



(12) 发明专利申请

(10) 申请公布号 CN 112602090 A

(43) 申请公布日 2021.04.02

(21) 申请号 201980056102.6

(22) 申请日 2019.06.19

(30) 优先权数据

62/693,237 2018.07.02 US

(85) PCT国际申请进入国家阶段日

2021.02.25

(86) PCT国际申请的申请数据

PCT/US2019/037952 2019.06.19

(87) PCT国际申请的公布数据

W02020/009800 EN 2020.01.09

(71) 申请人 奇跃公司

地址 美国佛罗里达州

(72) 发明人 G·韦迪希

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 杨晓光 于静

(51) Int.Cl.

G06K 9/00 (2006.01)

G06K 9/62 (2006.01)

G06T 19/00 (2011.01)

G06T 13/40 (2011.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

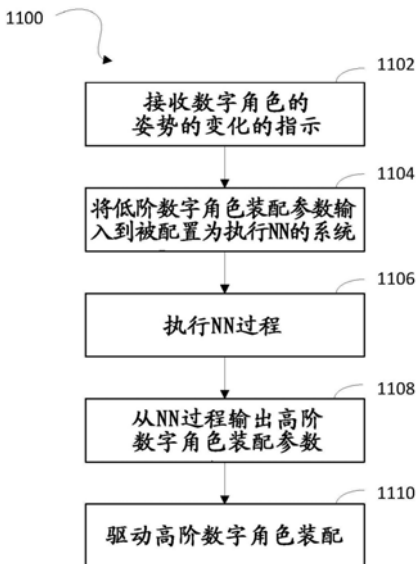
权利要求书2页 说明书44页 附图23页

(54) 发明名称

用于插值不同输入的方法和系统

(57) 摘要

提供用于对不同输入进行插值的系统和方法。径向基函数神经网络(RBFNN)可以用于对数字角色的姿势进行插值。RBFNN的输入参数可以按数据类型(例如,角度相对于线性)来分离,并且可以通过特定于数据类型的距离函数(例如,使用角度距离函数用于角度输入数据)在RBFNN内进行操作。可以将权重应用于每个距离,以补偿表示不同变量(例如,锁骨相对于肩部)的输入数据。RBFNN的输出参数可以是一组独立值,其可以组合成组合值(例如,表示S0(3)空间中的x、y、z、w角度值)。



1. 一种方法,包括:
接收包括至少一个角度分量和至少一个线性分量的输入数据;
将所述输入数据作为输入提供给至少一个神经网络NN,所述至少一个神经网络NN已经被训练为不同地评估所述至少一个角度分量和所述至少一个线性分量;以及
接收由所述至少一个NN基于对所述至少一个角度分量和所述至少一个线性分量的不同评估而生成的输出数据。
2. 根据权利要求1所述的方法,其中,所述至少一个角度分量处于三维(SO(3))空间中的特殊正交群。
3. 根据权利要求1所述的方法,其中,所述神经网络中的至少一个是前馈神经网络FFNN。
4. 根据权利要求3所述的方法,其中,至少一个FFNN是全连接网络。
5. 根据权利要求4所述的方法,其中,所述FFNN包括单个隐藏层。
6. 根据权利要求5所述的方法,其中,所述FFNN包括修正线性单元激活函数。
7. 根据权利要求6所述的方法,其中,所述隐藏层是残差NN块。
8. 根据权利要求1所述的方法,其中,所述NN中的至少一个是径向基函数神经网络RBFNN。
9. 根据权利要求1所述的方法,其中,所述输入数据描述数字角色的姿势。
10. 根据权利要求1所述的方法,其中,所述输入数据表示数字角色的低阶骨架,并且所述输出数据表示数字角色的高阶骨架。
11. 根据权利要求1所述的方法,其中,所述输出数据描述数字角色的姿势。
12. 根据权利要求1所述的方法,其中,所述输入数据和所述输出数据中的一个或多个还包括第三分量。
13. 根据权利要求12所述的方法,其中,所述角度分量、所述线性分量、以及所述第三分量分别是不同的运动分量。
14. 根据权利要求13所述的方法,其中:
所述至少一个角度分量描述旋转运动;
所述至少一个线性分量描述平移运动;以及
所述第三分量描述缩放。
15. 根据权利要求12所述的方法,其中,所述至少一个NN包括评估所述至少一个角度分量的第一FFNN和评估所述至少一个线性分量的第二FFNN。
16. 根据权利要求15所述的方法,其中,所述至少一个NN包括评估所述第三分量的第三FFNN。
17. 根据权利要求1所述的方法,其中,所述至少一个NN包括多个样本节点,其中,每个样本节点对应于训练姿势,并且其中,所述训练姿势中的至少一个包括至少一个角度分量和至少一个线性分量。
18. 根据权利要求1所述的方法,其中,所述至少一个NN通过评估处于三维SO(3)空间中的特殊正交群的所述至少一个角度分量以及通过利用欧氏距离公式评估所述至少一个线性分量,来不同地评估所述至少一个角度分量和所述至少一个线性分量。
19. 根据权利要求1所述的方法,还包括:

训练一个或多个NN,其中,所述训练包括:

接收包括训练输入数据和训练输出数据的训练数据,其中,所述训练输入数据和所述训练输出数据表示一个或多个训练姿势,其中,所述一个或多个训练姿势中的至少一个包括输入角度分量、输入线性分量、输出角度分量、以及输出线性分量;

将来自一个或多个姿势的每个姿势的所述输入角度分量分组为输入角度分量组;

将来自一个或多个姿势中的每个姿势的所述输入线性分量分组为输入线性分量组;以及

提供所述训练输入数据作为输入以训练所述至少一个NN,其中,所述输入角度分量组和所述输入线性分量组被不同地评估,其中,所述评估导致所述输出角度分量和所述输出线性分量。

20. 一种系统,包括:

至少一个处理器,其执行至少一个神经网络NN;以及

存储器,其被通信耦合到所述至少一个处理器,所述存储器存储指令,当所述指令在由所述至少一个处理器执行时使所述至少一个处理器执行包括以下的操作:

接收包括至少一个角度分量和至少一个线性分量的输入数据;

将所述输入数据作为输入提供给至少一个神经网络NN,所述至少一个神经网络NN已经被训练为不同地评估所述至少一个角度分量和所述至少一个线性分量;以及

接收由所述至少一个NN基于对所述至少一个角度分量和所述至少一个线性分量的不同评估而生成的输出数据。

用于插值不同输入的方法和系统

[0001] 相关申请的交叉引用

[0002] 本申请要求于2018年7月2日提交的题为“METHODS AND SYSTEMS FOR INTERPOLATION OF DISPARATE INPUTS (用于插值不同输入的方法和系统)”的第62/693,237号美国专利申请的优先权,该申请通过引用整体并入在此。

技术领域

[0003] 本公开涉及虚拟现实和增强现实,包括混合现实、成像和可视化系统,并且更具体地涉及用于使虚拟角色(诸如,化身或其他数字角色)动画化的装配(rigging)系统和方法。

背景技术

[0004] 现代计算和显示技术已经促进了所谓的“虚拟现实”、“增强现实”和“混合现实”体验的系统的发展,其中数字再现的图像以一种它们似乎是或可能被感知是真实的方式呈现给用户。虚拟现实(VR)场景通常涉及计算机生成的虚拟图像信息的呈现,而不对其他实际现实世界的视觉输入透明。增强现实(AR)场景通常涉及虚拟图像信息的呈现,作为对于用户周围的实际世界的可视化的增强。混合现实(MR)是一种增强现实,其中物理和虚拟对象可以共存并实时交互。本文公开的系统和方法解决了与VR、AR和MR技术相关的各种挑战。

发明内容

[0005] 执行各种姿势的主体的成像扫描可以用来创建对于该主体的栩栩如生的三维(3D)的虚拟表示(例如,化身或数字角色)。扫描通常是在主体执行训练姿势时进行的(例如,显示主体的不同的面部表情和/或动作的快照,诸如微笑、皱眉、眨眼、举起胳膊、走路等,其中每个姿势都由一组数字角色装配(rig)参数表示)。扫描可以用来使主体的虚拟化身动画化。虚拟化身只是数字角色的一个示例。这些术语可以在本公开中互换使用,但不应以任何方式限制本发明的范围。姿势的每个扫描都需要时间和精力来建立和捕获图像,并由数据表示。由于资源限制(例如,时间、人力、模型等)或数据存储限制,通常不可能捕获到数字角色可以执行的每个可能姿势。替代地,捕获表示整个姿势空间(例如,包含数字角色可以假设的所有可能的姿势的体积)的一组样本或训练的姿势,并使用插值来确定准确的角色姿势。

[0006] 用于对数字角色的姿势进行插值的方法和系统的各种实施例可以基于对表示数字角色的姿势的线性参数和角度参数进行分离。在用于插值引擎的输入姿势数据被分离成线性参数和角度参数时,可以开始这种分离。在一些实施例中,当在插值引擎内,使用更适合于线性空间的处理(例如,数学操纵)来操纵线性参数,并且使用更适合于角度空间的处理来操纵角度参数时,分离继续。在一些实施例中,可以在插值引擎内应用权重,使得能够更精确地组合不同类型(例如,锁骨类型、肩部类型等)的数据。插值引擎可以输出表示线性值和角度值两者的多个独立参数。然后,角度值可以被组合,以完全表示姿势的唯一角度值,这可以用于使数字角色装配动画化。

[0007] 该方法和系统的实施例可以使用径向基函数神经网络 (RBFNN) 来对数字角色的姿势进行插值。RBFNN的实现在由两者均独立的线性分量和3D角度组成的空间上使用多个距离度量,其中的线性分量服从于欧氏距离度量,其中的3D角度的距离在S0(3)中直接计算。线性距离和角度距离分开计算,然后利用一组权重组合为单个整体距离,然后将其用于插值。类似地,本文公开的RBFNN实施例利用了如下的一种方法来计算输出:将所有输出指定为独立的值,然后将角度值组合成一组完全定义的角度值。在各种实施例中,本文描述的方法和系统可以用于通过利用神经网络(诸如RBFNN、前馈神经网络 (FFNN)、或任何其他合适的神经网络 (NN)) 对包括不同类型的数据(例如,锁骨数据或肩部数据)的任何数据组进行插值。在一些实施例中,插值引擎的输出为使其可以容易地从编写引擎转移到实时引擎。

[0008] 在一些实施例中,公开了一种方法,包括:将输入参数接收到RBFNN中,其中输入参数包括第一数据类型和第二数据类型,其中RBFNN包括:与第一数据类型对应的第一距离函数和与第二数据类型对应的第二距离函数。

[0009] 在一些实施例中,公开了一种方法,包括:将表示第一输入类型的训练数据与表示第二输入类型的训练数据分离,定义对于第一输入类型的第一距离函数,定义对于第二输入类型的第二距离函数,并且将表示第一输入类型的训练数据和表示第二输入类型的训练数据输入到RBFNN,其中RBFNN包括第一距离函数和第二距离函数。

[0010] 在一些实施例中,公开了一种方法,包括:从RBFNN接收表示一组独立输出值的数据作为输出,从表示一组独立输出值的数据中分离表示角度值的数据,将表示角度值的数据组合为包括x、y、z和w值的四元数向量,并向被配置为驱动旋转的系统提供该组四元数向量。

[0011] 在一些实施例中,一种系统包括一个或多个计算机和一个或多个存储可执行指令的存储设备,当一个或多个计算机执行这些指令时,使一个或多个计算机执行包括以下的操作:接收来自低阶数字角色装配的姿势,将其作为输入提供给RBFNN,从RBFNN接收表示高阶数字角色装配的独立值的数据组作为输出,并且根据来自独立值的数据组的表示角度值的数据子集来定义一个或多个四分量的四元数向量。

[0012] 在一些实施例中,只要将角度分量和线性分量分离,任何合适的神经网络 (NN) 都可以用作插值引擎。在一些实施例中,可以使用前馈神经网络 (FFNN),例如具有单个隐藏层的全连接网络,使用修正线性单元 (ReLU) 激活函数。在一些实施例中,该隐藏层可以被合并为残差神经网络 (resnet) 块。

附图说明

[0013] 说明书中描述的主题的一种或多种实施方式的细节在附图和以下描述中阐述。根据说明书、附图和权利要求书,其它特征、方面和优点将变得显而易见。

[0014] 图1描绘了具有某些虚拟现实对象以及人观看的某些物理对象的混合现实场景的图示。

[0015] 图2示意性地示出可穿戴系统的示例。

[0016] 图3示意性地示出可穿戴系统的示例组件。

[0017] 图4示意性地示出用于将图像信息输出给用户的可穿戴设备的波导堆叠的示例。

[0018] 图5是用于与虚拟用户界面进行交互的方法的示例的过程流程图。

- [0019] 图6A是可包括化身处理和渲染系统的可穿戴系统的另一个示例的框图。
- [0020] 图6B示出化身处理和渲染系统的示例组件。
- [0021] 图7是包括进入可穿戴系统的各种输入的可穿戴系统的示例的框图。
- [0022] 图8是关于所识别的对象渲染虚拟内容的方法的示例的过程流程图。
- [0023] 图9A示意性地示出描绘彼此交互的多个可穿戴系统的整体系统视图。
- [0024] 图9B示出示例远程呈现会话。
- [0025] 图10示出可穿戴系统的用户所感知的化身的示例。
- [0026] 图11示出了根据一些实施例驱动数字角色装配的过程的示例概述。
- [0027] 图12A示出了根据一些实施例的具有装配参数的示例低阶数字角色装配。
- [0028] 图12B示出了根据一些实施例的具有装配参数的示例高阶数字角色装配。
- [0029] 图13A示出了根据一些实施例的低阶数字角色的示例姿势。
- [0030] 图13B示出了根据一些实施例的用于数字角色的中立姿势的示例。
- [0031] 图14示出了根据一些实施例的在RBFNN内的每个样本节点内所包含的各种组件。
- [0032] 图15A至图15C示出了根据一些实施例三个潜在的训练姿势,这些姿势可以表示在人弯曲其手臂的一系列姿势中的各种姿势。
- [0033] 图16示出了根据一些实施例的用于根据RBFNN对数字角色的输出姿势进行插值的系统图,该示例性数据在系统内进行通信。
- [0034] 图17示出了根据一些实施例的用于训练RBFNN的示例过程。
- [0035] 图18示出了根据一些实施例的使用经过训练的RBFNN来插值数字角色的新姿势的示例过程。
- [0036] 图19示出了根据本发明的一些实施例的用于插值来自FFNN的数字角色的输出姿势的系统示意图。
- [0037] 图20示出了根据本发明的至少一些实施例的用于训练NN的示例过程。
- [0038] 图21示出了根据本发明的至少一些实施例的用于使用NN的示例过程。
- [0039] 贯穿附图,可以重复使用附图标记来表示参考元件之间的对应关系。提供这些附图是为了说明本文所描述的示例实施例,并且并不旨在限制本公开的范围。

具体实施方式

[0040] 概述

[0041] 虚拟化身可以是真实或虚构的人(或生物或拟人化的物体)在AR/VR/MR环境中的虚拟表示。例如,在两个AR/VR/MR用户彼此交互的呈现会话期间,观看者可以在观看者的环境中感知到另一用户的化身,从而在观看者的环境中创造出另一用户存在的有形感觉。化身还可以为用户提供彼此交互并在共享的虚拟环境中一起做事情的方式。例如,参加在线课程的学生可以感知虚拟教室里其他学生或老师的化身,并与之互动。再例如,在AR/VR/MR环境中玩游戏的用户可以看到游戏中其他玩家的化身并与之交互。

[0042] 所公开的系统和方法的实施例可提供改进的化身以及可穿戴系统的用户与用户的环境中的化身之间更真实的交互。尽管本公开中的示例描述了使人形化身动画化,类似的技术也可以应用于其他数字角色,诸如动物、虚构的生物、物体等。因此,本文描述的成像扫描和对准或稳定技术的主体可以是人、动物或其他可变形的物体或角色。在本文描述的

许多示例中,主体是人(例如,以创建人形化身或数字角色),但这只是为了说明,并不限制所公开的方法和系统。此外,这里描述的许多技术可以对动物主体特别有利,因为动物可以具有有限的注意力跨度,相对于动物可以呈现的潜在姿势的数量,只有相对较少量的姿势可以被捕获到,以表示整个姿势空间。

[0043] 要创造令人信服的人的数字化身,可能需要扫描该人的三维(3D)相似度来创建数字几何结构。例如,可以通过将主体置于包括围绕并指向该主体的多个相机(例如,两个或更多相机)的摄影测量捕获台来进行成像扫描。这些相机可以彼此同步来捕获图像,这些图像可以转换成执行姿势的主体的3D扫描。

[0044] 根据静态突发(burst)(基本上同时捕获的多个图像),可以生成单个3D扫描。然而,对于可以移动并且表现感情的化身,可以在主体执行训练姿势时进行扫描,该训练姿势包括基本表达单元(例如,微笑、皱眉等)和预定义的静态姿势,其中一些训练姿势可以表示动画序列内的静态姿势(例如,由五个静态姿势表示的手臂抬起——手臂笔直向下(straight down)、手臂向上抬起(up)45度、手臂笔直向外(straight out)90度、手臂向上抬起135度、手臂向上抬起180度)。这些扫描可以用来创建训练组,该训练组用于构建动画装配(如下所述),其可以用来创建动作中的人的令人信服的数字化身。

[0045] 在这个过程中可能遇到的一个困难是需要大量的数据来准确描述数字角色的整个姿势空间。大量的数据可能需要大量的计算资源来在切实可行的时间量内分析数据(例如,对一些应用来说是实时的)。在一些实施例,这个问题可以通过如下方式来解决:利用跨越整个姿势空间但是包括代表性姿势的较小子集的姿势(例如,训练姿势或样本姿势),然后当数字角色需要假设中间姿势时在代表性训练姿势之间进行插值。为了执行插值,可以利用神经网络,例如FFNN或RBFNN。因此,各种此类实施例可以有利地提高计算速度或利用较少的计算资源来计算数字角色的姿势。

[0046] 径向基函数神经网络(RBFNN)是一种常用的插值方式。RBFNN广泛应用于许多学科(地质学、海洋学等),因此很容易理解,并且已经开发了许多优化,使它们执行得更快、更准确。RBFNN是解决诸如高阶数字角色姿势的复杂系统的相对简单的方式。

[0047] 然而,利用一个或多个NN(诸如RBFNN)对于数字角色的新姿势进行插值表现出其自身的一系列挑战。例如,当RBFNN用来对数字角色的姿势进行插值时,RBFNN将会周期性地错误地计算姿势。例如,RBFNN可能会引入抖动(例如,数字角色网格中的单个或一组顶点的地点或位置的小且频繁的变化)、身体部位在错误的地方的弹出(例如,数字角色网格中的单个或一组顶点的地点或位置的突然的大变化),混合形状(blendshape)迅速打开/关闭等等。换句话说,RBFNN在大多数情况下都能很好地工作,但会周期性且非常明显地输出不正确的姿势,给人一种不稳定和不可预测的系统的印象。修复这些周期性错误需要人工干预,从而降低了系统的整体效率。

[0048] 这些周期性错误至少部分地是由三个主要原因引起的。首先,RBFNN通常将距离计算为欧氏距离。然而,使数字角色动画化主要是关于旋转骨骼——平移和旋转数字角色的骨架,以确定覆盖骨架的网格的正确相应变形和其他高阶参数。其次,来自RBFNN的输出是单值的。然而,骨架运动的角度值可能需要用四个元素来表示(例如x、y、z、w)。第三,RBFNN可以对所有输入进行相同的处理。虽然输入数据表示单个数字角色,但数字角色装配可以有几种不同类型的数据。所公开的系统和技术的实施例可以用于通过解决这些问题来自动

确定数字角色的稳定且可预测的中间姿势。

[0049] 第一个问题,角度值在线性空间中的处理可以导致不正确的RBFNN距离计算,因为角度值将被定义为线性分析的函数。可能会发生诸如万向节锁定、快速翻转、符号和/或区域的改变的行为,其中区域是单个角度旋转可以用多种方式表示的位置。这是因为,不仅地点/位置很重要,取向也很重要。考虑下面的示例。化身的目标是在开始时其手臂放在侧面,在结束时其手臂笔直在其前面。化身开始于站立位置,右臂下垂在其侧面,手掌平放在臀部。化身将其手臂伸直在侧面,然后将整个手臂向身体前方旋转,手掌朝下(例如,在整个旋转过程中手掌面向地面)。现在,化身仍然站着,开始时左臂下垂在其侧面,手掌平放在臀部(例如,相同的起始位置)。化身将手臂笔直地伸向前方。注意化身的两个手臂都是在其前方笔直的(例如,具有相同的目标结束姿势),但是化身的右手掌心是向下的,而化身的左手掌心是向右的。在线性空间中描述旋转(诸如,如上所述的手臂旋转)的数字角色装配可能以两个手臂取向中的任一个结束,而在角度空间中描述的旋转只能有一个取向。

[0050] 旋转可以用许多不同的方式恰当地(例如,充分地)描述。例如,欧拉、三维的特殊正交群($SO(3)$),也称为旋转群 $SO(3)$ 、旋转矩阵、轴-角、四元数表示是角度空间的几个示例。为了便于描述,后续关于旋转的讨论将在 $SO(3)$ 空间中进行描述,但可以使用任何合适的方法。 $SO(3)$ 空间是由四个值定义的——由 x 、 y 、 z 定义的轴,以及绕轴的旋转 w 。轴可以位于具有绕该轴的旋转(例如,“ w ”)的空间中的任何位置,从而实现 $SO(3)$ 空间中数字角色的所有转换或操作的唯一或单值的解。

[0051] 第二个问题,角度值的单值输出是不清楚的,可能有针对单个期望结束姿势所计算的几个解/输出/取向。为了解决这个问题,可以为移动数字角色装配所需要的每个输出旋转计算四个独立的值。然后,四个独立的输出值被组合成组合值(x, y, z, w),该组合值可以作为旋转直接提供给数字角色装配。

[0052] 第三个问题,组合不同类型的旋转参数,可以导致较小的移动相对于较大的移动变得微不足道。不同类型的旋转参数可以是锁骨型、肩关节型、指骨型、跖骨型等。示例参数类型的列表仅用于说明,可以使用其他类型或组。每个组可以包括一个装配骨或多个装配骨。每个组都被分配了权重,该权重表示其对数字角色的整体运动的相对贡献,并能够平衡各个参数的重要性。例如,锁骨的权重可以更大,因此它的小运动并不比相对较大的肩部运动小(dwarf)。通过确保装配的所有部分都被考虑,即使它对数字角色的运动有相对较小的影响,平衡旋转之间的权重也可以减少平滑插值所需的样本总数。这些权重可以由用户定义,或者至少部分由用户定义。权重可以在几个不同的装配上重复使用,或者根据需要改变,这取决于将如何使用装配或者角色的类型。例如,所有的人类化身可以使用完全相同的一组权重,但对于各种动物或拟人化的物体可以进行微调。

[0053] 本文描述的系统和方法的另一益处是,将复杂的系统抽象成可序列化的数据,这些数据可以从编写(例如,直接内容创建(DCC))应用转移到实时引擎中。在NN的训练阶段期间,使用一组特定的参数来描述数字角色的姿势。线性输出由单值定义,旋转输出由一组4个值定义(如 $SO(3)$ 空间)。相同的装配设置可以用于所有角色,从而创建可用于驱动广泛的数字角色的标准化输出组。典型的装配系统建立在编写应用中,该编写应用具有用于建立数字角色装配的自己唯一的一组程序和系统。如果需要将装配转移到实时运行的引擎上,一些DCC应用功能会在转换过程中丢失,转移后的装配不会与最初所做的完全相同地来执

行。

[0054] 可穿戴系统的3D显示的示例

[0055] 可穿戴系统(在此也称为增强现实(AR)系统)可以被配置为向用户呈现2D或3D虚拟图像。图像可以是静止图像、视频的帧或视频的组合等。可穿戴系统的至少一部分可以在可穿戴设备上实现,该可穿戴设备可以针对用于交互单独或组合地呈现VR、AR或MR环境。可穿戴设备可以作为AR设备(ARD)被互换地使用。此外,出于本公开的目的,术语“AR”与术语“MR”可被互换地使用。

[0056] 图1描绘了具有某些虚拟现实对象以及由人观看的某些物理对象的混合现实场景的图示。在图1中,描绘了MR场景100,其中MR技术的用户看到了真实世界的公园状设置110,其以人、树、背景中的建筑物以及混凝土平台120为特征。除了这些项目之外,MR技术的用户还可以感知到他“看到”了站在真实世界平台120上的机器人雕像130,以及似乎是大黄蜂的拟人化的飞行着的卡通般化身角色140,尽管这些元素在真实世界中不存在。

[0057] 为了使3D显示器产生真实的深度感,且更具体地,产生表面深度的模拟感,可能希望显示器视野中的每个点都生成与其虚拟深度相对应的调节响应。如果对显示点的调节响应不符合该点的虚拟深度(如由会聚和立体视觉的双目深度提示确定),则人眼可能会遇到调节冲突,从而导致成像不稳定,有害的眼睛疲劳,头痛,以及在缺乏调节信息的情况下的几乎完全的表面深度缺失。

[0058] VR、AR和MR体验可以通过具有显示器的显示系统来提供,在该显示器中,与多个深度平面相对应的图像被提供给观看者。对于每个深度平面,图像可能不同(例如,提供场景或对象的稍微不同的呈现),并且可能由观看者的眼睛分别聚焦,从而基于使位于不同深度平面上的场景的不同图像特征聚焦所需的眼睛调节,或基于观察在不同深度平面上失焦的不同图像特征,为用户提供深度提示。如在此其它地方所讨论的,这种深度提示提供了可靠的深度感知。

[0059] 图2示出可被配置为提供AR/VR/MR场景的可穿戴系统200的示例。可穿戴系统200还可称为AR系统200。可穿戴系统200包括显示器220以及支持显示器220的功能的各种机械和电子模块和系统。显示器220可以耦合至框架230,该框架230可以由用户、佩戴者或观看者210佩戴。显示器220可以放置于用户210的眼睛前面。显示器220可以向用户呈现AR/VR/MR内容。显示器220可以包括戴在用户的头上的头戴式显示器(HMD)。

[0060] 在一些实施例中,扬声器240被耦合至框架230并且邻近用户的耳道放置(在一些实施例中,未示出的另一个扬声器邻近用户的另一耳道放置以提供立体声/可塑的声音控制)。显示器220可包括用于检测来自环境的音频流并捕获环境声音的音频传感器(例如,麦克风)232。在一些实施例中,未示出的一个或多个其它音频传感器被放置以提供立体声接收。立体声接收可以用来确定声源的位置。可穿戴系统200可以对音频流执行声音或语音识别。

[0061] 可穿戴系统200可以包括面向外成像系统464(图4中所示),该面向外成像系统464观察用户周围环境中的世界。可穿戴系统200还可以包括可以跟踪用户的眼睛运动的面向内成像系统462(图4中所示)。面向内成像系统可以跟踪一只眼睛的运动或两只眼睛的运动。面向内成像系统462可以附接至框架230,并且可以与处理模块260或270电通信,该处理模块260或270可以处理由面向内成像系统获取的图像信息以确定用户210的例如眼睛的瞳

孔直径或取向的眼睛运动或眼睛姿势。面向内成像系统462可以包括一个或多个相机。例如,至少一个相机可以用于对每只眼睛成像。由相机获取的图像可用于分别确定每只眼睛的瞳孔大小或眼睛姿势,从而允许将图像信息呈现给每只眼睛以针对该只眼睛动态地调整。

[0062] 作为示例,可穿戴系统200可使用面向外成像系统464或面向内成像系统462来获取用户的姿势的图像。图像可以是静止图像、视频的帧,或视频。

[0063] 显示器220可以诸如通过有线引线或无线连接可操作地耦合250到本地数据处理模块260,该本地数据处理模块260可以以各种配置安装,诸如固定地附接至框架230,固定地附接至由用户佩戴的头盔或帽子,嵌入耳机中或以其它方式可移除地附接至用户210(例如,以背包式配置、皮带耦合式配置)。

[0064] 本地处理和数据模块260可以包括硬件处理器以及数字存储器,诸如非易失性存储器(例如,闪存),二者都可以用于辅助数据的处理、缓存和存储。数据可以包括如下数据:
a) 从传感器(其可以例如可操作地耦合至框架230或以其它方式附接至用户210)捕获,诸如图像捕获设备(例如,面向内成像系统或面向外成像系统中的相机)、音频传感器(例如,麦克风)、惯性测量单元(IMU)、加速度计、罗盘、全球定位系统(GPS)单元、无线电设备或陀螺仪;或b) 使用远程处理模块270或远程数据存储库280获取或处理,可能在这种处理或取得之后传递给显示器220。本地处理和数据模块260可以通过通信链路262或264(诸如经由有线或无线通信链路)可操作地耦合至远程处理模块270或远程数据存储库280,使得这些远程模块可用作本地处理和数据模块260的资源。此外,远程处理模块280和远程数据存储库280可以可操作地彼此耦合。

[0065] 在一些实施例中,远程处理模块270可以包括被配置为分析和处理数据或图像信息的一个或多个处理器。在一些实施例中,远程数据存储库280可以包括数字数据存储设施,其可以通过互联网或以“云”资源配置的其它联网配置来使用。在一些实施例中,所有数据被存储并且所有计算在本地处理和数据模块中执行,从而允许来自远程模块的完全自主的使用。

[0066] 可穿戴系统的示例组件

[0067] 图3示意性地示出可穿戴系统的示例组件。图3示出可穿戴系统200,其可包括显示器220和框架230。放大视图202示意性地示出可穿戴系统200的各个组件。在某些实现方式中,图3中所示的一个或多个组件可以是显示器220的一部分。单独或组合的各种组件可以收集与可穿戴系统200的用户或用户的环境相关联的各种数据(诸如例如,音频或视觉数据)。应当理解,取决于可穿戴系统所用于的应用,其它实施例可以具有更多或更少的组件。尽管如此,图3提供了可以通过可穿戴系统收集、分析和存储的数据的各种成分和类型中的一些的基本思想。

[0068] 图3示出示例可穿戴系统200,该示例可穿戴系统200可以包括显示器220。显示器220可以包括可以安装到用户的头部或与框架230相对应的壳体或框架230上的显示透镜226。显示透镜226可以包括一个或多个透明镜或衍射光学元件,该透明镜或衍射光学元件由壳体230定位在用户的眼睛302、304的前面,并且可以被配置为将投射的光338引导到眼睛302、304中并有助于光束成形,同时也允许来自本地环境的至少一些光的透射。投射光束338的波前可以被发散以与投射光的期望焦距一致。如图所示,两个广视场机器视觉相机

316(也称为世界相机)可以被耦合到壳体230以对用户周围的环境成像。这些相机316可以是双重捕获可见光/不可见(例如,红外)光相机。相机316可以是图4中所示的面向外成像系统464的一部分。由世界相机316获取的图像可以由姿势处理器336处理。例如,姿势处理器336可以实现一个或多个对象识别器708(例如,图7所示)以识别用户或在该用户环境中另一人的姿势,或识别该用户环境中的物理对象。

[0069] 继续参考图3,示出了一对光投影仪模块,其具有被配置为将光338引导到眼睛302、304中的显示光学器件和透镜。所描绘的视图还示出了与红外光(诸如发光二极管“LED”)配对的两个微型红外相机324,其被配置为能够跟踪用户的眼睛302、304以支持渲染和用户输入。相机324可以是图4中所示的面向内成像系统462的一部分。可穿戴系统200可进一步具有传感器组件339,该传感器组件可包括X、Y和Z轴加速度计功能以及磁罗盘以及X、Y和Z轴陀螺仪功能,优选以相对较高的频率(诸如200Hz)提供数据。传感器组件339可以是参考图2A描述的IMU的一部分。所描绘的系统200还可以包括头部姿势处理器336,诸如ASIC(专用集成电路)、FPGA(现场可编程门阵列)或ARM处理器(先进的精简指令集机器),该头部姿势处理器可被配置为根据从捕获设备316输出的宽视场图像信息来计算实时或接近实时的用户头部姿势。头部姿势处理器336可以是硬件处理器,并且可以实现为图2A中所示的本地处理和数据模块260的一部分。

[0070] 可穿戴系统还可包括一个或多个深度传感器234。深度传感器234可被配置为测量环境中的对象到可穿戴设备之间的距离。深度传感器234可包括激光扫描仪(例如,激光雷达)、超声深度传感器或深度感测相机。在某些实现方式中,在相机316具有深度感测能力的情况下,相机316也可以被认为是深度传感器234。

[0071] 还示出了处理器332,其被配置为执行数字或模拟处理以从来自传感器组件339的陀螺仪、罗盘或加速度计数据中得出姿势。处理器332可以是图2中所示的本地处理和数据模块260的一部分。如图3中所示的可穿戴系统200还可包括诸如例如GPS 337(全球定位系统)的定位系统,以辅助姿势和定位分析。另外,GPS可以进一步提供关于用户环境的基于远程(例如,基于云)的信息。该信息可用于识别用户环境中的对象或信息。

[0072] 可穿戴系统可以将由GPS 337获取的数据与可以提供有关用户环境的更多信息的远程计算系统(诸如例如,远程处理模块270、另一个用户的ARD等)相结合。作为一个示例,可穿戴系统可以基于GPS数据确定用户的位置并取得世界地图(例如,通过与远程处理模块270通信),该世界地图包括与用户的位置相关联的虚拟对象。作为另一个示例,可穿戴系统200可以使用世界相机316(其可以是图4中所示的面向外成像系统464的一部分)来监测环境。基于由世界相机316获取的图像,可穿戴系统200可以检测环境中的对象(例如,通过使用图7中所示的一个或多个对象识别器708)。可穿戴系统可以进一步使用由GPS 337获取的数据来解释特征。

[0073] 可穿戴系统200还可以包括渲染引擎334,该渲染引擎334可以被配置为提供用户本地的渲染信息,以促进扫描仪的操作和成像到用户的眼睛中,以便用户观看世界。渲染引擎334可以由硬件处理器(诸如例如,中央处理单元或图形处理单元)实现。在一些实施例中,渲染引擎是本地处理和数据模块260的一部分。渲染引擎334可以通信地耦合(例如,经由有线或无线链路)到可穿戴系统200的其它组件。例如,渲染引擎334可以经由通信链路274耦合到眼睛相机324,并且经由通信链路272耦合到投影子系统318(其可以以类似于视

网膜扫描显示器的方式经由扫描的激光装置将光投射到用户的眼睛302、304中)。渲染引擎334还可以分别经由链路276和294与其它处理单元(诸如例如传感器姿势处理器332和图像姿势处理器336)通信。

[0074] 相机324(例如,微型红外相机)可用于跟踪眼睛姿势以支持渲染和用户输入。一些示例眼睛姿势可以包括用户正在看的地方或他或她正在聚焦的深度(这可以通过眼睛的聚散来估计)。GPS 337、陀螺仪、罗盘和加速度计339可用于提供粗略或快速的姿势估计。相机316中的一个或多个可以获取图像和姿势,该图像和姿势可以与来自关联的云计算资源的数据一起被用来映射本地环境并与其他人共享用户视图。

[0075] 图3中描绘的示例组件仅用于说明目的。为了便于说明和描述,一起示出了多个传感器和其它功能模块。一些实施例可以仅包括这些传感器或模块中的一个或子集。此外,这些组件的位置不限于图3中所示的位置。一些组件可以安装到其它组件或容纳在其它组件(诸如皮带安装组件、手持组件或头盔组件)内。作为一个示例,图像姿势处理器336、传感器姿势处理器332和渲染引擎334可以放置在腰带包中,并被配置为经由无线通信(诸如超宽带、Wi-Fi、蓝牙等)或经由有线通信与可穿戴系统的其它组件通信。所描绘的壳体230优选地是用户可头戴的和可穿戴的。然而,可穿戴系统200的一些组件可能会穿戴到用户身体的其它部分。例如,扬声器240可以被插入用户的耳朵中以向用户提供声音。

[0076] 关于光338到用户的眼睛302、304中的投影,在一些实施例中,相机324可以被用来测量用户的眼睛的中心在几何上趋向的位置,该位置通常与眼睛的焦点位置或“焦点深度”重合。眼睛趋向的所有点的三维表面可以称为“两眼视界(horopter)”。焦距可以采取一定数量的深度,或者可以无限变化。从聚散距离投射的光似乎聚焦到主体眼睛302、304,而在聚散距离之前或之后的光则是模糊的。在美国专利公开No.2016/0270656中也描述了本公开的可穿戴设备和其它显示系统的示例,其全部内容通过引用合并于此。

[0077] 人的视觉系统是复杂的,并且提供对深度的真实感是具有挑战性的。由于聚散和调节的组合,对象的观看者可以将对象感知为三维的。两只眼睛相对于彼此的聚散运动(例如,用于会聚眼睛的视线以固定在对象上的瞳孔朝向或远离彼此的转动运动)与眼睛晶状体的聚焦(或“调节”)密切相关。在正常情况下,改变眼睛晶状体的焦点或调节眼睛,以在被称作“调节-聚散反射”的关系下将焦点从一个对象改变为在另一距离处的另一对象来自动将聚散变化匹配到相同距离。同样,在正常情况下,聚散的变化将触发调节的匹配变化。在调节和聚散之间提供更好匹配的显示系统可以形成更逼真的和更舒适的三维图像模拟。

[0078] 此外,人眼可以正确地分辨束直径小于约0.7毫米的空间相干光,而不管眼睛聚焦在哪里。因此,为了产生适当的焦深的错觉,可以用相机324跟踪眼睛的聚散,并且渲染引擎334和投影子系统318可用于渲染处于聚焦的聚散上或附近的所有对象,以及处于变化的分辨程度的所有其它对象。优选地,系统220以大约每秒60帧或更高的帧速率向用户渲染。如上所述,优选地,相机324可以用于眼睛跟踪,并且软件可以被配置为不仅拾取聚散几何形状而且拾取焦点位置提示以用作用户输入。优选地,此类显示系统配置有适合白天或晚上使用的亮度和对比度。

[0079] 在一些实施例中,显示系统优选地对于视觉对象对准具有小于约20毫秒的延迟,小于约0.1度的角度对准和约1弧分的分辨率,不受理论的限制,据信这大约是人眼的极限。显示系统220可以与定位系统集成在一起,该定位系统可以涉及GPS元件、光学跟踪、罗盘、

加速度计或其它数据源,以帮助位置和姿势确定;定位信息可用于促进用户观看相关世界时的准确渲染(例如,此类信息将有助于眼镜了解它们关于真实世界的位置)。

[0080] 在一些实施例中,可穿戴系统200被配置为基于用户的眼睛的调节来显示一个或多个虚拟图像。在一些实施例中,不同于迫使用户聚焦在图像被投影的地方的现有3D显示方法,可穿戴系统被配置为自动改变所投影虚拟内容的焦点,以允许更舒适地观看被呈现给用户的一个或多个图像。例如,如果用户的眼睛具有1m的当前焦点,则图像可以被投影为与用户的焦点重合。如果用户将焦点移至3m,则图像被投影为与新焦点重合。因此,一些实施例的可穿戴系统200允许用户的眼睛以更自然的方式起作用而不是迫使用户达到预定的焦点。

[0081] 这样的可穿戴系统200可以消除或减少通常关于虚拟现实设备观察到的眼睛疲劳、头痛和其它生理症状的发生。为了实现这一点,可穿戴系统200的各种实施例被配置为通过一个或多个可变聚焦元件(VFE)以变化的焦距投影虚拟图像。在一个或多个实施例中,可通过多平面聚焦系统来实现3D感知,该多平面聚焦系统在远离用户的固定焦平面处投影图像。其它实施例采用可变平面焦点,其中,焦平面在z方向前后移动,以与用户当前的焦点状态一致。

[0082] 在多平面聚焦系统和可变平面聚焦系统二者中,可穿戴系统200可以采用眼睛跟踪来确定用户的眼睛的聚散,确定用户的当前焦点以及将虚拟图像投影在所确定的焦点处。在其它实施例中,可穿戴系统200包括光调制器,该光调制器通过光纤扫描仪或其它光发生源将可变焦点的光束以光栅图案可变地投射到整个视网膜上。因此,如在美国专利No.2016/0270656(其全部内容通过引用合并于此)中进一步描述的,可穿戴系统200的显示器以变化的焦距投影图像的能力不仅减轻了用户观看3D对象的调焦,而且还可以用于补偿用户的眼部异常。在一些其它实施例中,空间光调制器可以通过各种光学组件将图像投影到用户。例如,如下文进一步所述,空间光调制器可将图像投影到一个或多个波导上,然后将图像传输给用户。

[0083] 波导堆叠组件

[0084] 图4示出用于向用户输出图像信息的波导堆叠的示例。可穿戴系统400包括波导的堆叠,或堆叠的波导组件480,其可以用于使用多个波导432b、434b、436b、438b、4400b向眼睛/大脑提供三维感知。在一些实施例中,可穿戴系统400可以对应于图2的可穿戴系统200,图4示意性地更详细地示出该可穿戴系统200的一些部分。例如,在一些实施例中,波导组件480可以被集成到图2的显示器220中。

[0085] 继续参考图4,波导组件480还可在波导之间包括多个特征458、456、454、452。在一些实施例中,特征458、456、454、452可以是透镜。在其它实施例中,特征458、456、454、452可以不是透镜。而是,它们可以仅仅是间隔物(例如,覆盖层或用于形成气隙的结构)。

[0086] 波导432b、434b、436b、438b、440b或多个透镜458、456、454、452可以被配置为以各种水平的波前曲率或光线发散将图像信息发送到眼睛。每个波导级可以与特定的深度平面相关联,并且可以被配置为输出与该深度平面相对应的图像信息。图像注入设备420、422、424、426、428可以被用来将图像信息注入波导440b、438b、436b、434b、432b,波导中的每个波导可以被配置为将入射光横跨每个相应的波导分布,用于朝向眼睛410输出。光离开图像注入设备420、422、424、426、428的输出表面,并且被注入到波导440b、438b、436b、434b、

432b的对应输入边缘中。在一些实施例中,单个光束(例如,准直光束)可以被注入到每个波导中,以输出样本出射光瞳束的整个场,该样本出射光瞳束以对应于与特定波导相关联的深度平面的特定角度(和发散量)导向眼睛410。

[0087] 在一些实施例中,图像注入设备420、422、424、426、428是分立的显示器,其各自产生图像信息以分别注入到对应的波导440b、438b、436b、434b、432b中。在一些其它实施例中,图像注入设备420、422、424、426、428是单个多路复用显示器的输出端,该显示器可以例如经由一个或多个光导管(诸如光缆)将图像信息传送到图像注入设备420、422、424、426、428中的每一个。

[0088] 控制器460控制显示器和图像注入设备420、422、424、426、428的操作。控制器460包括调节定时和向波导440b、438b、436b、434b、432b提供图像信息的编程(例如,非暂态计算机可读介质中的指令)。在一些实施例中,控制器460可以是单个集成设备,或者是通过有线或无线通信信道连接的分布式系统。在一些实施例中,控制器460可以是处理模块260或270(在图2中示出)的一部分。

[0089] 波导440b、438b、436b、434b、432b可以被配置为通过全内反射(TIR)在每个相应波导内传播光。波导440b、438b、436b、434b、432b可各自是平面的或具有另一种形状(例如,弯曲的),具有主要的顶部表面和底部表面以及在那些主要的顶部表面和底部表面之间延伸的边缘。在所示的配置中,波导440b、438b、436b、434b、432b可各自包括光学元件440a、438a、436a、434a、432a,其被配置为通过衍射或其它方式重定在每个相应波导内传播的光而将光从波导中耦合出。出耦合光由波导在波导中传播的光撞击光重定向元件(例如,衍射光栅)的位置处输出。光学元件(440a、438a、436a、434a、432a)可以例如是反射或衍射光学特征。尽管为了便于描述和清楚起见而示出设置在波导440b、438b、436b、434b、432b的底部主表面上,但是在一些实施例中,光学元件440a、438a、436a、434a、432a可以设置在顶部主表面或底部主表面,或者直接设置在波导440b、438b、436b、434b、432b的体积中。在一些实施例中,光学元件440a、438a、436a、434a、432a可以形成为材料层,该材料层附接至透明基板上以形成波导440b、438b、436b、434b、432b。在一些其它实施例中,波导440b、438b、436b、434b、432b可以是单片材料,并且光学元件440a、438a、436a、434a、432a可以形成在该片材料的表面上或内部中。

[0090] 继续参考图4,如在此讨论的,在一些实施例中,每个波导440b、438b、436b、434b、432b被配置为输出光以形成与特定深度平面相对应的图像。例如,最靠近眼睛的波导432b可以被配置为将注入到这种波导432b中的准直光传送到眼睛410。准直光可以代表光学无限远焦平面。下一个波导向上434b可以被配置为在通过第一透镜452(例如,负透镜)的准直光到达眼睛410之前发出该准直光。第一透镜452可以被配置为产生轻微的凸波前曲率,使得眼睛/大脑将来自下一个波导向上434b的光解释为来自光学无限远朝内更加靠近眼睛410的第一焦平面。类似地,第三向上波导436b在到达眼睛410之前将其输出光通过第一透镜452和第二透镜454二者。第一透镜452和第二透镜454的组合光焦度可以配置为产生另一增量的波前曲率,使得眼睛/大脑将来自第三波导436b的光解释为来自第二焦平面,该第二焦平面比来自下一个波导向上434b的光从光学无限远朝内更加靠近人。

[0091] 其它波导层(例如,波导438b、440b)和透镜(例如,透镜456、458)被类似地配置,堆叠中的最高波导440b通过其与眼睛之间的所有透镜发送其输出,以用于代表相对于人的最

接近焦平面的聚合光焦度。为了在观看/解释来自堆叠波导组件480另一侧上的世界470的光时补偿透镜458、456、454、452的堆叠,可以在堆叠的顶部处设置补偿透镜层430,以补偿下面的透镜堆叠458、456、454、452的总焦度。此类配置提供与可用的波导/透镜对一样多的感知焦平面。波导的光提取光学元件和透镜的聚焦方面二者都可以是静态的(例如,不是动态的或电激励的)。在一些替代实施例中,其一者或二者可以动态使用电激活特征。

[0092] 继续参考图4,光提取光学元件440a、438a、436a、434a、432a可被配置为既将光重定向到其相应的波导之外,又针对与波导相关联的特定深度平面以适当的发散或准直量输出该光。结果,具有不同相关联的深度平面的波导可以具有光提取光学元件的不同配置,该光提取光学元件取决于相关联的深度平面输出具有不同发散量的光。在一些实施例中,如在此所讨论的,光提取光学元件440a、438a、436a、434a、432a可以是体积或表面特征,该体积或表面特征可以被配置为以特定角度输出光。例如,光提取光学元件440a、438a、436a、434a、432a可以是体积全息图、表面全息图和/或衍射光栅。在2015年6月25日公开的美国专利公开号2015/0178939中描述了诸如衍射光栅的光提取光学元件,其通过引用整体并入在此。

[0093] 在一些实施例中,光提取光学元件440a、438a、436a、434a、432a是形成衍射图案的衍射特征,或“衍射光学元件”(在此也称为“DOE”)。优选地,DOE具有相对较低的衍射效率,使得光束的光的仅一部分在DOE的每个相交处朝向眼睛410偏转离开,而其余的光经由全内反射继续移动通过波导。承载图像信息的光因此可被分成多个相关的出射束,该出射束在多个位置处离开波导,并且结果是针对在波导内到处弹跳的该特定的准直光束而朝向眼睛304的相当均匀的出射发射的图案。

[0094] 在一些实施例中,一个或多个DOE可以在它们主动衍射的“开”状态与它们不明显衍射的“关”状态之间切换。例如,可切换的DOE可以包含聚合物分散液晶层,其中微滴在主体介质中包含衍射图案,并且微滴的折射率可以切换为与主体材料的折射率基本上匹配(在这种情况下,该图案不会明显衍射入射光),或者微滴可以切换到与主体介质的折射率不匹配的折射率(在这种情况下,该图案主动衍射入射光)。

[0095] 在一些实施例中,可以基于观看者的眼睛的瞳孔大小或取向来动态地改变深度平面或景深的数量和分布。景深可能与观看者的瞳孔大小成反比。结果,随着观看者眼睛的瞳孔的大小减小,景深增加,使得因为该平面的位置超出了眼睛的焦点深度而无法辨别的一个平面可能会变得可辨别,并且随着瞳孔大小的减小和景深的相应增加而显得越来越集中。同样地,可以随着瞳孔大小的减小而减少用于向观看者呈现不同图像的间隔开的深度平面的数量。例如,观看者可能在不将眼睛的调节从一个深度平面调节到另一个深度平面的情况下无法在一个瞳孔大小下清楚地感知第一深度平面和第二深度平面二者的细节。然而,这两个深度平面可以同时在不改变调节的情况下以另一个瞳孔大小对用户充分聚焦。

[0096] 在一些实施例中,显示系统可以基于对瞳孔大小或取向的确定,或者基于接收指示特定瞳孔大小或取向的电信号,来改变接收图像信息的波导的数量。例如,如果用户的眼睛无法区分与两个波导相关联的两个深度平面,则控制器460(其可以是本地处理和数据模块260的实施例)可以被配置或编程以停止向这些波导中的一个波导提供图像信息。有利地,这可以减少系统上的处理负担,从而增加系统的响应性。在波导的DOE可在开和关状态之间切换的实施例中,当波导确实接收到图像信息时,DOE可被切换到关状态。

[0097] 在一些实施例中,可能希望使出射束满足直径小于观看者的眼睛直径的条件。然而,鉴于观看者瞳孔大小的可变性,满足该条件可能是一个挑战。在一些实施例中,通过响应于观看者的瞳孔大小的确定而改变出射束的大小,在宽范围的瞳孔大小上满足该条件。例如,随着瞳孔大小减小,出射束的大小也可以减小。在一些实施例中,可以使用可变孔径来改变出射束的大小。

[0098] 可穿戴系统400可以包括对世界470的一部分进行成像的面向外成像系统464(例如,数码相机)。世界470的该部分可以被称为世界相机的视场(FOV),并且成像系统464有时被称为FOV相机。世界相机的FOV可以与观看者210的FOV相同或可以不相同,该观看者210的FOV涵盖观看者210在给定时间感知的世界470的一部分。例如,在一些情况下,世界相机的FOV可以大于可穿戴系统400的观看者210的观看者210。观看者可用于观看或成像的整个区域可以被称为关注场(FOR)。FOR可以包括围绕可穿戴系统400的立体角为 4π 的弧度,因为佩戴者可以移动他的身体、头部或眼睛以感知空间中的基本上任何方向。在其它上下文中,佩戴者的动作可能会更受限制,并且因此佩戴者的FOR可能会产生更小的立体角。从面向外成像系统464获得的图像可以用于跟踪用户做出的手势(例如,手或手指的手势),检测用户前面的世界470中的对象等。

[0099] 可穿戴系统400可以包括音频传感器232,例如麦克风,以捕获环境声音。如上所述,在一些实施例中,可以放置一个或多个其它音频传感器以提供用于确定语音源的位置的立体声接收。作为另一示例,音频传感器232可以包括定向麦克风,该定向麦克风还可以提供关于音频源位于何处的这样的有用的定向信息。可穿戴系统400可以在定位语音源时使用来自面向外成像系统464和音频传感器230二者的信息,或者在特定时刻确定活动说话者等。例如,可穿戴系统400可以单独或者与说话者的反射图像(例如,如在镜子中看到)结合使用语音识别以确定说话者的身份。作为另一示例,可穿戴系统400可以基于从定向麦克风获取的声音来确定说话者在环境中的位置。可穿戴系统400可以采用语音识别算法来解析来自说话者位置的声音,以确定语音的内容,并使用语音识别技术来确定说话者的身份(例如,姓名或其他人口统计信息)。

[0100] 可穿戴系统400还可以包括面向内成像系统466(例如,数码相机),其观察用户的运动,诸如眼睛运动和面部运动。面向内成像系统466可以用于捕获眼睛410的图像以确定眼睛304的瞳孔的大小和/或取向。面向内成像系统466可以用于获得图像以用于确定用户正在看的方向(例如,眼睛姿势)或用户的生物度量识别(例如,经由虹膜识别)。在一些实施例中,可以为每只眼睛使用至少一个相机,以独立地确定每只眼睛单独的瞳孔大小或眼睛姿势,从而允许向每只眼睛呈现图像信息以针对该只眼睛被动态地调整。在一些其它实施例中,仅单只眼睛410的瞳孔直径或取向被确定(例如,每双眼仅使用单个相机),并且假定对于用户的两只眼睛是相似的。可以分析由面向内成像系统466获得的图像以确定用户的眼睛姿势或心情,可穿戴系统400可以使用该眼睛姿势或心情来确定应该向用户呈现哪些音频或视觉内容。可穿戴系统400还可以使用诸如IMU、加速度计、陀螺仪等的传感器来确定头部姿势(例如,头部位置或头部取向)。

[0101] 可穿戴系统400可以包括用户输入设备466,用户可以通过该用户输入设备466向控制器460输入命令以与可穿戴系统400交互。例如,用户输入设备466可以包括触控板、触摸屏、游戏杆、多自由度(DOF)控制器、电容感测设备、游戏控制器、键盘、鼠标、方向垫(D-

pad)、棒、触觉设备、图腾(例如,充当虚拟用户输入设备)等。多DOF控制器可以以控制器的一些或所有可能的平移(例如,向左/向右、向前/向后或向上/向下)或旋转(例如,偏航、俯仰或滚动)来感测用户输入。支持平移运动的多DOF控制器可以称为3DOF,而支持平移和旋转的多自由度控制器可以称为6DOF。在一些情况下,用户可以使用手指(例如,拇指)在触敏输入设备上按下或滑动以向可穿戴系统400提供输入(例如,向由可穿戴系统400提供的用户界面提供用户输入)。用户输入设备466可以在可穿戴系统400的使用期间被用户的手握住。用户输入设备466可以与可穿戴系统400进行有线或无线通信。

[0102] 可穿戴系统的其它组件

[0103] 在许多实现方式中,除了上述可穿戴系统的组件之外或作为其替代,可穿戴系统可包括其它组件。可穿戴系统可以例如包括一个或多个触觉设备或组件。触觉设备或触觉组件可以可操作以向用户提供触感。例如,当触摸虚拟内容(例如,虚拟物体、虚拟工具、其它虚拟构造)时,触觉设备或组件可以提供压力或纹理的触感。触感可以复制虚拟对象所代表的物理对象的感觉,或者可以复制虚拟内容所代表的想象的对象或角色(例如,龙)的感觉。在一些实现方式中,触觉设备或组件可以由用户佩戴(例如,用户可佩戴的手套)。在一些实现方式中,触觉设备或组件可以由用户握持。

[0104] 可穿戴系统可以例如包括可由用户操纵以允许输入可穿戴系统或与可穿戴系统交互的一个或多个物理对象。这些物理对象在此可以称为图腾。一些图腾可能采取无生命物体的形式,诸如例如一块金属或塑料、墙壁、桌子表面。在一些实现方式中,图腾实际上可能没有任何物理输入结构(例如,键、触发器、操纵杆、轨迹球、摇杆开关)。取而代之的是,图腾可以简单地提供物理表面,并且可穿戴系统可以渲染用户界面,以使用户看起来在图腾的一个或多个表面上。例如,可穿戴系统可以将计算机键盘和触控板的图像渲染为看起来驻留在图腾的一个或多个表面上。例如,可穿戴系统可以将虚拟计算机键盘和虚拟触控板渲染为出现在用作图腾的铝制矩形薄板的表面上。矩形板本身没有任何物理键、触控板或传感器。然而,可穿戴系统可以将用户采用矩形板操纵或交互或触摸检测为经由虚拟键盘或虚拟触控板进行的选择或输入。用户输入设备466(图4中所示)可以是图腾的实施例,其可以包括触控板、触摸板、触发器、操纵杆、跟踪球、摇杆或虚拟开关、鼠标、键盘、多自由度控制器或另一物理输入设备。用户可以单独或与姿势结合使用图腾与可穿戴系统或其他用户进行交互。

[0105] 可与本公开的可穿戴设备、HMD和显示系统一起使用的触觉设备和图腾的示例在美国专利公开No.2015/0016777中进行了描述,其全部内容通过引用合并于此。

[0106] 用户与可穿戴系统交互的示例过程

[0107] 图5是用于与虚拟用户界面进行交互的方法500的示例的过程流程图。方法500可以由在此描述的可穿戴系统执行。方法500的实施例可以由可穿戴系统用来检测可穿戴系统的FOV中的人或文档。

[0108] 在框510处,可穿戴系统可以识别特定的UI。UI的类型可以由用户预先确定。可穿戴系统可以识别特定的UI需要基于用户输入(例如,手势、视觉数据、音频数据、感觉数据、直接命令等)来填充。UI可以特定于安全场景,在该安全场景中,系统的佩戴者正在观察向佩戴者(例如,在旅行检查站)呈现文档的用户。在框520处,可穿戴系统可以生成用于虚拟UI的数据。例如,可以生成与UI的边界、总体结构、形状等相关联的数据。另外,可穿戴系统

可以确定用户的物理位置的地图坐标,使得可穿戴系统可以显示与用户的物理位置有关的UI。例如,如果UI以身体为中心,则可穿戴系统可以确定用户的身体姿势、头部姿势或眼睛姿势的坐标,使得可以在用户周围显示环形UI,或者可以在墙上或在用户前面显示平面UI。在此所述的安全性上下文中,可以将UI显示为好像UI围绕正在向系统的佩戴者呈现文档的旅行者一样,使得佩戴者可以在观看旅行者和旅行者的文档时轻松观看UI。如果UI以手为中心,则可以确定用户的手的地图坐标。这些地图点可以通过FOV相机接收的数据、感官输入或任何其它类型的收集数据得出。

[0109] 在框530处,可穿戴系统可以将数据从云发送到显示器,或者可以将数据从本地数据库发送到显示器组件。在框540处,基于所发送的数据向用户显示UI。例如,光场显示器可以将虚拟UI投射到用户的一只或两只眼睛中。一旦已经创建了虚拟UI,则在框550处,可穿戴系统可以简单地等待来自用户的命令以在虚拟UI上生成更多虚拟内容。例如,UI可以是围绕用户的身体或用户环境中的人(例如旅行者)的身体的以身体为中心的环。然后,可穿戴系统可以等待命令(手势、头部或眼睛的运动、语音命令、来自用户输入设备的输入等),并且如果识别出该命令(框560),则可以与该命令相关联的虚拟内容向用户显示(框570)。

[0110] 混合现实中的化身渲染的示例

[0111] 可穿戴系统可以采用各种映射相关技术,以便在渲染的光场中实现高景深。在映射出虚拟世界时,了解真实世界中的所有特征和点以相对于真实世界准确地描绘虚拟对象是有利的。为此,可以通过包含传达有关真实世界的各个点和特征的信息的新图片,将从可穿戴系统用户捕获的FOV图像添加到世界模型。例如,可穿戴系统可以收集一组地图点(例如2D点或3D点)并找到新的地图点以渲染世界模型的更准确版本。可以将第一用户的世界模型(例如,通过诸如云网络的网络)传送给第二用户,使得第二用户可以体验围绕第一用户的世界。

[0112] 图6A是可穿戴系统的另一示例的框图,该可穿戴系统可包括混合现实环境中的化身处理和渲染系统690。可穿戴系统600可以是图2中所示的可穿戴系统200的一部分。在该示例中,可穿戴系统600可以包括地图620,该地图620可以包括地图数据库710(图7中所示)中的数据的一部分。该地图可以部分地驻留在可穿戴系统本地上,并且可以部分地驻留在可通过有线或无线网络访问的联网存储位置(例如,在云系统中)。姿势过程610可以在可穿戴计算架构(例如,处理模块260或控制器460)上执行,并利用来自地图620的数据来确定可穿戴计算硬件或用户的位置和取向。当用户正在体验系统并在世界范围内操作时,可以根据飞行中收集的数据计算姿势数据。数据可以包括图像、来自传感器的数据(诸如惯性测量单元,其通常包括加速度计和陀螺仪组件)以及与真实或虚拟环境中的对象有关的表面信息。

[0113] 稀疏点表示可以是同时定位和映射(例如,SLAM或vSLAM,是指其中输入仅是图像/视觉的配置)过程的输出。该系统可以配置为不仅找出各个组件在世界中的位置,而且找出世界的构成。姿势可以是实现许多目标的构建块,包括填充地图和使用来自地图的数据。

[0114] 在一个实施例中,稀疏点位置本身可能并不完全足够,并且可能需要进一步的信息来产生多焦点AR、VR或MR体验。通常指代深度图信息的密集表示可以被用来至少部分地填充该间隙。可以从称为“立体”640的过程中计算此类信息,其中,使用诸如三角测量或飞行时间感测的技术来确定深度信息。图像信息和活动图案(诸如使用活动投影仪创建的红

外图案)、从图像相机获取的图像或手势/图腾650可用作立体处理640的输入。大量深度图信息可以被融合在一起,并且其中的一些可以采用表面表示来概括。例如,在数学上可定义的表面可能是其它处理设备(如游戏引擎)的有效(例如,相对于大的点云而言)且可理解的输入。因此,立体过程(例如,深度图)640的输出可以在融合过程630中被组合。姿势610也可以是该融合过程630的输入,并且融合630的输出变成填充地图过程620的输入。诸如在地形图绘制中,子表面可以彼此连接以形成较大的表面,并且该地图变成点和表面的大型混合体。

[0115] 为了解决混合现实过程660中的各个方面,可以利用各种输入。例如,在图6A中所示的实施例中,可以输入游戏参数来确定系统的用户正在与各个位置处的一个或多个怪物、在各个位置处的墙壁或其它物体等玩怪物战斗游戏,怪物在各种条件下死亡或逃跑(诸如,如果用户射击了怪物)。世界地图可以包括关于对象的位置的信息或对象的语义信息(例如,诸如对象是平坦还是圆形,水平还是竖直,桌子还是灯等的分类),并且世界地图可以是对于混合现实的另一个有价值的输入。相对于世界的姿势也变成输入,并且在几乎所有交互系统中都扮演着关键角色。

[0116] 来自用户的控制或输入是可穿戴系统600的另一输入。如在此所述,用户输入可包括视觉输入、手势、图腾、音频输入、感官输入等。为了四处移动或玩游戏,例如,用户可能需要指示可穿戴系统600有关他或她想做什么。除了在空间中移动自身外,还存在可以利用的各种形式的用户控件。在一个实施例中,图腾(例如用户输入设备)或诸如玩具枪的对象可以由用户握持并由系统跟踪。该系统优选地将被配置为了解用户正在握持该物品并了解用户正在与该物品进行什么类型的交互(例如,如果图腾或对象是枪,则系统可以被配置为了解位置和取向,以及用户是否点击了触发器或可能配备有传感器(诸如IMU)的其它感应按钮或元件,即使当这种活动不在任何相机的视野之内时,该传感器也可以辅助确定发生了什么)。

[0117] 手势跟踪或识别也可以提供输入信息。可穿戴系统600可以被配置为跟踪和解释用于按钮按压、用于手势向左或向右、停止、抓握、握持等的手势。例如,在一种配置中,用户可能想要在非游戏环境中翻阅电子邮件或日历,或者与另一个人或玩家进行“拳头碰撞”。可穿戴系统600可以被配置为利用最小量的手势,该最小量的手势可以是动态的或可以不是动态的。例如,手势可以是简单的静态手势,如张开手表示停止,上竖大拇指表示可以,下竖大拇指表示不可以;或手向右或向左或向上/向下翻转以表示方向命令。

[0118] 眼睛跟踪是另一种输入(例如,跟踪用户正在看向控制显示技术以在特定深度或范围进行渲染的位置)。在一个实施例中,可以使用三角测量确定眼睛的聚散,并且然后使用针对该特定人开发的聚散/调节模型来确定调节。眼睛跟踪可由眼睛相机执行,以确定眼睛注视(例如,一只眼睛或两只眼睛的方向或取向)。其它技术可以被用于眼睛跟踪,诸如例如,通过放置在眼睛附近的电极(例如,眼电位计)来测量电势。

[0119] 语音跟踪可以是另一种输入,并且可以单独使用或与其它输入(例如,图腾跟踪、眼睛跟踪、手势跟踪等)组合使用。语音跟踪可以包括单独或组合的语音识别、声音识别。系统600可以包括从环境接收音频流的音频传感器(例如,麦克风)。系统600可以结合声音识别技术来确定谁正在说话(例如,语音是来自ARD的佩戴者还是另一个人或语音(例如,语音是来自ARD佩戴者还是其他人还是语音(例如,环境中扬声器发送的录制语音)))以及语音识

别技术来确定在说什么。本地数据和处理模块260或远程处理模块270可以处理来自麦克风的音频数据(或诸如例如用户正在观看的视频流的另一个流中的音频数据),以通过应用各种语音识别算法,诸如例如隐马尔可夫模型、基于动态时间规整(DTW)的语音识别、神经网络、深度学习算法(诸如深度前馈和递归神经网络)、端到端自动语音识别、机器学习算法(参考图7进行描述),或使用声学建模或语言建模的其它算法等,来识别语音内容。

[0120] 本地数据和处理模块260或远程处理模块270也可以应用语音识别算法,该算法可以识别说话者的身份,诸如说话者是可穿戴系统600的用户210还是用户正在交流的其他人。一些示例语音识别算法可以包括频率估计、隐马尔可夫模型、高斯混合模型、模式匹配算法、神经网络、矩阵表示、矢量量化、说话者二义化、决策树和动态时间规整(DTW)技术。语音识别技术还可以包括反说话者技术,诸如同类群组模型和世界模型。频谱特征可用于表示说话者特征。本地数据和处理模块或远程数据处理模块270可以使用参考图7描述的各种机器学习算法来执行语音识别。

[0121] 可穿戴系统的实现方式可以经由UI使用这些用户控件或输入。UI元素(例如,控件、弹出窗口、气泡、数据输入字段等)可以用于例如关闭信息(例如对象的图形或语义信息)的显示。

[0122] 关于相机系统,图6A中所示的示例可穿戴系统600可以包括三对相机:布置在用户面部两侧的相对较宽的FOV或被动SLAM相机对;取向为用户前方的不同对的相机,其用于处理立体成像过程640并捕获在用户面部前方跟踪的手势和图腾/对象。FOV相机和用于立体过程640的相机对可以是面向外成像系统464(图4中所示)的一部分。可穿戴系统600可以包括朝向用户的眼睛取向的眼睛跟踪相机(其可以是图4中所示的面向内成像系统462的一部分),以便对眼睛矢量和其它信息进行三角测量。可穿戴系统600还可包括一个或多个纹理光投影仪(诸如红外(IR)投影仪)以将纹理注入场景。

[0123] 可穿戴系统600可以包括化身处理和渲染系统690。化身处理和渲染系统690可以被配置为基于上下文信息生成、更新、动画化和渲染化身。化身处理和渲染系统690中的一些或全部可以单独地或组合地实现为本地处理和数据模块260或远程处理模块262、264的一部分。在各个实施例中,多个化身处理和渲染系统690(例如,如在不同的可穿戴设备上实现)可以用于渲染虚拟化身670。例如,第一用户的可穿戴设备可以用于确定第一用户的意图,而第二用户的可穿戴设备可以基于从第一用户的可穿戴设备接收的意图来确定化身的特征并渲染第一用户的化身。如将参考图9A和9B所描述的,第一用户的可穿戴设备和第二用户的可穿戴设备(或其它此类可穿戴设备)可以例如经由网络进行通信。

[0124] 图6B示出示例化身处理和渲染系统690。示例化身处理和渲染系统690可以单独或组合地包括3D模型处理系统680、上下文信息分析系统688、化身自动缩放器692、意图映射系统694、解剖结构调节系统698、刺激响应系统696。系统690旨在示出用于化身处理和渲染的功能,而并非旨在进行限制。例如,在某些实现方式中,这些系统中的一个或多个系统可以是另一系统的一部分。例如,上下文信息分析系统688的部分单独地或组合地可以是化身自动缩放器692、意图映射系统694、刺激响应系统696或解剖结构调节系统698的一部分。

[0125] 上下文信息分析系统688可以被配置为基于参考图2和图3描述的一个或多个设备传感器来确定环境和对象信息。例如,上下文信息分析系统688可以使用由用户或用户化身的观看者的面向外成像系统464获取的图像,分析环境和用户环境或渲染用户化身的环境

的对象(包括物理或虚拟对象)。上下文信息分析系统688可以单独或者与从位置数据或世界地图(例如,地图620、710、910)获取的数据组合来分析此类图像,以确定环境中对象的位置和布局。上下文信息分析系统688通常还可以访问用户或人类的生物学特征,以现实地对虚拟化身670制作动画。例如,上下文信息分析系统688可以生成不适感曲线,该不适感曲线可以应用于化身,使得用户的化身身体的一部分(例如头部)关于用户的身体的其它部分没有处于不舒适(或不现实)的位置(例如,化身的头部未转动270度)。在某些实现方式中,一个或多个对象识别器708(图7中所示)可以被实现为上下文信息分析系统688的一部分。

[0126] 化身自动缩放器692、意图映射系统694和刺激响应系统696以及解剖结构调节系统698可以被配置为基于上下文信息来确定化身的特征。化身的一些示例特征可以包括大小、外观、位置、取向、运动、姿势、表情等。化身自动缩放器692可以被配置为自动缩放化身,使得用户不必以不舒服的姿势看化身。例如,化身自动缩放器692可以增加或减小化身的大小以将化身放到用户的视线高度,使得用户不需要分别向下看化身或向上看化身。意图映射系统694可以确定用户交互的意图,并基于渲染化身所在的环境将意图映射到化身(而不是确切的交互)。例如,第一用户的意图可以是在远程呈现会话中与第二用户通信(例如,参见图9B)。通常,两个人在交流时会面对面。第一用户的可穿戴系统的意图映射系统694可以确定在远程呈现会话期间存在此类面对面的意图,并且可以使第一用户的可穿戴系统将第二用户的化身渲染为面对第一用户。如果第二用户要在物理上转身,而不是将第二用户的化身渲染到转身的位置(这将导致第二用户的化身的背面被渲染给第一用户),则第一用户的意图映射系统694可以继续将第二化身的面部渲染给第一用户,这是远程呈现会话的推断意图(例如,在该示例中的面对面意图)。

[0127] 刺激响应系统696可以识别环境中的关注对象,并确定化身对关注对象的响应。例如,刺激响应系统696可以识别化身的环境中的声源,并自动转动化身以看着声源。刺激响应系统696还可以确定阈值终止条件。例如,刺激响应系统696可以在声源消失之后或者经过一段时间之后,使化身回到其原始姿势。

[0128] 解剖结构调节系统698可以被配置为基于生物学特征来调节用户的姿势。例如,解剖结构调节系统698可以被配置为基于不适感曲线来调节用户的头部与用户的躯干之间或用户的上半身与下半身之间的相对位置。

[0129] 3D模型处理系统680可以被配置为对显示器220制作动画并使其渲染虚拟化身670。3D模型处理系统680可以包括虚拟角色处理系统682和运动处理系统684。虚拟角色处理系统682可以被配置为生成和更新用户的3D模型(用于创建虚拟化身并对其制作动画)。运动处理系统684可以被配置为对化身制作动画,诸如例如通过改变化身的姿势,通过在用户的环境中四处移动化身或对化身的面部表情制作动画等。如将在此进一步描述的,可以使用装配技术对虚拟化身制作动画。在一些实施例中,化身以两个部分表示:用于渲染虚拟化身的外部外观的表面表示(例如,可变形的网格)和用于对网格制作动画的互连关节(例如,核心骨架)的分层集合。在一些实现方式中,虚拟角色处理系统682可以被配置为编辑或生成表面表示,而运动处理系统684可以用于通过移动化身,使网格变形等来对化身制作动画。

[0130] 化身处理和渲染系统690可以接收由本文描述的神经网络(例如,RBFNN或FFNN)生成的数字角色装备参数,并且可以使用这些参数来至少部分地使化身动画化。例如,化身处

理和渲染系统690可以执行参考图11描述的过程1100,参考图18描述的过程1800,或者参考图21描述的过程2100。

[0131] 映射用户环境的示例

[0132] 图7是MR环境700的示例的框图。MR环境700可以被配置为从一个或多个用户可穿戴系统(例如,可穿戴系统200或显示系统220)或固定房间系统(例如,房间相机等)接收输入(例如,来自用户的可穿戴系统的视觉输入702,诸如室内相机的固定输入704,来自各种传感器的感官输入706,手势,图腾,眼睛跟踪,来自用户输入设备466的用户输入等)。可穿戴系统可以使用各种传感器(例如,加速度计、陀螺仪、温度传感器、运动传感器、深度传感器、GPS传感器、面向内成像系统、面向外成像系统等)来确定用户环境的位置和各种其它属性。该信息可以进一步补充有来自房间中固定相机的信息,该固定相机可以从不同的视角提供图像或各种提示。由相机(诸如房间相机和/或面向外成像系统的相机)获取的图像数据可以减少到一组映射点。

[0133] 一个或多个对象识别器708可以爬行通过所接收的数据(例如,点的集合)并且识别或映射点、标签图像,借助于地图数据库710将语义信息附加到对象。地图数据库710可以包括随时间推移收集的各种点及其对应的对象。各种设备和地图数据库可以通过网络(例如,LAN、WAN等)彼此连接以访问云。

[0134] 基于该信息和地图数据库中的点的集合,对象识别器708a至708n可以识别环境中的对象。例如,对象识别器可以识别面部、人物、窗户、墙壁、用户输入设备、电视、文档(例如,本安全示例中在此所述的旅行机票、驾照、护照)、用户环境中的其它对象等。一个或多个对象识别器可以专门用于具有某些特征的对象。例如,对象识别器708a可用于识别面部,而另一对象识别器可用于识别文档。

[0135] 可以使用多种计算机视觉技术来执行对象识别。例如,可穿戴系统可以分析由面向外成像系统464(图4中所示)获取的图像以执行场景重构、事件检测、视频跟踪、对象识别(例如人或文档)、对象姿势估计、面部识别(例如,来自环境中的人或文档上的图像)、学习、索引、运动估计或图像分析(例如,在文档中识别标记,诸如照片、签名、标识信息、旅行信息等)等。可以使用一种或多种计算机视觉算法来执行这些任务。计算机视觉算法的非限制性示例包括:尺度不变特征变换(SIFT)、加速鲁棒特征(SURF)、定向FAST和旋转的BRIEF(ORB)、二进制鲁棒不变可扩展关键点(BRISK)、快速视网膜关键点(FREAK)、Viola-Jones算法、Eigenfaces方法、Lucas-Kanade算法、Horn-Schunk算法、均值漂移算法、视觉同时定位和映射(vSLAM)技术、顺序贝叶斯估计器(例如,卡尔曼过滤器、扩展卡尔曼过滤器等)、束调节、自适应阈值(和其它阈值技术)、迭代最近点(ICP)、半全局匹配(SGM)、半全局块匹配(SGBM)、特征点直方图、各种机器学习算法(诸如例如支持矢量机、k最近邻算法、朴素贝叶斯、神经网络(包括卷积或深度神经网络)或其它监督/无监督模型等)等。

[0136] 对象识别可以另外地或可替代地通过各种机器学习算法来执行。训练后,机器学习算法可以由HMD存储。机器学习算法的一些示例可以包括有监督或无监督的机器学习算法,包括回归算法(诸如例如普通最小二乘回归)、基于实例的算法(诸如例如学习矢量量化)、决策树算法(诸如例如分类和回归树)、贝叶斯算法(诸如例如朴素贝叶斯)、聚类算法(诸如例如k均值聚类)、关联规则学习算法(诸如例如先验算法)、人工神经网络算法(诸如例如Perceptron)、深度学习算法(诸如例如Deep Boltzmann Machine或深度神经网络)、降

维算法(诸如例如,主要成分分析)、整体算法(例如,Stacked Generalization)和/或其它机器学习算法。在一些实施例中,可以为各个数据集定制单独的模型。例如,可穿戴设备可以生成或存储基本模型。基本模型可以用作生成特定于数据类型(例如,远程呈现会话中的特定用户)、数据集(例如,用户在远程呈现会话中获得的一组附加图像的集合)、有条件的情况或其它变化的附加模型的起点。在一些实施例中,可穿戴HMD可以被配置为利用多种技术来生成用于分析聚合数据的模型。其它技术可以包括使用预定义的阈值或数据值。

[0137] 基于该信息和地图数据库中的点的集合,对象识别器708a至708n可以识别对象并向对象补充语义信息以赋予对象生命。例如,如果对象识别器将一组点识别为门,则系统可以附加一些语义信息(例如,门具有铰链并且围绕铰链具有90度的运动)。如果对象识别器将一组点识别为镜子,则系统可以附加该镜子具有可以反射房间中对象图像的反射表面的语义信息。语义信息可以包括在此所述的对象的功能可见性。例如,语义信息可以包括对象的法线。系统可以分配矢量,其方向指示对象的法线。随着时间推移,地图数据库会随着系统(其可驻留在本地或可通过无线网络访问)从世界收集更多数据而增长。一旦对象被识别,信息就可以被传输到一个或多个可穿戴系统。例如,MR环境700可以包括关于在加利福尼亚州发生的场景的信息。环境700可被传输给纽约的一个或多个用户。基于从FOV相机和其它输入接收的数据,对象识别器和其它软件组件可以映射从各种图像收集的点,识别对象等,使得可以将场景准确地“传递”给第二用户,该第二用户可能在世界的不同部分。环境700也可以将拓扑图用于定位目的。

[0138] 图8是关于所识别的对象渲染虚拟内容的方法800的示例的过程流程图。方法800描述了可以如何向可穿戴系统的用户呈现虚拟场景。用户可以在地理位置上远离场景。例如,用户可能在纽约,但可能想要观看当前在加利福尼亚州发生的场景,或者可能想与居住在加利福尼亚州的朋友一起散步。

[0139] 在框810处,可穿戴系统可以从用户和其他用户接收关于用户的环境的输入。这可以通过各种输入设备以及地图数据库中已经拥有的知识来实现。在框810处,用户的FOV相机、传感器、GPS、眼睛跟踪等将信息传达到系统。在框820处,系统可以基于该信息来确定稀疏点。稀疏点可以用于确定姿势数据(例如,头部姿势、眼睛姿势、身体姿势或手势),该姿势数据可以用于显示和理解用户周围环境中的各种对象的取向和位置。在框830处,对象识别器708a-708n可爬行通过这些收集的点并使用地图数据库识别一个或多个对象。然后,在框840处,该信息可以被传达到用户的单独的可穿戴系统,并且在框850处,所需的虚拟场景可以因此被显示给用户。例如,所需的虚拟场景(例如,加利福尼亚州的用户)可以相对于纽约的用户的各种对象和其它周围环境以适当的取向、位置等显示。

[0140] 多个可穿戴系统之间的示例通信

[0141] 图9A示意性地示出描绘彼此交互的多个用户设备的整体系统视图。计算环境900包括用户设备930a、930b、930c。用户设备930a、930b和930c可以通过网络990彼此通信。用户设备930a-930c可以各自包括网络接口,以经由网络990与远程计算系统920(也可以包括网络接口971)通信。网络990可以是LAN、WAN、对等网络、无线电、蓝牙或任何其它网络。计算环境900也可以包括一个或多个远程计算系统920。远程计算系统920可以包括集群的并且位于不同地理位置的服务器计算机系统。用户设备930a、930b和930c可以经由网络990与远程计算系统920通信。

[0142] 远程计算系统920可以包括远程数据仓库980,该远程数据仓库980可以维护关于特定用户的物理和/或虚拟世界的信息。数据存储库980可以存储与用户、用户的环境(例如,用户环境的世界地图)或用户的化身的配置有关的信息。远程数据存储库可以是图2中所示的远程数据存储库280的实施例。远程计算系统920还可以包括远程处理模块970。远程处理模块970可以是图2中所示的远程处理模块270的实施例。远程处理模块970可以包括一个或多个处理器,该处理器可以与用户设备(930a、930b、930c)和远程数据仓库980通信。处理器可以处理从用户设备和其它来源获得的信息。在一些实现方式中,处理或存储的至少一部分可以由本地处理和数据模块260(如图2中所示)提供。远程计算系统920可以使给定用户能够与另一用户共享关于特定用户自己的物理和/或虚拟世界的信息。

[0143] 用户设备可以是单独或组合的可穿戴设备(诸如HMD或ARD)、计算机、移动设备或任何其它设备。例如,用户设备930b和930c可以是可以被配置为呈现AR/VR/MR内容的图2中所示的可穿戴系统200(或图4中所示的可穿戴系统400)的实施例。

[0144] 一个或多个用户设备可以与图4中所示的用户输入设备466一起使用。用户设备可以获得关于用户和用户的环境的信息(例如,使用图4中所示的面向外成像系统464)。用户设备和/或远程计算系统1220可以使用从用户设备获得的信息来构造、更新和构建图像、点和其它信息的集合。例如,用户设备可以处理获取的原始信息,并将处理的信息发送到远程计算系统1220以进行进一步处理。用户设备还可以将原始信息发送到远程计算系统1220以进行处理。用户设备可以从远程计算系统1220接收处理的信息,并在投影到用户之前提供最终处理。用户设备还可以处理获得的信息,并将处理的信息传递给其他用户设备。用户设备可以在处理获取的信息的同时与远程数据存储库1280通信。多个用户设备和/或多个服务器计算机系统可以参与所获取图像的构造和/或处理。

[0145] 关于物理世界的信息可以随时间推移发展,并且可以基于由不同用户设备收集的信息。虚拟世界的模型也可随时间推移而发展,并基于不同用户的输入。此类信息和模型有时在此可以被称为世界地图或世界模型。如参考图6和图7所述,由用户设备获取的信息可以用于构造世界地图910。世界地图910可以包括图6A中描述的地图620的至少一部分。各种对象识别器(例如708a,708b,708c...708n)可用于识别对象和标记图像,以及将语义信息附加到对象。这些对象识别器也在图7中描述。

[0146] 远程数据存储库980可以用于存储数据并且促进世界地图910的构造。用户设备可以不断更新关于用户环境的信息并接收关于世界地图910的信息。世界地图910可以由用户或其他人创建。如在此所讨论的,用户设备(例如930a、930b、930c)和远程计算系统920可以单独或组合地构造和/或更新世界地图910。例如,用户设备可以与远程处理模块970和远程数据仓库980通信。用户设备可以获取和/或处理关于用户和用户环境的信息。远程处理模块970可以与远程数据存储库980和用户设备(例如,930a、930b、930c)通信,以处理关于用户和用户的环境的信息。远程计算系统920可以修改由用户设备(例如,930a、930b、930c)获取的信息,诸如例如,选择性地裁剪用户的图像,修改用户的背景,将虚拟对象添加到用户的环境,采用辅助信息注释用户的语音等。远程计算系统920可以将处理的信息发送到相同和/或不同的用户设备。

[0147] 远程呈现会话的示例

[0148] 图9B描绘了其中相应可穿戴系统的两个用户正在进行远程呈现会话的示例。该图

中示出两个用户(在该示例中分别名为Alice 912和Bob 914)。两个用户正在佩戴他们相应的可穿戴设备902和904,该可穿戴设备902和904可以包括参考图2描述的HMD(例如,系统200的显示设备220),用于表示远程呈现会话中的另一用户的虚拟化身。两个用户可以使用可穿戴设备进行远程呈现会话。注意,图9B中将两个用户分开的竖直线旨在示出当Alice 912和Bob 914经由远程呈现进行通信的同时(例如,Alice可能在她位于亚特兰大的办公室内而Bob在波士顿户外时)可能(但是不是必须)位于两个不同的位置。

[0149] 如参考图9A所描述的,可穿戴设备902和904可以彼此通信或者与其他用户设备和计算机系统通信。例如,Alice的可穿戴设备902可以例如经由网络990(在图9A中示出)与Bob的可穿戴设备904通信。可穿戴设备902和904可以跟踪用户的环境以及环境中的运动(例如,经由相应的面向外成像系统464,或者一个或多个位置传感器)和语音(例如,经由相应的音频传感器232)。可穿戴设备902和904还可以基于由面向内成像系统462获取的数据来跟踪用户的眼睛运动或注视。在一些情况下,可穿戴设备还可以捕获或跟踪用户的面部表情或其它身体运动(例如手臂或腿部运动),其中用户靠近反射表面,并且面向外成像系统464可以获得用户的反射图像,以观察用户的面部表情或其它身体运动。

[0150] 可穿戴设备可以使用从第一用户和环境中的获取的信息来对虚拟化身制作动画,该虚拟化身将由第二用户的可穿戴设备渲染,以在第二用户环境中产生第一用户的有形的存在感。例如,可穿戴设备902和904、远程计算系统920可以单独或组合地处理由Bob的可穿戴设备904呈现的Alice的图像或运动,或者可以处理由Alice的可穿戴设备902呈现的Bob的图像或运动。如在此进一步描述,可以基于诸如例如用户的意图、用户的环境或其中渲染化身的环境或人的其它生物学特征的上下文信息来渲染化身。

[0151] 尽管示例仅涉及两个用户,但是在此描述的技术不应限于两个用户。使用可穿戴设备(或其它远程呈现设备)的多个用户(例如,两个、三个、四个、五个、六个或更多)可以参加远程呈现会话。在远程呈现会话期间,特定用户的可穿戴设备可以向该特定用户呈现其他用户的化身。此外,虽然该图中的示例将用户示出为站在环境中,但是不需要用户站立。在远程呈现会话期间,任何用户都可以站立、坐着、跪着、躺下、步行或奔跑,或者处于任何位置或运动。除了在此示例中所述,用户还可能处于物理环境中。在进行远程呈现会话的同时,用户可以处于单独的环境中,或可以处于相同的环境中。并非所有用户都需要在远程呈现会话中佩戴其相应的HMD。例如,在Bob 914佩戴可穿戴设备904的同时,Alice 912可以使用其它图像采集和显示设备,诸如网络摄像头和计算机屏幕。

[0152] 虚拟化身的示例

[0153] 图10示出可穿戴系统的用户所感知的化身的示例。图10中所示的示例化身1000可以是站在房间中的物理植物后面的Alice 912(图9B中所示)的化身。化身可包括各种特征,诸如例如大小、外观(例如肤色、面色、发型、衣服、面部特征,诸如皱纹、痣、斑点、粉刺、酒窝等)、位置、取向、运动、姿势、表情等。这些特征可以基于与化身相关联的用户(例如,Alice的化身1000可以具有实际人Alice 912的一些或全部特征)。如在此进一步描述的,化身1000可以基于上下文信息被动动画化,其可以包括对化身1000的一个或多个特征的调节。尽管在此通常被描述为表示人(例如Alice)的身体外观,但这是为了说明而不是限制。Alice的化身可以表示除了Alice、拟人化的对象、生物或任何其它真实或虚拟的表示之外的另一个真实或虚构的人的外观。此外,图10中的植物不必是物理的,而可以是由可穿戴系统向用

户呈现的植物的虚拟表示。此外,与图10中所示的内容相比,附加的或不同的虚拟内容也可以呈现给用户。

[0154] 用于虚拟角色的装配(rigging)系统的示例

[0155] 诸如人类化身的动画虚拟角色可以在计算机图形中被全部或部分地表示为多边形网格。多边形网格或简称为“网格”,是在所建模的三维空间中的点的集合。网格可以形成多面体对象,其表面定义虚拟角色(或其一部分)的主体或形状。尽管网格可以包括任何数量的点(在可用计算能力可能会施加的实际限制内),但具有更多点的更精细的网格通常能够描绘出具有可以更接近现实生活中的人、动物、物体等的更精细细节的更逼真的虚拟角色。图10示出化身1000的眼睛周围的网格1010的示例。化身

[0156] 网格中的每个点都可以通过所建模的三维空间中的坐标定义。所建模的三维空间可以是例如由(x,y,z)坐标寻址的笛卡尔空间。网格中的点是构成多面体对象的多边形的顶点。每个多边形表示多面体对象的表面或面,并由有序的一组顶点定义,每个多边形的边是连接有序的一组顶点的直线边缘。在一些情况下,网格中的多边形顶点可与几何多边形不同,其中它们在3D图形中不一定共面。另外,网格中多边形的顶点可以是共线的,在这种情况下,多边形具有零面积(被称为退化多边形)。

[0157] 在一些实施例中,网格由三顶点多边形(例如,三角形或简称为“tris”)或四顶点多边形(例如,四边形或称为“quad”)组成。然而,也可以在一些网格中使用高阶多边形。在直接内容创建(DCC)应用(例如,诸如Maya(可从Autodesk, Inc.获得)或Houdini(可从Side Effects Software Inc.获得)的应用)中,网格通常基于四边形,其主要设计用于创建和操纵3D计算机图形),而网格在实时应用中通常是基于三角形的。

[0158] 为了对虚拟角色制作动画,可以通过在各个时刻将虚拟角色的一些或全部顶点移动到空间中的新位置来变形虚拟角色的网格。变形既可以表示大规模运动(例如,四肢运动)又可以表示精细运动(例如,面部运动)。这些变形和其它变形可以基于真实世界的模型(例如,对执行身体运动,关节、面部扭曲,表情等的真实人类进行摄影测量扫描)、艺术指导的开发(其可基于真实世界的采样)、相同或其它技术的组合。在计算机图形学的早期,可以通过独立设定顶点的新位置来手动完成网格变形,但是鉴于现代网格的大小和复杂性,通常期望使用自动化系统和过程来产生变形。用于产生这些变形的控制系统、过程和技术称为装配,或简称为“装配(the rig)”。图6B的示例化身处理和渲染系统690包括可实现装配的3D模型处理系统680。

[0159] 用于虚拟角色的装配可以使用骨架系统来辅助网格变形。骨架系统包括关节的集合,该关节对应于至少一部分网格的关节。在装配的上下文中,关节有时被称为“骨骼”而不管在解剖学意义上使用时的这些术语之间的差异。骨架系统中的关节可以根据可以被应用于关节的变换彼此相对移动或以其它方式改变。变换可以包括空间的平移或旋转以及其它操作。可以为关节分配彼此之间的层次关系(例如,父子关系)。这些层次关系可以允许一个关节从另一个关节继承变换或其它特征。例如,骨架系统中的子关节可以继承被分配给其父关节的变换,以使子关节中的运动与父关节中的运动相对应。

[0160] 可以采用适当位置(例如解剖学参考)处的关节以及采用适当的局部旋转轴、自由度等来定义虚拟角色的骨架系统,以允许要被执行的期望的一组网格变形。一旦为虚拟角色定义了骨架系统,就可以在被称为“蒙皮”的过程中为每个关节分配对网格中各个顶点的

影响量。这可以通过为骨架系统中每个关节的每个顶点分配权重值来完成。当将变换应用于任何给定关节时,可以根据该关节变换对受该关节影响的顶点自动地移动或以其它方式改变可取决于它们相应的权重值的量。

[0161] 装配可以包括多个骨架系统。一种类型的骨架系统是核心骨架(也被称为低阶骨架),其可用于控制虚拟角色的大规模运动。例如,在人类化身的情况下,核心骨架可能类似于人的解剖骨架。尽管用于装配目的的核心骨架可能无法精确地映射到在解剖学上正确的骨架,但它可能在相似位置中具有带有类似取向和运动特性的关节子集。

[0162] 如上所述,关节的骨架系统可以是分层的,例如,关节之间具有父子关系。当将变换(例如,位置和/或取向的变化)应用于骨架系统中的特定关节时,可以将相同的变换应用于相同层次结构内的所有其它较低级别的关节。例如,在用于人类化身的装配的情况下,核心骨架可包括用于化身的肩部、肘部和腕部的单独关节。其中,可以将肩关节分配给层次结构中的最高级别,而可以将肘关节分配为肩关节的子级,并且可以将腕关节分配为肘关节的子级。因此,当将特定的平移和/或旋转变换被应用于肩关节时,相同的变换也可以被应用于肘关节和腕关节,使得它们以与肩膀相同的方式被平移和/或旋转。

[0163] 尽管其名称具有含义,但是装配中的骨架系统不必表示解剖骨架。在装配中,骨架系统可以表示用于控制网格变形的各种层次结构。例如,头发可以被表示为层次链中的一系列关节;由于化身的面部扭曲(其可表示诸如微笑、皱眉、笑、说话、眨眼等的表情)引起的皮肤运动可以由一系列由面部装配控制的面部关节来表示;肌肉变形可以由关节建模;以及衣服的运动可以由关节网格来表示。

[0164] 虚拟角色的索装可以包括多个骨架系统,其中的一些可以驱动其它骨架的运动。低阶骨架系统是驱动一个或多个高阶骨架系统的系统。相反,高阶骨架系统是由低阶骨架系统驱动或控制的系统。例如,虽然角色的核心骨架的运动可能由动画师手动控制,但核心骨架可以进而驱动或控制高阶骨架系统的运动。例如,更高阶的辅助关节(在物理骨架中可能没有解剖类似物)可以被提供以改进由核心骨架的运动引起的网格变形。可以从被应用于低阶骨架的变换中,在算法上得出被应用于高阶骨架系统中的这些关节和其它关节的变换。高阶骨架可以表示例如肌肉、皮肤、脂肪、衣服、头发或不需要直接动画控制的任何其它骨架系统。通常,在一些实施例,低阶骨架系统可以表示数字角色的大规模运动,而高阶骨架系统可以表示数字角色的小规模运动。

[0165] 如已经讨论的那样,可以将变换应用于骨架系统中的关节,以便执行网格变形。在装配的上下文中,变换包括接受3D空间中的一个或多个给定点并产生一个或多个新3D点的输出的函数。例如,变换可以接受定义关节的一个或多个3D点,并可以输出指定变换的关节的一个或多个新3D点。关节变换可以包括例如平移分量、旋转分量和缩放分量。

[0166] 平移是如下变换,该变换将所建模的3D空间中的一组一个或多个指定点移动指定量,而该点集的取向或大小不变。旋转是将所建模的3D空间中的一组一个或多个指定点围绕指定轴旋转指定量(例如,将网格中的每个点围绕z轴旋转45度)的变换。仿射变换(或6自由度(DOF)变换)是包括平移和旋转的一种变换。仿射变换的应用可以被认为是移动一组一个或多个空间点而不改变其大小,但是取向可以改变。

[0167] 同时,缩放变换是一种通过以指定值缩放所建模的3D空间中的一个或多个指定点的相应坐标来修改该指定点的变换。这将改变所变换的一组点的大小和/或形状。均匀缩放

变换可以将每个坐标缩放相同的量,而非均匀缩放变换可以独立缩放指定点的(x,y,z)坐标。例如,可以使用非均匀缩放变换来提供挤压和拉伸效果,诸如可能由肌肉动作产生的效果。另一类型的变换是剪切变换。剪切变换是通过基于该坐标距轴的距离将点的坐标平移不同量来修改建模3D空间中一组一个或多个指定点的变换。

[0168] 当将变换被应用于关节以使其移动时,在该关节的影响下的顶点也会被移动。这导致网格变形。如上所述,分配权重以量化每个关节对每个顶点的影响的过程称为蒙皮(或有时称为“权重绘画”或“蒙皮加权”)。权重通常是介于0(表示无影响)和1(表示完全影响)之间的值。网格中的一些顶点可能仅受单个关节影响。在那种情况下,这些顶点针对该关节被分配1的权重值,并且它们的位置基于被分配给该特定关节而不是其它关节的变换而改变。网格中的其它顶点可受到多个关节影响。在该情况下,将为每个影响关节的那些顶点分配0到1之间的单独的权重,每个顶点的权重之和等于1。基于被分配给所有的那些影响关节的变换来改变这些顶点的位置。例如,如果数字角色举起手臂,则在其总共50个低阶骨头/关节(脚、腿、手指关节等)中,只有两个(锁骨和肩关节)权重可能不为零。锁骨可以具有比肩部权重(例如,0.3)更大的权重(例如,0.7),以确保相对较小的锁骨运动足以使分配给锁骨和肩部的顶点变形。

[0169] 对网格中的所有顶点进行权重分配可能会非常费力,特别是当关节数量增加时。对于即使是训练有素的艺术家来说,响应于被应用于关节的变换来平衡权重以实现期望的网格变形也可能非常困难。在实时应用的情况下,许多实时系统还对可被加权到特定顶点的关节数量(通常为8个或更少)施加了限制的事实可使任务进一步复杂化。通常出于图形处理单元(GPU)的效率施加这种限制。

[0170] 术语蒙皮还指基于应用于骨架系统中关节的变换,使用所分配的权重而实际使网格变形的过程。例如,动画师可以指定一系列核心骨架关节变换,以生成期望的角色运动(例如,跑步运动或舞步)。将变换应用于一个或多个关节时,在变换的关节的影响下为顶点计算新位置。通常,将任何给定顶点的新位置计算为影响该特定顶点的所有关节变换的加权平均值。存在用于计算该加权平均值的许多算法,但是最常见的并且由于其简单易用且易于控制而在大多数实时应用中使用的算法是线性混合蒙皮(LBS)。在线性混合蒙皮中,使用针对其的顶点具有非零权重的每个关节变换来计算每个顶点的新位置。然后,将这些关节变换中的每一个关节变换产生的新顶点坐标与被分配给关节中的每个关节的其顶点的相应权重成比例地平均。在实践中,存在着对于LBS众所周知的限制,而制作高质量装配的许多工作都致力于发现并克服这些限制。为此专门设计了许多辅助关节系统。

[0171] 除了蒙皮,“混合形状”也可以用在装配中以产生网格变形。混合形状(有时也称为“变形目标”或简称为“形状”)是一种应用于网格中一组顶点的变形,其中,该一组顶点中的每个顶点都基于权重在指定方向中移动指定量。对于特定的混合形状,集合中的每个顶点可能具有其自己的自定义运动,并且同时移动集合中的顶点将生成期望的形状。可以由“增量(delta)”指定混合形状中每个顶点的自定义运动,该“增量”是表示施加到该顶点的XYZ运动的量和方向的矢量。仅举几种可能性,混合形状可用于产生例如使眼睛、嘴唇、眉毛、鼻子、酒窝等移动的面部变形。可以通过任何方法来创建混合形状,该方法包括针对一组顶点将中立姿势定义为零,将目标姿势定义为一,其中混合形状权重使该组顶点在中立姿势与目标姿势之间移动权重或分数量。蒙皮和混合形状都需要权重,并且它们都使网格变形,但

是蒙皮和混合形状都有各自单独的权重。一般来说,混合形状可以建立在蒙皮之上用于附加的网格变形。

[0172] 混合形状可用于以一种艺术指导的方式使网格变形。它们可以提供很大的控制力,因为可以通过扫描模型来雕刻或捕获精确的形状。但是,混合形状的益处是以必须存储混合形状中所有顶点的增量为代价的。对于具有精细网格和许多混合形状的动画角色,增量数据量可能很大。

[0173] 通过使用混合形状权重,可以将每个混合形状应用到指定程度。这些权重通常在0(根本没有应用混合形状)到1(其中混合形状完全激活)的范围内。例如,移动角色眼睛的混合形状可以以较小的权重被施加以少量移动眼睛,或者可以施加较大的权重以产生较大的眼睛运动。

[0174] 装配可以彼此组合地施加多个混合形状以实现期望的复杂变形。例如,为了产生微笑,装配可能会应用混合形状以拉动唇角,抬高上唇和降低下唇,以及移动眼睛、眉毛、鼻子和酒窝。从两个或多个混合形状组合而成的期望形状称为组合形状(或简称为“组合”)。

[0175] 组合应用两个混合形状可能导致的一个问题是,混合形状可能在一些相同的顶点上操作。当两个混合形状都激活时,结果称为双重变换或“脱离模型”。该问题的解决方案通常是校正混合形状。校正混合形状是一种特殊的混合形状,其关于当前施加的变形表示期望的变形,而不是关于中性表示期望的变形。可以基于他们正在校正的混合形状的权重来应用校正混合形状(或仅称为“校正”)。例如,可以使校正混合形状的权重与触发混合校正形状的应用的基础混合形状的权重成比例。

[0176] 校正混合形状也可以被用于校正蒙皮异常或改进变形的质量。例如,关节可以表示特定肌肉的运动,但作为单个变换,它不能表示皮肤、脂肪和肌肉的所有非线性行为。在肌肉激活时应用一种校正或一系列校正可以导致更令人愉悦和令人信服的变形。

[0177] 装配内置于层中,较低较简单的层通常驱动较高阶的层。这适用于骨架系统和混合形状变形两者。例如,如已经提到的,用于动画虚拟角色的装配可以包括由低阶骨架系统控制的高阶骨架系统。存在基于低阶骨架控制高阶骨架或混合形状的许多方法,包括约束、逻辑系统和基于姿势的变形。

[0178] 约束通常是一种系统,其中特定对象或关节变换控制应用于另一个关节或对象的变换的一个或多个分量。存在许多不同类型的约束。例如,目标约束将目标变换的旋转改变为指向特定方向或特定对象。父约束充当变换对之间的虚拟父子关系。位置约束将变换约束到特定点或特定对象。取向约束将变换限制为对象的特定旋转。

[0179] 逻辑系统是数学方程式的系统,其在给定一组输入的情况下产生一些输出。这些是指定的而不是学习的。例如,混合形状值可能被定义为其它两个混合形状的乘积(这是称为组合或组合形状的校正形状的示例)。

[0180] 基于姿势的变形也可以用于控制高阶骨架系统或混合形状。骨架系统的姿势由该骨架系统中所有关节的变换(例如旋转和平移)集合定义。也可以为骨架系统中关节的子集定义姿势。例如,可以通过应用于肩部、肘部和腕部的关节的变换来定义手臂姿势。姿势空间变形器(PSD)是一种系统,该系统用于基于该姿势与定义的姿势之间的一个或多个“距离”来确定特定姿势的变形输出。这些距离可以是表征一个姿势与另一个姿势如何不同的度量。PSD可以包括姿势插值节点,例如,该姿势插值节点接受一组关节旋转(定义姿势)作

为输入参数,并且进而输出归一化的每姿势权重以驱动变形器,诸如混合形状。姿势插值节点可以以多种方式实现,包括采用径向基函数(RBF)。RBF可以执行机器学习的函数数学近似。可以使用一组输入及其关联的预期输出来训练RBF。训练数据可以是例如多组关节变换(其定义特定姿势)和响应于这些姿势而应用的对应混合形状。一旦学习了该函数,就可以给出新的输入(例如姿势),并且可以有效地计算它们的预期输出。RBF是人工神经网络的子类型。RBF可用于基于较低级别组件的状态来驱动装配的较高级别组件。例如,核心骨架的姿势可以以较高级别驱动辅助关节和校正。

[0181] 这些控制系统可以被链接在一起以执行复杂的行为。例如,眼睛装配可以包含两个“环视”值以进行水平和竖直旋转。这些值可以通过一些逻辑传递,以确定眼睛关节变换的精确旋转,然后可以将其用作RBF的输入,该RBF控制改变眼睑的形状以匹配眼睛的位置的混合形状。这些形状的激活值可用于使用附加逻辑来驱动面部表情的其它组件,依此类推。

[0182] 装配系统的目标通常是提供一种基于简单的人类可理解的控制系统产生令人愉悦的高保真变形的机制。在实时应用的情况下,目标通常是提供足够简单以在例如VR/AR/MR系统200上实时运行的装配系统,同时尽可能对最终质量做出很少的妥协。在一些实施例中,3D模型处理系统680执行装配系统以实时地在混合现实环境100中对化身制作动画以(与VR/AR/MR系统的用户)交互,并在用户的环境中提供适当的上下文化身行为(例如,基于意图的行为)。

[0183] NN方法概述

[0184] 图11示出了使用本文描述的系统和方法驱动数字角色装配的过程1100的示例概述。过程1100可以通过与可穿戴系统200相关联的硬件处理器来执行,例如本地处理和数据模块260、远程处理模块270、远程计算系统920、或其组合。

[0185] 如上所述,数字角色由数字角色装配表示。在一些实施例中,化身由低阶数字角色装配表示。该化身可以是佩戴可穿戴设备的用户的化身,如图9B所示的902或904。用户可以与可穿戴设备的另一用户进行化身聊天,其中第一用户(例如爱丽丝(Alice))看到另一用户(例如鲍勃(Bob))的虚拟数字化身。在一个示例中,爱丽丝和鲍勃可能正在举办一个虚拟舞会,其中化身会随着YMCA歌曲而跳舞。化身可以类似于图9B中所示的那样,以一个中立或A姿势开始舞蹈会话,然后用他们的胳膊摆出字母Y,从而请求新的装配姿势。

[0186] 系统接收数字角色的姿势的变化的指示(例如,新的舞蹈姿势)1102。与舞蹈姿势相关联的低阶数字角色装配参数被提供给NN 1104。NN过程在被配置为执行NN的系统上执行1106,然后输出表示新舞蹈姿势的高阶数字角色装配参数1108。输出接着可以被用来驱动高阶数字角色装配1110,该高阶数字角色装配在本示例中可以使爱丽丝和鲍勃的化身摆出字母Y。例如,可穿戴显示系统200可以使用数字角色装配参数来渲染通过显示器220提供给系统200的用户210的化身。

[0187] 输入和输出

[0188] 图12A示出了将为图11中的1104处的输入提供参数的示例低阶数字角色装配,以及图12B示出如图11中1110处的示例高阶数字角色装配。如上所述,低阶骨架1200A可以表示看起来类似于解剖骨架的简化版本骨架。骨头和关节的位置大致与解剖骨架相似。例如,c和b之间的线1202可以表示锁骨,e 1204可以表示肩部关节。

[0189] 图12A中的低阶骨架有9个输入参数,由9个关节/骨头表示,标记为a-j。这九个输入参数仅为了说明而示出,不应限制本发明的范围。在实践中,可以使用更少或更多的输入参数,从1、2、5、10、100到1000或更多,可以取决于需要至少半独立于其他身体部位而进行移动的身体部位的数量(因此可以取决于被操纵的特定数字角色)。输入参数可以是关节、骨头、混合形状和/或表示低阶骨架的任何其他方法。

[0190] 如上所述,高阶骨架1200B可以表示韧带、肌腱和/或可以添加到低阶骨架上的附加骨头。这使变形显得更逼真和栩栩如生,具有更好的肌肉张力和整体形状。图12B中的高阶骨架有6个高阶输出参数k-p。六个高阶输出参数仅用于说明目的,不应限制本发明的范围。在实践中,可以使用更少或更多的输入参数,范围从1、5、10、100到1000,或更多。输出参数可以是关节、骨头、高阶校正(corrective)和/或表示高阶骨架的任何其他方法。

[0191] 在替代实施例中,输入和输出可以二者表示高阶骨架表示,或者它们可以二者表示低阶骨架表示。在一些示例实施例中,本文所述方法和系统用作在已知的训练姿势或样本姿势之间进行插值的方式。

[0192] 在实践中,给定的骨头可以有多个线性分量(x、y、z上的平移、骨头长度、尺度等),但为了便于说明,所公开的实施例利用由单个值表示的线性分量进行描述。此外,根据情景,每个所请求的数字角色姿势1104可能没有针对每个参数的角度分量和/或线性分量。例如,如果所请求的数字角色姿势1104与中立姿势基本相同,那么大多数值将为零,只有从中立姿势偏移的一组关节具有非零值。

[0193] 在一些实施例中,输入是低阶数字角色装配参数,输出是高阶数字角色装配参数。从低阶输入开始可以有利于本发明存在于更广泛的系统。通常,与高阶装配参数相比,低阶装配参数生成得更容易且更快,因为高阶装配参数是根据低阶装配参数计算而得的。在此描述的实施例无论输入或输出是高阶还是低阶都能发挥作用。从低阶输入开始但具有高阶输出可以提供更快的处理速度的额外益处,因此,在一些实施例中,实现实时性能。

[0194] 图13A至图13B进一步说明和定义了用于数字角色装配的姿势。输入姿势表示针对给定姿势的一组关节、混合形状和/或骨头的值仍然是正确的,但在一些实施例中,姿势可以相对于中立位置进行描述。中立位置可以是任何姿势,但常用的是A姿势(如图12A至图12B所示)或T姿势。在实践中,提供到NN中的输入姿势可以是姿势1300A与所选择的中立姿势1300B之间的差。输入姿势1300A可以是化身随着YMCA跳舞,从而摆出字母Y,如图11所示。因此,输入姿势可以表示相对于中立姿势的姿势的相对平移和旋转。

[0195] 在一些实施例中,低阶角色装配1300A可以用九个输入参数来表示。这九个输入参数可以各自具有角度分量和线性分量,总共有18个参数描述输入姿势。在 $S^0(3)$ 角度空间中,角度参数可以表示为包括四个值(x,y,z,w)的向量。线性参数可以使用线性分量可以由一个值来表示的简化假设。这意味着,在一些实施例中,可能有9个线性输入参数+9个角度参数 $\times 4$ (x,y,z,w)=45个数字表示数字角色相对于中立姿势的输入姿势。

[0196] 在一些实施例中,输入姿势不相对于中立位置,而是可以相对于固定坐标系或任何其他合适的参考系。

[0197] RBFNN内容

[0198] 被配置为对一个或多个数字角色姿势进行插值的神经网络的第一示例实施例是被配置为分离线性分量和角度分量的RBFNN,如下文进一步详细描述。

[0199] 图14示出了RBFNN 1604内的每个样本节点1400中包含的各种组件(参考图16进行描述)。尽管一些实施例可以包含更多或更少的组件,将描述四个主要的组件,一组训练姿势1402、径向基函数1404、一组距离权重, γ ,1406、和一组基权重, w ,1408。

[0200] 在一些实施例中,每个样本节点可以包含一组训练姿势1402。这些训练姿势可用于任何类型的神经网络,而且不限于RBFNN。这些训练姿势表示数字角色在姿势空间中的一组姿势。训练姿势可以被组织为从期望的动画序列取得的一组静态姿势,或者他们可以是单独的姿势。例如,一些训练姿势可以表示抬腿。第一姿势可以是左腿和右腿并拢,垂直向下。第二、第三和第四姿势可以是右腿抬起20度,然后60度,然后90度。可以有一系列的训练姿势,其中伸直腿的状态下使腿移动通过这些角度,也可以有不同的一系列训练姿势,其中在膝盖弯曲的状态下以前面提到的角度抬起腿。左腿也可以重复这个过程。任何数量的训练姿势可以被选择和使用,并且不限制本发明的范围。如上所述,理想情况下,所需的训练姿势数量越少就越好,以减少收集输入、训练和运行NN所需的时间和精力。在系统需要实时运行的情况下,重要的是使训练姿势的数量最小化。然而,这需要与如下认识相平衡:如果使用太少的训练姿势,系统可能不能准确地对姿势进行插值。

[0201] 图15A至图15C示出了可能使用的三个潜在训练姿势。图15A至图15C可以表示人在弯曲手臂的情况下一系列姿势中的各种姿势。

[0202] 在一些实施例中,每个样本节点可以包含径向基函数1404。这个函数对于每个样本节点都是相同的(对于每个RBF,每个训练姿势都创建一个样本节点)。它由距离函数和基函数组成,其中的基函数应用于距离函数的输出。这个函数组合是标准的RBF公式。然而,RBF中的距离函数通常是简单的函数,如欧氏范数。这是可能的,因为输入通常是相同类型的。然而,用于数字角色装配的输入有多种类型,包括线性空间和角度空间中的参数。

[0203] 在一些实施例中,每个样本节点可以包含一组距离权重 γ 1406。对于 n 个输入(低阶关节、骨头、混合形状等),有 n 个线性和 n 个角度距离权重,总共有 $2 * n$ 个距离权重。这些距离权重在RBF距离函数中使用,如下所述,并在训练期间由用户分配、部分地由用户分配和/或计算。由于有多种输入类型,并且需要将不同类型的输入组合在一起,所以需要这些距离权重。一个示例输入类型是踝关节类型,第二输入类型是膝关节。膝关节的运动范围比踝关节大。如果数字角色需要移动他们的下半身,比如踢球,则这组距离权重将有助于确保较小的踝关节运动相比于较大的膝关节运动不被视为舍入误差。可以使用更多或更少的权重,以匹配输入的形式。如前所述,一个简化假设是假设线性参数只有一个值(相对于角度参数有四个值)。在替代实施例中,如果系统替代地利用由三个值组成的线性参数来建立,则需要相等数量的距离权重——每个输入参数值一个距离权重。替代实施例可以具有表示每个线性输入和每个角度输入的任意数量的值,并且可以具有与每个输入参数值对应的相等数量的距离权重。

[0204] 在一些实施例中,每个样本节点可以包含一组基权重 w 1408。对于 m 个输出(高阶关节、骨头、高阶校正等),可以有 m 个线性和 $4 * m$ 个角度基权重。在本发明的系统和方法的一些实施例中,可以针对每个角度输出计算四个值,因此对于驱动数字角色装配的每个角度输出有多个基权重。这些基权重在网络的训练阶段被计算,并在评估阶段被固定(例如,RBFNN的使用阶段,与训练阶段相对)。采用RBFNN的标准技术进行了基权重的计算。对于典型的RBFNN,每个输出被认为是独立的,并且有与输出一样的基权重。因为我们的一些输

出是角度的,需要多于一个实数(例如,可以表示 x 、 y 、 z 、 w 的四个值),我们对每个输出都有多个基权重,并执行一个额外的步骤,如下面的描述,把这些值组合到角度输出中。可以计算任意数量的输出。在一些实施例中,可以计算1、5、10、100或1000或更多的输出参数。无论RBFNN系统提供多少输出,每个输出都可以有一个基权重 w 。

[0205] 用于根据RBFNN对数字角色的输出姿势进行插值的系统的示意图

[0206] 图16是用于根据RBFNN对数字角色的输出姿势进行插值的系统的示意图,示例性数据在系统内被传输。示意图1600从1602开始,其中用于数字角色装配的新姿势由 n 个输入参数表示。这个新姿势可以是化身做出YMCA舞蹈序列中的Y。在一些实施例中,输入姿势可以表示为中立参考姿势之间的线性和角度差异。在一些实施例中,线性参数可以由单个实数来表示,角度参数可以由任何合适的角度量来表示,如欧拉角、旋转矩阵、轴角、四元数等。只要馈入到RBFNN的数据被分离成线性组和角度组,输入可以在被系统接收之前,或者作为系统中的步骤,被分离成一组线性分量和角度分量。对于 n 个输入,有 n 个线性参数和 n 个角度参数,总共有 $2 * n$ 个输入参数表示该姿势。每个参数可以由一个或多个值组成,例如每个角度参数包括四个值,每个线性参数包括一个值。每个参数可以包括比在这里提供的示例更多或更少的值,并且不应该被解释为限制本发明的范围。分离的线性和角度输入参数被传递给RBFNN 1604,并输入到每个样本节点1606。每个训练姿势有一个样本节点。到目前为止,一个新颖的地方是把角度参数与线性参数分开。所有的输入参数(分离的线性和角度参数)都被传递给所有的样本节点。

[0207] 每个样本节点1606可以包含参照图14描述的节点1400中指定的四个组件1402、1404、1406、1408,但可能包含更多或更少的组件。

[0208] 样本节点1606可以各自计算一组距离,每个输入值对应一个距离。对于线性输入,距离是输入值与来自RBFNN中包含的每个训练(例如,样本)姿势的对应值之间的欧氏距离。在替代实施例中,可以使用任何其他适当的线性距离方程。对于角度输入,距离可以是SO(3)空间中计算的角度距离。通常,RBFNN只有线性输入,并且只使用欧氏范数。

[0209] 样本节点1606也可以各自计算出考虑到不同类型数据的最终距离。最终距离是单个的实数,使用上面描述的权重 γ 1406计算为加权的欧氏范数。这个权重是一个新颖的点,在典型的RBFNN中是不存在的。

[0210] 样本节点1606可以各自计算基值。该值是至少部分根据最终距离计算出来的,是基函数的输出,对所有RBFNN都是标准的。一般来说,径向基函数是径向函数的网络。这类函数已被证明在利用距离和径向函数的适当选择来近似常规且未知函数中是有用的。径向函数是一种其值是基于到原点的距离而定义的函数,虽然可以使用原点以外的参考点,其中的距离通常被定义为欧氏距离。径向基函数可以是多个径向函数的和。

[0211] 样本节点1606也可以各自计算以 w 加权的基值。这是样本节点1606的输出。基值可以乘以每个 w 基权重,并被传递到相应的累积(accumulation)节点1608或1610。可以有两个累积节点,一个用于线性输出1608,一个用于角度输出1610。线性输出可以计算为从每个样本传递的加权值的总和。每个样本节点向每个角度累积节点传递4个值(图16中用圆表示单个实数值)。这些值可以以与线性输出相同的方式累积,得到四个和。前三个和可用作旋转轴向量的 x 、 y 、 z 分量。该旋转轴向量可以缩放为单位长度。第四个也是最后一个值可以用作绕旋转轴的角度旋转值,提供了3D旋转的共同的角-轴表示,其可以输出到相应的关节、骨

头、高阶校正或其他高阶装配表示而作为角度输出。最后,所有的输出1612被传递到外部系统来驱动数字角色。

[0212] RBFNN的示例训练

[0213] 图17示出了用于训练此处公开的RBFNN的示例过程。过程1700示出了在数字角色装配的上下文中训练RBFNN的过程,但是应该理解其他上下文也是可能的。从本文公开的发明中获益最大的上下文是由两种或两种以上不同类型的数据组成的上下文,和/或包括角度输入的上下文,如下文将进一步公开。过程1700可以由适合于训练NN的硬件处理器来执行。在一些实施例中,可以通过远程处理模块270或远程计算系统920执行训练。训练后的RBFNN可以通过网络990被传送到可穿戴系统200。在一些实施例中,可穿戴系统200可以获得附加的训练数据,并且RBFNN可以进一步由例如本地处理和数据模块260、远程处理模块270或远程计算系统920来训练,这可以有助于个性化RBFNN到可穿戴系统200的用户210观看的特定的姿势或数字角色。

[0214] 在1702块,接收数字角色姿势的训练组,其中输入可以表示低阶数字角色装配姿势参数,输出可以表示高阶数字角色装配姿势参数。如上所述,替代实施例可以使用其他输入/输出训练数据组,例如都表示高阶数字角色装配的输入和输出,或都表示低阶数字角色装配的输入和输出。在替代实施例中,输入数据可以不表示数字角色装配,但仍然可以具有由多于一个类型组成和/或包括角度数据的输入。

[0215] 在1704块,训练数据被分为角度参数和线性参数。如上所述,为了改善插值结果而将角度参数和线性参数分离。当使用线性方程或处理方法对角度参数进行插值时,可能会出现不希望的输出,如万向节锁定、快速翻转和/或符号或区域的改变,其中单个角度旋转可以用多种方式表示。在替代实施例中,只要数据子集可以受益于后续的特殊处理,例如对每个子集进行唯一的距离计算,就可以将训练数据分离为两个或多个数据子集,如下所述。角度数据和线性数据可以用任何适当的方法分离。例如,可以创建脚本、计算机程序或其他代码,其可以识别角度数据的形式(例如,包含四个元素的向量)和线性数据的形式(例如,单个值),并根据输入数据的格式自动分组数据。可替代地,可以根据输入数据源预先标记输入数据,这样系统可以简单地读取标记并根据预定义的方法来操纵输入数据。在一些实施例中,输入数据的格式定义输出数据的所需格式,并且可以添加任意数量的前处理或后处理步骤,以便操作该数据来匹配期望的数据格式。

[0216] 在1708和1710块,可以定义RBFNN的距离函数。1710块可以定义对于线性参数的距离函数。这是通常定义为欧式范数,虽然其他距离函数也是可行的,并可写成 $\Delta(q_1, q_2) = |q_1 - q_2|$, 或可以替代地(更常见)写成 $\Psi(x) = \Psi(|x|)$, 其中 Ψ 是相对于原点来定义距离时的径向函数,或者可以在相对于原点以外的点来定义距离时替代地被写成 $\Psi(x) = \Psi(|x - c|)$ 。线性距离函数可以替代使用任何距离函数 Δ 来一般化为 $\Psi(x) = \Psi(\Delta(x, c))$ 。块1708可以定义对于角度参数的距离函数。两个旋转之间的角度距离可以在0到180度或0到 π 弧度范围内。两个角之间的距离可以由 $\Delta(q_1, q_2) = \arccos(2\langle q_1, q_2 \rangle^2 - 1)$ 来定义,其中在 q_1 和 q_2 可以是表示为四元数的角,并且 $\langle q_1, q_2 \rangle$ 表示相应的四元数的内积,例如, $q_1 = [x_1, y_1, z_1, w_1]$ 且 $q_2 = [x_2, y_2, z_2, w_2]$ 且 $\langle q_1, q_2 \rangle = x_1x_2 + y_1y_2 + z_1z_2 + w_1w_2$ 。在该实施例中,距离度量是对称的,并且 $\Delta(q_1, q_2) + \Delta(q_2, q_3) \neq \Delta(q_1, q_3)$, 这可能是前面描述的手臂抬起示例的数学表示,该示例演示了两个手臂如何都处于在用户前方抬起的目标位置,但是由于每只手臂进行了不同

组的平移/旋转,两只手的取向不同。虽然块1708和1710指定了角度距离和线性距离函数,但替代实施例可以利用更适合被操作的数据类型的其他函数。

[0217] 块1712和1714在过程1700是可选步骤,并示出了为线性参数和角度参数分配距离权重 γ 的步骤。如上文参考图14在1406中提到的,对于 n 个输入(低阶关节、骨头、混合形状等),有 n 个线性和 n 个角度距离权重 γ ,总共有 $2 * n$ 个距离权重。这些距离权重在RBF距离函数中用于计算最终的距离函数,并由用户分配、部分地由用户分配和/或在训练期间计算。由于有多种输入类型,并且需要将不同类型的输入组合在一起,所以需要这些距离权重。一个示例输入类型是锁骨类型,第二输入类型是肩关节类型。肩关节与锁骨相比具有更大的运动范围,所以一组距离权重 γ 可以有助于确保较小的锁骨运动在当与较大的肩关节运动相比较时是不可忽视的。

[0218] 在1716块,最终距离函数可以定义为从块1708和块1710定义的方程输出的每个距离的 γ 加权的欧氏范数。

[0219] 在1718块,训练输入/输出组被输入到RBFNN中。

[0220] 在1720块,RBFNN调整一组基权重 w ,以便在使用相应输入时,将来自RBFNN的计算输出与来自训练组的已知输出相匹配。基权重 w 由RBFNN自动计算,并且只在训练阶段进行调整。如上关于1408所述,对于 m 个输出(高阶关节、骨头、高阶校正等),可以有 m 个线性和 $4 * m$ 个角度权重值。在本发明的一些系统和方法中,对于每个角度输出计算了四个值,因此对于驱动数字角色装配的每个角度输出存在多个基权重。采用RBFNN的标准技术来进行基权重的计算。

[0221] 在1720块之后,RBFNN是固定的(例如,在训练组中所有的训练输入/输出对已经被输入到RBFNN之后)。这意味着不再允许变化基权重 w ,基权重 w 现在被认为是固定值,除非提供附加的训练数据。RBFNN现在已经准备好对于来自新输入的新输出进行插值。在一些实施例中,这可以意味着数字角色的新高阶姿势可以基于低阶数字角色输入姿势进行插值,例如当化身随着YMCA歌曲跳舞的同时摆出Y时。

[0222] 使用经训练的RBFNN的示例过程

[0223] 图18示出了使用经训练的RBFNN对于数字角色的新姿势进行插值的示例过程1800。过程1800可以通过与可穿戴系统200相关联的硬件处理器执行,例如,本地处理和数据模块260、远程处理模块270、远程计算系统920或其组合。

[0224] 在块1802,RBFNN系统接收由数字角色装配表示的数字角色的低阶姿势,其中的数字角色装配由关节、骨头、混合形状或描述数字角色装配的其他方法来定义。如上面关于图13所提到的,这些输入可以具有角度参数和线性参数,每个角度参数或线性参数可以包含一个或多个值。如上参考图11所述,当前的低阶姿势可以是一个人在随着YMCA歌曲跳舞时用他/她的手臂摆出字母Y。

[0225] 在块1804,当前的姿势被分离为角度参数和线性参数。在块1806,所有的输入参数都被传递到RBFNN,在块1808,RBFNN将这些输入传递到每个样本节点。

[0226] 在块1810,每个样本节点确定输入姿势和样本姿势之间的距离。这意味着可以计算出共 $(n+n*4)*S$ 个值,其中两个是因为有两种类型的计算(例如,一个用于线性和一个用于角度), n 是输入的数量(关节、骨头、混合形状等等),并且 S 是RBFNN内包含的样本的数量。对于图13A至图13B和图16中的特定示例,这意味着在1810块处计算出 $(9+9*4)*3=$

135个距离值。

[0227] 在1812块,通过利用距离权重 γ ,如上面在1716中所讨论的形成加权的欧氏范数来计算最最终距离。这意味着所有训练姿势(例如,样本节点)上的单个关节的所有对应距离被组合成单个距离度量。对于图13的示例,这意味着我们可以有9个线性值和9个角度值,每个关节a至j有一个线性值和一个角度值。

[0228] 在1814块,标准基函数应用于1812块的输出。这个步骤将值从相对于n的输入值转换为相对于m的输出值。标准基函数可以是许多选项之一,如高斯基函数、三次(cubic)基函数、多重二次(multiquadratic)基函数或逆多重二次基函数。在过程的这一阶段,相对于图13的示例,可以有6个线性和 $4 * 6 = 24$ 个角度无关的输出。更广泛地说,每个输出参数(关节、骨头、高阶校正等)可以有 $5 * m$ 个独立的值。

[0229] 在块1816,RBFNN训练阶段学习到的基权重w可以应用到块1814的输出中。

[0230] 在块1818,值根据输出类型和关节进行分组。例如,角度值可以组合成表示 $S^0(3)$ 空间的向量,其中向量包含四个值 $= [x, y, z, w]$ 。可选地,单个关节的线性输出和角度输出可以配对。

[0231] 在块1820,经过插值的高阶数字角色姿势被发送到高阶角色装配来驱动角色变形。在图11的示例中,这意味着数字角色以随着YMCA歌曲跳舞时呈现Y形手臂的姿势来进行渲染。这个过程可以反复重复,以创建数字角色的动画。过程1800的后续循环可以显示数字角色从Y形状的手臂,到M形状的手臂,再到C形状的手臂,然后是A形状的手臂的转换。

[0232] 用于根据FFNN对数字角色的输出姿势进行插值的系统的示意图

[0233] 被配置为对一个或多个数字角色姿势进行插值的神经网络的第二示例实施例是被配置为分离线性分量和角度分量的前馈神经网络(FFNN),下文将进一步详细描述。

[0234] 图19示出了根据本发明的一些实施例,用于对数字角色的输出姿势进行插值的FFNN架构1900的一部分。如前所述,驱动关节(如低阶骨架)与辅助关节(如高阶骨架)之间的关系可以是复杂的,但可以利用神经网络来求解。在一些实施例中,可以使用多层FFNN,并且可以需要更少的参数来求解一个或多个输出。这可以导致计算时间的减少,例如从RBFNN实施例的每帧4.5毫秒减少到FFNN实施例的每帧0.9毫秒,这可以导致显著的性能改进,特别是在计算有限的设备上。在一些实施例中,可以针对运动的每个分量分别训练FFNN。运动的三个示例分量可以是旋转、平移和缩放。在一些实施例总,FFNN的输入数据可以被预处理以使位置参数(例如,旋转、缩放)可以缩放到单位范围(例如,在0到1之间)。在一些实施例中,旋转参数可以从动画编写系统(例如,DCC系统)中使用的欧拉角转换为单位四元数。

[0235] 在一些实施例中,FFNN可以是使用修正线性单元(ReLU)激活函数的具有单个隐藏层的全连接网络。在一些实施例中,隐藏层可以作为RESNET块合并。在一些实施例中,要使用的隐藏维度的数量最好是通过超参数调试来发现。在一些实施例中,可以使用下列公式来确定要使用的隐藏维度的数量:

[0236] $10 * \max(\text{驱动器的数量}, \text{辅助器的数量})$

[0237] 作为示例,数字角色装配的小腿系统可能包括膝关节和足关节,它们共同驱动位于小腿中部的单个关节。当数字角色的腿移动时,小腿关节可以使数字角色的小腿变形。在这个例子中,膝关节和足关节是驱动关节(或称为驱动器),小腿中部关节是辅助关节(或称

为辅助器)。每个驱动器可以向NN提供三个平移参数X、Y、Z,并且单个辅助关节可能需要三个平移。所以对于小腿这个示例,输入层有六个神经元(三个参数乘以两个驱动关节),隐藏层有二十个神经元(十乘以两个驱动器),输出层有三个神经元(三个参数乘以一个辅助关节)。在一些实施例中,缩放网络架构可以是相同的,因为缩放也可以具有三个参数X、Y、Z。在一些实施例中,每一个旋转,其可以是四元数旋转,可以具有四个元素,因此旋转网络架构1900的输入层1902可以具有八个神经元,隐藏层1904可以有二十个神经元,输出层1906可以有四个神经元。

[0238] 在一些实施例中,FFNN系统可以能够实现与更大的RBFNN系统相同的结果。与RBFNN类似,FFNN可以具有输入层1902,该输入层1902可以接受来自驱动关节的控制参数。输出层可以产生辅助关节可能被设置的值。

[0239] NN的示例训练

[0240] 图20示出了用于训练本文公开的NN的示例过程。过程2000示出了在数字角色装配上下文中训练NN的过程,但是应该理解其他上下文也是可能的。从本文公开的发明中获益最大的上下文是由两种或两种以上不同类型的数据组成,和/或如上所述包括角度输入的上下文。过程2000可以由适合训练NN的硬件处理器来执行。在一些实施例中,可以通过远程处理模块270或远程计算系统920执行训练。经训练的NN可以通过网络990传送到可穿戴系统200。在一些实施例中,可穿戴系统200可以获得附加的训练数据,并且NN可以进一步由例如本地处理和数据模块260、远程处理模块270或远程计算系统920来进行训练,这可以有助于使NN个性化至由可穿戴系统200的用户210观看的特定的姿势或数字角色。

[0241] 在块2002,接收数字角色姿势的训练组,其中的输入可以表示低阶数字角色装配姿势参数,输出可以表示高阶数字装配姿势参数。如上所述,可选实施例可以使用其他输入/输出训练数据组,例如输入和输出都表示高阶数字角色装配,或输入和输出都表示低阶数字角色装配。在替代实施例中,输入数据可以不表示数字角色装配,但仍然可以具有由多于一个类型组成和/或包括角度数据的输入。

[0242] 在2004块,训练数据被分离为每个运动分量。在一些实施例中,分离可以分为角度参数和线性参数。在一些实施例中,分离可以分为旋转、平移和/或缩放分量。如上所述,为了改善插值结果而分离参数/分量。当使用线性方程或处理方法对角度参数进行插值时,可能会出现不期望的输出,如万向节锁定、快速翻转和/或符号或区域的改变,其中单个角度旋转可以用多种方式表示。在替代实施例中,只要数据子集可以受益于后续的特殊处理,例如对每个子集进行唯一的距离计算,就可以将训练数据分离为两个或多个数据子集,如下所述。参数/分量可以用任何合适的方法分离。例如,可以创建脚本、计算机程序或其他代码,其可以识别角度数据的形式(例如,包含四个元素的向量)和线性数据的形式(例如,单个值),并自动按输入数据的格式对数据进行分组。替代地,可以根据输入数据源预先标记输入数据,这样系统就可以简单地读取标记并根据预定义的方法操作输入数据。在一些实施例中,输入数据的格式定义输出数据的期望格式,并且可以添加任意数量的前处理或后处理步骤,以便操作数据来匹配期望的数据格式。

[0243] 在块2018,将训练输入/输出集输入到NN中。

[0244] 在块2018之后,NN是固定的(例如,训练组内的所有训练输入/输出对已经输入到NN之后),训练阶段可以完成。现在,NN已经准备好根据新的输入对新的输出进行插值。在一

些实施例中,这可以意味着数字角色的新高阶姿势可以基于低阶数字角色输入姿势进行插值,例如当化身随着YMCA歌曲跳舞的同时摆出Y时。

[0245] 使用经训练的NN的示例过程

[0246] 图21示出了使用经训练的NN对数字角色的新姿势进行插值的示例过程2100。在一些实施例中,NN可以是FFNN或RBFNN。过程2100可以通过与可穿戴系统200相关联的硬件处理器执行,例如,本地处理和数据模块260、远程处理模块270、远程计算系统920或其组合。

[0247] 在块2102,NN系统接收由数字角色装配表示的数字角色的低阶姿势,其中的数字角色装配由关节、骨头、混合形状或描述数字角色装配的其他方法来定义。如上面关于图13所提到的,这些输入可以具有角度参数和线性参数,每个角度参数或线性参数可以包含一个或多个值。如上关于图11所提到的,当前的低阶姿势可以是一个人在随着YMCA歌曲跳舞时用其手臂摆出字母Y。

[0248] 在步骤2104,当前的姿势可以被分离成角度参数和线性参数,分离成运动的分量,或任何其他合适的分离方法。在块2106,所有的输入参数都可以被传递到NN中,然后NN可以在步骤2108执行。

[0249] 在步骤2018,可以根据输出类型和关节对值进行分组。在一些实施例中,可以组合角度输出和线性输出以表示单个关节的运动。在一些实施例中,可以组合运动的分量以表示单个关节的旋转、平移和/或缩放运动。

[0250] 在步骤2020,可将经插值的高阶数字角色姿势发送到高阶角色装配以驱动角色变形。在图11的示例中,这意味着数字角色是以随着YMCA歌曲跳舞时Y形手臂的姿势进行渲染。这个过程可以反复重复,以创建数字角色的动画。过程2000的后续循环可以显示数字角色从Y形状的手臂转换到M形状的手臂,再到C形状的手臂,然后是A形状的手臂。

[0251] 实时的可移植性

[0252] 虚拟角色的装配可以包括用于基于各种输入自动控制装配元素(如高阶装配元素)的控制系统,并且可以包括RBFNN、FFNN或其他神经NN系统。在一些实施例中,在特定虚拟角色的装配中有多个交错层的控制系统。在本公开的上下文中,装配控制系统包括一个或多个规则(如逻辑规则、数学规则等等),其基于与低阶装配元素或其他源相关联的输入来确定用于控制(如调节、调整、指定、选择、调用或其它影响)装配元素的输出,诸如高阶骨架系统或高阶混合形状。由于装配控制系统可以非常复杂,已经创建了各种工具来编写和实现它们。这些工具通常利用特定于应用的装配控制协议。特定的装配控制协议可以包括一组特定于应用的数据格式、数据结构、函数、计算单元(例如,节点、类等)和/或编程语言。当不同的应用使用不同的装配控制协议时,不可能直接转移已经在一个应用中编写的装配控制系统,然后在另一个应用中执行它。

[0253] 数字内容创建(DCC)应用经常用于开发针对虚拟角色的装配,包括装配控制系统。这类DCC应用的一个示例是Autodesk **Maya®**。DCC应用可以提供各种工具来定义装配元素,如多边形网格、骨架系统、混合形状等,并定义控制系统来执行这些装配元素的自动控制。装配控制系统是在DCC应用中使用其特定于应用的装配控制协议来构建的。

[0254] 虽然DCC应用非常适合创建虚拟角色及其关联的装配,但它们通常不适合在例如游戏或VR/AR/MR应用中实时显示虚拟角色(包括关联角色装配的实时执行)。替代地,已经为此开发了实时引擎。这种实时引擎的一个例子是从Epic Games可获得的**Unreal®**引擎。

实时引擎的主要设计目的是显示完成的装配资产,并不是很适合编写复杂的装配资产。虽然实时引擎可以包含用于创建装配元素和/或装配控制系统的工具,但它们通常不像DCC应用中存在的工具那么鲁棒;在DCC应用中指定、构建和测试新的装配资产(包括装配控制系统)要容易得多。因此,实时引擎通常依赖于DCC应用来提供虚拟角色及其相关联的装配。这需要将装配资产从DCC应用转移到实时引擎。

[0255] 存在诸如Filmbox (**FBX®**) 格式的工具,用于在应用之间转移网格、混合形状、骨架系统和动画。然而,没有类似的工具可以在应用之间转移装配控制系统。虽然网格、混合形状、骨架系统和动画可以更容易地在应用之间移植的格式来很好地定义,但装配控制系统通常是紧密耦合到特定应用的数据结构和约定(例如,编写应用所使用的装配控制协议)。因此,装配控制系统从编写应用转移到实时应用需要在实时应用环境中重新实现控制系统,上述重现实现的方式与软件开发人员可能以多个软件语言重新实现算法从而在多个应用中和/或多个计算平台上运行的方式非常相同。然而,重新实现虚拟角色的装配控制系统可能是一个非常困难、耗时且昂贵的提议,由于它们的复杂性和应用之间验证校验的困难。

[0256] 上述NN系统和方法能够克服从第一应用转移到第二应用的挑战,同时保留通常为离线渲染而保留的高保真生物力学动画和变形。在一些实施例中,一旦NN被设计并用于DCC应用,诸如Maya,权重和其他参数就可以序列化为任何方便的数据格式(如另一种标记语言(YAML)、JavaScript对象表示法(JSON)、可扩展标记语言(XML)、Filmbox (FBX) 等等),从而系统可以读取到其他应用中,诸如实时引擎。在目标应用中只需要实现RBF节点和导入逻辑,这与移植NN通常所基于的整个复杂的装配系统的传统方法不同,然后就会遇到上面描述的转移问题。换句话说,这里描述的过程能够转移复杂的装配系统,因为创建神经网络的训练阶段,如图17和图20所示,将复杂的装配逻辑集成到NN中。换句话说,通过转移神经网络,这些过程正在传递复杂的装配系统。装配封装性和可移植性也在第62/658,415号美国专利申请中进行了描述,该专利申请通过引用全部并入本文中。

[0257] NN系统和方法实现了应用之间的可移植性。在一些实施例中,这可以通过将复杂的、分层的、分层的、应用特定的装配控制系统改变为数学构造(例如NN)来实现,因此消除了应用特定的构造(换句话说,数学是任何系统都可以进行编程以理解的通用语言)。

[0258] 附加方面

[0259] 1. 一种方法,包括:接收包括至少一个角度分量和至少一个线性分量的输入数据;将所述输入数据作为输入提供给至少一个神经网络NN,所述至少一个神经网络NN已经被训练为不同地评估所述至少一个角度分量和所述至少一个线性分量;以及接收由所述至少一个NN基于对所述至少一个角度分量和所述至少一个线性分量的不同评估而生成的输出数据。

[0260] 2. 根据方面1所述的方法,其中,所述至少一个角度分量处于三维(S0(3))空间中的特殊正交群。

[0261] 3. 根据方面1或方面2所述的方法,其中,所述神经网络中的至少一个包括前馈神经网络FFNN。

[0262] 4. 根据方面3所述的方法,其中,至少一个FFNN是全连接网络。

[0263] 5. 根据方面3或方面4所述的方法,其中,所述前馈神经网络包括单个隐藏层。

- [0264] 6. 根据方面4或方面5所述的方法,其中,所述FFNN包括修正线性单元激活函数。
- [0265] 7. 根据方面5或方面6所述的方法,其中,所述隐藏层是残差NN块。
- [0266] 8. 根据方面1至方面7中的任一项所述的方法,其中,所述NN中的至少一个包括径向基函数神经网络RBFNN。
- [0267] 9. 根据方面1至方面8中的任一项所述的方法,其中,所述输入数据描述数字角色的姿势。
- [0268] 10. 根据方面1至方面9中的任一项所述的方法,其中,所述输入数据表示数字角色的低阶骨架,并且所述输出数据表示数字角色的高阶骨架。
- [0269] 11. 根据方面1至方面10中的任一项所述的方法,其中,所述输出数据描述数字角色的姿势。
- [0270] 12. 根据方面1至方面11中的任一项所述的方法,其中,所述输入数据和所述输出数据中的一个或多个还包括第三分量。
- [0271] 13. 根据方面12所述的方法,其中,所述角度分量、所述线性分量、以及所述第三分量分别是不同的运动分量。
- [0272] 14. 根据方面12或方面13所述的方法,其中:所述至少一个角度分量描述旋转运动;所述至少一个线性分量描述平移运动;以及所述第三分量描述缩放。
- [0273] 15. 根据方面12至方面14中的任一项所述的方法,其中,所述至少一个NN包括评估所述至少一个角度分量的第一FFNN和评估所述至少一个线性分量的第二FFNN。
- [0274] 16. 根据方面15所述的方法,其中,所述至少一个NN包括评估所述第三分量的第三FFNN。
- [0275] 17. 根据方面1至方面16中的任一项所述的方法,其中,所述至少一个NN包括多个样本节点,其中,每个样本节点对应于训练姿势,其中,所述训练姿势中的至少一个包括至少一个角度分量和至少一个线性分量。
- [0276] 18. 根据方面1至方面17中的任一项所述的方法,其中,所述至少一个NN通过评估 $S^0(3)$ 空间中的所述至少一个角度分量以及通过利用欧氏距离公式评估所述至少一个线性分量,来不同地评估所述至少一个角度分量和所述至少一个线性分量。
- [0277] 19. 根据方面1至方面18中的任一项所述的方法,还包括:训练一个或多个NN,其中,所述训练包括:接收包括训练输入数据和训练输出数据的训练数据,其中,所述训练输入数据和所述训练输出数据表示一个或多个训练姿势,其中,所述一个或多个训练姿势中的至少一个包括输入角度分量、输入线性分量、输出角度分量、以及输出线性分量;将来自一个或多个姿势的每个姿势的所述输入角度分量分组为输入角度分量组;将来自一个或多个姿势中的每个姿势的所述输入线性分量分组为输入线性分量组;以及提供所述训练输入数据作为输入以训练所述至少一个NN,其中,所述输入角度分量组和所述输入线性分量组被不同地评估,其中,所述评估导致所述输出角度分量和所述输出线性分量。
- [0278] 20. 根据方面19所述的方法,其中,所述输入角度分量和输出角度分量在 $S^0(3)$ 或四元数空间中。
- [0279] 21. 根据方面19或方面20所述的方法,其中,所述神经网络是前馈神经网络(FFNN)。
- [0280] 22. 根据方面21所述的方法,其中,所述FFNN是全连接网络。

- [0281] 23. 根据方面21或方面22所述的方法, 其中, 所述FFNN包括单个隐藏层。
- [0282] 24. 根据方面21至方面23中的任何一个所述的方法, 其中, 所述FFNN包括修正线性单元激活函数。
- [0283] 25. 根据方面23或方面24所述的方法, 其中, 所述隐藏层是残差NN块。
- [0284] 26. 根据方面19至方面25中的任何一个所述的方法, 其中, 所述NN是径向基函数神经网络(RBFNN)。
- [0285] 27. 根据方面19至方面26中的任何一个所述的方法, 其中, 所述训练数据是数字角色的姿势。
- [0286] 28. 根据方面19至方面27中的任何一个所述的方法, 其中, 所述训练数据包括数字角色的低阶姿势和数字角色的高阶姿势。
- [0287] 29. 根据方面19至方面28中的任何一个所述的方法, 其中, 所述输出角度分量和所述输出线性分量表示数字角色的姿势。
- [0288] 30. 根据方面19至方面29中的任何一个所述的方法, 其中, 所述一个或多个训练姿势中的至少一个包括第三输入分量和第三输出分量。
- [0289] 31. 根据方面30所述的方法, 其中, 所述第三输入分量和第三个输出分量是缩放。
- [0290] 32. 根据方面19至方面31中的任何一个所述的方法, 其中, 所述输入角度数据和输出角度数据表示数字角色的运动的旋转分量。
- [0291] 33. 根据方面19至方面32中的任何一个所述的方法, 其中, 所述输入线性数据和输出线性数据表示数字角色的运动的平移分量。
- [0292] 34. 一种系统, 包括: 至少一个处理器, 其执行至少一个神经网络NN; 以及通信耦合到所述至少一个处理器的存储器, 所述存储器存储指令, 所述指令在由所述至少一个处理器执行时使所述至少一个处理器执行包括以下的操作: 接收包括至少一个角度分量和至少一个线性分量的输入数据; 将所述输入数据作为输入提供给至少一个神经网络NN, 所述至少一个神经网络NN已经被训练为不同地评估所述至少一个角度分量和所述至少一个线性分量; 以及接收由所述至少一个NN基于对所述至少一个角度分量和所述至少一个线性分量的不同评估而生成的输出数据。
- [0293] 35. 根据方面34所述的系统, 其中, 所述至少一个角度分量处于三维(S0(3))空间中的特殊正交群。
- [0294] 36. 根据方面34或方面35所述的系统, 其中, 所述神经网络中的至少一个是前馈神经网络FFNN。
- [0295] 37. 根据方面36所述的系统, 其中, 至少一个FFNN是全连接网络。
- [0296] 38. 根据方面36或方面37所述的系统, 其中, 所述FFNN包括单个隐藏层。
- [0297] 39. 根据方面36至方面38中的任何一个所述的系统, 其中, 所述FFNN包括修正线性单元激活函数。
- [0298] 40. 根据方面38或方面39所述的系统, 其中, 所述隐藏层是残差NN块。
- [0299] 41. 根据方面34至方面40中的任一项所述的系统, 其中, 所述NN中的至少一个是径向基函数神经网络RBFNN。
- [0300] 42. 根据方面34至方面41中的任一项所述的系统, 其中, 所述输入数据描述数字角色的姿势。

[0301] 43. 根据方面34至方面42中的任一项所述的系统,其中,所述输入数据表示数字角色的低阶骨架,并且所述输出数据表示数字角色的高阶骨架。

[0302] 44. 根据方面34至方面43中的任一项所述的系统,其中,所述输出数据描述数字角色的姿势。

[0303] 45. 根据方面34至方面44中的任一项所述的系统,其中,所述输入数据和所述输出数据中的一个或多个还包括第三分量。

[0304] 46. 根据方面45所述的系统,其中,所述角度分量、所述线性分量、以及所述第三分量分别是不同的运动分量。

[0305] 47. 根据方面45或方面46所述的系统,其中:所述至少一个角度分量描述旋转运动;所述至少一个线性分量描述平移运动;以及所述第三分量描述缩放。

[0306] 48. 根据方面45至方面47中的任一项所述的系统,其中,所述至少一个NN包括评估所述至少一个角度分量的第一FFNN和评估所述至少一个线性分量的第二FFNN。

[0307] 49. 根据方面48所述的系统,其中,所述至少一个神经网络包括评估所述第三分量的第三FFNN。

[0308] 50. 根据方面34至方面49中的任一项所述的系统,其中,所述至少一个NN包括多个样本节点,其中,每个样本节点对应于训练姿势,其中,所述训练姿势中的至少一个包括至少一个角度分量和至少一个线性分量。

[0309] 51. 根据方面34至方面50中的任一项所述的系统,其中,所述至少一个NN通过评估S0(3)空间中的所述至少一个角度分量以及通过利用欧氏距离公式评估所述至少一个线性分量,来不同地评估所述至少一个角度分量和所述至少一个线性分量。

[0310] 52. 根据方面34至方面51中的任一项所述的系统,其中,所述至少一个处理器还被编程为:训练一个或多个NN,其中,为了训练所述至少一个处理器被编程为:接收包括训练输入数据和训练输出数据的训练数据,其中,所述训练输入数据和训练输出数据表示一个或多个训练姿势,其中,所述一个或多个训练姿势中的至少一个包括输入角度分量、输入线性分量、输出角度分量、以及输出线性分量;将来自一个或多个姿势的每个姿势的所述输入角度分量分组成输入角度分量组;将来自一个或多个姿势中的每个姿势的所述输入线性分量分组成输入线性分量组;以及提供所述训练输入数据作为到至少一个神经网络NN的输入,其中,所述输入角度分量组和所述输入线性分量组被不同地评估,其中,所述评估导致所述输出角度分量和所述输出线性分量。

[0311] 53. 根据方面52所述的系统,其中,所述输入角度分量和输出角度分量在S0(3)空间中。

[0312] 54. 根据方面52或方面53所述的系统,其中,所述神经网络是前馈神经网络。

[0313] 55. 根据方面54所述的系统,其中,所述前馈神经网络是全连接网络。

[0314] 56. 根据方面54或方面55所述的系统,其中,所述前馈神经网络包括单个隐藏层。

[0315] 57. 根据方面54至方面56中的任何一个所述的系统,其中,所述前馈神经网络包括修正线性单元激活函数。

[0316] 58. 根据方面56或方面57所述的系统,其中,所述隐藏层是残差神经网络块。

[0317] 59. 根据方面52至方面58中的任何一个所述的系统,其中,所述神经网络是径向基函数神经网络(RBFNN)。

[0318] 60. 根据方面52至方面59中的任何一个所述的系统,其中,所述训练数据是数字角色的姿势。

[0319] 61. 根据方面60所述的系统,其中,所述训练数据包括数字角色的低阶姿势和数字角色的高阶姿势。

[0320] 62. 根据方面52至方面61中的任何一个所述的系统,其中,所述输出角度分量和所述输出线性分量表示数字角色的姿势。

[0321] 63. 根据方面52至方面62中的任何一个所述的系统,其中,所述一个或多个训练姿势中的至少一个包括第三输入分量和第三输出分量。

[0322] 64. 根据方面63所述的系统,其中,所述第三输入分量和第三个输出分量是缩放。

[0323] 65. 根据方面52至方面64中的任何一个所述的系统,其中,所述输入角度数据和输出角度数据表示数字角色的运动的旋转分量。

[0324] 66. 根据方面52至方面65中的任何一个所述的系统,其中,所述输入线性数据和输出线性数据表示数字角色的运动的平移分量。

[0325] 67. 一种方法,包括:接收到径向基函数神经网络RBFNN的至少一个输入参数;根据所述RBFNN中的一个或多个距离函数计算一组距离;并且将一组距离权重应用到根据所述RBFNN中的所述一个或多个距离函数所计算的该组距离,其中,该组距离权重包括一个或多个距离权重,其中,该组权重包括到所述RBFNN的每个输入参数的距离权重。

[0326] 68. 根据方面67所述的方法,其中,该组距离权重中的至少一个距离权重由用户定义。

[0327] 69. 根据方面67或方面68所述的方法,其中,该组距离权重中的至少一个距离权重由RBFNN定义。

[0328] 70. 根据方面67至方面69中的任何一个所述的方法,还包括:将所述输入参数分离为第一输入数据类型和第二输入数据类型。

[0329] 71. 根据方面70所述的方法,其中,所述第一输入数据类型是表示线性参数的数据。

[0330] 72. 根据方面70或方面71所述的方法,其中,所述第二输入数据类型是表示角度参数的数据。

[0331] 73. 根据方面67至方面72中的任何一个所述的方法,其中,所述一个或多个距离函数包括至少一个角度距离函数。

[0332] 74. 根据方面73所述的方法,其中,所述角度距离函数是处于三维($S^0(3)$)空间中的特殊正交群。

[0333] 75. 根据方面73或方面74所述的方法,其中,所述角度距离函数在四元数空间中。

[0334] 76. 根据方面67至方面75中任何一个所述的方法,其中,所述一个或多个距离函数包括至少一个线性距离函数。

[0335] 77. 根据方面76所述的方法,其中,所述线性距离函数是欧氏距离函数。

[0336] 78. 根据方面67至方面77中任何一个所述的方法,还包括:将所述至少一个输入参数分离为第一输入数据类型和第二输入数据类型。

[0337] 79. 根据方面78所述的方法,其中,所述第一输入数据类型是表示线性参数的数据。

[0338] 80. 根据方面78或方面79所述的方法, 其中, 所述第二输入数据类型是表示角度参数的数据。

[0339] 81. 根据方面80所述的方法, 还包括: 使用所述第二输入数据类型和所述角度距离函数计算角度距离, 以及使用所述第一输入数据类型和所述线性距离函数计算线性距离。

[0340] 82. 根据方面67至方面81中任何一个所述的方法, 其中, 所述RBFNN被配置为对数字角色的姿势进行插值。

[0341] 83. 根据方面67到方面82中任何一个所述的方法, 其中, 所述一个或多个输入参数表示低阶数字角色装配。

[0342] 84. 根据方面67至方面83中的任何一个所述的方法, 其中, 所述RBFNN被配置为计算表示高阶数字角色装配的输出参数。

[0343] 85. 根据方面67至方面84中任何一个所述的方法, 其中, 所述RBFNN被配置为计算具有独立值的多个输出参数。

[0344] 86. 根据方面85所述的方法, 还包括: 将多个独立值中的至少两个分组为组合值。

[0345] 87. 根据方面86所述的方法, 其中, 所述组合值表示包括x、y、z和w的四元数向量。

[0346] 88. 一种系统, 包括: 径向基函数神经网络RBFNN, 其在一个或多个计算机上实现; 以及一个或多个存储设备, 其存储指令, 当由所述一个或多个计算机执行时, 所述指令可操作以使所述一个或多个计算机执行: 接收到所述RBFNN的一组输入参数中的至少一个输入参数; 根据所述RBFNN中的一个或多个距离函数计算一组距离; 并且将一组距离权重应用到根据所述RBFNN中的所述一个或多个距离所计算的该组距离, 其中, 该组距离权重包括到所述RBFNN的每个输入参数的距离权重。

[0347] 89. 根据方面88所述的系统, 其中, 该组距离权重中的至少一个距离权重由用户定义。

[0348] 90. 根据方面88或方面89所述的系统, 其中, 该组距离权重中的至少一个距离权重由所述RBFNN定义。

[0349] 91. 根据方面88至方面90中的任何一个所述的系统, 还包括: 将所述输入参数分离为第一输入数据类型和第二输入数据类型。

[0350] 92. 根据方面91所述的系统, 其中, 所述第一输入数据类型是表示线性参数的数据。

[0351] 93. 根据方面91或方面92所述的系统, 其中, 所述第二输入数据类型是表示角度参数的数据。

[0352] 94. 根据方面88至方面93中的任何一个所述的系统, 其中, 所述一个或多个距离函数包括至少一个角度距离函数。

[0353] 95. 根据方面94所述的系统, 其中, 所述角度距离函数是处于三维(SO(3))空间中的特殊正交群。

[0354] 96. 根据方面94或方面95所述的系统, 其中, 所述角度距离函数在四元数空间中。

[0355] 97. 根据方面88至方面96中任何一个所述的系统, 其中, 所述一个或多个距离函数包括至少一个线性距离函数。

[0356] 98. 根据方面97所述的系统, 其中, 所述线性距离函数是欧氏距离函数。

[0357] 99. 根据方面88至方面98中任何一个所述的系统, 还包括: 将所述输入参数分离为

第一输入数据类型和第二输入数据类型。

[0358] 100.根据方面99所述的系统,其中,所述第一输入数据类型是表示线性参数的数据。

[0359] 101.根据方面99或方面100所述的系统,其中,所述第二输入数据类型是表示角度参数的数据。

[0360] 102.根据方面99至方面101中任何一个所述的系统,还包括:使用所述第二输入数据类型和所述角度距离函数计算角度距离,以及使用所述第一输入数据类型和所述线性距离函数计算线性距离。

[0361] 103.根据方面88至方面102中任何一个所述的系统,其中,所述RBFNN被配置为对数字角色的姿势进行插值。

[0362] 104.根据方面88到方面103中任何一个所述的系统,其中,所述一个或多个输入参数表示低阶数字角色装配。

[0363] 105.根据方面88至方面104中的任何一个所述的系统,其中,所述RBFNN被配置为计算表示高阶数字角色装配的输出参数。

[0364] 106.根据方面88至方面105中任何一个所述的系统,其中,所述RBFNN被配置为计算具有独立值的多个输出参数。

[0365] 107.根据方面106所述的系统,还包括:将多个独立值中的至少两个分组为组合值。

[0366] 108.根据方面107所述的系统,其中,所述组合值表示包括x、y、z和w的四元数向量。

[0367] 其它注意事项

[0368] 在此描述和/或附图中描绘的过程、方法和算法中的每一个可以在由一个或多个物理计算系统、硬件计算机处理器、专用电路和/或配置为执行专用和特定计算机指令的电子硬件执行的代码模块中体现,并且完全或部分地由代码模块自动执行。例如,计算系统可以包括采用特定计算机指令或专用计算机、专用电路等编程的通用计算机(例如服务器)。代码模块可以被编译并链接到可执行程序中,安装在动态链接库中,或者可以用解释的编程语言编写。在一些实现方式中,特定操作和方法可以由特定于给定功能的电路执行。

[0369] 此外,本公开的功能的某些实现方式在数学上、计算上或技术上足够复杂,以使得专用硬件或一个或多个物理计算设备(利用适当的专用可执行指令)可能有必要例如由于所涉及的计算的量或复杂性而执行功能或者基本上实时地提供结果。例如,动画或视频可以包括许多帧,每个帧具有数百万个像素,并且需要专门编程的计算机硬件来处理视频数据,以在商业上合理的时间量内提供期望的图像处理任务、神经网络任务或应用。

[0370] 代码模块或任何类型的数据可以被存储在任何类型的非暂态计算机可读介质上,诸如包括硬盘驱动器、固态存储器、随机存取存储器(RAM)、只读存储器(ROM)、光盘、易失性或非易失性存储设备,与它们相同或类似物的组合等的物理计算机存储装置。方法和模块(或数据)还可以作为所生成的数据信号(例如,作为载波或其它模拟或数字传播信号的一部分)在各种计算机可读传输介质(包括基于无线的和有线的/基于电缆的介质)上发送,并且可以采用多种形式(例如,作为单个或多路复用模拟信号的一部分,或者作为多个离散数字分组或帧)。所公开的过程或过程步骤的结果可以持久地或以其它方式存储在任何类型

的非暂态有形计算机存储装置中,或者可以经由计算机可读传输介质传送。

[0371] 在此描述和/或附图中描绘的流程图中的任何过程、框、状态、步骤或功能应当理解为可能表示代码模块、代码段或代码部分,其包括用于实施特定功能(例如,逻辑或算术)或过程中的步骤的一个或多个可执行指令。各种过程、块、状态、步骤或功能可以与在此提供的说明性示例组合、重新排列、添加、删除、修改或以其它方式改变。在一些实施例中,附加或不同的计算系统或代码模块可以执行在此描述的一些或全部功能。在此描述的方法和过程也不限于任何特定序列,并且与其相关的块、步骤或状态可以以适当的其它顺序执行,例如,以串行、并行或以一些其它方式。可以向所公开的示例实施例添加任务或事件或从中移除任务或事件。此外,在此描述的实现方式中的各种系统组件的分离是出于说明性目的,并且不应被理解为在所有实现方式中都需要这种分离。应当理解,所描述的程序部件、方法和系统通常可以一起集成在单个计算机产品中或打包成多个计算机产品。许多实现方式变化都是可能的。

[0372] 过程、方法和系统可以在网络(或分布式)计算环境中实现。网络环境包括企业范围的计算机网络、内联网、局域网(LAN)、广域网(WAN)、个人局域网(PAN)、云计算网络、众包计算网络、互联网和万维网。网络可以是有线或无线网络或任何其它类型的通信网络。

[0373] 本公开的系统和方法各自具有若干创新方面,其中没有一个单独地对在此公开的期望属性负责或要求。以上所述的各种特征和过程可以彼此独立地使用,或者可以以各种方式组合。所有可能的组合和子组合都旨在落入本公开的范围。对本领域技术人员来说,对本公开中描述的实现方式的各种修改是显而易见的,并且在不脱离本公开的精神或范围的情况下,在此定义的一般原理可以应用于其它实现方式。因此,权利要求不旨在限于在此所示的实现方式,而是与符合在此公开的本公开、原理和新颖特征的最宽范围相一致。

[0374] 在单独实现方式的上下文中在本说明书中描述的某些特征也可以在单个实现方式中组合实施。相反,在单个实现方式的上下文中描述的各种特征也可以单独地或以任何合适的子组合在多个实现方式中实现。此外,尽管特征可以在上面被描述为以某些组合起作用并且甚至最初如此要求保护,但是在一些情况下可以从组合中排除来自所要求保护的组合的一个或多个特征,并且所要求保护的组合可以用于子组合或子组合的变体。对于每个和所有实施例,没有单个特征或特征组是必需或必不可少的。

[0375] 除非另有说明,或者在所使用的上下文中以其它方式理解,否则在此使用的条件语言,诸如“能够”、“可以”、“可能”、“可”、“例如”等通常旨在传达某些实施例包括某些特征、元件和/或步骤,而其它实施例不包括某些特征、元件和/或步骤。因此,这种条件语言通常不旨在暗示对于一个或多个实施例以任何方式需要特征、元素和/或步骤,或者一个或多个实施例必须包括用于决定是否在任何特定实施例中包括或将要执行这些特征、元件和/或步骤(无论是否有作者输入或提示)的逻辑。术语“包含”、“包括”、“具有”等是同义的并且以开放式方式包含使用,并且不排除附加元件、特征、动作、操作等。此外,术语“或”在其包含意义上(而不是在其独有意义上)使用,因此当使用时,例如,为了连接元素列表,术语“或”表示列表中的一个、一些或全部元素。另外,除非另有说明,否则本申请和所附权利要求中使用的冠词“一”、“一个”和“该”应理解为表示“一个或多个”或“至少一个”。

[0376] 如在此所使用的,指代项目列表中的“至少一个”的短语是指那些项目的任何组合,包括单个成员。例如,“A、B或C中的至少一个”旨在涵盖:A、B、C、A和B、A和C、B和C,以及A、

B和C。除非另外特别说明,否则诸如短语“X、Y和Z中的至少一个”的联合语言以其它方式通过上下文理解为通常用于传达项目、术语等可以是X、Y或Z中的至少一个。因此,这种联合语言通常不旨在暗示某些实施例需要X中的至少一个、Y中的至少一个和Z中的至少一个各自存在。

[0377] 类似地,虽然可以以特定顺序在附图中描绘操作,但应认识到,不需要以所示的特定顺序或按顺序执行这些操作,或者执行所有示出的操作,以实现期望的结果。此外,附图可以以流程图的形式示意性地描绘一个或多个示例过程。然而,未示出的其它操作可以包含在示意性示出的示例方法和过程中。例如,可以在任何所示操作之前、之后、同时或之间执行一个或多个附加操作。另外,可以在其它实现方式中重新排列或重新排序操作。在某些情况下,多任务处理和并行处理可能是有利的。此外,上述实现方式中的各种系统组件的分离不应被理解为在所有实现方式中都需要这种分离,并且应当理解,所描述的程序组件和系统通常可以在单个软件产品中集成在一起或者被打包到多种软件产品。另外,其它实现方式在以下权利要求的范围内。在一些情况下,权利要求中记载的动作可以以不同的顺序执行并且仍然实现期望的结果。

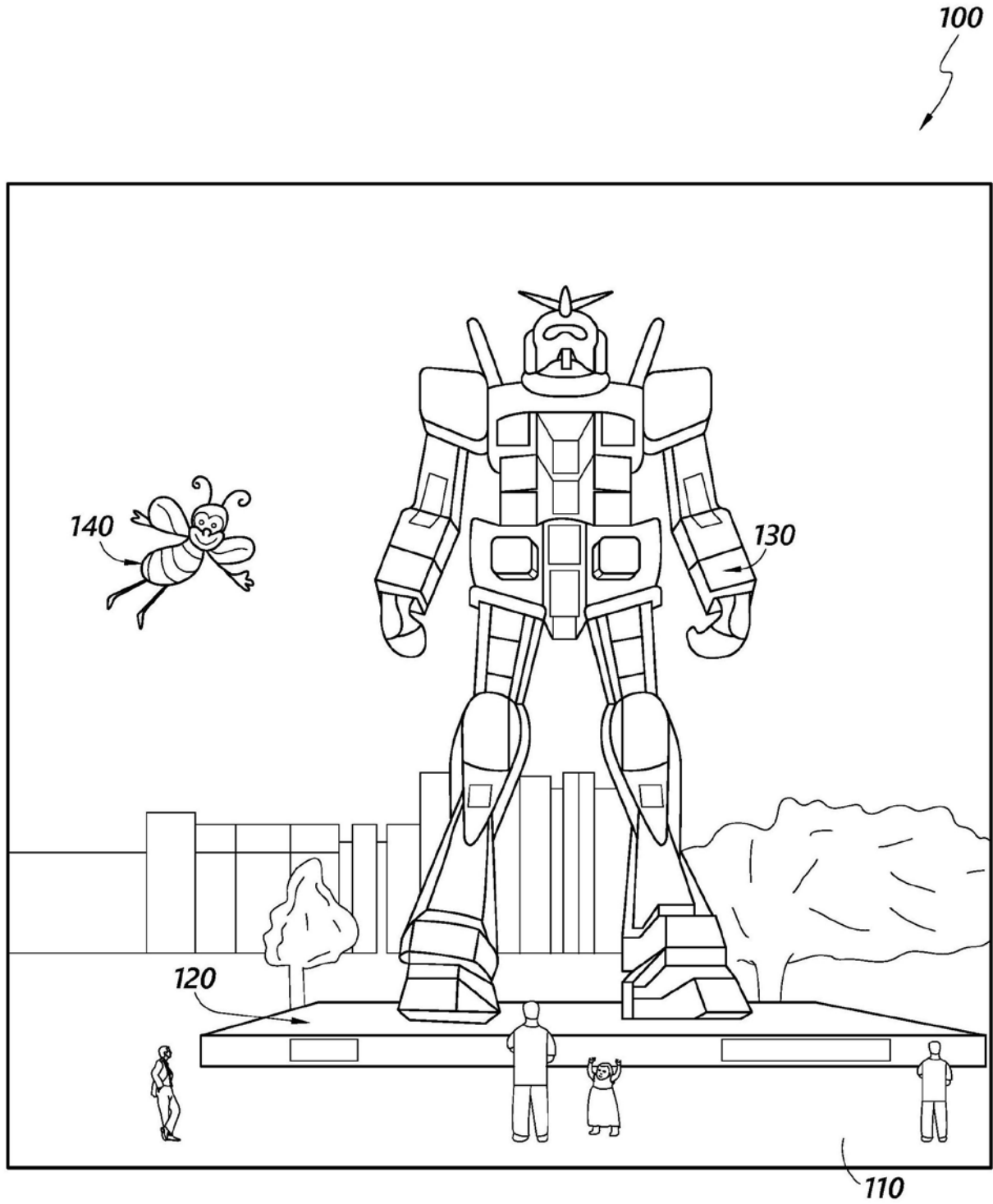


图1

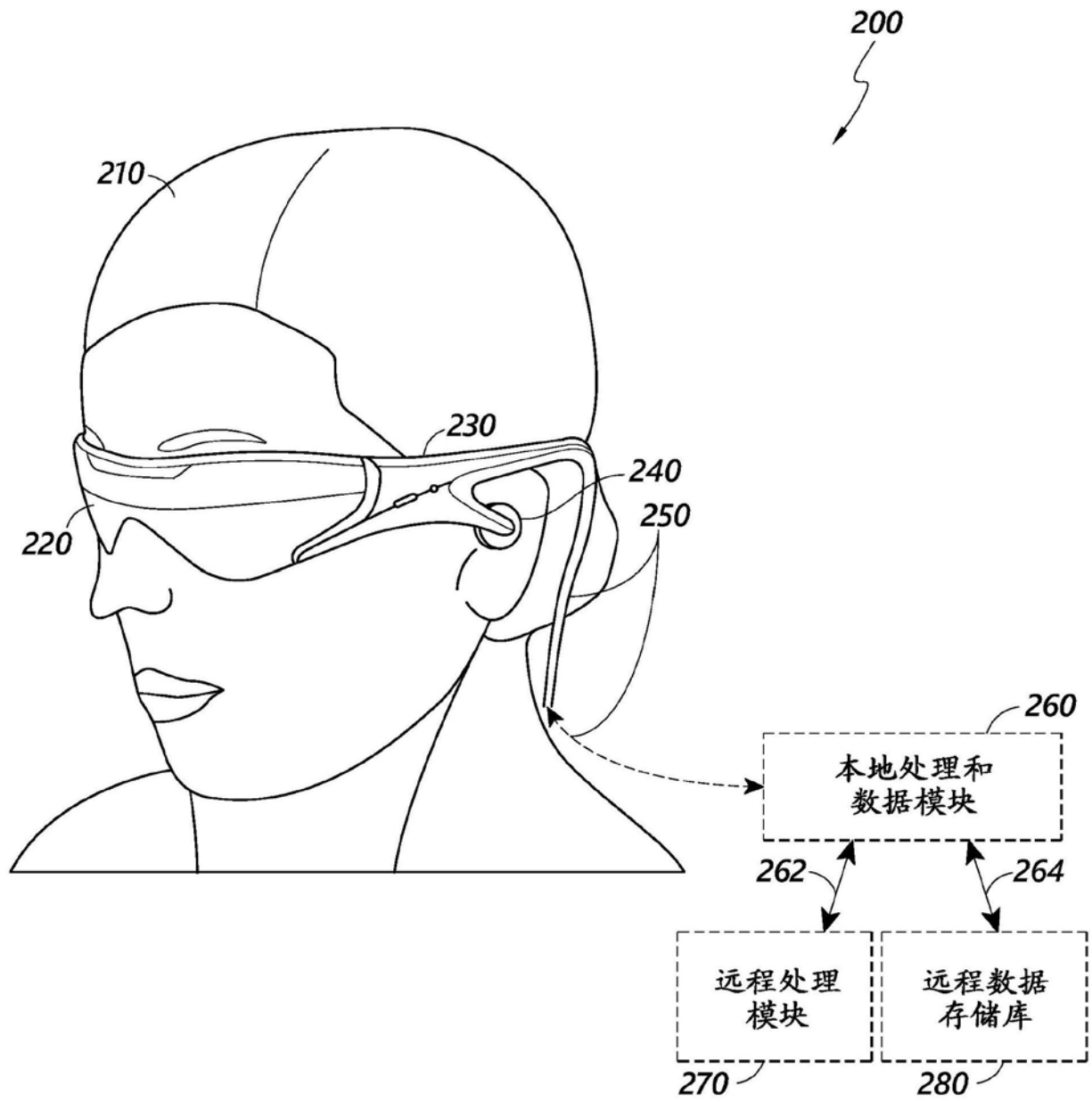


图2

可穿戴设备的示例传感器

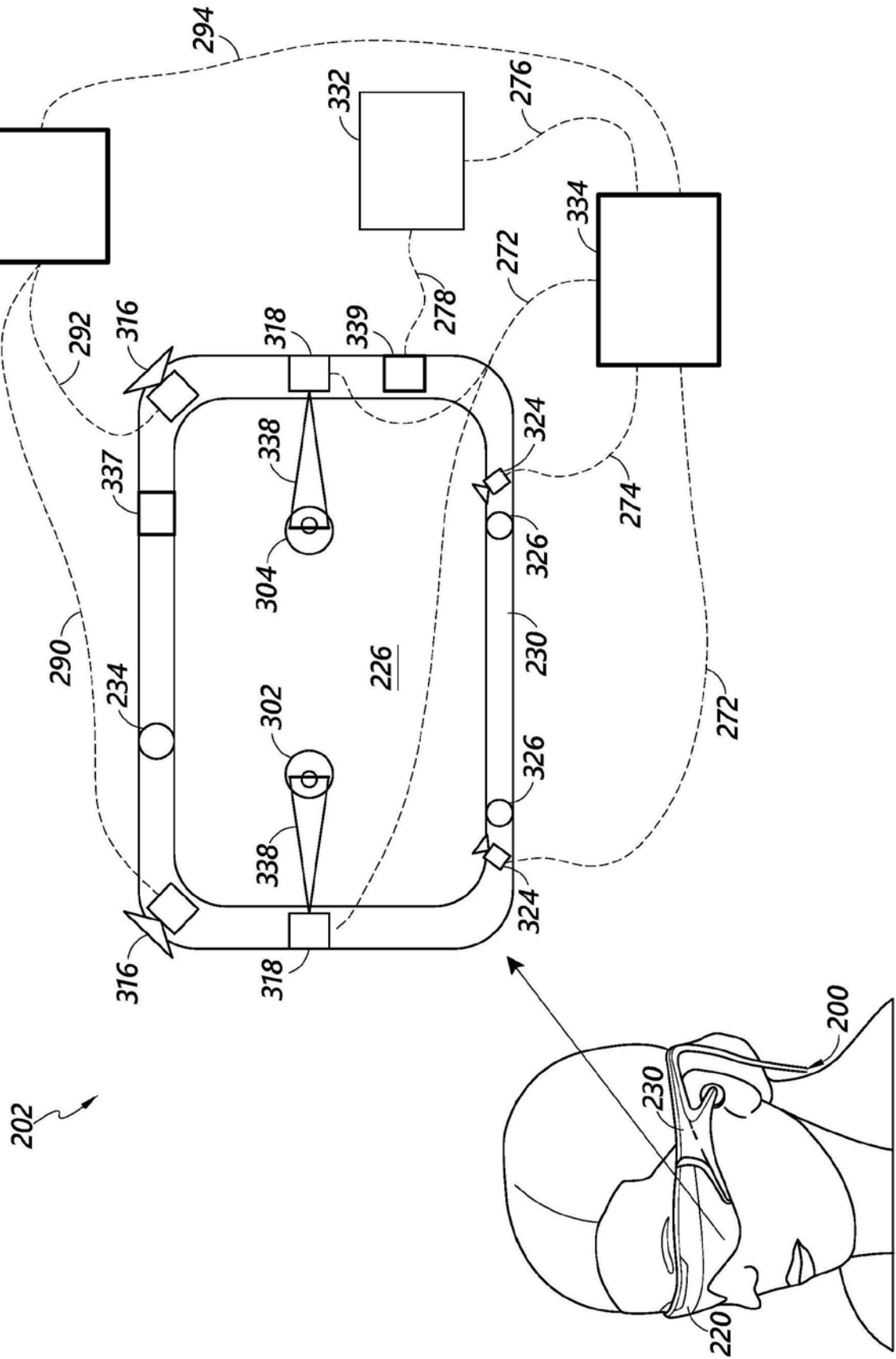


图3

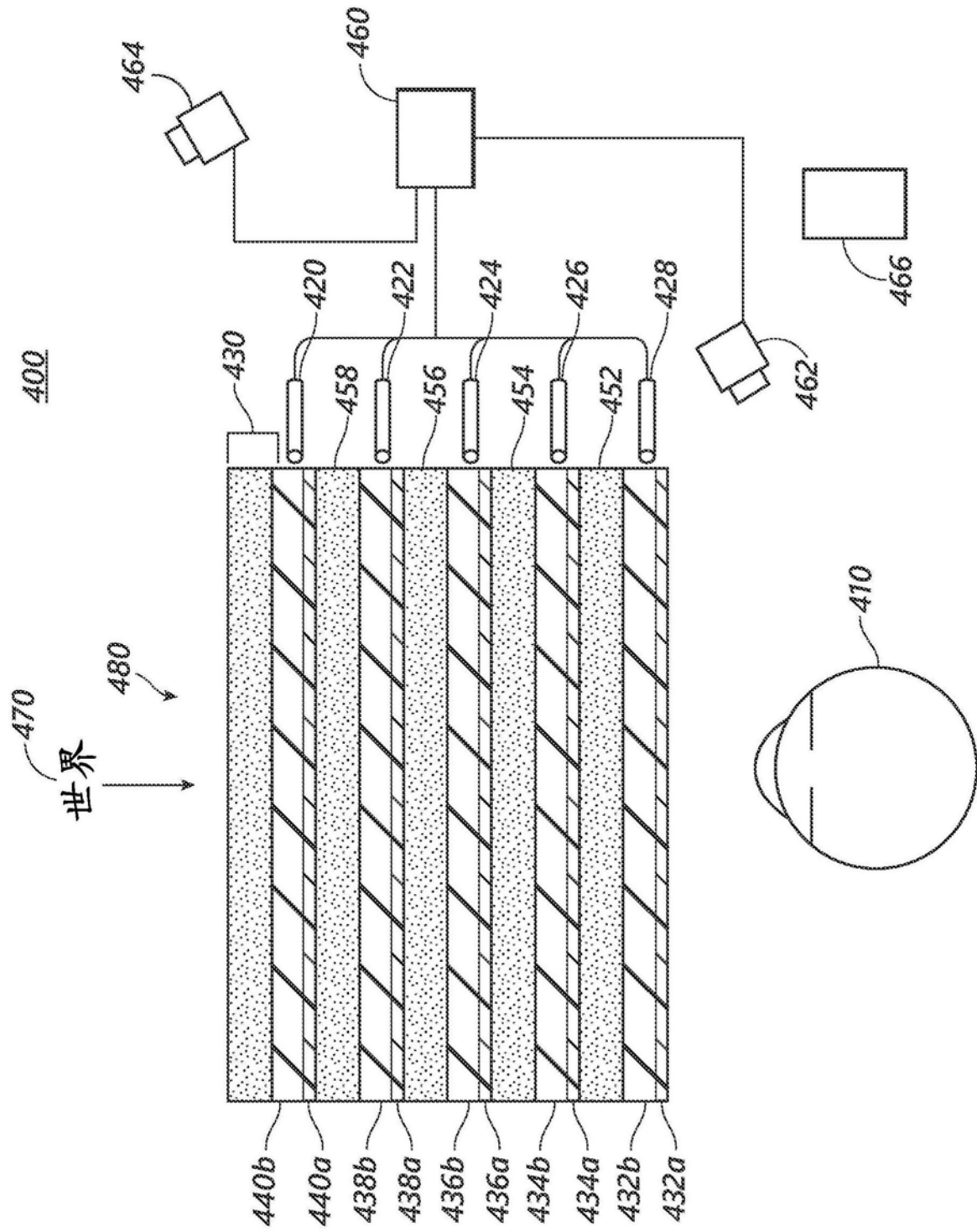


图4

用户与可穿戴设备的交互的示例

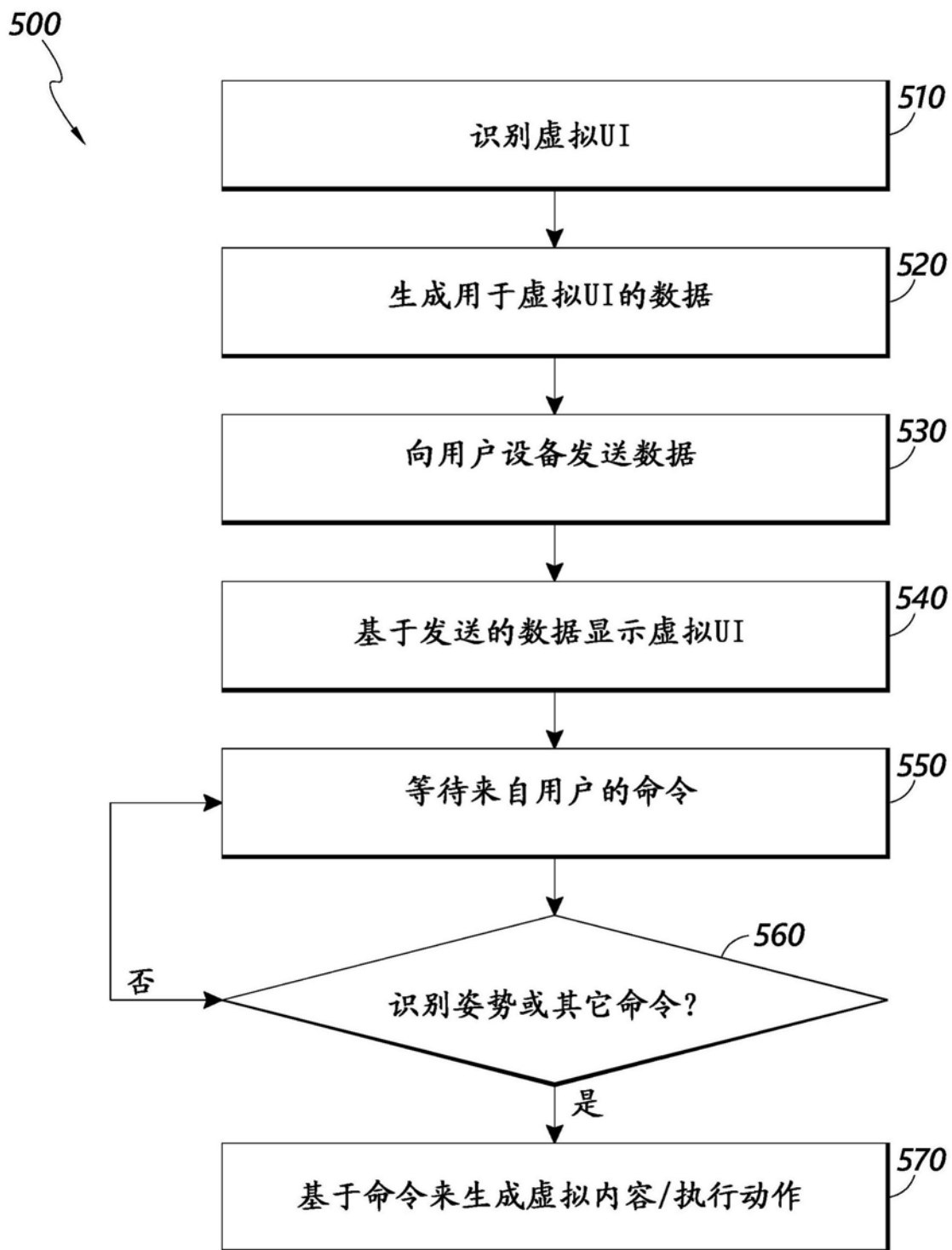


图5

示例混合现实过程

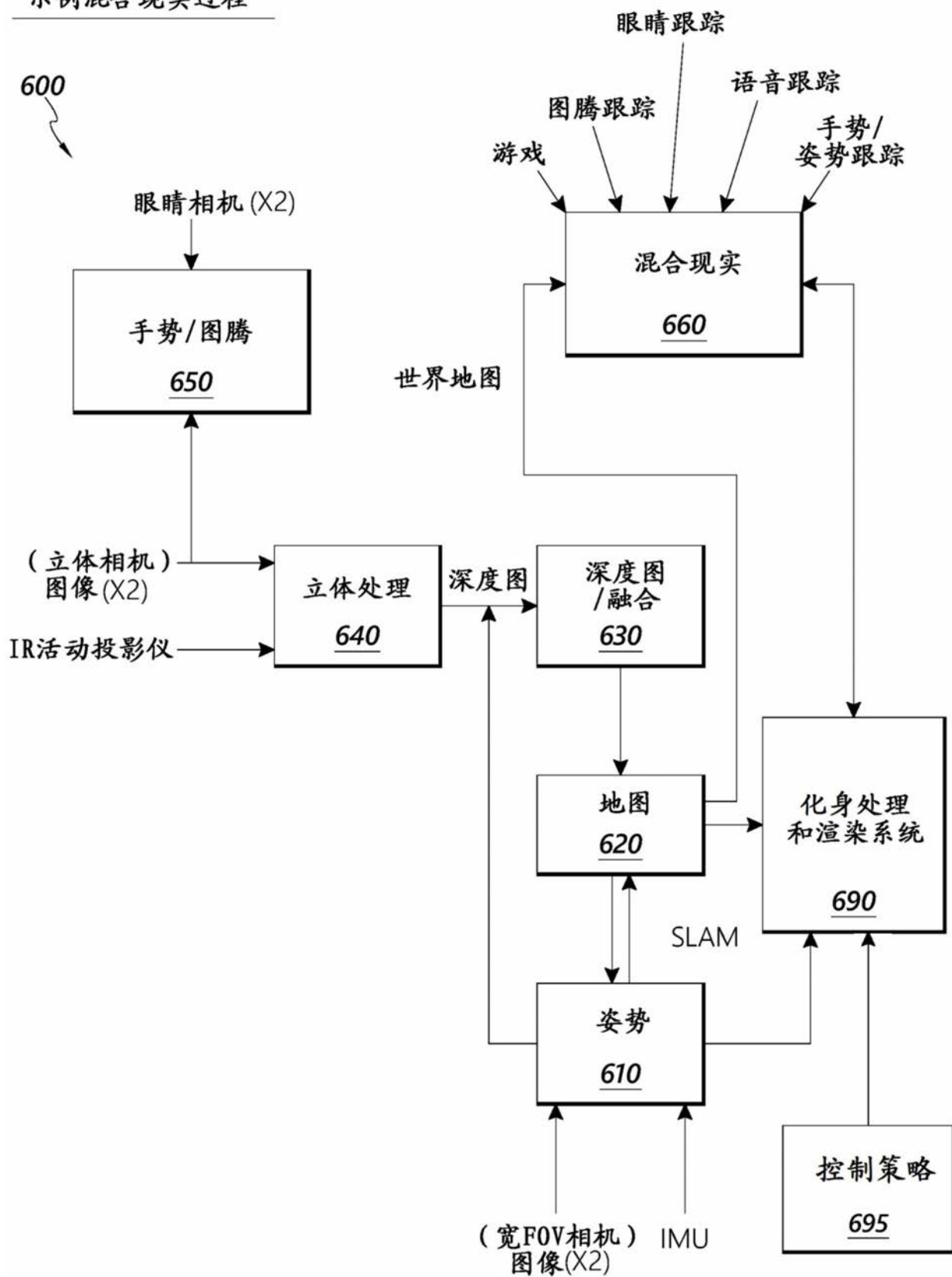


图6A

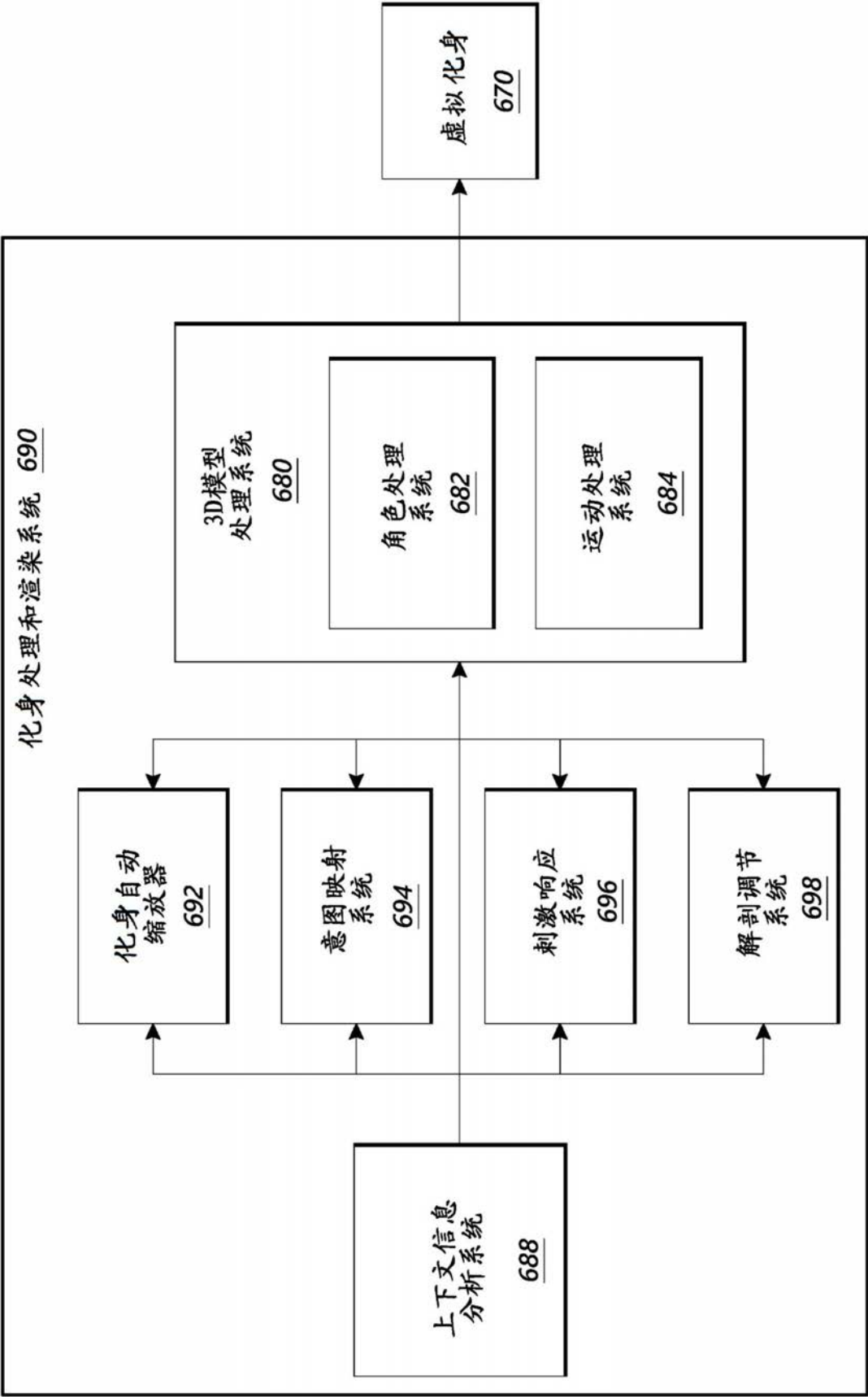


图6B

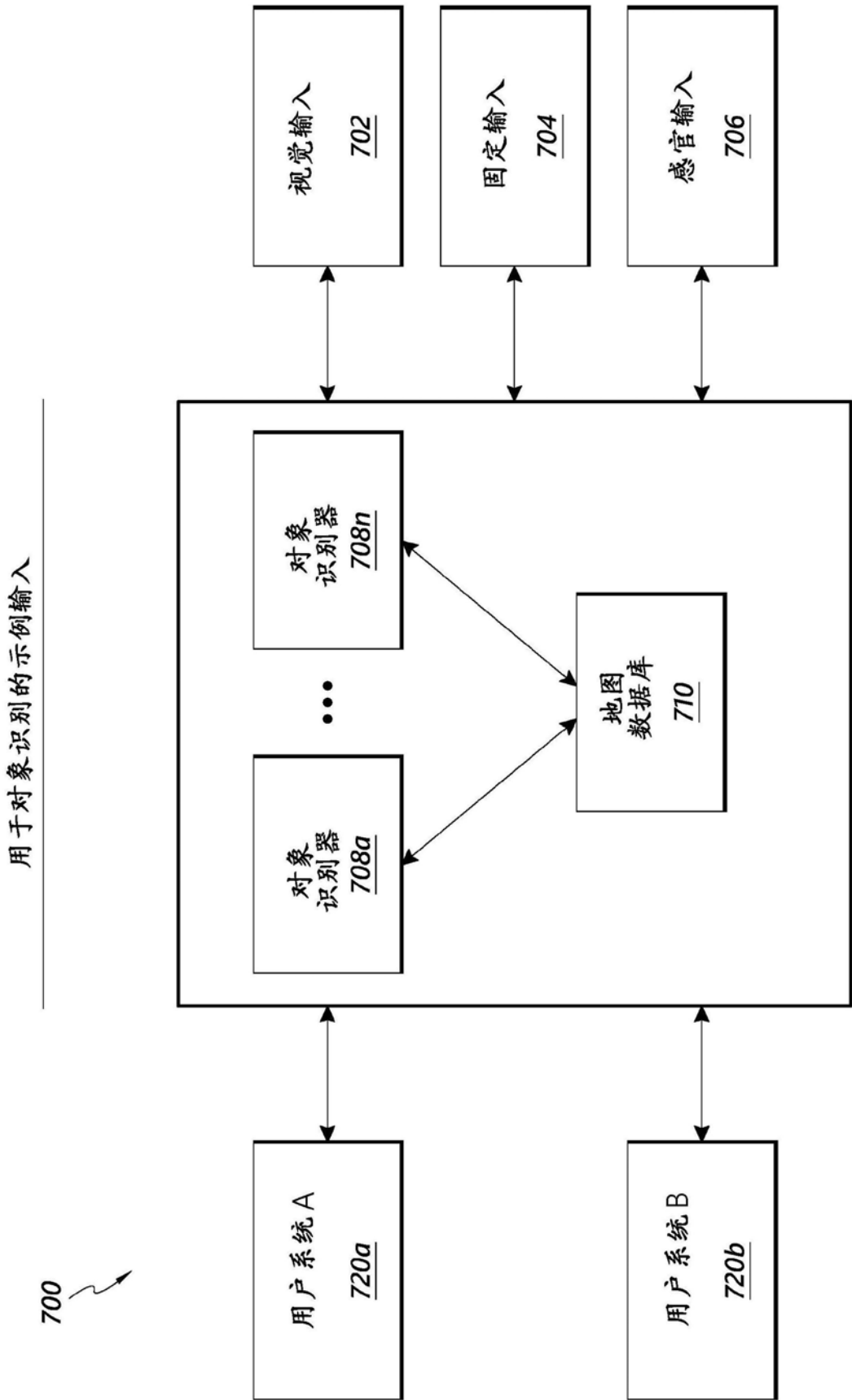


图7

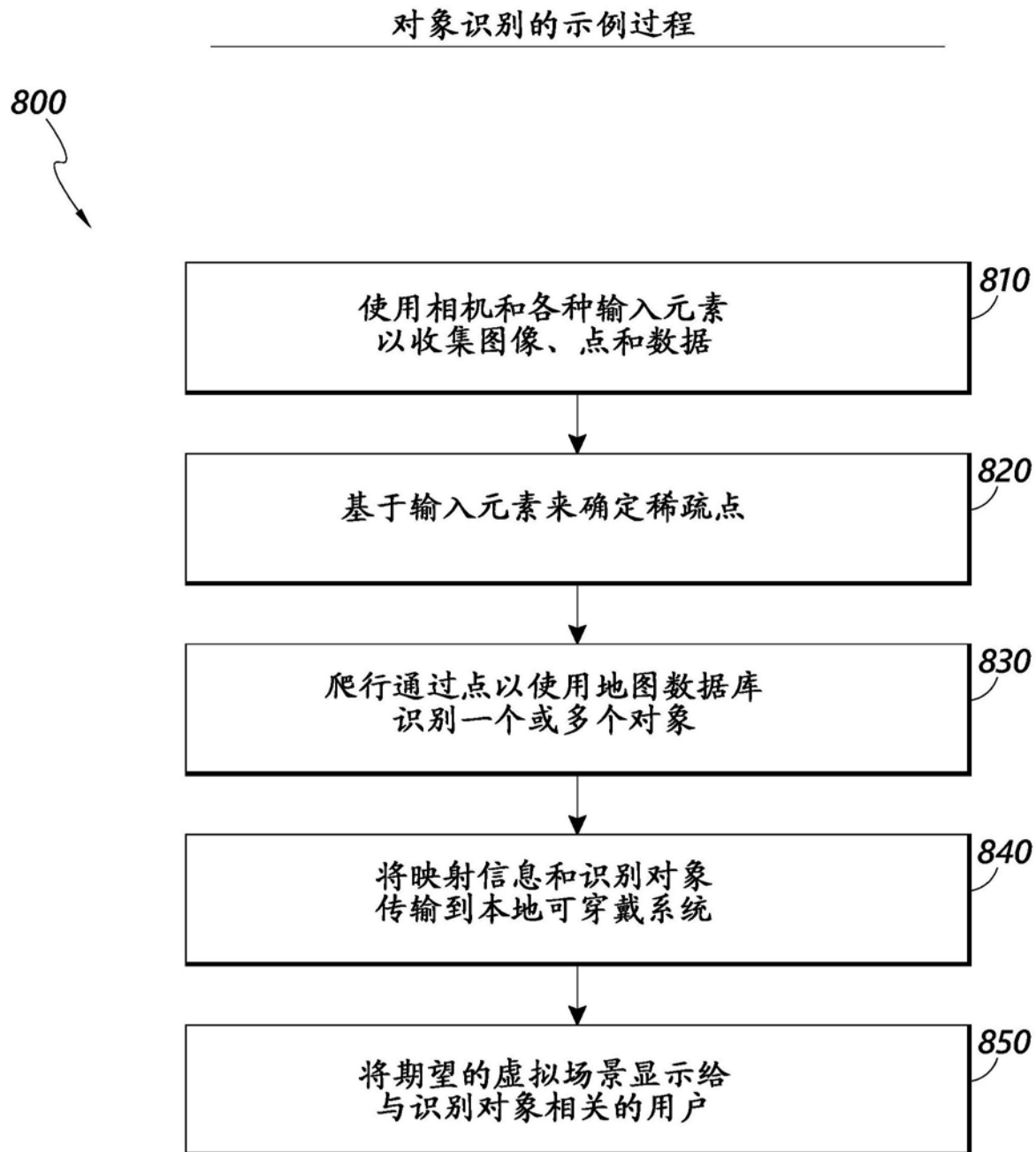


图8

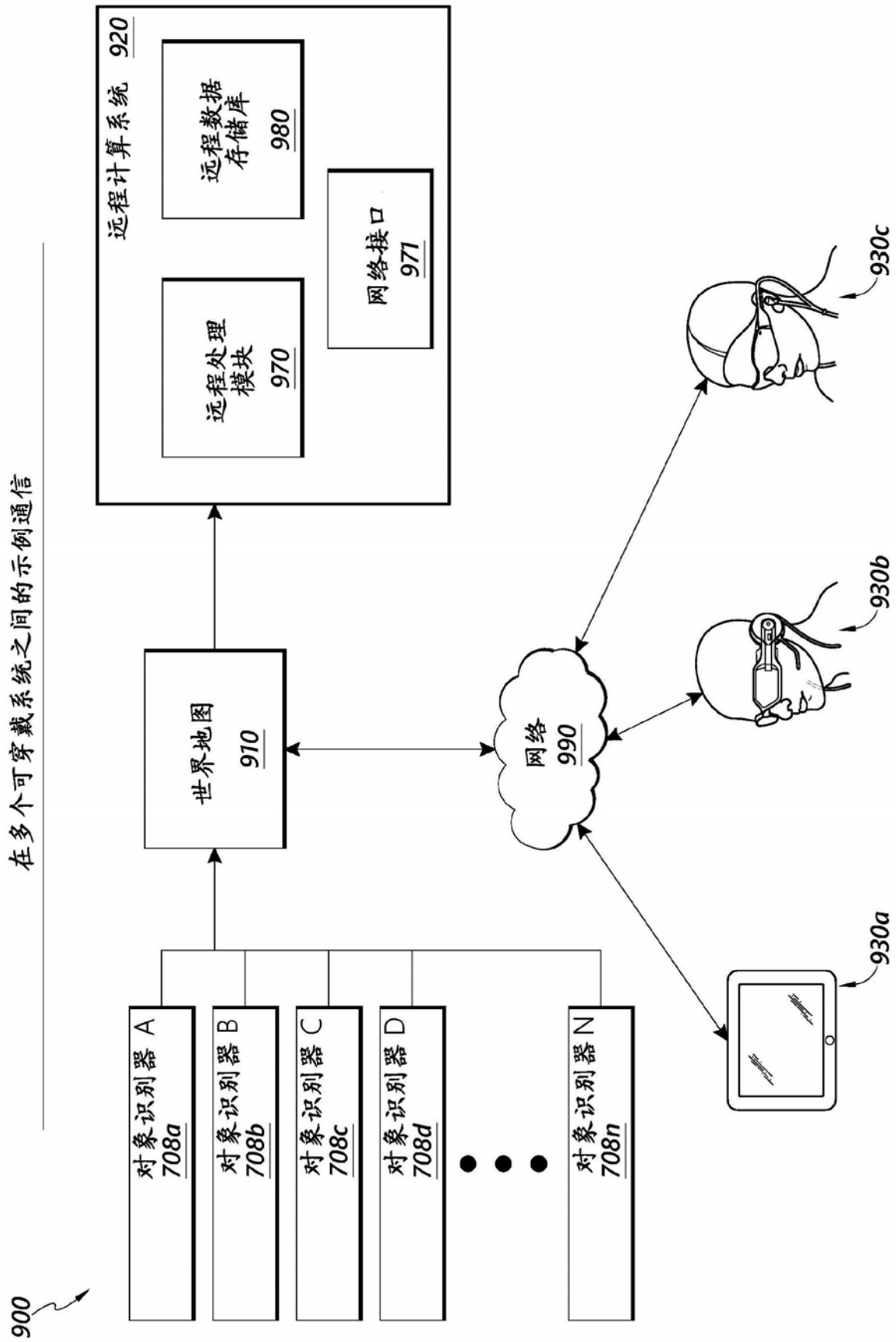


图9A

示例远程呈现会话

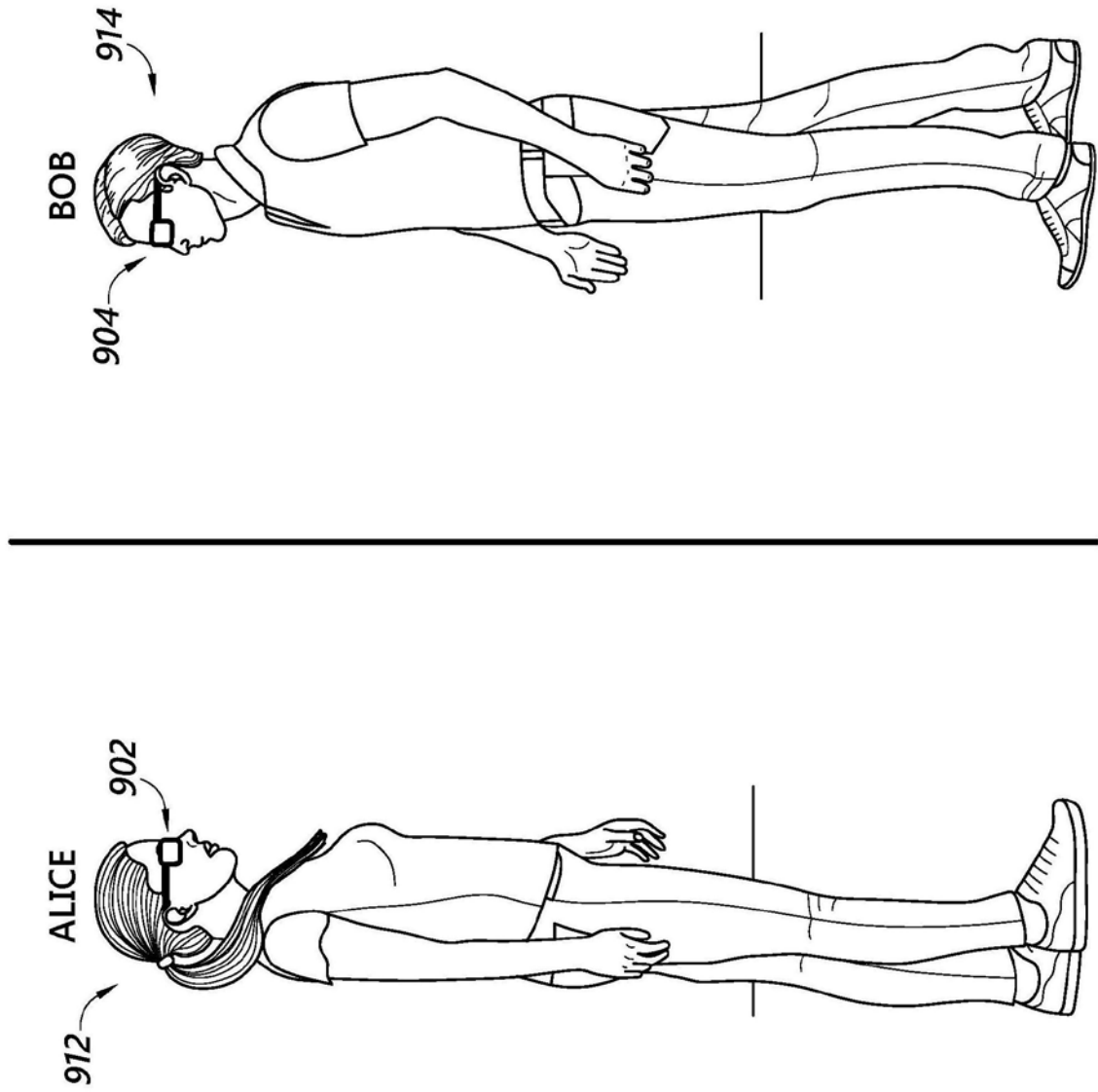


图9B

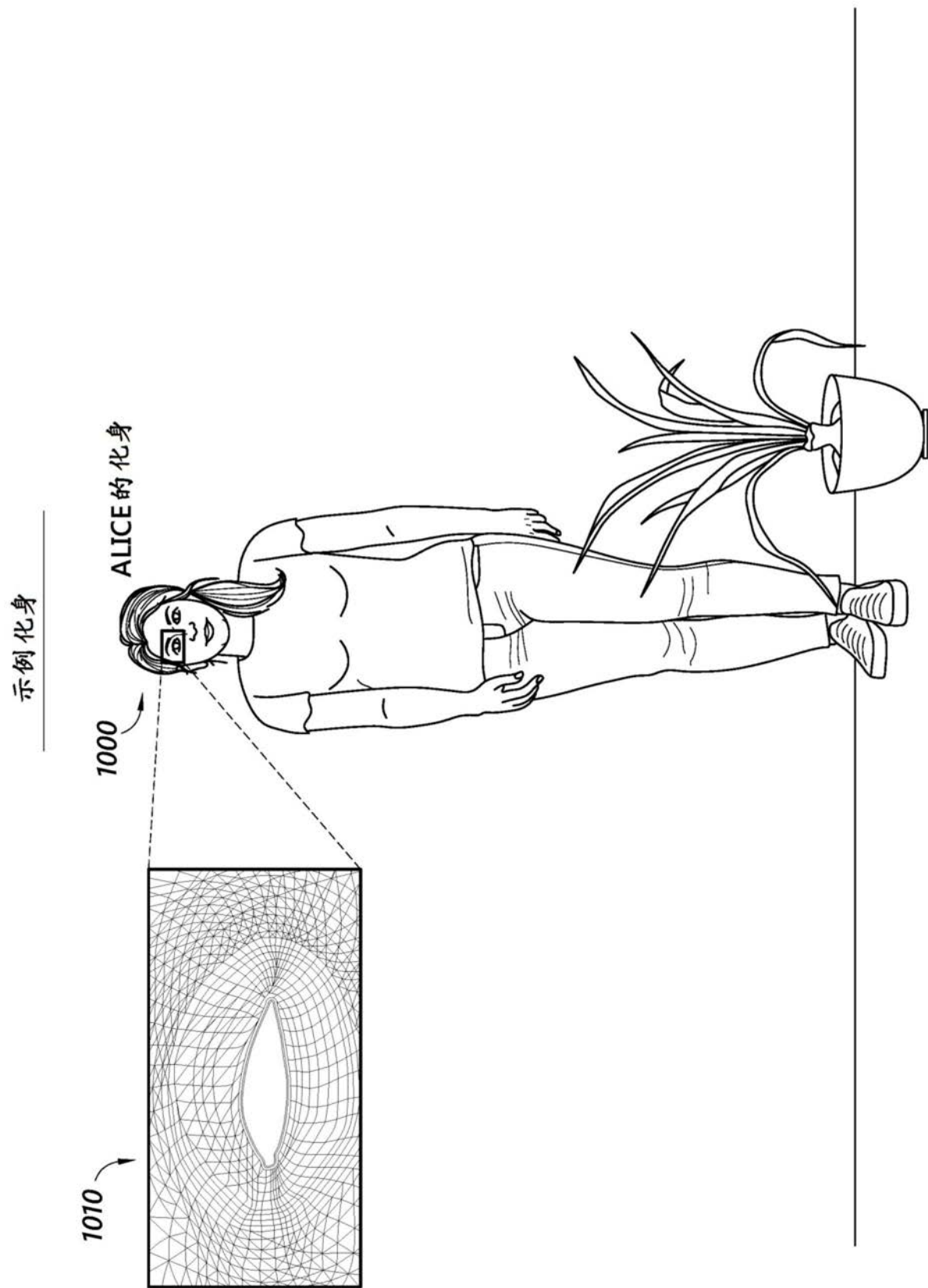


图10

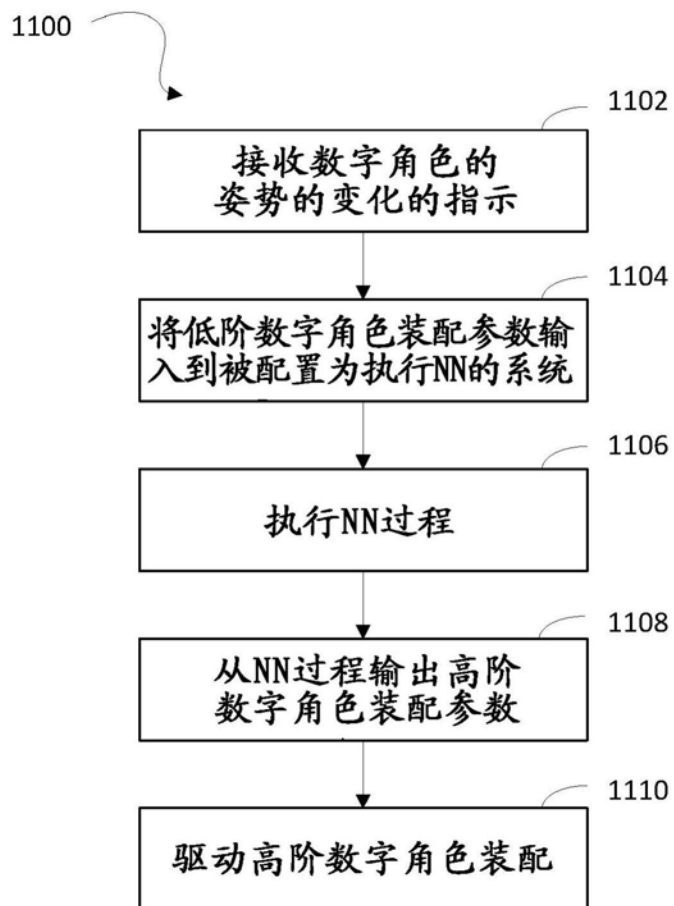


图11

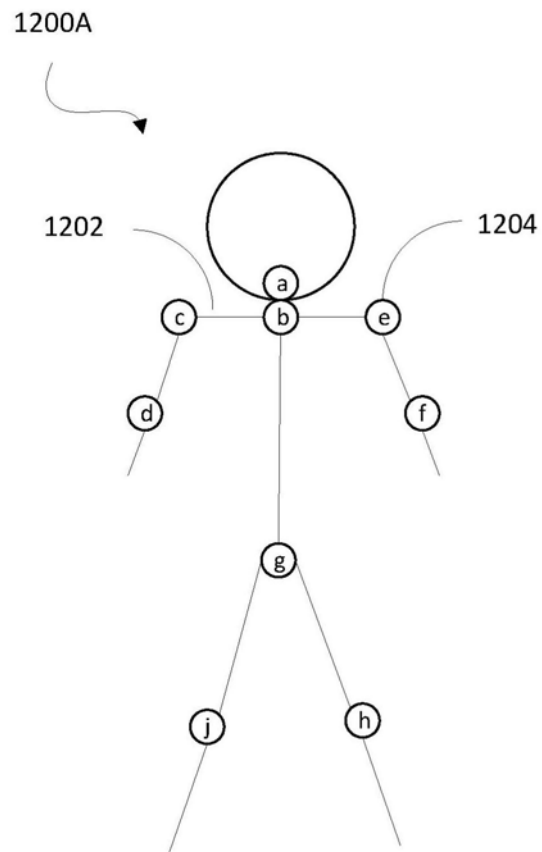


图12A

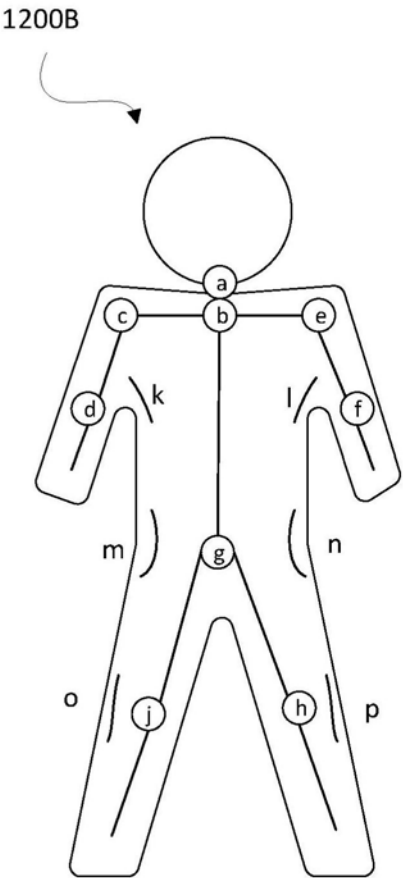
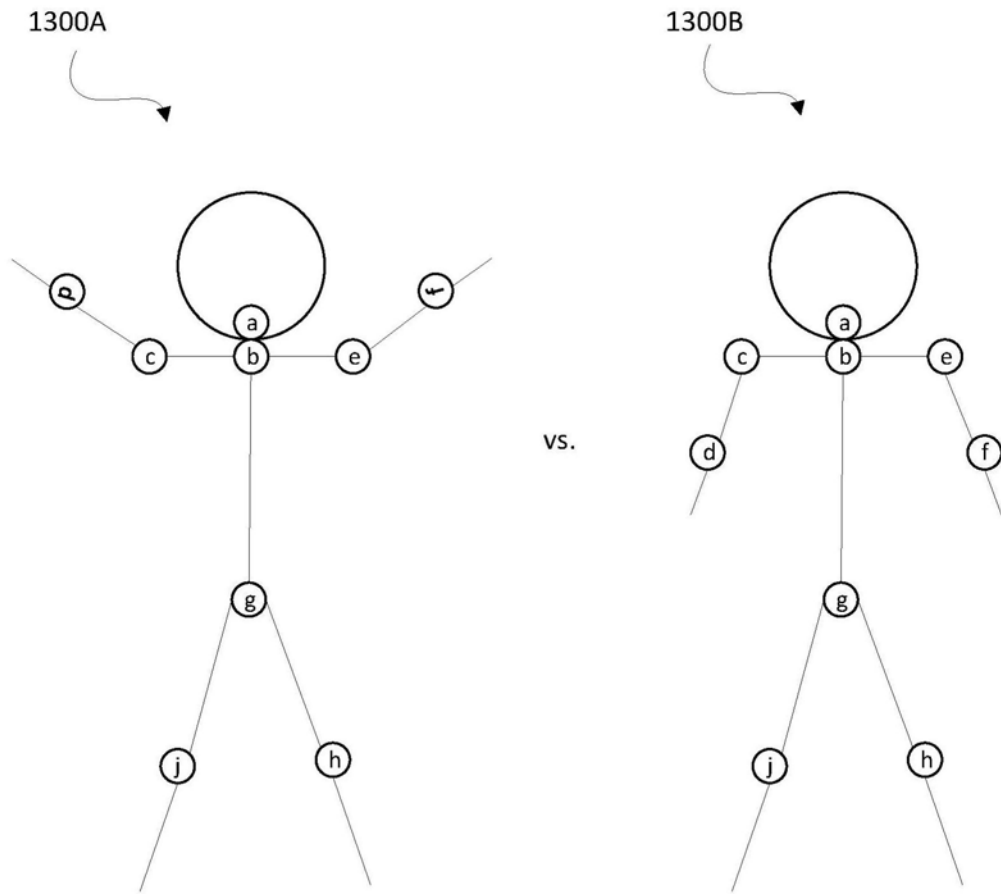


图12B



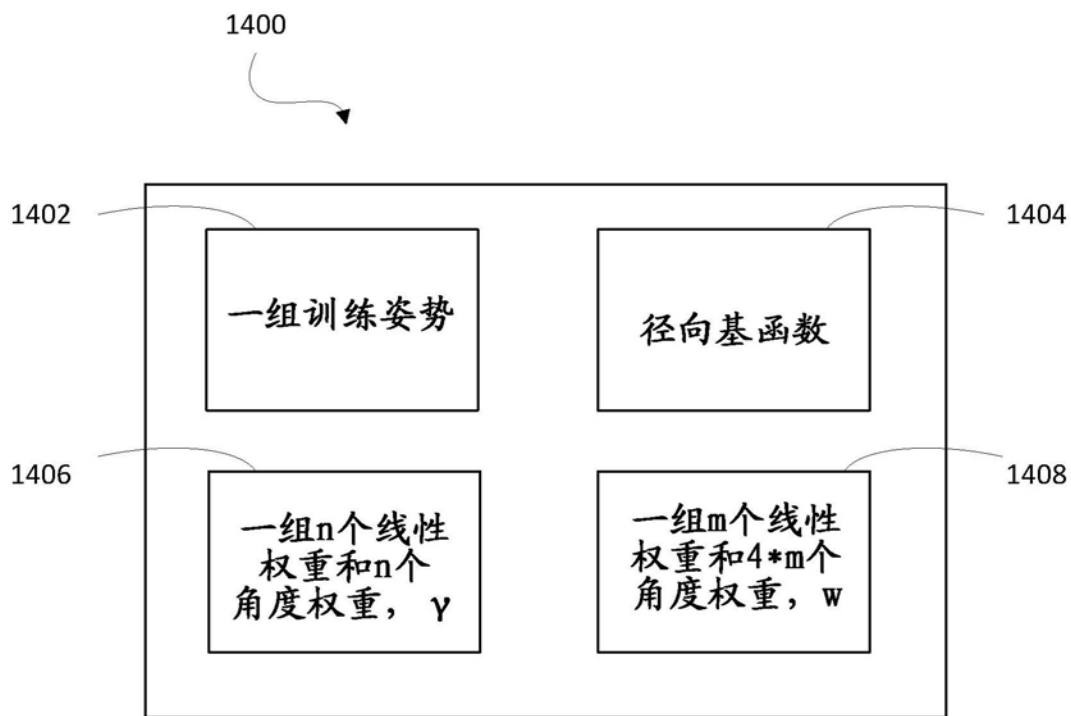


图14

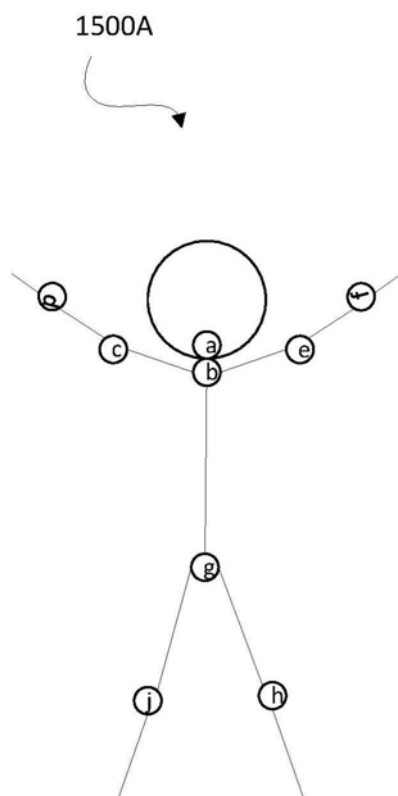


图15A

1500B

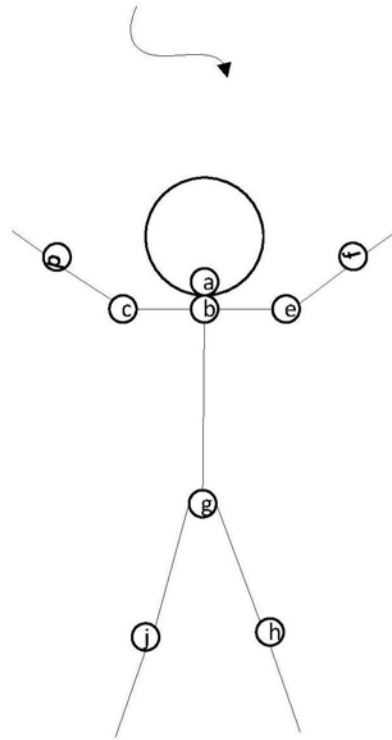


图15B

1500C

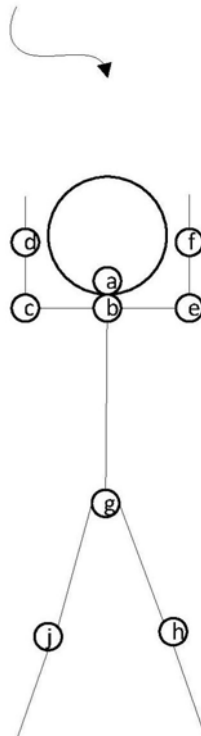


图15C

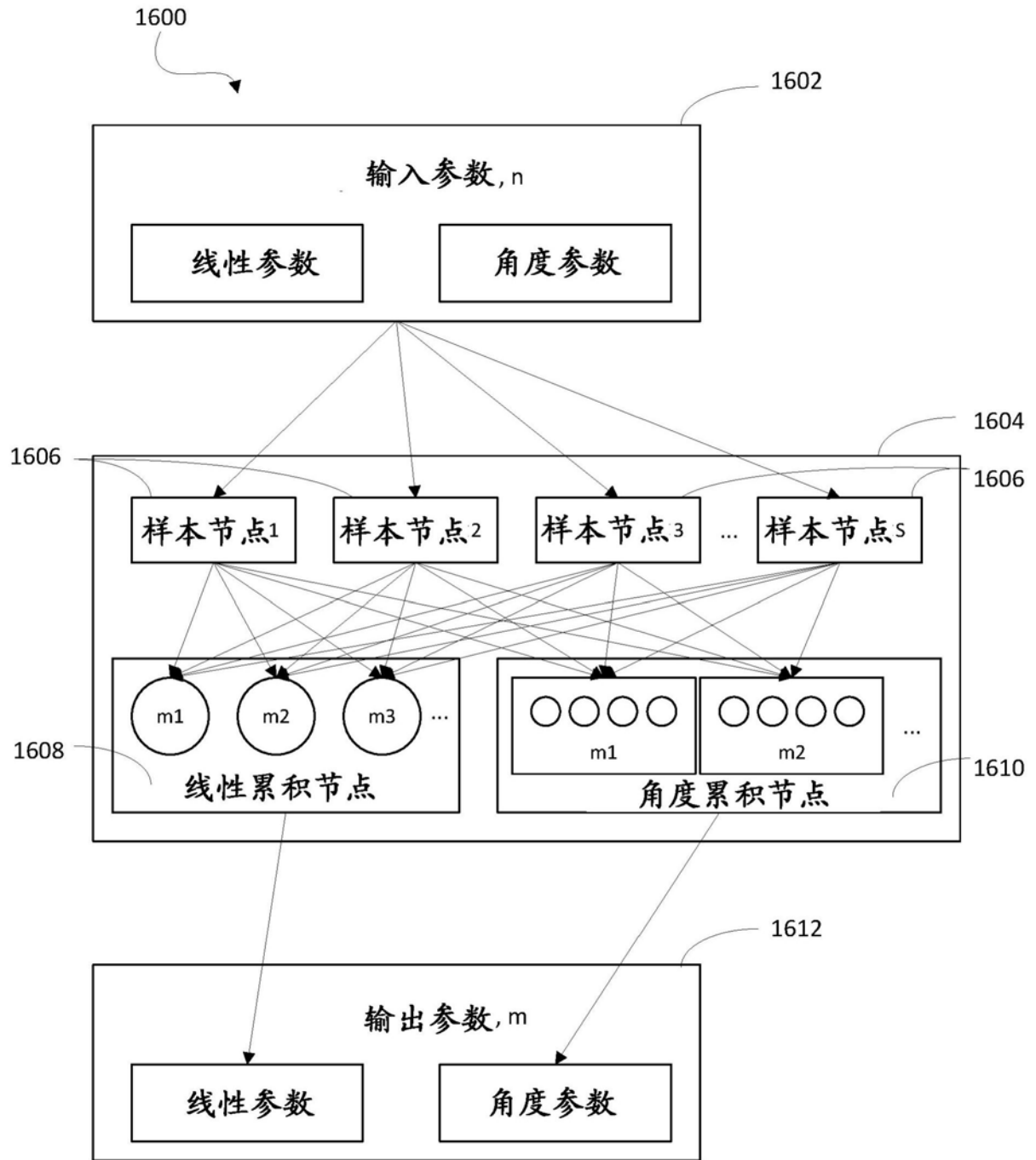


图16

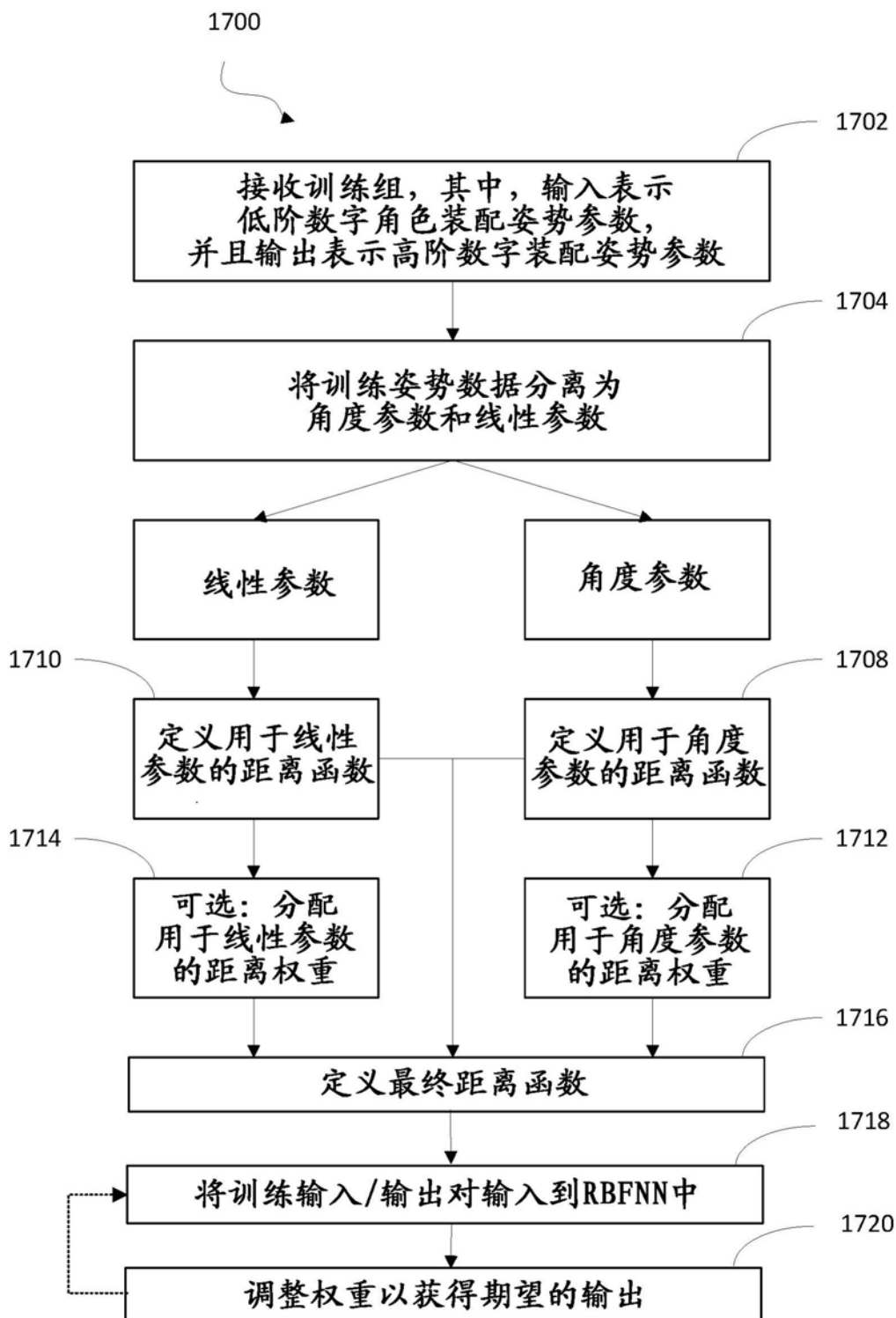


图17

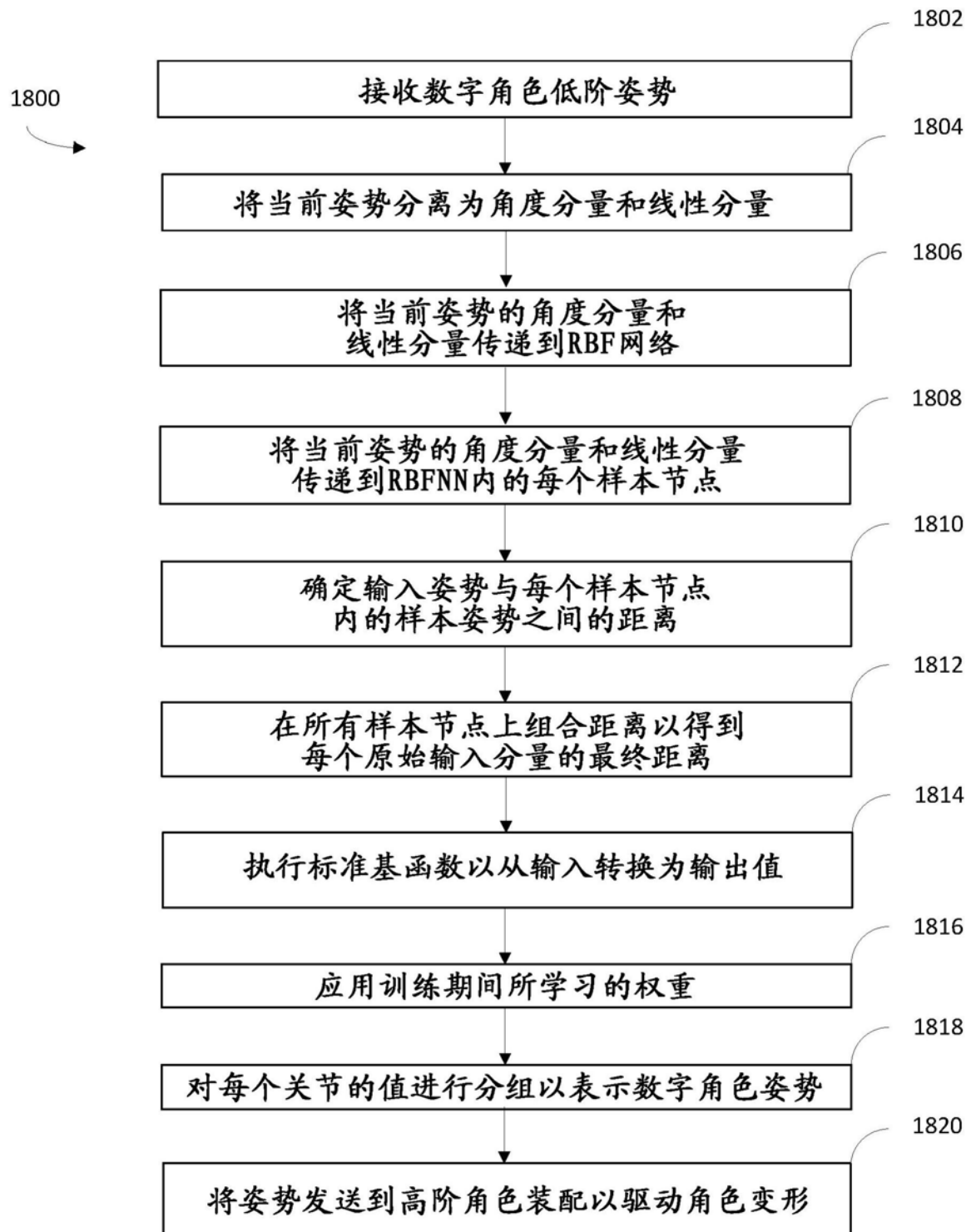


图18

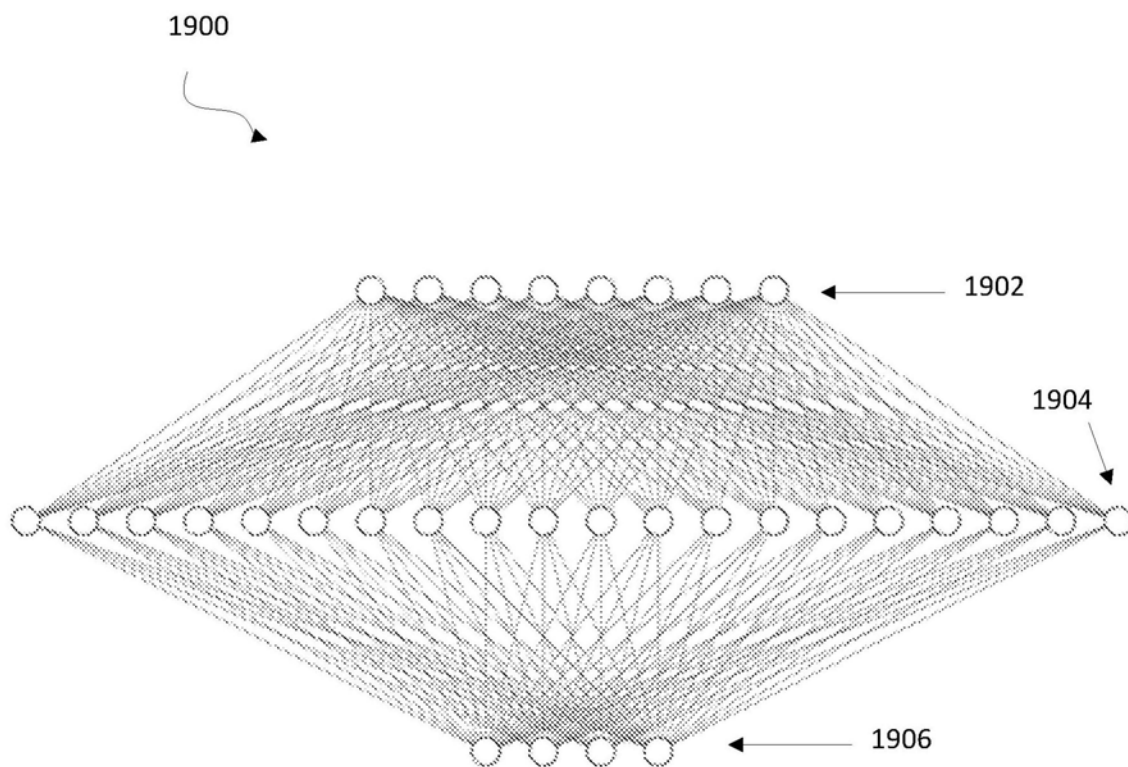


图19

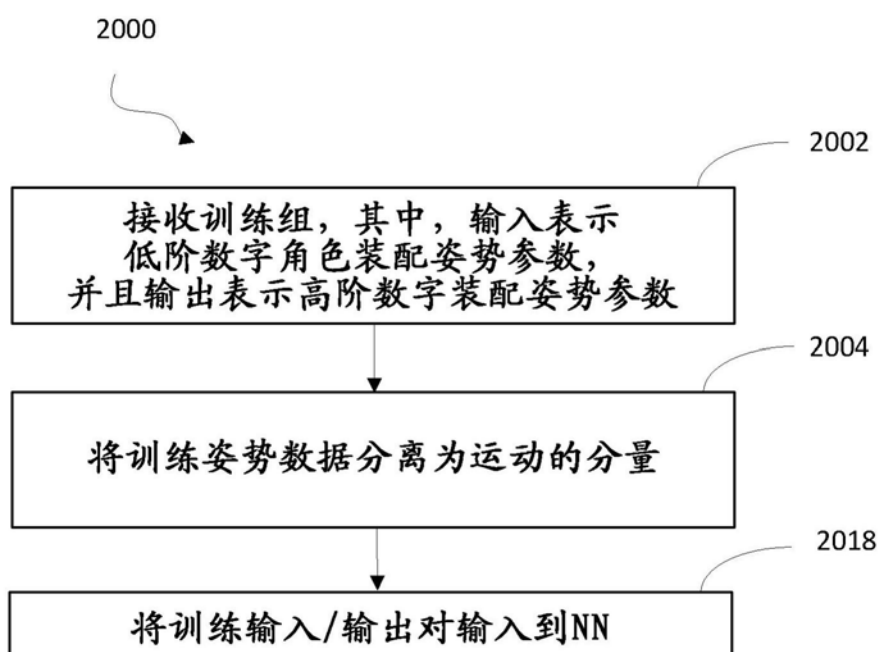


图20

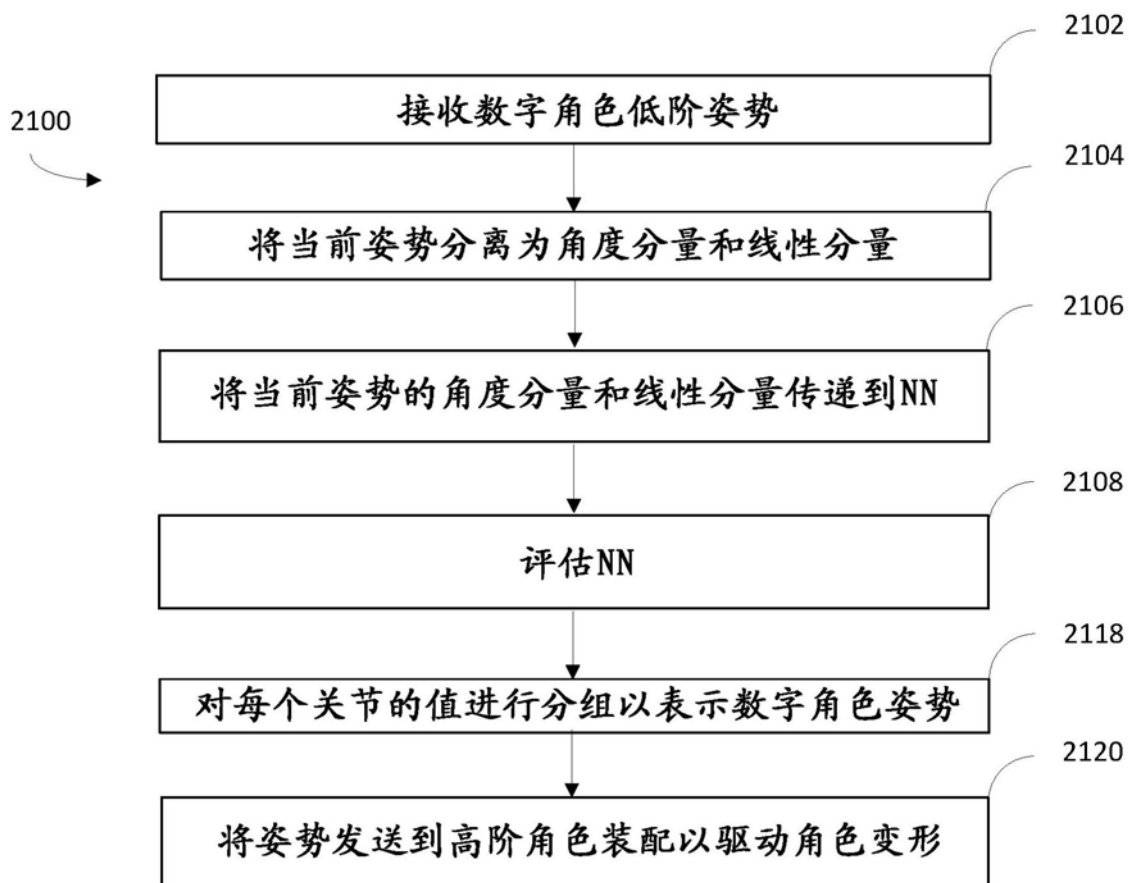


图21