



(12) 发明专利申请

(10) 申请公布号 CN 116830077 A

(43) 申请公布日 2023. 09. 29

(21) 申请号 202280014048.0 奚锦文 D·C·伯格 E·S·钟

(22) 申请日 2022.01.20 (74) 专利代理机构 北京市金杜律师事务所
11256
专利代理师 李辉 辛鸣

(30) 优先权数据
63/148,086 2021.02.10 US
17/361,263 2021.06.28 US

(85) PCT国际申请进入国家阶段日
2023.08.08

(86) PCT国际申请的申请数据
PCT/US2022/013086 2022.01.20

(87) PCT国际申请的公布数据
W02022/173572 EN 2022.08.18

(71) 申请人 微软技术许可有限责任公司
地址 美国华盛顿州

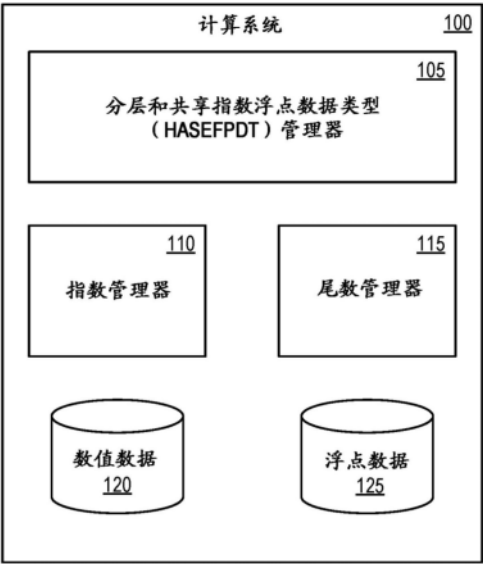
(72) 发明人 B·达尔维什·鲁哈尼 V·埃兰戈
R·沙菲普尔 J·弗沃斯 刘明罡

权利要求书2页 说明书11页 附图9页

(54) 发明名称
分层和共享指数浮点数据类型

(57) 摘要

本公开的实施例包括用于提供分层和共享指数浮点数据类型的系统和方法。第一共享指数值和第二共享指数值基于多个浮点值的指数值而被确定。第三共享指数值基于第一共享指数值和第二共享指数值而被确定。第一差值和第二差值基于第一共享指数值、第二共享指数值和第三共享指数值而被确定。符号值和尾数值针对多个浮点值而被确定。针对多个浮点值中的每个浮点值的符号值和尾数值、第三共享指数值、第一差值和第二差值被存储在针对共享指数浮点数据类型的数据结构中。



1. 一种机器可读介质, 存储由设备的至少一个处理单元可执行的程序, 所述程序包括用于以下各项的指令集:

基于多个浮点值的第一子集中的每个浮点值的指数值来确定第一共享指数值;

基于所述多个浮点值的第二子集中的每个浮点值的指数值来确定第二共享指数值;

基于所述第一共享指数值和所述第二共享指数值来确定第三共享指数值;

基于所述第一共享指数值和所述第三共享指数值来确定第一差值;

基于所述第二共享指数值和所述第三共享指数值来确定第二差值;

针对所述多个浮点值中的每个浮点值确定符号值和尾数值; 以及

在针对共享指数浮点数据类型的数据结构中存储针对所述多个浮点值中的每个浮点值的所述符号值和所述尾数值、所述第三共享指数值、所述第一差值和所述第二差值,

其中针对神经网络使用所述共享指数浮点数据增加所述神经网络的效率。

2. 根据权利要求1所述的机器可读介质, 其中所述第一共享指数值、所述第二共享指数值和所述第三共享指数值形成共享指数值的层级。

3. 根据权利要求2所述的机器可读介质, 其中所述第一共享指数值和所述第二共享指数值是所述共享指数值的层级中的第一级指数值, 其中所述第三共享指数值是所述共享指数值的层级中的第二级指数值。

4. 根据权利要求2所述的机器可读介质, 其中所述共享指数值的层级由所述第三共享指数值、所述第一差值和所述第二差值表示。

5. 根据权利要求1所述的机器可读介质, 其中所述程序还包括用于以下各项的指令集:

基于所述第一共享指数值和所述多个浮点值的所述第一子集中的每个浮点数的所述指数值来确定差值的第一集合;

基于所述第二共享指数值和所述多个浮点值的所述第二子集中的每个浮点数的所述指数值来确定差值的第二集合; 以及

在针对所述共享指数浮点数据类型的所述数据结构中存储所述差值的第一集合和所述差值的第二集合。

6. 根据权利要求1所述的机器可读介质, 其中确定所述第一共享指数值包括: 确定多个浮点值的所述第一子集中的浮点值的、具有最高值的指数值, 以及使用所述指数值作为所述第一共享指数值。

7. 根据权利要求1所述的机器可读介质, 其中确定所述第二共享指数值包括: 确定多个浮点值的所述第二子集中的浮点值的、具有最高值的指数值, 以及使用所述指数值作为所述第二共享指数值。

8. 根据权利要求1所述的机器可读介质, 其中确定所述第三共享指数值包括: 在确定所述第一共享指数值大于或等于所述第二共享指数值时, 使用所述第一共享指数值作为所述第三指数值, 以及在确定所述第一共享指数值小于所述第二共享指数值时, 使用所述第二共享指数值作为所述第三指数值。

9. 根据权利要求1所述的机器可读介质, 其中被用于表示被存储在针对所述共享指数浮点数据类型的所述数据结构中的所述尾数值的比特的数目小于被用于表示所述第三指数值的比特的数目。

10. 一种方法, 包括:

基于多个浮点值的第一子集中的每个浮点值的指数值来确定第一共享指数值；
基于所述多个浮点值的第二子集中的每个浮点值的指数值来确定第二共享指数值；
基于所述第一共享指数值和所述第二共享指数值来确定第三共享指数值；
基于所述第一共享指数值和所述第三共享指数值来确定第一差值；
基于所述第二共享指数值和所述第三共享指数值来确定第二差值；
针对所述多个浮点值中的每个浮点值确定符号值和尾数值；以及

在针对共享指数浮点数据类型的数据结构中存储针对所述多个浮点值中的每个浮点值的所述符号值和所述尾数值、所述第三共享指数值、所述第一差值和所述第二差值，
其中针对神经网络使用所述共享指数浮点数据增加所述神经网络的效率。

11. 根据权利要求10所述的方法，其中所述第一共享指数值、所述第二共享指数值和所述第三共享指数值形成共享指数值的层级。

12. 根据权利要求11所述的方法，其中所述第一共享指数值和所述第二共享指数值是所述共享指数值的层级中的第一级指数值，其中所述第三共享指数值是所述共享指数值的层级中的第二级指数值。

13. 根据权利要求11所述的方法，其中所述共享指数值的层级由所述第三共享指数值、所述第一差值和所述第二差值表示。

14. 根据权利要求10所述的方法，还包括：

基于所述第一共享指数值和所述多个浮点值的所述第一子集中的每个浮点数的所述指数值来确定差值的第一集合；

基于所述第二共享指数值和所述多个浮点值的所述第二子集中的每个浮点数的所述指数值来确定差值的第二集合；以及

在针对所述共享指数浮点数据类型的数据结构中存储所述差值的第一集合和所述差值的第二集合。

15. 一种系统，包括：

一组处理单元；以及

机器可读介质，存储指令，所述指令在由所述一组处理单元中的至少一个处理单元执行时，使所述至少一个处理器单元：

基于多个浮点值的第一子集中的每个浮点值的指数值来确定第一共享指数值；

基于所述多个浮点值的第二子集中的每个浮点值的指数值来确定第二共享指数值；

基于所述第一共享指数值和所述第二共享指数值来确定第三共享指数值；

基于所述第一共享指数值和所述第三共享指数值来确定第一差值；

基于所述第二共享指数值和所述第三共享指数值来确定第二差值；

针对所述多个浮点值中的每个浮点值确定符号值和尾数值；以及

在针对共享指数浮点数据类型的数据结构中存储针对所述多个浮点值中的每个浮点值的所述符号值和所述尾数值、所述第三共享指数值、所述第一差值和所述第二差值，
其中针对神经网络使用所述共享指数浮点数据增加所述神经网络的效率。

分层和共享指数浮点数据类型

背景技术

[0001] 本公开涉及计算硬件。更具体地,本公开涉及浮点数据类型。

[0002] 神经网络是一种用于各种不同应用(例如,图像分类、计算机视觉、自然语言处理、语音识别、书写识别等)的机器学习模型。通过神经网络运行数据集,将来自神经网络的结果与已知结果进行比较,并且基于差值来更新网络,可以针对特定目的来训练神经网络。

[0003] 神经网络的高效训练和使用神经网络在低保真度数据类型下进行推理可能需要开发使每个位的保真度最大化同时使计算成本最小化的数据类型。这可以被公式化为优化问题,其中目标是使量化信噪比(QSNR)度量最大化,同时使硬件点积单元的面积开销最小化。

附图说明

[0004] 本公开的各种实施例在附图的图中以示例而非限制的方式示出。

[0005] 图1示出了根据一些实施例的计算系统。

[0006] 图2示出了根据一些实施例的将浮点值转换为分层和共享指数浮点数据类型的示例。

[0007] 图3示出了根据一些实施例的用于存储图2所示的分层和共享指数浮点数据类型的示例数据结构。

[0008] 图4示出了根据一些实施例的将浮点值转换为分层和共享指数浮点数据类型的另一示例。

[0009] 图5示出了根据一些实施例的用于存储图4所示的分层和共享指数浮点数据类型的示例数据结构。

[0010] 图6示出了根据一些实施例的将分层和共享指数浮点数据类型变量相乘的示例。

[0011] 图7示出了根据一些实施例的用于创建分层和共享指数浮点数据类型变量的过程。

[0012] 图8描绘了根据一些实施例的示例计算机系统的简化框图。

[0013] 图9示出了根据一些实施例的神经网络处理系统。

具体实施方式

[0014] 在以下描述中,出于解释的目的,为了提供对本公开的全面理解,列举了很多示例和具体细节。这样的示例和细节不应当被解释为不适当地将权利要求的元素或所要求保护的主体限制为一个整体。基于不同权利要求的语言,本领域技术人员将很清楚的是,所要求保护的主体可以单独地或组合地包括这些示例中的一些或全部特征,并且还可以包括本文中描述的特征和技术的修改和等同方案。

[0015] 这里描述了用于提供分层和共享指数浮点数据类型的技术。在一些实施例中,计算系统可以被配置为创建和管理分层和共享指数浮点数据类型数据。例如,计算系统可以

从若干数字创建分层和共享指数浮点数据类型数据,如果需要,这些数字已经转换为浮点数(例如,半精度浮点数、单精度浮点数和双精度浮点数等)。为此,计算系统首先确定每个浮点数的指数值。然后,计算系统基于共享指数值的层级来量化这些指数值。接下来,计算系统基于量化的指数值来量化每个浮点数的尾数。最后,计算系统将共享指数值的层级和量化尾数存储在用于分层和共享指数浮点数据类型的数据结构中。

[0016] 在一些实施例中,超窄位宽(例如,尾数的1-4位)下的高效训练/推理可能需要开发数据类型,该数据类型使每个位的保真度最大化,同时使计算成本最小化。这可以被公式化为优化问题,其中目标是使量化信噪比(QSNR)最大化,同时使硬件点积单元的面积开销最小化。如将要讨论的,本文中描述的分层和共享指数浮点数据类型可以进一步推动QSNR与乘累加(MAC)区域Pareto边界,并且与其他浮点数据类型相比挤出更多效率。在这种超窄位宽情况下,高效QSNR的这种增加可以直接转化为训练神经网络和使用它们进行推理的更好的可实现精度。

[0017] 在一些实施例中,本文中描述的数据类型是共享指数数据类型。也就是说,指数是在若干元素之间共享的。元素的数目可以被称为瓦片大小或块大小。从硬件的角度来看,使用小的瓦片大小可能不是最佳选择,因为随着尾数位的数目减少,整数乘法器的成本可以基本消除,并且累加器开始主导乘法器累加器(MAC)单元的总面积。因此,采用相对较大瓦片大小允许累加器硅成本摊销,同时避免由于瓦片的私有指数与共享指数之间的差值而将瓦片内的元素过度映射为零。

[0018] 在一些实施例中,本文中描述的数据类型旨在在超窄位宽情况下提供可接受瓦片大小与编码效率之间的平衡。编码效率可以由以下等式定义:

$$[0019] \quad QSNR_{dB} = 20 \log_{10} \frac{\|x\|}{\|x-Q(x)\|} = 10 \log_{10} \frac{\|x\|^2}{\|x-Q(x)\|^2} = 10 \log_{10} \frac{\text{信号能量}}{\text{噪声能量}}$$

[0020] 其中 $Q(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是n维向量x的量子化方法, $\|x\|$ 是欧几里得范数。n表示分层和共享指数浮点数据类型中的瓦片大小。QSNR可以通过在很多向量x上对上述比率求平均来测量。在一些实施例中,该定义可以修改以测量MAC单元的信噪比,其中x被替换为 $x=Q(y)*Q(z^T)$, 其中y和z指示到矩阵乘法单元的输入。

[0021] 在一些实施例中,分层和共享指数浮点数据类型将每个瓦片分解为嵌套的子瓦片,并且允许每个子瓦片的共享指数与其下一级子瓦片相差标度值。这些标度可以在点积单元中部署的加法器树中实现为右移。

[0022] 在一些实施例中,本文中描述的数据类型具有若干可调谐参数。一个参数是瓦片大小,它是共享全局指数的元素的数目。另一参数是尾数位的数目。在一些实施例中,尾数值使用2的补码来表示,而在其他实施例中,尾数值使用符号幅度格式来表示。尾数位的数目可以取大于或等于1的任何值。

[0023] 还有一个参数是共享指数的表示。在一些实施例中,共享指数可以使用单个级别来表示。对于单级表示,一个参数是子瓦片大小,它是共享局部标度以调节全局指数的元素的数目。每个瓦片可以被拆分成具有大小相等m的多个子瓦片,其中每个子瓦片共享公共子标度。如果 $m=1$,则可以将右移器添加到点积单元中的加法器树的叶子。m的值越大,右移器就会移向加法器树的根。例如,如果 $m=16$,则右移器被添加到加法器树的第 $\log_2(16)=4$

级。单级表示的另一参数是子瓦片标度,子瓦片标度指定每元素允许的右移量。该标度对全局共享指数与子瓦片内的所选择的指数之间允许的最大差值进行编码。例如,加法器树的叶子处的为1的标度表示每个输入元素可以移位0或1,从而产生2位组合mux。在某些情况下,每瓦片的最大指数可以用作全局共享指数。因此,标度始终是大于或等于零的正数。在一些实施例中,可以使用其他指数选择技术。在全局共享指数不是瓦片的最大指数的情况下,所选择的标度可以是正值或负值。

[0024] 在其他实施例中,共享指数以分层方式表示。不是使用大的单级移位来缩放共享指数,而是将移位分布到加法器树的多个级别,并且限制每移位的标度。分层表示中的一个参数是分层级别,分层级别是加法器树中用于调节共享指数的分层级别的数目。索引从树的叶开始。例如,对于为32的瓦片大小,加法器树中将最多有5个级别。分层级别3表示在加法器树的前三个级别内从叶子开始使用右移。分层表示的另一参数是子瓦片标度,子瓦片标度是每个级别下允许的每元素最大标度。该标度使用下一级共享指数(不一定是全局指数)对差值进行编码。共享指数可以是也可以不是最大指数。例如,如果计算瓦片的最大和最小指数,则最大与最小之间的差值给出最大标度差。如果分层级别设置为5,则该(max-min)指数差可以上限为5(因为有5个级别的子瓦片),并且全局指数可以设置为(max-min).clip(5)/2。这会按层级将瓦片拆分为子瓦片,并且在每个级别,它会检查为-1/0/1的标度是否必须基于子瓦片的指数来设置。

[0025] 累加器位宽参数指定加法器树的末尾处用于累加的数据类型。这样的数据类型的示例包括半精度浮点数据类型、单精度浮点数据类型、双精度浮点数据类型等。在一些实施例中,默认数据类型可以是单精度浮点值。截断位宽可以是另一参数。该参数指定每次整数相加之后的截断量,其可以高效地减少区域开销。高效截断方案的一个示例是将加法器树的叶处的整数乘法器的输出截断为8位尾数,并且允许加法器树的每一级处的整数加法器的输出的位宽增长1。例如,这种方法可以减少面积开销,同时对可实现的QSNR(即,精度)的影响最小。

[0026] 图1示出了根据一些实施例的计算系统100。如图所示,计算系统100包括分层和共享指数浮点数据类型(HASEFPDT)管理器105、指数管理器110、尾数管理器115、数值数据存储装置120和浮点数据存储装置125。数值数据存储装置120存储可以被转换为分层和共享指数浮点数据类型的数值数据。数值数据的示例包括浮点数、整数等。浮点数据存储装置125被配置为存储分层和共享指数浮点数据类型。以下将详细描述这种数据类型的示例。此外,浮点数据存储装置125可以存储分层和共享指数浮点数据类型的定义。

[0027] HASEFPDT管理器105负责管理分层和共享指数浮点数据类型以及根据这样的数据类型而存储的数据。例如,HASEFPDT管理器105可以通过从浮点数据存储装置125中检索分层和共享指数浮点数据类型的定义来创建存储为分层和共享指数浮点数据类型的数据。接下来,HASEFPDT管理器105可以从数值数据存储装置120中检索数值数据,并且将数值数据转换为浮点数(例如,半精度浮点数、单精度浮点数和双精度浮点数等)。HASEFPDT管理器105中检索的数字数据的数目可以基于在分层和共享指数浮点数据类型的定义中指定的内容。然后,HASEFPDT管理器105向指数管理器110发送转换后的数值数据、分层和共享指数浮点数据类型的定义、以及确定数值数据的共享指数值的请求。当HASEFPDT管理器105从指数管理器110接收到共享指数值时,HASEFPDT管理器105向尾数管理器115发送数字数据、共享

指数值、分层和共享指数浮点数据类型的定义、以及确定数字数据的尾数值的请求。一旦HASEFPDT管理器105从尾数管理器115接收到尾数值,HASEFPDT管理器105就将尾数值和共享指数值存储在数据结构中,该数据结构被配置为存储分层和共享指数浮点数据类型的数据。

[0028] HASEFPDT管理器105还可以从分层和共享指数浮点数据类型数据中读取值。例如HASEFPDT管理器105可以接收(例如,从在计算系统100上操作的组件、从在另一计算系统上操作的组件等)从根据分层和共享指数浮点数据类型而存储的数据(例如,用于根据分层和共享指数浮点数据类型来存储数据的数据结构)中读取特定值的请求。作为响应,HASEFPDT管理器105从分层和共享指数浮点数据类型数据中检索请求值的尾数值,基于存储在分层和共享指数浮点数据类型数据中的共享指数值来确定请求值的指数值,并且基于尾数值和指数值来生成请求值。

[0029] 指数管理器110被配置为确定分层和共享指数浮点数据类型的指数值。例如,指数管理器110可以从HASEFPDT管理器105接收数字数据、分层和共享指数浮点数据类型的定义、以及确定数字数据的指数值的请求。作为响应,指数管理器110基于数值数据以及分层和共享指数浮点数据类型的定义来确定指数值的层级。然后,指数管理器110将指数值的层级发送给HASEFPDT管理器105。

[0030] 尾数管理器115处理分层和共享指数浮点数据类型的尾数值的确定。例如,尾数管理器115可以从HASEFPDT管理器105接收数字数据、共享指数值、分层和共享指数浮点数据类型的定义、以及确定数字数据的尾数值的请求。基于共享指数值以及分层和共享指数浮点数据类型的定义,尾数管理器115确定数值数据的尾数值。然后,尾数管理器115将尾数值发送给HASEFPDT管理器105。

[0031] 图2示出了根据一些实施例的将浮点值转换为分层和共享指数浮点数据类型的示例。具体地,图2示出了表200,该表200包括在将浮点值转换为分层和共享指数浮点数据类型的整个过程中的各种值。将参考图1来描述该示例。对于该示例,行205中的四个值将被转换为分层和共享指数浮点数据类型,该数据类型使用具有三个级别的共享指数值的层级来存储共享指数值。如图所示,这四个值以十进制值表示。HASEFPDT管理器105通过从浮点数据存储装置125中检索分层和共享指数浮点数据类型的定义开始。接下来,HASEFPDT管理器105从数值数据存储装置120中检索行205中的值,并且将它们转换为浮点值(例如,半精度浮点值、单精度浮点值和双精度浮点值等)。然后,HASEFPDT管理器105向指数管理器110发送浮点值、分层和共享指数浮点数据类型的定义、以及确定浮点值的指数值的请求。

[0032] 在接收到浮点值、分层和共享指数浮点数据类型的定义、以及请求之后,指数管理器110确定每个浮点值的指数值,如表200的行210所示。在该示例中,这些指数值被认为是0级指数值,如行215中所示。指数管理器110然后确定每对指数值的1级共享指数值。对于该示例,指数管理器110通过确定每对指数值中的具有最高值的指数值并且将其用作1级共享指数值来确定1级共享指数值。如行220所示,第一对值的指数值之间的最高指数值是-3。第二对值的指数值之间的最高指数值是-2。接下来,指数管理器110确定每对0级指数值和对对应1级共享指数值的1级标度值。这里,指数管理器110通过确定0级指数值与对应的1级共享指数值之间的差值来确定1级标度值。如果该差值大于1,则指数管理器110将1级标度值近似为1。

[0033] 对于第一对和第二对0级指数值以及对应1级共享指数值和第二值,指数管理器110将1级标度值每个确定为0(即,-3与-3之间的差值为0)。对于第三对0级指数值和对应1级共享指数值,指数管理器110将0级指数值与1级共享指数值之间的差值确定为2。由于这大于1,指数管理器110将1级标度值近似为1。对于第四对0级指数值和对应1级共享指数值,指数管理器110将1级标度值确定为0(即,-2与-2之间的差值为0)。1级标度值在表200的行225中描绘。

[0034] 在确定1级共享指数值和标度值之后,指数管理器110确定2级共享指数值。在该示例中,指数管理器110基于1级共享指数值来确定2级共享指数值。特别地,指数管理器110通过确定1级共享指数值中的具有最高值的指数值并且将其用作2级共享指数值来确定2级共享指数值。这里,指数管理器110将2级共享指数值确定为-2,如行230中所示。

[0035] 接下来,指数管理器110确定每对1级共享指数值和对应2级共享指数值的2级标度值。对于该示例,指数管理器110通过确定1级共享指数值与对应2级共享指数值之间的差值来确定2级标度值。如果该差值大于1,则指数管理器110将2级标度值近似为1。对于第一对1级共享指数值和对应2级共享指数值,指数管理器110将2级标度值确定为1(即,-2与-3之间的差值为1)。对于第二对1级共享指数值和对应2级共享指数值,指数管理器110将2级标度值确定为0(即,-2与-2之间的差值为0)。2级标度值在行235中示出。行240描绘了每个浮点值的总标度值,其是1级标度值和2级标度值的总和。行245示出了每个浮点值的总共享指数标度值,其是通过从2级共享指数值中减去相应总标度值而为浮点值而确定的指数值。行250描绘了每个浮点值的所确定的指数值与原始指数值之间的差值。行255示出了浮点值的尾数的量化幅度。在该示例中,2位尾数与用于符号的第一比特(表200中未示出)和用于幅度的第二比特一起使用。行260示出了每个浮点值的量化指数值,其与行245中的共享指数值相同。最后,行265示出了在将原始浮点值转换为分层和共享指数浮点数据类型之后的量化浮点值的十进制表示。

[0036] 在指数管理器110确定2级共享指数、2级标度值和1级标度值之后,指数管理器将它们发送给HASEFPDT管理器105。然后,HASEFPDT管理器105向尾数管理器115发送行205中的浮点值、2级共享指数、2级标度值、1级标度值、分层和共享指数浮点数据类型的定义、以及确定浮点值的尾数值的请求。当尾数管理器115接收到该数据和请求时,尾数管理器115确定每个浮点值的尾数值。如果指数差(在行250中示出)大于0,则尾数管理器115将相应浮点值的尾数的幅度量化为0。否则,尾数管理器115使用可用比特(在该示例中,符号为一个比特,幅度为一个比特)和相应量化指数值(在行260中示出)来最佳地近似原始浮点值。这里,尾数管理器115将每个浮点值的量化幅度值确定为行255中所描绘的那些。接下来,尾数管理器115将所确定的尾数值发送给HASEFPDT管理器105。

[0037] 一旦HASEFPDT管理器105从尾数管理器115接收到尾数值,HASEFPDT管理器105就将尾数值、2级共享指数、2级标度值和1级标度值存储在数据结构中,该数据结构被配置为存储分层和共享指数浮点数据类型的数据。图3示出了根据一些实施例的用于存储图2所示的分层和共享指数浮点数据类型的示例数据结构300。数据结构300可以基于分层和共享指数浮点数据类型的定义来生成。这里,数据类型的定义将指定四个浮点值的四个2位尾数值、四个1位1级标度值、两个1位2级标度值和一个8位2级共享指数值。如图所示,数据结构300包括四个属性:尾数属性、1级标度属性、2级标度属性和共享指数属性。尾数属性存储为

每个浮点值而确定的2位尾数值。1级标度属性存储为每个浮点值而确定的1位1级标度值。2级标度值属性存储为每对浮点值而确定的1位2级标度值。共享指数属性值存储为浮点值而确定的8位2级共享指数值。在该示例中,2级共享指数值的值使用为127的偏差来表示。因此,对于该示例,值125(即, $127+(-2)$)存储在共享指数属性中。

[0038] 图4示出了根据一些实施例的将浮点值转换为分层和共享指数浮点数据类型的另一示例。特别地,图4示出了表400,该表400包括在将浮点值转换为不同分层和共享指数浮点数据类型的整个过程中的各种值。还将参考图1来描述该示例。在该示例中,在上面的示例中使用的行405中描绘的相同的四个值将被转换为分层和共享指数浮点数据类型,该数据类型使用具有两个级别的共享指数值的层级来存储共享指数值。该示例开始于HASEFPDT管理器105从浮点数据存储装置125中检索分层和共享指数浮点数据类型的定义。然后,HASEFPDT管理器105从数值数据存储装置120中检索行405中的值,并且将其转换为浮点值(例如,半精度浮点值、单精度浮点值和双精度浮点值等)。接下来,HASEFPDT管理器105向指数管理器110发送浮点值、分层和共享指数浮点数据类型的定义、以及确定浮点值的指数值的请求。

[0039] 在接收到浮点值、分层和共享指数浮点数据类型的定义、以及请求之后,指数管理器110确定每个浮点值的指数值,如行410中所示。对于该示例,指数管理器110将每对原始指数值的共享指数值确定为0级共享指数。在该示例中,指数管理器110通过确定每对原始指数值中的具有最高值的指数值并且将其用作0级共享指数值来确定0级共享指数值。如行415所示,第一对值的指数值之间的最高指数值是-3,第二对值的指数值之间的最高指数值是-2。指数管理器110然后基于0级共享指数值来确定1级共享指数。对于该示例,指数管理器110通过确定具有最高值的0级共享指数值并且将其用作1级共享指数值来确定1级共享指数。这里,指数管理器110确定1级共享指数值是-2,如行420中所示。接下来,指数管理器110确定每对0级共享指数值 and 对应1级共享指数值的1级标度值。在该示例中,指数管理器110通过确定0级共享指数值与对应1级共享指数值之间的差值来确定1级标度值。如果该差值大于1,则指数管理器110将1级标度值近似为1。

[0040] 对于第一对0级共享指数值 and 对应1级共享指数值,指数管理器110将1级标度值确定为1(即,-2与-3之间的差值为1)。对于第二对0级共享指数值 and 对应1级共享指数值,指数管理器110将1级标度值确定为0(即,-2与-2之间的差值为0)。1级标度值在行425中示出。行430描绘了与1级标度值相同的总标度值。行435示出了每个浮点值的总共享指数标度值,其是通过从1级共享指数值中减去相应总标度值而为浮点值而确定的指数值。行440描绘了每个浮点值的所确定的指数值与原始指数值之间的差值。行445示出了浮点值的尾数的量化幅度。在该示例中,3位尾数与用于符号的第一比特(表400中未示出)和用于幅度的第二比特一起使用。行450示出了每个浮点值的量化指数值,其与行435中的共享指数值相同。然后,行455描绘了在将原始浮点值转换为分层和共享指数浮点数据类型之后的量化浮点值的十进制表示。

[0041] 在确定1级共享指数和1级标度值之后,指数管理器110将它们发送给HASEFPDT管理器105。HASEFPDT管理器105向尾数管理器115发送行405中的浮点值、1级共享指数、1级标度值、分层和共享指数浮点数据类型的定义、以及确定浮点值的尾数值的请求。一旦尾数管理器115接收到该数据和请求,尾数管理器115就确定每个浮点值的尾数值。如果指数差(在

行440中示出)大于1,则尾数管理器115将相应浮点值的尾数的幅度量化为0。否则,尾数管理器115使用可用比特(在该示例中,符号为一个比特,幅度为一个比特)和相应量化指数值(在行450中示出)来最佳地近似原始浮点值。尾数管理器115将每个浮点值的量化幅度值确定为行445中所示的那些。然后,尾数管理器115将所确定的尾数值发送给HASEFPDT管理器105。

[0042] 当HASEFPDT管理器105从尾数管理器115接收尾数值时,HASEFPDT管理器105将尾数值、1级共享指数和1级标度值存储在数据结构中,该数据结构被配置为存储分层和共享指数浮点数据类型的数据。图5示出了根据一些实施例的用于存储图4所示的分层和共享指数浮点数据类型的示例数据结构500。数据结构500可以基于分层和共享指数浮点数据类型的定义来生成。对于该示例,数据类型的定义将指定四个浮点值的四个2位尾数值、两个2位1级标度值和一个8位1级共享指数值。如图所示,数据结构500包括三个属性:尾数属性、1级标度属性和共享指数属性。尾数属性存储为每个浮点值而确定的2位尾数值。1级标度属性存储为每个浮点值而确定的2位1级标度值。共享指数属性值存储为浮点值而确定的8位1级共享指数值。在该示例中,1级共享指数值的值使用为127的偏差来表示。因此,对于该示例,值125(即, $127+(-2)$)存储在共享指数属性中。

[0043] 上面参考图1-图5描述的示例演示了如何将浮点数转换为两种不同的分层和共享指数浮点数据类型。本领域普通技术人员将理解,在一些实施例中,可以使用附加的和/或不同的分层和共享指数浮点数据类型。例如,分层和共享指数浮点数据类型可以表示不同数目的浮点值,这些浮点值共享指数(上面的示例每个表示四个浮点值),使用不同数目的比特来表示尾数值、级别标度值和共享指数值,在指数值的层级中使用不同数目的级别(例如,不同数目的标度值级别),等等。

[0044] 此外,上面参考图2-图5解释的示例示出了共享指数值的层级如何由最高级别共享指数值和级别标度值表示。例如,图2所示的示例使用具有三个不同级别的共享指数的层级:0级指数值、1级共享指数值和2级共享指数值。如图3所示,共享指数值的层级由2级共享指数值、2级标度值和1级标度值表示。浮点值的量化指数值(行260中的值)可以根据2级共享指数值、2级标度值和1级标度值来确定。作为另一示例,图4所示的示例使用具有两个不同级别的共享指数的层级:0级指数值和1级共享指数值。如图5所示,共享指数值的层级由1级共享指数值和1级标度值表示。浮点值的量化指数值(行450中的值)可以根据1级共享指数值和1级标度值来确定。HASEFPDT管理器105在读取根据分层和共享指数浮点数据类型而存储的浮点值(例如,确定如行265和455中所示的量化值)时以这种方式确定量化指数值。

[0045] 此外,上面参考图2-图5描述的示例通过在较低级别指数/共享指数中选择最高值来确定共享指数值。本领域普通技术人员将理解,可以使用附加的和/或不同的统计度量来确定特定级别的共享指数值。这样的度量的示例包括加权平均值、均值、加上定义数目的标准偏差(例如,一个标准偏差、两个标准偏差和三个标准偏差等)。

[0046] 分层和共享指数浮点数据类型的示例应用

[0047] 上述示例解释了如何创建分层和共享指数浮点数据类型。以下示例说明了这些数据类型的很多应用中的一个。图6示出了根据一些实施例的将分层和共享指数浮点数据类型变量相乘的示例。具体地,图6示出了用于在根据分层和共享指数浮点数据类型而存储的两个变量之间执行点积的硬件架构600。在一些实施例中,硬件架构600可以用于实现人工

智能(AI)加速器,该AI加速器用于训练神经网络并且使用神经网络进行推理。如图所示,硬件架构600包括乘法器605-1-605-n、移位器610-1-610-n和620-1-620-m、加法器615-1-615-m、625-1-625-k和630、以及截断/舍入模块635。

[0048] 对于该示例,硬件架构600被配置为接收两个输入A和B。输入A和B中的每个是存储n个浮点值(例如,数据结构300、数据结构500等)的分层和共享指数浮点数据类型变量。每个变量中浮点值的尾数值被输入到对应乘法器605中。例如,变量A中的第一浮点值的尾数值和变量B中的第一浮点值的尾数值被输入到乘法器605-1,变量A的第二浮点值的尾数值和变量B的第二浮点值的尾数值被输入到乘法器605-2,变量A中的第三浮点值的尾数值和变量B中的第三浮点值的尾数值被输入到乘法器605-3,等等。如图所示,每个乘法器605中的输入具有位宽dw。每个乘法器605将两个输入尾数值相乘,并且输出具有位宽2dw的乘积。乘法器605的输出被输入到对应移位器610,该移位器610被配置为基于在移位器610的顶部描绘的控制输入来执行右移。移位器610的控制输入是乘法器605的每个输入的1级标度值的总和。移位器610执行与控制输入的值相等的多个右移操作。移位器610的输出具有位宽2dw+3。两个移位器610的输出被输入到对应加法器615,该加法器615将两个输入相加在一起以产生位宽为2dw+4的和。该过程继续使用对应级别标度值来控制移位器(例如,2级标度值的总和用于控制移位器620-1-620-m)。级别的数目由 $\log_2(n)$ 决定。如上所述,n是存储在变量中的浮点值的数目。最后一级的输出具有位宽 $2dw+4\log_2(n)$ 。

[0049] 当截断/舍入模块635接收到来自最后一级的输出时,截断/舍入数模块635基于每个变量的共享指数值的总和对其输出的值执行移位操作。例如,如果共享指数值的总和是正值,则截断/舍入模块635对等于共享指数值的总和的值执行多个左移位操作。如果共享指数值的总和是负值,则截断/舍入模块635对等于共享指数值的总和的绝对值的值执行多个右移位操作。然后,截断/舍入模块635将该值截断为定义数目的比特,以产生输出值C。

[0050] 图7示出了根据一些实施例的用于创建分层和共享指数浮点数据类型变量的过程700。在一些实施例中,计算系统100执行过程700。过程700开始于在710处基于多个浮点值的第一子集中的每个浮点值的指数值来确定第一共享指数值。参考图1和图4作为示例,指数管理器110可以基于第一对浮点值的原始指数值来确定0级共享指数值。

[0051] 接下来,在720,过程700基于多个浮点值的第二子集中的每个浮点值的指数值来确定第二共享指数值。参考图1和图4作为示例,指数管理器110可以基于第二对浮点值的原始指数值来确定0级共享指数值。

[0052] 然后,在730,过程700基于第一共享指数值和第二共享指数值来确定第三共享指数值。参考图1和图4作为示例,指数管理器110可以基于0级共享指数值来确定1级共享指数值。

[0053] 在740,过程700基于第一共享指数值和第三共享指数值来确定第一差值。参考图1和图4作为示例,指数管理器110可以确定第一对0级共享指数值和对应1级共享指数值的1级标度值。然后,在750,过程700基于第二共享指数值和第三共享指数值来确定第二差值。参考图1和图4作为示例,指数管理器110可以确定第二对0级共享指数值和对应1级共享指数值的1级标度值。

[0054] 接下来,在760,过程700针对多个浮点值中的每个浮点值确定符号值和尾数值。参考图1和图4作为示例,尾数管理器115可以确定尾数的量化幅度值,如表400的行445中所

示。在770,过程700最终在针对共享指数浮点数据类型的数据结构中存储针对多个浮点值中的每个浮点值的符号值和尾数值、第三共享指数值、第一差值和第二差值。参考图1和图5作为示例,HASEFPDT管理器105将尾数值、1级共享指数和1级标度值存储在数据结构500中。

[0055] 上述技术可以在被配置为处理神经网络的各种计算机系统中实现。图8描绘了示例计算机系统800的简化框图,其可以用于实现前述公开中描述的技术(例如,计算系统100)。如图8所示,计算机系统800包括一个或多个处理器802,处理器802经由总线子系统804与多个外围设备通信。这些外围设备可以包括存储子系统806(例如,包括存储器子系统808和文件存储子系统810)和网络接口子系统816。一些计算机系统还可以包括用户接口输入设备812和/或用户接口输出设备814。

[0056] 总线子系统804可以提供一种机制以用于让计算机系统800的各种组件和子系统按照预期彼此通信。尽管总线子系统804被示意性地示出为单个总线,但是总线子系统的替代实施例可以利用多个总线。

[0057] 网络接口子系统816可以用作计算机系统800与其他计算机系统或网络之间的数据通信的接口。网络接口子系统816的实施例可以包括例如以太网、Wi-Fi和/或蜂窝适配器、调制解调器(电话、卫星、电缆、ISDN等)、数字订户线(DSL)单元等。

[0058] 存储子系统806包括存储器子系统808和文件/磁盘存储子系统810。本文中描述的子系统808和810以及其他存储器是可以存储提供本公开的实施例的功能的可执行程序代码和/或数据的非暂态计算机可读存储介质的示例。

[0059] 存储器子系统808包括多个存储器,包括用于在程序执行期间存储指令和数据的主随机存取存储器(RAM) 818、以及其中存储有固定指令的只读存储器(ROM) 820。文件存储子系统810可以为程序和数据文件提供持久(例如,非易失性)存储,并且可以包括磁或固态硬盘驱动器、光学驱动器、以及相关联的可移动介质(例如,CD-ROM、DVD、蓝光等)、基于可移动闪存的驱动器或卡、和/或本领域已知的其他类型的存储介质。

[0060] 应当理解,计算机系统800是示例性的,并且具有比系统800更多或更少组件的很多其他配置也是可能的。

[0061] 图9示出了根据一些实施例的神经网络处理系统。在各种实施例中,根据本公开的神经网络可以在包括一个或多个神经网络处理器的硬件环境中实现和训练。神经网络处理器可以是指各种图形处理单元(GPU)(例如,用于处理由Nvidia Corp®生产的神经网络的GPU)、现场可编程门阵列(FPGA)(例如,用于处理由Xilinx®生产的神经网络的FPGA)、或者包括例如针对神经网络计算而优化的硬件架构的各种专用集成电路(ASIC)或神经网络处理器。在该示例环境中,一个或多个服务器902(其可以包括上图8所示的架构)可以通过通信网络901(例如,交换机、路由器等)耦合到多个控制器910(1)-910(M)。控制器910(2)-910(M)还可以包括上图8中示出的架构。每个控制器910(1)-910(M)可以耦合到一个或多个NN处理器,诸如处理器911(1)-911(N)和912(1)-912(N)。NN处理器911(1)-911(N)和912(1)-912(N)可以包括针对神经网络处理(诸如训练或推理)而优化的各种配置的功能处理块和存储器。NN处理器针对神经网络计算被优化。服务器902可以用NN模型以及到模型的输入数据来配置控制器910,这些模型可以由例如NN处理器911(1)-911(N)和912(1)-912(N)并行加载和执行。例如,模型可以包括如上所述的层和相关联的权重。NN处理器可以加载模型并且应用输入以产生输出结果。NN处理器还可以实现本文中描述的训练算法。

[0062] 另外的示例实施例

[0063] 在各种实施例中,本公开包括用于提供分层和共享指数浮点数据类型的系统、方法和装置。本文中描述的技术可以体现在一种存储由计算机系统可执行的程序的非暂态机器可读介质中,该程序包括用于执行本文中描述的技术的指令集。在一些实施例中,一种系统包括一组处理单元和存储指令的非暂态机器可读介质,该指令在由该组处理单元中的至少一个处理单元执行时,使该至少一个处理器执行上述技术。在一些实施例中,该非暂态机器可读介质可以是存储器,例如,该存储器可以耦合到一个或多个控制器或一个或多个人工智能处理器。

[0064] 以下技术可以单独体现或以不同组合体现,并且也可以进一步与本文中描述的其他技术一起体现。

[0065] 例如,在一个实施例中,本公开包括一种存储由设备的至少一个处理单元可执行的程序的非暂态机器可读介质。所述程序包括用于以下各项的指令集:基于多个浮点值的第一子集中的每个浮点值的指数值来确定第一共享指数值;基于所述多个浮点值的第二子集中的每个浮点值的指数值来确定第二共享指数值;基于所述第一共享指数值和所述第二共享指数值来确定第三共享指数值;基于所述第一共享指数值和所述第三共享指数值来确定第一差值;基于所述第二共享指数值和所述第三共享指数值来确定第二差值;针对所述多个浮点值中的每个浮点值确定符号值和尾数值;以及在针对共享指数浮点数据类型的数据结构中存储针对所述多个浮点值中的每个浮点值的所述符号值和所述尾数值、所述第三共享指数值、所述第一差值和所述第二差值。

[0066] 在一个实施例中,所述第一共享指数值、所述第二共享指数值和所述第三共享指数值形成共享指数值的层级。

[0067] 在一个实施例中,所述第一共享指数值和所述第二共享指数值是所述共享指数值的层级中的第一级指数值,其中所述第三共享指数值是所述共享指数值的层级中的第二级指数值。

[0068] 在一个实施例中,所述共享指数值的层级由所述第三共享指数值、所述第一差值和所述第二差值表示。

[0069] 在一个实施例中,本公开还基于所述第一共享指数值和所述多个浮点值的所述第一子集中的每个浮点数的所述指数值来确定差值的第一集合;基于所述第二共享指数值和所述多个浮点值的所述第二子集中的每个浮点数的所述指数值来确定差值的第二集合;以及在针对所述共享指数浮点数据类型的所述数据结构中存储所述差值的第一集合和所述差值的第二集合。

[0070] 在一个实施例中,确定所述第一共享指数值包括:确定多个浮点值的所述第一子集中的、具有最高值的浮点值的指数值,以及使用所述指数值作为所述第一共享指数值。

[0071] 在一个实施例中,确定所述第二共享指数值包括:确定多个浮点值的所述第二子集中的、具有最高值的浮点值的指数值,以及使用所述指数值作为所述第二共享指数值。

[0072] 在一个实施例中,确定所述第三共享指数值包括:在确定所述第一共享指数值大于或等于所述第二共享指数值时,使用所述第一共享指数值作为所述第三指数值,以及在确定所述第一共享指数值小于所述第二共享指数值时,使用所述第二共享指数值作为所述第三指数值。

[0073] 在一个实施例中,被用于表示被存储在针对所述共享指数浮点数据类型的所述数据结构中的所述尾数值的比特的数目小于被用于表示所述第三指数值的比特的数目。

[0074] 以上描述说明了本公开的各种实施例、以及如何实现特定实施例的各方面的示例。上述示例不应当被认为是唯一的实施例,并且被呈现以说明如以下权利要求所限定的特定实施例的灵活性和优点。基于以上公开和以下权利要求,在不脱离权利要求所限定的本公开的范围的情况下,可以采用其他布置、实施例、实现和等同方案。

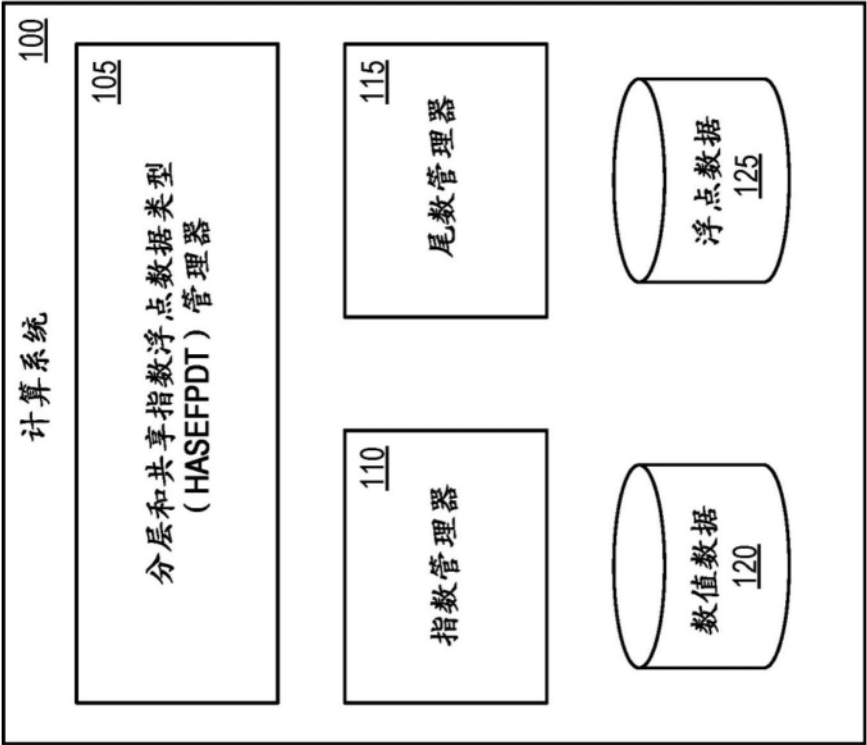


图1

200

205	值	0.1946	0.1535	0.0904	0.4727
210	原始指数	-3	-3	-4	-2
215	0级指数	-3	-3	-4	-2
220	1级共享指数	-3			-2
225	1级标度	0	0	1	0
230	2级共享指数	-2			
235	2级标度	1			0
240	总标度	1	1	1	0
245	共享指数标度	-3	-3	-3	-2
250	指数差	0	0	1	0
255	量化的1位幅度	1	1	0	1
260	量化指数	-3	-3	-3	-2
265	量化值	0.125	0.125	0	0.25

图2

300



尾数 (1位符号, 1位幅度)	1级标度	2级标度	共享指数
2位	1位	1位	8位
2位	1位		
2位	1位	1位	
2位	1位		

图3

400

405	值	0.1946	0.1535	0.0904	0.4727
410	原始指数	-3	-3	-4	-2
415	0级共享指数	-3	-3	-2	
420	1级共享指数	-2			
425	1级标度	1	1	0	0
430	总标度	1	1	0	0
435	共享指数标度	-3	-3	-2	-2
440	指数差	0	0	2	0
445	量化的2位幅度	3	2	0	3
450	量化指数	-3	-3	-2	-2
455	量化值	0.1875	0.125	0	0.375

图4

500



尾数 (1位符号, 1位幅度)		1级标度	共享指数
2位	2位	2位	8位
2位	2位	2位	
2位	2位	2位	

图5

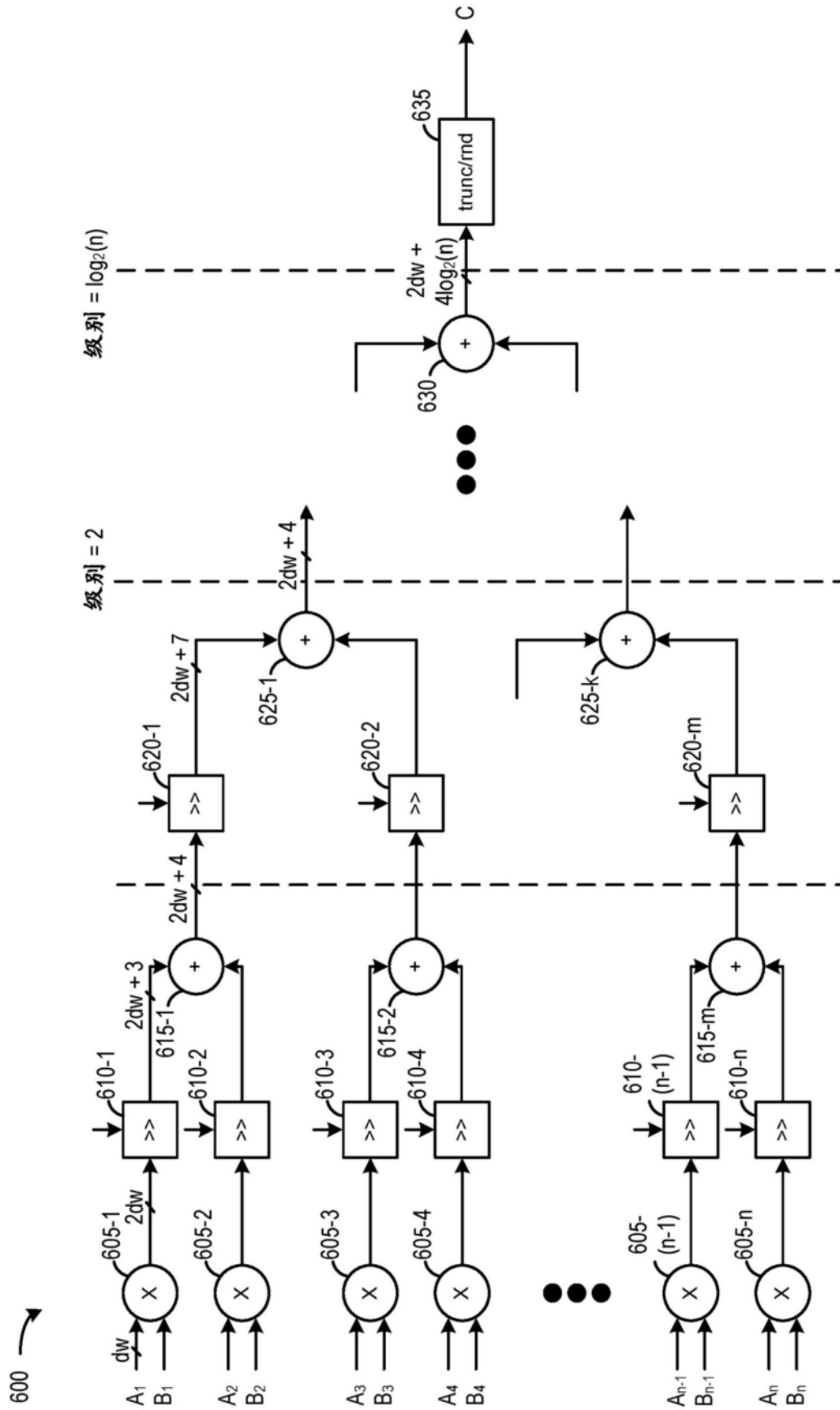


图6

700

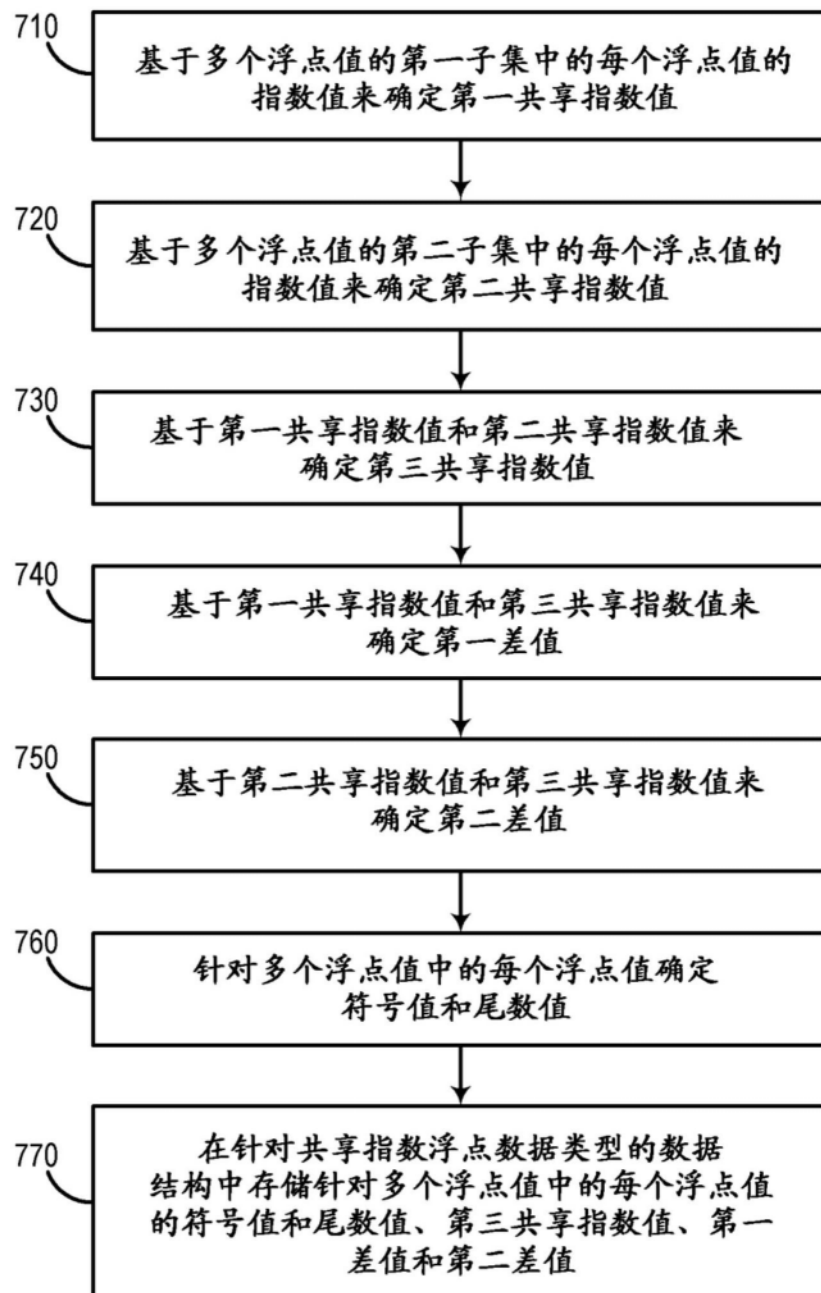


图7

800 →

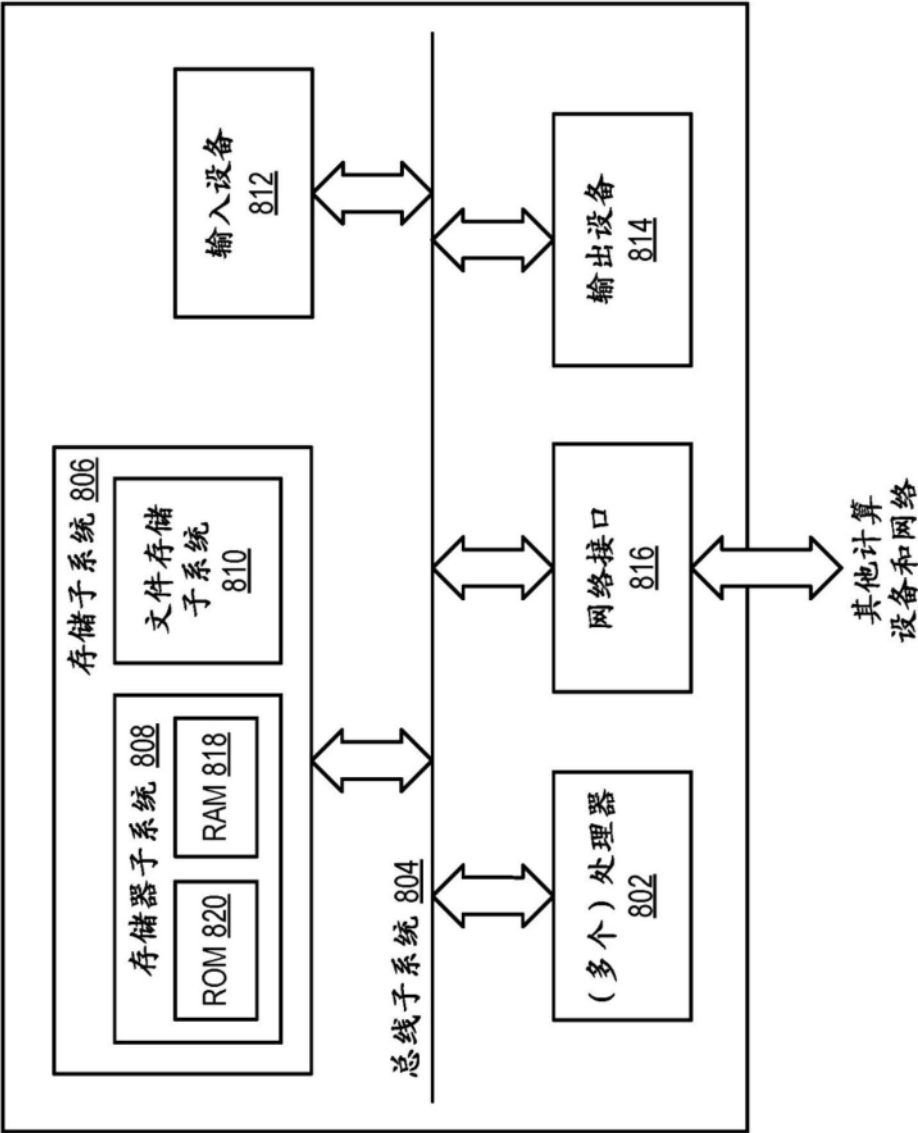


图8

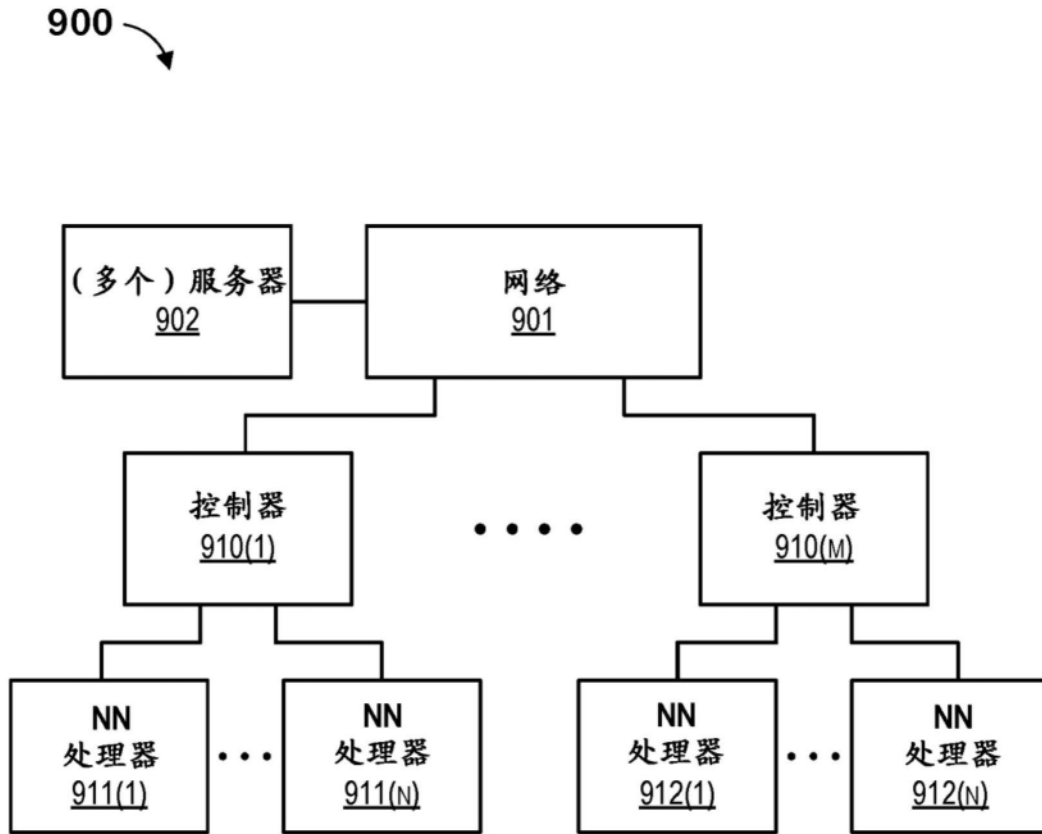


图9