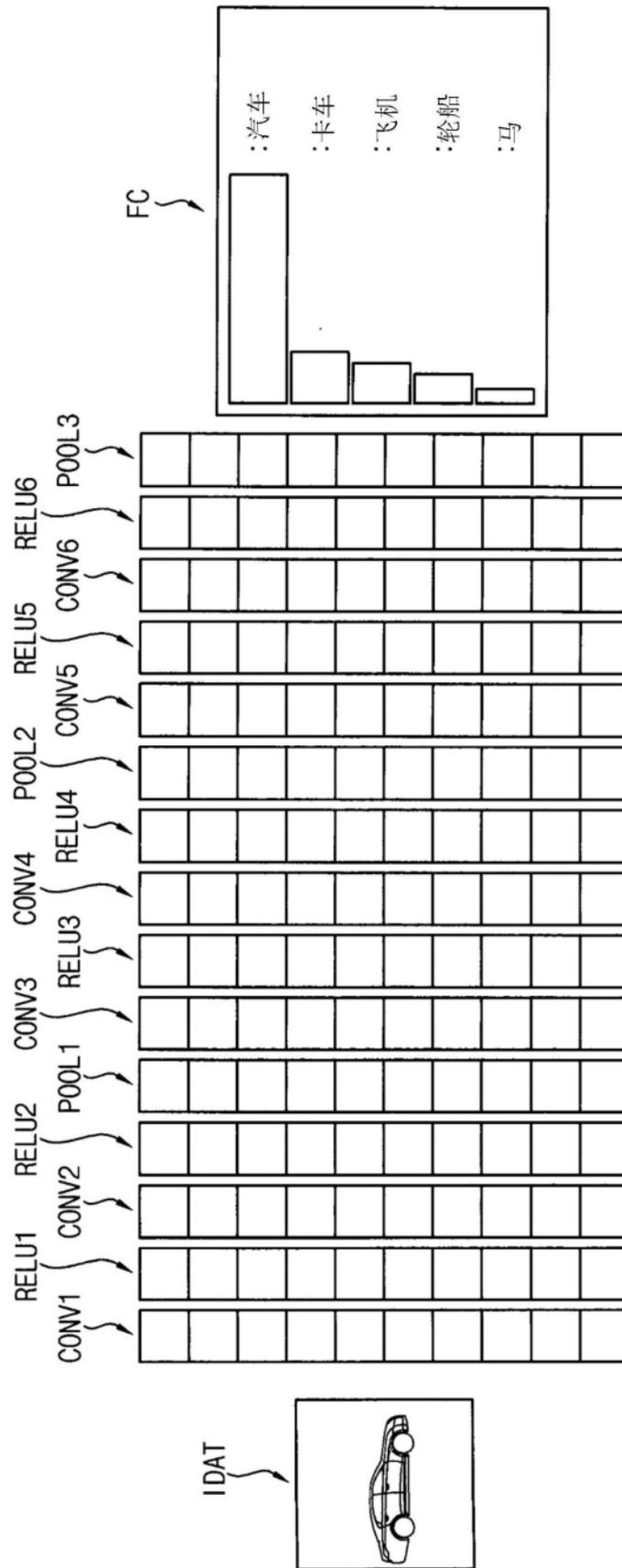


图5B

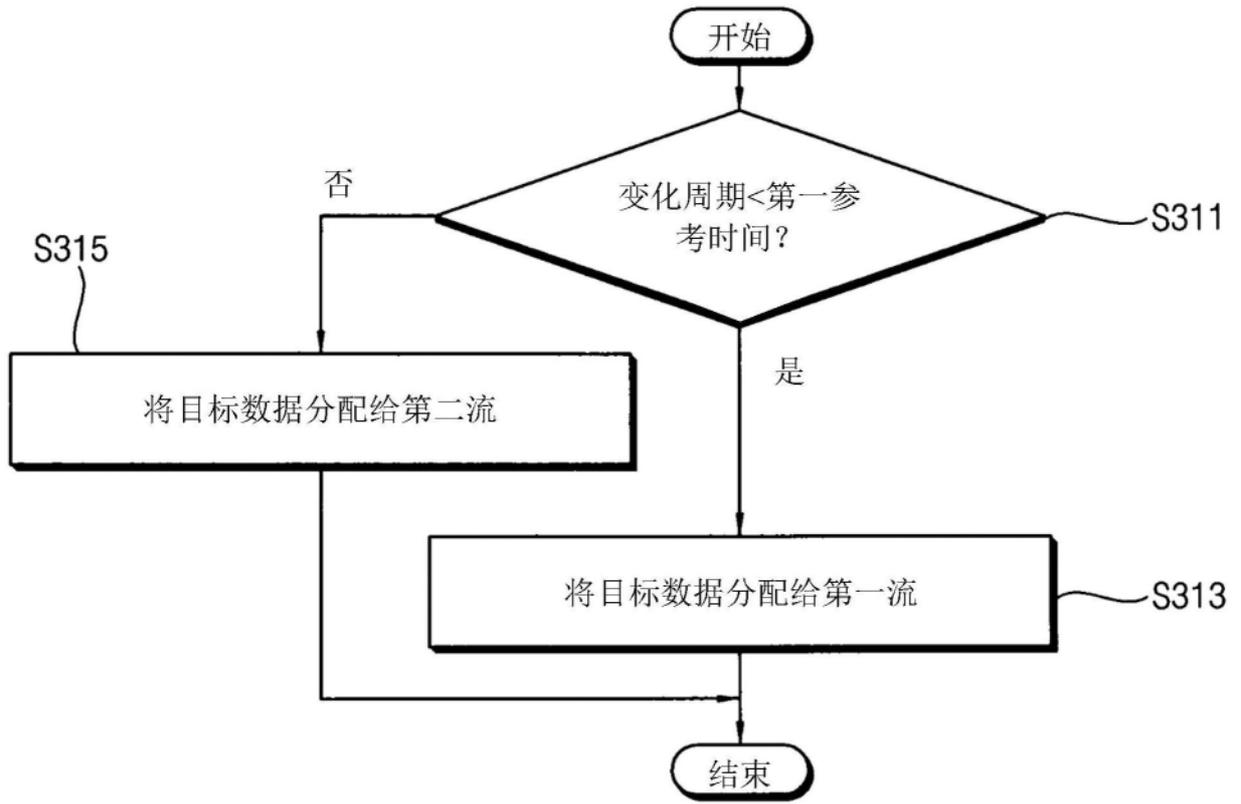


图6

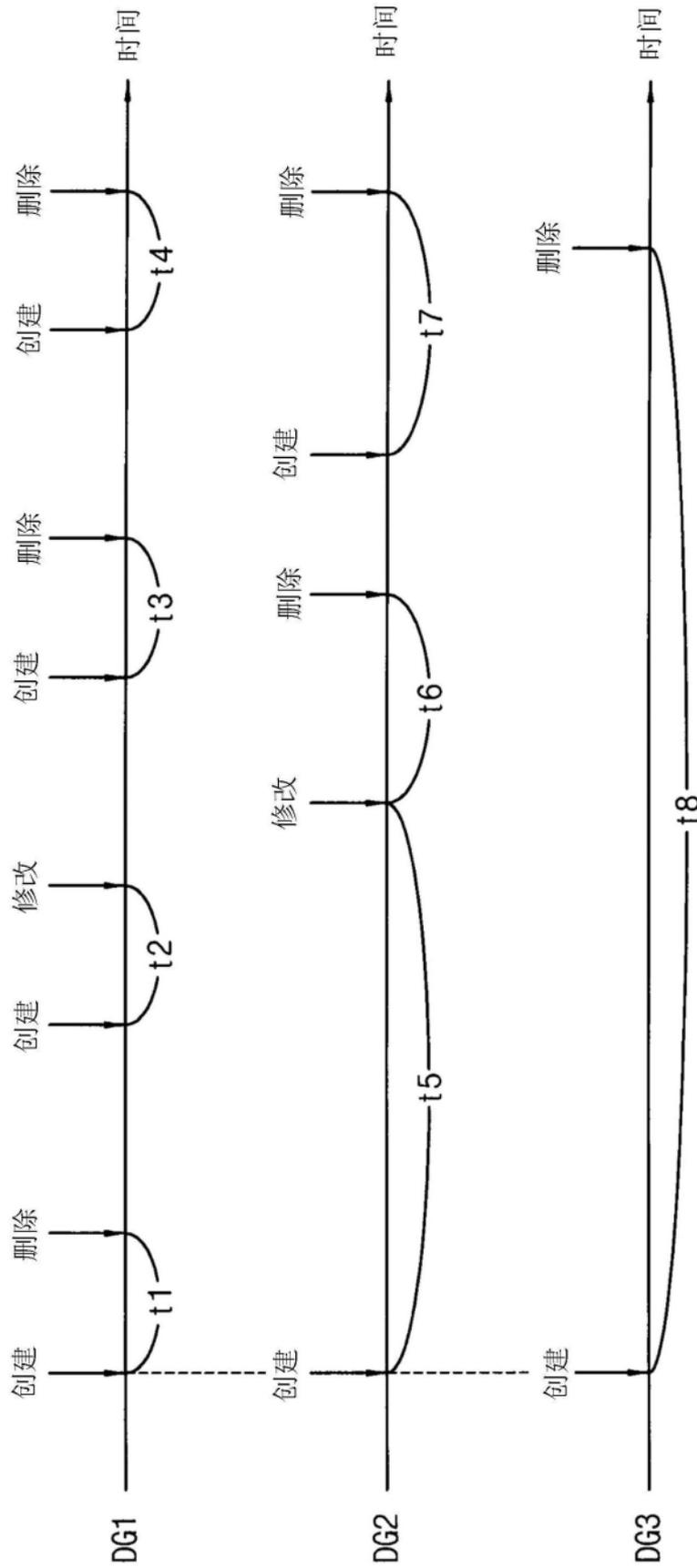


图7

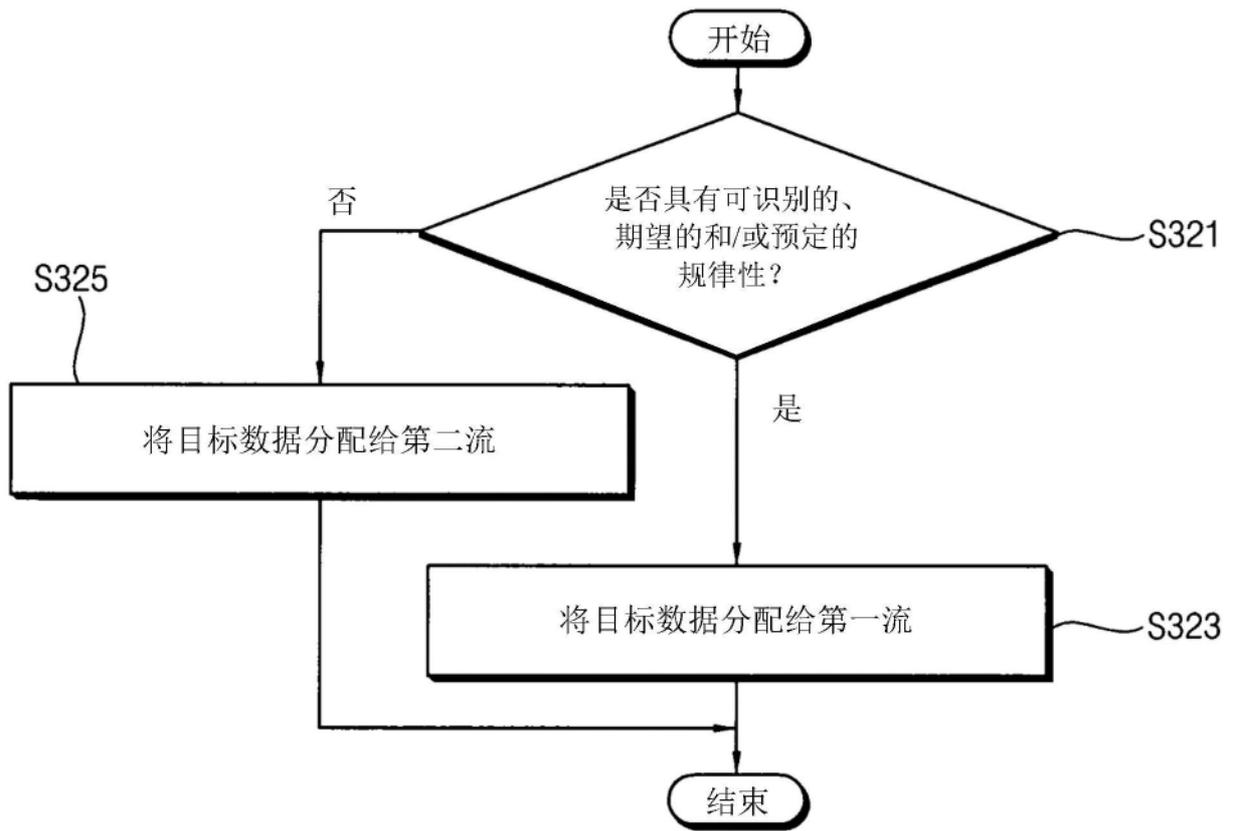


图8

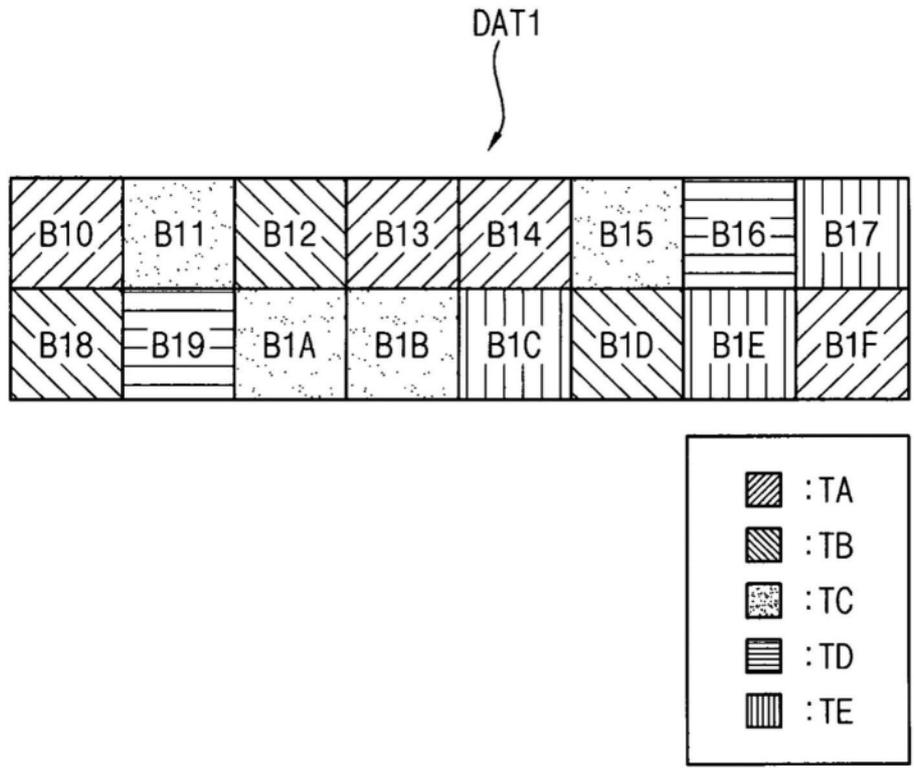


图9A

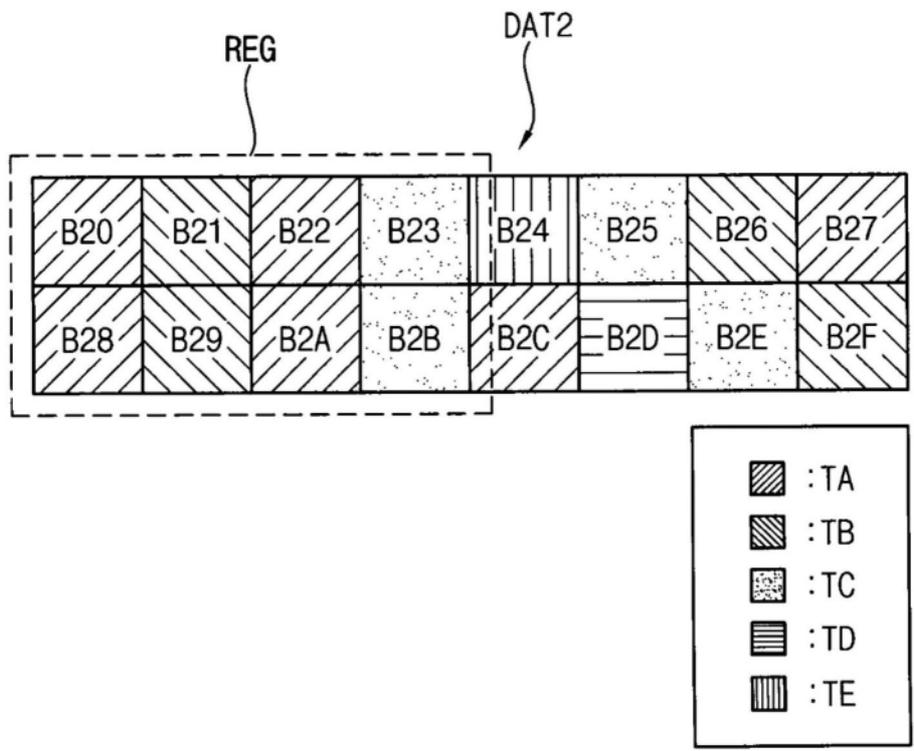


图9B

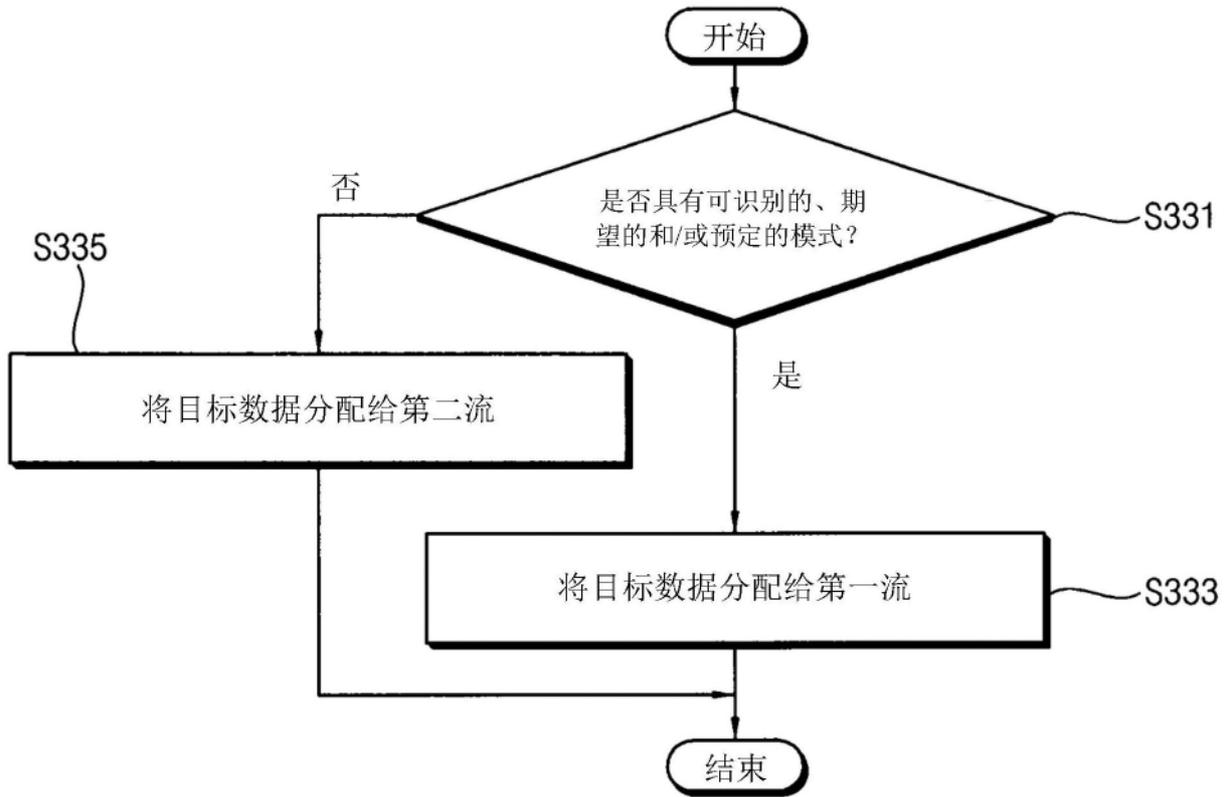


图10

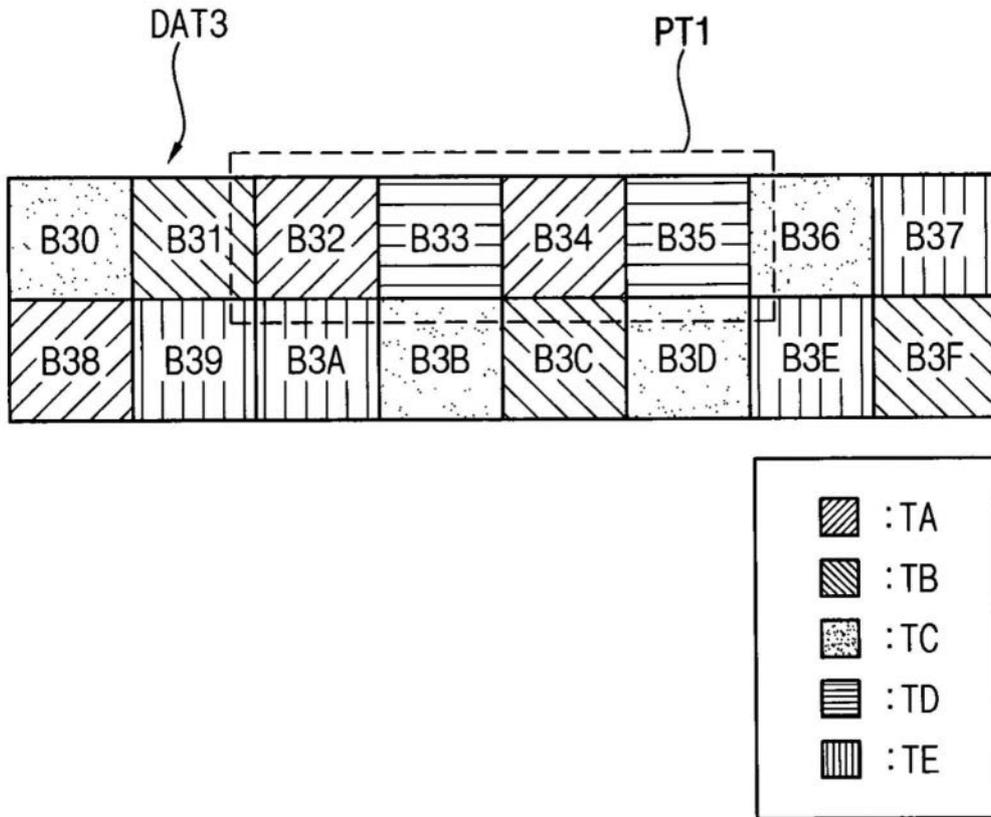


图11

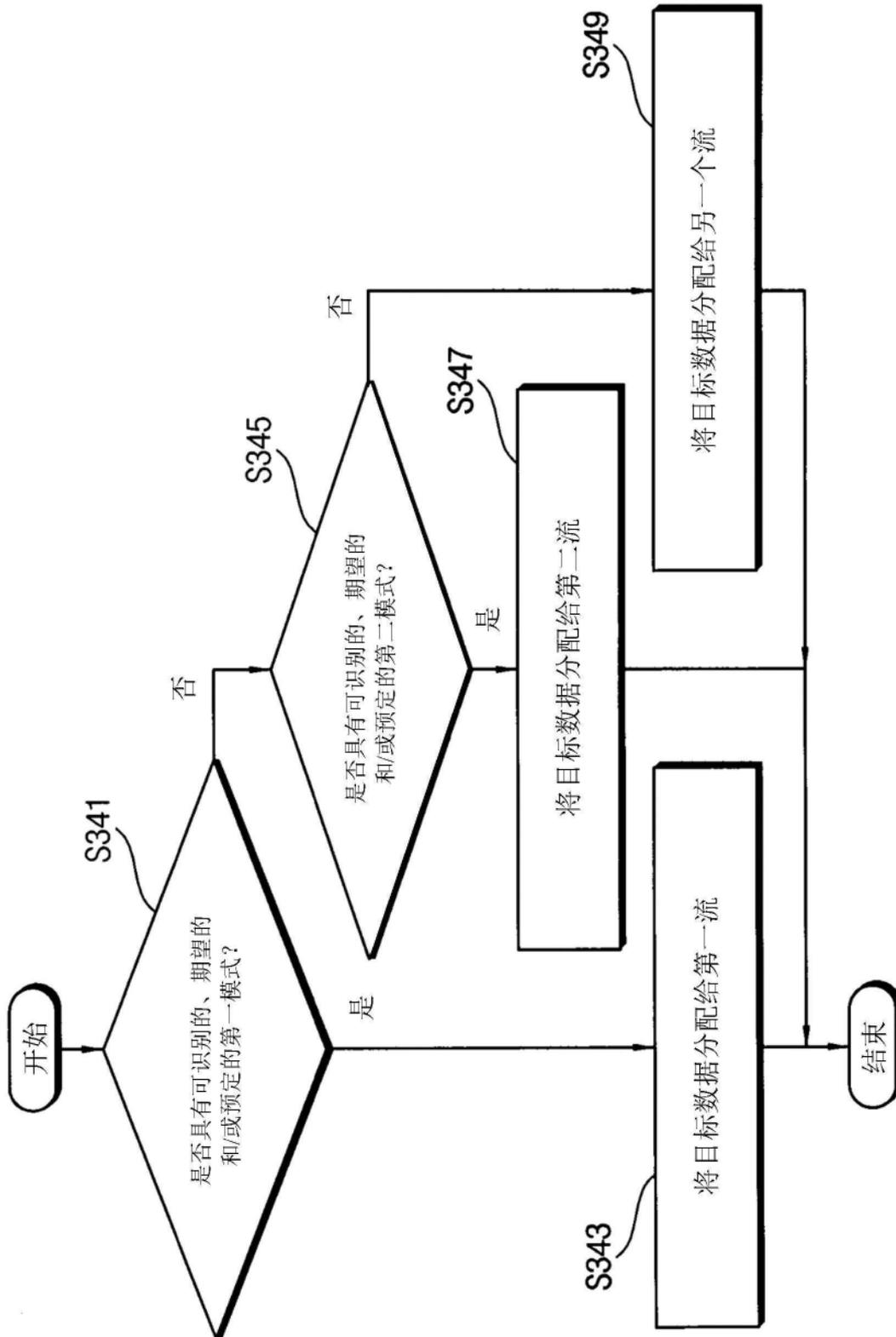


图12

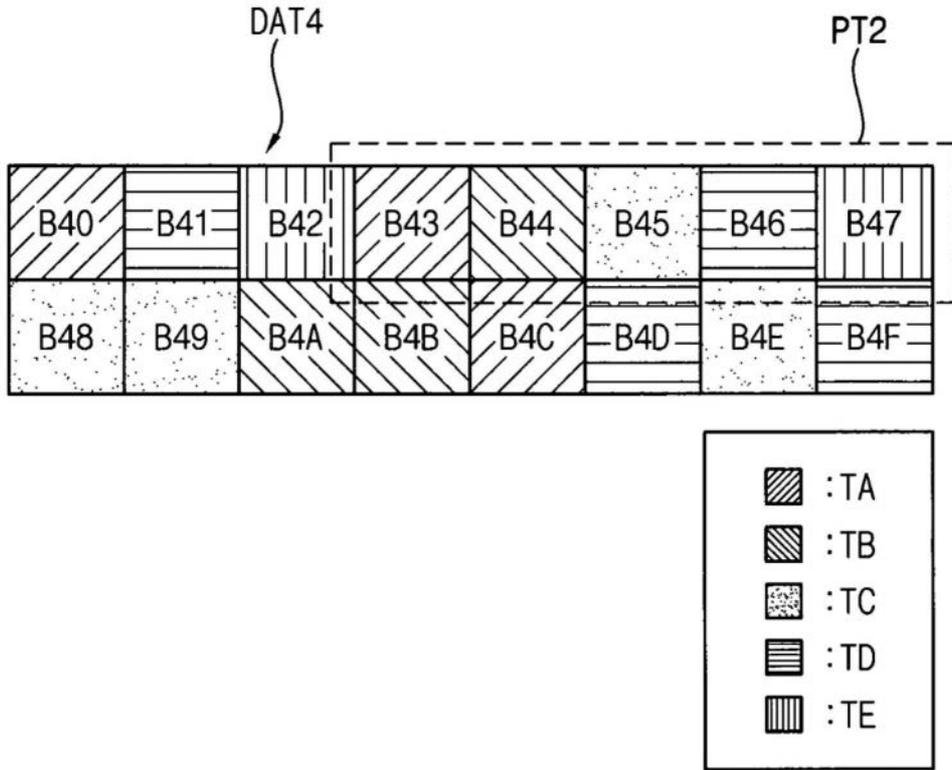


图13

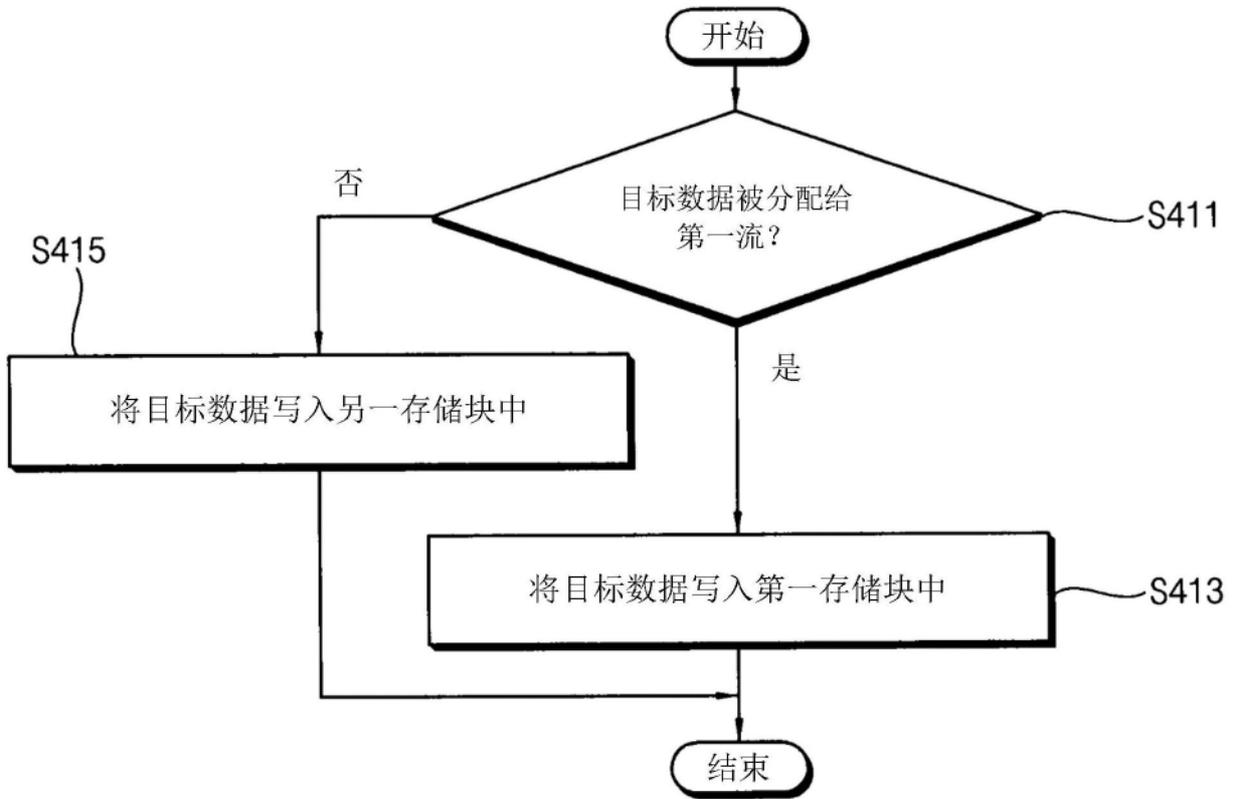


图14

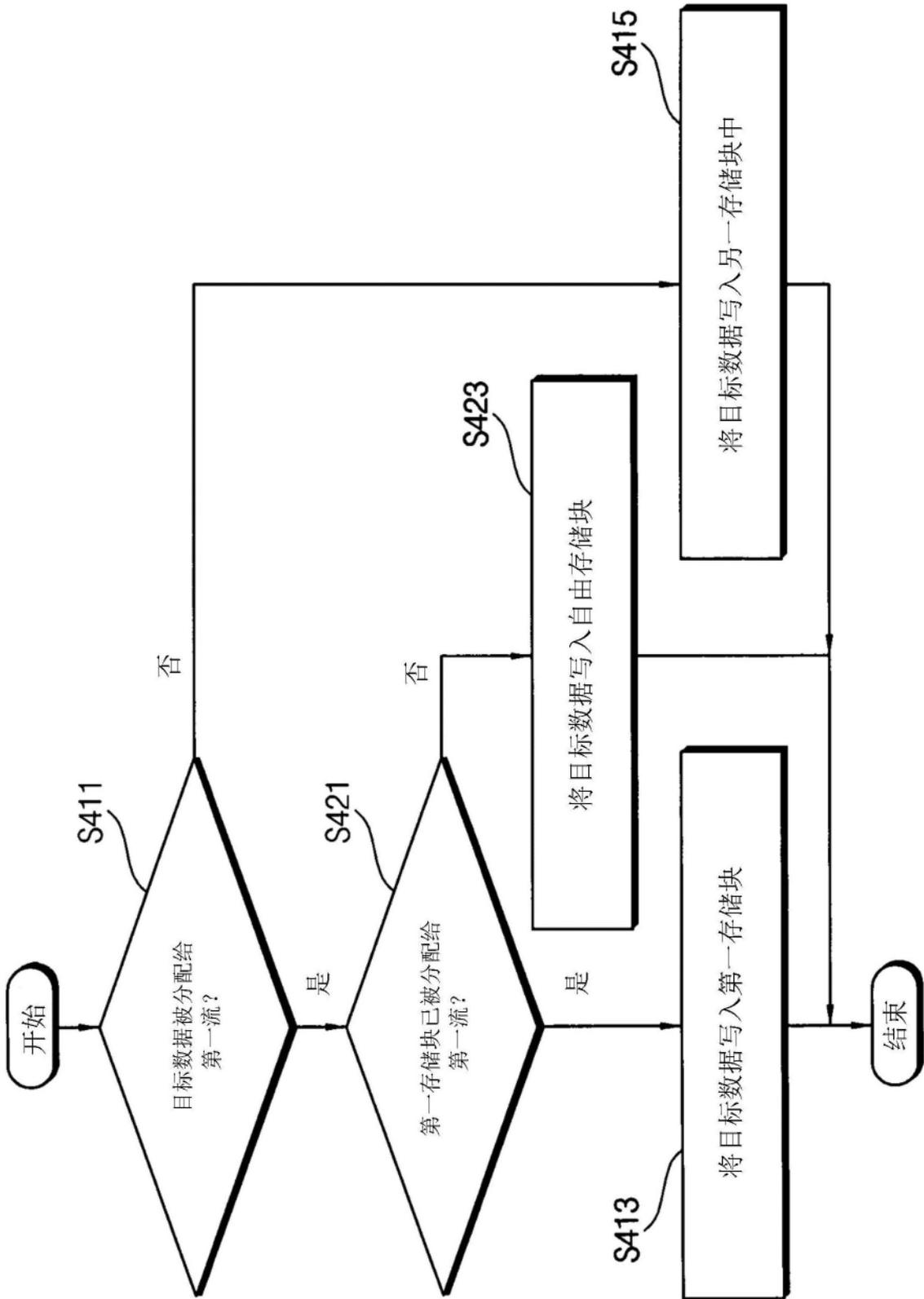


图15

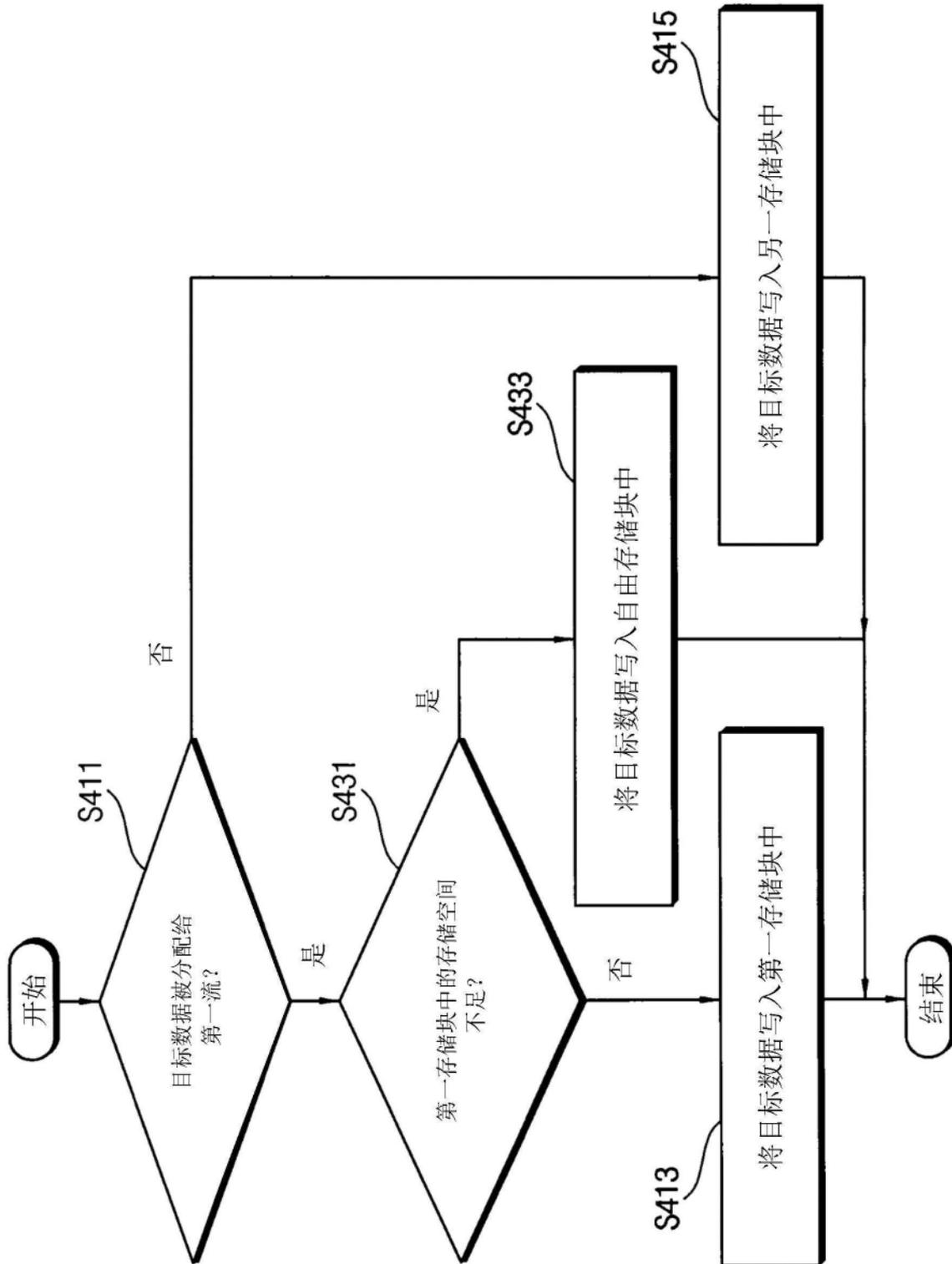


图16

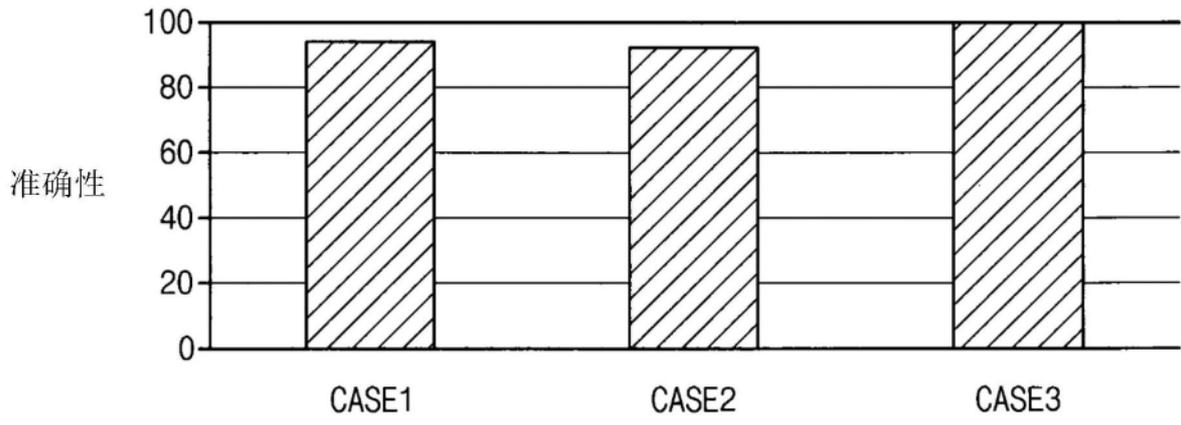


图17

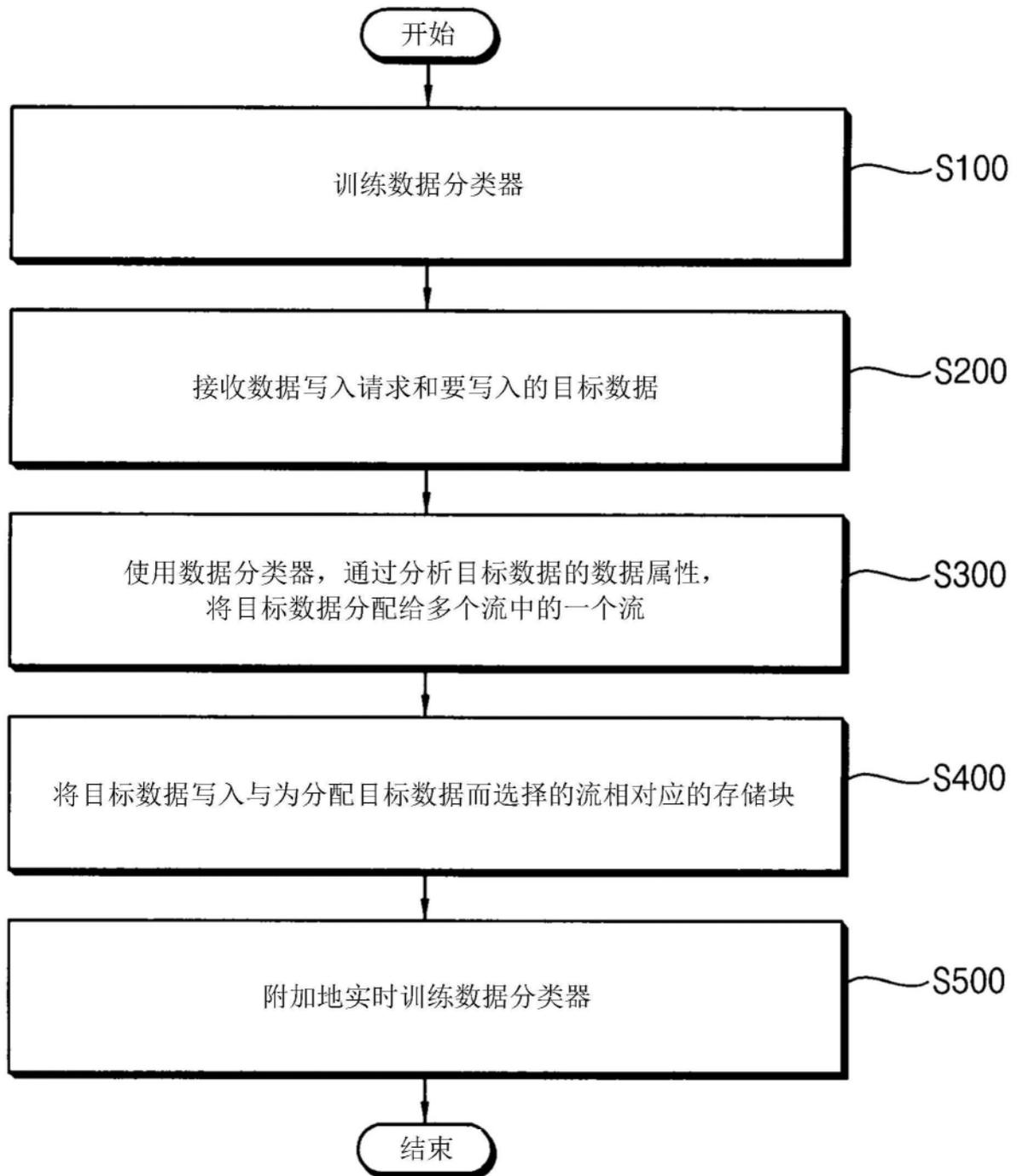


图18

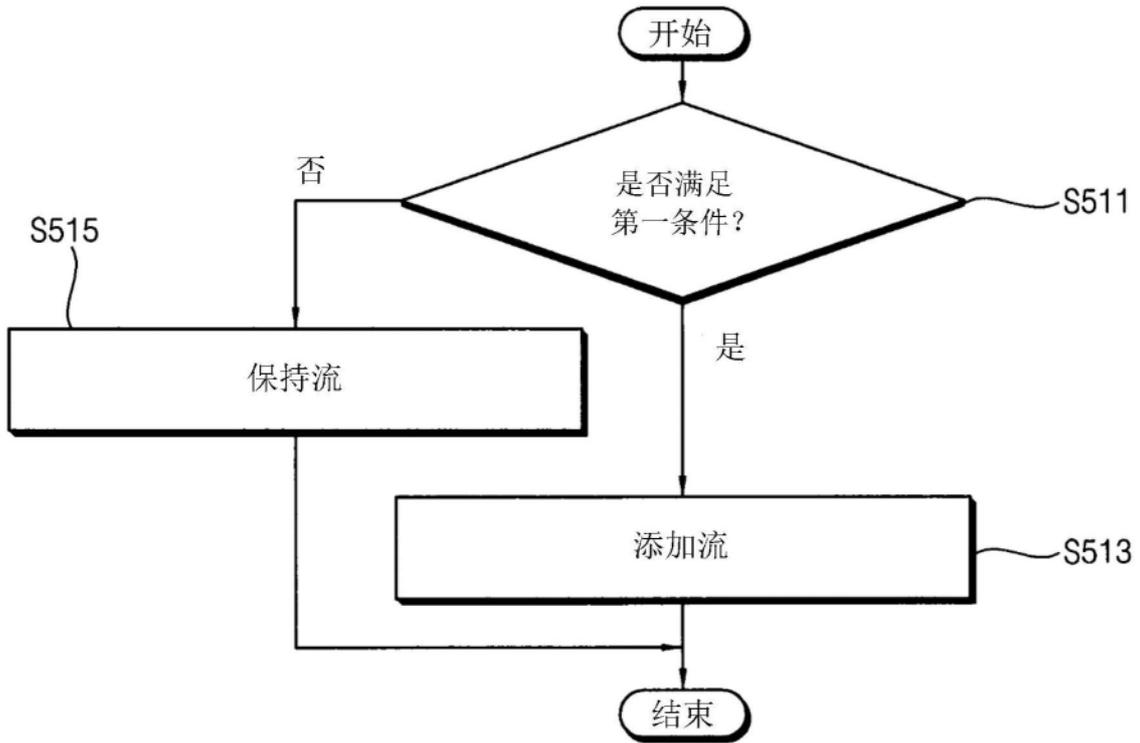


图19

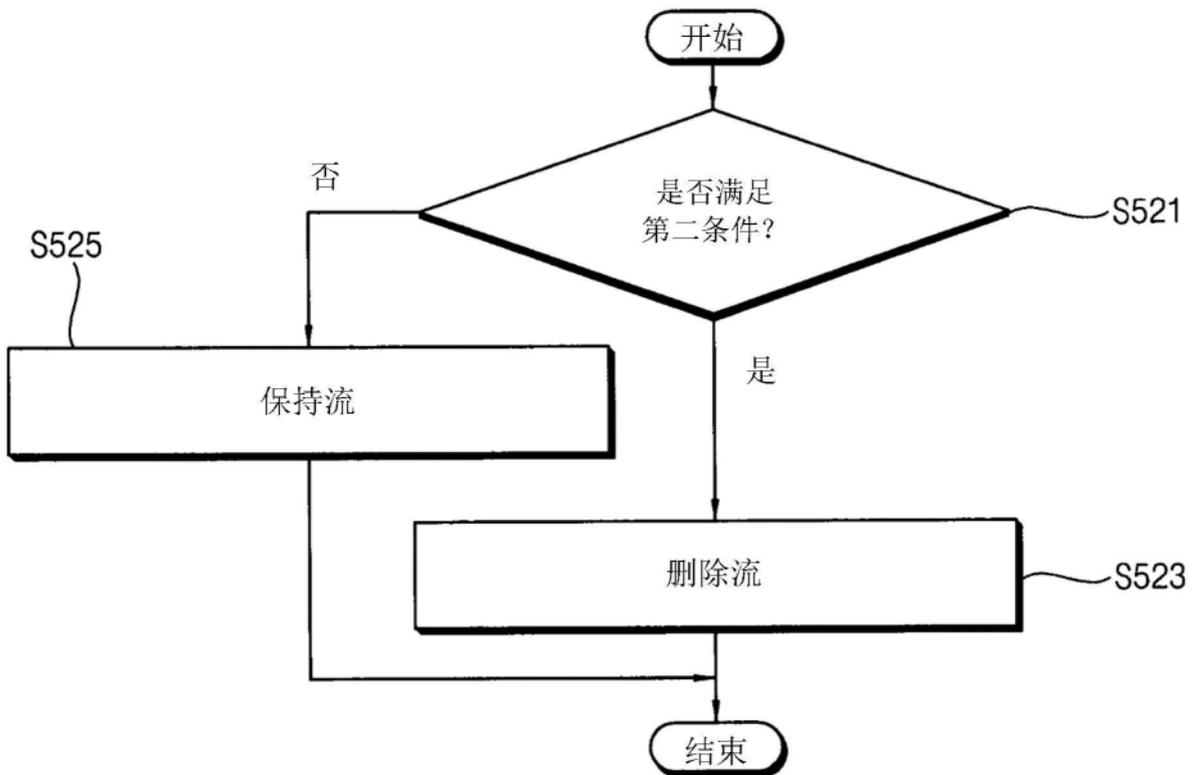


图20

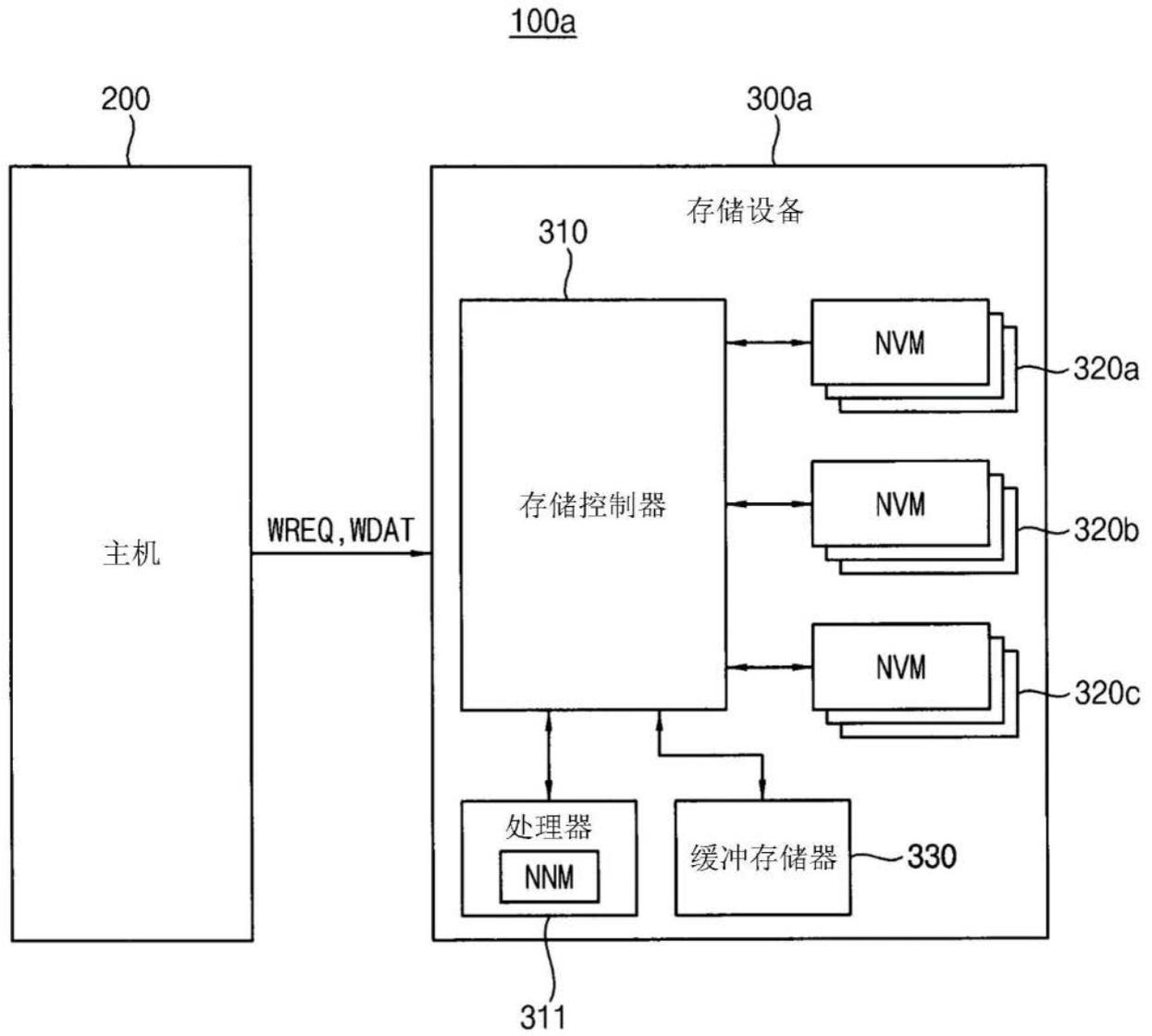


图21

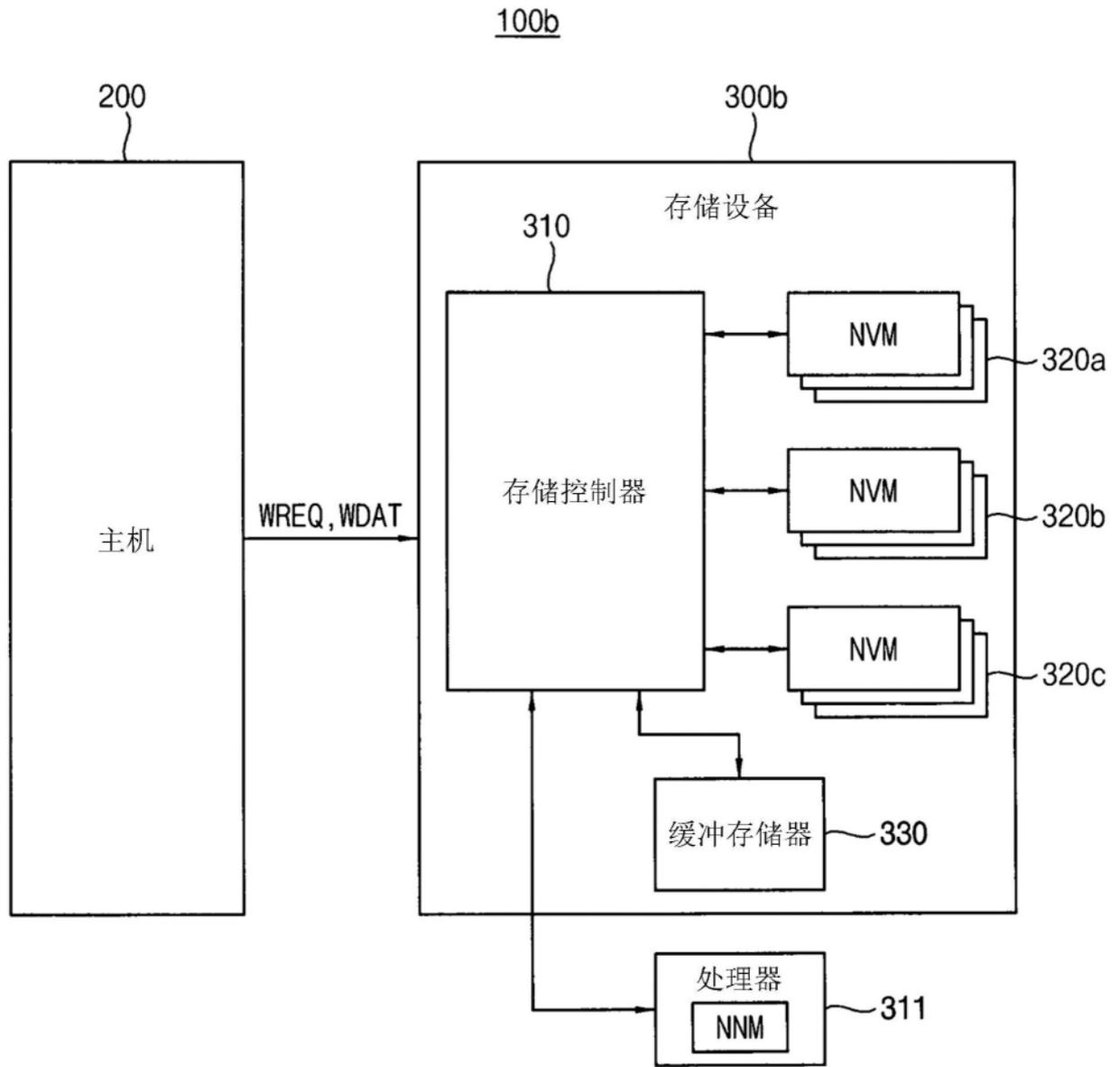


图22

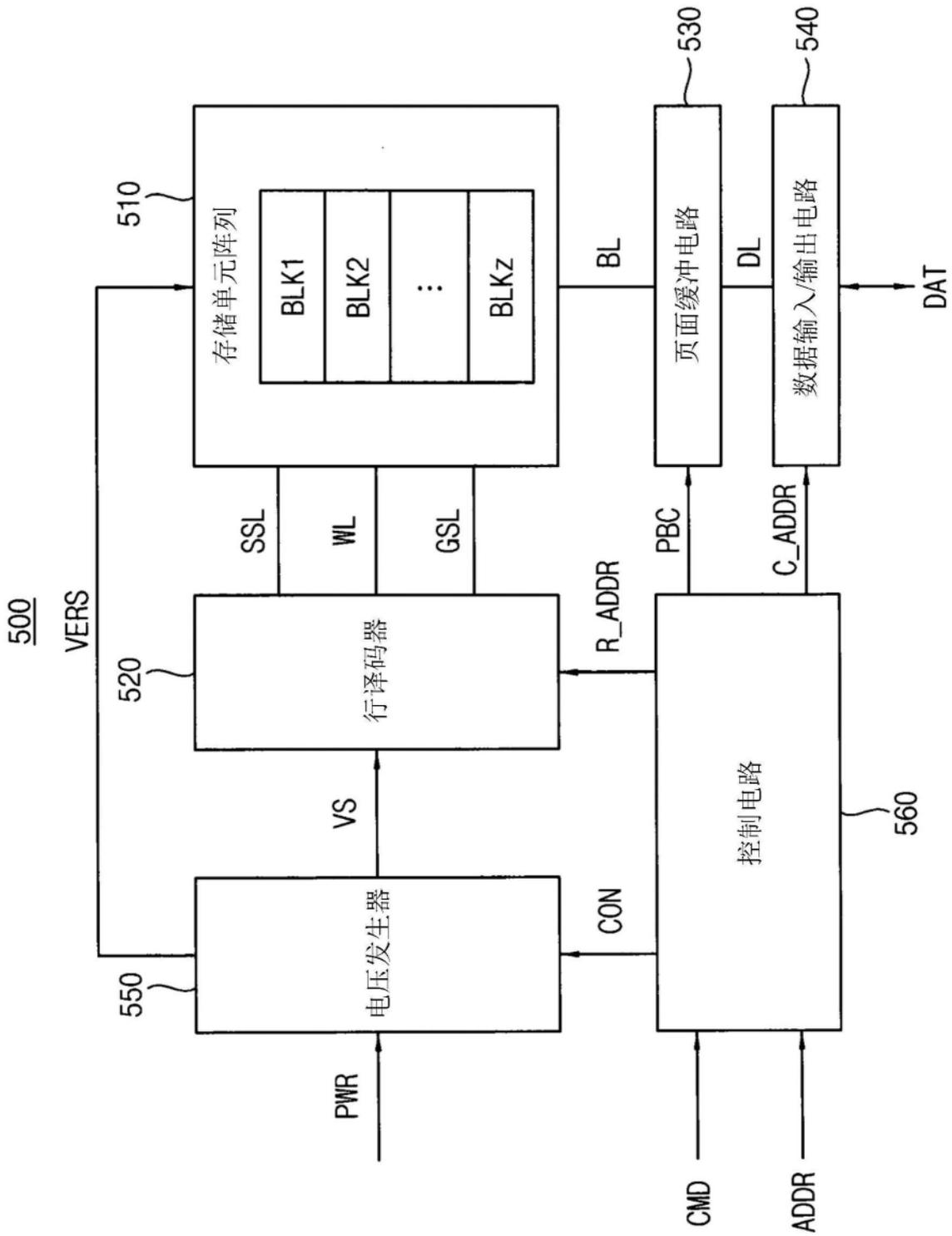


图23

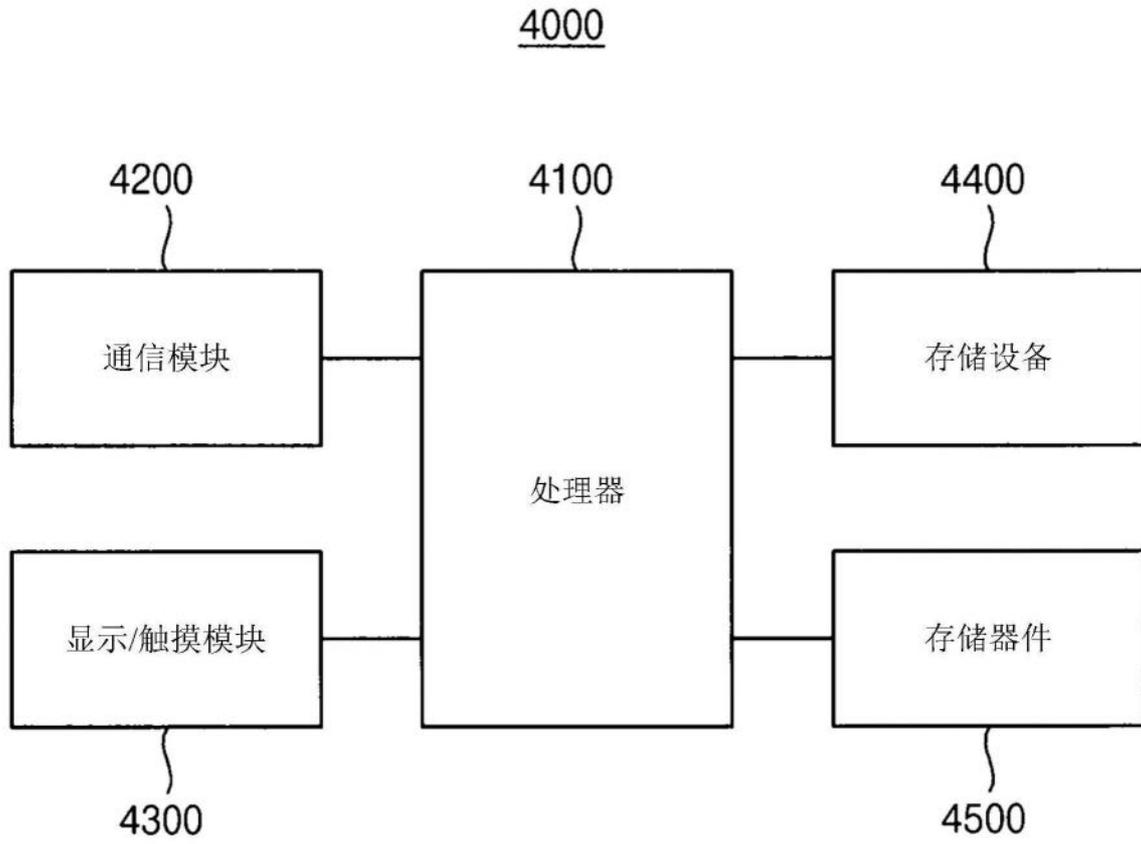


图24