

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2023 年 3 月 30 日 (30.03.2023)



(10) 国际公布号
WO 2023/045114 A1

- (51) 国际专利分类号:
G06F 15/78 (2006.01) **G06N 3/04** (2006.01)
G06K 9/62 (2022.01) **G06N 3/08** (2006.01)
- (21) 国际申请号: PCT/CN2021/137844
- (22) 国际申请日: 2021 年 12 月 14 日 (14.12.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202111109178.6 2021 年 9 月 22 日 (22.09.2021) CN
- (71) 申请人: 清华大学 (TSINGHUA UNIVERSITY)
[CN/CN]; 中国北京市海淀区清华园,
Beijing 100084 (CN)。
- (72) 发明人: 吴华强 (WU, Huaqiang); 中国北京市
海淀区清华园, Beijing 100084 (CN)。 江燕海
(JIANG, Yanhai); 中国北京市海淀区清华园,

Beijing 100084 (CN)。 姚鹏 (YAO, Peng); 中国
北京市海淀区清华园, Beijing 100084 (CN)。
高滨 (GAO, Bin); 中国北京市海淀区清华园,
Beijing 100084 (CN)。 唐建石 (TANG, Jianshi);
中国北京市海淀区清华园, Beijing 100084
(CN)。 钱鹤 (QIAN, He); 中国北京市海淀区
清华园, Beijing 100084 (CN)。

(74) 代理人: 北京市柳沈律师事务所 (LIU, SHEN &
ASSOCIATES); 中国北京市海淀区彩和坊路 10
号 1 号楼 10 层, Beijing 100080 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家
保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,
BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU,
CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT,
JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,
LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,

(54) Title: STORAGE AND COMPUTATION INTEGRATED CHIP AND DATA PROCESSING METHOD

(54) 发明名称: 存算一体芯片及数据处理方法

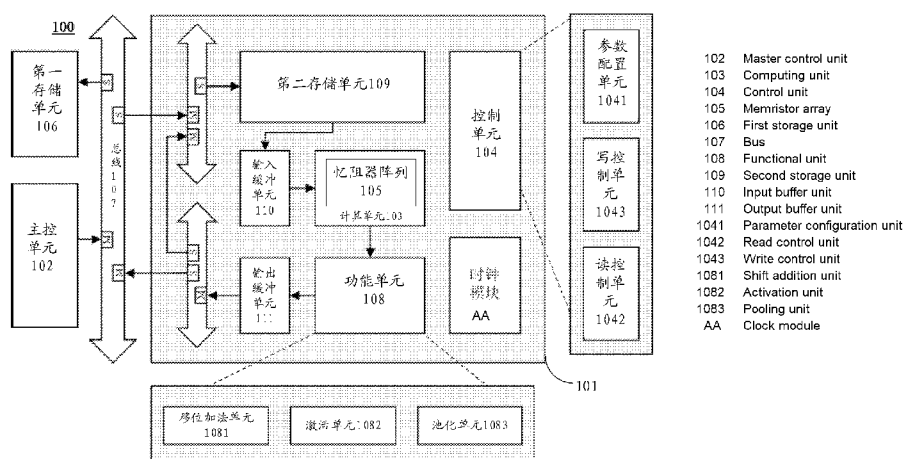


图 3A

(57) Abstract: A storage and computation integrated chip and a data processing method applied to same. The storage and computation integrated chip comprises: a plurality of processing units (101) and a master control unit (102). Each processing unit (101) comprises a computing unit (103) and a control unit (104). The computing unit (103) is configured to store a parameter matrix, receive input data, and compute the input data by using the parameter matrix to obtain output data. The control unit (104) is configured to communicate with the computing unit (103), and is configured to write the parameter matrix into the computing unit (103) to obtain the configured computing unit (103), and control the configured computing unit (103) to compute the input data. The master control unit (102) is configured to communicate with a plurality of processing units (101), and provide, to respective control units (104) of the plurality of processing units, the parameter matrixes for the plurality of processing units (101) according to an overall parameter set. According to the storage and computation integrated chip, an efficient distributed architecture is utilized, and high computing power and high energy efficiency are achieved while ensuring the universality and flexibility.

MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 一种存算一体芯片及应用该存算一体芯片的数据处理方法。该存算一体芯片包括: 多个处理单元 (101) 以及主控单元 (102)。每个处理单元 (101) 包括计算单元 (103) 以及控制单元 (104)。计算单元 (103) 配置为存储参数矩阵、接收输入数据并采用参数矩阵对输入数据进行计算得到输出数据; 控制单元 (104) 配置为与计算单元 (103) 相通信, 且配置为将参数矩阵写入计算单元 (103) 得到配置后的计算单元 (103), 并且控制配置后的计算单元 (103) 对输入数据进行计算。主控单元 (102) 配置为与多个处理单元 (101) 通信, 且根据整体参数集合将分别用于多个处理单元 (101) 的参数矩阵提供至多个处理单元各自的控制单元 (104)。该存算一体芯片利用高效的分布式架构, 在保证通用性和灵活性的同时, 实现高算力和高能效。

存算一体芯片及数据处理方法

本申请要求于 2021 年 9 月 22 日递交的中国专利申请第 202111109178.6 号的优先权，在此全文引用上述中国专利申请公开的内容以作为本申请的一部分。

技术领域

本公开的实施例涉及一种存算一体芯片及应用该存算一体芯片的数据处理方法。

背景技术

各种深度学习算法的识别率不断改善，随之而来的是数据量和运算量的指数级增长。基于忆阻器的存算一体技术有望突破经典计算系统的冯诺依曼架构的瓶颈，带来硬件算力、能效的爆发式增长，进一步促进人工智能的发展和落地，是最具潜力的下一代硬件芯片技术之一。基于忆阻器的存算一体系统凭借其高性能、低功耗的架构优势，有望成为突破冯诺依曼瓶颈的后摩尔时代智能计算的硬件载体。

发明内容

本公开至少一实施例提供一种分布式控制架构的存算一体芯片，包括：多个处理单元，每个处理单元包括：计算单元，配置为存储参数矩阵、接收输入数据并采用参数矩阵对输入数据进行计算得到输出数据；以及控制单元，配置为与计算单元相通信，且配置为将参数矩阵写入计算单元得到配置后的计算单元，并且控制配置后的计算单元采用参数矩阵对输入数据进行计算；以及主控单元，配置为与多个处理单元通信，且根据整体网络参数集合将分别用于多个处理单元的参数矩阵提供至多个处理单元各自的控制单元。

例如，本公开至少一实施例提供的存算一体芯片还包括：第一存储单元，配置为与多个处理单元通信以及与主控单元通信，存储整体输入数据和/或整体输出数据，在主控单元的控制下，向多个处理单元分别提供多个处理单元所需的输入数据，和/或，从多个处理单元分别接收多个处理单元产生的输出数

据,其中,整体输入数据包括用于多个计算单元的输入数据,整体输出数据包括多个计算单元的输出数据。

例如,本公开至少一实施例提供的存算一体芯片还包括:片上互连,配置为连接主控单元、第一存储单元、以及多个处理单元以进行通信。

- 5 例如,在本公开至少一实施例提供的存算一体芯片中,计算单元包括:至少一个忆阻器阵列,其中,该至少一个忆阻器阵列配置为存储参数矩阵且对输入数据进行计算。

10 例如,在本公开至少一实施例提供的存算一体芯片中,控制单元包括:参数配置单元,配置为将参数矩阵配置到计算单元中,和/或从计算单元获取计算单元当前具有的参数矩阵的值。

 例如,在本公开至少一实施例提供的存算一体芯片中,控制单元还包括:读控制单元,配置为控制输入数据的读取以将输入数据提供到计算单元中;写控制单元,配置为控制输出数据的输出。

- 15 例如,在本公开至少一实施例提供的存算一体芯片中,每个处理单元还包括:功能单元,其中,该功能单元配置为对计算单元的计算结果进行处理以得到输出数据。

20 例如,在本公开至少一实施例提供的存算一体芯片中,功能单元包括:移位加法单元,配置为获取来自计算单元的计算结果,并对计算结果进行移位运算和加法运算;或激活单元,配置为对计算结果进行激活函数处理;或池化单元,配置为对计算结果进行池化处理。

 例如,在本公开至少一实施例提供的存算一体芯片中,每个处理单元还包括:第二存储单元,配置为从第一存储单元获取输入数据或向第一存储单元发送输出数据;输入缓冲单元,配置为缓存输入数据;输出缓冲单元,配置为缓存输出数据。

- 25 例如,在本公开至少一实施例提供的存算一体芯片中,每个处理单元还包括:互连接口,配置为与主控单元以及第一存储单元通信。

30 本公开至少一实施例提供一种数据处理方法,应用于本公开至少一实施例提供的存算一体芯片,该数据处理方法包括:由主控单元接收整体网络参数集合,根据整体参数集合将分别用于多个处理单元的参数矩阵提供至多个处理单元各自的控制单元;多个处理单元各自的控制单元将接收的参数矩阵配置到多个处理单元各自的计算单元中;为多个处理单元提供输入数据;由多个

处理单元独立地根据接收的输入数据采用接收的参数矩阵对输入数据进行计算以得到输出数据。

附图说明

5 为了更清楚地说明本公开实施例的技术方案，下面将对实施例的附图作简单地介绍，显而易见地，下面描述中的附图仅仅涉及本公开的一些实施例，而非对本公开的限制。

图 1A 示出了矩阵向量乘法的示意图；

图 1B 为本公开至少一实施例提供的一种忆阻器阵列的示意性结构图；

10 图 2A 示出了以主控单元为中央调度单元的架构的控制流的示意图；

图 2B 示出了本公开至少一实施例提供的基于总线机制的分布式控制架构的示意图；

图 2C 示出了本公开至少一实施例提供的分布式控制架构的控制流的示意图；

15 图 2D 示出了在片上路由机制下传统的存算一体架构和本公开所提供的分布式控制架构的控制流的示意图；

图 3A 为本公开至少一实施例提供的一种分布式控制架构的存算一体芯片的示意图；

20 图 3B 示出了以忆阻器阵列输出 8 比特为例的功能单元的数据流的示意图；

图 4 示出了本公开至少一实施例提供的数据处理方法的流程图；

图 5 示出了本公开至少一实施例提供的存算一体芯片的工作流程图；

图 6 示出了一个示例性的读控制过程的流程图；以及

图 7 示出了将数据从第二存储单元读取到输入缓冲单元的示意图。

25

具体实施方式

为了使得本公开实施例的目的、技术方案和优点更加清楚，下面将结合本公开实施例的附图，对本公开实施例的技术方案进行清楚、完整地描述。显然，所描述的实施例是本公开的一部分实施例，而不是全部的实施例。基于所描述的本公开的实施例，本领域普通技术人员在无需创造性劳动的前提下所获得的所有其他实施例，都属于本公开保护的范围。

30

除非另外定义，本公开使用的技术术语或者科学术语应当为本公开所属领域内具有一般技能的人士所理解的通常意义。本公开中使用的“第一”、“第二”以及类似的词语并不表示任何顺序、数量或者重要性，而只是用来区分不同的组成部分。“包括”或者“包含”等类似的词语意指出现该词前面的元件或者物件涵盖出现在该词后面列举的元件或者物件及其等同，而不排除其他元件或者物件。“连接”或者“相连”等类似的词语并非限定于物理的或者机械的连接，而是可以包括电性的连接，不管是直接的还是间接的。“上”、“下”、“左”、“右”等仅用于表示相对位置关系，当被描述对象的绝对位置改变后，则该相对位置关系也可能相应地改变。

10 为了保持本公开实施例的以下说明清楚且简明，本公开省略了部分已知功能和已知部件的详细说明。

忆阻器（例如，阻变存储器、相变存储器、导电桥存储器等）是一种可以通过施加外部激励，调节其电导状态的非易失型器件。忆阻器作为一种二端器件，具有电阻可调节且非挥发的特性，因此被广泛应用于存算一体技术。根据
15 基尔霍夫电流定律和欧姆定律，由忆阻器构成的阵列可以并行的完成乘累加计算，且存储和计算都发生在忆阻器阵列的各个忆阻器中。基于这种计算架构，可以实现不需要大量数据搬移的存算一体计算。

目前，由忆阻器构成的交叉阵列可以并行完成矩阵矢量乘法操作，并广泛应用于加速神经网络算法中，因此存算一体芯片可以采用忆阻器阵列完成计
20 算。现有的存算一体芯片往往是面向特定的神经网络进行定制化的架构设计，能对目标任务实现高效处理，但是由于没有高效的数据流控制单元，不能够同时有效处理多种算法模型。

图 1A 示出了矩阵向量乘法的示意图。如图 1A 所示，输入向量 X （例如，包括图 1A 所示出的 x_1 、 x_2 、……、 x_m ）和参数矩阵 G （例如，包括图 1A 所
25 示出的 g_{11} 、 g_{12} 、……、 g_{1m} ，……， g_{n1} 、 g_{n2} 、……、 g_{nm} ）相乘得到输出向量 I （例如，包括图 1A 所示出的 I_1 、 I_2 、……、 I_n ）。例如， $I_1 = x_1 g_{11} + x_2 g_{12} + \dots + x_m g_{1m}$ 。

图 1B 为本公开至少一实施例提供的一种忆阻器阵列的示意性结构图。

如图 1B 所示，忆阻器阵列结构可以包括以阵列排布的多个忆阻器，将图 1A 所示的输入向量 X （例如，输入向量可以为已编码幅值、宽度或脉冲个数的电压）提供至忆阻器阵列的输入端，将图 1A 所示的参数矩阵 G 编码为忆
30 阻器电导值，利用高并行、低功耗的阵列读操作获得输出电流 I ，即，图 1A

所示的输出向量 I ，可以实现深度学习中普遍的乘累加计算，进而对矩阵向量乘进行加速。

例如，根据基尔霍夫定律，忆阻器阵列结构的输出电流可以根据下述公式得出： $I = G \times X$ 。例如， $I_1 = x_1 g_{11} + x_2 g_{12} + \dots + x_m g_{1m}$ 。上述乘累加计算过程利用物理定律实现，不同于布尔逻辑的数字电路实现方式，不需要频繁访存搬移权重数据，解决了经典计算系统的冯诺依曼瓶颈，可以实现高算力、高能效的智能计算任务。

图 1B 所示的忆阻器阵列中的阵列排布的忆阻器可以为 1T1R 结构或者 2T2R 结构，其中，1T1R 结构的忆阻器单元包括一个开关晶体管和一个忆阻器，2T2R 结构的忆阻器单元包括两个开关晶体管和两个忆阻器，本公开的实施例对采用的忆阻器的类型不作限定。

需要说明的是，本公开的实施例中采用的晶体管均可以为薄膜晶体管或场效应晶体管（例如 MOS 场效应晶体管）或其他特性相同的开关器件。这里采用的晶体管的源极、漏极在结构上可以是对称的，所以其源极、漏极在结构上可以是没有区别的。本公开的实施例对采用的晶体管的类型不作限定。

为实现通用性，可以采用基于主控单元的架构，主控单元作为中央调度单元，完成对多个计算单元的控制、配置和调度，而计算单元本身不具有调度功能。然而，在基于主控单元的架构中，主控单元和计算单元之间频繁的调度和数据交互会消耗大量的时间，导致计算效率低下。

图 2A 示出了以主控单元为中央调度单元的架构的控制流的示意图。

如图 2A 所示，在以主控单元 102 为中央调度单元的架构中，由主控单元 102 完成数据调度和任务分配，并由计算单元 103 完成计算。图 2A 中的虚线箭头表示该中央控制架构的控制流，控制流从主控单元 102 流经总线 107 到达计算单元 103。在以数据为中心的神经网络算法中，主控单元 102 和计算单元 103 之间频繁的调度和数据交互会消耗大量的时间，导致计算效率低下。

本公开至少一实施例提供一种分布式控制架构的存算一体芯片。该存算一体芯片包括：多个处理单元，每个处理单元包括：计算单元，配置为存储参数矩阵、接收输入数据并采用参数矩阵对输入数据进行计算得到输出数据；以及控制单元，配置为与计算单元相通信，且配置为将参数矩阵写入计算单元得到配置后的计算单元，并且控制配置后的计算单元采用参数矩阵对输入数据进行计算；以及主控单元，配置为与多个处理单元通信，且根据整体网络参数

集合将分别用于多个处理单元的参数矩阵提供至多个处理单元各自的控制单元。

本公开至少一实施例还提供应用于上述存算一体芯片的数据处理方法。

5 本公开上述实施例提供的分布式控制架构的存算一体芯片能够利用高效的分布式架构，在保证通用性和灵活性的同时，实现高算力和高能效。

与采用中央控制架构的存算一体芯片相比，本公开实施例所提出的存算一体芯片采用分布式控制架构，即，在图 2A 所示的传统的计算单元 103 中集成本地的控制单元，从而组成一个处理单元，但是在本公开提出的分布式控制架构中，数据调度和任务分配的工作由本地的控制单元完成，通过并行的分布
10 式调度和配置，大幅提高计算效率。

下面结合附图对本公开的实施例进行详细说明，但是本公开并不限于这些具体的实施例。

图 2B 示出了本公开至少一实施例提供的基于总线机制的分布式控制架构的示意图。

15 如图 2B 所示，主控单元 102、第一存储单元 106 以及多个处理单元 101 通过总线 107 连接，通过在多个处理单元 101 内部集成控制单元 104（参见图 2C），可以实现本地的控制和灵活的矩阵参数配置、数据流调度，将工作任务分布到每个处理单元 101 中执行，从而在兼顾了灵活性和通用性的同时，克服了图 2A 所示的中央控制架构的低效性。

20 图 2C 为图 2B 所示的分布式控制架构的控制流的示意图。

如图 2C 所示，每个处理单元 101 包括计算单元 103 和控制单元 104，与图 2A 相比，计算单元 103 不由主控单元 102 直接进行数据调度和任务分配，而是由同一个处理单元中与其相对应的控制单元 104 进行数据调度和任务分配。虚线箭头表示控制流，在图 2C 的示例中，对于分布式控制架构，控制流
25 在处理单元 101 内部，能够提高计算效率和能效。

实际上，多个处理单元 101、主控单元 102 以及第一存储单元 106 之间除了通过总线进行通信之外，还可以基于其他片上互连机制进行通信。该片上互连机制包括开关网络、片上网络（片上路由）等方式，本公开的实施例对此不作限制。

30 图 2D 示出了在片上路由机制下，传统的存算一体架构和本公开所提供的分布式控制架构的控制流的示意图。图 2D 示出的片上路由系统中，多个处理

单元 101、主控单元 102 以及第一存储单元 106 分别经过网络接口与网络节点处的特定的路由器相连，将这些单元之间的通信转换为路由器与路由器的通信。相比于其它片上系统的通信方式，片上路由机制不再受限于总线架构，因此可以扩展任意数量的路由节点。此外，在需要对片上路由系统的功能进行扩展时，只需要将设计好的功能模块通过网络接口接入，而无需重新设计网络的整体架构。对于传统的存算一体架构，控制流如标号为 1 的虚线箭头所示，控制流从主控单元 102 连接的路由节点经过多个处理单元 101 直到主控单元 102 需要控制的单元 101。对于本公开所提供的分布式架构，控制流如标号为 2 的虚线箭头所示，控制流在处理单元 101 内部，从而避免了主控单元 102 和处理单元 101 之间频繁的调度和数据交互，提高了计算效率。

由于分布式控制架构的高效率，因此本公开提出了一种采用分布式控制架构的存算一体芯片，例如用于人工智能运算。

图 3A 为本公开至少一实施例提供的一种分布式控制架构的存算一体芯片 100 的示意图。

如图 3A 所示，存算一体芯片 100 用电路（数字电路以及模拟电路）实现，包括多个处理单元 101（图中仅示出了一个处理单元）和主控单元 102，每个处理单元 101 包括计算单元 103 和控制单元 104。计算单元 103 被配置为存储参数矩阵、接收输入数据并采用参数矩阵对输入数据进行计算得到输出数据，控制单元 104 被配置为与计算单元 103 相通信，且配置为将参数矩阵写入计算单元 103 得到配置后的计算单元 103，并且控制配置后的计算单元 103 采用参数矩阵对输入数据进行计算，主控单元 102 被配置为与多个处理单元 101 通信，且根据整体网络参数集合将分别用于多个处理单元 101 的参数矩阵提供至多个处理单元 101 各自的控制单元 104。

根据本公开的至少一实施例，计算单元 103 包括至少一个忆阻器阵列 105，至少一个忆阻器阵列 105 配置为存储参数矩阵且对输入数据进行计算。

例如，忆阻器阵列 105 的结构如图 1B 所示，参数矩阵例如为图 1A 所示的参数矩阵 G ，以电导的形式存储在计算单元 103 中的至少一个忆阻器阵列 105 中。计算单元 103 接收输入数据（例如，输入数据排列对应于如图 1A 所示的输入向量 X ），采用参数矩阵对输入数据进行计算（例如，图 1A 所示的矩阵向量乘法运算）得到输出数据（例如，对应于图 1A 所示的输出向量 I ）。

在此实施例中，在计算单元 103 包括多个忆阻器阵列 105 的情形，每个

忆阻器阵列 105 的行列数量可以相同，然而，每个忆阻器阵列 105 的行和列的数量、忆阻器阵列 105 的数量均可以根据实际情况自由设计。由于忆阻器阵列中的计算是基于模拟电压和模拟电流实现的，而在芯片中数据的传输是以数字信号的形式传输的，因此一种实现方式是在忆阻器阵列的输入端和输出端分别引入数模转换器（DAC）和模数转换器（ADC）等外围电路；同时，为了节省功耗与面积开销，DAC 和 ADC 可以采用分时复用的方式设计，本公开对 DAC 和 ADC 的要求与设计不做描述。

根据本公开的实施例，控制单元 104 包括参数配置单元 1041，该参数配置单元 1040 被配置为将参数矩阵配置到计算单元 103 中，和/或根据需要从计算单元 103 获取计算单元 103 当前具有的参数矩阵的值，从而可以对存入的电导值进行校核。

例如，控制单元 104 基于参数矩阵的参数与计算单元 103 中的忆阻器阵列 105 的映射关系，将参数矩阵的参数与忆阻器阵列 105 对应的部分映射到忆阻器阵列 105 的多个忆阻器，从而实现将参数矩阵配置到计算单元 103 中。

例如，对于忆阻器阵列 105，其中包含的忆阻器面临着可靠性问题，具有不可避免的波动、噪声和状态漂移等，从而造成计算误差，影响系统的正常功能。在使用忆阻器的电导值进行计算时，由于忆阻器的波动性等非理想特性，例如忆阻器随机涨落、弛豫特性、保持特性等，会造成实际的电导值偏离理想电导值，从而使得计算结果出现偏差。因此，需要参数配置单元 1040 从计算单元 103 中获取计算单元 103 当前具有的参数矩阵的值，并与实际需要配置

的参数矩阵的值进行比较来校正计算单元 103 中的参数矩阵的值。

根据本公开的实施例，控制单元 104 还包括读控制单元 1042 和写控制单元 1043，读控制单元 1042 被配置为控制输入数据的读取以将输入数据提供到计算单元 103 中，写控制单元 1043 被配置为控制输出数据的输出。例如，读控制单元 1042 读取输入数据的过程如图 6 所示，该读控制过程主要包含四部分：读地址更新、数据读取、输入缓冲单元填充、输入比特控制，具体的将在图 6 进行说明。

例如，在输入数据的读取和输出数据的输出过程中，均需要存储单元来存储数据。根据本公开的实施例，存算一体芯片 100 还包括第一存储单元 106，该第一存储单元 106 配置为与多个处理单元 101 通信以及与主控单元 102 通信，存储整体输入数据和/或整体输出数据，在主控单元 102 的控制下，向多

个处理单元 101 分别提供多个处理单元 101 所需的输入数据, 和/或, 从多个处理单元 101 分别接收多个处理单元 101 产生的输出数据, 其中, 整体输入数据包括用于多个计算单元 103 的输入数据, 整体输出数据包括多个计算单元 103 的输出数据。第一存储单元 106 可以是半导体存储装置, 例如, 非易失存储器等。

例如, 第一存储单元 106 可以通过总线与多个处理单元 101 以及主控单元 102 连接以进行通信, 根据本公开的实施例, 存算一体芯片 100 还包括片上互连, 例如总线 107。总线是计算机各种功能部件之间传送信息的公共通信干线, 它是由导线组成的传输线束, 可以用于传输数据、数据地址和控制信号, 本公开的实施例对其类型和实现方式不做限制。

根据本公开的实施例, 每个处理单元 101 还包括互连接口, 互连接口被配置为与主控单元 102 以及第一存储单元 106 通信。例如, 连接在总线 107 上的多个处理单元 101 与总线 107 的连接电路称为总线接口, 总线接口是多个处理单元 101 与主控单元 102 以及与第一存储单元 106 交换数据的接口; 在其他片上互连类型中, 互连接口则是其他类型的接口, 用于进行数据、指令交互。

例如, 整体输入数据是指多个处理单元 101 所需的输入数据的集合, 整体输出数据是指多个处理单元 101 产生的输出数据的集合。多个处理单元 101 各自所需的输入数据以及各自产生的输出数据需要多个处理单元 101 内部的存储单元来存储。

根据本公开的实施例, 每个处理单元 101 还包括第二存储单元 109、输入缓冲单元 110 以及输出缓冲单元 111, 第二存储单元 109 被配置为从第一存储单元 106 获取输入数据或向第一存储单元 106 发送输出数据, 输入缓冲单元 110 被配置为缓存输入数据, 输出缓冲单元 111 被配置为缓存输出数据。第二存储单元 109、输入缓冲单元 110 以及输出缓冲单元 111 的大小可以根据实际情况自行设定。

例如, 在主控单元 102 的控制下, 第一存储单元 106 将输入数据经由总线 107 提供至第二存储单元 109, 第二存储单元 109 可以为各种类型的存储装置, 例如为双口 SRAM (Static Random-Access Memory, 静态随机存取存储器), 双口 SRAM 包含读端口和写端口, 写端口为总线接口 107, 读端口与忆阻器阵列 105 的输入缓冲单元 110 相连。为了提高数据读效率, 该 SRAM 的

读端口可以设置成较大位宽（例如，512 比特）。输入缓冲单元 110 的作用是将第二存储单元 109 送来的数据暂时存放，输出缓冲单元 111 的作用是将忆阻器阵列 105 产生的输出数据暂时存放，从而可以起到协调缓冲的作用，实现数据传送的同步。

- 5 整体网络参数集合由主控单元进行处理，并分发到各个处理单元；整体网络参数集合例如是对应于某一神经网络的处理层的权重矩阵的全部或部分，将其拆分后得到用于各个处理单元的参数矩阵。

例如，整体网络参数集合例如包括网络层数、卷积核尺寸、网络类型等，主控单元 102 用于根据编译好的整体网络参数集合，将分别用于多个处理单元 101 的参数矩阵提供至多个处理单元 101 的控制单元 104。下面的表 1 列出了一个示例性的具体的神经网络结构需要配置的网络参数及其含义。

表 1

参数符号	参数含义
NUM_LAYERS	卷积层的数量，也代表忆阻器阵列的分组数量，每组映射一层参数矩阵
NUM_ARRAYS[i]	第 i 层卷积层用到的忆阻器阵列个数 ($0 \leq i < \text{NUM_LAYERS}$)
SIZE_IMAGE[i]	第 i 层卷积层的输入特征图的尺寸
NUM_CHANNELS[i]	第 i 层卷积层的输入特征图的通道 (channel) 数
SIZE_KERNEL[i]	第 i 层卷积层的单个卷积核 (kernel) 尺寸
NUM_KERNELS[i]	第 i 层卷积层的卷积核的个数
STRIDE[i]	第 i 层卷积的步幅
POOLING_MODE[i]	第 i 层卷积层之后的池化模式，可选无池化、最大值和平均值池化
FC_MODE[i]	当其有效时表示用于计算全连接层

这里，对表 1 中的一些网络参数的概念进行解释。

- 15 卷积神经网络是一类包含卷积计算且具有深度结构的前馈神经网络。卷积神经网络的结构主要包括输入层、隐含层和输出层，隐含层包含卷积层、池化层和全连接层 3 类常见构筑。卷积层的功能是对输入数据进行特征提取，其内部包含多个卷积核。在卷积神经网络的术语中，3*3 的矩阵叫做“滤波器”，通过在图像上滑动滤波器并计算点乘得到的矩阵叫做“特征图”。对于同样的

输入图像，不同值的滤波器将会生成不同的特征图。在此实施例中，卷积层的数量即为忆阻器阵列的分组数量，每一组忆阻器阵列被映射一层卷积层对应的参数矩阵，每层卷积层用到的忆阻器阵列的数量的范围是 $0 \leq i < \text{NUM_LAYERS}$ (NUM_LAYERS 代表卷积层的数量)。

5 特征图的通道数用于表示图像的某种组成。例如，一个标准数字相机拍摄的图像会有三个通道（红、绿和蓝），每个通道代表一个颜色，每个通道的像素值在 0 到 255 的范围内，可以把该图像的三个通道看作是堆在一起的深度为 3 的二维矩阵。

卷积层参数包括卷积核尺寸、卷积的步幅和填充，三者共同决定了卷积层的输出特征图的尺寸。卷积核的个数和输出特征图的通道数相同。卷积核尺寸可以被指定为小于输入图像尺寸的任意值，例如 3，表示卷积核的高和宽均为 3，卷积核尺寸越大，可提取的输入特征越复杂。卷积的步幅定义了卷积核相邻两次扫过特征图时的位置的距离，卷积的步幅为 1 时，卷积核会逐个扫过特征图的元素，卷积的步幅为 n 时，卷积核会在下一次扫描跳过 $n-1$ 个像素。
10 在卷积层进行特征提取后，输出的特征图会被传递至池化层进行特征选择和信息过滤，池化过程即为降采样过程，常见的池化模式包括最大值池化、平均值池化等。全连接层位于卷积神经网络隐含层的最后部分，全连接层的作用是对提取的特征进行非线性组合以得到输出。

根据本公开的实施例，每个处理单元 101 还包括功能单元 108，功能单元
20 108 被配置为对计算单元 103 的计算结果进行处理以得到输出数据。

例如，功能单元 108 包括移位加法单元 1081 或激活单元 1082 或池化单元 1083，移位加法单元 1081 被配置为获取来自计算单元 103 的计算结果，并对计算结果进行移位运算和加法运算，激活单元 1082 被配置为对计算结果进行激活函数处理，池化单元 1083 被配置为对计算结果进行池化处理。

25 例如，输入缓冲单元 110 中缓存的输入数据会提供至对应的忆阻器阵列 105 中进行计算，可以将输入数据分成两次提供至忆阻器阵列 105 中，第一次提供输入数据的高比特位，第二次提供输入数据的低比特位，这样设计的目的是为了保证计算的精度。例如，输入数据为 8 比特，则先将高 4 比特作为忆阻器阵列 105 的输入数据，再将低 4 比特作为忆阻器阵列 105 的输入数据。
30 忆阻器阵列输出的两次计算结果分别对应高比特位输入数据的计算结果和低比特位输入数据的计算结果，然后需要通过移位加法单元 1081 对两次计算结

果进行移位加法，以将两次计算结果组合成完整的输出数据。

图 3B 示出了以忆阻器阵列输出 8 比特为例的功能单元的数据流的示意图。

如图 3B 所示，两次结果分别对应高低比特位输入数据的 8 比特有符号计算结果，将高比特位输入数据的计算结果左移 4 比特与低比特位输入数据的计算结果相加，得到 13 比特有符号的计算结果，该 13 比特的计算结果表示阵列的完整矩阵乘法结果。在此实施例中，一个卷积层对应一组忆阻器阵列，一组忆阻器阵列可以包含多个忆阻器阵列，将多个忆阻器阵列的矩阵乘法结果进行累加，得到该卷积层的最终计算结果。然后通过右移和 Clip（将结果限制在一定范围内）等操作，将结果量化为 9 比特有符号的计算结果。接着进行激活函数处理，将结果变成 8 比特无符号的计算结果。最后进行池化处理，池化结果通过上述写控制单元被送至第一存储单元或第二存储单元。

激活函数和池化是神经网络中的概念。如果不引入激活函数，神经网络中的每一层输出都是上一层输入的线性函数，无论神经网络有多少层，输出都是输入的线性组合，容易验证。因此引入非线性函数作为激活函数，这样神经网络的输出不再是输入的线性组合，从而可以逼近任意函数。在此实施例中，激活单元 1082 对计算结果进行激活函数处理可以增加计算结果的非线性。池化过程即为降采样过程，常见的池化方法包括最大值池化、平均值池化等，池化单元 1083 对计算结果进行池化处理可以使神经网络中的参数和计算数量更加可控的减小，从而可以避免过拟合。

需要说明的是，图 2 中的功能单元 108 仅是示例性的，公开的实施例对于具体的实现不做限制，可以根据实际需要增加或减少所包含的移位加法单元 1081、激活单元 1082 以及池化单元 1083。功能单元 108 可以包括除了上述移位加法单元 1081、激活单元 1082 以及池化单元 1083 以外的其它单元，或者仅包括移位加法单元 1081、激活单元 1082 以及池化单元 1083 中的部分单元，或者不包括上述单元。

图 4 示出了本公开至少一实施例提供的数据处理方法的流程图，该数据处理方法应用于图 3 所示的存算一体芯片。

如图 4 所示，该数据处理方法包括步骤 S411~S414。

步骤 S411：由主控单元接收整体网络参数集合，根据整体网络参数集合将分别用于多个处理单元的参数矩阵提供至多个处理单元各自的控制单元。

例如，整体网络参数集合的内容可以参考表 1 所示的参数符号和参数含义。

步骤 S412：多个处理单元各自的控制单元将接收的参数矩阵配置到多个处理单元各自的计算单元中。

- 5 例如，计算单元包括至少一个忆阻器阵列，控制单元基于参数矩阵的参数与计算单元中的忆阻器阵列的映射关系，将参数矩阵的参数与忆阻器阵列对应的部分映射到忆阻器阵列的多个忆阻器。

步骤 S413：为多个处理单元提供输入数据。

- 10 例如，第一存储单元用于存储整体输入数据，在主控单元的控制下，多个处理单元各自所需的输入数据经由总线从第一存储单元提供至多个处理单元。

步骤 S414：由多个处理单元独立地根据接收的输入数据采用接收的参数矩阵对输入数据进行计算以得到输出数据。

- 15 例如，在每个处理单元中，计算单元中包括的至少一个忆阻器阵列的结构如图 1B 所示，忆阻器阵列中存储有接收的参数矩阵，采用接收的参数矩阵对输入数据进行计算得到输出数据。

图 5 示出了本公开至少一实施例提供的存算一体芯片的工作流程图。

- 20 如图 5 所示，首先，存算一体芯片处于初始状态，判断整体网络参数集合是否已经被编译好，如果已经编译好，则由主控单元根据编译好的整体网络参数集合将分别用于多个处理单元的参数矩阵提供至多个处理单元各自的控制单元。

接着，由多个处理单元各自的控制单元将参数矩阵配置到多个处理单元各自的计算单元中。

接着，由多个处理单元各自的控制单元中的读控制单元进行读控制，从而将输入数据读取到多个处理单元各自的计算单元中的至少一个忆阻器阵列。

- 25 接着，由多个处理单元各自的计算单元中的至少一个忆阻器阵列对读取到的输入数据进行计算。

接着，将至少一个忆阻器阵列的计算结果提供到功能单元中进行计算，得到输出数据。

接着，由多个处理单元各自的控制单元中的写控制单元进行写控制。

- 30 例如，在写控制过程中，池化单元输出的前 NUM_KERNELS（卷积核数量）个通道的数据会被存放到输出缓冲单元（其它通道的数据均是无效数据），

然后通过总线接口写到第二存储单元或第一存储单元中。如果正在计算的是最后一层卷积层,则输出数据会被写到第一存储单元中,否则会被写到第二存储单元。

最后,判断是否是最后一层卷积层,如果是,则对存算一体芯片进行复位,5 否则继续对下一层卷积层进行读控制。

图 6 示出了一个示例性的读控制过程的流程图。如图 6 所示,该过程主要包含四个部分:读地址更新、数据读取、输入缓冲单元填充、输入比特控制,图 6 还展示了描述各部分的伪代码。

首先,多个处理单元各自的控制单元会给出输入数据的首地址(StartA),10 并且每计算完一次完整的矩阵运算后,输入数据的首地址都会更新(伪代码中,SIZE_IMAGE 表示输入图片的尺寸,K 表示卷积核尺寸,STRIDE 表示卷积的步幅,V 表示图片的通道数, $L=V*SIZE_IMAGE$,这些参数的含义已经在上文进行了说明)。而当一层卷积层全部计算完成后,便会开始下一层卷积层的读地址更新,此时以上提到的 SIZE_IMAGE、K、STRIDE、V 和 L 等参15 数会更新为下一层卷积层对应的值。

接着,当读地址对应的第二存储单元中的 ready 信号有效时,便会开始执行数据读取。数据读取的过程如图 7 所示。

在图 7 中,输入数据(输入图片)按照一行一行(或者一列一列)连续地存储在第二存储单元 109 中,对于尺寸为 $K*K$ 的卷积核,总共需要读取 K 笔20 数据(每笔数据的读取由数据首地址和数据长度表征),一笔数据的长度为 $K*V$ (V 对应输入图片的通道数),每笔数据对应卷积核中每行的输入数据。

第二存储单元 109 被划分为许多单元(称为块),对于写入第二存储单元109 的数据,会计算写入的数据的字节数,当写入的数据的字节数等于块长度或者写入的数据为最后一笔数据时,该块的 ready 信号被置为有效,此时不再25 允许对该块进行写入操作。这样设计的目的是为了保证在控制单元的读控制过程中能够及时读取到有效数据,理论上第二存储单元 109 划分的单元长度越小越好(能够缩短取得有效数据的时间),但是如此会增加第二存储单元 109 的面积和功耗,因此应当合理划分第二存储单元 109 的单元长度。

从第二存储单元 109 中读取数据时,控制单元会计算当前所需数据的首30 地址和末地址,然后查询首地址和末地址所对应的块的 ready 信号,当块的 ready 信号均为有效时执行数据读取操作,然后启动对应的忆阻器阵列的计算

操作。当第二存储单元 109 的某个块的数据在计算过程中不再使用时，控制单元发起该块的清除操作，并将该块的 ready 信号置为无效，此时该块允许写入新的数据。

回到图 6，接着，对输入缓冲单元进行填充。每个忆阻器阵列都有相应的输入缓冲单元，K 笔数据将会平均分配到该层卷积层计算所需的每个忆阻器阵列的输入缓冲单元中（伪代码中 NUM_ARRAYS 表示该层卷积层计算需要用到的忆阻器阵列的数量）。

最后，输入缓冲单元中的数据会提供至对应的忆阻器阵列中进行计算，数据可以分两次输入，第一次输入高比特位，第二次输入低比特位。

本公开实施例的分布式控制架构的存算一体芯片可以用于其他电子装置，例如人工智能处理器（AI 处理器）等，不但具有通用性和灵活性，且同时能够实现高算力和高能效。

对于本公开，还有以下几点需要说明：

（1）本公开实施例附图只涉及到与本公开实施例涉及到的结构，其他结构可参考通常设计。

（2）在不冲突的情况下，本公开的实施例及实施例中的特征可以相互组合以得到新的实施例。

以上所述仅为本公开的具体实施方式，但本公开的保护范围并不局限于此，本公开的保护范围应以所述权利要求的保护范围为准。

权利要求书

1、一种分布式控制架构的存算一体芯片，包括：

多个处理单元，每个处理单元包括：

5 计算单元，配置为存储参数矩阵、接收输入数据并采用所述参数矩阵对所述输入数据进行计算得到输出数据；以及

 控制单元，配置为与所述计算单元相通信，且配置为将所述参数矩阵写入所述计算单元得到配置后的计算单元，并且控制所述配置后的计算单元采用所述参数矩阵对所述输入数据进行计算；以及

10 主控单元，配置为与所述多个处理单元通信，且根据整体网络参数集合将分别用于所述多个处理单元的参数矩阵提供至所述多个处理单元各自的控制单元。

2、根据权利要求1所述的存算一体芯片，还包括：

 第一存储单元，配置为与所述多个处理单元通信以及与所述主控单元通
15 信，存储整体输入数据和/或整体输出数据，在所述主控单元的控制下，向所述多个处理单元分别提供所述多个处理单元所需的输入数据，和/或，从所述多个处理单元分别接收所述多个处理单元产生的输出数据，

 其中，所述整体输入数据包括用于所述多个计算单元的输入数据，所述整体输出数据包括所述多个计算单元的输出数据。

20 3、根据权利要求2所述的存算一体芯片，还包括：

 片上互连，配置为连接所述主控单元、所述第一存储单元、以及所述多个处理单元以进行通信。

4、根据权利要求1-3任一项所述的存算一体芯片，其中，所述计算单元包括：

25 至少一个忆阻器阵列，其中，所述至少一个忆阻器阵列配置为存储所述参数矩阵且对所述输入数据进行计算。

5、根据权利要求4所述的存算一体芯片，其中，所述控制单元包括：

 参数配置单元，配置为将所述参数矩阵配置到所述计算单元中，和/或从所述计算单元获取所述计算单元当前具有的参数矩阵的值。

30 6、根据权利要求5所述的存算一体芯片，其中，所述控制单元还包括：

 读控制单元，配置为控制所述输入数据的读取以将所述输入数据提供到

所述计算单元中；

写控制单元，配置为控制所述输出数据的输出。

7、根据权利要求 1-6 任一所述的存算一体芯片，其中，每个所述处理单元还包括功能单元，

5 其中，所述功能单元配置为对所述计算单元的计算结果进行处理以得到所述输出数据。

8、根据权利要求 7 所述的存算一体芯片，其中，所述功能单元包括：

移位加法单元，配置为获取来自所述计算单元的计算结果，并对所述计算结果进行移位运算和加法运算；或

10 激活单元，配置为对所述计算结果进行激活函数处理；或

池化单元，配置为对所述计算结果进行池化处理。

9、根据权利要求 2 或 3 所述的存算一体芯片，其中，每个所述处理单元还包括：

15 第二存储单元，配置为从所述第一存储单元获取所述输入数据或向所述第一存储单元发送所述输出数据；

输入缓冲单元，配置为缓存所述输入数据；

输出缓冲单元，配置为缓存所述输出数据。

10、根据权利要求 2 或 3 所述的存算一体芯片，其中，每个所述处理单元还包括：

20 互连接口，配置为与所述主控单元以及所述第一存储单元通信。

11、一种数据处理方法，应用于权利要求 1-10 任一项所述的存算一体芯片，所述数据处理方法包括：

25 由所述主控单元接收所述整体网络参数集合，根据所述整体参数集合将分别用于所述多个处理单元的参数矩阵提供至所述多个处理单元各自的控制单元；

所述多个处理单元各自的控制单元将接收的参数矩阵配置到所述多个处理单元各自的计算单元中；

为所述多个处理单元提供所述输入数据；

30 由所述多个处理单元独立地根据接收的输入数据采用所述接收的参数矩阵对所述输入数据进行计算以得到所述输出数据。

$$\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1m} \\ g_{21} & g_{22} & \cdots & g_{2m} \\ \cdots & \cdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nm} \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_m \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \\ \cdots \\ I_n \end{pmatrix}^T$$

输入矩阵

输入向量

输出向量

图 1A

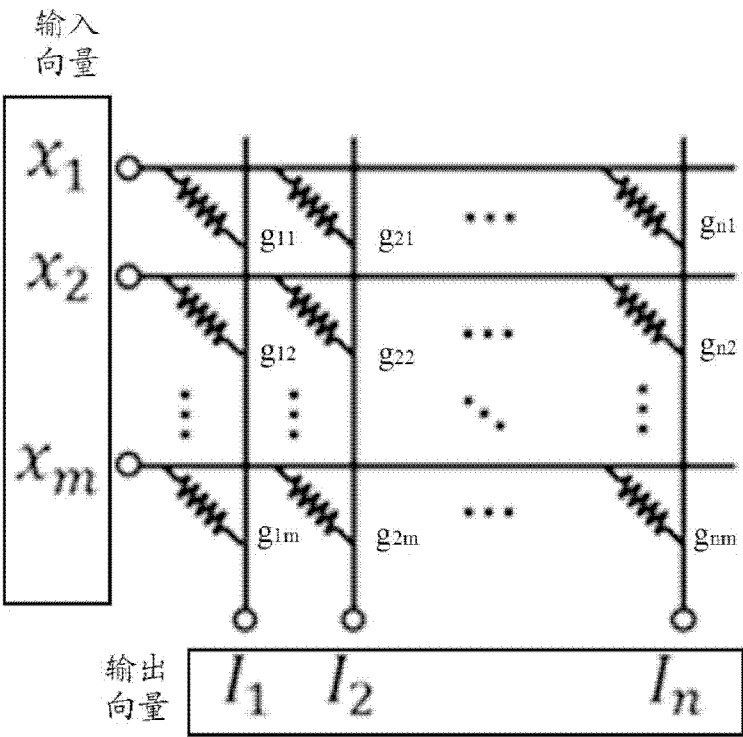


图 1B

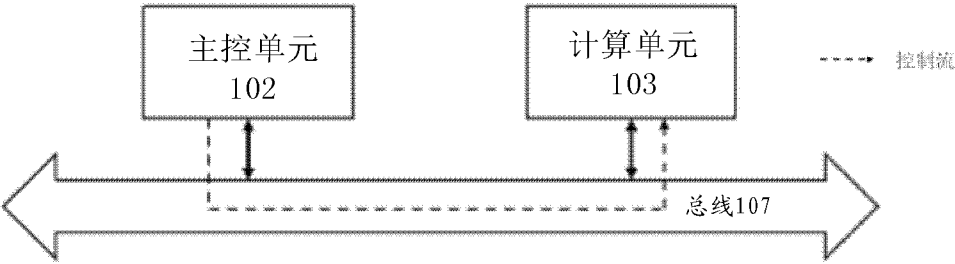


图 2A

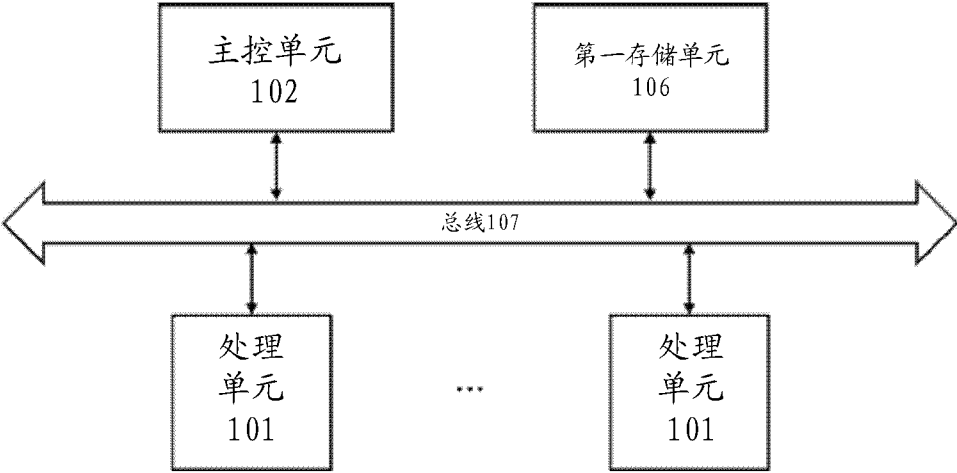


图 2B

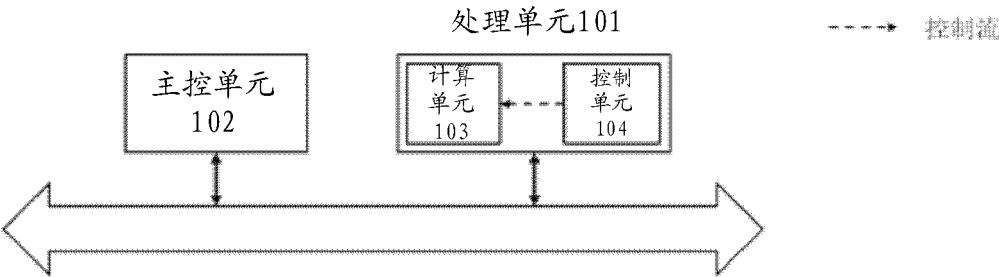


图 2C

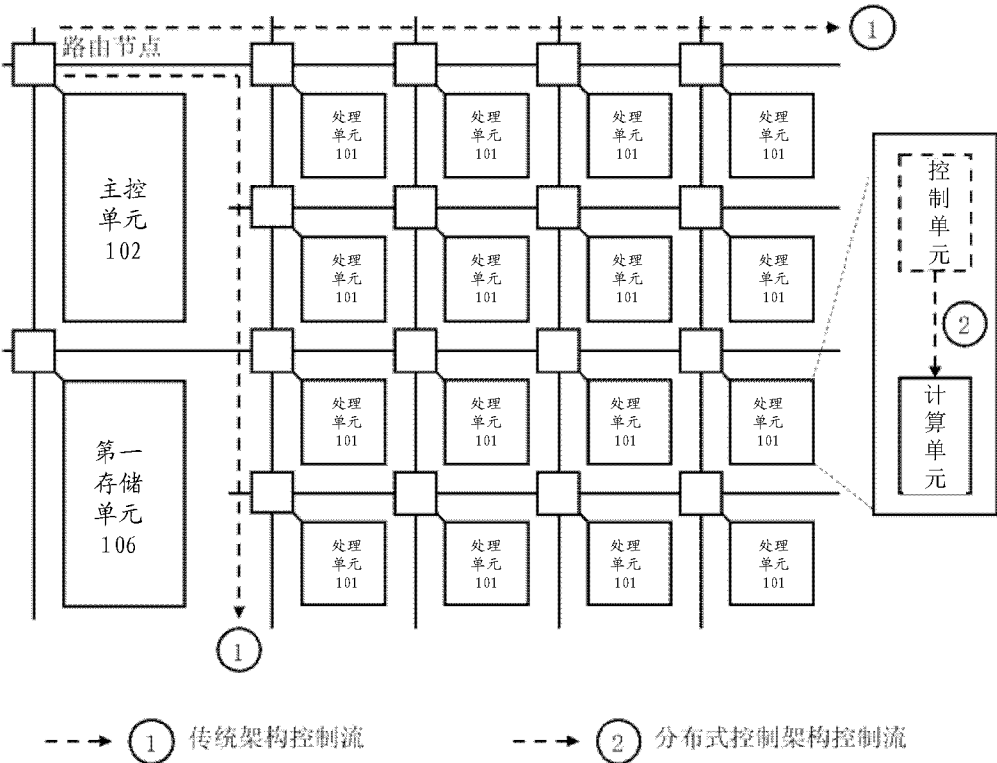


图 2D

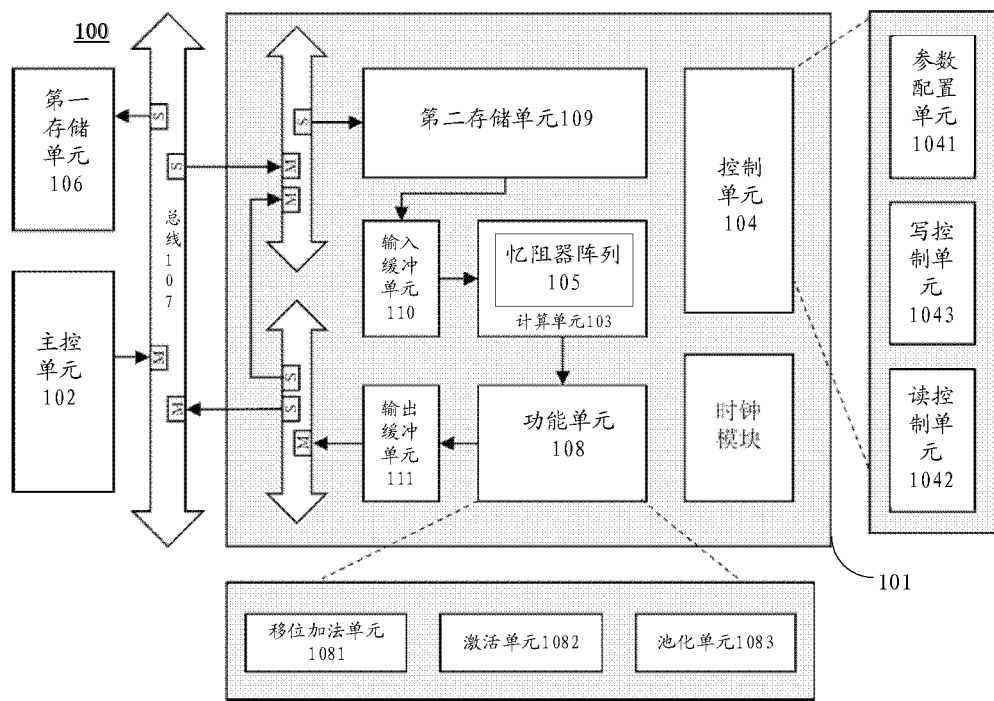


图 3A

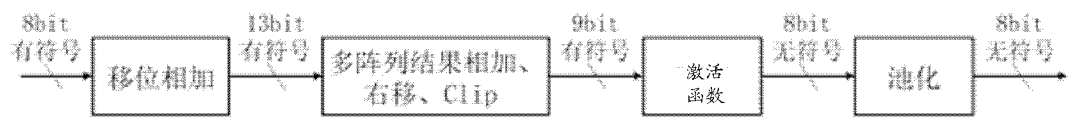


图 3B

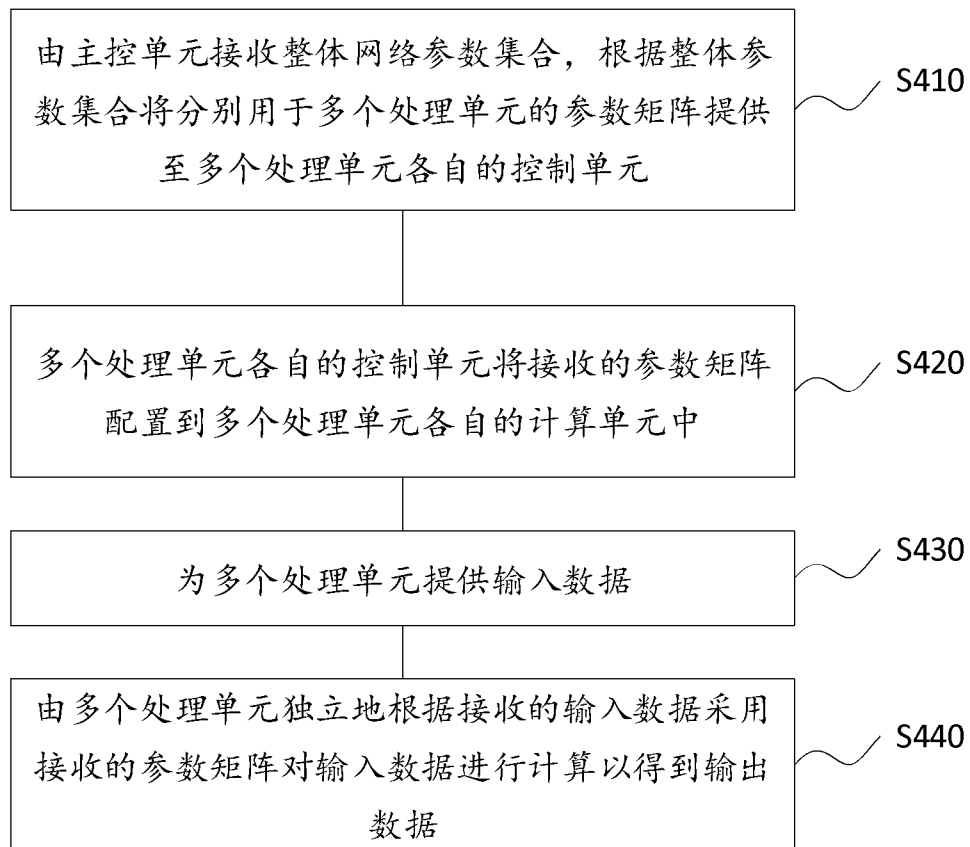


图 4

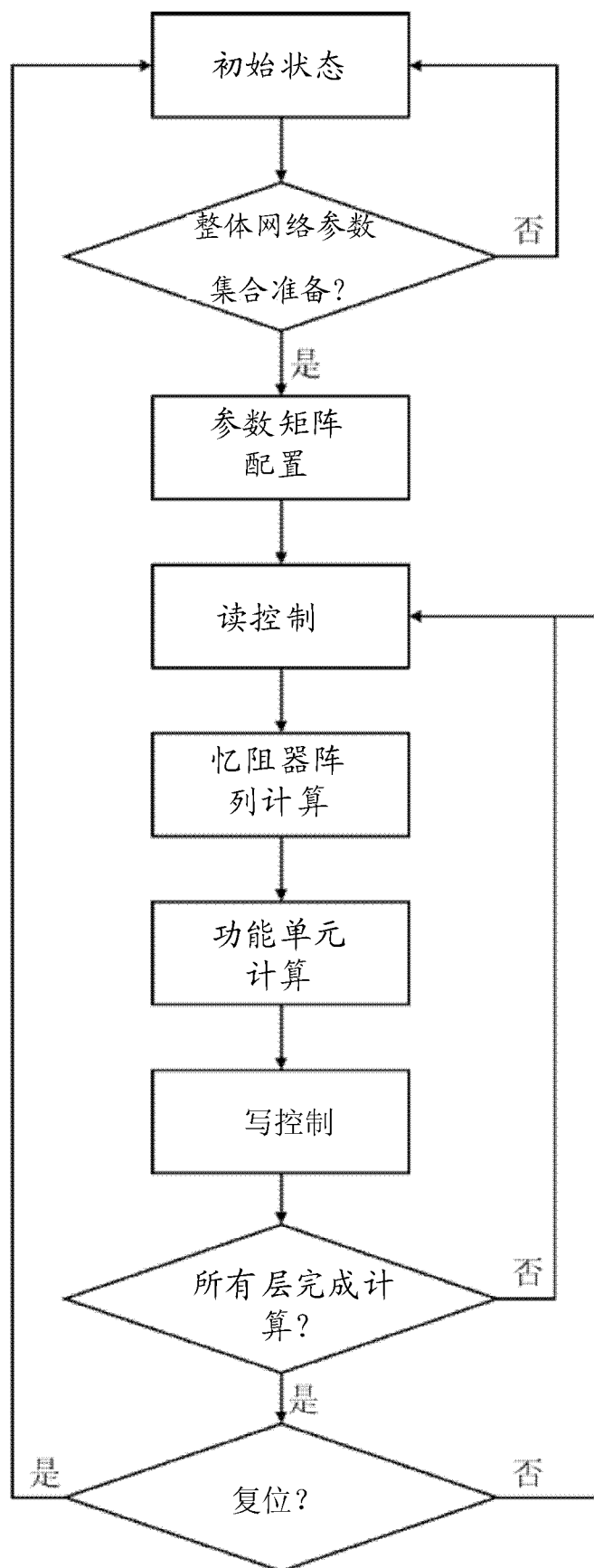


图 5

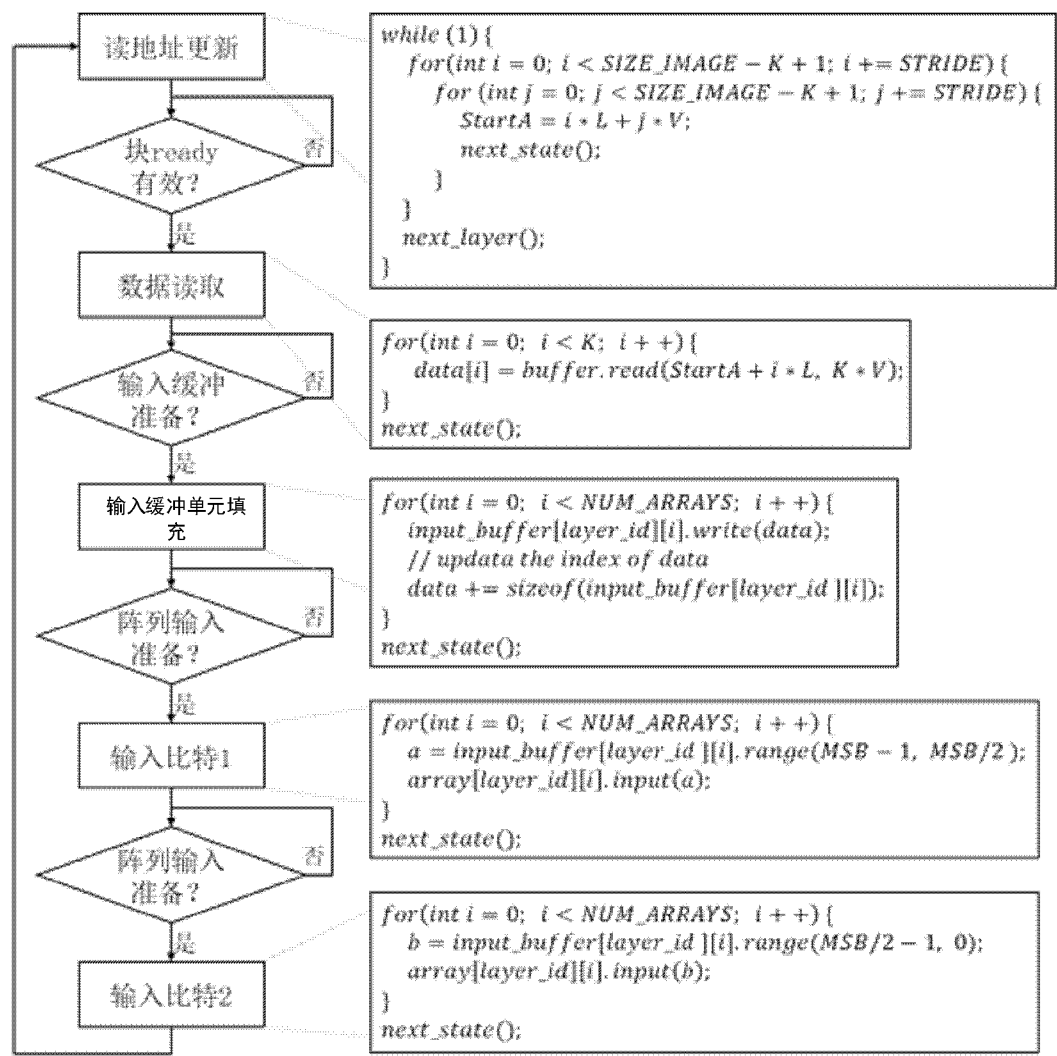


图 6

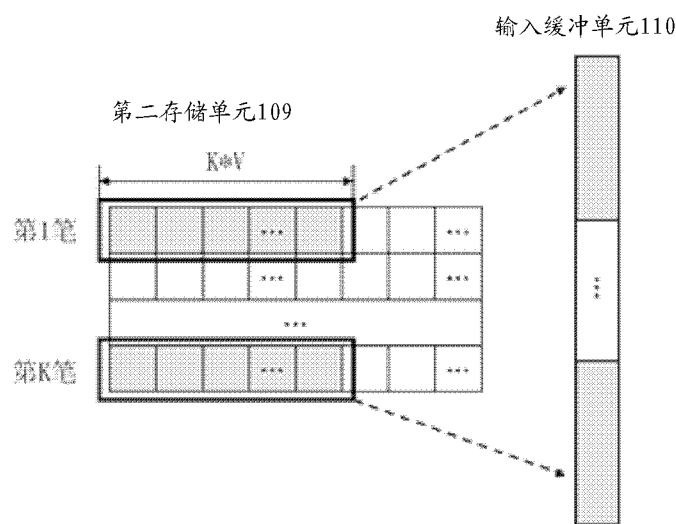


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/137844

A. CLASSIFICATION OF SUBJECT MATTER

G06F 15/78(2006.01)i; G06K 9/62(2022.01)i; G06N 3/04(2006.01)i; G06N 3/08(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F G06K G06N

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

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

CNPAT, CNKI, WPI, EPODOC, IEEE: 分布式, 存算一体, 芯片, 参数, 矩阵, 控制, 计算, 主控, 忆阻器, 阵列, distributed architecture, storage and calculation integrated, chip, parameter, matrix, calculate, control, memristor, array

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 112836814 A (TSINGHUA UNIVERSITY) 25 May 2021 (2021-05-25) description, paragraphs [0004]-[0017] and [0055]	1-11
Y	CN 113077829 A (TSINGHUA UNIVERSITY) 06 July 2021 (2021-07-06) description, paragraph [0004]	1-11
A	CN 112825153 A (HUAWEI TECHNOLOGIES CO., LTD. et al.) 21 May 2021 (2021-05-21) entire document	1-11
A	CN 110990060 A (BEIJING HANNUO SEMICONDUCTOR TECHNOLOGY CO., LTD. et al.) 10 April 2020 (2020-04-10) entire document	1-11
A	WO 2021115262 A1 (NANJING SIMIND SEMICONDUCTOR LTD.) 17 June 2021 (2021-06-17) entire document	1-11
A	US 2020342301 A1 (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY) 29 October 2020 (2020-10-29) entire document	1-11



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

09 June 2022

Date of mailing of the international search report

22 June 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/137844

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112836814	A	25 May 2021	None			
CN	113077829	A	06 July 2021	None			
CN	112825153	A	21 May 2021	WO	2021098821	A1	27 May 2021
CN	110990060	A	10 April 2020	None			
WO	2021115262	A1	17 June 2021	TW	202123032	A	16 June 2021
				CN	113033759	A	25 June 2021
US	2020342301	A1	29 October 2020	WO	2020052342	A1	19 March 2020
				CN	109460817	A	12 March 2019

国际检索报告

国际申请号

PCT/CN2021/137844

A. 主题的分类

G06F 15/78(2006.01)i; G06K 9/62(2022.01)i; G06N 3/04(2006.01)i; G06N 3/08(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F G06K G06N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT, CNKI, WPI, EPDOC, IEEE: 分布式, 存算一体, 芯片, 参数, 矩阵, 控制, 计算, 主控, 忆阻器, 阵列, distributed architecture, storage and calculation integrated, chip, parameter, matrix, calculate, control, memristor, array

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 112836814 A (清华大学) 2021年5月25日 (2021 - 05 - 25) 说明书第[0004]-[0017], [0055]段	1-11
Y	CN 113077829 A (清华大学) 2021年7月6日 (2021 - 07 - 06) 说明书第[0004]段	1-11
A	CN 112825153 A (华为技术有限公司 等) 2021年5月21日 (2021 - 05 - 21) 全文	1-11
A	CN 110990060 A (北京瀚诺半导体科技有限公司 等) 2020年4月10日 (2020 - 04 - 10) 全文	1-11
A	WO 2021115262 A1 (NANJING SIMIND SEMICONDUCTOR LTD.) 2021年6月17日 (2021 - 06 - 17) 全文	1-11
A	US 2020342301 A1 (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY) 2020年10月29日 (2020 - 10 - 29) 全文	1-11

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2022年6月9日

国际检索报告邮寄日期

2022年6月22日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

郭婉莹

电话号码 86-(10)-53961361

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/137844

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	112836814	A	2021年5月25日	无			
CN	113077829	A	2021年7月6日	无			
CN	112825153	A	2021年5月21日	WO	2021098821	A1	2021年5月27日
CN	110990060	A	2020年4月10日	无			
WO	2021115262	A1	2021年6月17日	TW	202123032	A	2021年6月16日
				CN	113033759	A	2021年6月25日
US	2020342301	A1	2020年10月29日	WO	2020052342	A1	2020年3月19日
				CN	109460817	A	2019年3月12日