(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118974698 A
(43)申请公布日 2024.11.15

(21)申请号 202380031660.3
(22)申请日 2023.03.13
(30)优先权数据
2205498.5 2022.04.13 GB
(85)PCT国际申请进入国家阶段日
2024.09.29
(86)PCT国际申请的申请数据
PCT/GB2023/050587 2023.03.13
(87)PCT国际申请的公布数据
WO2023/199015 EN 2023.10.19
(71)申请人 ARM有限公司
地址 英国剑桥
(72)发明人 伊莲娜·米兰诺维奇
克劳迪奥·马蒂诺
奈杰尔·约翰·斯蒂芬斯
阿诺德·菲利普·克劳德·格拉塞特
贾亚斯里·桑卡拉纳拉亚南
(74)专利代理机构 北京东方亿思知识产权代理有限责任公司 11258
专利代理师 杨佳婧
(51)Int.Cl.
*G06F 9/30*(2006.01)

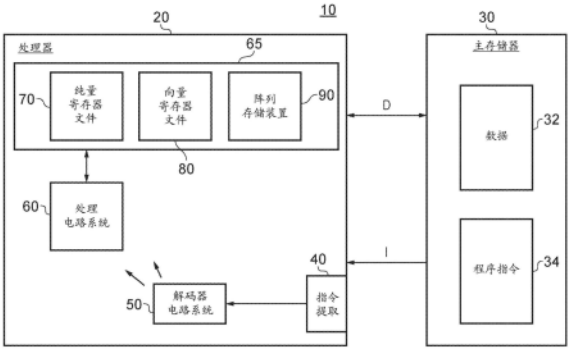
权利要求书3页 说明书17页 附图18页

(54)发明名称

用于处置存储在阵列存储装置中的数据元素的技术

(57)摘要

提供了一种设备,包括:处理电路系统,该处理电路系统用以执行操作;指令解码器电路系统,该指令解码器电路系统用以解码指令以控制该处理电路系统,从而执行由该指令指定的该操作;和阵列存储装置,该阵列存储装置包括用以存储数据元素的存储元件。该阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行该操作时该处理电路系统能够访问该数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的。该指令解码器电路系统被布置成响应于解码识别在该阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令,还解码被布置成对所识别多个数据元素向量进行操作的后续累加指令,以及控制该处理电路系统以执行由该累加指令指定的累加运算的非累加变体,以产生用于存储在该阵列存储装置内的该所识别多个向量中的结果数据元素。




CN 118974698 A

1.一种设备,所述设备包括:

处理电路系统,所述处理电路系统用以执行操作;

指令解码器电路系统,所述指令解码器电路系统用以解码指令,从而控制所述处理电路系统执行由所述指令指定的所述操作;和

阵列存储装置,所述阵列存储装置包括用以存储数据元素的存储元件,所述阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行所述操作时所述处理电路系统能够存取所述数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;

其中所述指令解码器电路系统被布置成响应于解码识别在所述阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令,还解码被布置成对所识别多个数据元素向量进行操作的后续累加指令,以及控制所述处理电路系统以执行由所述累加指令指定的累加运算的非累加变体,以产生用于存储在所述阵列存储装置内的所识别多个向量中的结果数据元素。

2.根据权利要求1所述的设备,其中:

所述阵列存储装置包括在第一阵列方向延伸的多个阵列向量寄存器,并且在所述阵列存储装置内的所识别多个向量是由所述阵列存储装置的群组的多个阵列向量寄存器提供;

所述给定二维阵列数据元素包括存储在所述群组的多个阵列向量寄存器内的所述数据元素;

所述后续累加指令指定处理运算,所述处理运算包括待对所述所识别多个数据元素向量执行的累加运算;并且

所述零向量指令与所述后续累加指令组合使用以实现由所述处理运算的非累加变体的所述处理电路系统的执行。

3.根据权利要求1或权利要求2所述的设备,其中所述零向量指令包括用以识别在所述阵列存储装置内的所述给定二维阵列数据元素的所述多个数据元素向量的向量识别字段。

4.根据权利要求1至3中任一项所述的设备,其中所述零向量指令包括用以识别预测信息的预测字段,所述预测信息用以识别多个所识别向量内的哪些存储元件待设定为逻辑零值。

5.根据权利要求4所述的设备,其中所述零向量指令还包括用以识别在所述多个所识别向量内的每个数据元素的大小的大小字段。

6.一种处置设备的阵列存储装置内的数据元素的方法,所述方法包括:

利用处理电路系统执行操作;

利用指令解码器电路系统解码指令,以控制所述处理电路系统执行由所述指令指定的所述操作;

在所述阵列存储装置内提供存储元件以存储数据元素,所述阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行所述操作时所述处理电路系统能够存取所述数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;以及

利用所述指令解码器电路系统以响应于解码识别在所述阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令,还解码被布置成对所识别多个数据元素向量进行操作的后续累加指令,以及控制所述处理电路系统以执行由所述累加指令指定的

累加运算的非累加变体,以产生用于存储在所述阵列存储装置内的所识别多个向量中的结果数据元素。

7.一种计算机程序,所述计算机程序用于控制主机数据处理设备以提供指令执行环境,所述计算机程序包括:

处理程序逻辑,所述处理程序逻辑用以执行操作;

指令解码程序逻辑,所述指令解码程序逻辑用以解码指令,从而控制所述处理程序逻辑执行由所述指令指定的所述操作;和

阵列存储装置仿真程序逻辑,所述阵列存储装置仿真程序逻辑用以仿真包括用以存储数据元素的存储元件的阵列存储装置,所述阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行所述操作时所述处理程序逻辑能够存取所述数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;

其中所述指令解码程序逻辑被布置成响应于解码识别在所述阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令,还解码被布置成对所识别多个数据元素向量进行操作的后续累加指令,以及控制所述处理程序逻辑以执行由所述累加指令指定的累加运算的非累加变体,以产生用于存储在所述阵列存储装置内的所识别多个向量中的结果数据元素。

8.一种设备,所述设备包括:

处理电路系统,所述处理电路系统用以执行操作;

指令解码器电路系统,所述指令解码器电路系统用以解码指令,从而控制所述处理电路系统执行由所述指令指定的所述操作;以及

阵列存储装置,所述阵列存储装置包括用以存储数据元素的存储元件,所述阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行所述操作时所述处理电路系统能够存取所述数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;

其中所述指令解码器电路系统被布置成响应于解码识别在所述阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令来控制所述处理电路系统,以将用以存储所识别多个向量的所述数据元素的所述阵列存储装置的所述存储元件设定为逻辑零值。

9.一种处置设备的阵列存储装置内的数据元素的方法,所述方法包括:

利用处理电路系统执行操作;

利用指令解码器电路系统解码指令,以控制所述处理电路系统执行由所述指令指定的所述操作;以及

在阵列存储装置内提供存储元件以存储数据元素,所述阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行所述操作时所述处理电路系统能够存取所述数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;

其中所述指令解码器电路系统响应于解码识别在所述阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令来控制所述处理电路系统,以将用以存储所识别多个向量的所述数据元素的所述阵列存储装置的所述存储元件设定为逻辑零值。

10.一种计算机程序,所述计算机程序用于控制主机数据处理设备以提供指令执行环

境,所述计算机程序包括:

处理程序逻辑,所述处理程序逻辑用以执行操作;

指令解码程序逻辑,所述指令解码程序逻辑用以解码指令,从而控制所述处理程序逻辑执行由所述指令指定的所述操作;以及

阵列存储装置仿真程序逻辑,所述阵列存储装置仿真程序逻辑用以仿真包括用以存储数据元素的存储元件的阵列存储装置,所述阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行所述操作时所述处理程序逻辑能够存取所述数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;

其中所述指令解码程序逻辑被布置成响应于解码识别在所述阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令来控制所述处理程序逻辑,以将用以存储所识别多个向量的所述数据元素的所述阵列存储装置的所述存储元件设定为逻辑零值。

# 用于处置存储在阵列存储装置中的数据元素的技术

## 背景技术

[0001] 本技术涉及数据处理领域,并且更具体地涉及处置存储在阵列存储装置中的数据元素。

[0002] 一些现代数据处理系统可提供用于存储一个或多个二维阵列数据元素的阵列存储装置,当执行数据处理操作时可由该数据处理系统的处理电路系统存取该数据元素。此可提供用于执行多种不同类型的操作的有效机制,例如包括累加函数的操作,其中该累加输出可维持在二维阵列数据元素内。

[0003] 然而,为了充分利用从使用此类阵列存储装置可能实现的效率增益,提供一种用以释放该阵列存储装置的资源以供结合后续操作使用的有效机制是有益的。

## 发明内容

[0004] 根据一个示例性布置,提供一种设备,该设备包括:处理电路系统,该处理电路系统用以执行操作;指令解码器电路系统,该指令解码器电路系统用以解码指令,从而控制该处理电路系统执行由该指令指定的该操作;和阵列存储装置,该阵列存储装置包括用以存储数据元素的存储元件,该阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行该操作时该处理电路系统能够存取该数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;其中该指令解码器电路系统被布置成响应于解码识别在该阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令,还解码被布置成对所识别多个数据元素向量进行操作的后续累加指令,以及控制该处理电路系统以接着执行由该累加指令指定的累加运算的非累加变体,以产生用于存储在该阵列存储装置内的该所识别多个向量中的结果数据元素。

[0005] 在另一示例性布置中,提供一种处置设备的阵列存储装置内的数据元素的方法,该方法包括:利用处理电路系统执行操作;利用指令解码器电路系统解码指令,以控制该处理电路系统执行由该指令指定的该操作;在该阵列存储装置中提供存储元件以存储数据元素,该阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行该操作时该处理电路系统能够存取该数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;以及利用该指令解码器电路系统以响应于解码识别在该阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令,还解码被布置成对该所识别多个数据元素向量进行操作的后续累加指令,以及控制该处理电路系统以执行由该累加指令指定的累加运算的非累加变体,以产生用于存储在该阵列存储装置内的该所识别多个向量中的结果数据元素。

[0006] 在更进一步示例性布置中,提供一种计算机程序,其用于控制主机数据处理设备以提供指令执行环境,该计算机程序包括:处理程序逻辑,该处理程序逻辑用以执行操作;指令解码程序逻辑,该指令解码程序逻辑用以解码指令,从而控制该处理程序逻辑执行由该指令指定的该操作;和阵列存储装置仿真程序逻辑,该阵列存储装置仿真程序逻辑用以仿真包括用以存储数据元素的存储元件的阵列存储装置,该阵列存储装置被布置成存储至

少一个二维阵列数据元素,当执行该操作时该处理程序逻辑能够存取该数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;其中该指令解码程序逻辑被布置成响应于解码识别在该阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令,还解码被布置成对该所识别多个数据元素向量进行操作的后续累加指令,以及控制该处理程序逻辑以接着执行由该累加指令指定的累加运算的非累加变体,以产生用于存储在该阵列存储装置内的该所识别多个向量中的结果数据元素。

[0007] 在另一示例性布置中,提供一种设备,该设备包括:处理电路系统,该处理电路系统用以执行操作;指令解码器电路系统,该指令解码器电路系统用以解码指令,从而控制该处理电路系统执行由该指令指定的该操作;和阵列存储装置,该阵列存储装置包括用以存储数据元素的存储元件,该阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行该操作时该处理电路系统能够存取该数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;其中该指令解码器电路系统被布置成响应于解码识别在该阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令来控制该处理电路系统,以将用以存储该所识别多个向量的该数据元素的该阵列存储装置的该存储元件设定为逻辑零值。

[0008] 在又另一示例性布置中,提供一种处置设备的阵列存储装置内的数据元素的方法,该方法包括:利用处理电路系统执行操作;利用指令解码器电路系统解码指令,以控制该处理电路系统执行由该指令指定的该操作;以及在阵列存储装置中提供存储元件以存储数据元素,该阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行该操作时该处理电路系统能够存取该数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;其中该指令解码器电路系统响应于解码识别在该阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令来控制该处理电路系统,以将用以存储该所识别多个向量的该数据元素的该阵列存储装置的该存储元件设定为逻辑零值。

[0009] 在更进一步示例性布置中,提供一种用于控制主机数据处理设备以提供指令执行环境的计算机程序,该计算机程序包括:处理程序逻辑,该处理程序逻辑用以执行操作;指令解码程序逻辑,该指令解码程序逻辑用以解码指令,从而控制该处理程序逻辑执行由该指令指定的该操作;和阵列存储装置仿真程序逻辑,该阵列存储装置仿真程序逻辑用以仿真包括用以存储数据元素的存储元件的阵列存储装置,该阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行该操作时该处理程序逻辑能够存取该数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的;其中该指令解码程序逻辑被布置成响应于解码识别在该阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令来控制该处理程序逻辑,以将用以存储该所识别多个向量的该数据元素的该阵列存储装置的该存储元件设定为逻辑零值。

## 附图说明

[0010] 本技术将仅通过图示、参照如随附图式中所示出的其示例来进一步地描述,其中:

[0011] 图1为根据一个示例性具体实施的设备的框图;

[0012] 图2显示可在设备内提供的架构寄存器的示例,该架构寄存器包括用于存储向量运算元的向量寄存器和用于存储2D阵列数据元素的阵列寄存器,该阵列寄存器包括阵列寄

存器的物理具体实施的示例;

[0013] 图3A和图3B示出其中给定2D阵列数据元素可为非方形或方形的示例;

[0014] 图4A和图4B示意性地示出根据一个示例性具体实施的如何对阵列存储装置内的方形2D阵列执行存取;

[0015] 图5A示意性地示出根据一个示例性具体实施的在移动和零指令内所提供的字段,并且图5B示意性地示出可用于一个特定示例性具体实施中以实施移动和零指令的向量识别字段的子字段;

[0016] 图6为示出根据一个示例性具体实施的可如何处置移动和零指令的流程图;

[0017] 图7示出可对在阵列存储装置内所提供的数据元素进行操作的一个示例性指令序列,其中该指令序列包括本文中所描述的移动和零指令的数个实例;

[0018] 图8示意性地示出可执行的有限脉冲响应(finite impulse response,FIR)滤波操作;

[0019] 图9A至图9D示出根据一个示例性具体实施的在执行2D图像滤波操作时可如何使用阵列存储装置;

[0020] 图10示出可对在阵列存储装置内所提供的数据元素进行操作的替代示例性指令序列,其中该指令序列包括本文中所描述的移动和零指令的数个实例;

[0021] 图11为示出根据一个示例性具体实施的可如何处置零向量指令的流程图;

[0022] 图12示意性地示出根据一个示例性具体实施的可在零向量指令内提供的字段;

[0023] 图13为示出根据替代示例性具体实施的可如何处置零向量指令的流程图;并且

[0024] 图14示出可使用的模拟器具体实施。

## 具体实施方式

[0025] 在一示例性布置中,提供一种设备,该设备具有:处理电路系统,该处理电路系统用于执行操作;和指令解码器电路系统,该指令解码器电路系统用于解码指令以控制该处理电路系统,从而执行由该指令指定的该操作。还提供一种阵列存储装置,该阵列存储装置包括用以存储数据元素的存储元件。该阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行该操作时该处理电路系统可存取该数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的。

[0026] 如前述,使用阵列存储装置可提供用于执行某些类型的操作的非常有效的机制,例如累加运算。累加运算可仅执行累加函数,但替代地,除了累加函数以外,还可并入额外处理(纯粹举例来说,累加运算可执行形式A＝A+B*C的乘积累加函数)。为了产生最大的可实现的潜在效率优势,可能需要具有用于在数据元素不再经受使用阵列存储装置执行的运算时将数据元素移出阵列存储装置的有效率机制,并且还需要释放阵列存储装置的相关联的存储元件使得它们可供与后续运算相关联地使用的有效率机制。

[0027] 根据一个示例性具体实施,提供可实现显著改善此类过程的效率的移动和零指令。具体来说,在一个示例性具体实施中,该指令解码器电路系统可被布置成响应于识别在该阵列存储装置内的给定二维阵列数据元素的一个或多个数据元素向量的移动和零指令来控制该处理电路系统,以将该一个或多个所识别向量的该数据元素从该阵列存储装置移动至目的地存储装置,并且将用以存储该一个或多个所识别向量的该数据元素的该阵列存

储装置的该存储元件设定为逻辑零值。

[0028] 因此,根据上述技术,可指定单一指令,当该单一指令执行时引起在给定二维阵列内的一个或多个所识别数据元素向量中的数据元素被移出该阵列存储装置,并且额外引起用于存储该数据元素的该阵列存储装置的该相关联存储元件被清除为逻辑零值,从而准备该存储元件以供后续运算中使用。

[0029] 这可显著改善性能。具体来说,将该数据元素移出该阵列存储装置的动作以及准备该相关联的存储元件以供再使用的动作本身不执行有用的运算,并且因此可视为与该阵列存储装置的使用相关联的负荷。通过允许单一指令引起该数据元素两者被移动并且将该相关联的存储元件清除为逻辑零值,此可显著降低该负荷。

[0030] 具体来说,先前已知技术应该需要执行至少一个移动指令以将所需数据元素从该阵列存储装置移出至指定目的地存储装置,并且其后将需要一个或多个额外移动指令以将逻辑零值的一个或多个向量从一个或多个来源向量寄存器移动至该阵列存储装置的相关存储元件中。此因此建立必须逐一执行的相依性指令序列。例如,仅考虑其中第一移动指令用于将该一个或多个数据元素向量从该阵列存储装置移出、且接着使用第二移动指令将逻辑零值的向量从指定来源向量寄存器移动至该阵列存储装置的该相关存储元件中的简单情况,明确地存在需要逐一执行的两个相依的移动指令,并且此指令相依性经由使用新的移动和零指令而移除。

[0031] 进一步,已发现,在一些具体实施中,与执行经组合的移动和归零操作相关联的硬件成本可与仅进行标准移动操作以将该数据元素向量从该阵列存储装置移出相关联的硬件成本相同,并且因此可有效地获得该相关存储元件的归零而无额外的硬件负荷。进一步,在一示例性具体实施中,已发现经组合的移动和归零操作的性能(执行速度)与仅执行单移动操作相同。

[0032] 此外,本技术的使用避免在向量寄存器内存储逻辑零值,否则将需要用作为用以将该逻辑零值移动至该阵列存储装置的移动指令的来源运算元,因此释放该向量寄存器文件内的一个或多个向量寄存器。

[0033] 除此之外,可以看到性能经改善,因为不需要分开的移动指令以执行归零功能性。

[0034] 已发现,此类方法高度有益于阵列存储装置的许多示例性使用案例。例如,阵列存储装置经常用以累加在执行累加运算的数次迭代时产生的结果,并且当产生最终累加结果时,该最终累加结果一般例如通过移动至设备内所提供的一个或多个向量寄存器而移出阵列存储装置。当阵列存储装置正在用于累加运算的执行时,仅如果该存储元件首先被设定为逻辑零值时才可能再使用正在存储该最终累加结果的该存储元件以用于一系列新的累加运算,并且使用如本文中所描述的移动和零指令可使此能够有效地实现。

[0035] 因此,在一示例性具体实施中,该处理电路系统可被布置成执行累加运算的多次迭代,并且使用该给定二维阵列数据元素,以维持在执行该累加运算时产生的累加结果,其中在该累加运算的给定迭代之后,在该给定二维阵列数据元素中的至少一个给定数据元素向量被布置成存储最终累加结果,同时该给定二维阵列数据元素中的剩余数据元素向量被布置成存储中间累加结果。在此类具体实施中,该移动和零指令可被布置成识别该至少一个给定数据元素向量且可在该累加运算的该给定迭代之后执行,以使该处理电路系统将该至少一个给定向量的该最终累加结果从该阵列存储装置移动至该目的地存储装置,并且清

除用以存储该至少一个给定向量的该最终累加结果的该阵列存储装置的该存储元件,以释放该存储元件以用于后续累加运算中。

[0036] 应注意,上文所提及的累加运算可仅执行累加函数(例如,形式A＝A+B),但更一般来说还可涉及除了累加函数之外的一些额外的处理运算。因此,累加运算可包括处理运算,该处理运算经执行以产生处理运算结果值,然后,该处理运算结果值与在该阵列存储装置的相关联的存储元件中的现有数据元素值累加,以建立待存储在该阵列存储装置的该相关联的存储元件内的新数据元素值。纯粹举例来说,上文所提及的累加运算可为乘积累加运算(例如,形式A＝A+B*C)。

[0037] 存在可使用上文所提及的累加功能性的各种类型的数据处理运算,并且对于其使用阵列存储装置可提供有效的具体实施技术。在一个特定示例性使用案例中,累加运算的该多次迭代为用于对阵列的输入数据元素实施有限脉冲响应(FIR)滤波操作的处理和累加运算,并且该阵列存储装置内的该给定二维阵列数据元素可用以维持在该FIR滤波操作的执行期间产生的阵列的输出数据元素。该处理电路系统可被布置成在该累加运算的每个迭代期间处理输入数据元素的单一向量,并且产生用于在该阵列的输出数据元素的多个向量内累加的输出数据元素。

[0038] 介于输入数据元素与输出数据元素之间的对应性可取决于具体实施而变化。例如,输入数据元素的一个向量可与输出数据元素的多个向量相关联。除此之外,输出数据元素的该多个向量可经布置在该阵列存储装置内的水平和垂直方向中的任一者或两者(以支持其中可在水平和垂直方向两者存取该向量的具体实施,该二维阵列数据元素一般将为二维方形阵列数据元素)。进一步,该输入数据元素与该输出数据元素的大小可不同。

[0039] 如上文所描述的此类技术可利用外积方法,以使用方形阵列数据元素来运算通过滑动窗技术所实施的FIR滤波。此类技术一般导致该方形阵列输出数据元素的一些向量在其它输出数据元素向量之前完成,并且因此使用上文所描述的移动和零指令可实现该经完成的输出数据元素向量被移出该阵列存储装置,其中相关联的存储元件被释放以供与其它输出数据元素向量相关联地使用。

[0040] 该阵列输入数据元素可采取多种形式,但在一示例性具体实施中可表示阵列像素值。然而,本文中所描述的技术同样适用于可例如不表示图像数据的其它阵列数据元素。

[0041] 在一些示例性具体实施中,通过处理一行输入数据元素所产生的输出数据元素的该多个向量可称为多“行”的输出数据元素。然而,如前文所提及,应注意,可依该阵列存储装置内的任何期望定向存储在输出数据元素的给定方形2D阵列数据元素(在本文中,此类方形2D阵列还可称为方形子阵列)内的累加的该行的输出数据元素。例如,行可存储作为在方形子阵列内的水平向量或作为在方形子阵列内的垂直向量,并且因此在本文中,用语“行”不应视为表示在阵列存储装置内的数据元素的任何特定定向。

[0042] 在一示例性布置中,该给定二维阵列数据元素为方形二维阵列数据元素,形成该方形二维阵列数据元素的该多个向量包括经布置在第一阵列方向的第一多个向量和经布置在第二阵列方向的第二多个向量,该第二阵列方向正交于该第一阵列方向,并且该移动和零指令的每个实例被布置成识别全部在该第一阵列方向延伸或全部在该第二阵列方向延伸的一个或多个数据元素向量。因此,此提供如何识别待将各种数据元素向量移出该阵列存储装置的大幅灵活性。

[0043] 取决于具体实施,存储在该阵列存储装置内的一个或多个二维阵列数据元素可采取多种形式。在一个特定示例性具体实施中,该处理电路系统被布置成对该方形二维阵列数据元素执行处理运算,在其期间该处理电路系统经启用以在该第一阵列方向和该第二阵列方向两者存取数据元素向量。

[0044] 在一示例性具体实施中,该阵列存储装置可被配置成包括在第一阵列方向中延伸的多个阵列向量寄存器。该处理电路系统可被布置成执行一个或多个累加运算,其中每个累加运算被布置成产生用于在该阵列存储装置的群组的多个阵列向量寄存器内累加的输出数据。因此,在此类具体实施中,该阵列存储装置被视为包括在单一方向内延伸的多个可分开寻址的阵列向量寄存器,并且前文所提及的给定二维阵列数据元素可被视为包括存储在上述所提及的该群组的多个阵列向量寄存器内的数据元素。

[0045] 在此类示例性具体实施中,该移动和零指令可经执行,当该一个或多个累加运算的执行已导致最终结果数据存在于由该移动和零指令所识别的该一个或多个所识别向量中,就引起该处理电路系统将该一个或多个所识别向量的该数据元素从该阵列存储装置移动至该目的地存储装置,并且将用以存储该一个或多个所识别向量的该数据元素的该群组的多个阵列向量寄存器内的每个阵列向量寄存器设定为逻辑零值。

[0046] 在一个特定示例性具体实施中,当该处理电路系统已完成该一个或多个累加运算的执行时,最终结果数据存在于该群组的多个阵列向量寄存器中的每个阵列向量寄存器中。接着,该移动和零指令的执行可引起将该最终结果数据从该群组的多个阵列向量寄存器移动至该目的地存储装置,并且将该群组的多个阵列向量寄存器的每个阵列向量寄存器清除为零。此接着使该处理电路系统能够再使用来自该群组的阵列向量寄存器的一个或多个阵列向量寄存器以用于任何期望后续处理运算(因此例如使用任何或所有该阵列向量寄存器来执行后续累加指令将由于该阵列向量寄存器的内容被清除为零而引起非累加变体被执行)。

[0047] 指定用于该移动和零指令的目的地存储装置可采取多种形式。在一示例性具体实施中,该设备可进一步提供包括多个向量寄存器的向量寄存器文件,并且该移动和零指令可被布置成指示在该向量寄存器文件内的一个或多个向量寄存器作为该目的地存储装置。存在其中该移动和零指令可被布置成识别该一个或多个向量寄存器的各种方式。例如,在其中数据元素的单一向量被移出该阵列存储装置的单一向量寄存器情况下,该移动和零指令可提供用以确定该单个向量寄存器的识别符。对于在其中数据元素的多个向量被移出该阵列存储装置的多个向量寄存器情况下,可使用用于该向量寄存器中的每一者的分开的识别符信息明确识别该多个向量寄存器,或替代地,可通过该指令来识别一个向量寄存器,其中该多个向量寄存器中的其它向量寄存器是隐含的。例如,该多个向量寄存器可为以已明确识别的该向量寄存器开始的一序列的相邻向量寄存器,或该多个向量寄存器可各自通过恒定跨步值(constant stride value)分开。

[0048] 然而,在替代具体实施中,如果需要,该移动和零指令可被布置成指示待存储该一个或多个所识别向量的该数据元素的存储器中的一个或多个位置作为该目的地存储装置。在此情况下,该移动和零指令可替代地称为存储和零指令。

[0049] 存在其中该移动和零指令可被布置成识别该一个或多个数据元素向量待移动至其的存储器位置的各种方式。例如,考虑到其中仅单一数据元素向量被移动至存储器的单

一向量情况,该移动和零指令可被布置成识别在存储器中的位置,其中该数据元素向量接着被写入至由该位置所识别的连续存储器地址(在此情况下,该位置可例如为该第一数据元素的该存储器地址)。如果多个向量待被移动,则可通过该指令来识别存储器中的多个离散位置,其中该数据元素向量中的每一者被移动至由该指定位置中的一者所识别的一序列的存储器地址。替代地,一个位置可由该指令所指定,并且其它位置可为隐含的(例如,识别在相距于该所识别位置的固定跨步/偏移处的存储器地址的该位置)。

[0050] 取决于所存取的该阵列方向和保持在该阵列内的所存取向量的该数据元素的本质,还可能需要将单一向量的个别元素存储至离散存储器位置。然而,在单一向量内的个别数据元素需要被存储至离散存储器位置的情况下,一般来说将是该数据元素向量在适当的时候被传输至存储器之前首先被移动至向量寄存器的情况。

[0051] 存在其中可通过该移动和零指令来识别待移动的该一个或多个数据元素向量的数种方法。在一示例性具体实施中,该移动和零指令可包括用以识别在该阵列存储装置内的该给定二维阵列数据元素的该一个或多个数据元素向量的向量识别字段。例如,当单一数据元素向量被移动时,可提供足以识别该单一向量的识别符。当多个数据元素向量待移动时,接着该向量识别字段可用以提供足够的信息以明确识别该多个向量中的每一者,或替代地,可明确识别一个数据元素向量,而其它向量接着是隐含的,例如相邻向量或有规则间隔的向量(常称为跨步存取(stride access))。在此后一情况下,数目可由该向量识别字段提供以识别待移动的向量数目。

[0052] 在可在任一阵列方向进行存取的前文提及的方法中,接着该向量识别字段还可用以提供足够信息以识别正在进行存取的该阵列方向。例如,在一个具体实施中,该向量识别字段可包括:第一子字段,该第一子字段用以识别该方形二维阵列;和第二子字段,该第二子字段提供用于识别该一个或多个向量的一个或多个行识别符和阵列方向指示。

[0053] 在一示例性具体实施中,该移动和零指令可包括用以识别预测信息的预测字段,该预测信息用以识别该一个或多个所识别向量的哪些数据元素待从该阵列存储装置移动至该目的地存储装置且将它们的相关联的存储元件设定为该逻辑零值。此可通过使功能性能够限制于特定向量内的特定数据元素而提供额外灵活性。

[0054] 在使用预测的一些具体实施中,该数据元素大小可能变化,并且在此情况下,该移动和零指令可包括大小字段以识别在该一个或多个所识别向量内的每个数据元素的大小。通过使指令能够提供此额外信息,可能允许该指令被用于在该系统中所处理的各种不同数据元素大小,同时实现对该一个或多个所识别向量内的总数据元素的子集执行该移动和零操作。

[0055] 根据本文所描述的另一技术,提供一种额外新形式的指令,该指令还可用将在该阵列存储装置内的数据元素向量归零,并且在执行使用此类阵列存储装置执行累加运算时提供性能改善。根据此技术,提供一种设备,该设备具有:处理电路系统,该处理电路系统用于执行操作;指令解码器电路系统,该指令解码器电路系统用于解码指令以控制该处理电路系统,从而执行由该指令指定的该操作;和阵列存储装置,该阵列存储装置包括用以存储数据元素的存储元件。如同先前所描述的技术,该阵列存储装置被布置成存储至少一个二维阵列数据元素,当执行该操作时该处理电路系统可存取该数据元素,每个二维阵列数据元素包括多个数据元素向量,其中每个向量是一维的。根据此额外技术,该指令解码器电路

系统被布置成响应于解码识别在该阵列存储装置内的给定二维阵列数据元素的多个数据元素向量的零向量指令,还解码被布置成对该所识别多个数据元素向量进行操作的后续累加指令,并且控制该处理电路系统以执行由该累加指令指定的累加运算的非累加变体,以产生用于存储在该阵列存储装置内的该所识别多个向量中的结果数据元素。

[0056] 通过使用上述方法,该零向量指令在由该指令解码器电路系统解码时可与后续累加指令融合/合并以实际上建立该累加指令的非累加变体,该后续累加指令指定与该零向量指令所指定相同的多个数据元素向量。已发现,可为期望的是因为指令编码空间一般相当受限制,此类方法可为高度有益的,并且同时提供数个不同的累加指令以供对该阵列存储装置内的多个向量执行累加运算,可为非期望的是由于指令编码空间量将通过提供该非累加变体而被耗用,寻求还提供该指令的非累加变体。通过使用本技术,不需要提供该指令的该非累加变体,并且替代地,它们可通过组合该零向量指令与后续累加指令以引起该处理电路系统执行该非累加变体来模拟。

[0057] 在一示例性具体实施中,该阵列存储装置可包括在第一阵列方向延伸的多个阵列向量寄存器,并且在该阵列存储装置内的该所识别多个向量是由该阵列存储装置的群组的多个阵列向量寄存器提供。在此类配置中,该给定二维阵列数据元素可包括存储在该群组的多个阵列向量寄存器内的该数据元素。进一步,该后续累加指令可指定处理运算,该处理运算包括待对该所识别多个数据元素向量(即,对与由该零向量指令所指定相同的该群组的多个阵列向量)执行的累加运算,并且该零向量指令可与该后续累加指令组合地使用以实现通过该处理电路系统执行该处理运算的非累加变体。

[0058] 在一示例性具体实施中,该零向量指令可包括用以识别在该阵列存储装置内的该给定二维阵列数据元素的该多个数据元素向量的向量识别字段。

[0059] 如同先前讨论的移动和零指令,如果需要,则该零向量指令可包括用以识别预测信息的预测字段,该预测信息用以识别该多个所识别向量内的哪些存储元件待设定为该逻辑零值。此类方法可实际上允许一些数据元素经受后续累加指令的非累加变体,而其它数据元素经受真实累加变体。再者,如果需要,该零向量指令可包括大小字段,该大小字段用以识别在该多个所识别向量内的每个数据元素的大小。

[0060] 现将参照图式讨论特定的示例性具体实施。

[0061] 图1示意地示出数据处理系统10,该数据处理系统包括耦合至存储器30的处理器20,该存储器存储数据值32和程序指令34。处理器20包括用于自存储器30提取程序指令34并将该经提取程序指令供应至指令解码器电路系统50的指令提取单元40。解码器电路系统50解码该经提取程序指令,并产生控制信号以控制处理电路系统60对寄存器存储装置65的存储元件内所保持的数据值执行处理运算,如由该经解码向量指令所指定。如图1所示,寄存器存储装置65可由多个不同块形成。例如,可提供纯量寄存器文件70,该纯量寄存器文件包括可通过指令指定的多个纯量寄存器,并且类似地,可提供向量寄存器文件80,该向量寄存器文件包括可通过指令指定的多个向量寄存器。

[0062] 如图1所示,处理器20可存取阵列存储装置90。在图1所示的示例中,阵列存储装置90经提供作为处理器20的部分,但此非必要。在各种示例中,该阵列存储装置可实施为以下中的任一者或多者:架构上可寻址寄存器;非架构上可寻址寄存器;高速暂存存储器;和快取存储器。

[0063] 在一示例性具体实施中,处理电路系统60可包括向量处理电路系统和纯量处理电路系统两者。纯量处理与向量处理之间的大致区别如下。向量处理可涉及施加单一向量处理指令至数据向量的数据元素,该数据向量在数据向量中的相应位置处具有多个数据元素。根据本技术,该处理电路系统还可执行向量处理以对存储在阵列存储装置90内的二维阵列数据元素(其还可称为子阵列)内的多个向量执行操作。纯量处理有效地对单一数据元素进行操作,但非对数据向量。向量处理可用于其中对待处理的数据的许多不同实例实行处理操作的实例中。在向量处理布置中,单一指令可同时施加于(数据向量的)多个数据元素。相较于纯量处理,此可改善数据处理的效率和产出量。

[0064] 处理器20可被布置成处理存储在阵列存储装置90中的二维阵列数据元素。在至少一些示例中,该二维阵列可作为一维向量数据元素在多个方向存取。在一示例性具体实施中,阵列存储装置90可被布置成存储一个或多个二维阵列数据元素,并且每个二维阵列数据元素可在存储器中形成较大或甚至较高维阵列数据元素的方形阵列部分。

[0065] 图2显示在一示例性具体实施中可提供的处理器20的架构寄存器65的示例。架构寄存器(如指令集架构(instruction set architecture,ISA)中所定义的)可包括一组纯量整数寄存器100,该组纯量整数寄存器作用为用于处理由处理电路系统60内的纯量处理电路系统执行的操作的通用寄存器。例如,可存在一定数目个通用寄存器100,例如在此示例中提供31个寄存器X0至X30(纯量寄存器字段的第32个编码可不对应于硬件中提供的寄存器,因为其依据默认可视为指示例如零的值,或可用以指示不是通用寄存器的专用类型的寄存器)。存取经映射至相同物理存储装置的不同大小的纯量寄存器是可行的。例如,寄存器标签X0至X30可为指64位寄存器,但相同寄存器还可作为32位寄存器存取(例如,使用在硬件中提供的每个64位寄存器的下部32位进行存取),在此情况下,寄存器标签W0至W30可用在组合程序代码(assembler code)中以参考相同寄存器。

[0066] 再者,可供由解码器50支持的ISA中的程序指令选择的架构寄存器可包括一定数目个向量寄存器105(在此示例中标记为Z0至Z31)。当然,提供图2中所示的数个纯量/向量寄存器不是必需的,并且其它示例可提供可由程序指令指定的不同数目个寄存器。每个向量寄存器可存储包括可变量目个数据元素的向量运算元,其中每个数据元素可表示独立数据值。响应于向量处理(SIMD)指令,处理电路系统可对存储在寄存器中的向量运算元执行向量处理以产生结果。例如,向量处理可包括逐通道运算,其中对应运算为对一个或多个运算元向量中的每个元素通道执行以产生结果向量的元素的对应结果。当执行向量或SIMD处理时,每个向量寄存器可具有一定的向量长度VL,其中向量长度是指在给定向量寄存器中的位数目。用于向量处理模式中的向量长度VL对于给定硬件具体实施可为固定的或可为可变的。由处理器20支持的ISA可支持可变向量长度,使得可选择不同处理器具体实施以实施不同大小的向量寄存器,但ISA与向量长度无关,使得指令经设计以使得程序代码可正确地作用而无论在执行该程序的给定CPU上实施的特定向量长度。

[0067] 向量寄存器Z0至Z31还可作用为用于存储向量运算元的运算元寄存器,它们提供输入至由处理电路系统60对存储在阵列存储装置90内的二维阵列数据元素执行的处理和累加运算。当向量寄存器用以提供输入至此类运算时,则向量寄存器具有向量长度MVL,该向量长度可与用于向量运算的向量长度VL相同,或可为不同的向量长度。

[0068] 如图2所示,架构寄存器还包括形成前文提及的阵列存储装置90(ZA0至ZA(NA-1))

的一定数目NA个阵列寄存器110。可将每个阵列寄存器视为用于存储单一2D阵列数据元素(例如,处理和累加运算的结果)的一组寄存器存储装置。然而,处理和累加运算可能不是唯一可使用阵列寄存器的运算。阵列寄存器还可用于存储方形阵列,同时执行存储器中的阵列结构的行/列方向的转置。当程序指令参照阵列寄存器110中的一者时,将其参照为使用阵列识别符ZAi的单一实体,但一些类型的指令(例如,数据传输指令)还可通过定义选择阵列的一部分的索引值而选择该阵列的子部分(例如,一个水平/垂直元素群组)。

[0069] 实际上,对应于该阵列寄存器的寄存器存储装置的物理具体实施可包括一定数目NR个阵列向量寄存器ZAR0至ZAR(NR-1),如图2所示。形成阵列寄存器存储装置110的阵列向量寄存器ZAR可为与用于SIMD处理和至阵列处理的向量输入的向量寄存器Z0至Z31不同的一组寄存器。阵列向量寄存器ZAR中的每一者可具有向量长度MVL,所以每个阵列向量寄存器ZAR可存储长度MVL的1D向量,该向量可逻辑地分割成可变量目的数据元素。例如,如果MVL为512个位,则此可为一组64个8位元素、32个16位元素、16个32位元素、8个64位元素、或4个128位元素。应理解,并非所有这些选项均需要在给定具体实施中支持。通过支持可变元件大小,此提供处置涉及不同精确度的数据结构的计算的灵活性。为了表示2D阵列数据,一群组的阵列向量寄存器ZAR0至ZAR(NR-1)可逻辑上视为指派阵列寄存器识别符ZA0至ZA(NA-1)的给定者的单一实体,使得2D阵列经形成有在对应于阵列的一维的单一向量寄存器内延伸的元件,以及在条串跨接(striped across)多个向量寄存器的阵列的另一维度中的元件。

[0070] 布置阵列寄存器ZA使得它们存储方形阵列数据可为有用的(虽然非必需的),其中在水平方向上的元件数目等于垂直方向的元件数目。此可有助于支持阵列的实时转置,其中通过对在水平方向或在垂直方向读取/写入阵列寄存器110提供支持,存储器中的阵列结构的行/列维度可在阵列寄存器110与存储器之间传输阵列结构时交换。通过对在水平方向或垂直方向上对2D阵列寄存器写入/读取数据提供支持,此可允许在一个方向(例如,逐行)上从存储器载入的数据在相对方向(例如,逐列)写回至存储器,可能会比使用若干个收集/分散载入/存储或置换运算(permute operation)以在存储器与向量寄存器之间传输数据更快。

[0071] 如上文所提及,给定2D阵列数据元素可在一些示例性具体实施中形成方形阵列,但此非必要。因此,如图3A中所示,在一示例性具体实施中,给定2D阵列数据元素115可形成非方形阵列。替代地,如图3B所示,给定2D阵列数据元素120可形成方形阵列。在图3A和图3B中的每一者中,个别方框表示数据元素,并且在一些具体实施中,数据元素大小可变化。在任一示例中,可以多种方式指定二维阵列数据元素,但在一示例性具体实施中,如图3A和图3B所示,可通过一序列的阵列向量寄存器(ZAR)来指定给定的2D阵列数据元素。

[0072] 如上文所讨论,处理电路系统60被布置成在由解码器电路系统50所解码的指令的控制下存取纯量寄存器70、向量寄存器80、和/或阵列存储装置90。现将参考图4A描述此后一布置的进一步细节,其仅提供如何可存取阵列存储装置的一个说明性示例,尤其考虑存取在阵列存储装置内的方形2D阵列。

[0073] 在所说明的示例中,阵列存储装置90内的方形2D阵列经布置为n x n个存储元件/位置200的阵列205,其中n为大于1的整数。在本示例中,n为16,其表示对于存储位置200的存取的粒度在水平或垂直阵列方向中的任一者中为总存储的第1/16。

[0074] 从处理电路系统的观点来说,n x n个位置的阵列为可存取为在第一方向(例如,如所示出的水平方向)上的n个线性(一维)向量以及在第二阵列方向(例如,如所示出的垂直方向)上的n个线性向量。因此,从处理电路系统60的观点来说,n x n个存储位置经布置或至少可存取为2n个线性向量,各具有n个数据元素。

[0075] 在与至少处理电路系统60和可选地与解码器电路系统50通讯的控制电路系统250的控制下,存储位置200的阵列可由存取电路系统210、220、列选择电路系统230和行选择电路系统240存取。

[0076] 参考图4B,在指定为"A1"(请注意,如下文讨论,在阵列存储装置90内可提供多于一个此类2D阵列,例如A0、A1、A2以此类推)的示例性方形2D阵列的情况下,第一方向(如所示出的水平或"H"方向)的n个线性向量可各有16个数据元素0...F(以十六进位标记法表示),并且可在此示例中提及为A1H0...A1H15。存储在图4B的阵列存储装置90A1的256个项目(16x 16个项目)中的相同基本数据可替代地在第二方向(如所示出的垂直或"V"方向)上参照为A1V0…A1V15。应注意,例如,数据元素260为参照为A1H0的项F,而非A1V15的项0。应注意,"H"和"V"的使用并不表示关于组成阵列存储装置90的数据元素的存储的任何空间或物理布局需求,也不具有阵列存储装置内的2D阵列在任何示例性应用中是否存储行数据或列数据的任何相关性。

[0077] 如先前所讨论,使用阵列存储装置90可显著改善对某些类型的运算(例如累加运算)的性能,其中可对在阵列存储装置90内的给定二维阵列数据元素执行此类累加运算的多次迭代,其中该二维阵列数据元素用以在执行该累加运算时累加结果。然而,当该累加运算完成,则用于将所得的该数据元素向量移出该阵列存储装置以及准备在阵列存储装置内的相关联的该存储元件使得它们可用于后续累加运算的有效机制将为期望的。

[0078] 如先前所讨论,在一示例性具体实施中,这是经由使用识别在阵列存储装置90内的给定二维阵列数据元素的一个或多个数据元素向量的移动和零指令来实现。当此移动和零指令经解码时,接着处理电路系统60经控制以将该一个或多个所识别向量的该数据元素从该阵列存储装置移动至目的地存储装置(其可为例如在向量寄存器文件80内的一个或多个向量寄存器),并且还将用于存储该一个或多个所识别向量的该数据元素的该阵列存储装置的该存储元件设定为逻辑零值。

[0079] 图5A示意性地示出根据一个示例性具体实施的可在该移动和零指令内提供的字段。操作码字段305用于将指令识别为移动和零指令。在一些示例性具体实施中,可能存在经提供的移动和零指令的不同变体,并且因此可能存在可识别移动和零指令的多于一个不同的操作码。举一个特定示例,当该移动和零指令用于将该阵列存储装置内的该所识别向量移动至向量寄存器文件80内的目标向量寄存器时,可存在针对该移动和零指令定义的一个变体,并且当该阵列存储装置内的一个或多个向量移动至存储器时,可提供不同的变量(在此后者情况下,该指令可例如称为存储和零指令)。

[0080] 还提供用以识别阵列存储装置内的待经受移动操作的一个或多个向量的向量识别字段310。在一些实例中,可识别仅单一向量,但在其它实例中,可通过此字段识别多个向量。在后者情况下,在一示例性具体实施中,可独立地识别所有该多个向量,但在另一示例性具体实施中,可例如从第一向量的指示和待移动的向量数目的指示来推断该多个向量。

[0081] 如图5A所示,在移动和零指令内还提供用以识别(该)向量应移动至其的目的地存

储装置的目的地存储装置识别字段315。在一个示例性具体实施中,此字段用以识别在向量寄存器文件80中的一个或多个向量寄存器,并且在其中识别多个此类向量寄存器的示例中,它们可以与通过该向量识别字段识别多个向量的类似方式来识别(例如,可识别第一向量寄存器,接着其它向量寄存器是基于形成目的地存储装置所需的向量寄存器数目的知识而为隐含的)。在一替代具体实施中,在该阵列存储装置内待移动至存储器的一个或多个向量,接着在目的地存储装置识别字段315中所提供的信息可被布置成识别该一个或多个向量的数据元素待存储至其的存储器中的位置。此可例如涉及识别其内容用以识别存储器中所需位置的一个或多个寄存器。

[0082] 如果需要,可在指令300内提供一个或多个可选的额外字段320。例如,预测字段可用以识别预测信息,该预测信息用以控制该一个或多个所识别向量的哪些数据元素待经受移动和零操作。此通过允许对某些数据元素施加操作、但对其它数据元素不施加操作而提供灵活性。举另一示例,数据元素大小指示可提供在指令内,由此允许对其数据元素大小不固定的向量施加指令。

[0083] 在一示例性具体实施中,该移动和零指令可被布置成对在第一阵列方向延伸的阵列向量寄存器进行操作,并且在该具体实施中,不需要使在水平和垂直方向皆能够在向量识别字段310内编码。然而,在阵列存储装置90内的2D阵列可在水平或垂直方向上存取的示例中(在一个此类具体实施中,2D阵列为方形阵列),则向量识别字段可采取图5B中所示的形式。具体来说,此向量识别字段310’可包括用以识别阵列存储装置90内待存取的给定方形2D阵列的第一子字段312,以及由两个部分313和314所形成的第二子字段。第一部分313提供用以识别方形2D阵列内的一行或多行数据元素的一个或多个行识别符,并且第二部分314提供阵列方向指示,因此实现确定由该行识别符所识别的数据元素行是否在水平方向或垂直方向延伸。应了解,第一部分313与第二部分314的组合实现识别方形2D阵列内的一个或多个向量。

[0084] 图6为示出移动和零指令的操作的流程图。在步骤350,当确定解码器电路系统50已遇到移动和零指令时,接着在步骤355,从在该移动和零指令的向量识别字段310中所提供的信息来识别在该阵列存储装置内的一个或多个数据元素向量。进一步,在步骤360,从该移动和零指令的目的地存储装置识别字段315识别待使用的目的地存储装置。如前文所提及,此步骤一般可引起向量寄存器文件80中的一个或多个向量寄存器被识别为移出阵列存储装置的数据元素的目的地,但替代地,在一些具体实施中,这可为所识别的目的地存储装置采取在存储器中的一个或多个位置的形式的情况。在一示例性具体实施中,步骤355和360可由解码器电路系统50执行,但在替代具体实施中,处理电路系统60可基于由解码器电路系统50所提供的信息来执行该确定步骤。

[0085] 在步骤365,处理电路系统60用以将每个所识别的数据元素向量移动至目的地存储装置,并且接着将该阵列存储装置的相关存储元件(即,用以存储现在已移动至目的地存储装置的数据元素的该存储元件)设定为零。

[0086] 图7示出可在一示例性具体实施中的可对阵列存储装置90进行操作的示例性指令序列。如图7的示例中所示,可执行一系列数据处理指令(在此示例中,三个指令),以执行在阵列存储装置内提供的给定2D阵列内的处理和累加运算。在这些多个指令的执行期间,该结果在给定2D阵列内累加。

[0087] 在此示例中,假设当第三数据处理指令已完成,接着在给定2D阵列内的第一垂直向量就存储最终累加结果,而在此阶段,给定2D阵列中的其它垂直向量仅存储中间累加结果。鉴于第一垂直向量存储最终累加结果,将该结果移出阵列存储装置以释放该第一垂直向量的该存储元件以供在后续处理和累加运算中使用将是有用的。

[0088] 如图7所示,此通过执行移动和零指令、识别垂直向量1、和定义该垂直向量的内容应移动至其的目的地向量寄存器而实现,在此示例中,该向量寄存器称为寄存器Zi。执行此指令引起在垂直向量1中的该最终累加结果被移动至所识别向量寄存器中,并且在给定2D阵列中的相关联存储元件(即,实施垂直向量1者)被清除为逻辑零值。由于执行此单一指令,不仅该最终累加结果已被移出该阵列存储装置,而基本存储元件也已经准备使得它们立即可供再使用在后续处理和累加运算中。具体来说,通过将它们的内容清除为0,它们可立即开始被指定为由后续指令产生的新累加结果的目的地。

[0089] 因此,如图7所示,当执行后续数据处理指令4时,此可累加至2D阵列中,并且如果需要可再使用垂直向量1。当已执行数据处理指令,则假设垂直向量2现在保持最终累加结果,并且据此可执行额外的移动和零指令,以将给定2D阵列内的垂直向量2的内容移出阵列存储装置而至目的地向量寄存器(在此示例中,寄存器Zi+x)中。再次,此指令的执行引起所识别向量的内容被移出阵列存储装置,并且对应的存储元件被清除为逻辑零值,因此释放该存储元件以供在后续处理和累加运算中使用。因此,如图7所示,接着可执行数据处理指令的后续迭代,并且如果需要,可再使用垂直列2。

[0090] 存在可使用阵列存储装置90内的给定2D阵列执行的各种类型的运算类型以累加结果,其中并非在给定2D阵列内的所有向量均将必然同时保持最终累加结果。在此类情况下,采用在图7中举例示意性地示出的方法以释放2D阵列中的资源以供再使用可为有用的。此类方法的一示例性使用情况是在使用滑动窗方法执行2D有限脉冲响应(FIR)滤波时。图8中示意性地显示此类方法,其中考虑输入图像400。具体来说,FIR滤波操作被施加于该输入图像以产生对应的输出图像430,其中输出图像430中的每个像素415、420是由于对应的滤波操作407、412而产生。对于该滤波操作中的每一者,考虑多个输入像素,并且对该多个输入像素施加滤波系数,以产生输出像素的值。

[0091] 在图8所示的示例中,假设每个输出像素是通过考虑3x3阵列的输入像素而产生。因此,第一3x3阵列的输入像素405被提供至滤波操作407,其中使用对应阵列的滤波系数执行滤波以产生像素415的输出值。类似地,第二3x3阵列的输入像素410经受使用一组对应的滤波系数的滤波操作412,以产生用于输出像素420的值。应了解,阵列410相对于阵列405被向右移位一个像素位置,并且随着重复上文所描述的过程,3x3个像素的滑动窗可经检索以作用为至每个滤波操作的输入。当到达行的结尾时,过程可返回至输入图像的左侧,但从图像中下方的一行开始,并且再次从图像的左至右进行。因此,应了解,在图8所示的布置中,有效地存在一个滑动窗首先沿“水平”方向移动,该滑动窗在输入图像上移动而检索3x3阵列的输入图像像素以在运算每个输出图像像素时使用。

[0092] 图9A至图9D示出如何使用阵列存储装置90内的方形2D阵列且通过对在2D阵列中保持的数据执行外积累加运算来有效地执行此类2D图像滤波操作。注意,在图9A至图9D中所示出的示例中,滑动窗将首先在“垂直”方向上移动,即,在正交于图8示例中所显示的方向。因此,如图9A至图9D中所示出的输入图像440可视为在相对于其正常检视方向的侧面

上。

[0093] 如图9A至图9D所示,一次处理输入图像440的一列(或图像的一部分),并且经受使用滤波系数向量进行滤波操作。滤波系数向量内的各块(参见例如图9A中的块465)表示来自3x3阵列系数的三个滤波系数。填补元素(参见例如图9A中的元素467)对应于零或未经定义的值,并且仅为所示出的具体实施的人工产物。具体来说,在所示出的具体实施中,所使用的指令可执行每结果四个乘法和累加,但在所显示的示例性具体实施中仅需要执行三个乘法和累加。

[0094] 在示例中,示出四组系数(参见例如图9A中的四个块468,该四组系数取自四组3x3个系数中的一行)用作为输入,以对于所提供的输入向量运算四个输出向量(在用以产生四个输出向量中的四个乘法运算中的每一者产生一组(3+1)个系数)。除此之外,图9A至图9D中所示出的过程仅使用来自每个3x3阵列系数的一列的三个系数,并且因此,采用三个指令以在任何特定组的四个向量内产生最终累加结果。

[0095] 当该过程正在进行并处于稳定状态时,接着如稍后将参考图9C和图9D讨论,该过程同时对三组的四个输出向量(参见例如图9C中所示的三组输出向量475、485、495)进行操作。

[0096] 如图9A所示,当使用系数向量460处理第一列470的输入图像440,此引起累加结果被存储在四个向量475内。如图9B所示,当使用系数向量482处理第二列480的输入图像440,此将累加结果填入在四个向量475和四个向量485两者内。接着,如图9C所示,当使用系数向量492处理第三列490时,此引起累加结果被填入在四个向量475、四个向量485和四个向量495内。从图8的前文论述,应了解,在此点,在前三列的输入图像内的所有像素将已被处理,并且据此,四个水平向量475的内容将表示第一列群组的输出图像的最终累加结果。据此,如图9C所示,四个寄存器475的内容可经受前文描述的移动和零指令,以将该内容移动至目的地存储装置(例如,向量寄存器文件80内的四个向量寄存器),并且清除形成2D阵列450内的四个向量475的存储元件,使得它们可供后续累加运算使用。

[0097] 因此,举例来说,如图9D所示,当使用系数向量460处理第四列500的输入图像440时,此可引起累加结果被填入在四个向量485、四个向量495、和四个向量475内(现在可由于已通过先前的移动和零指令而将该向量内的存储元件清除为逻辑零值而可再使用)。也如图9D所示,四个向量485现在将存储最终累加结果,该最终累加结果表示第二列群组的输出图像(因为在此点,将已经处理第二输入列、第三输入列和第四输入列中的每一者)。据此,存储在四个水平向量485内的数据元素可被移出阵列而至向量寄存器文件的向量寄存器中,并且接着基本存储元件被清除以允许它们在后续累加运算中再使用。

[0098] 虽然在图9A至图9D中,经由移动和零指令的任何向量皆立即被再使用(因为此可导致更容易地程序化),但是并未要求经清除的向量立即被再使用,并且替代地处理可继续耗用在群组的向量475、485、495下方的向量,如果需要,则当到达阵列450的底部时仅返回至开头。

[0099] 在一示例性具体实施中,可在水平或垂直方向存取阵列存储装置90内的给定方形2D阵列。然而,在一些具体实施中,存在可使用在阵列存储装置90内的2D阵列执行的某些处理运算,其中仅在该方向的一者存取向量。因此,举例来说,返回参照先前讨论的图2,可存在具体识别在第一阵列方向延伸穿过阵列的阵列向量寄存器ZAR的一些处理指令。这些指

令可通过指定对其执行相关联的处理和累加运算的多个ZAR寄存器而允许高效执行某些处理和累加运算。然而,当已执行一系列该指令,则所有所识别的阵列向量寄存器ZAR一般将包括最终累加结果,但该寄存器用于后续处理和累加运算将是不可行的,直到该结果已被移出该阵列存储装置且形成该阵列向量寄存器的存储元件的目前内容已被清除为逻辑零值。

[0100] 图10示意性地示出可如何使用前文描述的移动和零指令以显著增加此类情况中的性能。具体来说,如图10所示,假设三个阵列向量寄存器ZAR2、ZAR3和ZAR4被初始化为0,并且接着执行上述类型的一系列数据处理指令以执行处理和累加运算,其中该累加结果维持在上述三个阵列向量寄存器内。当已完成所需的该系列数据处理指令(在此示例中,假设存在两个此类数据处理指令被执行),则所有上述三个阵列向量寄存器将存储最终累加结果。前文所提及的移动和零指令可因此用于指定该三个阵列向量寄存器作为其数据元素应移动至目的地存储装置的向量,并且还可识别待用作为该目的地存储装置的存储装置,在此示例中,假设使用向量寄存器文件80内的三个相邻向量寄存器。因此,执行移动和零指令将引起所有该累加结果被移出该阵列存储装置而至该向量寄存器文件的该所识别向量寄存器中,并且还引起形成该三个阵列向量寄存器的存储元件被清除为逻辑零值。因此,如图10所示,该过程可接着立即继续以执行还将累加至相同系列的阵列向量寄存器ZAR2、ZAR3和ZAR4中的一序列的后续数据处理指令。此提供高效率具体实施。

[0101] 根据本文中所描述的另一技术,提供额外的新形式的指令(在本文中称为零向量指令),该指令还可用以将在该阵列存储装置内的数据元素向量归零,并且在执行使用此类阵列存储装置时执行累加运算时提供性能改善(当与需要使用移动指令将零从一个或多个向量寄存器移动至阵列存储装置的期望向量、且需要保留一个或多个向量寄存器以保持该零值的具体实施相比较时)。根据此额外技术,指令解码器电路系统50被布置成响应于解码此零向量指令(其被布置成识别在该阵列存储装置内的给定二维阵列数据元素的多个数据元素向量),还解码被布置成对该所识别多个数据元素向量进行操作的后续累加指令。接着,引起该处理电路系统将用以存储该所识别多个向量的该数据元素的该阵列存储装置的该存储元件设定为逻辑零值,并且接着执行由该累加指令指定的累加运算以产生用于存储在该阵列存储装置内的该所识别多个向量中的结果数据元素。

[0102] 通过使用上述方法,该零向量指令在由该指令解码器电路系统解码时可与后续累加指令合并以实际上建立该累加指令的一非累加变体,该后续累加指令指定与该零向量指令所指定相同的多个数据元素向量。此可为高度有益的,因为指令编码空间通常是高价的(high premium),并且指定可经定义以对在该阵列存储装置内的多个数据元素向量进行操作的各种累加指令的非累加变体可能不是可行的。

[0103] 图11为示出根据一个示例性具体实施的此类零向量指令的处置的流程图。当在步骤520,解码器电路系统50遭遇到零向量指令时,接着在步骤525,参考零向量指令的向量识别字段来识别阵列存储装置内的多个数据元素向量。例如,可在向量识别字段中识别多个阵列向量寄存器ZAR。接着,在步骤530,解码器电路系统确定待解码的下一指令是否为对与由该零向量指令所识别的相同向量进行操作的累加指令。

[0104] 如果否,则在步骤535,处理电路系统经控制以将在步骤525所确定的用于存储该所识别多个向量的该数据元素的该阵列存储装置的存储元件设定为逻辑零值,并且其后处

理仅继续下一指令的执行。

[0105] 然而,在步骤530,如果确定下一指令为对与由该零向量指令所识别的相同向量进行操作的累加指令,则解码器有效融合该两个指令,并且在步骤540,控制该处理电路系统执行由该累加指令所指定的该累加运算的非累加变体(一般来说,此涉及处理运算和后续累加两者),以产生用于存储在所识别的多个向量的每一者中的结果。如前文所提及,通过此类方法,不需要具体编码被布置成对阵列存储装置内的多个向量进行操作的任何累加指令的非累加变体,因为此非累加变体可经由使用零向量指令、后续接着使用所需累加指令,通过上述融合过程而有效地实施(由此实施由该累加指令所定义的处理运算,但其中该累加函数有效地变成无效(nullified))。

[0106] 图12为示意性地示出在一示例性具体实施中的可在零向量指令内提供的字段的图式。具体来说,零向量指令550包括用以识别指令实际上为零向量指令的操作码字段555。进一步,提供用以识别阵列存储装置中的多个向量的向量识别字段560。此字段中的信息大致采用前文在描述图5A的移动和零指令300的向量识别字段310时所讨论的形式,虽然在一示例性具体实施中,后续累加指令为被布置成对在第一阵列方向延伸的阵列向量寄存器进行操作者,并且因此一般不需要使水平和垂直方向皆在向量识别字段560内编码。

[0107] 如果需要,则如方框565所示,可提供特定可选的额外字段,诸如如前文参考移动和零指令示例所论述的预测信息字段和数据元素大小字段。

[0108] 图13为示出根据另一示例性具体实施的零向量指令的处置的流程图。当在步骤570,解码器电路系统50遭遇到零向量指令时,接着在步骤575,参考零向量指令的向量识别字段来识别阵列存储装置内的多个数据元素向量。例如,可在向量识别字段中识别多个阵列向量寄存器ZAR。

[0109] 接着在步骤580,处理电路系统经控制以将在步骤575所确定的用于存储该所识别多个向量的该数据元素的该阵列存储装置的存储元件设定为逻辑零值,并且其后处理仅继续下一指令的执行。

[0110] 甚至在此具体实施中,当未发生融合以组合零向量指令与后续累加指令时,仍可实现显著的益处。具体来说,不需要执行多个移动指令,其中每个移动指令将零向量从该向量寄存器文件的向量寄存器移动至该阵列存储装置的所识别向量中。进一步,与必须实施(零)移动向量功能性相比,此类归零功能性以硬件建构更简单且更便宜。此外,因为不需要保留在向量寄存器文件中的一个或多个向量寄存器以保持逻辑零值(在基于所利用的移动指令的上述具体实施将需要),所以有额外节省。

[0111] 图14示出可使用的模拟器具体实施。虽然先前所述的示例以用于操作支持所关注技术的特定处理硬件的设备和方法来实施本发明,但还可能根据本文所述的示例提供指令执行环境,该指令执行环境是经由使用计算机程序实施。此类计算机程序常称为模拟器,因为它们提供硬件架构的基于软件的具体实施。模拟器计算机程序的种类包括仿真器、虚拟机、模型和二进制转译器(包括动态二进制转译器)。一般来说,模拟器具体实施可在可选地运行主机操作系统610、支持模拟器程序605的主机处理器615上运行。在一些布置中,在硬件与所提供的指令执行环境和/或相同的主机处理器上提供的多个相异指令执行环境之间可有多层模拟。历史上,已需要强大的处理器来提供模拟器具体实施,该处理器以合理速度执行,但此种方法在某些情况下可为有正当理由的,诸如当因为兼容性或再使用原因此需

要执行另一处理器原生的程序代码时。例如,模拟器具体实施可提供具有不为主机处理器硬件所支持的额外功能性的指令执行环境,或提供一般与不同的硬件架构相关联的指令执行环境。模拟的综述于“Some Efficient Architecture Simulation Techniques”(Robert Bedichek,Winter 1990USENIX Conference,第53至63页)中给出。

[0112] 在先前已参照特定硬件架构或特征描述实施的情况下,在模拟具体实施中,等效功能性可通过合适的软件架构或特征提供。例如,可在模拟具体实施中将特定电路系统提供作为计算机程序逻辑。类似地,存储器硬件(诸如寄存器或快取)可在模拟具体实施中提供作为软件数据结构。再者,可将用于存取硬件设备10中的存储器30的物理地址空间仿真为通过模拟器605映射至由主机操作系统610使用的虚拟地址空间的模拟地址空间。在先前描述的示例中提及的硬件元件的一者或多者存在于主机硬件(例如主机处理器615)上的布置中,一些模拟具体实施可(在适当处)利用主机硬件。

[0113] 模拟器程序605可存储在计算机可读存储介质(其可为非暂态介质)上,并提供虚拟硬件接口(指令执行环境)给目标码600(其可包括应用程序、操作系统、和超管理器),该虚拟硬件接口与通过模拟器程序605模型化的硬件架构的硬件接口相同。因此,目标码600的程序指令可在指令执行环境内使用模拟器程序605执行,使得实际上不具有上文所讨论的设备10的硬件特征的主机计算机615可仿真该特征。模拟器程序可包括:处理程序逻辑620,该处理程序逻辑仿真处理电路系统60的行为;指令解码程序逻辑625,该指令解码程序逻辑仿真指令解码器电路系统50的行为;和阵列存储装置仿真程序逻辑622,该阵列存储装置仿真程序逻辑维持数据结构以仿真阵列存储装置90。因此,本文所述的技术在图14的示例中可通过模拟器程序605以软件执行。

[0114] 在本申请案中,用语“被配置成……”是用以意指设备的元件具有能够实行该经定义操作的配置。在此上下文中,“配置”意指硬件或软件的互连的布置或方式。例如,该设备可具有专用硬件,该专用硬件提供经定义的操作,或者处理器或其它处理装置可经程序化以执行该功能。“被配置成”并不意味着设备元件需要以任何方式改变以提供所定义的操作。

[0115] 虽然本文中已参照附图详细描述说明性示例,但应明白,本发明不限于该精确示例,并且所属技术领域中具有通常知识者可实行各种变化、新增与修改于其中,而不脱离如随附权利要求书所定义的本发明的范围与实质。例如,可用独立权利要求的特征在不脱离本发明的范围的情况下作出与从属权利要求的特征的各种组合。

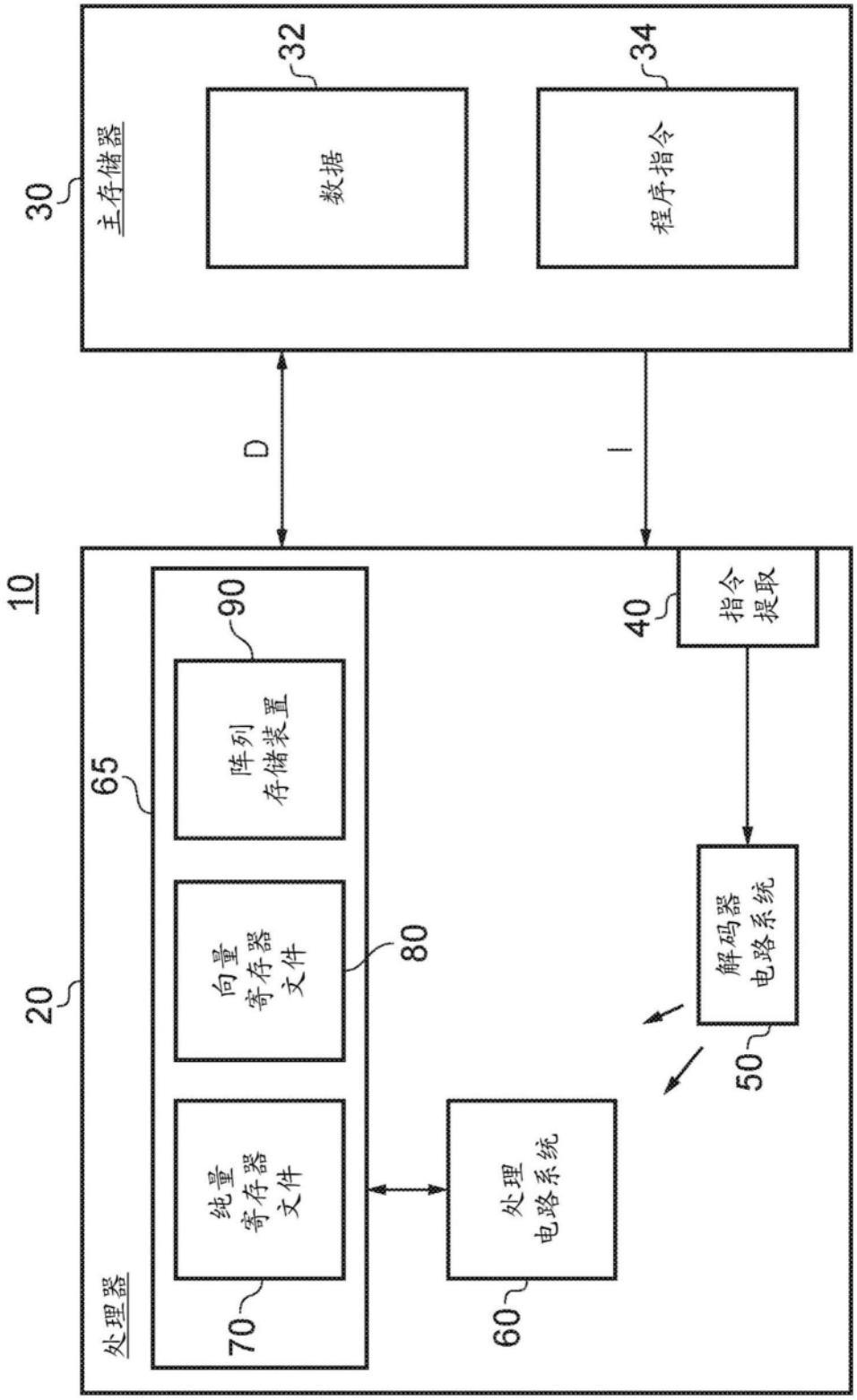

30
主存储器
32
数据
34
程序指令
D
I
10
20
处理器
65
90
阵列
存储装置
向量
寄存器
文件
80
纯量
寄存器
文件
70
处理
电路系统
60
40
指令
提取
解码器
电路系统
50

图1

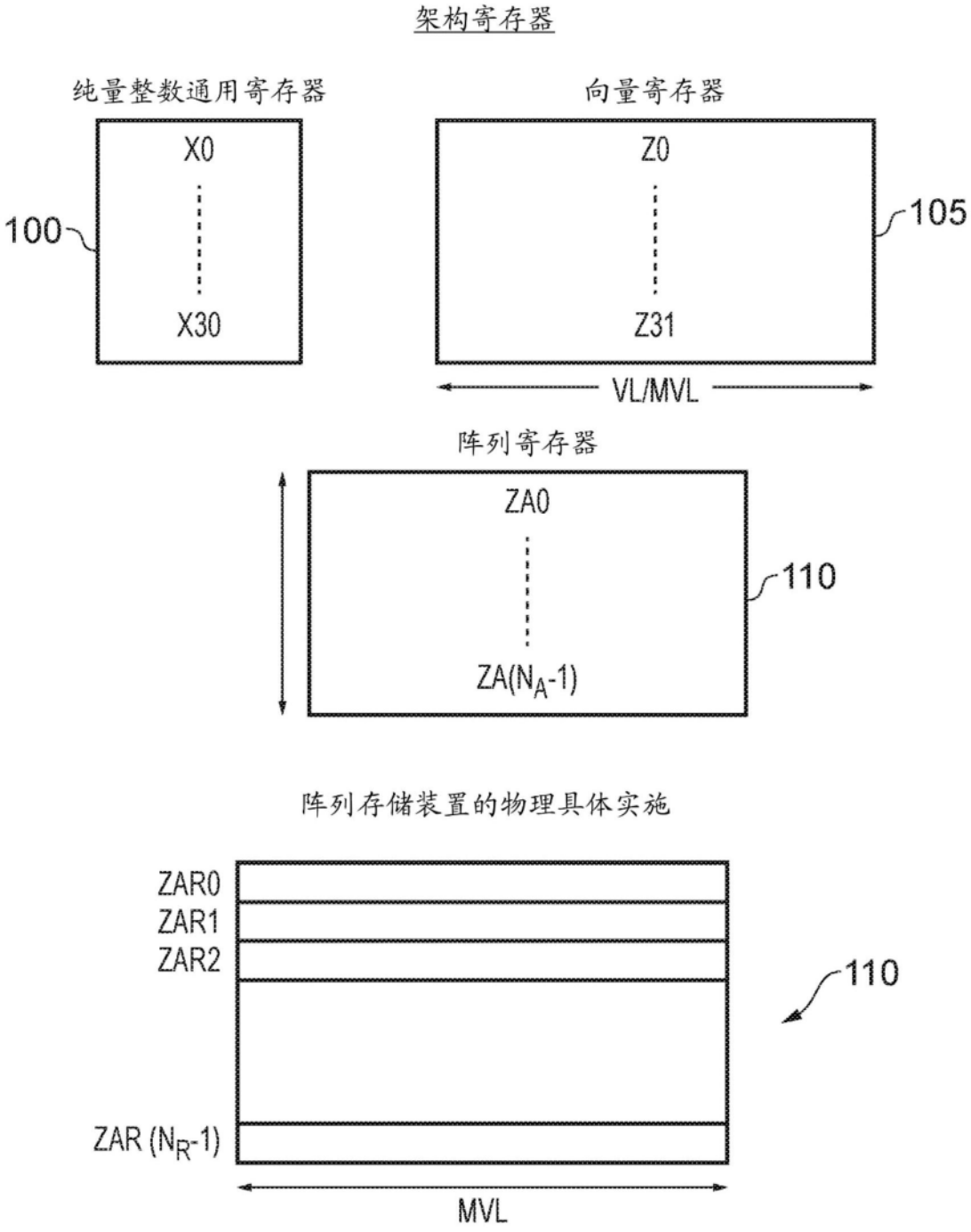

架构寄存器
纯量整数通用寄存器
向量寄存器
X0
X30
100
Z0
Z31
105
VL/MVL
阵列寄存器
ZA0
ZA(NA-1)
110
阵列存储装置的物理具体实施
ZAR0
ZAR1
ZAR2
ZAR (NR-1)
110
MVL

图2


非方形2D阵列数据元素
ZAR
ZAR
ZAR
ZAR
115

图3A

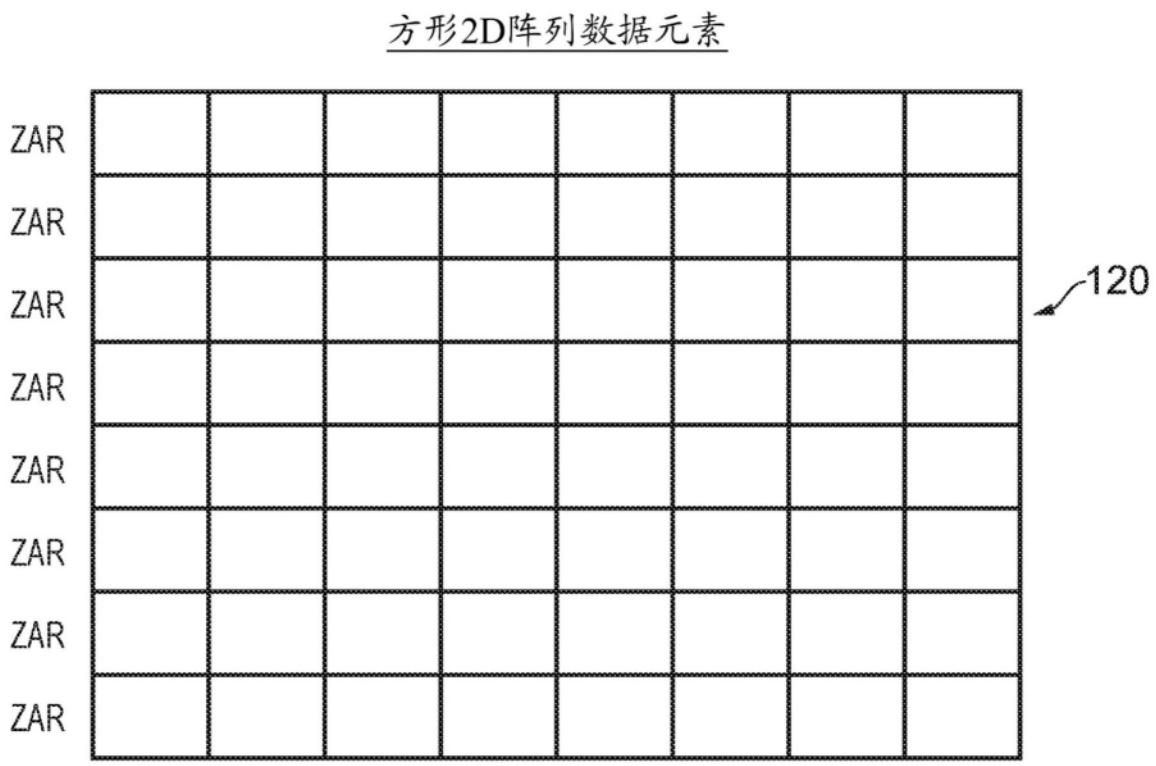

方形2D阵列数据元素
ZAR
ZAR
ZAR
ZAR
ZAR
ZAR
ZAR
ZAR
120

图3B

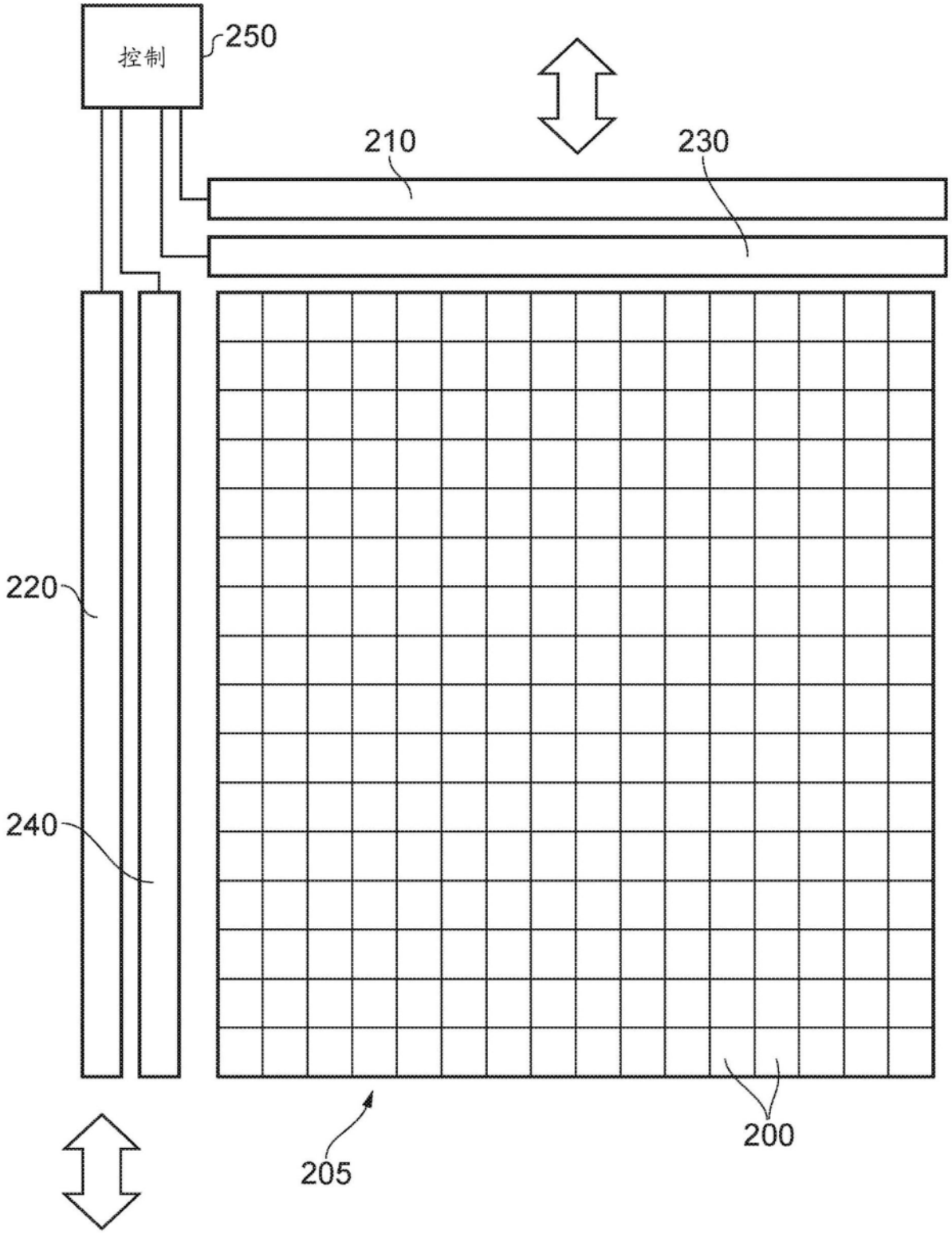



图4A

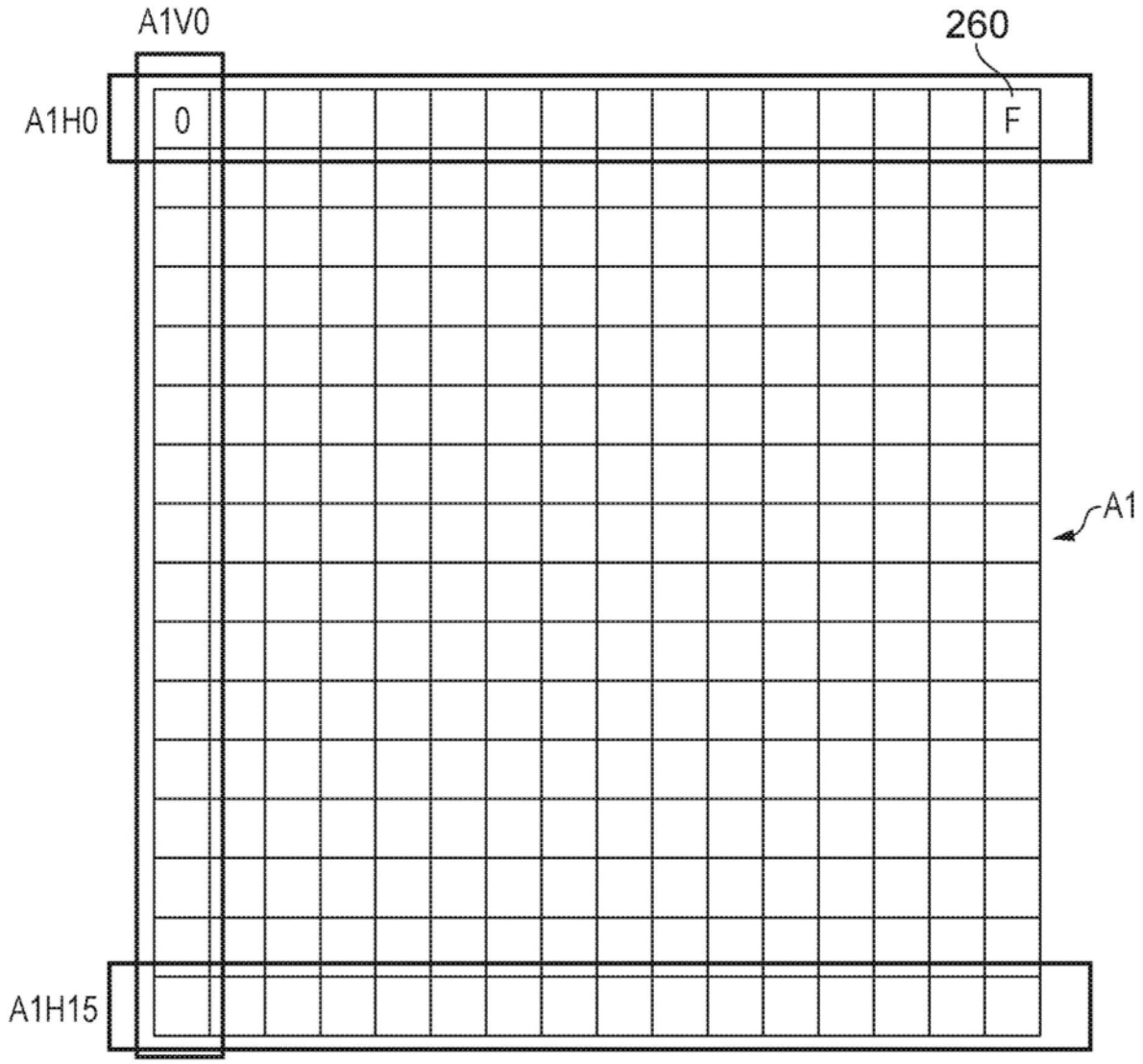



图4B

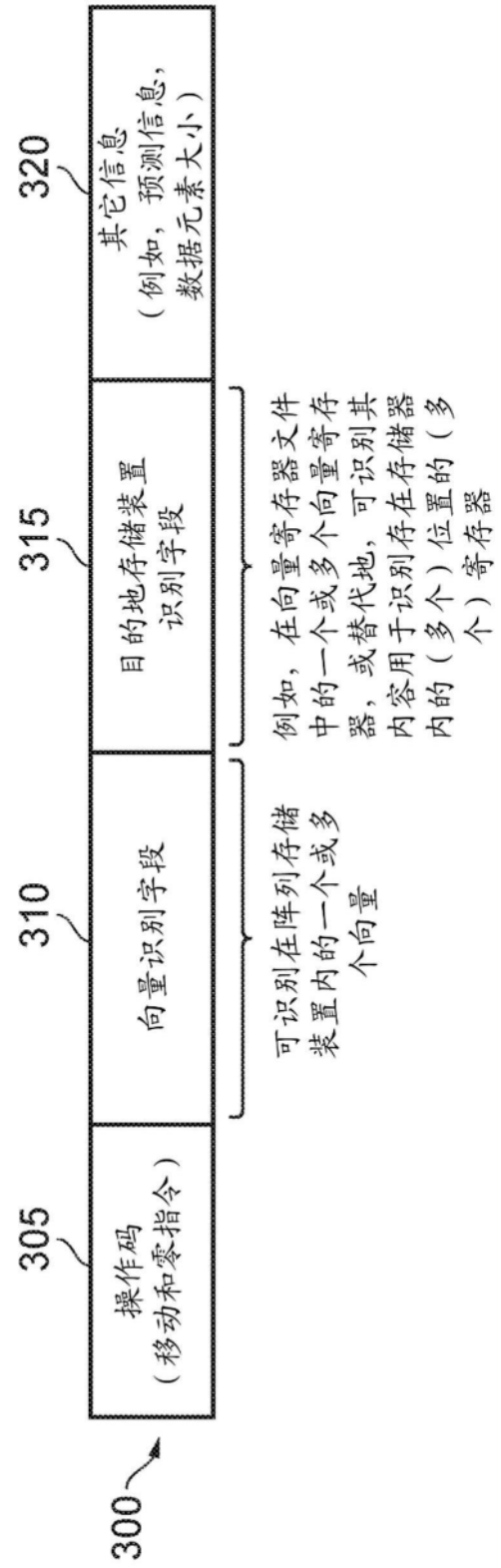

300
305
操作码
（移动和零指令）
310
向量识别字段
315
目的地存储装置
识别字段
320
其它信息
（例如，预测信息，
数据元素大小）
可识别在阵列存储
装置内的一个或多
个向量
例如，在向量寄存器文件
中的一个或多个向量寄存
器，或替代地，可识别其
内容用于识别存在存储器
内的（多个）位置的（多
个）寄存器

图5A

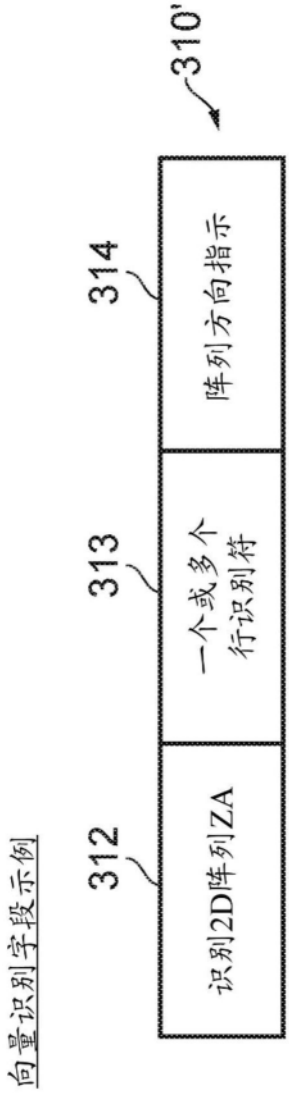



图5B

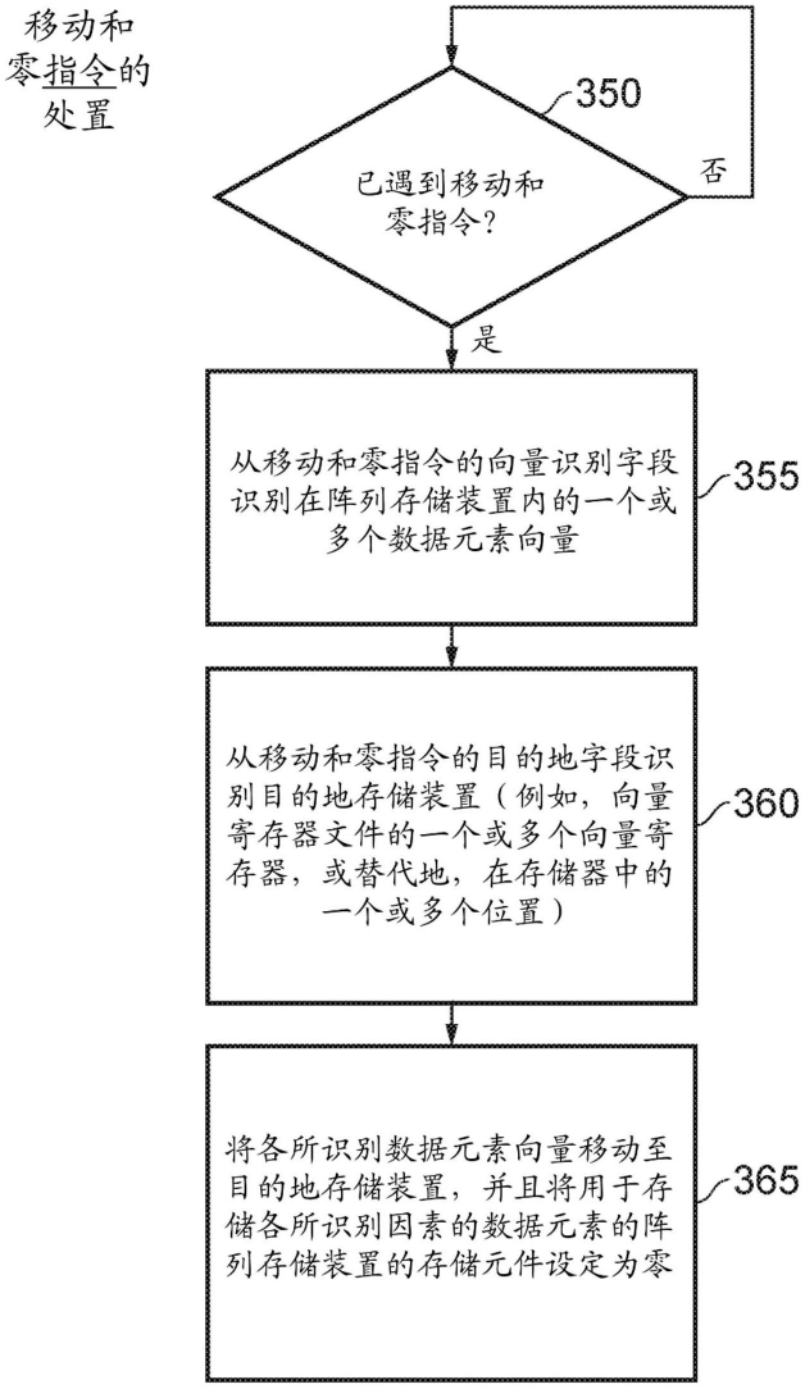

移动和零指令的处置
350
已遇到移动和零指令?
否
是
从移动和零指令的向量识别字段识别在阵列存储装置内的一个或多个数据元素向量
355
从移动和零指令的目的地字段识别目的地存储装置（例如，向量寄存器文件的一个或多个向量寄存器，或替代地，在存储器中的一个或多个位置）
360
将各所识别数据元素向量移动至目的地存储装置，并且将用于存储各所识别因素的数据元素的阵列存储装置的存储元件设定为零
365

图6

对阵列存储装置进行操作的示例性指令序列

1) 数据处理指令1（累加到2D阵列）

2) 数据处理指令2（累加到2D阵列）

3) 数据处理指令3（累加到2D阵列）

4) 移动和零指令（垂直向量1、向量寄存器$Z_i$）

- 垂直向量1保持最终累加结果，并且移动和零指令的执行将那些最终累加结果移动到$Z_i$并清除用于垂直向量1的存储元件

5) 数据处理指令4（累加到2D阵列-可再使用垂直列1）

6) 移动和零指令（垂直向量2、向量寄存器$Z_{i+x}$）

- 垂直向量2保持最终累加结果，并且移动和零指令的执行将那些最终累加结果移动到$Z_{i+x}$并清除用于垂直向量2的存储元件

7) 数据处理指令5（累加到2D阵列-可再使用垂直列2）

图7

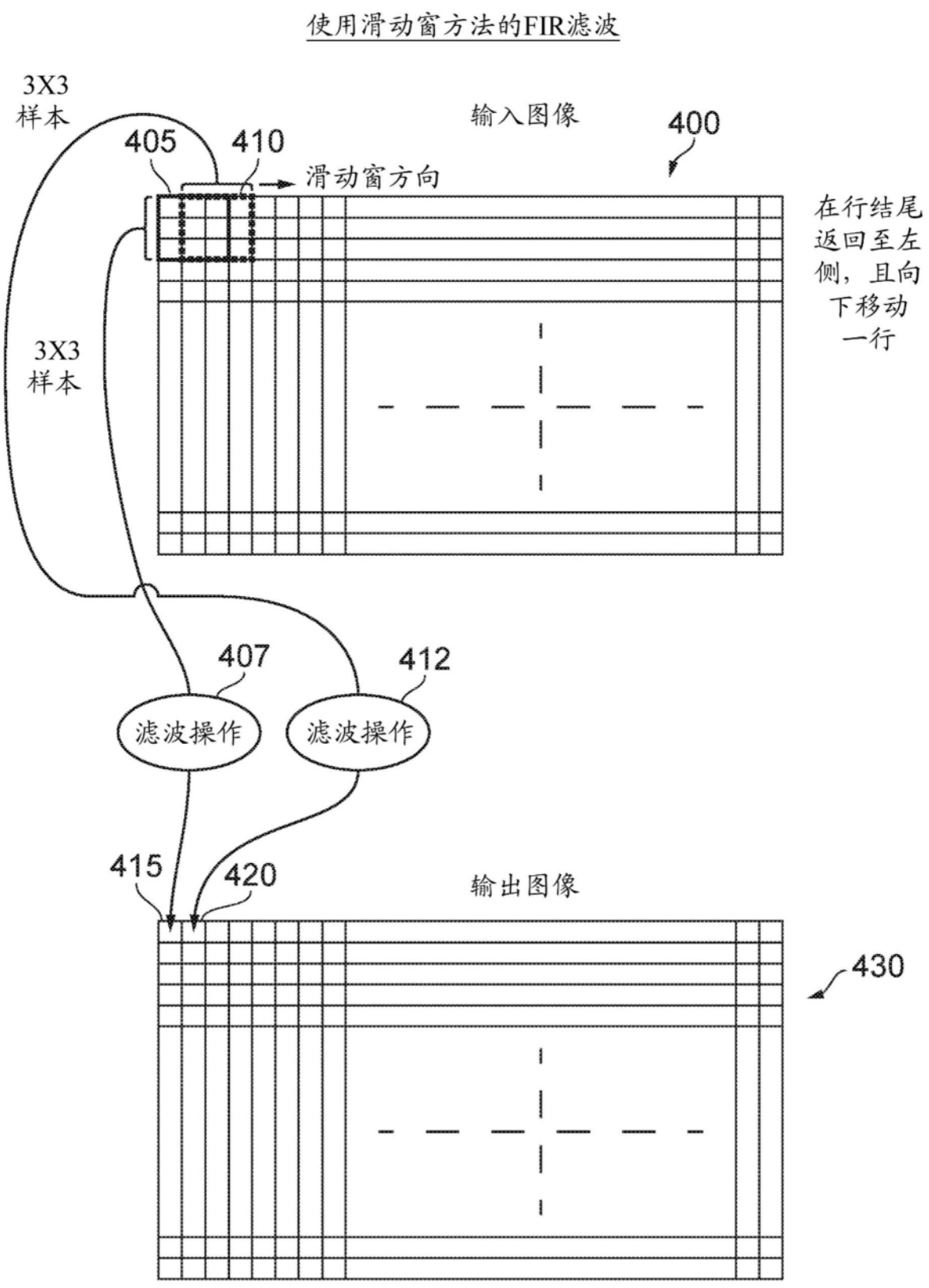

使用滑动窗方法的FIR滤波
3X3
样本
405
410
输入图像
400
滑动窗方向
在行结尾
返回至左
侧，且向
下移动
一行
3X3
样本
407
412
滤波操作
滤波操作
415
420
输出图像
430

图8

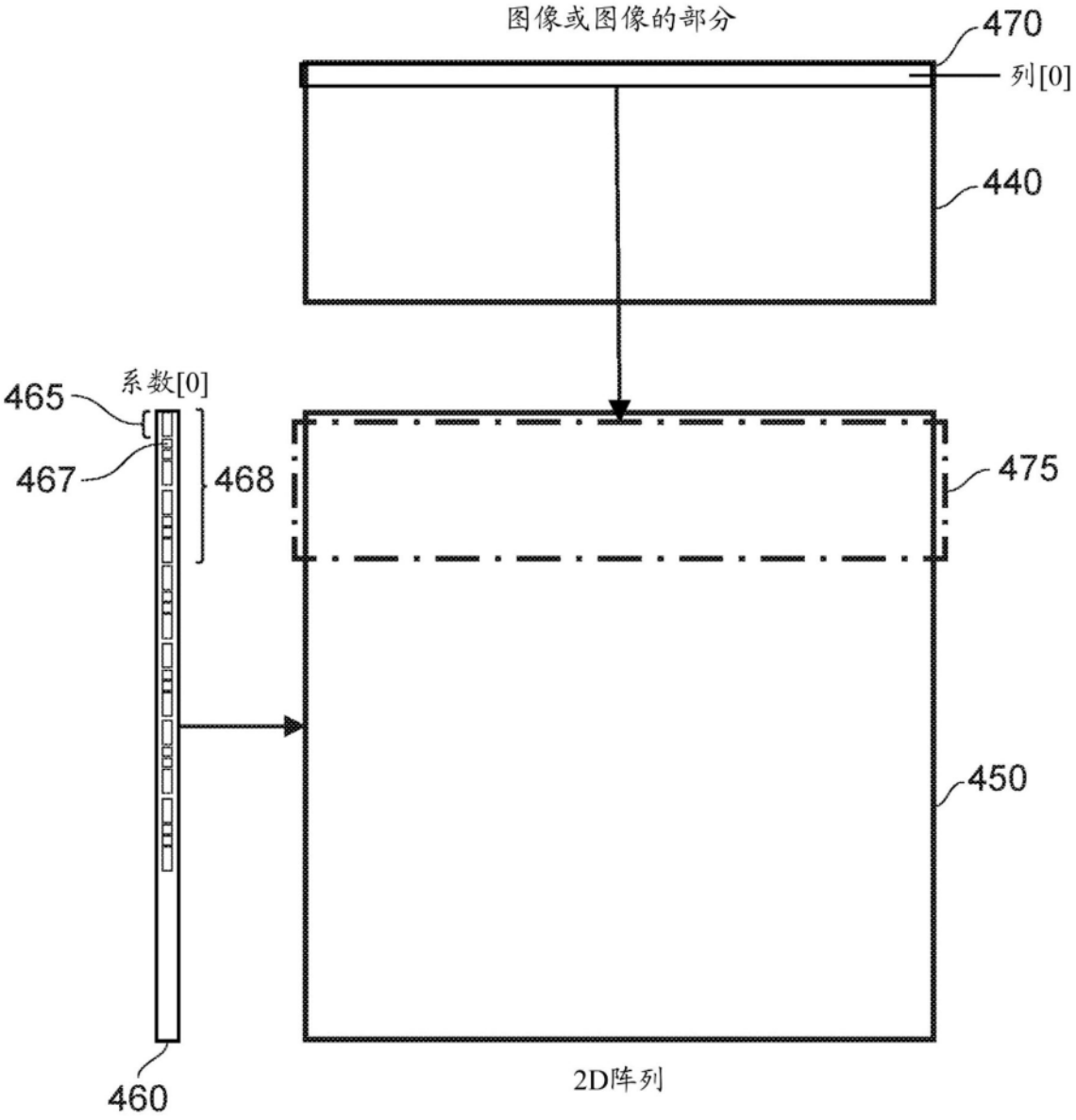



图9A

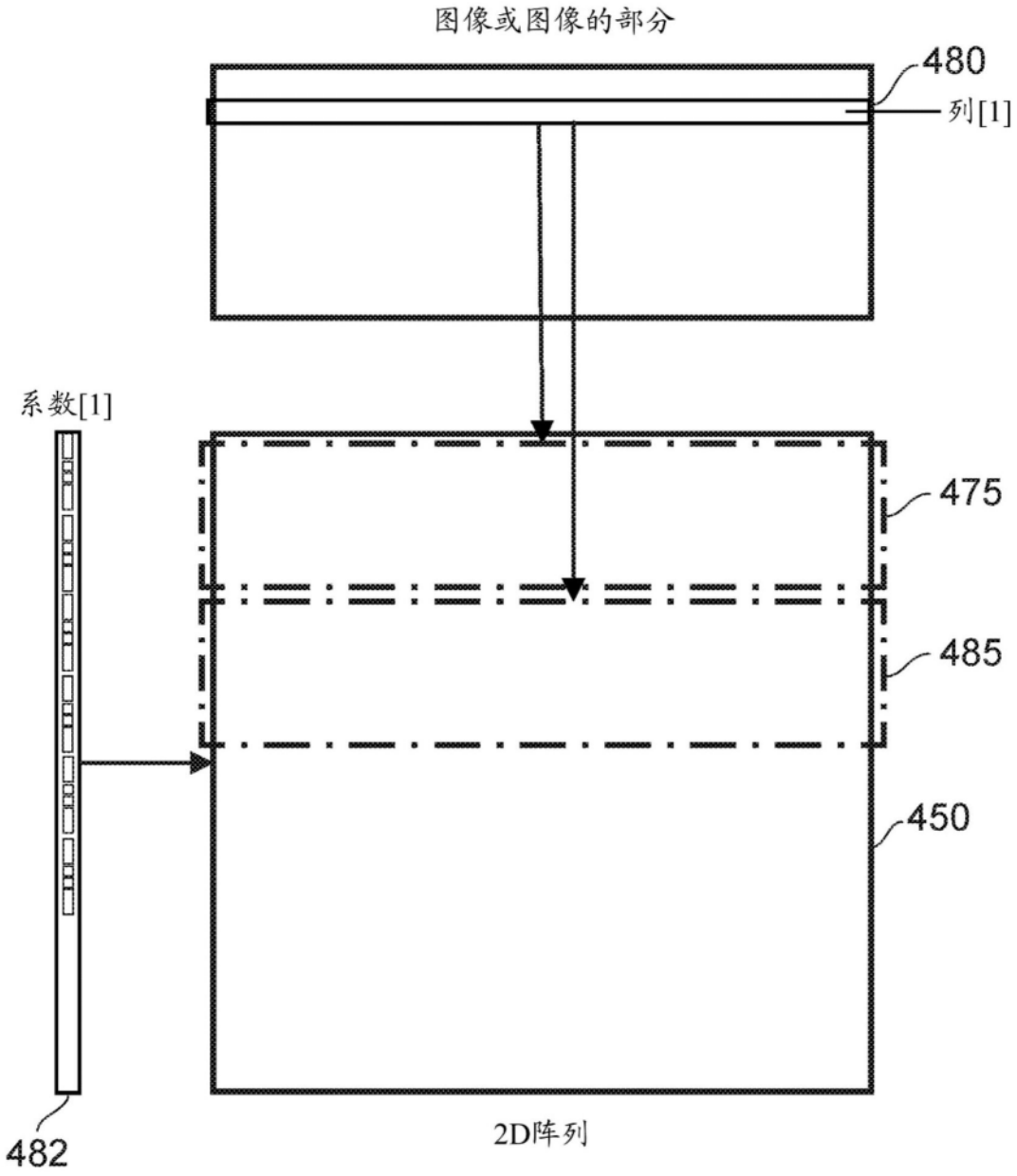



图9B

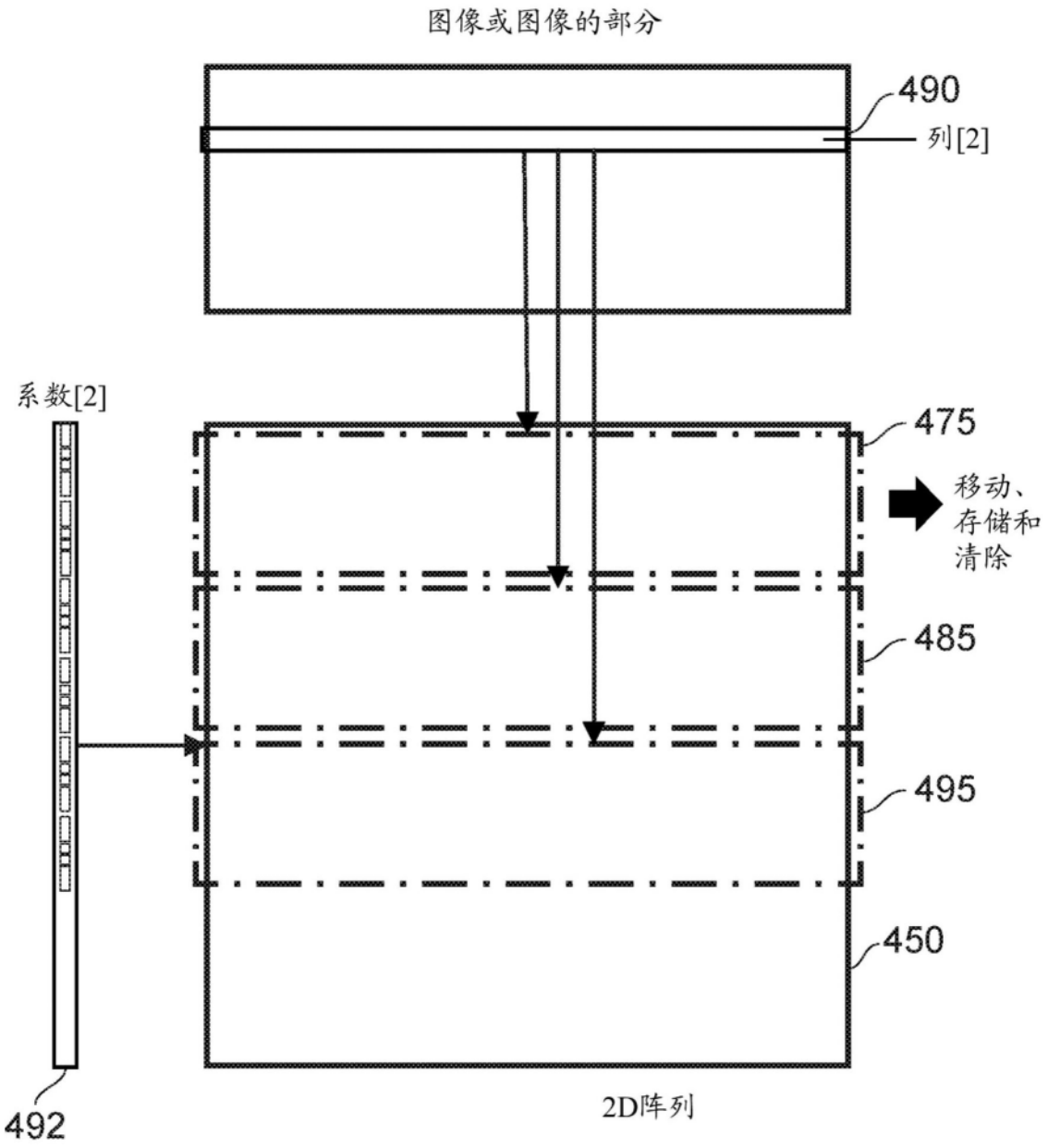



图9C

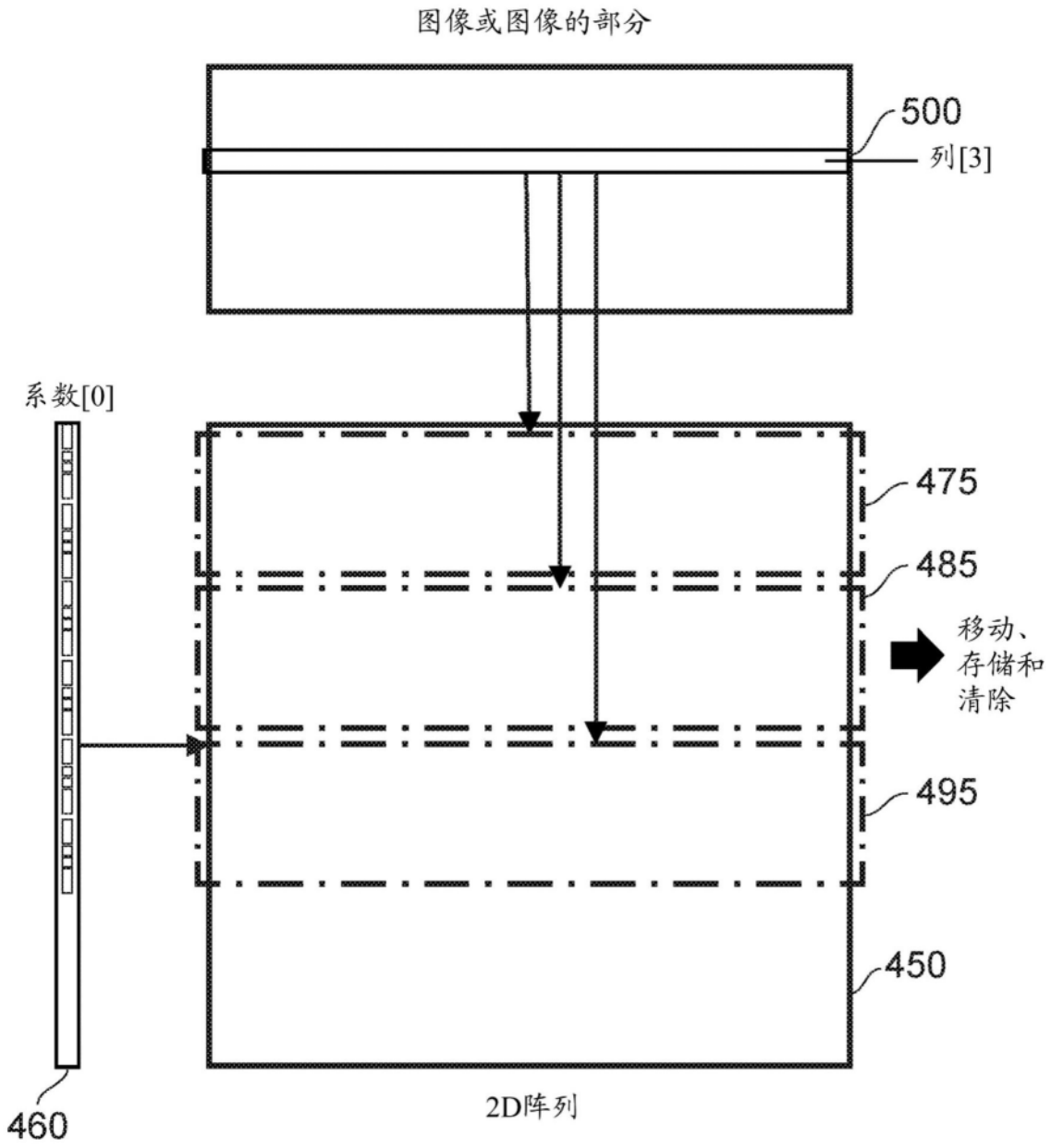



图9D

对阵列存储装置进行操作的替代示例性指令序列

假设ZAR 2、ZAR 3和ZAR 4被初始化为零

1) 数据处理指令（累加到ZAR 2、ZAR 3和ZAR 4）

2) 数据处理指令（累加到ZAR 2、ZAR 3和ZAR 4）

3) 移动和零指令（ZAR 2、ZAR 3、ZAR 4到$Z_i$、$Z_{i+1}$、$Z_{i+2}$）

- 所有累加结果移出而到向量寄存器文件中的向量寄存器，并且清除ZAR 2、ZAR 3和ZAR 4）

4) 数据处理指令（累加到ZAR 2、ZAR 3和ZAR 4）

执行有效引起所定义的运算的非累加变体被执行，其中结果经产生并存储于ZAR 2、ZAR 3和ZAR 4中的每一者中）

5) 数据处理指令（累加到ZAR 2、ZAR 3和ZAR 4）

图10

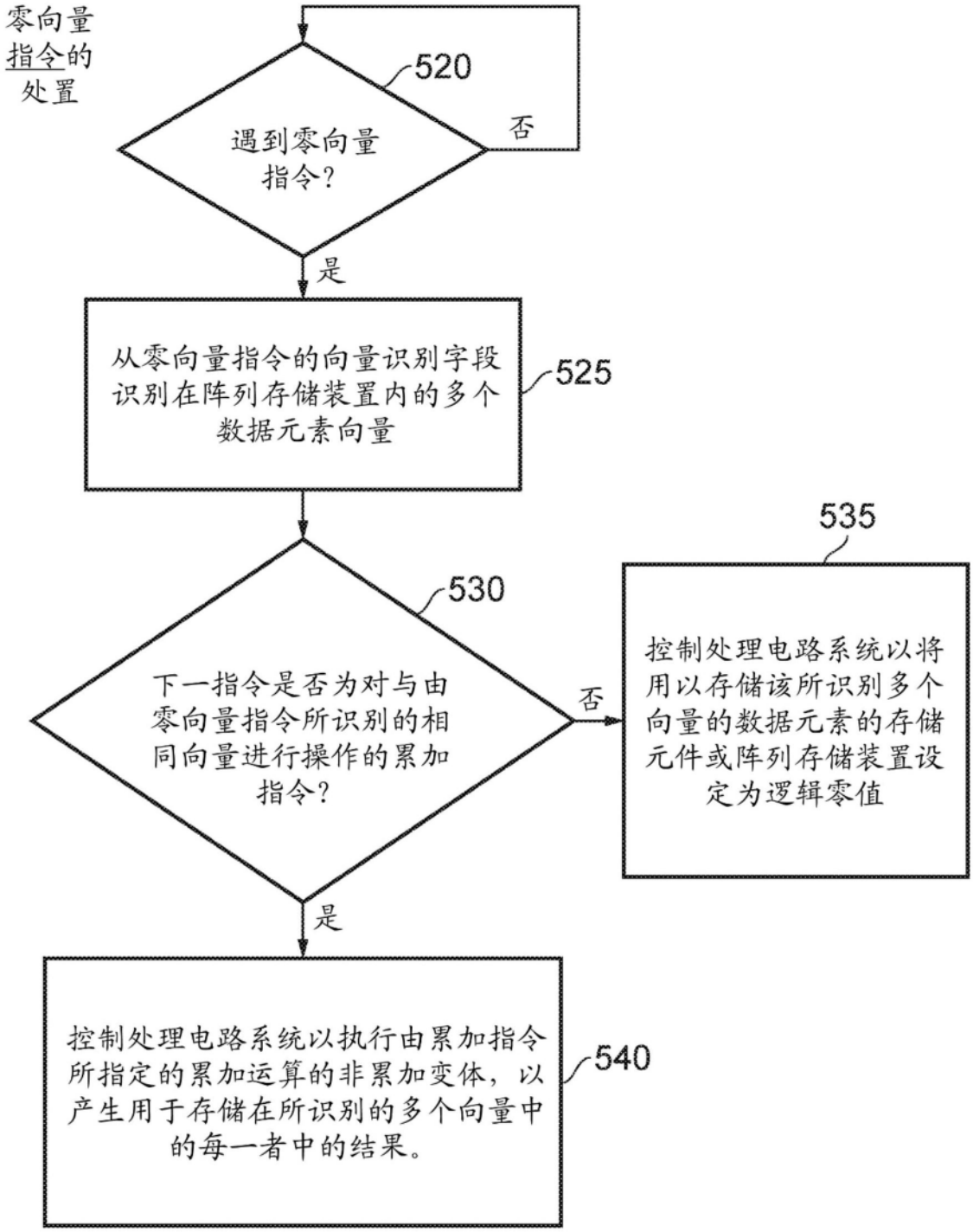

零向量指令的处置
520
遇到零向量指令?
否
是
从零向量指令的向量识别字段识别在阵列存储装置内的多个数据元素向量
525
530
下一指令是否为对与由零向量指令所识别的相同向量进行操作的累加指令?
否
535
控制处理电路系统以将用以存储该所识别多个向量的数据元素的存储元件或阵列存储装置设定为逻辑零值
是
控制处理电路系统以执行由累加指令所指定的累加运算的非累加变体，以产生用于存储在所识别的多个向量中的每一者中的结果。
540

图11

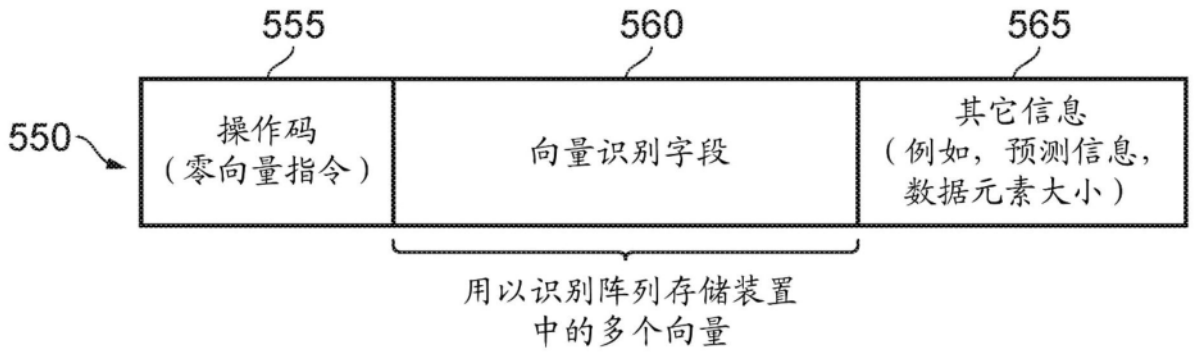

555
560
565
550
操作码
（零向量指令）
向量识别字段
其它信息
（例如，预测信息，
数据元素大小）
用以识别阵列存储装置
中的多个向量

图12

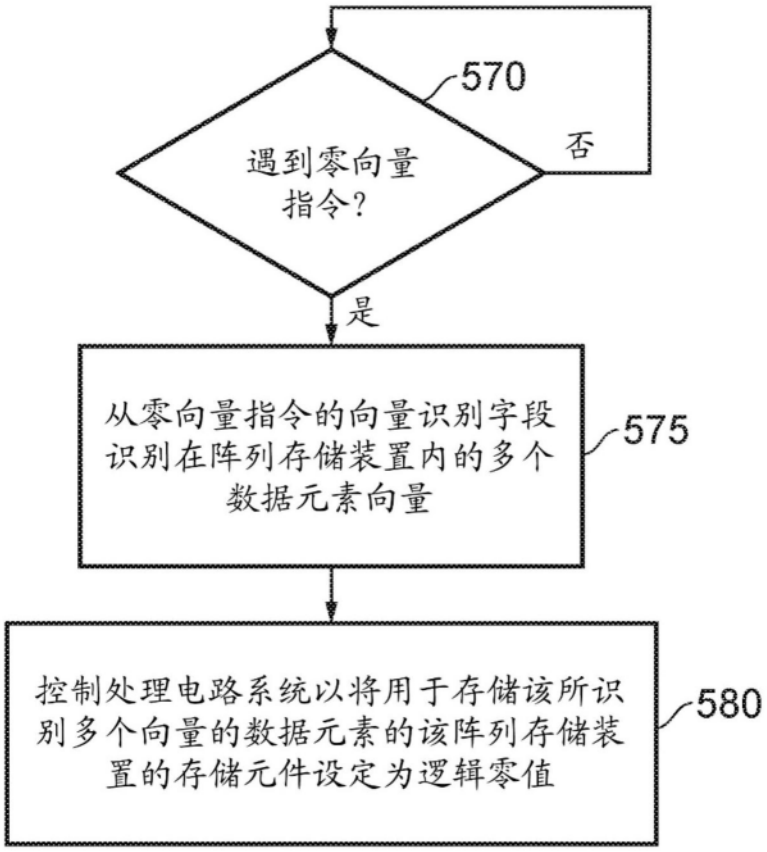

570
遇到零向量
指令?
否
是
从零向量指令的向量识别字段
识别在阵列存储装置内的多个
数据元素向量
575
控制处理电路系统以将用于存储该所识
别多个向量的数据元素的该阵列存储装
置的存储元件设定为逻辑零值
580

图13

模拟器
具体实施

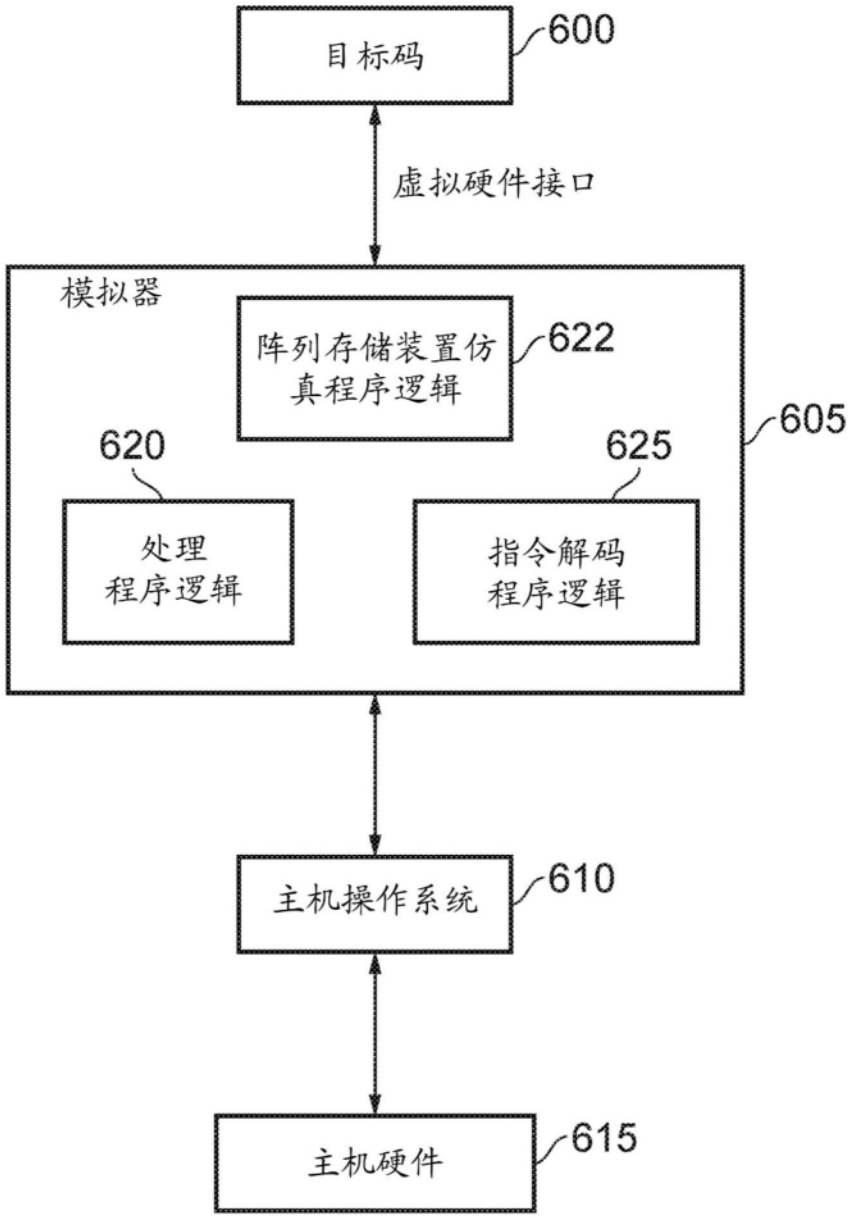

目标码
600
虚拟硬件接口
模拟器
阵列存储装置仿真程序逻辑
622
605
620
625
处理程序逻辑
指令解码程序逻辑
主机操作系统
610
主机硬件
615

图14