

(12) 发明专利申请

(10) 申请公布号 CN 102665837 A

(43) 申请公布日 2012. 09. 12

(21) 申请号 201080055319. 4

代理人 胡利鸣

(22) 申请日 2010. 12. 06

(51) Int. Cl.

(30) 优先权数据

A63F 13/00 (2006. 01)

12/632, 599 2009. 12. 07 US

G06T 17/00 (2006. 01)

(85) PCT申请进入国家阶段日

2012. 06. 06

(86) PCT申请的申请数据

PCT/US2010/059094 2010. 12. 06

(87) PCT申请的公布数据

W02011/071815 EN 2011. 06. 16

(71) 申请人 微软公司

地址 美国华盛顿州

(72) 发明人 R·M·盖斯

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

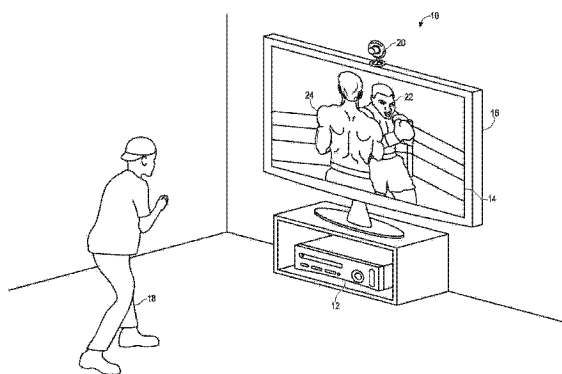
权利要求书 1 页 说明书 23 页 附图 22 页

(54) 发明名称

虚拟目标跟踪

(57) 摘要

一种用于跟踪目标的方法包括：利用一个或多个像素例对具有像素地址的像素进行分类。基于一个或多个所观察到的或合成的值对所述像素进行分类。像素地址的所观察到的值的一个示例包括从深度相机获得的所观察到的深度值。像素地址的合成值的示例包括：通过对所述目标的模型进行栅格化计算出的合成深度值；估计与该像素地址相对应的身体部位的一个或多个身体部位索引；以及估计与该像素地址相对应的目标的一个或多个玩家索引。基于像素例为像素计算一个或多个力矢量，并且将所述力矢量映射到表示所述目标的模型的一个或多个受力位置以将表示所述目标的模型调整为经调整的姿态。



1. 一种用于跟踪人类目标的方法,该方法包括:
从源接收所述人类目标的所观察到的深度图像,所述所观察到的深度图像包括多个所观察到的像素;
获得表示所述人类目标的模型的合成深度图像,所述合成深度图像包括多个合成像素;
用像素例对所述多个合成像素中的合成像素进行分类;
基于所述合成像素的像素例为所述合成像素计算力矢量;以及
将所述力矢量映射到表示所述人类目标的模型的一个或多个受力位置,以将表示所述人类目标的模型调整为经调整的姿态。
2. 如权利要求 1 所述的方法,其特征在于,至少部分通过将表示所述人类目标的模型栅格化来获得所述合成深度图像。
3. 如权利要求 1 所述的方法,其特征在于,至少基于所述合成像素的身体部位索引来用所述像素例对所述合成像素进行分类。
4. 如权利要求 3 所述的方法,其特征在于,至少部分基于所述合成像素的原始身体部位索引与所述合成像素的咬合身体部位索引或所述合成像素的增长身体部位索引之间的差异来用所述像素例对所述合成像素进行分类。
5. 如权利要求 1 所述的方法,其特征在于,至少基于所述合成像素的玩家索引来用所述像素例对所述合成像素进行分类。
6. 如权利要求 5 所述的方法,其特征在于,至少部分基于所述合成像素的原始玩家索引与所述合成像素的咬合玩家索引或所述合成像素的增长玩家索引之间的差异来用所述像素例对所述合成像素进行分类。
7. 如权利要求 1 所述的方法,其特征在于,至少基于所述合成像素与相应的所观察到的像素之间的深度值差来用所述像素例对所述合成像素进行分类。
8. 如权利要求 1 所述的方法,其特征在于,所述合成像素具有像素地址,并且其中基于针对所述像素地址的原始玩家索引、咬合玩家索引、增长玩家索引、原始身体部位索引、咬合身体部位索引、增长身体部位索引、原始合成深度值、咬合深度值以及所观察到的深度值中的一个或多个来用所述像素例对所述像素地址处的合成像素进行分类。
9. 如权利要求 1 所述的方法,其特征在于,所述像素例选自一组定义的像素例,所述定义的像素例包括:精制 z 像素例、磁性像素例、推像素例、拉像素例、自遮挡推像素例、以及自遮挡拉像素例。
10. 如权利要求 1 所述的方法,其特征在于,映射到所述受力位置的力矢量是映射到所述受力位置的多个力矢量之一,每个这样的力矢量都对应于合成像素和经分类的像素例,并且每个这样的力矢量都作用于表示所述人类目标的模型到经调整的姿态的调整。

虚拟目标跟踪

[0001] 背景

[0002] 许多计算机游戏和其他计算机视觉应用利用复杂的控制来允许用户操纵游戏人物或应用的其他方面。这些控制可能难以学习,从而对许多游戏或其他应用造成了进入市场壁垒。此外,这些控制可能与这些控制所用于的实际游戏动作或其他应用动作非常不同。例如,使得游戏人物挥动棒球拍的游戏控制可能完全不类似于挥动棒球拍的实际运动。

[0003] 概述

[0004] 提供本概述以便以简化的形式介绍将在以下详细描述中进一步描述的一些概念。本概述并不旨在标识所要求保护的的主题的关键特征或必要特征,也不旨在用于限制所要求保护的的主题的范围。此外,所要求保护的的主题不限于解决在本公开的任一部分中提及的任何或所有缺点的实现。

[0005] 这里讨论了与视觉目标跟踪有关的各种实施例。所公开的实施例包括用包括多个部分的模型表示所述目标,每个部分都与同所述目标的一部分相对应的部位索引相关联。可以基于下列各项之一或二者来用像素例对模型的一部分进行分类:与模型的该部分相关联的部位索引;以及所述模型的该部分的所建模的位置与所述目标的相应部分的所观察到的位置之间的差异。然后,根据模型的该部分的像素例来调整该模型。

[0006] 附图简述

[0007] 图 1A 示出了跟踪玩拳击游戏的游戏玩家的示例性目标识别、分析和跟踪系统的实施例。

[0008] 图 1B 示出了图 1A 的游戏玩家挥出重拳,该重拳被跟踪并且被解释成致使玩家化身在游戏空间中挥出重拳的游戏控制。

[0009] 图 2 示意性地示出根据本公开的一实施例的计算系统。

[0010] 图 3 示出了用于表示人类目标的示例性身体模型。

[0011] 图 4 示出了用于表示人类目标的示例性骨架模型的基本上正面的视图。

[0012] 图 5 示出了用于表示人类目标的示例性骨架模型的斜视图。

[0013] 图 6 示出了用于表示人类目标的示例性网格模型。

[0014] 图 7 示出了在视觉上跟踪目标的示例方法的流程图。

[0015] 图 8 示出了示例性的观察到的深度图像。

[0016] 图 9 示出了示例性的合成的深度图像。

[0017] 图 10 示意性地示出了构成合成的深度图像的像素中的某一些。

[0018] 图 11A 示意性地示出了向模型的受力位置施加力。

[0019] 图 11B 示意性地示出了向图 11A 的模型的受力位置施加力的结果。

[0020] 图 12A 示出了从图 11A 的模型渲染的玩家化身。

[0021] 图 12B 示出了从图 11B 的模型渲染的玩家化身。

[0022] 图 13 示意性地示出了将合成的深度图像与相应的观察到的深度图像进行比较。

[0023] 图 14 示意性地示出了标识图 13 的较不匹配的合成的像素的区域。

[0024] 图 15 示意性地示出了合成的深度图像和相应的观察到的深度图像的另一比较,

其中不匹配的像素的区域对应于各种像素例。

[0025] 图 16 示意性地示出了拉像素例(pull pixel case)的示例实施例。

[0026] 图 17 示意性地示出了推像素例(push pixel case)的示例实施例。

[0027] 图 18 示出了详细列出各种像素例和骨架模型关节之间的示例关系的表。

[0028] 图 19 示出了向表示目标的模型的约束的施加。

[0029] 图 20 示出了向表示目标的模型的约束的另一施加。

[0030] 图 21 示出了向表示目标的模型的约束的又一施加。

[0031] 详细描述

[0032] 本公开涉及目标识别、分析和跟踪。具体来说,公开了使用深度相机或其他源来获取一个或多个目标的深度信息。如下文详细描述,这些深度信息然后可被用于高效且准确地对一个或多个目标进行建模以及跟踪。这里所述的目标识别、分析和跟踪提供了一种健壮的平台,其中可以相对较快的帧率来一致地跟踪一个或多个目标,即使目标移动成已被认为是使用其他方法难于分析的姿态(例如,当两个或更多个目标部分地互相重叠和/或遮挡时;当目标的一部分自遮挡同一目标的另一部分时;当目标改变其局部解剖外观(如某人触摸他的或她的头)时,等等)。

[0033] 图 1A 示出目标识别、分析和跟踪系统 10 的一非限制性示例。具体而言,图 1A 示出了计算机游戏系统 12,该计算机游戏系统 12 可以用于玩各种各样不同的游戏、播放一个或多个不同的媒体类型、和/或控制或操纵非游戏应用。图 1A 还示出了高清晰度电视即 HDTV 16 形式的显示器 14,该高清晰度电视可以用于向诸如游戏玩家 18 之类的游戏玩家呈现游戏视觉。此外,图 1A 示出了深度相机 20 形式的捕获设备,其可以用于视觉地监视诸如游戏玩家 18 之类的一个或多个游戏玩家。图 1A 中所示的示例是非限制性的。如下面参考图 2 所述,在不偏离本公开的范围的情况下可以使用各种不同类型的目标识别、分析和跟踪系统。

[0034] 目标识别、分析和跟踪系统可以用于识别、分析和/或跟踪诸如游戏玩家 18 之类的一个或多个目标。图 1A 示出了一种场景,其中使用深度相机 20 跟踪游戏玩家 18,使得游戏玩家 18 的移动可以被游戏系统 12 解释成可以用于影响由游戏系统 12 执行的游戏的控制。换言之,游戏玩家 18 可以使用他的移动来控制游戏。游戏玩家 18 的移动可以被解释成实际上任何类型的游戏控制。

[0035] 图 1A 中示出的示例场景示出了游戏玩家 18 正在玩正由游戏系统 12 执行的拳击游戏。游戏系统使用 HDTV 16 来在视觉上向游戏玩家 18 呈现拳击对手 22。此外,游戏系统使用 HDTV 16 来在视觉上呈现游戏玩家 18 用他的移动控制的玩家化身 24。如图 1B 中所示,游戏玩家 18 可以在物理空间中挥出重拳来作为对玩家化身 24 在游戏空间中挥出重拳的指令。游戏系统 12 和深度相机 20 可以用于识别和分析游戏玩家 18 在物理空间中的重拳,从而使得该重拳可以被解释成使得游戏化身 24 在游戏空间中挥出重拳的游戏控制。例如,图 1B 示出了 HDTV 16 响应于游戏玩家 18 在物理空间中挥出重拳而在视觉上呈现游戏化身 24 挥出袭击拳击对手 22 的重拳。

[0036] 游戏玩家 18 的其他移动也可以被解释成其他控制,诸如上下快速摆动、闪避、滑步、格挡、用拳猛击或挥动各种不同力度的重拳等控制。此外,一些移动可以被解释成用于除了控制游戏化身 24 之外的目的的控制。例如,玩家可以使用移动来结束、暂停或保存游

戏、选择级别、查看高分、与朋友交流等。

[0037] 在某些实施例中，目标可包括人类和物体。在这些实施例中，例如，电子游戏的玩家可能正手持物体，从而玩家和物体的运动被用来调整和 / 或控制电子游戏的参数。例如，可以跟踪并利用手持球拍的玩家的运动来控制电子体育游戏中的屏幕上球拍。在另一示例中，可以跟踪并利用手持物体的玩家的运动来控制电子格斗游戏中的屏幕上武器。

[0038] 目标识别、分析和跟踪系统可以用于将目标移动解释成游戏领域之外的操作系统和 / 或应用控制。实际上操作系统和 / 或诸如图 1A 和 1B 中所示的拳击游戏之类的应用的任何可控制的方面都可以由诸如游戏玩家 18 之类的目标的移动来控制。所示出的拳击场景是作为示例来提供的，但决不意味着以任何方式进行限制。相反，所示出的场景旨在展示可以在不背离本公开的范围的情况下应用于各种各样不同的应用的一般概念。

[0039] 在此所述的方法和过程可以绑定到各种不同类型的计算系统。图 1A 和 1B 示出了游戏系统 12、HDTV 16 和深度相机 20 形式的非限制性示例。作为另一更一般的示例，图 2 示意性地示出可以执行这里所述的目标识别、跟踪和分析方法和过程中的一个或多个的计算系统 40。计算系统 40 可以采取各种不同的形式，尤其包括但不限于：游戏控制台、个人计算游戏系统、军用跟踪和 / 或定标系统、以及提供绿屏或者运动捕获功能的特征采集系统。

[0040] 计算系统 40 可包括逻辑子系统 42、数据保持子系统 44、显示子系统 46 和 / 或捕捉设备 48。计算系统可任选地包括未示出在图 2 中的组件，并且 / 或者示出在图 2 中的某些组件可以是未被集成在计算系统中的外围组件。

[0041] 逻辑子系统 42 可包括被配置为执行一个或多个指令的一个或多个物理设备。例如，逻辑子系统可被配置为执行一个或多个指令，该一个或多个指令是一个或多个程序、例程、对象、组件、数据结构或其他逻辑构造的一部分。可实现这样的指令以执行任务、实现数据类型、变换一个或多个设备的状态、或以其他方式得到所希望的结果。逻辑子系统可包括被配置为执行软件指令的一个或多个处理器。附加地或替代地，逻辑子系统可包括被配置成执行硬件或固件指令的一个或多个硬件或固件逻辑机。逻辑子系统可任选地包括分布在两个或更多个设备上的独立组件，这些独立组件在一些实施例中可被远程放置。

[0042] 数据保持子系统 44 可包括一个或多个物理设备，该一个或多个设备被配置成保持逻辑子系统可执行以实现此处所述的方法和过程的数据和 / 或指令。在实现这样的方法和过程时，可以变换数据保持子系统 44 的状态（例如，以保持不同的数据）。数据保持子系统 44 可包括可移动介质和 / 或内置设备。数据保持子系统 44 可以包括光学存储器设备、半导体存储器设备（如 RAM、EEPROM、闪存等）、和 / 或磁存储器设备等。数据保持子系统 44 可包括具有以下特性中的一个或多个特性的设备：易失性、非易失性、动态、静态、读 / 写、只读、随机存取、顺序存取、位置可寻址、文件可寻址、以及内容可寻址。在某些实施例中，可以将逻辑子系统 42 和数据保持子系统 44 集成到一个或更多个常见设备中，如专用集成电路或片上系统。

[0043] 图 2 还示出采用计算机可读可移动介质 50 形式的数据保持子系统的一方面，该方面可用于存储和 / 或传送可执行来实现此处所述的方法和过程的数据和 / 或指令。

[0044] 显示子系统 46 可用于呈现由数据保持子系统 44 所保持的数据的可视表示。在此处所描述的方法和过程改变由数据保持子系统保持的数据，并由此变换数据保持子系统的状态时，同样可以变换显示子系统 46 的状态以在视觉上表示底层数据的改变。作为一个非

限制性示例,可通过显示子系统 46 以响应于游戏玩家在物理空间中的移动来改变游戏空间中的姿态的游戏人物的形式来反映这里所述的目标识别、跟踪和分析。显示子系统 46 可以包括使用实际上任何类型的技术的一个或多个显示设备。可将这些显示设备与逻辑子系统 42 和 / 或数据保持子系统 44 一起组合在共享外壳中,或这些显示设备可以是外围显示设备,如图 1A 和 1B 所示。

[0045] 计算系统 40 还包括被配置成获得一个或多个目标的深度图像的捕捉设备 48。捕捉设备 48 可被配置成通过任何合适的技术(例如飞行时间、结构化光、立体图像等)捕捉具有深度信息的视频。如此,捕捉设备 48 可包括深度相机、摄像机、立体相机、和 / 或其他合适的捕捉设备。

[0046] 例如,在飞行时间分析中,捕捉设备 48 可以向目标发射红外光,然后可使用传感器来检测从目标的表面反向散射的光。在一些情况下,可以使用脉冲式红外光,其中可以测量出射光脉冲和相应的入射光脉冲之间的时间并将该时间用于确定从该捕捉设备到目标上的特定位置的物理距离。在一些情况下,出射光波的相位可以与入射光波的相位相比较以确定相移,并且该相移可以用于确定从该捕捉设备到目标上的特定位置的物理距离。

[0047] 在另一示例中,飞行时间分析可用于通过经由诸如快门式光脉冲成像之类的技术分析反射光束随时间的强度,来间接地确定从该捕捉设备到目标上的特定位置的物理距离。

[0048] 在另一示例中,捕捉设备 48 可利用结构化光分析来捕捉深度信息。在这样的分析中,图案化光(即被显示为诸如网格图案或条纹图案之类的已知图案的光)可以被投影到目标上。在落到目标的表面上后,作为响应,该图案可能变形,并且图案的这种变形可以被研究以确定从捕捉设备到目标上的特定位置的物理距离。

[0049] 在另一示例中,捕捉设备可以包括从不同的角度查看目标的两个或更多个物理上分开的相机,以获得视觉立体数据。在这些情形中,可解析视觉立体数据以生成深度图像。

[0050] 在其他实施例中,捕捉设备 48 可利用其他技术来测量和 / 或计算深度值。此外,捕捉设备 48 可以将计算出的深度信息组织为“Z 层”,即与从深度相机沿其视线延伸到观察者的 Z 轴垂直的层。

[0051] 在某些实施例中,可将两个或更多个相机整合到一个集成捕捉设备中。例如,可将深度相机和摄像机(例如 RGB 摄像机)整合到共同的捕捉设备中。在某些实施例中,可协同使用两个或更多个分开的捕捉设备。例如,可使用深度相机和分开的摄像机。当使用摄像机时,该摄像机可用于提供目标跟踪数据、对目标跟踪进行纠错的确认数据、图像捕捉、面部识别、对手指(或其他小特征)的高精度跟踪、光感测和 / 或其他功能。

[0052] 要理解,至少一些目标分析和跟踪操作可以由一个或多个捕捉设备的逻辑机来执行。捕捉设备可以包括被配置成执行一个或多个目标分析和 / 或跟踪功能的一个或多个板载处理单元。捕捉设备可包括固件以帮助更新这样的板载处理逻辑。

[0053] 计算系统 40 可任选地包括诸如控制器 52 和控制器 54 之类的一个或多个输入设备。输入设备可被用于控制计算系统的操作。在游戏的上下文中,诸如控制器 52 和 / 或控制器 54 之类的输入设备可被用于控制游戏的那些不是通过这里所述的目标识别、跟踪和分析方法和过程来控制的方面。在某些实施例中,诸如控制器 52 和 / 或控制器 54 之类的输入设备可包括可用于测量控制器在物理空间中的移动的加速计、陀螺仪、红外目标 / 传感

器系统等中的一个或多个。在某些实施例中,计算系统可任选地包括和 / 或利用输入手套、键盘、鼠标、跟踪垫、轨迹球、触摸屏、按钮、开关、拨盘、和 / 或其他输入设备。如将理解的,目标识别、跟踪和分析可被用于控制或扩充游戏或其他应用的常规上由诸如游戏控制器之类的输入设备所控制的方面。在某些实施例中,这里所述的目标跟踪可被用作对其他形式的用户输入的完全替代,而在其他实施例中,这种目标跟踪可被用于补充一个或多个其他形式的用户输入。

[0054] 计算系统 40 可被配置成执行这里所述的目标跟踪方法。然而,应理解,计算系统 40 是作为可执行这种目标跟踪的设备的非限制性示例而提供的。其他设备也在本公开的范围之内。

[0055] 计算系统 40 或另一合适的设备可被配置成用模型来表示每个目标。如下面更详细地描述的,从这样一种模型导出的信息可与从诸如深度相机之类的捕捉设备获得的信息相比较,使得模型的基本比例或形状以及其当前姿态可被调整成更准确地表示被建模的目标。模型可由一个或多个多边形网格、由一组数学图元、和 / 或通过被建模的目标的其他合适的机器表示来表示。

[0056] 图 3 示出了示例身体模型 70 的非限制性视觉表示。身体模型 70 是被建模的目标(例如,图 1A 和 1B 的游戏玩家 18)的机器表示。身体模型可包括一个或多个数据结构,该一个或多个数据结构包括一组变量,该组变量一起以游戏或其他应用 / 操作系统的语言来定义被建模的目标。

[0057] 目标的模型可在不背离本公开的范围的情况下不同地被配置。在某些示例中,模型(例如,机器可读模型)可包括将目标表示为包括刚性和 / 或可变形形状,即身体部位,的三维模型的一个或多个数据结构。每一身体部位可被表征为数学图元,其示例包括但不限于球体、各向异性地缩放的球体、圆柱体、各向异性圆柱体、平滑圆柱体、方框、斜面方框、棱柱等。

[0058] 此外,可用包括多个部分的模型来表示该目标,每个部分和与该目标的某一部位相对应的部位索引相关联。从而,对于该目标是人类目标的情况,该部位索引可以是与人类目标的某一部位相对应的身体部位索引。例如,图 3 中的身体模型 70 包括身体部位 bp1 到 bp14,这些身体部位中的每一个都表示被建模的目标的一不同部分。每一身体部位都是三维形状。例如,bp3 是表示被建模的目标的左手的矩形棱柱,而 bp5 是表示被建模的目标的左上臂的八角棱柱。身体模型 70 是示例性的,因为身体模型可包含任意数量的身体部位,其各自可以是被建模的目标的相应部位的任何机器可理解的表示。

[0059] 包括两个或更多身体部位的模型还可包括一个或多个关节。每一关节可允许一个或多个身体部位相对于一个或多个其他身体部位移动。例如,表示人类目标的模型可包括多个刚性和 / 或可变形身体部位,其中一些身体部位可表示人类目标的相应解剖学身体部位。此外,该模型的每一身体部位可包括一个或多个结构件(即,“骨骼”),而关节位于相邻骨骼的交叉点处。应当理解,一些骨骼可对应于人类目标中的解剖学骨骼,和 / 或一些骨骼在人类目标中可能不具有相应的解剖学骨骼。

[0060] 作为示例,可将人类目标建模成包括多个骨架点的骨架,每一骨架点具有在世界空间中的三维位置。各个骨架点可与人类目标的实际关节、人类目标的肢端的末端、和 / 或不与人类目标在解剖学上直接链接的点相对应。每一骨架点具有至少三个自由度(例如,世

界空间 x, y, z)。如此,该骨架可完全由 $3 \times \lambda$ 个值定义,其中 λ 等于该骨架中所包括的骨架点的总数。例如,具有 33 个骨架点的骨架可由 99 个值来定义。如下面更详细地描述的,一些骨架点可说明轴向滚动角。

[0061] 骨骼和关节可共同构成骨架模型,它们可以是该模型的构成元素。骨架模型可包括每一身体部位的一个或多个骨架构件以及相邻骨架构件之间的关节。示例性骨架模型 80 和示例性骨架模型 82 分别如图 4 和 5 所示。图 4 示出了从正面查看的具有关节 j_1 到 j_{33} 的骨架模型 80。图 5 示出了从斜视图看的也具有关节 j_1 到 j_{33} 的骨架模型 82。骨架模型 82 还包括滚动关节 j_{34} 至 j_{47} ,其中每一滚动关节可用于跟踪轴向滚动角。例如,轴向滚动角可用于定义肢相对于其源肢和 / 或躯干的旋转定向。例如,如果骨架模型示出手臂的轴向旋转,则滚动关节 j_{40} 可用于指示相关联的腕所指向的方向(例如,手掌向上)。由此,尽管关节可受力并调整骨架模型,但如下所述,可改为构造并利用滚动关节来跟踪轴向滚动角。更一般地,通过检查肢相对于其源肢和 / 或躯干的定向,可确定轴向滚动角。例如,如果检查小腿,则可检查小腿相对于相关联大腿和髌部的定向以确定轴向滚动角。

[0062] 如上所述,某些模型可包括用作被建模的目标的机器表示的骨架和 / 或身体部位。在一些实施例中,模型可另选地或另外地包括线框网格,该线框网格可包括刚性多边形网格、一个或多个可变形网格、或两者的任意组合的分层结构。作为一个非限制性示例,图 6 示出了模型 90,该模型 90 包括排列成限定身体模型的形状的网格的多个三角形(例如,三角形 92)。这一网格可包括每一多边形边处的弯曲限制。当使用网格时,可选择共同构成该网格的三角形和 / 或其他多边形的数量来达到质量和计算开销之间的所需平衡。更多三角形可提供更高的质量和 / 或更准确的模型,而更少三角形可在计算上需求更少。包括多边形网格的身体模型不需要包括骨架,但在某些实施例中可包括骨架。

[0063] 上述身体部位模型、骨架模型和多边形网格是可用作被建模的目标的机器表示的非限制性示例类型的模型。其他模型也在本公开的范围之内。例如,某些模型可包括曲片(patch)、非均匀有理 B 样条、细分表面、或其他高阶表面。模型还可包括表面纹理和 / 或其他信息来更准确地表示被建模的目标的衣着、头发和 / 或其他方面。模型可任选地包括与当前姿态、一个或多个过去的姿态和 / 或模型物理学有关的信息。要理解,可摆姿态并然后被光栅化成合成的深度图像(或以其他方式由合成的深度图像呈现或表达)的任何模型与这里所述的目标识别、分析和跟踪兼容。

[0064] 如上所述,模型用作诸如图 1A 和 1B 中的游戏玩家 18 等目标的表示。当目标在物理空间中移动时,来自诸如图 1A 和 1B 中的深度相机 20 等捕捉设备的信息可用于调整模型的姿态和 / 或基本大小 / 形状,从而它更准确地表示该目标。具体地,可向该模型的一个或多个受力方面施加一个或多个力,来将该模型调整成更接近地对应于物理空间中的目标的姿态。取决于所使用的模型的类型,可向模型的关节、身体部位的质心、三角形的顶点、或任何其他合适的受力方面施加力。此外,在某些实施例中,在确定力的方向和 / 或大小时可使用两个或更多个不同的计算。如以下更详细描述的,捕捉设备所检索的目标的观察到的图像与模型的光栅化(即,合成的)图像之间的差异可用于确定施加于该模型的力以将身体调整到不同的姿态。

[0065] 图 7 示出了使用模型(例如,图 3 的身体模型 70)来跟踪目标的示例方法 100 的流程图。在某些实施例中,该目标可以是人类,并且该人类可以是所跟踪的两个或更多个目标

中的一个。如此,在某些实施例中,方法 100 可由计算系统(如图 1 所示的游戏系统 12 和 / 或图 2 所示的计算系统 40)执行来跟踪与计算系统上正在玩的电子游戏交互的一个或多个玩家。如以上所介绍的,对玩家的跟踪允许这些玩家的物理移动用作调整和 / 或控制电子游戏的参数的实时用户接口。例如,所跟踪的玩家的运动可用于移动电子角色扮演游戏中的屏幕上人物或化身。在另一示例中,所跟踪的玩家的运动可用于控制电子赛车游戏中的屏幕上车辆。在又一示例中,所跟踪的玩家的运动可用于控制虚拟环境中的物体的建筑或组织。

[0066] 在 102,方法 100 包括从源接收目标的观察到的深度图像。在某些实施例中,源可以是配置成通过诸如飞行时间分析、结构化光分析、立体视觉分析或其他合适的技术之类的合适的技术获得关于目标的深度信息的深度相机。观察到的深度图像可包括多个观察到的像素,其中每一观察到的像素具有观察到的深度值。观察到的深度值包括从源查看到的目标的深度信息。知道了深度相机的水平和竖直视野,以及一像素的深度值和该像素的像素地址,就可确定该像素所成像的表面的世界空间位置。为了方便,可将该像素所成像的表面的世界空间位置称为该像素的世界空间位置。

[0067] 图 8 示出了示例性的观察到的深度图像 140 的视觉表示。如所示,观察到的深度图像 140 捕捉手臂举起的站立着的人(例如,游戏玩家 18)的示例性的观察到的姿态。

[0068] 如图 7 的 104 处所示,在接收到观察到的深度图像之后,方法 100 可任选地包括将观察到的深度图像降采样到较低的处理分辨率。降采样到较低的处理分辨率可允许以更少的计算开销来更容易地使用和 / 或更快速的处理观察到的深度图像。

[0069] 如 106 处所示,在接收到观察到的深度图像之后,方法 100 可任选地包括从观察到的深度图像中移除非玩家背景元素。移除这些背景元素可包括将观察到的深度图像的各个区域分隔成背景区域和目标的图像所占据的区域。背景区域可从图像中被移除或被标识以使其可在一个或多个后续处理步骤期间被忽略。实际上可使用任何背景移除技术,且来自跟踪(以及来自前一帧)的信息可任选地被用于帮助和提高背景移除的质量。

[0070] 如 108 处所示,在接收到观察到的深度图像之后,方法 100 可任选地包括从观察到的深度图像中移除和 / 或平滑掉一个或多个高变度和 / 或含噪声的深度值。观察到的深度图像中的此类高变度和 / 或含噪声的深度值可源自多个不同的源,诸如在图像捕捉过程期间发生的随机和 / 或系统性误差、因捕捉设备而导致的缺陷和 / 或畸变等。由于此类高变度和 / 或含噪声的深度值可以是图像捕捉过程的伪像,因此在对图像的任何未来的分析中包括这些值可使结果偏斜和 / 或使计算变慢。由此,移除此类值可为将来的计算提供更好的数据完整性。

[0071] 还可过滤掉其他深度值。例如,下文参考步骤 118 所述的增长操作的准确度可通过选择性地移除满足一个或多个移除准则的像素来得到增强。例如,如果一深度值处于手和手所遮挡的躯干之间的中间,移除该像素可防止在后续处理步骤期间增长操作将一个身体部位溢出(spill)在另一身体部位之上。

[0072] 如 110 处所示,方法 100 可任选地包括填入和 / 或重构丢失和 / 或移除的深度信息的各部分。此类回填可通过对最近邻居求平均、过滤和 / 或任何其他合适的方法来实现。

[0073] 如图 7 的 112 处所示,方法 100 可包括获得模型(例如,图 3 的身体模型 70)。如上所述,该模型可包括骨架,该骨架包括多个骨架点、一个或多个多边形网格、一个或多个数

学图元、一个或多个高阶表面、和 / 或用于提供目标的机器表示的其他特征。此外,该模型可作为计算系统上存在的一个或多个数据结构的实例而存在。

[0074] 在方法 100 的一些实施例中,该模型可以是从前一时间步(即,帧)获得的摆姿态的模型。例如,如果方法 100 是连续执行的,则可获得从方法 100 的前一次迭代得到的对应于前一时间步的摆姿态的模型。以此方式,基于当前帧的观察到的深度图像和来自前一帧的模型,可将该模型从一个帧调整到下一帧。在某些情况下,可通过动量计算投影前一帧的模型以产生估计的模型以与当前观察到的深度图像进行比较。无需从数据库中查找模型或相反从头开始每一帧就可以完成这一操作。相反,可在连续帧中对该模型进行递增改变。

[0075] 在某些实施例中,姿态可由一个或多个算法来确定,这些算法可分析深度图像并以粗略水平标识所感兴趣的的目标(例如人)位于哪里和 / 或这(些)目标的姿态。在初始迭代期间或每当相信算法可选择比前一时间步期间计算的姿态更准确的姿态时,可使用算法选择姿态。

[0076] 在某些实施例中,该模型可以从数据库和 / 或其他程序获得。例如,模型可能在方法 100 的第一次迭代期间不可用,在这一情况下该模型可从包括一个或多个模型的数据库获得。在这一情况下,可使用被设计成选择展示与目标的姿态相似的姿态的模型的搜索算法来从数据库选择模型。即使来自前一时间步的模型是可用的,也可使用来自数据库的模型。例如,如果目标已改变姿态超过预定阈值,和 / 或根据其他准则,则可在特定帧数之后使用来自数据库的模型。

[0077] 在其他实施例中,可合成模型或其各部分。例如,如果目标的身体核心(躯干、上腹部和髋部)是由可变形的多边形模型来表示的,则可使用观察到的深度图像的内容来最初地构造该模型,其中该图像中的目标的外形(即,轮廓)可用于成形 X 和 Y 维度中的网格。另外,在这一方法中,观察到的深度图像的该区域中观察到的深度值可用于在模型的 XY 方向以及 Z 方向上“塑造”网格,以更令人满意地表示目标的身体形状。

[0078] 在 2009 年 10 月 21 日提交的美国专利申请 12/603,437 中描述了用于获得模型的另一方法,通过引用将该申请的内容全部并入于此。

[0079] 方法 100 还可包括使用合适的方法来表示出现在目标上的任何衣物。这一合适的方法可包括向模型添加图元或多边形网格形式的辅助几何形状,以及任选地基于姿态来调整该辅助几何形状来反映重力、衣料仿真等等。这一方法可方便将模型塑造成目标的更真实表示。

[0080] 如 114 所示,方法 100 可任选地包括向模型应用动量算法。由于目标的各个部位的动量可以预测图像序列中的变化,因此这一算法可助于获得模型的姿态。动量算法可使用模型的每一关节或顶点在多个固定数量的先前帧上的轨迹,以帮助获得模型。

[0081] 在某些实施例中,关于目标的不同部分可以在一时间帧(例如,一秒的 1/30 或 1/60)内移动有限距离的知识可用作获得模型时的约束。这一约束可用于在前一帧已知时排除特定姿态。

[0082] 在图 7 的 116,方法 100 还可包括将模型光栅化成合成的深度图像。光栅化允许由数学图元、多边形网格或其他对象描述的模型被转换成由多个像素描述的合成的深度图像。

[0083] 光栅化可以使用一个或多个不同的技术和 / 或算法来进行。例如,光栅化模型可

包括将模型的代表投影到二维平面上。在包括多个身体部位形状 of 模型(例如,图 3 的身体模型 70)的情况下,光栅化可包括将身体部位形状的集合投影并光栅化到二维平面上。对于模型投影到其上的二维平面中的每一像素,可存储各种不同类型的信息。

[0084] 图 9 示出了对应于图 3 的身体模型 70 的示例性合成的深度图像的视觉表示 150。图 10 示出了同一合成的深度图像的一部分的像素矩阵 160。如在 170 所指示的,合成的深度图像中的每一合成的像素可包括合成的深度值。如光栅化期间所确定的,给定的合成的像素的合成深度值可以是来自该合成的像素表示的模型的对应部位的深度值。换言之,如果前臂身体部位(例如,图 3 的前臂身体部位 bp4)的一部分被投影到二维平面上,则对应的合成的像素(例如,图 10 的合成像素 162)可被给予等于前臂身体部位的该部分的深度值的合成的深度值(例如,图 10 的合成深度值 164)。在所示示例中,合成的像素 162 具有合成的深度值 382cm。同样,如果相邻的手身体部位(例如,图 3 的手身体部位 bp3)被投影到二维平面上,则对应的合成的像素(例如,图 10 的合成像素 166)可被给予等于手身体部位的该部分的深度值的合成的深度值(例如,图 10 的合成深度值 168)。在所示示例中,合成的像素 166 具有合成的深度值 383cm。相应的观察到的深度值是在同一像素地址处由深度相机观察到的深度值。可以理解,以上是作为示例提供的。合成的深度值可以用任何度量单位来保存,或作为无量纲数来保存。

[0085] 如 170 处所指示的,合成的深度图像中的每一合成的像素可包括光栅化期间确定的原始身体部位索引。这一原始身体部位索引可指示像素对应于模型的哪些身体部位。在图 10 所示的示例中,合成的像素 162 具有原始身体部位索引 bp4,而合成的像素 166 具有原始身体部位索引 bp3。在某些实施例中,如果合成的像素不对应于目标的身体部位(例如,如果合成的像素是背景像素),则合成的像素的原始身体部位索引可以为 0 (nil)。在某些实施例中,不对应于某一身体部位的合成的像素可被给予一不同类型的索引。身体部位索引可以是一离散值或概率分布,指示出像素属于两个或更多个不同的身体部位的可能性。

[0086] 如 170 处所指示的,合成的深度图像中的每一合成的像素可包括光栅化期间确定的原始玩家索引,其中该原始玩家索引对应于该目标。例如,如果有两个目标,则对应于第一目标的合成的像素将具有第一玩家索引,而对应于第二目标的合成的像素将具有第二玩家索引。在所示示例中,像素矩阵 160 仅对应于一个目标,因此合成的像素 162 具有原始玩家索引 P1,而合成的像素 166 具有原始玩家索引 P1。可使用其他类型的索引系统而不背离本公开的范围。

[0087] 如 170 处所指示的,合成的深度图像中的每一合成的像素可包括一像素地址。该像素地址可定义像素相对于其他像素的位置。在所示示例中,合成的像素 162 具有像素地址 [5, 7],而合成的像素 166 具有像素地址 [4, 8]。可以理解,可使用其他定址方案而不背离本公开的范围。

[0088] 如在 170 处所指示的,每个合成的像素可任选地包括其他类型的信息,这些信息中的一些可在光栅化后获得。例如,每个合成的像素可包括更新的身体部位索引,该更新的身体部位索引可作为光栅化期间执行的咬合(snap)操作的一部分来确定,如下所述。每个合成的像素可包括更新的玩家索引,该更新的玩家索引可作为光栅化期间执行的咬合操作的一部分来确定每个合成的像素可包括更新的身体部位索引,该更新的身体部位索引可作为增长/修补(fix)操作的一部分来获得,如下所述。每个合成的像素可包括更新的玩家

索引,该更新的玩家索引可作为增长/修补(fix)操作的一部分来获得,如下所述。每个合成的像素可包括更新后的合成的深度值,该值可作为咬合操作的一部分来获得。

[0089] 以上提供的示例类型的像素信息并非限制。可存储各种不同类型的信息来作为每一像素的一部分。这样的信息可包括从深度图像获得的信息、从光栅化机器可读模型而获得的信息、和/或从一个或多个处理操作(例如,咬合操作、增长操作等)得出的信息。这些信息可被存储为公共数据结构的一部分,或者不同类型的信息可被存储在可(例如通过像素地址)被映射到具体像素位置的不同数据结构。作为一示例,作为光栅化期间的咬合操作的一部分获得的玩家索引和/或身体部位索引可被存储在光栅化映射和/或咬合映射中,而作为光栅化之后的增长/修补操作的一部分获得的玩家索引和/或身体部位索引可被存储在增长映射中,如下所述。可被分配给每一像素的其他类型的像素信息的非限制示例包括但不限于,关节索引、骨骼索引、顶点索引、三角形索引、质心索引等等。

[0090] 尽管在观察到的像素和合成的像素之间做出了区分,然而要理解,这些区分仅是为了便于描述而做出的。在每个像素地址处,可使用数据来表示从深度相机或其他源获得的观察到的信息。同样,在每个像素地址处,可使用数据来表示被光栅化、得出、计算或以其他方式合成的信息。在考虑某一像素的观察到的数据(例如,观察到的深度值)时,可将该像素称为观察到的像素。在考虑同一像素的合成的数据(例如,合成的深度值)时,可将该同一像素称为合成的像素。如此,通过将该像素地址处的观察到的数据与该像素地址处的合成的数据进行比较,可以在该相同像素地址处的观察到的像素和合成的像素之间做出比较。

[0091] 在 118,图 7 的方法 100 可任选地包括咬合和/或增长身体部位索引和/或玩家索引。换言之,可扩大合成的深度图像以使某些像素的身体部位索引和/或玩家索引改变来试图更接近地对应于被建模的目标。在提及身体部位索引或玩家索引而没有明确提及从光栅化最初得出的索引、从咬合操作得出的索引、或从增长操作得出的索引时,要理解,可以使用这些索引中的任何一个或多个,以及从估算该像素所属的玩家和/或身体部位的其他适当方法获得的其他索引。

[0092] 在执行上述光栅化时,可构造一个或多个 Z 缓存和/或身体部位/玩家索引映射。作为非限制示例,这一缓存/映射的第一版本可通过执行如下的 Z 测试来构造:选择每个像素处的最接近观看者(例如,深度相机)的最前方表面并且将与该表面相关联的身体部位索引和/或玩家索引写入对应的像素。此映射可被称为光栅化映射或原始合成的深度映射,并且此映射可包括每个像素的原始身体部位索引。这一缓存/映射的第二版本可通过执行如下的 Z 测试来构造:选择该模型的最接近该像素处的观察到的深度值的表面,并且将与该表面相关联的身体部位索引和/或玩家索引写入对应的像素。这可以被称为咬合映射,并且此映射可包括每个像素的咬合身体部位索引。可约束这些测试,以便拒绝合成的深度值和观察到的深度值之间超过预定阈值的 Z 距离。在某些实施例中,可维护两个或更多个 Z 缓存和/或两个或更多个身体部位/玩家索引映射,从而允许进行两个或更多个上述测试。

[0093] 缓存/映射的第三版本可通过增长和/或纠正身体部位/玩家索引映射来构造。这可被称为增长映射。以上述咬合映射的副本开始,可在预定 Z 距离内在任何“未知的”值上增长各值,使得被目标占据但未被身体模型占据的空间可以用适当的身体部位/玩家索引来填充。这一方法还可包括如果标识了更满意的匹配则取代已知值。

[0094] 增长映射可以从在咬合映射的合成的像素上过一遍来检测具有带有不同身体部

位 / 玩家索引的相邻像素的像素开始。这些可被认为是“边缘”像素,即,可任选地沿其来对值进行传播的边界。如上所述,增长像素值可包括增长成“未知的”或“已知的”像素。对于“未知的”像素,例如,在一个场景中,身体部位 / 玩家索引值之前可能为 0,但是现在可具有非零相邻像素。在这一情况下,可检查四个直接相邻像素,并且可选择具有更接近地类似于所感兴趣的像素的深度值的观察到的深度值的相邻像素并将其分配给所感兴趣的像素。在“已知的”像素的情况下,可能的是如果具有已知非零身体部位 / 玩家索引值的像素的一个相邻像素具有在光栅化期间写入的、比该像素的合成的深度值更接近地匹配所感兴趣的像素的观察到的深度值的深度值,则可不取(overtake)该具有已知非零身体部位 / 玩家索引值的像素。

[0095] 另外,出于效率的目的,更新合成的像素的身体部位 / 玩家索引值可包括将其相邻的四个像素添加到要在后一遍重新被访问的像素队列。如此,值可沿着边界持续传播而不用在所有像素上完整地过一遍。作为另一优化,可跟踪感兴趣的目标所占据的不同的 $N \times N$ 的像素块(例如 16×16 的像素块),使得不被感兴趣的目标占据的其他块可被忽略。可在光栅化之后在目标分析期间的任何点上用各种形式来应用这种优化。

[0096] 然而,要注意的是增长操作可采用各种不同的形式。例如,可首先执行各种泛色填充以标识相似值的区域,然后可决定哪些区域属于哪些身体部位。此外,任何身体部位 / 玩家索引对象(例如,图 3 的左前臂身体部位 bp4)可增长的像素数量可基于预期这一对象占据多少像素(例如,给定其形状、距离和角度)对比咬合映射中有多少像素被分配该身体部位 / 玩家索引来得到限制。另外,上述方法可包括对于某些姿态添加优点或缺点来偏移对某些身体部位的增长,使得增长可正确。

[0097] 如果确定来自某一身体部位的像素的分布被分组成一个深度,而来自同一身体部位的像素的另一分布被分组成另一深度,从而在这两个分布之间存在间隙,则可以对咬合映射作出渐进咬合调整。例如,在躯干前且靠近躯干挥动的手臂可“溢出到”(spill into)躯干中。这一情况可产生一组具有一身体部位索引的躯干像素,该身体部位索引指示这些像素是手臂像素,而实际上它们应是躯干像素。通过检查小臂中的合成的深度值的分布,可以确定某些手臂像素可被分组成一个深度,而其余可被分组成另一深度。这两组深度值之间的间隙指示手臂像素和应该是躯干像素的像素之间的跳跃。由此,响应于标识这一间隙,然后可通过向溢出像素分配躯干身体部位索引来补救该溢出。作为另一示例,渐进咬合调整在“手臂在背景对象上”(arm-over-background-object)情况中可能是有帮助的。在该情况中,可使用直方图来标识感兴趣的像素(即被认为属于手臂的像素)的观察到的深度中的间隙。基于这一间隙,一组或多组像素可被标识为正确地属于手臂,和 / 或其他组可作为背景像素而被拒绝。直方图可基于各种度量,例如绝对深度;深度误差(合成的深度 - 观察到的深度)等。在任何增长操作之前,可在光栅化期间内联地执行渐进咬合调整。

[0098] 在 120 处,图 7 的方法 100 可任选地包括从上述三个阶段的处理处的观察到的深度图像、合成的深度图像、以及身体部位 / 玩家索引映射创建高度映射。这一高度映射的梯度和 / 或这一高度映射的模糊版本可在确定要对模型作出的调整的方向时使用,如下文所描述的。然而,高度映射仅仅是一种优化;另选地或另外地,可执行所有方向上的搜索来标识可应用调整的最近的关节和 / 或要作出这些调整的方向。当使用高度映射时,可以在以下描述的像素类确定之前、之后或同时创建高度映射。当被使用时,高度映射被设计成将玩

家的实际身体设为低标高(elevation),而将背景元素设为高标高。然后可使用分水岭样式的技术来跟踪高度映射中的“下坡”,以找到玩家上离背景最近的点,或反之亦然(即搜寻高度映射中的“上坡”以找到最接近于给定玩家像素的背景像素)。

[0099] 合成的深度图像和观察到的深度图像可能不相同的,且因此合成的深度图像可使用调整和 / 或修改以使其更接近地匹配观察到的深度图像且因此更准确地表示目标。要理解,可通过首先对模型作出调整(例如改变模型的姿态),然后将经调整的模型合成为合成的深度图像的新版本来对合成的深度图像作出调整。

[0100] 可采取多种不同方法来修改合成的深度图像。在一种方法中,可获得并光栅化两个或更多个不同的模型来产生两个或更多个合成的深度图像。每一合成的深度图像然后可以按照预定的一组比较度量来与观察到的深度图像进行比较。可选择展示出与观察到的深度图像最接近地匹配的合成的深度图像,且该过程可任选地被重复来改进该模型。当被使用时,该过程可尤其有用于精制该身体模型以匹配玩家的体型和 / 或身体尺寸。

[0101] 在另一方法中,可经由内插或外插来混合两个或更多个合成的深度图像以产生混合的合成的深度图像。在又一方法中,可按混合技术和参数跨混合的合成的深度图像变化的方式来混合两个或更多个合成的深度图像。例如,如果第一合成的深度图像在一个区域中令人满意地匹配到观察到的深度图像,而第二合成的深度图像在第二区域中令人满意地匹配,则混合的合成的深度图像中所选的姿态可以是类似于用于创建第一区域中的第一合成的深度图像的姿态和用于创建第二区域中的第二合成的深度图像的姿态的混合。

[0102] 在再一方法中,如图 7 的 122 处所示,可将合成的深度图像与观察到的深度图像进行比较。合成的深度图像的每一合成的像素可基于比较的结果来分类。这一分类可被称为为每一像素确定像素例(pixel case)。用于创建合成的深度图像的模型(例如,图 3 的身体模型 70)可根据所确定的像素例来系统地被调整。

[0103] 如上所述,可基于与和每个合成的像素具有相同像素地址的观察到的图像的相应像素的比较,选择该合成的像素的一个或多个像素例。在某些实施例中,该比较可基于一个或多个因素,这些因素包括但不限于 - 合成的像素的观察到的深度值和合成的深度值之间的差;合成的像素的原始身体部位索引、(咬合)身体部位索引、和 / 或(增长)身体 / 部位索引之间的差;和 / 或合成的像素的原始玩家索引、(咬合)玩家索引和 / 或(增长)玩家索引之间的差。相应地,在某些实施例中,可从一组定义的像素例中选择该像素例,如参考图 7 的 124-136 更详细地描述的。

[0104] 作为示例,图 13 示出合成的深度图像(例如,图 9 的合成的深度图像 150)与相应的观察到的深度图像(例如,图 8 的观察到的深度图像 140)通过分析进行比较的示例,以确定像素不匹配并从而标识像素例。与该模型相对应的合成的深度图像 150 的合成的像素在图 13 中由用实线描绘的合成的轮廓 200 表示,而与该目标相对应的观察到的深度图像 140 的观察到的像素在图 13 中由用虚线描绘的观察到的轮廓 202 表示。可以理解,尽管将这种比较示意性地描绘为视觉比较,然而在实践中这种比较可以是与诸如在图 10 示出的每个像素地址相对应的信息的通过比较的分析。

[0105] 在将每个合成的像素或合成的像素的组与相应的观察到的像素或观察到的像素的组进行比较后,可将每个合成的像素与像素例相关联。例如,对于每个合成的像素,可从诸如精制 z 像素例、磁性像素例、推像素例、拉像素例、自遮挡推和 / 或拉像素例等一组定义

的像素例中选择一像素例。

[0106] 图 14 示出在诸如在 204 示出的对角线阴影中指示的具有像素不匹配(例如,观察到的深度图像的的深度值与合成的图像的的深度图的不同大于阈值量)的合成的轮廓 200 的合成的像素的示例区域。在标识合成的深度图像的哪些合成的像素关于具有观察到的图像中的相同像素地址的像素不匹配之后,可调整以合成的深度图像所表示的模型,使得该模型更好地表示该目标。

[0107] 图 15 示出合成的深度图像和相应的观察到的深度图像的另一示例比较 206,其中对合成的像素的不同区域选择了不同的像素例。区域 208 包括该模型中在深度方向上从观察到的深度图像的相应的一个或多个部分向前或向后移位(例如, Z- 移位)的一个或多个部分。作为示例,区域 208 可与具有精制 z 像素例的像素相对应。由对角线阴影标识的区域(诸如示例区域 210)指示该模型中的从观察到的深度图像中的人类目标的轮廓移开的部分。作为示例,区域 210 可与具有推像素例的像素相对应。由水平线阴影标识的区域(诸如示例区域 212)指示观察到的深度图像中的从该模型的轮廓移开的部分。作为示例,区域 212 可与具有拉像素例的像素相对应。由交叉线阴影标识的区域(诸如示例区域 214)指示该模型中的与具有磁性像素例的像素相对应的部分,诸如臂和 / 或手。

[0108] 返回图 7,如上所述,对于每个合成的像素,可从诸如精制 z 像素例、磁性像素例、推像素例、拉像素例、自遮挡推和 / 或拉像素例等一组定义的像素例中选择一像素例。在这样做之后,该模型中的具有这些像素不匹配的合成的像素可随后通过将该模型调整为更接近地匹配观察到的图像来被纠正。例如,通过向该模型施加力来将该模型重新定位到更接近地匹配观察到的图像的不同姿态,可以做出这样的调整。在某些实施例中,可经由具有大小和方向的力矢量来施加力,该力可被施加到该模型的受力位置,如在图 7 的 141、142 和 144 处指示的。每个力矢量的计算和施加可基于该像素例。这种力矢量可从单一像素地址或从一组两个或更多个相关的像素地址(例如,具有匹配值——身体部位索引、玩家索引等——的相邻的像素地址)得出。后文更详细地讨论像素例和相关联的力矢量的示例。

[0109] 如图 7 的 124 处所示,确定像素例可包括选择精制 z (refine-z)像素例。可在观察到的深度图像的一观察到的像素(或多个观察到的像素的区域中)的观察到的深度值不匹配合成的深度图像中的合成的深度值、但足够接近而可能属于两个图像中的同一对象,且身体部位索引匹配(或在某些情况下对应于相邻身体部位或区域)时,选择精制 z 像素例。如果合成的像素的观察到的深度值和合成的深度值之间的差在预定范围内,且(任选地)该合成的像素的(增长)身体部位索引对应于还未被指定用于接收磁性力的身体部位,则可对该合成的像素选择精制 z 像素例。作为另一示例,如果合成的深度值不与观察到的深度值相匹配,并且合成的深度值和观察到的深度值之间的绝对差小于预先确定的阈值,则可用精制 z 像素例来分类所感兴趣的合成的像素。

[0110] 该精制 z 像素例对应于可对模型施加力来将模型移至正确位置的计算出的力矢量。换言之,可将精制 z 力矢量施加到该模型的一个或多个受力位置以将该模型的一部分朝观察到的深度图像的相应部分移动(例如,在沿着 Z 轴并垂直于图像平面的方向上)。计算出的力矢量可沿着垂直于图像平面的 Z 轴、沿着模型的某一方面(例如,相应身体部位的脸部)的法向的矢量、和 / 或沿着附近的观察到的像素的法向的矢量来施加。在某些实施例中,可将计算出的力矢量施加到相应身体部位的面部的法向的矢量和附近的观察到的像素

的法向的矢量的组合。作为非限制性示例,这种组合可以是平均、加权平均、线性内插等。力矢量的大小基于观察到的深度值和合成的深度值的差,越大的差对应于越大的力。换言之,在某些实施例中,该力矢量可与合成的深度值和观察到的深度值之间的绝对差成比例地增加。施加力的受力位置可被选为所感兴趣的像素的最近的合格受力位置(如最近的躯干关节),或该力可被分布在多个最近的合格受力位置的加权混合之间。最近的合格受力位置可被选择,但是,在某些情况下,应用偏移可能是有帮助的。例如,如果某一像素位于大腿向下一半位置处且已确定髌关节比膝更不易于活动(或灵活),则可能有帮助的是将对于中间腿像素的关节力偏移成作用于膝部而不是髌部。后文描述偏移的其他示例。

[0111] 在某些实施例中,可通过将感兴趣的合成的像素和每个合格受力位置之间的距离进行比较来确定精制 z 像素例的最近的合格受力位置。例如通过比较感兴趣的合成的像素和与该感兴趣的合成的像素的身体部位索引相关联的身体部位上的每个合格受力位置之间的距离,可以确定最近的合格受力位置。作为另一示例,该力矢量可以是被施加到多个最近的合格受力位置的加权混合的多个力矢量中的一个。进而,可将该力矢量例如朝相对更易于移动的合格受力位置偏移。例如,可将该力矢量的施加偏移 to 比最近的合格受力位置更易于移动的次于最近的合格受力位置。

[0112] 利用或不利用上述的偏移,对于哪个受力位置最接近于感兴趣的像素(即,感兴趣的合成的像素)的确定可通过强力搜索来找到。为了加速搜索,所搜索的受力位置的集合可被限于仅仅位于或接近于与该像素的身体部位索引相关联的身体部位的那些受力位置。每次姿态变化时,也可设置 BSP(二叉空间划分)树来帮助加速这些搜索。身体上的每个区域或对应于某一身体部位索引的每个身体部位可被给予其自己的 BSP 树。如果这样,则可对每个身体部位不同地施加偏移,这进一步允许对适当的受力位置的明智选择。

[0113] 如图 7 的 126 处所示,确定像素例可包括选择磁性像素例。当增长映射中正被检查的合成的像素对应于某一预定子集的身体部位(如手臂或图 3 的 bp3、bp4、bp5、bp7、bp8 和 bp9)时,可使用磁性像素例。尽管提供手臂作为示例,但在某些场景中,诸如腿或整个身体之类的其他身体部位可任选地与磁性像素例相关联。同样,在某些场景中,手臂可能不与磁性像素例相关联。

[0114] 为磁性例标记的像素可被分组成区域,每个区域与特定的身体部位相关联(在该示例中例如左上臂、左下臂、左手等等)。例如,诸如上面描述的增长操作可在处理磁性像素之前完成。在增长操作期间,可用该目标的最可能与该像素相对应的身体部位来“标记”每个像素。然而,在增长操作期间,可能的是可能用错误的身体部位标记一个或多个像素(即,误标记)。作为示例,在臂和/或手的快速移动期间,运动预测器可能不能完成预测,并且因此快速移动的手部像素可能没有被添加到咬合映射中,然而肩膀附近的较慢的上臂像素可能仍旧被添加到咬合映射中。在这种情况下,距肩更远的肢像素可能在分配身体部位索引时具有相对更多的误差。在这些像素被误标记的情况下,在增长操作期间,小臂像素可能向下增长到手区域中。作为另一示例,如果小臂像素和手像素均没有被添加到咬合映射中,则被添加到咬合映射中的上臂像素可被向下增长到小臂和手区域中。因此,例如,与人类目标的手相对应的像素可能被标记为“小臂”或所有的臂像素可能被标记为“上臂”。因此,在处理磁性时丢弃该信息可能是有用的,更详细地描述如下。

[0115] 尽管增长操作可能不正确地标识了这些像素属于肢的哪个部位,然而被分配到标

识肢本身的像素的原始身体部位往往具有更高的置信度。换言之,尽管小臂像素可能不正确地与上臂相关联,然而该像素与臂的某个部位相对应这一事实仍旧是正确的。相应地,在增长操作期间分配的子集分类可被丢弃。如此,磁性像素可被分组为更广的分类(即,“池”),诸如“左臂”、“右臂”和“其他”。左臂和右臂池中的像素可随后被标记为属于磁性像素例。上面是非限制性示例,并且可以使用标识臂像素或属于灵活身体部位的其他像素的其他方法。

[0116] 对于为磁性例标记的每个像素(例如,左臂池的像素),可将该像素的位置从具有 X、Y 像素位置和深度值的屏幕空间位置转换到具有在三维空间中标识该位置的坐标的世界空间位置。可以理解,这仅是处理像素的一个实施例。在其他实施例中,可将该像素的屏幕空间位置转换到世界空间位置。

[0117] 继续左臂和右臂磁性像素中的每一个的处理,可将该像素投影到构成该模型的臂并且用线段表示的“骨骼”上。与所有像素例一样,可将该像素投影到该骨骼的当前的、最佳猜测版本上。骨骼的最佳猜测版本可来自前一帧的最终姿态(用或者不用动量),或者可以用在当前帧期间的任何调整来更新(例如,运行精制 z 像素例来调整骨骼,并随后使用调整后的骨骼来用于磁性像素)。换言之,可在帧的处理期间的任何点处渐进地更新关节,并且可使用更新后的关节进行当前或后续帧中的后续处理。

[0118] 作为磁性处理的示例,该臂可包括三个骨骼分段,即上臂骨骼、下臂骨骼和手。对于每个像素,可通过分析确定该有限线段上的最靠近该像素的点。在某些实施例中,这可包括将像素与三维关节位置(所述三维关节位置在 Z 方向上以该关节处的该肢的估算的半径被向前拉)进行比较,使得该比较是两个表面值而不是一个表面值和一个内部值的比较。

[0119] 随后可将该像素分配到该最近的线段。在某些实施例中,如果确定最近的线段可能为不正确的,则可将该像素分配到不同的线段。例如,如果目标的臂是伸出的而模型的臂处于“鸡翅”位置,则距肩膀足够远的像素(例如,上臂长度的 1.5 倍)可使最近的线段被取代为小臂骨骼。在确定该像素可与哪个骨骼相关联之后,可将该像素的位置添加到该骨骼的“近”和“远”质心,如下文更详细地描述的。

[0120] 对于这些磁性区域中的每一个,可计算属于该区域的像素的质心。这些质心可以是正统的(所有起作用的像素被同等地加权),或偏移的,其中某些像素带有比其他像素更重的权重。例如,对于上臂,可跟踪三个质心:1) 无偏移的质心;2) “近”质心,其起作用的像素在更接近于肩部时被更重地加权;以及 3) “远”质心,其起作用的像素在更接近于肘时被更重地加权。这些权重可以是线性的(如 $2X$) 或非线性的(如 x^2) 或遵循任何曲线。

[0121] 一旦计算出这些质心,各种选项是可用的(并可被动态地选择)以用于计算感兴趣的体部位的位置和定向,即使某些被部分地遮挡。例如,当试图确定肘的新位置时,如果该区域中的质心足够可见(如果起作用的像素的权重的和超过预定阈值),则该质心本身标记该肘(估计 #1)。然而,如果肘区域不可见(可能由于它被某一其他物体或身体部位遮挡),肘位置仍然可被常常确定,如下面的非限制性示例所述。如果上臂的远质心可见,则可从肩部出发做出通过该质心达上臂的长度的投影,以获得肘的非常可能的位置(估计 #2)。如果下臂的近质心可见,则可从腕向上做出通过该质心达下臂的长度的投影,以获得肘的非常可能的位置(估计 #3)。

[0122] 向具有更高可见性、置信度、像素计数或任何数量的其他度量的估计给予优先级

(或更高的权重),可作出对三个可能的估计之一的选择,或可作出三个可能的估计之间的混合。最后,在该示例中,可在肘的位置向模型施加单个磁性力矢量;然而,它可被更重地加权(当与产生于其他像素例但作用于该相同受力位置的像素力矢量累计时),以表示许多像素被用于构造它的事实。当施加时,计算出的磁性力矢量可移动模型,使得相应的模型可更令人满意地匹配观察到的图像中所示的目标。磁性像素例的一个优点是其对诸如手臂等高度灵活的身体部位表现良好的能力。

[0123] 在某些实施例中,可仅使用磁性像素例来调整没有定义的关节或身体部位的模型。

[0124] 如图 7 的 128 和 130 处所示,确定像素例可包括选择推像素例和 / 或拉像素例。这些像素例可在合成的深度值和观察到的深度值可能在同一像素地址处严重失配的轮廓处被调用。要注意,拉像素例和推像素例也可在原始玩家索引不匹配(增长)玩家索引时被使用。是推还是拉的确如下。如果同一像素地址处合成的深度图像包含比观察到的深度图像中的深度值更大(更远)的深度值,例如大超过一阈值量,则该模型可被拉向增长的图像中所见的真实轮廓。换言之,对于从该模型的轮廓移开的观察到的深度图像的各部分,可在 XY 平面中朝观察到的深度图像中的该目标的轮廓拉该模型。作为示例,可使用施加到该模型的一个或多个受力位置的拉力矢量来“拉”该模型。图 16 示出了这种拉像素例的示例,并且更详细地描述如下。

[0125] 图 16 示意性地示出了与示例合成的深度图像 222 相比的示例观察到的深度图像 220,如在 224 所指示的。如此,合成的深度图像 222 的像素地址与观察到的深度图像 220 的像素地址相对应。为了更清晰地示出该示例,图 16 示出了夸大的示例,其中观察到的深度图像 220 和合成的深度图像 222 明显不匹配。然而,可以理解,在实践中,两个图像可仅以相对小的量不匹配,而如图所示那样严重的不匹配可能难以解决。

[0126] 观察到的深度图像 220 包括观察到的人类目标(例如,游戏玩家)的图像(即,玩家图像 226),其中该玩家图像 226 具有轮廓(即,玩家轮廓 228),使得玩家轮廓 228 内的像素是玩家图像 226 的像素而玩家轮廓 228 外的像素是观察到的背景 230 的像素。类似地,合成的深度图像 222 包括表示观察到的游戏玩家的模型 232,其中模型 232 具有轮廓(即,模型轮廓 234),使得模型轮廓 234 内的像素是模型 232 的像素而模型轮廓 234 外的像素是合成的背景 236 的像素。

[0127] 在将合成的深度图像 222 与观察到的深度图像 220 进行比较之后,不匹配的存在变得更加显然,使得在同一像素地址处的像素对应于每个深度图像的不同部位。例如,为讨论而选择一示例像素,即感兴趣的合成的像素 238。如图所示,感兴趣的合成的像素 238 与合成的深度图像 222 的合成的背景 236 相对应。然而,相应的观察到的深度图像中的同一像素地址对应于与玩家图像 226 相关联的观察到的像素 240。在该特定示例中,感兴趣的合成的像素 238 具有比相应的观察到的像素 240 更大的深度值,因为背景处于比游戏玩家更大的深度处(即,更远离深度相机)。如此,可将模型 232 朝感兴趣的合成的像素 238(即,朝玩家轮廓 228)拉,如箭头 240 所示。

[0128] 相反,如果原始合成的图像包含比观察到的深度图像中的深度值更小(更近)的深度值,例如小超过一阈值量,则该模型可被推出玩家不再占据的空间之外(且朝向增长的图像中的真实轮廓)。换言之,对于从观察到的深度图像中的人类目标的轮廓移开的该模型的

各部分,可在 XY 平面中朝观察到的深度图像中的该人类目标的轮廓推该模型。作为示例,可使用施加到该模型的一个或多个受力位置的推力矢量来“推”该模型。图 17 示出了这种推像素例的示例,并且更详细地描述如下。

[0129] 图 17 示意性地示出了与图 16 中所示的类似的比较,即,合成的深度图像 222 和观察到的深度图像 220 的比较。然而,对于图 17 中示出的示例,检查一不同的感兴趣的合成的像素,即,感兴趣的合成的像素 250。感兴趣的合成的像素 250 对应于合成的深度图像 222 的模型 232,而在相应的观察到的深度图像 220 中的同一像素地址处,观察到的像素 252 与观察到的背景 230 相关联。在该特定示例中,感兴趣的合成的深度像素 250 具有比相应的观察到的像素 252 更小的深度值,因为模型处于比背景更小的深度处(即,更接近深度相机)。如此,可推模型 222 远离感兴趣的合成的像素 250(即,朝玩家轮廓 228),如箭头 254 所示。

[0130] 在任一种情况中(例如,图 16 的拉像素例或图 17 的推像素例),对于这些像素或像素区域中的每一个,可对模型施加二维或三维计算出的力矢量以纠正轮廓失配,将身体模型各部位推到或拉到更精确地匹配观察到的深度图像中的目标的位置的位置。这一推和/或拉的方向通常主要在 XY 平面中,但在某些场景中可向该力添加 Z 分量。相应地,在某些示例中,推力矢量和/或拉力矢量可以是包括 Z 分量的三维矢量。

[0131] 例如,对于图 16 中示出的拉例,可向模型 232 的受力位置施加拉力矢量来将模型 232 朝观察到的深度图像中的玩家轮廓 228 拉。拉力矢量的大小可与观察到的深度图像的一部分从该模型的轮廓移开的拉偏移距离成比例。换言之,拉偏移距离 D_1 可被定义为感兴趣的合成的像素(例如,像素 238)和模型轮廓 234 的最近的合格像素之间的距离。如此,拉力矢量的大小 D_2 可以是拉偏移距离 D_1 的函数,如后文更详细地描述的。此外,拉力矢量的方向可平行于从模型轮廓 234 上的最近的合格像素向感兴趣的合成的像素 238 延伸的矢量。

[0132] 对于图 17 中示出的推像素例,可向模型 232 的受力位置施加推力矢量来将模型 232 朝观察到的深度图像 220 中的玩家轮廓 228 推。推力矢量的大小可与该模型的一部分从该玩家轮廓 228 移开的推偏移距离成比例。换言之,推偏移距离 D_1 可被定义为感兴趣的合成的像素(例如,像素 250)和玩家轮廓 228 的最近的合格像素之间的距离。在某些实施例中,推力矢量的大小 D_2 可以是推偏移距离 D_1 的函数,如后文更详细地描述的。此外,推力矢量的方向可平行于从感兴趣的合成的像素 250 向玩家轮廓 228 上的最近的合格像素延伸的矢量。

[0133] 为了对推或拉例产生适当的力矢量,合成的深度图像中玩家轮廓(即,模型轮廓)(对于拉例)或观察到的深度图像中玩家的轮廓(对于推例)上的最近的合格点可首先被找到。对于每个源像素(或对于每组源像素),可通过对最近点(在所希望的轮廓上)执行满足下列准则的强力穷举 2D 搜索来找到该点。在拉像素例中,具有原始映射中与增长映射中的玩家索引(在源像素或区域处)相匹配的玩家索引(在搜寻位置处)的最近像素被找到。在推像素例中,具有增长映射中与原始映射中的玩家索引(在源像素或区域处)相匹配的玩家索引(在搜寻位置处)的最近像素被找到。

[0134] 然而,强力搜索在计算上花费可能是非常高的,可使用优化来减少计算花费。用于更高效地找到该点的一个非限制性示例优化是遵循上述高度映射的梯度或其模糊版本,

以及仅在梯度方向中检查直线中的像素。在该高度映射中,在玩家索引在原始玩家索引映射和增长玩家索引映射中相同时,高度值低,而在玩家索引(在两个映射中)为零时,高度值高。在任一给定像素处,梯度可被定义为在该高度映射中指向“下坡”的矢量。如上所述,然后拉像素和推像素两者可沿该梯度(下坡)搜寻,直到它们到达它们的相应停止条件。如此,可使用沿模糊的高度映射的梯度的一维搜索来找到模型轮廓 234 上的最近的合格像素和 / 或玩家轮廓 228 上的最近的合格像素。进而,可通过测试使用该一维搜索找到的轮廓像素附近的模型轮廓像素来找到模型轮廓 234 上的最近的合格像素。同样,可通过测试使用该一维搜索找到的轮廓像素附近的玩家轮廓像素来找到玩家轮廓 228 上的最近的合格像素。

[0135] 这种搜寻操作的其他基本优化包括使用间隔对分或使用基于斜率的逼近来跳过像素;随着搜寻的进行,每隔一定间隔重新采样梯度;以及一旦满足停止准则则在附近检查以找到更好 / 更近的匹配(不是直接沿着梯度)。某些搜索策略可从满足一个或多个选择准则的候选像素的子集(诸如具有某个身体部位索引的像素)中选择最近的合格像素。

[0136] 不管使用什么技术来找到感兴趣的轮廓上的最近点,可使用行进的距离(源像素和轮廓像素之间的距离) D_1 来计算将推或拉该模型的力矢量的大小(长度) D_2 。在某些实施例中, D_2 可以是与 D_1 线性或非线性地相关的(如 $D_2=2*D_1$ 或 $D_2=D_1^2$)。作为一个非限制性示例,可使用下列公式: $D_2=(D_1-0.5 \text{ 像素}) * 2$ 。如上所述, D_1 可以是拉偏移距离或推偏移距离。相应地, D_2 可以分别是拉力矢量或推力矢量的大小。可使用如上所述的沿模糊的高度映射的梯度的一维搜索来找到拉偏移距离和 / 或推偏移距离。

[0137] 例如,如果在两个深度图像中的轮廓之间有 5 像素的间隙,则该间隙中的每个像素可执行小“搜寻”并产生力矢量。在真实轮廓附近的像素可仅按 1 个像素搜寻以到达轮廓,使得那些像素处的力大小为 $(1 - 0.5) * 2 = 1$ 。远离真实轮廓的像素可按 5 个像素搜寻,使得力大小将为 $(5 - 0.5) * 2 = 9$ 。一般来说,从最接近真实轮廓的像素到最远的那些,搜寻距离将为 $D_1 = \{1, 2, 3, 4, 5\}$, 且产生的力大小将为: $D_2 = \{1, 3, 5, 7, 9\}$ 。该例中的 D_2 的平均为 5, 如所希望的一产生的力矢量的平均大小相当于轮廓(接近每个受力位置)之间的距离,它是模型可被移动以将模型放在适当的位置的距离。

[0138] 对于每个源像素,然后可用方向和大小(即长度)来构造最终力矢量。对于拉像素,方向是由从轮廓像素到源像素的矢量确定的;对于推像素,它是反向矢量。该力矢量的长度为 D_2 。在每个像素,然后可将该力施加于最合格的(如最近的)受力位置(或分布在若干受力位置之间),且在每个受力位置可将这些力平均以产生身体模型的适当的局部移动。尽管在图 16-17 中未示出,然而在某些实施例中,受力位置可以是该模型的关节。

[0139] 如图 7 的 132 和 134 处所示,确定像素例可包括选择自遮挡推和 / 或拉像素例。尽管在上述推像素例和拉像素例中身体部位可在前景中相对于背景或另一目标移动,但自遮挡推像素例和拉像素例考虑了该身体部位在同一目标的另一身体部位的前方(例如,一条腿在另一条腿前方,手臂在躯干前方等等)的情况。可在像素的(咬合)玩家索引匹配其对应的(增长)玩家索引,但该(咬合)身体部位索引不匹配其对应的(增长)身体部位索引时标识这些例。在这些例中,搜寻方向(以找到轮廓)可用若干方式得出。作为非限制性示例,可执行强力 2D 搜索;可对该例调整第二组“遮挡”高度映射使得梯度可指导 1D 搜索;或可将方向设置为朝向最近的骨架构件上的最近点。这两种例的细节在其他方面与标准拉例和推例类似。

[0140] 如果合成的像素的(增长)身体部位索引对应于还未被指定来接收磁性力的身体部位,则可对该合成的像素选择推、拉、自遮挡推和 / 或自遮挡拉例。

[0141] 要理解,在某些情况下,单个像素可负责一个或多个像素例。作为一非限制性示例,一像素可负责自遮挡推像素力和精制 z 像素力两者,其中自遮挡推像素力施加于进行遮挡的身体部位上的受力位置,而精制 z 像素力施加于身体部位上被遮挡的受力位置。

[0142] 如图 7 的 136 处所示,确定像素例可包括对合成的像素不选择像素例。经常会不需要对合成的深度图像的所有合成的像素计算力矢量。例如,更远离合成的深度图像中所示的身体模型的合成的像素以及更远离观察到的深度图像中所示的目标的观察到的像素(即,背景像素)可能不会影响任何受力位置或身体部位。不需要对这些像素确定像素例,但在某些场景中也可以确定。作为另一示例,合成的像素的观察到的深度值和合成的深度值之间的差异可能低于预定阈值(例如,模型已经匹配观察到的图像)。由此,不需要对这些像素确定像素例,但在某些场景中也可以确定。

[0143] 在 141,图 7 的方法 100 包括对已为其确定了像素例的每一合成的像素,基于为该合成的像素选择的像素例来计算力矢量。如上所述,每一像素例对应于用于选择力矢量的大小、方向和 / 或受力位置的不同算法和 / 或方法。具体地,可基于所确定的像素例来为每个合成的像素计算力矢量(大小和方向),并且取决于模型的类型,可将所计算的力矢量施加于最近的合格关节、身体部位的质心、身体部位的某一点、三角形的顶点、或模型的用于生成合成的深度图像的另一预定受力位置。在某些实施例中,归因于给定像素的力可被分布在模型上的两个或更多个受力位置之间。

[0144] 力矢量可在任何坐标空间,如世界空间、屏幕空间(预 Z 划分)、投影空间(后 Z 划分)、模型空间等中计算和 / 或累积。例如,如上对于推和 / 或拉像素例所述,推和 / 或拉力矢量的大小可分别与推偏移距离和 / 或拉偏移距离成比例。对于精制 z 像素例,精制 z 力矢量的大小可基于合成的深度值和观察到的深度值之间的绝对差,使得精制 z 力矢量与该绝对差成比例地增加。对于磁性例,力矢量可取决于合成的像素对骨骼段的临近度,以及相应肢的质心。

[0145] 在 142,方法 100 包括将每一计算出的力矢量映射到模型的一个或多个受力位置。映射可包括将所计算的力矢量映射到“最匹配的”受力位置。对模型的最匹配的受力位置的选择取决于为对应的像素选择的像素例。最匹配的受力位置可以是例如最接近的关节、顶点或质心。在某些实施例中,可向模型施加力矩(即,旋转力)。

[0146] 在某些情况下,单一像素可负责两个或更多个不同的力矢量。作为非限制性示例,某像素可在咬合操作后被标识为遮挡躯干的肢像素,而同一像素随后在增长操作后可被标识为躯干像素(即,肢已从该像素地址移开)。在这种情况下,该像素可负责用于将肢推开的对肢的推力以及用于将躯干向适当的深度移动的对躯干的精制 z 力。作为另一示例,若某单个像素位于两个或更多个关节之间,则两个或更多个像素力可来自该单个像素。例如,小腿肚(mid-calf)像素可移动踝和膝两者。

[0147] 一般而言,在每个像素地址处,可使用给定像素地址的原始玩家索引、咬合玩家索引、增长玩家索引、原始身体部位索引、咬合身体部位索引、增长身体部位索引、合成的深度值、咬合深度值、观察到的深度值和 / 或其他观察到的或合成的数据的组合来确定该像素地址的一个或多个像素例。

[0148] 图 18 示出详细描述上面描述的像素例和图 5 的骨架模型 82 中示出的可向其映射力矢量的关节之间的示例关系的表格。在该表中,像素例 1-7 被如下缩写:1-拉(规则的)、2-拉(遮挡)、3-推(规则的)、4-推(遮挡)、5-精制 Z、6-磁性拉、以及 7-遮挡(无动作)。“受力?”列中的“是”条目指示该行的关节可接收来自力矢量的力。像素例列中的“X”条目表示该行的关节可接收来自对应于该列的像素例的力矢量的力。可以理解,该表格是作为示例提供的。其不被认为是限制。可建立模型和像素例之间的其他关系而不背离本公开的范围。

[0149] 一般而言,平移可由具有相似方向的力作用于模型的受力位置而引起,而旋转可由不同方向的力作用于模型的受力位置而引起。对于可变形对象,力矢量的某些分量可用于将模型在其变形限制内变形,并且力矢量的其余分量可用于平移和 / 或旋转模型。

[0150] 在某些实施例中,力矢量可被映射到最匹配的刚性或可变形对象、子对象、和 / 或对象的多边形集。因此,力矢量的一些分量可用于将模型变形,而力矢量的其余分量可用于执行模型的刚性平移。这一技术可导致“断开的”模型(例如,手臂可从身体切断)。如以下更详细讨论的,然后可使用校正步骤来将平移变换成旋转和 / 或应用约束以沿低能量路径将身体部位连回一起。

[0151] 此外,在某些实施例中,方法 100 的 142 包括映射超过一个力矢量。例如,具有与人类目标的臂相对应的身体部位索引的第一合成的像素可能已用第一像素例来分类,而具有与人类目标的躯干相对应的身体部位索引的第二合成的像素可能已用第二像素例来分类。在这种情况下,第一合成的像素的第一力矢量可按照第一像素例来计算,而第二合成的像素的第二力矢量可按照第二像素例来计算。相应地,第一力矢量可被映射到该模型的第一受力位置,其中该第一受力位置对应于该人类目标的臂。相应地,第二力矢量可被映射到该模型的第二受力位置,其中该第二受力位置对应于该人类目标的躯干。

[0152] 图 11A 和 11B 示出了向模型——所示示例中为骨架模型 180 施加力矢量的非常简化的示例。为简明起见,所示示例中仅示出了两个力矢量。每一这样的力矢量可以从两个或更多不同像素的像素例确定和力矢量计算中所得的两个或更多不同力矢量相加的结果。通常,模型将通过许多不同的力矢量来调整,每一力矢量是从许多不同的像素的像素例确定和力矢量计算所得的许多不同的力矢量之和。

[0153] 图 11A 示出了骨架模型 180,其中出于使骨架模型 180 的一条手臂伸直来更接近地匹配观察到的深度图像的目的,力矢量 182 被施加于关节 j18 (即,肘),而力矢量 184 被施加于关节 j20 (即,腕)。图 11B 示出了在施加力后的骨架模型 180。图 11B 示出所施加的力如何调整模型的姿态。如图 11B 所示,可保持骨架构件的长度。如进一步示出的,关节 j2 的位置保持在骨架模型的肩部,这是对于人类伸直他们的手臂的情况所预期的。换言之,在施加了力之后,骨架模型保持完整。在施加力时维持骨架模型的完整性是由所应用的一个或多个约束导致的,如以下更详细讨论的。可实施各种不同的约束来维持不同的可能模型类型的完整性。

[0154] 在 144,图 7 的方法 100 任选地包括将模型校正到满足一个或多个约束的姿态。如上所述,在收集了计算出的力矢量并将计算出的力矢量映射到模型的受力位置之后,然后可将计算出的力矢量施加到模型。如果是在没有约束的情况下执行的,则这可能“破坏”模型,将其不成比例地拉长和 / 或将身体部位移至对于目标的实际身体无效的配置。然后可

使用各种功能的迭代来将新模型位置“放松”到“附近的”合法配置。在校正模型的每一次迭代期间,约束可被逐渐地和 / 或逐步地应用于姿态,以便将该组姿态限制到在物理上可由一个或多个目标的一个或多个实际身体来表达的姿态。在其他实施例中,这一校正步骤可以按非迭代方式来完成。

[0155] 在某些实施例中,约束可包括以下各项中的一个或多个:骨架构件长度约束、关节角度约束、多边形边角约束、以及碰撞测试,如下文所讨论的。

[0156] 作为使用骨架模型的一示例,可应用骨架构件(即,骨骼)长度约束。可被检测的力矢量(即,在关节和 / 或身体部位可见且未被遮挡的位置处的力矢量)可沿着骨架模型的骨架构件的网络来传播。通过应用骨架构件长度约束,一旦所有骨架构件是可接受的长度,所传播的力就可“安顿下来”(settle in)。在某些实施例中,允许一个或多个骨架构件长度在预定范围内可变。例如,构成躯干两侧的骨架构件的长度可以变化以模拟可变形的上腹部。作为另一示例,构成上臂的骨架构件的长度可以变化以模拟复杂的肩窝。

[0157] 骨架模型可另外地或另选地通过基于目标计算每一骨架构件的长度,使得这些长度可被用作校正期间的约束来得到约束。例如,从身体模型得知所希望的骨骼长度;可评估当前骨骼长度(即新关节位置之间的距离)和所希望的骨骼长度之间的差。可调整该模型以减少所希望的长度和当前长度之间的任何误差。可对被认为是更重要的某些关节和 / 或骨骼以及当前比其他关节或身体部位更可见的关节或身体部位给予优先级。还可相对于低幅度变化给予高幅度变化优先级。

[0158] 作为示例,图 19 示出向表示目标的模型施加一个或多个约束。可以理解,尽管图 19 提供了视觉图示,然而在实践中,约束的施加可以是分析性的,并且可包括例如修改诸如图 10 中所示的像素数据。对于图 19 中描绘的示例,向模型 260 施加力矢量可能导致“断开的”模型。例如,该目标可对自身重新定位以将臂举过头顶。在跟踪该运动时,可向该臂的一个或多个受力位置施加力矢量以模仿该目标的运动。然而,这样做可能导致“打断”该臂,如在 262 处所描绘,和 / 或改变该臂的比例,如在 264 处所示。因为在本示例中该模型表示人类目标,所以这两个场景对于人类的臂在物理上是不可能的。相应地,可施加约束以确保对该模型的调整是物理上适当的。例如,可施加约束(如在 266 处所示)以确保前臂和上臂在肘处保持衔接。作为另一示例,可向前臂施加骨骼长度约束(如在 268 处所示)以确保前臂保持大致相同的长度。在施加这些约束后,如在 270 处所示,该模型保持其物理完整性。

[0159] 关节可见性和 / 或置信度可在 X、Y 和 Z 维度中单独被跟踪,以允许对骨骼长度约束的更准确的应用。例如,如果骨骼将胸连接到左肩,且胸关节的 Z 位置是高置信度的(即许多精制 z 像素对应于该关节)以及肩的 Y 位置是高置信度的(许多推 / 拉像素对应于该关节),则可纠正骨骼长度中的任何误差,同时部分地或完全地限制肩在 Y 方向中的移动或胸在 Z 方向中的移动。

[0160] 在某些实施例中,在校正之前的关节位置可以与校正之后的关节位置进行比较。如果确定在每一帧中正对骨架模型作出一组连贯的调整,则方法 100 可使用该信息来对骨架和 / 或身体模型执行“渐进式精制”。例如,通过比较校正之前和之后的关节位置,可以确定在校正期间在每一帧中肩正被推得相距更宽。这一连贯的调整暗示骨架模型的肩要小于所表示的目标的肩,并且因此,在校正期间每一帧调整肩宽来纠正此问题。在这一情况下,

可作出诸如增大骨架模型的肩宽等渐进式精制来纠正骨架和 / 或身体模型以更好地匹配目标。

[0161] 关于关节角度约束,可以限制某些肢和身体部位相对于相邻身体部位的运动范围。另外,该运动范围可基于相邻身体部位的定向来改变。因此,应用关节角度约束可允许在给定源肢和 / 或身体部位的定向的情况下将肢段约束到可能的配置。例如,小腿可被配置成向后弯曲(在膝部),但不向前弯曲。如果检测到非法角度,则违反常理的身体部位和 / 或它们的源(parent)(或在网格模型的情况下,违反常理的三角形和它们的邻居)被调整成将姿态保持在预定可能性的范围内,从而帮助避免模型折叠成被认为不可接受的姿态的情况。图 20 示出施加了一个或多个关节角度约束以将在 282 处示出的不正确的关节角度校正为诸如在 284 处示出的可接受运动范围内的模型 280 的示例。在极限角度违反的某些情况下,姿态可被认为是相反的,即被跟踪为胸的实际上是玩家的背;左手实际上是右手;等等。当这种不可能的角度明显可见(且足够异乎寻常)时,这可被解释成意味着该姿态已被反向映射到玩家的身体上,且该姿态可被翻转以正确地为目标建模。

[0162] 可应用碰撞测试来防止模型穿透其自身。例如,碰撞测试可防止前臂 / 手的任何部位穿透到躯干中,或防止前臂 / 手彼此穿透。在其他示例中,碰撞测试可防止一条腿穿透另一条腿。在某些实施例中,碰撞测试可被应用于两个或更多个玩家的模型来防止类似的场景在模型之间出现。要理解,这可以通过模型几何形状的许多不同表示来实现——例如,可使用多边形壳用于核心身体,而使用参数胶囊(圆柱,其可在相对端处具有不同半径)作为肢段。在某些实施例中,碰撞测试可被应用于身体模型和 / 或骨架模型。在某些实施例中,碰撞测试可被应用于网格模型的某些多边形。作为示例,图 21 示出模型 290,其中模型 290 的手和前臂已穿透躯干,如在 292 处所示。在应用碰撞测试后,这种穿透可被检测和纠正,如在 294 处所示。

[0163] 碰撞测试可以按任何合适的方式来被应用。一种方法检查一个“体线段”(volumetric line segment)与另一个的碰撞,其中体线段可以是具有在 3-D 中向外延伸的半径的线段。这一碰撞测试的一示例可以是检查一条前臂对另一条前臂。在某些实施例中,体线段在该线段的每一端可具有不同的半径。

[0164] 另一种方法检查体线段与摆好姿态的多边形对象的碰撞。这一碰撞测试的一示例可以是检查一条前臂对躯干。在某些实施例中,摆好姿态的多边形对象可以是变形的多边形对象。

[0165] 在某些实施例中,关于目标的不同部分可以在一时间帧(例如,一秒的 1/30 或 1/60)内移动有限距离的知识可用作约束。这一约束可用于排除由将力施加于模型的像素接收位置而引起的某些姿态。

[0166] 如 145 处所示,在调整并任选地约束了模型之后,该过程可循环回到开始新一次将模型光栅化成新的合成的深度图像,该新的合成的深度图像然后可以与观察到的深度图像进行比较以便对模型作出进一步的调整。以此方式,该模型可被渐进地调整以更接近地表示所建模的目标。每一帧可完成实际上任何次数的迭代。更多迭代可得到更准确的结果,但更多的迭代也可能需要更多的计算开销。在许多情况下,每帧两次或三次迭代被认为是合适的,但是在某些实施例中一次迭代可能就足够了。

[0167] 在 146,图 7 的方法 100 任选地包括响应于对模型的改变,如图 11B 所示的改变,

来改变屏幕上人物(例如,图 12A 的玩家化身 190)的视觉外观。例如,游戏控制台可如此处所述地跟踪正在游戏控制台(例如,图 1A 和 1B 的游戏系统 12)上玩电子游戏的用户。具体地,可使用包括骨架模型(例如,图 11A 的骨架模型 180)的身体模型(例如,图 3 的身体模型 70)来对目标游戏玩家建模,并且可使用该身体模型来渲染屏幕上玩家化身。当游戏玩家伸直一条手臂时,游戏控制台可跟踪该运动,然后响应于所跟踪的运动来如图 11B 所示地调整模型 180。游戏控制台还可如上所述地应用一个或多个约束。在作出此类调整并应用此类约束之后,游戏控制台可显示经调整的玩家化身 192,如图 12B 所示。这也在图 1A 中作为示例示出,其中玩家化身 24 被示为响应于游戏玩家 18 在真实空间中挥出重拳而用拳猛击拳击对手 22。

[0168] 如上所述,可出于除了改变屏幕上人物或化身的视觉外观之外的其他目的来执行视觉目标识别。由此,不需要在所有实施例中都改变屏幕上人物或化身的视觉外观。如上所讨论的,目标跟踪可被用于实际上无限的不同目的,其中许多都不会导致屏幕上人物的改变。目标跟踪和 / 或被调整的模型的姿态可被用作参数来影响诸如游戏之类的应用的实际上任何元素。

[0169] 如 147 处所示,可对后续的帧重复上述过程。

[0170] 应该理解,此处所述的配置和 / 或方法在本质上是示例性的,且不应认为这些具体实施例或示例是局限性的,因为众多变体是可能。此处所述的具体例程或方法可表示任何数量的处理策略中的一个或多个。由此,所示出的各个动作可以按所示顺序执行、按其他顺序执行、并行地执行、或者在某些情况下被省略。同样,可以改变上述过程的次序。

[0171] 本公开的主题包括各种过程、系统和配置、此处所公开的其他特征、功能、动作、和 / 或特性、以及其任何和全部等效物的所有新颖和非显而易见的组合和子组合。

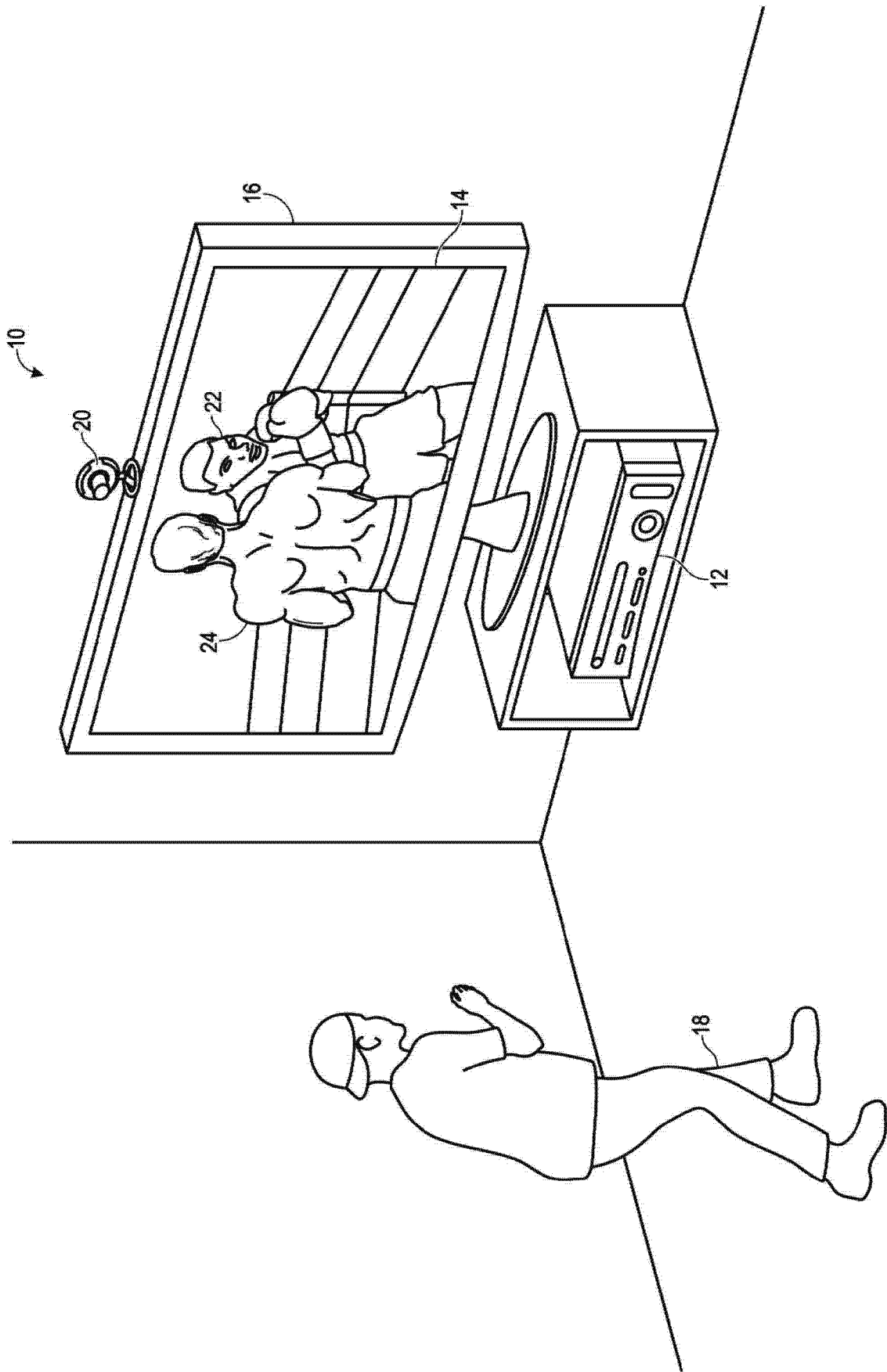


图 1

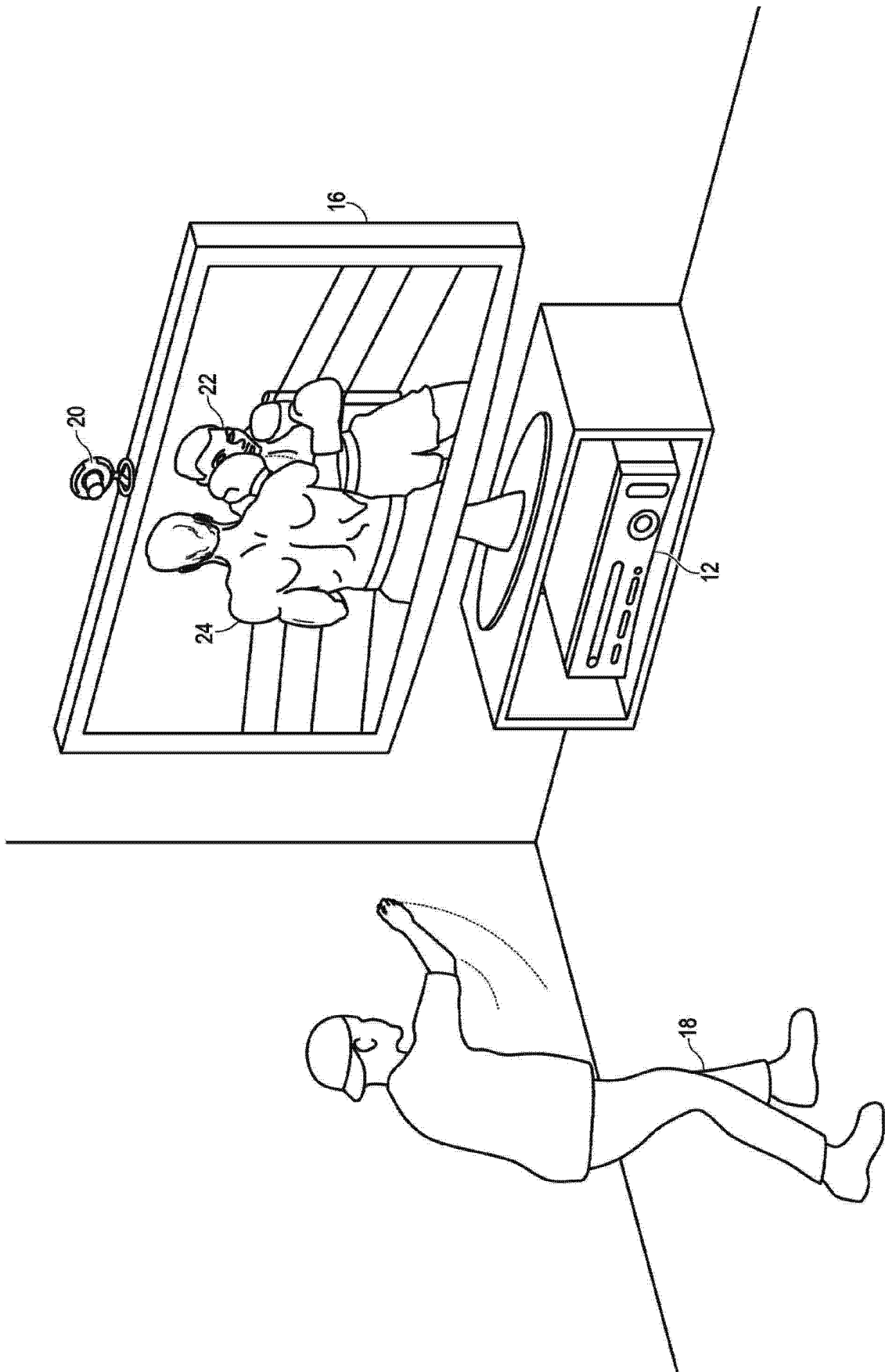


图 2

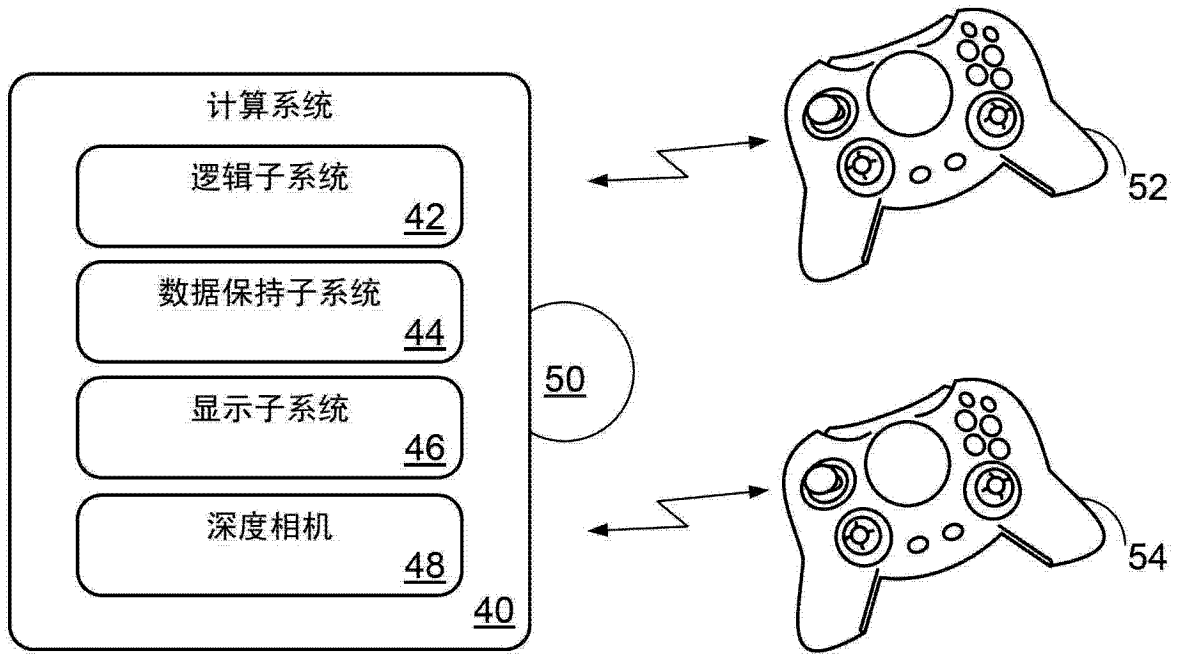


图 2

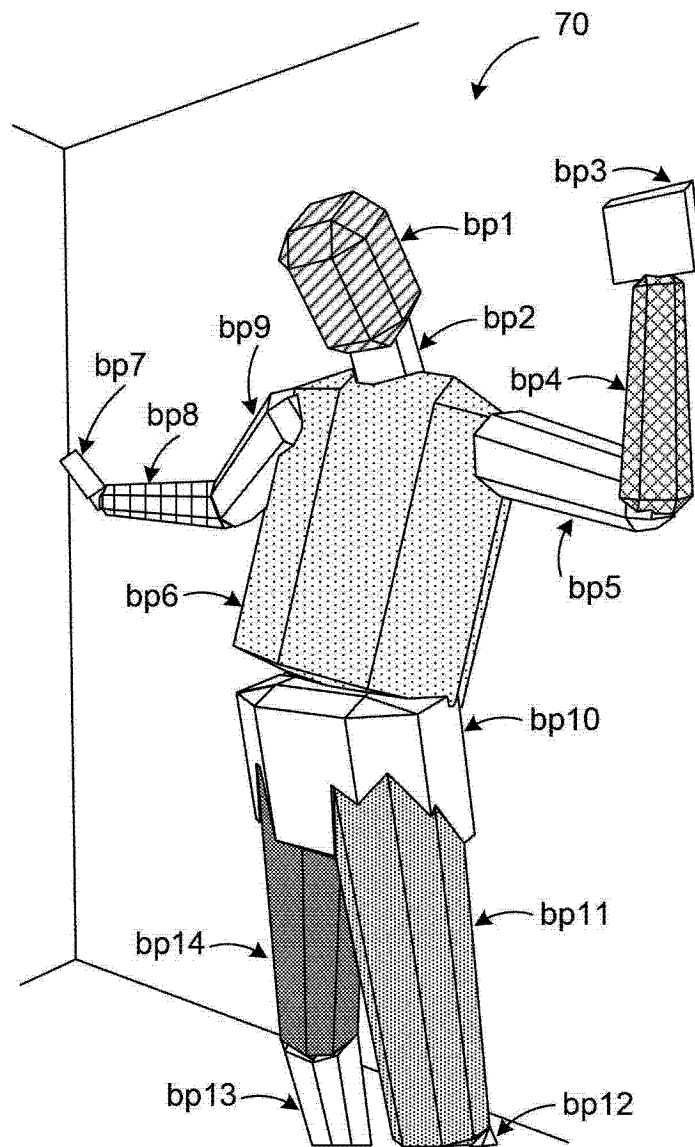


图 3

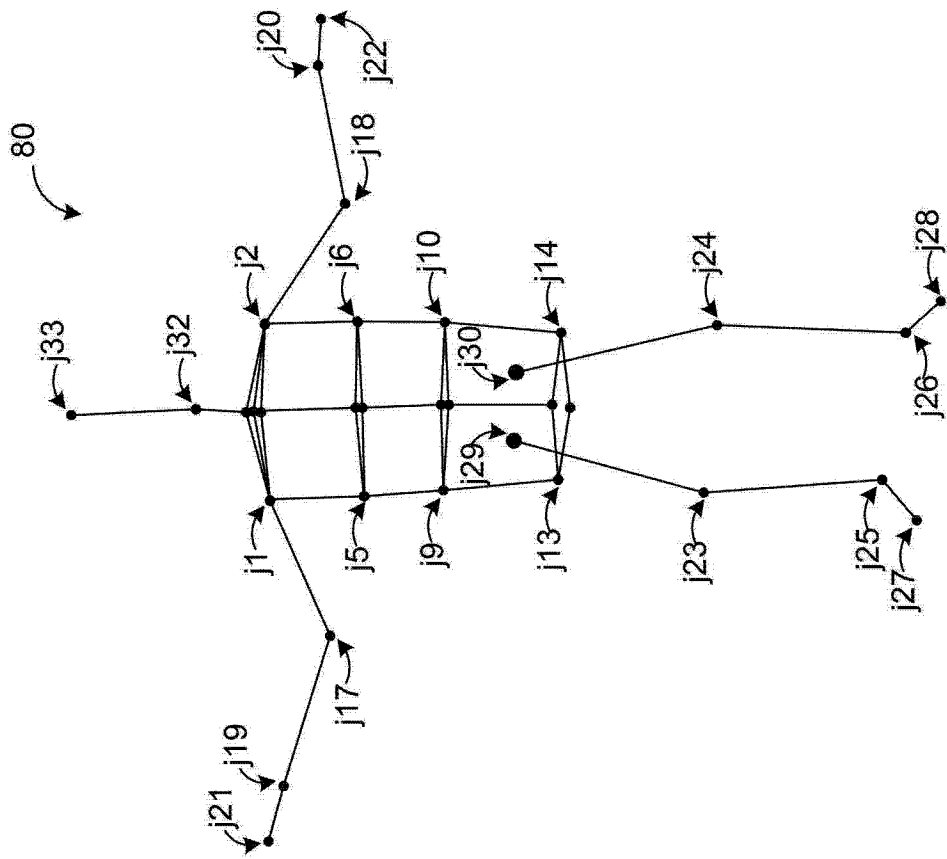


图 4

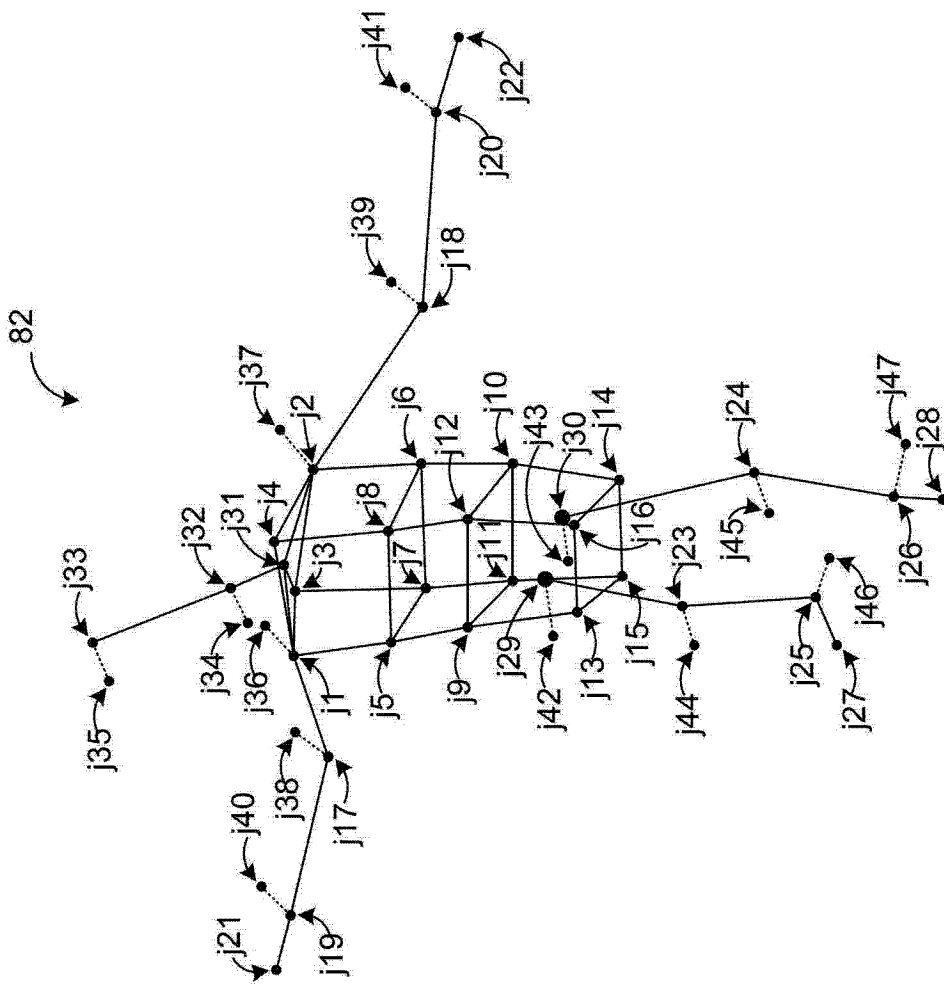


图 5

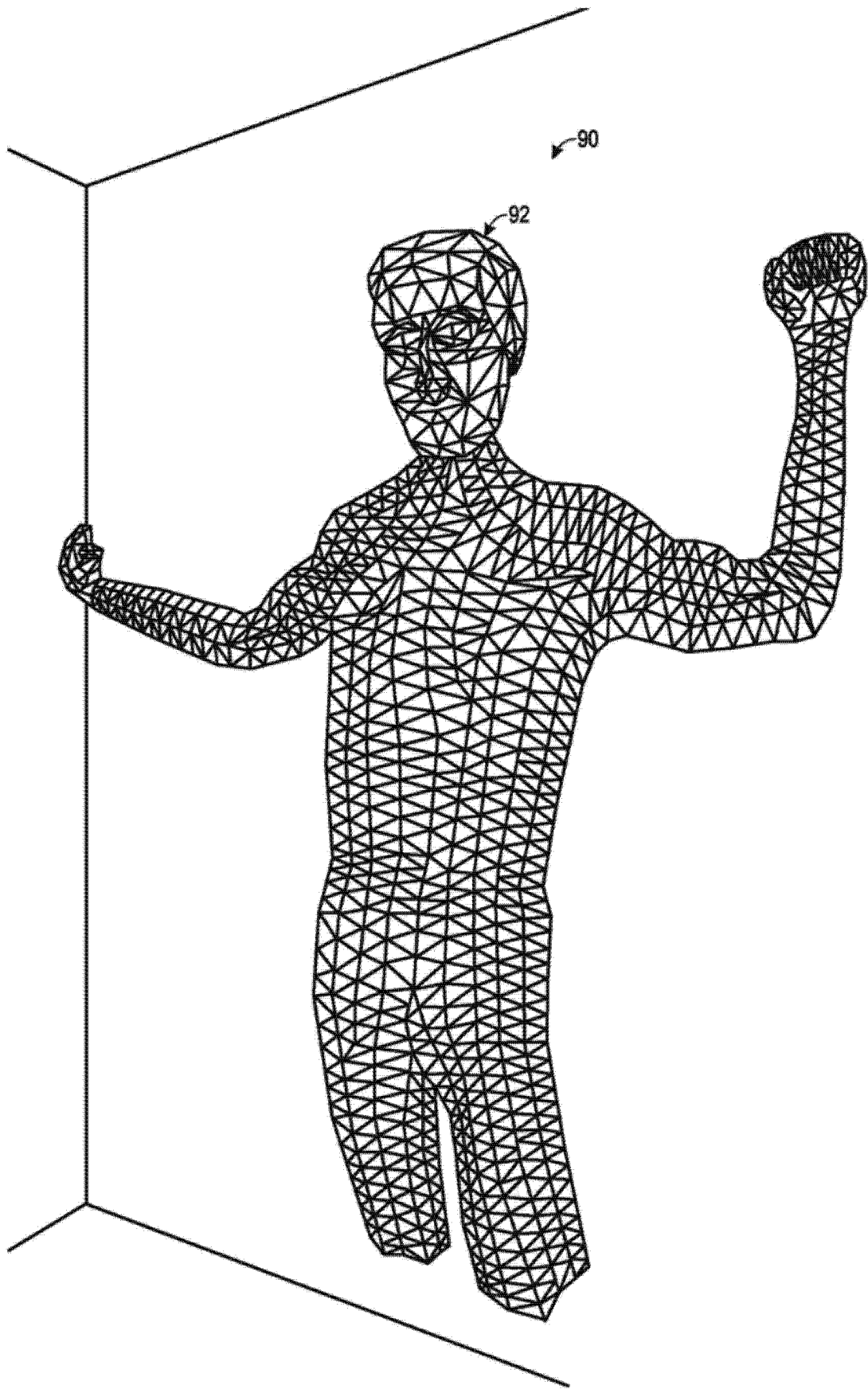


图 6

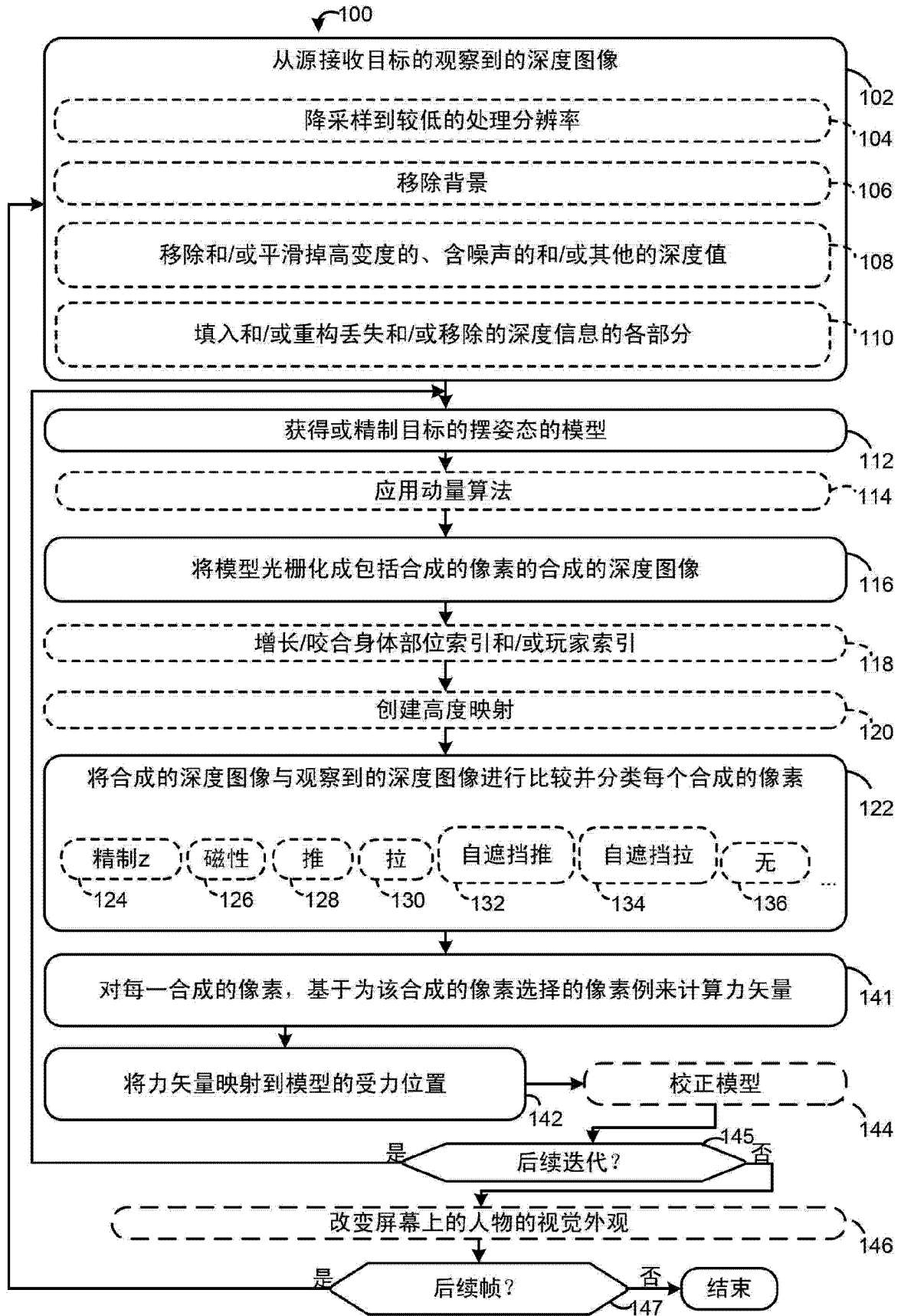


图 7

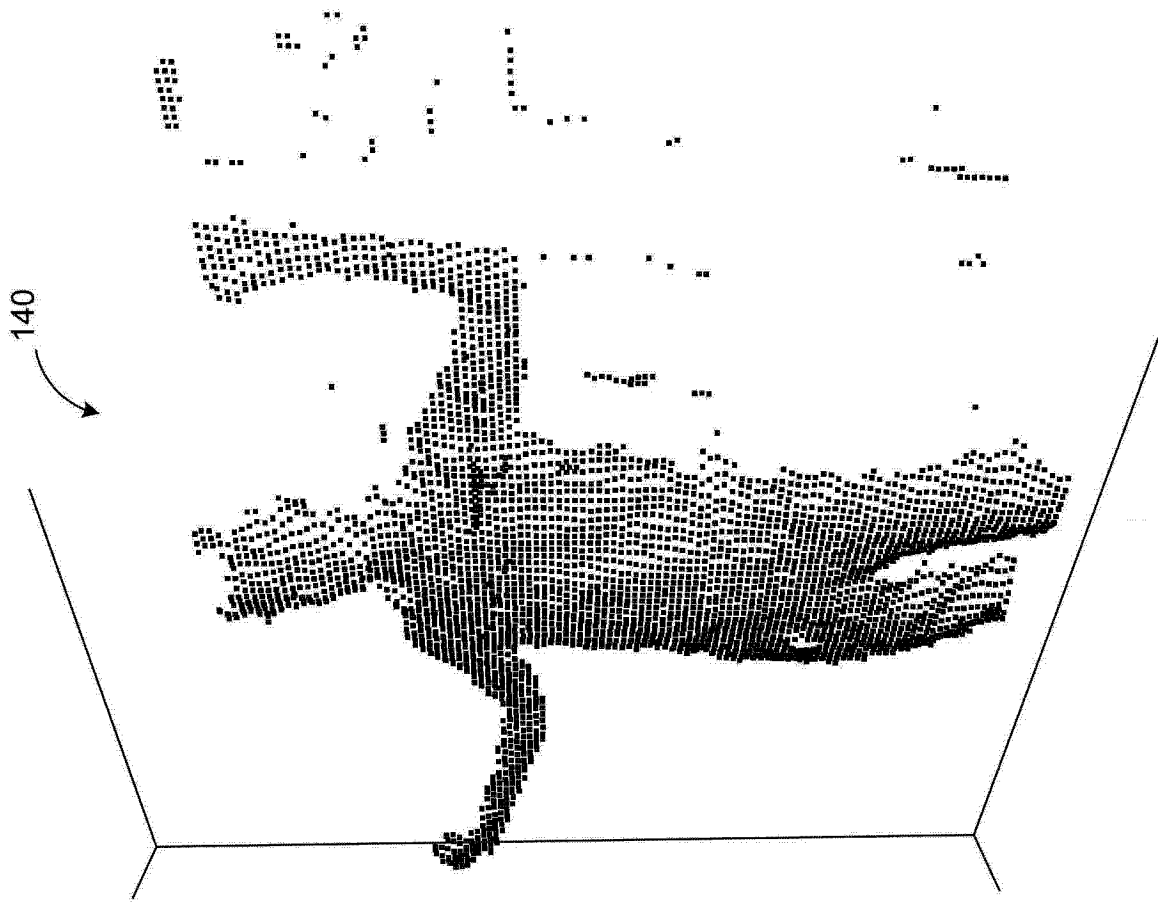


图 8

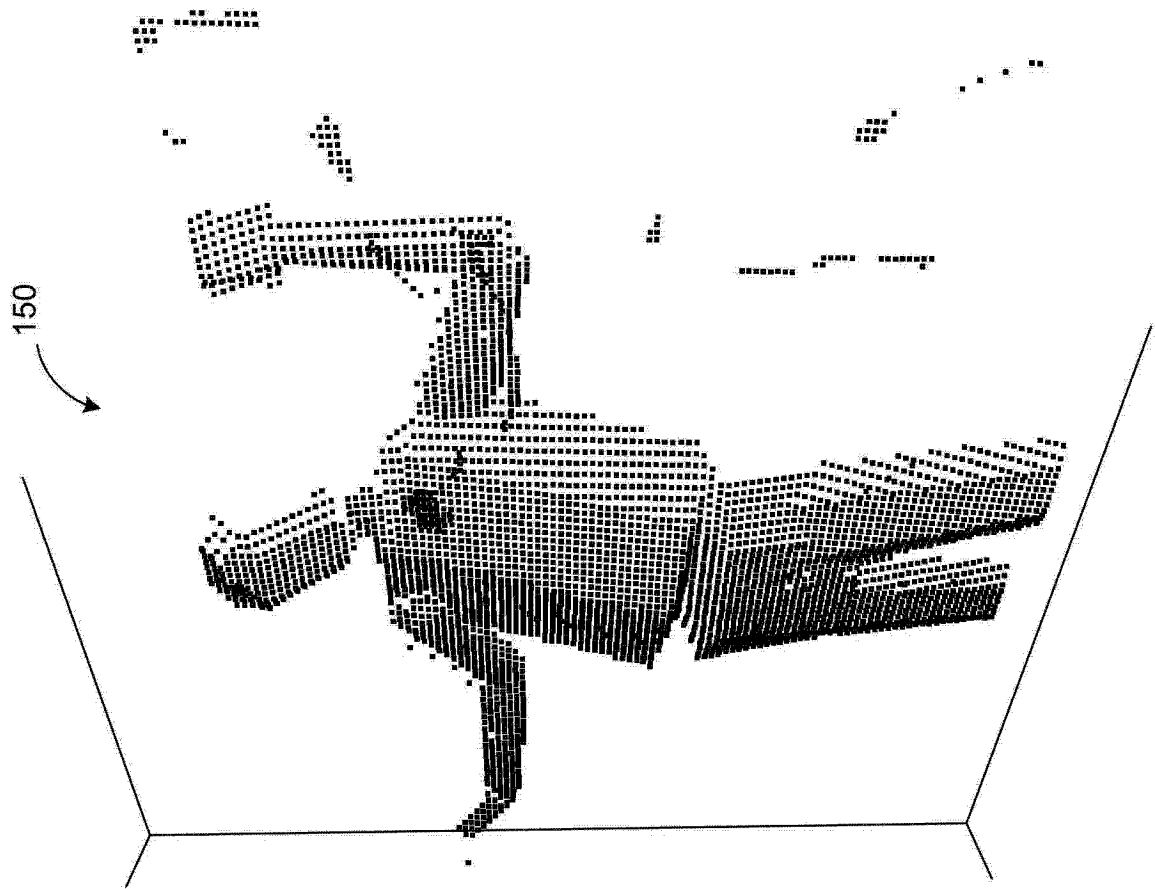


图 9

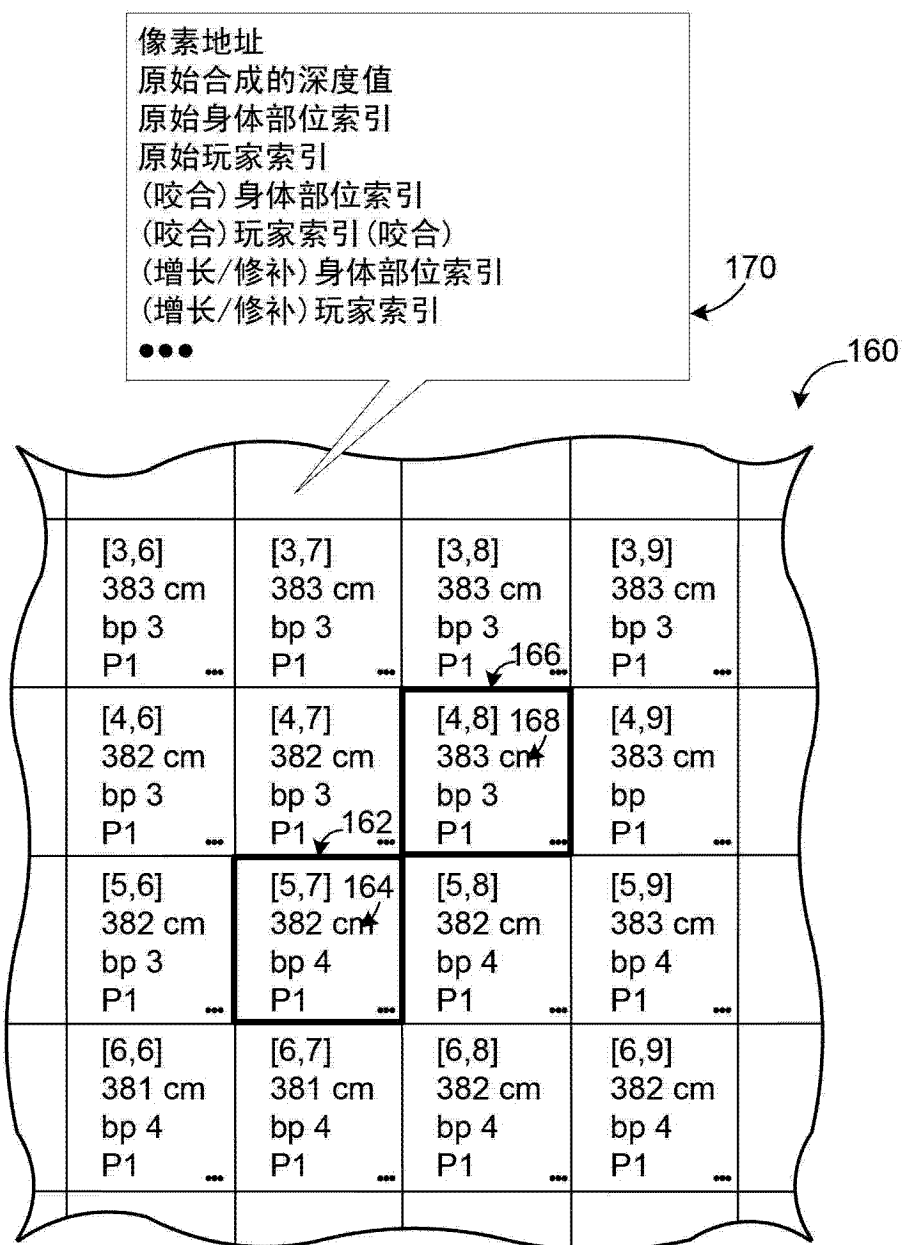


图 10

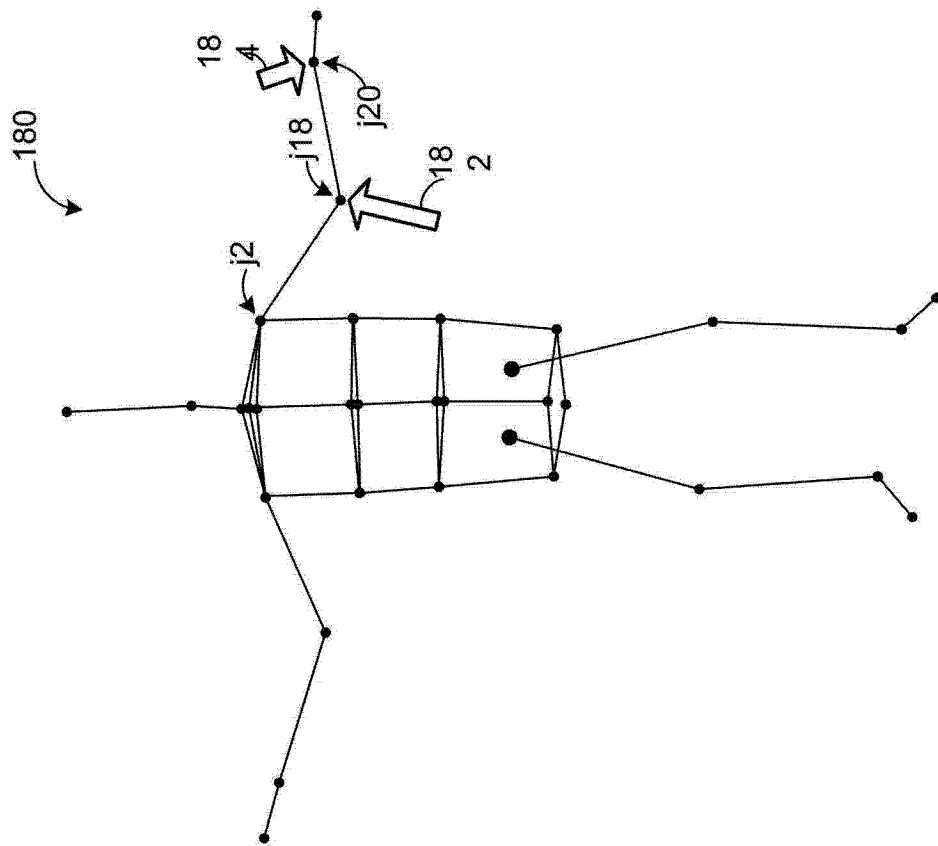


图 11A

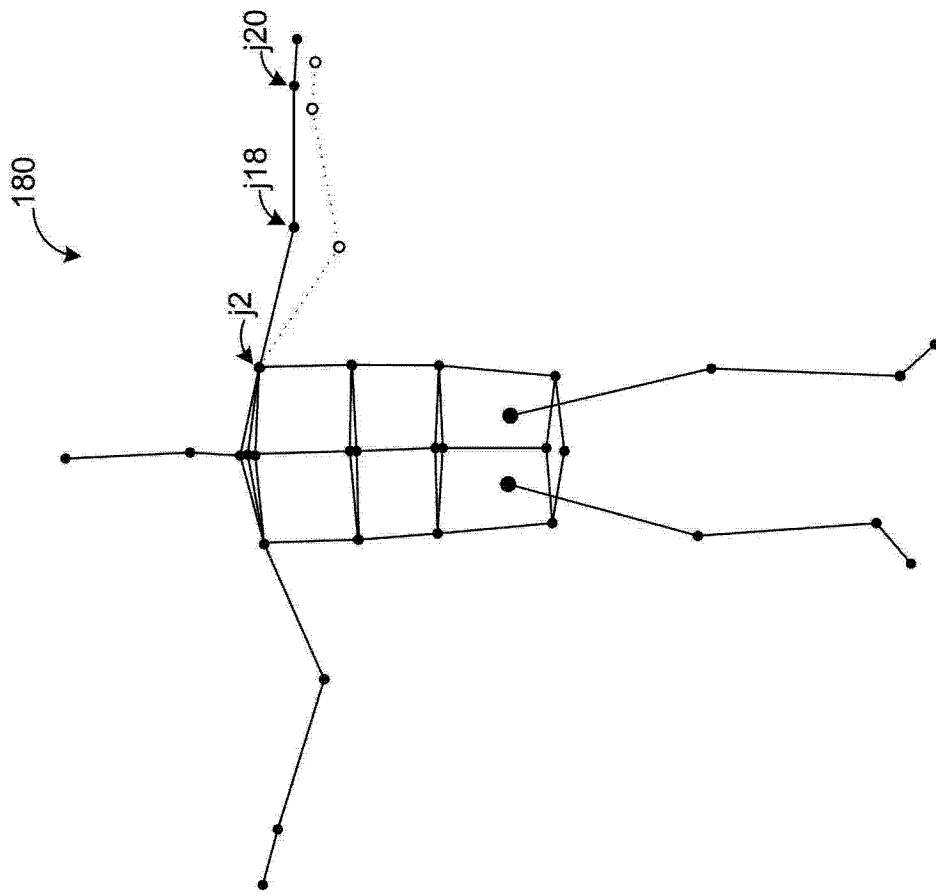


图 11B

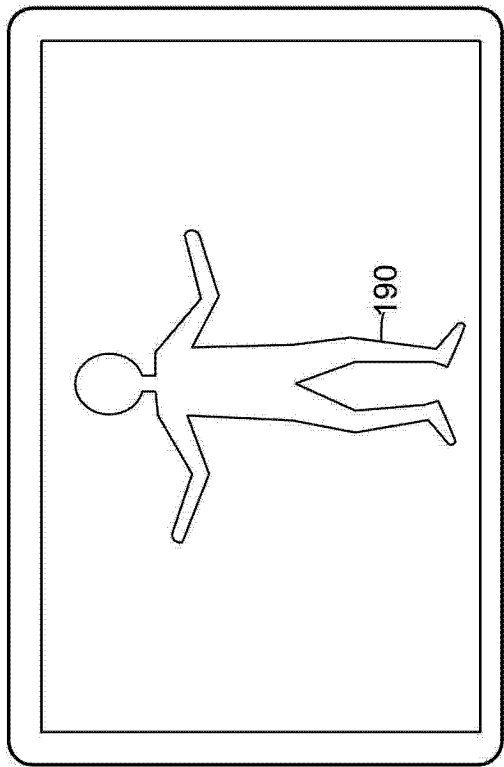


图 12A

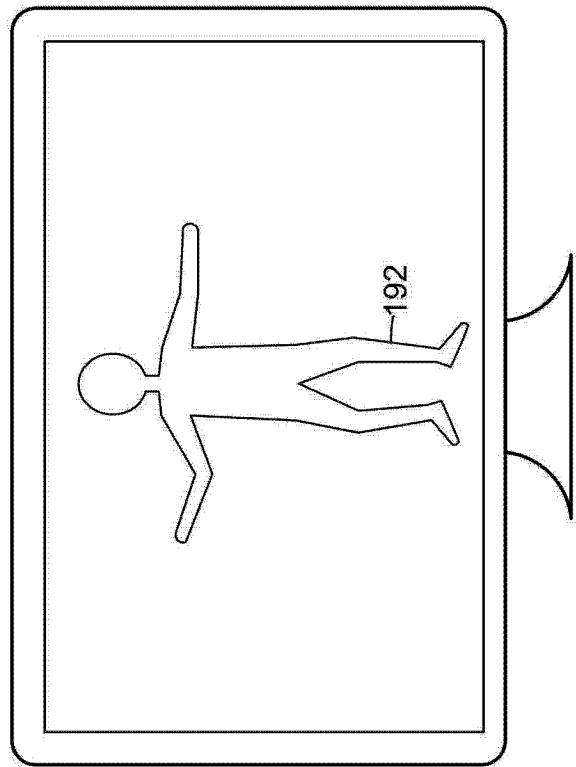


图 12B

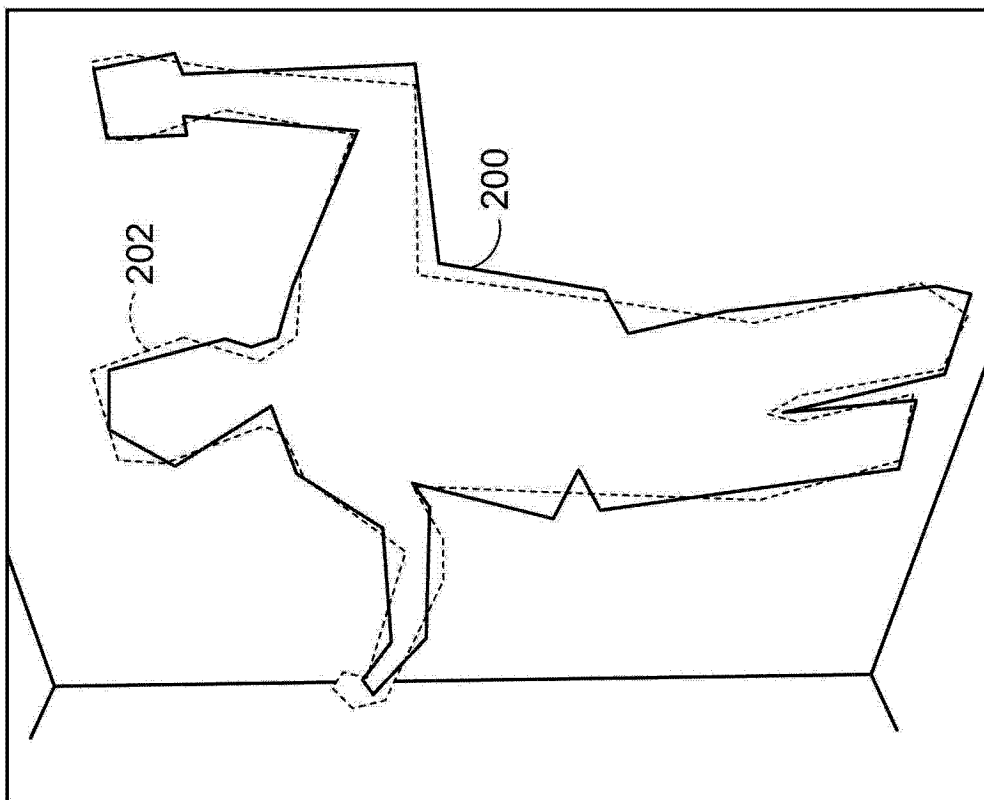


图 13

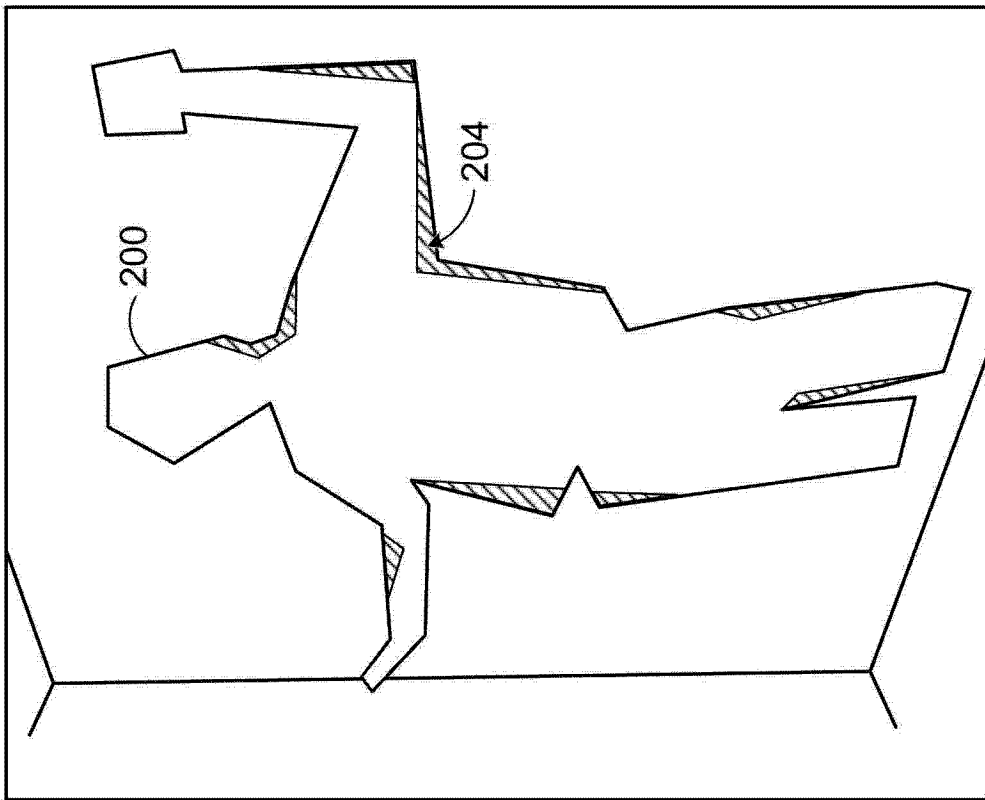


图 14

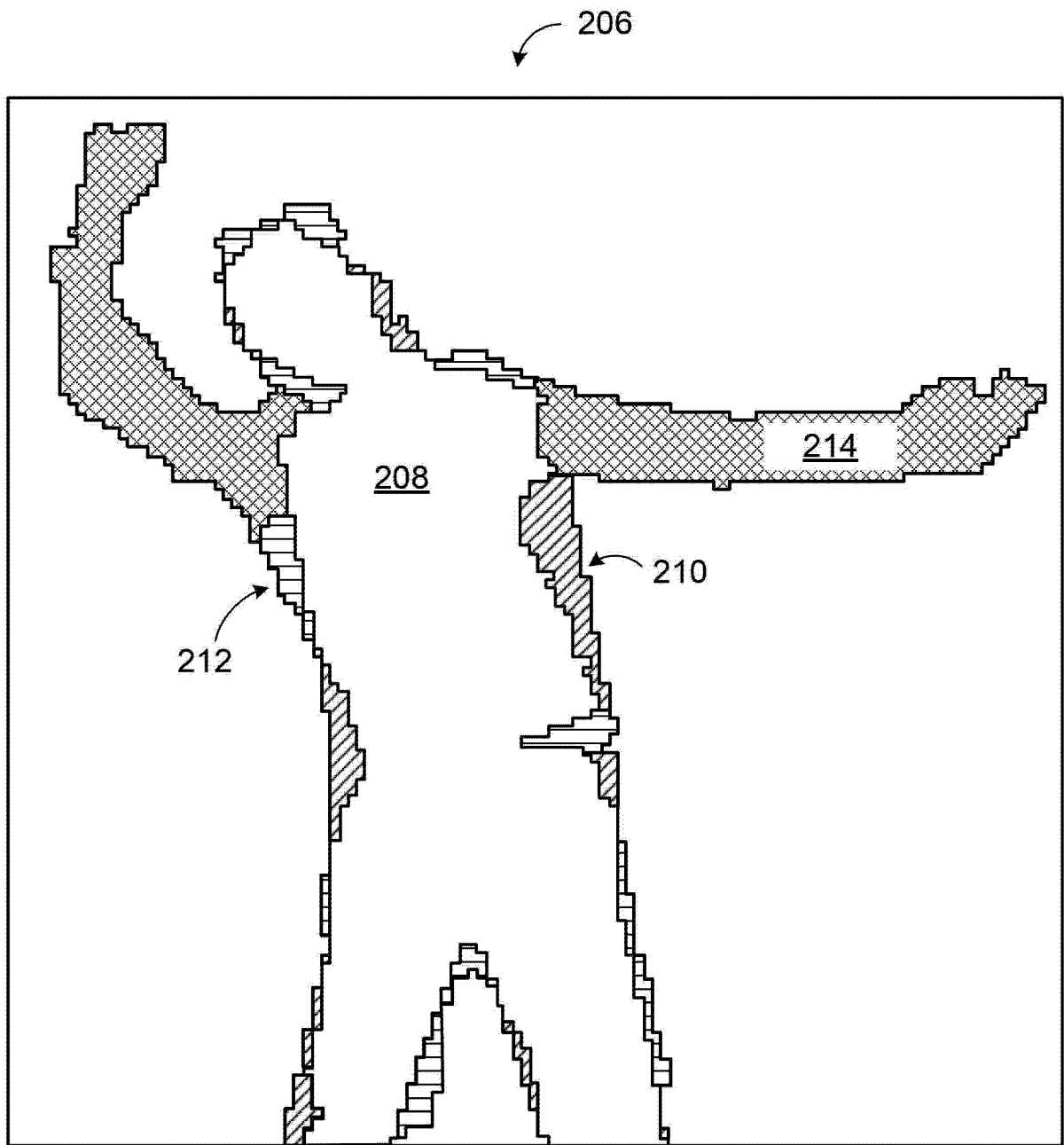


图 15

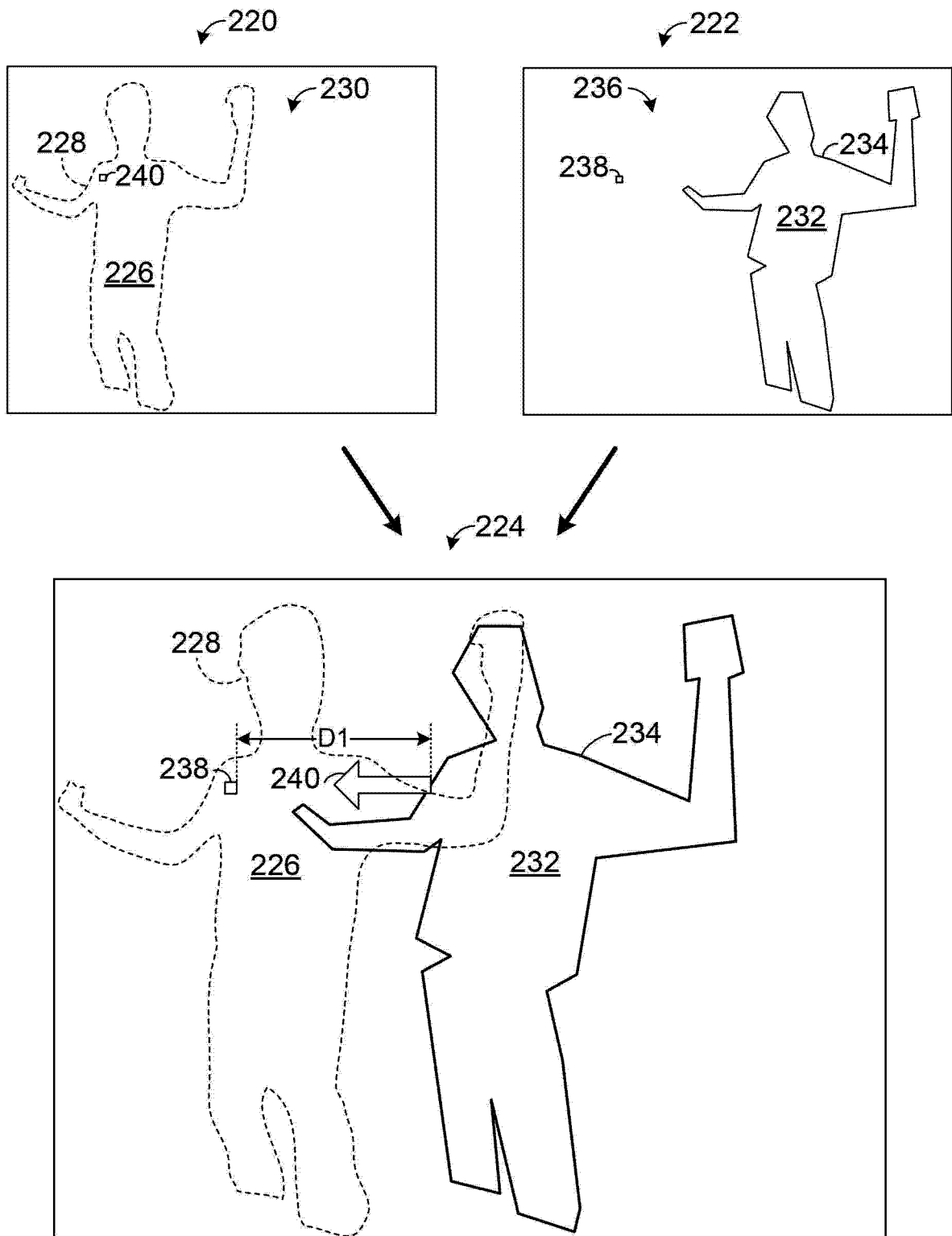


图 16

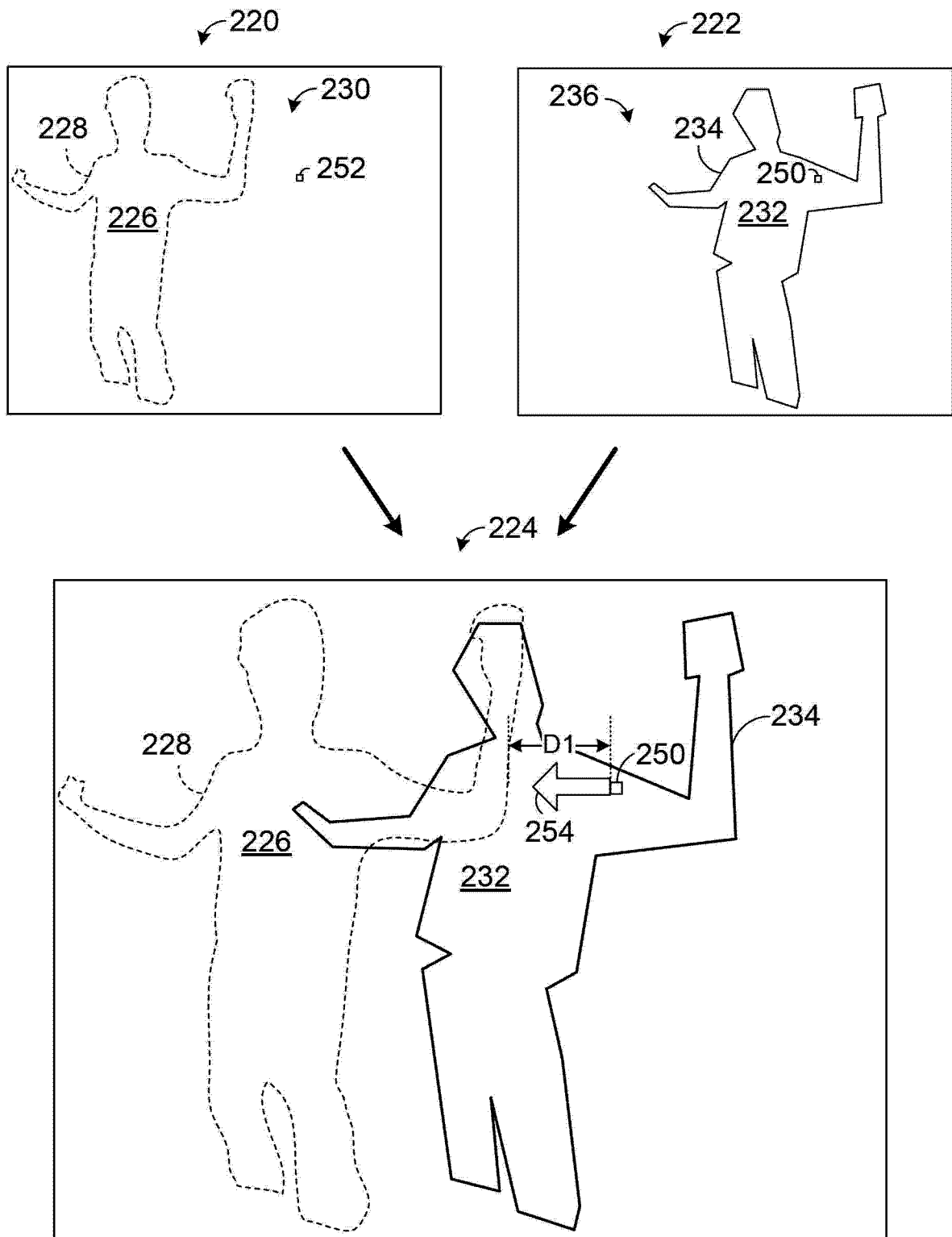


图 17

关节	受力?	像素例						
		1	2	3	4	5	6	7
j1	是	X	X	X	X	X		X
j2	是	X	X	X	X	X		X
j3	是	X	X	X	X	X		X
j4	是	X	X	X	X	X		X
j5	是	X	X	X	X	X		X
j6	是	X	X	X	X	X		X
j7	是	X	X	X	X	X		X
j8	是	X	X	X	X	X		X
j9	是	X	X	X	X	X		X
j10	是	X	X	X	X	X		X
j11	是	X	X	X	X	X		X
j12	是	X	X	X	X	X		X
j13	是	X	X	X	X	X		X
j14	是	X	X	X	X	X		X
j15	是	X	X	X	X	X		X
j16	是	X	X	X	X	X		X
j17	是						X	X
j18	是						X	X
j19	是						X	X
j20	是						X	X
j21	是						X	X
j22	是						X	X
j23	是	X	X	X	X	X		X
j24	是	X	X	X	X	X		X
j25	是	X	X	X	X	X		X
j26	是	X	X	X	X	X		X
j27	是	X	X	X	X	X		X
j28	是	X	X	X	X	X		X
j29	否	-	-	-	-	-	-	-
j30	否	-	-	-	-	-	-	-
j31	否	-	-	-	-	-	-	-
j32	是	X	X	X	X	X		X
j33	是	X	X	X	X	X		X
j34 - j47	否	-	-	-	-	-	-	-

图 18

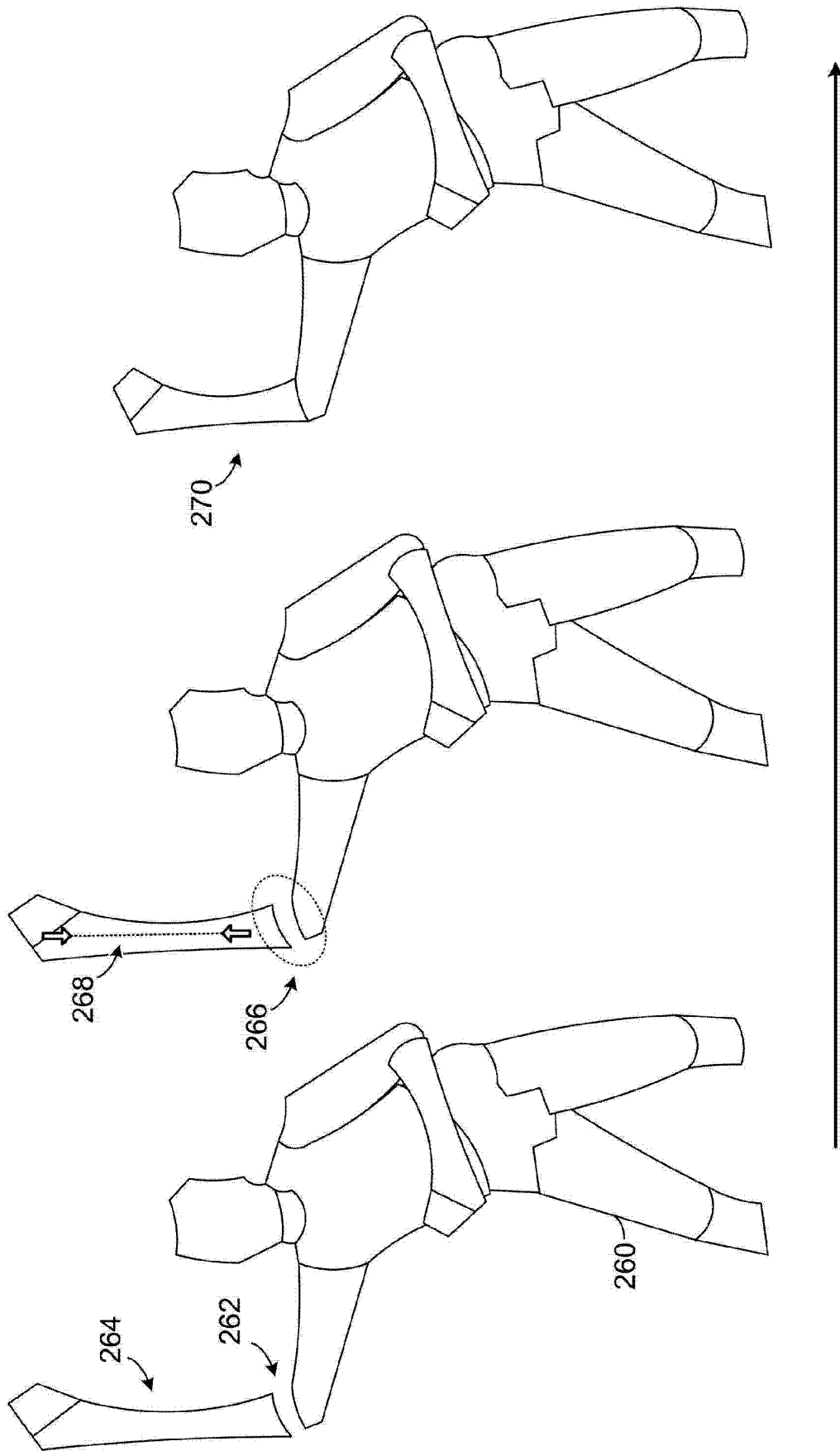


图 19

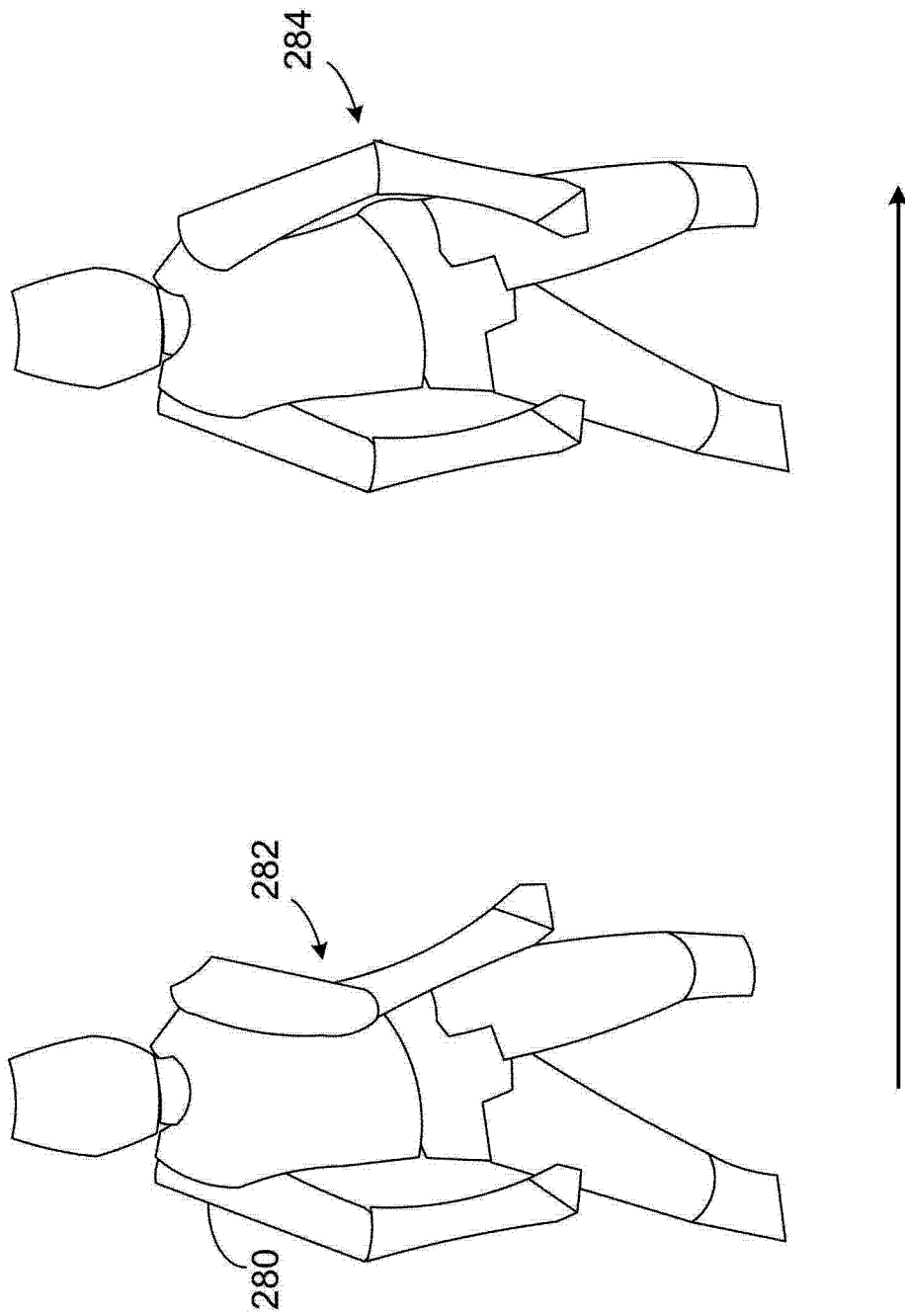


图 20

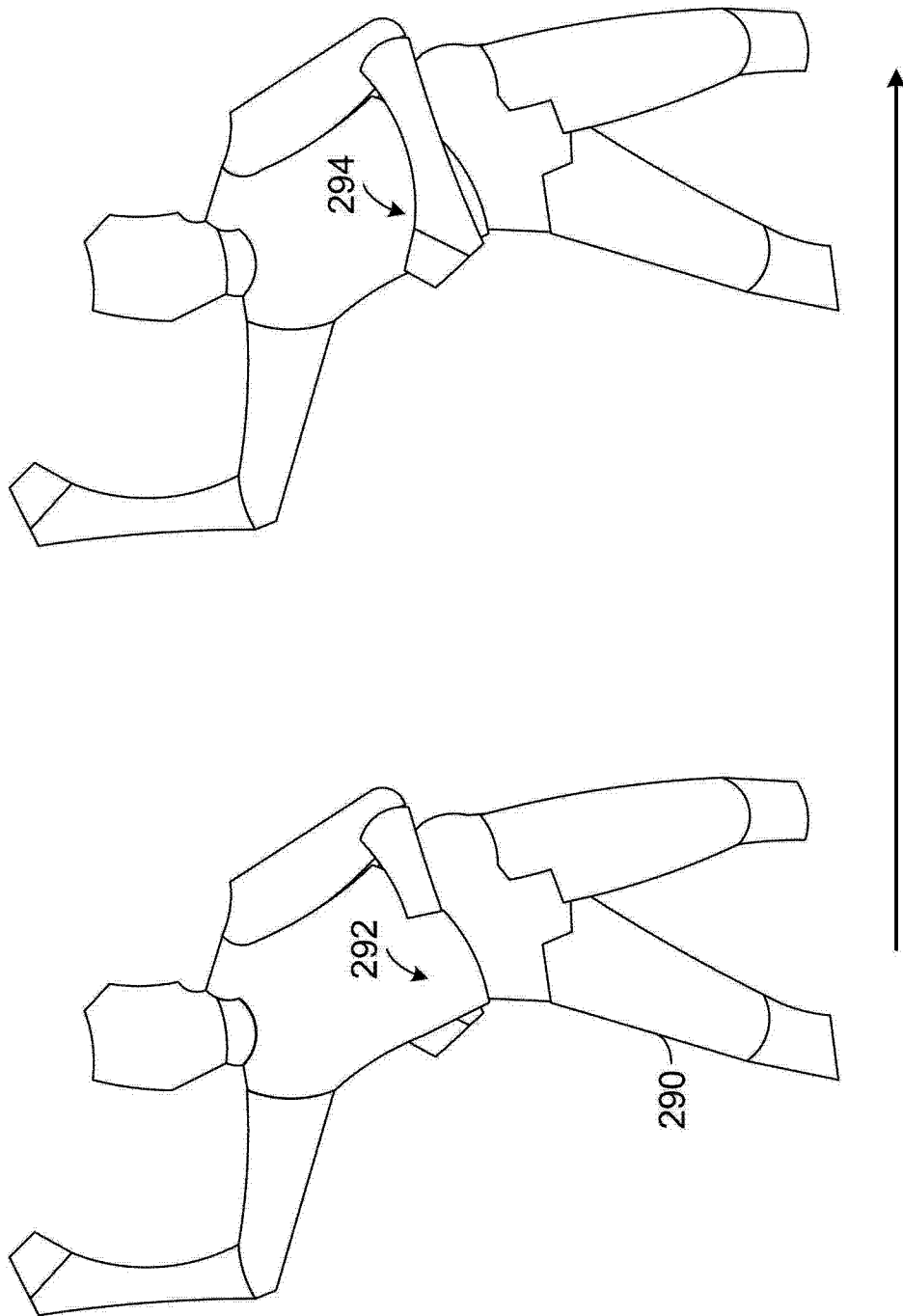


图 21