



(12) 发明专利

(10) 授权公告号 CN 102301311 B

(45) 授权公告日 2013. 07. 24

(21) 申请号 201080006306. 8

G06F 15/00 (2006. 01)

(22) 申请日 2010. 01. 22

(56) 对比文件

(30) 优先权数据

61/148, 884 2009. 01. 30 US

12/391, 150 2009. 02. 23 US

US 2007/0143715A1 A1, 2007. 06. 21, 全文.

JP 特开 2008-250823 A, 2008. 10. 16, 全文.

CN 1764931 A, 2006. 04. 26, 全文.

(85) PCT申请进入国家阶段日

审查员 俞立文

2011. 07. 28

(86) PCT申请的申请数据

PCT/US2010/021883 2010. 01. 22

(87) PCT申请的公布数据

W02010/088156 EN 2010. 08. 05

(73) 专利权人 微软公司

地址 美国华盛顿州

(72) 发明人 S·G·拉塔 K·兹努达 K·盖斯那

R·马尔科维奇 D·贝内特

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 陈斌

(51) Int. Cl.

G06F 3/00 (2006. 01)

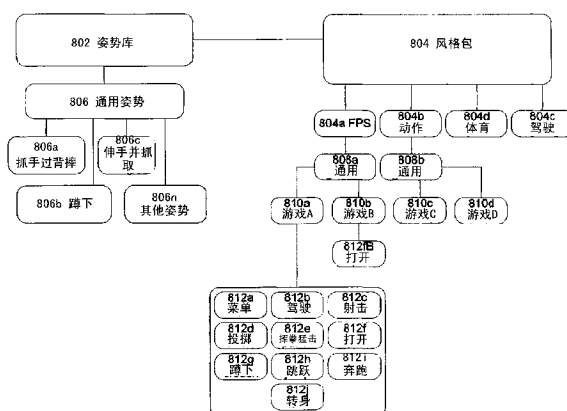
权利要求书2页 说明书15页 附图12页

(54) 发明名称

标准姿势

(57) 摘要

公开了用于将标准姿势的互补集分组成姿势库的系统、方法和计算机可读介质。姿势可以是互补的, 在于它们在某一上下文中经常被一起使用或在于它们的参数是相互关联的。在用第一值设置了某一姿势的参数, 该姿势的以及姿势包中依赖于该第一值的其他姿势的所有其他参数可用使用第一值确定的它们自己的值来设置。



1. 一种用于向应用提供互补姿势过滤器的包的方法,包括:

提供包括多个过滤器的包,每个过滤器包括关于姿势的信息,至少一个过滤器与包中的至少一个其他过滤器互补(902);

接收将第一值赋予第一过滤器的一参数的指示(904);

将所述值赋予所述参数(906);以及

将第二值赋予第二过滤器的第二参数,所述第二值是使用所述第一值确定的(908)。

2. 如权利要求1所述的方法,其特征在于,所述第一过滤器的所述参数表示身体部位、空间体、速度、移动方向、角度、二维(2D)平面、或移动发生的场所。

3. 如权利要求1所述的方法,其特征在于,所述姿势包括蹲下、跳跃、倚靠、抓手过背摔、投掷、挥臂、躲避、踢或阻挡。

4. 如权利要求1所述的方法,其特征在于,过滤器在以下情况下与包中的至少一个其他过滤器互补,(i)该过滤器具有基于包中的所述至少一个其他过滤器的参数确定的至少一个参数;(ii)该过滤器表示通常由用户在包中的所述至少一个其他过滤器所表示的姿势的短时间段内作出的姿势;或(iii)该过滤器所表示的姿势能够与包中的所述至少一个其他过滤器所表示的姿势同时作出。

5. 如权利要求1所述的方法,其特征在于,将所述第一值赋予所述参数的指示是作为用户作出所述姿势的结果而接收的。

6. 如权利要求1所述的方法,其特征在于,所述第一值是固定值、范围或具有容错的值。

7. 如权利要求1所述的方法,其特征在于,所述第二值是使用所述第一值基于比例关系、反比例关系、线性关系、指数关系、或取第一值作为输入的函数来确定的。

8. 如权利要求1所述的方法,其特征在于,所述包包含某一风格的互补姿势过滤器,其中该风格是第一人称射击、动作、驾驶或体育风格之一。

9. 如权利要求1所述的方法,其特征在于,所述应用是视频游戏或操作系统。

10. 如权利要求1所述的方法,其特征在于,还包括:

在将第二值赋予第二过滤器的第二参数之后,接收将第三值赋予所述第二参数的指示;以及

将所述第三值赋予所述第二参数。

11. 一种用于向应用提供互补过滤器的包的方法,包括:

将过滤器包提供给所述应用,所述过滤器包被包括在包含过滤器的至少一个包的过滤器库中,每个过滤器包括关于姿势的信息,所述至少一个包中的至少一个过滤器与所述包中的至少一个其他过滤器互补(802);

接收对应于用户的至少一部分的至少一个图像的数据,

确定该数据是否对应于至少两个匹配的过滤器,

确定匹配的过滤器中的第一过滤器如何与匹配的过滤器中的至少一个其他过滤器一起操作,以及

将对应于每个匹配的过滤器的输出发送到所述应用(12)。

12. 如权利要求11所述的方法,其特征在于,所述输出包括置信度水平,通过基于至少一个其他匹配的过滤器的置信度水平来改变所述第一过滤器的置信度水平来确定第一匹

配的过滤器如何与所述至少一个其他匹配的过滤器一起操作。

13. 如权利要求 11 所述的方法,其特征在于,

在对包中不存在的过滤器解析所述数据之前,对包中的每个过滤器解析所述数据以确定所述数据是否指示与所述包中的一个或多个过滤器的匹配。

14. 如权利要求 11 所述的方法,其特征在于,

在对不能与对应于所述数据的所述至少一个过滤器同时解析的过滤器解析所述数据之前,对所述包中可与对应于所述数据的所述至少一个过滤器同时解析的每个其他过滤器解析所述数据。

15. 一种用于向应用提供互补姿势过滤器的包的方法,包括:

提供包括多个过滤器的包,每个过滤器包括关于姿势的信息,每个过滤器与包中的至少一个其他过滤器互补,姿势是由用户在作出与由距离相机捕捉的该姿势相关联的运动或姿态时输入的(902);

接收将第一值赋予第一过滤器的一参数的指示(904);

将所述值赋予所述参数(906);以及

将第二值赋予第二过滤器的第二参数,所述第二值是使用所述第一值确定的(908)。

标准姿势

[0001] 发明背景

[0002] 诸如计算机游戏、多媒体应用、办公应用等的许多计算应用使用控制来允许用户操纵游戏角色或应用的其他方面。通常使用,例如,控制器、遥控器、键盘、鼠标等等,来输入这样的控制。不幸的是,这些控制可能是难以学习的,由此造成了用户和这些游戏及应用之间的障碍。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不同。例如,使得游戏角色挥动棒球拍的游戏控制可能不与挥动棒球拍的实际动作相对应。

[0003] 发明概述

[0004] 这里公开了用于接收反映用户的骨架移动的数据以及从该数据确定用户是否已经执行了一个或多个姿势的系统和方法。公开了标准姿势包,应用开发者可从标准姿势包将姿势识别并入到他们的应用中。

[0005] 在一示例实施例中,姿势库包括多个姿势。在这些姿势彼此互补的情况下,它们可被分组成姿势包。这些姿势包然后被提供给应用供姿势识别器引擎使用。应用可利用一个或多个姿势包。

[0006] 应用可向姿势的第一参数赋一值,例如要被识别的美式足球投掷姿势所必须达到的手臂速度最低阈值。识别器引擎用该值设置第一参数,并设置该姿势或姿势包中依赖于该第一姿势的值的任何其他姿势的任何其他参数的值。例如,在姿势包是包括美式足球投掷姿势的运动姿势包的情况下,该包还可包括与美式足球投掷姿势相互关联的曲线球棒球投掷姿势以及快速球棒球投掷姿势。在已经确定曲线球棒球投掷姿势应具有美式足球投掷姿势的手臂速度最低阈值的 80%、以及快速球棒球投掷姿势应具有美式足球投掷姿势的手臂速度最低阈值的 90%的情况下,可分别将那些参数设置为该值的 80%和 90%。

[0007] 本领域技术人员将理解,本发明的一个或多个方面可包括但不限于用于实现本发明的本文所述方面的电路和 / 或编程;该电路和 / 或编程实质上可以是配置成实现本文所述方面的硬件、软件和 / 或固件的任何组合,这取决于系统设计者的设计选择。

[0008] 以上是概述,并且因此必然包含细节的简化、一般化及省略。本领域技术人员将明白,本概述只是说明性的并且决不旨在是限制性的。

[0009] 附图简述

[0010] 参考附图来进一步描述根据本说明书的用于互补姿势的姿势包的系统、方法和计算机可读介质,在附图中:

[0011] 图 1A 和 1B 示出了伴随用户玩游戏的目标识别、分析和跟踪系统的示例实施例。

[0012] 图 2 示出了可在目标识别、分析和跟踪系统中使用的捕捉设备的示例实施例。

[0013] 图 3A 示出了可用于在目标识别、分析和跟踪系统中解释一个或多个姿势的计算环境的示例实施例。

[0014] 图 3B 示出了可用于在目标识别、分析和跟踪系统中解释一个或多个姿势的计算环境的另一示例实施例。

[0015] 图 4A 示出了从图 2 的目标识别、分析和跟踪系统生成的用户的骨架映射。

[0016] 图 4B 示出了图 2 所示的姿势识别器架构的进一步细节。

[0017] 图 5A 和 5B 示出了可如何堆叠姿势过滤器来创建更复杂的姿势过滤器。

[0018] 图 6A、6B、6C、6D 和 6E 示出了用户 502 在美式足球视频游戏中可以作出的用以发出“安全接球”(fair catch) 的信号示例姿势。

[0019] 图 7A、7B、7C、7D 和 7E 示出了在解析了图像数据的每一帧来产生用户的骨架映射时图 6A-E 的示例“安全接球”姿势。

[0020] 图 8 示出了来自姿势库的通用姿势如何被分组成用于某一任务的互补姿势的风格包。

[0021] 图 9 示出了用于在应用为一个姿势的一个参数提供值时调谐姿势包中的互补姿势的示例操作过程。

[0022] 说明性实施例的详细描述

[0023] 如本文中描述的, 用户可通过执行一个或多个姿势来控制诸如游戏控制台、计算机等计算环境上执行的应用。根据一个实施例, 姿势可由例如捕捉设备来接收。例如, 捕捉设备可捕捉场景的深度图像。在一个实施例中, 捕捉设备可以确定场景中的一个或多个目标或对象是否对应于诸如用户等人类目标。为了确定场景中的目标或对象是否对应于人类目标, 可对每个目标进行泛色填充并将其与人体模型的图案作比较。然后可扫描匹配人体模型的每一目标或对象来生成与其相关联的骨架模型。该骨架模型然后可被提供给计算环境, 使得计算环境可跟踪该骨架模型, 呈现与该骨架模型相关联的化身, 以及可基于例如用户的已经从该骨架模型中识别出的姿势来确定要在计算环境上执行的应用中执行那些控制。姿势识别器引擎(其架构在下文中更完整地描述)被用于确定用户何时做出了某一姿势。

[0024] 图 1A 和 1B 示出伴随用户 18 玩拳击游戏的目标识别、分析和跟踪系统 10 的配置的示例实施例。在一示例实施例中, 目标识别、分析和跟踪系统 10 可用于识别、分析和 / 或跟踪诸如用户 18 等的人类目标。

[0025] 如图 1A 所示, 目标识别、分析和跟踪系统 10 可包括计算环境 12。计算环境 12 可以是计算机、游戏系统或控制台等等。根据一示例实施例, 计算环境 12 可包括硬件组件和 / 或软件组件, 使得计算环境 12 可用于执行诸如游戏应用、非游戏应用等的应用。

[0026] 如图 1A 所示, 目标识别、分析和跟踪系统 10 还可包括捕捉设备 20。捕捉设备 20 可以是例如相机, 该相机可用于在视觉上监视诸如用户 18 等一个或多个用户, 以使得可以捕捉、分析并跟踪一个或多个用户所执行的姿势, 以执行应用中的一个或多个控制或动作, 如将在下面更详细地描述的。

[0027] 根据一实施例, 目标识别、分析和跟踪系统 10 可连接到可向诸如用户 18 等的用户提供游戏或应用视觉和 / 或音频的视听设备 16, 如电视机、监视器、高清电视机 (HDTV) 等。例如, 计算环境 12 可包括诸如图形卡等视频适配器和 / 或诸如声卡等音频适配器, 这些适配器可提供与游戏应用、非游戏应用等相关联的视听信号。视听设备 16 可从计算环境 12 接收视听信号, 然后可向用户 18 输出与该视听信号相关联的游戏或应用视觉和 / 或音频。根据一个实施例, 视听设备 16 可经由例如, S- 视频电缆、同轴电缆、HDMI 电缆、DVI 电缆、VGA 电缆等连接到计算环境 12。

[0028] 如图 1A 和 1B 所示, 目标识别、分析和跟踪系统 10 可用于识别、分析和 / 或跟踪诸如用户 18 等的人类目标。例如, 可使用捕捉设备 20 来跟踪用户 18, 以使得可将用户 18 的

移动解释为可用于影响由计算机环境 12 执行的的应用的控制。因而,根据一个实施例,用户 18 可移动他或她的身体来控制应用。

[0029] 如图 1A 和 1B 所示,在一示例实施例中,在计算环境 12 上执行的应用可以是用户 18 可能正在玩的拳击游戏。例如,计算环境 12 可使用视听设备 16 来向用户 18 提供拳击对手 22 的视觉表示。计算环境 12 还可使用视听设备 16 来提供用户 18 可通过他的或她的移动来控制的玩家化身 24 的视觉表示。例如,如图 1B 所示,用户 18 可在物理空间中挥重拳来使得玩家化身 24 在游戏空间中挥重拳。因此,根据一示例实施例,目标识别、分析和跟踪系统 10 的计算环境 12 和捕捉设备 20 可用于识别和分析用户 18 在物理空间中的重拳,从而使得该重拳可被解释为对游戏空间中的玩家化身 24 的游戏控制。

[0030] 用户 18 的其他移动也可被解释为其他控制或动作,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,某些移动可被解释为可对应于除控制玩家化身 24 之外的动作的控制。例如,玩家可以使用运动来结束、暂停或保存游戏、选择级别、查看高分、与朋友交流等。

[0031] 在各示例实施例中,诸如用户 18 等的人类目标可具有一物体。在这些实施例中,电子游戏的用户可手持物体从而可以使用玩家和物体的运动来调整和 / 或控制游戏的参数。例如,可以跟踪并利用玩家手持球拍的运动来控制电子运动游戏中的屏幕上球拍。在另一示例实施例中,可以跟踪并利用玩家手持物体的运动来控制电子格斗游戏中的屏幕上武器。

[0032] 根据其他示例实施例,目标识别、分析和跟踪系统 10 还可用于将目标移动解释为游戏领域之外的操作系统和 / 或应用控制。例如,事实上操作系统和 / 或应用的任何可控方面可由诸如用户 18 等目标的移动来控制。

[0033] 图 2 示出可在目标识别、分析和跟踪系统 10 中使用的捕捉设备 20 的示例实施例。根据一示例实施例,捕捉设备 20 可被配置成经由任何合适的技术,包括例如飞行时间、结构化光、立体图像等来捕捉包括深度图像的带有深度信息的视频,该深度信息可包括深度值。根据一实施例,捕捉设备 20 可将所计算的深度信息组织为“Z 层”,或可与从深度相机沿其视线延伸的 Z 轴垂直的层。

[0034] 如图 2 所示,捕捉设备 20 可包括图像相机组件 22。根据一个示例实施例,图像相机组件 22 可以是捕捉场景的深度图像的深度相机。深度图像可包括被捕捉的场景的二维 (2-D) 像素区域,其中 2-D 像素区域中的每一个像素都可以 (例如以厘米、毫米等等为单位) 表示来自相机的被捕捉的场景中的对象的长度。

[0035] 如图 2 所示,根据一示例实施例,图像相机组件 22 可包括可用于捕捉场景的深度图像的 IR 光组件 24、三维 (3-D) 相机 26、和 RGB 相机 28。例如,在飞行时间分析中,捕捉设备 20 的 IR 光组件 24 可以将红外光发射到场景上,然后,可以使用传感器 (未示出),用例如 3-D 相机 26 和 / 或 RGB 相机 28,来检测从场景中的一个或多个目标和对象的表面反向散射的光。在某些实施例中,可以使用脉冲式红外光从而可以测量出射光脉冲和相应的入射光脉冲之间的时间差并将其用于确定从捕捉设备 20 到场景中的目标或物体上的特定位置的物理距离。另外,在其他示例实施例中,可将入射光波的相位与出射光波的相位进行比较来确定相移。然后可以使用相移来确定从捕捉设备到目标或对象上的特定位置的距离。

[0036] 根据另一示例实施例,可使用飞行时间分析,通过经由包括例如快门式光脉冲成

像的各种技术来分析反射光束随时间的强度变化以间接地确定从捕捉设备 20 到目标或物体上的特定位置的物理距离。

[0037] 在另一示例实施例中,捕捉设备 20 可使用结构化光来捕捉深度信息。在该分析中,图案化光(即,被显示为诸如网格图案或条纹图案等已知图案的光)可经由例如 IR 光组件 24 被投影到场景上。在撞击到场景中的一个或多个目标或物体的表面时,作为响应,图案可变形。图案的这种变形可由例如 3-D 相机 26 和 / 或 RGB 相机 28 来捕捉,然后可被分析来确定从捕捉设备到目标或对象上的特定位置的物理距离。

[0038] 根据另一实施例,捕捉设备 20 可包括两个或更多个物理上分开的相机,这些相机可从不同角度查看场景来获得可被解析以生成深度信息的视觉立体数据。

[0039] 捕捉设备 20 还可包括话筒 30。话筒 30 可包括可接收声音并将其转换成电信号的换能器或传感器。根据一个实施例,话筒 30 可以被用来减少目标识别、分析和跟踪系统 10 中的捕捉设备 20 和计算环境 12 之间的反馈。另外,话筒 30 可用于接收也可由用户提供的音频信号,以控制可由计算环境 12 执行的诸如游戏应用、非游戏应用等应用。

[0040] 在一示例实施例中,捕捉设备 20 还可以包括可与图像相机组件 22 进行可操作的通信的处理器 32。处理器 32 可包括可执行指令的标准处理器、专用处理器、微处理器等,这些指令可包括用于接收深度图像的指令、用于确定合适的目标是否可被包括在深度图像中的指令、用于将合适的目标转换成该目标的骨架表示或模型的指令、或任何其他合适的指令。

[0041] 捕捉设备 20 还可以包括存储器组件 34,该存储器组件 34 可以存储可以由处理器 32 执行的指令,由 3-D 相机或 RGB 相机捕捉到的图像或图像的帧,或任何其他合适的信息、图像等等。根据一个示例实施例,存储器组件 34 可包括随机存取存储器 (RAM)、只读存储器 (ROM)、高速缓存、闪存、硬盘或任何其他合适的存储组件。如图 2 所示,在一个实施例中,存储器组件 34 可以是与图像捕捉组件 22 和处理器 32 进行通信的单独的组件。根据另一实施例,存储器组件 34 可被集成到处理器 32 和 / 或图像捕捉组件 22 中。

[0042] 如图 2 所示,捕捉设备 20 可经由通信链路 36 与计算环境 12 进行通信。通信链路 36 可以是包括例如 USB 连接、火线连接、以太网电缆连接等的有线连接和 / 或诸如无线 802.11b、802.11g、802.11a 或 802.11n 连接等无线连接。根据一个实施例,计算环境 12 可以向捕捉设备 20 提供时钟,可以使用该时钟来通过通信链路 36 确定何时捕捉,例如,场景。

[0043] 另外,捕捉设备 20 可以通过通信链路 36 向计算环境 12 提供深度信息和由例如 3-D 相机 26 和 / 或 RGB 相机 28 捕捉到的图像,以及可以由捕捉设备 20 生成的骨架模型。计算环境 12 然后可使用该骨架模型、深度信息和捕捉的图像来例如识别用户姿势以及作为响应来控制诸如游戏或文字处理程序等应用。例如,如图 2 所示,计算环境 12 可包括姿势识别器引擎 190。姿势识别器引擎 190 可包括姿势过滤器集合,每一姿势过滤器包括关于骨架模型(在用户移动时)可执行的姿势的信息。可将由相机 26、28 和设备 20 捕捉的骨架模型形式的数据以及与其相关联的移动同姿势识别器引擎 190 中的姿势过滤器进行比较来标识用户(如骨架模型所表示的)何时执行了一个或多个姿势。那些姿势可与应用的各种控制相关联。因此,计算环境 12 可使用姿势识别器引擎 190 来解释骨架模型的移动并基于该移动来控制应用。

[0044] 图 3A 示出了可用于在目标识别、分析和跟踪系统中解释一个或多个姿势的计算

环境的示例实施例。上面参考图 1A-2 所描述的诸如计算环境 12 等的计算环境可以是诸如游戏控制台等的多媒体控制台 100。如图 3A 所示,多媒体控制台 100 具有含有一级高速缓存 102、二级高速缓存 104 和闪存 ROM(只读存储器)106 的中央处理单元(CPU)101。一级高速缓存 102 和二级高速缓存 104 临时存储数据并因此减少存储器访问周期数,由此改进处理速度和吞吐量。CPU101 可以设置成具有一个以上的内核,以及由此的附加的一级和二级高速缓存 102 和 104。闪存 ROM 106 可存储在当多媒体控制台 100 通电时的引导过程的初始阶段期间加载的可执行代码。

[0045] 图形处理单元(GPU)108 和视频编码器/视频编解码器(编码器/解码器)114 形成用于高速、高分辨率图形处理的视频处理流水线。数据经由总线从图形处理单元 108 输送到视频编码器/视频编解码器 114。视频处理流水线向 A/V(音频/视频)端口 140 输出数据,以便传输到电视机或其他显示器。存储器控制器 110 连接到 GPU 108 以方便处理器访问各种类型的存储器 112,诸如但不限于 RAM(随机存取存储器)。

[0046] 多媒体控制台 100 包括较佳地在模块 118 上实现的 I/O 控制器 120、系统管理控制器 122、音频处理单元 123、网络接口控制器 124、第一 USB 主控制器 126、第二 USB 控制器 128 和前面板 I/O 子部件 130。USB 控制器 126 和 128 用作外围控制器 142(1)-142(2)、无线适配器 148、和外置存储器设备 146(例如闪存、外置 CD/DVD ROM 驱动器、可移动介质等)的主机。网络接口 124 和/或无线适配器 148 提供对网络(例如,因特网、家庭网络等)的访问并且可以是包括以太网卡、调制解调器、蓝牙模块、电缆调制解调器等各种不同的有线或无线适配器组件中任何一种。

[0047] 提供系统存储器 143 来存储在引导过程期间加载的应用数据。提供了介质驱动器 144,其可以包括 DVD/CD 驱动器、硬盘驱动器,或其他可移动介质驱动器等。介质驱动器 144 可以是多媒体控制台 100 内部或外部的。应用数据可经由介质驱动器 144 访问,以由多媒体控制台 100 执行、回放等。介质驱动器 144 经由诸如串行 ATA 总线或其他高速连接(例如 IEEE 1394)等总线连接到 I/O 控制器 120。

[0048] 系统管理控制器 122 提供涉及确保多媒体控制台 100 的可用性的各种服务功能。音频处理单元 123 和音频编解码器 132 形成具有高保真度和立体声处理的对应的音频处理流水线。音频数据经由通信链路在音频处理单元 123 与音频编解码器 132 之间传输。音频处理流水线将数据输出到 A/V 端口 140 以供外置音频播放器或具有音频能力的设备再现。

[0049] 前面板 I/O 子部件 130 支持暴露在多媒体控制台 100 的外表面上的电源按钮 150 和弹出按钮 152 以及任何 LED(发光二极管)或其他指示器的功能。系统供电模块 136 向多媒体控制台 100 的组件供电。风扇 138 冷却多媒体控制台 100 内的电路。

[0050] CPU 101、GPU 108、存储器控制器 110、和多媒体控制台 100 内的各个其他组件经由一条或多条总线互连,包括串行和并行总线、存储器总线、外围总线、和使用各种总线架构中任一种的处理器或局部总线。作为示例,这些架构可以包括外围部件互连(PCI)总线、PCI-Express 总线等。

[0051] 当多媒体控制台 100 通电时,应用数据可从系统存储器 143 加载到存储器 112 和/或高速缓存 102、104 中并在 CPU 101 上执行。应用可呈现在导航到多媒体控制台 100 上可用的不同媒体类型时提供一致的用户体验的图形用户界面。在操作中,介质驱动器 144 中包含的应用和/或其他媒体可从介质驱动器 144 启动或播放,以向多媒体控制台 100 提供

附加功能。

[0052] 多媒体控制台 100 可通过将该系统简单地连接到电视机或其他显示器而作为独立系统来操作。在该独立模式中,多媒体控制台 100 允许一个或多个用户与该系统交互、看电影、或听音乐。然而,随着通过网络接口 124 或无线适配器 148 可用的宽带连接的集成,多媒体控制台 100 还可作为较大网络社区中的参与者来操作。

[0053] 当多媒体控制台 100 通电时,可以保留设定量的硬件资源以供多媒体控制台操作系统作系统使用。这些资源可包括存储器的保留量(诸如,16MB)、CPU 和 GPU 周期的保留量(诸如,5%)、网络带宽的保留量(诸如,8kbs),等等。因为这些资源是在系统引导时间保留的,所保留的资源对于应用视角而言是不存在的。

[0054] 具体地,存储器保留较佳地足够大,以包含启动内核、并发系统应用和驱动程序。CPU 保留较佳地为恒定,使得若所保留的 CPU 用量不被系统应用使用,则空闲线程将消耗任何未使用的周期。

[0055] 对于 GPU 保留,通过使用 GPU 中断来显示由系统应用生成的轻量消息(例如,弹出窗口),以调度代码来将弹出窗口呈现为覆盖图。覆盖图所需的存储器量取决于覆盖区域大小,并且覆盖图较佳地与屏幕分辨率成比例缩放。在并发系统应用使用完整用户界面的情况下,优选使用独立于应用分辨率的分辨率。定标器可用于设置该分辨率,从而无需改变频率并引起 TV 重新同步。

[0056] 在多媒体控制台 100 引导且系统资源被保留之后,就执行并发系统应用来提供系统功能。系统功能被封装在上述所保留的系统资源中执行的一组系统应用中。操作系统内核标识是系统应用线程而非游戏应用线程的线程。系统应用优选地被调度为在预定时间并以预定时间间隔在 CPU 101 上运行,以便为应用提供一致的系统资源视图。进行调度是为了把由在控制台上运行的游戏应用所引起的高速缓存分裂最小化。

[0057] 当并发系统应用需要音频时,则由于时间敏感性而异步调度音频处理给游戏应用。多媒体控制台应用管理器(如下所述)在系统应用活动时控制游戏应用的音频水平(例如,静音、衰减)。

[0058] 输入设备(例如,控制器 142(1) 和 142(2))由游戏应用和系统应用共享。输入设备不是所保留的资源,但却在系统应用和游戏应用之间切换以使其各自具有设备的焦点。应用管理器较佳地控制输入流的切换,而无需知晓游戏应用的知识,并且驱动程序维护有关焦点切换的状态信息。相机 26、28 和捕捉设备 20 可以为控制台 100 定义额外的输入设备。

[0059] 图 3B 示出了可用于在目标识别、分析和跟踪系统中解释一个或多个姿势的计算环境的计算环境 220 的另一示例实施例,该计算环境可以是图 1A-2 所示的计算环境 12。计算系统环境 220 只是合适的计算环境的一个示例,并且不旨在对所公开的主题的使用范围或功能提出任何限制。也不应该将计算环境 220 解释为对示例性操作环境 220 中示出的任一组件或其组合有任何依赖性 or 要求。在某些实施例中,所描绘的各种计算元素可包括被配置成实例化本公开的各具体方面的电路。例如,本公开中使用的术语电路可包括被配置成通过固件或开关来执行功能的专用硬件组件。在其他示例实施方式中,术语电路可包括由实施可用于执行功能的逻辑的软件指令配置的通用处理单元、存储器等。在其中电路包括硬件和软件的组合的示例实施例中,实施者可以编写体现逻辑的源代码,且源代码可以

被编译为可以由通用处理单元处理的机器可读代码。因为本领域技术人员可以明白现有技术已经进化到硬件、软件或硬件 / 软件组合之间几乎没有差别的地步,因而选择硬件或是软件来实现具体功能是留给实现者的设计选择。更具体地,本领域技术人员可以明白软件进程可被变换成等价的硬件结构,而硬件结构本身可被变换成等价的软件进程。因此,对于硬件实现还是软件实现的选择是设计选择并留给实现者。

[0060] 在图 3B 中,计算环境 220 包括计算机 241,其通常包括各种计算机可读介质。计算机可读介质可以是能由计算机 241 访问的任何可用介质,而且包含易失性和非易失性介质、可移动和不可移动介质。系统存储器 222 包括易失性和 / 或非易失性存储器形式的计算机存储介质,如只读存储器 (ROM) 223 和随机存取存储器 (RAM) 260。基本输入 / 输出系统 224 (BIOS) 包括如在启动时帮助在计算机 241 内的元件之间传输信息的基本例程,它通常储存在 ROM 223 中。RAM 260 通常包含处理单元 259 可以立即访问和 / 或目前正在操作的数据和 / 或程序模块。作为示例而非局限,图 3B 示出了操作系统 225、应用程序 226、其它程序模块 227 和程序数据 228。

[0061] 计算机 241 还可以包括其他可移动 / 不可移动、易失性 / 非易失性计算机存储介质。仅作为示例,图 3B 示出了对不可移动、非易失性磁介质进行读写的硬盘驱动器 238,对可移动、非易失性磁盘 254 进行读写的磁盘驱动器 239,以及对诸如 CD ROM 或其它光学介质等可移动、非易失性光盘 253 进行读写的光盘驱动器 240。可以在示例性操作环境中使用的其他可移动 / 不可移动、易失性 / 非易失性计算机存储介质包括但不限于,磁带盒、闪存卡、数字多功能盘、数字录像带、固态 RAM、固态 ROM 等等。硬盘驱动器 238 通常由不可移动存储器接口,诸如接口 234 连接至系统总线 221,磁盘驱动器 239 和光盘驱动器 240 通常由可移动存储器接口,诸如接口 235 连接至系统总线 221。

[0062] 以上讨论并在图 3B 中示出的驱动器及其相关联的计算机存储介质为计算机 241 提供了对计算机可读指令、数据结构、程序模块和其他数据的存储。在图 3B 中,例如,硬盘驱动器 238 被示为存储操作系统 258、应用程序 257、其他程序模块 256 和程序数据 255。注意,这些组件可以与操作系统 225、应用程序 226、其他程序模块 227 和程序数据 228 相同,也可以与它们不同。在此给操作系统 258、应用程序 257、其他程序模块 256、以及程序数据 255 提供了不同的编号,以说明至少它们是不同的副本。用户可以通过输入设备,诸如键盘 251 和定点设备 252 (通常被称为鼠标、跟踪球或触摸垫),向计算机 241 输入命令和信息。其他输入设备 (未示出) 可以包括话筒、操纵杆、游戏手柄、圆盘式卫星天线、扫描仪等等。这些和其他输入设备通常由耦合至系统总线的用户输入接口 236 连接至处理单元 259,但也可以由其他接口和总线结构,诸如并行端口、游戏端口或通用串行总线 (USB),来连接。相机 26、28 和捕捉设备 20 可以为控制台 100 定义额外的输入设备。监视器 242 或其他类型的显示设备也通过接口,诸如视频接口 232,连接至系统总线 221。除监视器以外,计算机也可以包括其他外围输出设备,诸如扬声器 244 和打印机 243,它们可以通过输出外围接口 233 连接。

[0063] 计算机 241 可使用至一个或多个远程计算机,诸如远程计算机 246 的逻辑连接在网络化环境中操作。远程计算机 246 可以是个人计算机、服务器、路由器、网络 PC、对等设备或其他常见的网络节点,且通常包括许多或所有以上相对于计算机 241 描述的元件,尽管在图 3B 中仅示出了存储器存储设备 247。图 3B 中所示的逻辑连接包括局域网 (LAN) 245

和广域网 (WAN) 249, 但也可以包括其他网络。这样的联网环境在办公室、企业范围计算机网络、内联网和因特网中是常见的。

[0064] 当在 LAN 联网环境中使用时, 计算机 241 通过网络接口或适配器 237 连接至 LAN 245。当在 WAN 联网环境中使用时, 计算机 241 通常包括调制解调器 250 或用于通过诸如因特网等 WAN 249 建立通信的其他手段。调制解调器 250 可以是内置或外置的, 它可以经由用户输入接口 236 或其他适当的机制连接至系统总线 221。在网络化环境中, 相对于计算机 241 所描述的模块或其部分可被存储在远程存储器存储设备中。作为示例而非限制, 图 3B 示出了远程应用程序 248 驻留在存储器设备 247 上。可以理解, 所示的网络连接是示例性的, 且可以使用在计算机之间建立通信链路的其他手段。

[0065] 图 4A 描绘了可从捕捉设备 20 生成的用户的示例骨架映射。在该实施方式中, 标识出各个关节和骨骼: 每一手 302、每一前臂 304、每一肘 306、每一二头肌 308、每一肩 310、每一髋 312、每一大腿 314、每一膝 316、每一小腿 318、每一足 320、头 322、躯干 324、脊椎的顶部 326 和底部 328, 以及腰 330。在跟踪更多点的情况下, 可标识出附加的特征, 比如手指或脚趾的骨骼和关节, 或脸部的各个特征, 如鼻和眼。

[0066] 用户可通过移动他的身体创建姿势。姿势包括用户的运动或姿态, 其可被捕捉为图像数据并解析其意义。姿势可以是动态的, 包括运动, 如模仿扔球。姿势可以是静态姿态, 如在一个人的躯干 324 前面交叉握住他的前臂 304。姿势也可结合道具, 如通过挥动仿制的剑。姿势可包括多于一个的身体部位, 如拍双手 302, 或是较微小的运动, 如撅起一个人的嘴唇。

[0067] 姿势可用于一般计算上下文中的输入。例如, 手 302 或其他身体部位的各种运动可对应于常见的系统级任务, 如在分层列表中向上或向下导航、打开文件、关闭文件和保存文件。姿势也可在视频游戏专用上下文中取决于游戏来使用。例如, 对于驾驶游戏, 手 302 和脚 320 的各种运动可对应于在一方向上操控车辆、换挡、加速和刹车。

[0068] 用户可通过自己在原地行走或奔跑来生成对应于行走或奔跑的姿势。用户可另选地提起并放下每一腿 312-320 来在不移动的情况下模拟行走。系统可通过分析每一髋 312 和每一大腿 314 来解析该姿势。当一个髋部 - 大腿角 (如相对于垂直线测量的, 其中站立的腿具有 0° 的髋部 - 大腿角, 而向前水平伸展的腿具有 90° 的髋部 - 大腿角) 超过相对于另一大腿的特定阈值时, 可识别一步。行走或奔跑可在交替腿的某一数量的连续步之后被识别。两个最近的步之间的时间可被认为是一周期。在不满足阈值角度达某一数量的周期之后, 系统可确定行走或奔跑姿势已停止。

[0069] 给定“行走或奔跑”姿势, 应用可为与该姿势相关联的应用确定的参数设定值。这些参数可包括上述阈值角度、发起行走或奔跑姿势所需的步数、结束姿势的没有发生步的周期数、以及确定姿势是行走还是奔跑的阈值周期。快周期可对应于奔跑, 因为用户将快速地移动他的腿, 而较慢的周期可对应于行走。

[0070] 姿势最初可以与一组默认参数相关联, 应用可用其自己的参数来覆盖该组默认参数。在这一场景中, 不迫使应用提供参数, 但是应用可改为使用一组允许在没有应用定义的参数的情况下识别姿势的默认参数。

[0071] 有各种可以与姿势相关联的输出。可以有姿势是否正在发生的基线“是或否”。还可以有置信度水平, 其对应于用户跟踪的移动对应于姿势的可能性。这可以是范围

为 0 和 1 之间（包括端点）的浮点数的线性标度。在接收该姿势信息的应用不能接受假肯定作为输入的应用中，它可仅使用具有高置信度水平，如至少 0.95 的已识别的姿势。在应用必须识别姿势的每一实例的情况下，即使以假肯定为代价，它可使用至少具有低得多的置信度水平的姿势，如仅仅大于 0.2 的那些姿势。姿势可具有关于两个最近步之间的时间的输出，并且在仅注册了第一步的情况下，这可被设为保留值，如 -1（因为任何两步之间的时间必须为正）。姿势也可具有关于在最近一步期间达到的最高大腿角的输出。

[0072] 另一示例性姿势是“脚跟提起跳”。在该姿势中，用户可通过将他的脚跟提离地面，但保持他的脚趾着地来创建该姿势。另选地，用户可跳向空中，其中他的脚 320 完全离开地面。该系统可通过分析肩 310、髌 312 和膝 316 的角度关系来解析该姿势的骨架，以查看它们是否在等于直立的对齐位置。然后，可监视这些点和较高 326 和较低 328 脊椎点来发现任何向上加速。足够的加速组合可触发跳跃姿势。

[0073] 给定该“脚跟提起跳”姿势，应用可为与该姿势相关联的应用确定的参数设定值。参数可包括上述加速阈值，其确定用户的肩 310、髌 312 和膝 316 的某种组合必须向上移动多快来触发该姿势；以及肩 310、髌 312 和膝 316 之间仍可触发跳跃的最大对齐角。

[0074] 输出可包括置信度水平，以及用户在跳跃时的身体角度。

[0075] 基于将接收姿势的应用的细节来为姿势设定参数对于准确地标识姿势而言是重要的。正确地标识姿势以及用户的意图极大地有助于创建积极的用户体验。在姿势识别器系统太敏感、甚至手 302 的微小向前运动也被解释为投掷的情况下，用户可能变得沮丧，因为在他没有作出姿势的意图的情况下却识别出姿势，因此他缺乏对系统的控制。在姿势识别器系统不够敏感的情况下，系统可能未识别出用户作出投掷姿势的刻意尝试，从而类似地使用户感到沮丧。在敏感度范围的任一端，用户都会变得沮丧，因为他不能正确地向系统提供输入。

[0076] 对于姿势的另一参数可以是移动的距离。在用户的姿势控制虚拟环境中的化身的动作的情况下，该化身可以是手臂离球的长度。如果用户希望与该球交互并抓住它，则这可要求用户伸展他的手臂 302-310 到全长同时作出抓握姿势。在这一情形中，用户仅部分地伸展他的手臂 302-310 的类似的抓握姿势可能无法达到与球交互的结果。

[0077] 姿势或其一部分可具有它必须在其中发生的空间体来作为参数。在姿势包括身体移动的情况下，该空间体通常可相对于身体来表达。例如，对于惯用右手的用户的美式足球投掷姿势可仅在不低于右肩 310a、且与投掷手臂 302a-310a 在头 322 的相同侧的空间体中识别。可能不必要定义空间体的所有边界，如对于该投掷姿势，其中从身体向外的边界留着不被定义，并且该空间体无限地向外延伸，或者延伸到正被监视的场景的边缘。

[0078] 图 4B 提供了图 2 的姿势识别器引擎 190 的一个示例性实施例的进一步细节。如图所示，姿势识别器引擎 190 可包括至少一个过滤器 418 来确定一个或多个姿势。过滤器 418 包括定义姿势 426（以下称为“姿势”）的信息，并还可包括用于该姿势的参数 428 或元数据。例如，包括一只手从身体背后越过身体前方的运动的投掷可被实现为包括表示用户的一只手从身体背后越过身体前方的运动的信息的姿势 426，该移动将由深度相机来捕捉。然后可为该姿势 426 设定参数 428。在姿势 426 是投掷的情况下，参数 428 可以是该手必须达到的阈值速度、该手必须行进的距离（绝对的，或相对于用户的整体大小）、以及识别器引擎对发生姿势的置信度评级。用于姿势 426 的这些参数 428 可以在各应用之间、在单个

应用的各上下文之间、或在一个应用的一个上下文内随着时间变化。

[0079] 过滤器可以是模块化的或是可互换的。在一个实施例中,过滤器具有多个输入,这些输入中的每一个具有一类型;以及多个输出,这些输出中的每一个具有一类型。在这一情形中,第一过滤器可用具有与第一过滤器相同数量和类型的输入和输出的第二过滤器来替换而不更改识别器引擎体系结构的其他方面。例如,可以有用于驾驶的第一过滤器,该第一过滤器取骨架数据作为输入并输出与该过滤器相关联的姿势正在发生的置信度以及转向角。在希望用第二驾驶过滤器来替换该第一驾驶过滤器的情况下——这可能是因为第二驾驶过滤器更高效且需要更少的处理资源——则可以通过简单地用第二过滤器替换第一过滤器来这样做,只要第二过滤器具有相同的输入和输出——骨架数据类型的一个输入、以及置信度类型和角度类型的两个输出。

[0080] 过滤器不需要具有参数。例如,返回用户的高度的“用户高度”过滤器可能不允许任何可调节的参数。替换的“用户高度”过滤器可具有可调节参数,如在确定用户的高度时是否考虑用户的鞋、发型、头饰以及体态。

[0081] 对过滤器的输入可包括诸如关于用户的关节位置的关节数据,像在关节处相交的骨所形成的角度、来自场景的 RGB 色彩数据、以及用户的某一方面的变化速率等内容。来自过滤器的输出可包括诸如正作出给定姿势的置信度、作出姿势运动的速度、以及作出姿势运动的时间等内容。

[0082] 上下文可以是文化上下文,并且可以是环境上下文。文化上下文指的是使用系统的用户的文化。不同的文化可使用相似的姿势来赋予显著不同的含义。例如,希望叫另一个用户“看”或“使用他的眼睛”的美国用户可将他的食指放在他头上靠近他的眼睛的远端处。然而,对意大利用户而言,该姿势可被解释为对黑手党的引用。

[0083] 类似地,在单个应用的不同环境之中可能有不同的上下文。取涉及操作摩托车的第一人称射击游戏。当用户在走路时,将手指朝向地面握拳并向前且从身体向外伸出拳头可表示重拳姿势。当用户在驾驶上下文中时,相同的运动可表示“换挡”姿势。可能还有一个或多个菜单环境,其中用户可保存他的游戏、在他的人物装备之间选择或执行类似的不包括直接玩游戏的动作。在该环境中,该游戏姿势可具有第三个含义,如选择某样东西或前进到另一屏幕。

[0084] 姿势识别器引擎 190 可具有向姿势过滤器 418 提供功能的基础识别器引擎 416。在一实施例中,识别器引擎 416 实现的功能包括跟踪所识别的姿势和其他输入的随时间输入 (input-over-time) 存档、隐马尔可夫模型实现 (其中模型化系统被假定为具有未知参数的马尔可夫过程—其中当前状态封装了确定将来状态所需的任何过去状态信息,因此不必为此目的而维护任何其它过去状态信息的过程,并且隐藏参数从可观察数据来确定)、以及求解姿势识别的特定实例所需的其它功能。

[0085] 过滤器 418 在基础识别器引擎 416 之上加载并实现,并且可利用引擎 416 提供给所有过滤器 418 的服务。在一实施例中,基础识别器引擎 416 处理所接收到的数据来确定它是否满足任何过滤器 418 的要求。由于这些诸如解析输入等所提供的服务是由基础识别器引擎 416 一次性提供而非由每一过滤器 418 提供的,因此这一服务在一段时间内只需被处理一次而不是对该时间段对每一过滤器 418 处理一次,由此减少了确定姿势所需的处理。

[0086] 应用可使用姿势识别器引擎 190 所提供的过滤器 418,或者它可提供其自己的过

滤器 418, 该过滤器被插入到基础识别器引擎 416 中。在一实施例中, 所有过滤器 418 具有启用该插入特性的通用接口。此外, 所有过滤器 418 可利用参数 428, 因此可使用如下所述的单个姿势工具来诊断并调节整个过滤器系统 418。

[0087] 这些参数 428 可由姿势工具 420 为应用或应用的上下文来调节。在一实施例中, 姿势工具 420 包括多个滑块 422 以及身体 424 的图表示, 每一滑块 422 对应于一参数 428。当用相应的滑块 422 来调整参数 428 时, 身体 424 可展示将用这些参数 428 被识别为姿势的动作以及使用这些参数 428 将不被识别为姿势的动作, 如所标识的。姿势的参数 428 的这一可视化提供了调试并细化调节姿势的有效手段。

[0088] 图 5 描绘了从堆叠的姿势或过滤器 418 创建的更复杂的姿势或过滤器 418。姿势可以彼此堆叠, 堆叠的过滤器于是可被认为是包括多个其他过滤器的过滤器。即, 用户一次可表达多于一个的姿势。例如, 并非在作出投掷姿势时不允许除了投掷之外的任何输入, 或要求用户除了该姿势的分量之外保持不动 (例如, 在作出仅涉及一条手臂的投掷姿势时站着不动)。在姿势堆叠时, 用户可同时作出跳跃姿势和投掷姿势, 并且这两个姿势都将被姿势引擎识别出。

[0089] 图 5A 描绘了根据堆叠范例的简单姿势过滤器 418。I 过滤器 (IFilter) 502 是可在每一姿势过滤器中使用的基本过滤器 418。I 过滤器 (IFilter) 502 取用户位置数据 504, 并输出姿势已发生的置信度水平 506。它还将该位置数据 504 馈入方向盘 (SteeringWheel) 过滤器 508, 该过滤器 508 取该位置数据作为输入, 并输出用户正在转向的角度 (例如, 向用户当前方向右侧 40 度) 510。

[0090] 图 5B 描绘了将过滤器 418 堆叠到图 5A 的姿势过滤器上的更复杂的姿势。除了 I 过滤器 502 和方向盘 508 之外, 还存在 I 跟踪 (ITracking) 过滤器 512, 该过滤器从 I 过滤器 502 接收位置数据 504, 并输出用户通过姿势 514 所作出的进展量。I 跟踪 512 还将位置数据 504 馈送到润滑油减轻 (GreaseLightning) 514 和 E 刹车 516, 这些是关于可在操作车辆时如使用紧急刹车作出的其他姿势的过滤器 418。

[0091] 存在用于堆叠姿势的其他实施例。在一实施例中, 位置数据 504 被直接, 而不是通过 I 过滤器 502, 传递给堆叠的姿势的全部过滤器, 且识别器引擎的组件确定每个过滤器如何互操作。例如, 利用跳跃过滤器和投掷过滤器, 当没有其他用户动作正在发生时, 每个可被独立地识别, 但是该组件将基于来自每个过滤器的各自的输出识别出跳跃和可能是投掷的某一用户输入正在同时发生。该组件然后可较不严厉地解释投掷过滤器 (例如, 通过扩展可接受的值的范围来满足某一参数), 并基于此可将“可能为投掷的某一输入”识别为投掷。

[0092] 图 6 描绘了用户 602 在美式足球视频游戏中可以作出的用以发出“安全接球”(fair catch) 信号的示例姿势。这些图描绘了各个时间点的用户, 其中图 6A 是第一时间点, 而图 6E 是最后一个时间点。这些图中的每一个可对应于由深度相机 402 捕捉的图像数据的一个快照或帧, 但不必是连续的图像数据帧, 因为深度相机 402 能够比用户可覆盖该距离更快地捕捉帧。例如, 该姿势可在 3 秒的时间段内发生, 并且在深度相机以每秒 40 帧来捕捉数据的情况下, 在用户 602 作出该安全接球姿势时它将捕捉 60 个图像数据帧。

[0093] 在图 6A 中, 用户 602 以他的手臂 604 向下放于两侧开始。他然后将手臂举起超过肩膀, 如图 6B 所描绘的, 然后进一步向上到大约他的头部的水平, 如图 6C 所描绘的。从那

里,他将他的手臂 604 降下到肩膀水平,如图 6D 所描绘的,然后再一次将手臂举起到大约他的头部的水平,如图 6E 所描绘的。在系统捕捉用户 602 的这些位置而没有可发出该姿势被取消或正在作出另一姿势的信号的任何居间位置的情况下,可将安全接球姿势识别为已由用户 602 作出。

[0094] 图 7 描绘了在解析每一图像数据帧来产生用户的骨架映射时图 5 的示例“安全接球”姿势。该系统在从用户的深度图像产生了骨架映射之后,现在可确定用户的身体如何随时间而移动,并且可从中解析姿势。

[0095] 在图 7A 中,用户的肩膀 310 在他的肘 306 上方,而肘 306 又在他的手 302 上方。肩 310、肘 306 和手 302 然后在图 7B 中处于一致水平。系统然后在图 7C 中检测到手 302 在肘上方,而肘在肩 310 上方。在图 7D 中,用户返回到图 7B 的位置,其中肩 310、肘 306 和手 302 处于一致水平。在该姿势的最终位置,图 7E 中示出,用户返回到图 7C 的位置,其中手 302 在肘上方,而肘在肩 310 上方。

[0096] 尽管深度相机 402 捕捉了一系列静止图像,使得在任何一个图像中用户看似为静止的,但用户在执行该姿势的过程中在移动(而非如上所述的静止姿势)。该系统能够取每一静止图像中的该一系列姿态,并从中确定用户正在作出的移动姿势。

[0097] 在执行姿势时,用户可能无法创建如他的右肩 310a、右肘 306a 和右手 302a 所形成的角度,例如在 140° 和 145° 之间。因此,使用用于安全接球姿势 428 的过滤器 418 的应用可调谐相关联的参数 426 来最佳地服务于应用的细节。例如,图 7C 和 7E 中的位置可在用户将其手 302 放在其肩 310 上方的任何时刻被识别出,而不关注肘 306 的位置。更严格的一组参数可能要求手 302 在头 310 上方,并且肘 306 既在肩 310 上方又在头 322 和手 302 之间。另外,用于安全接球姿势 428 的参数 426 可要求用户在诸如 1.5 秒等一段指定时间内从图 7A 的位置移动到图 7E 的位置,并且如果用户花费了多于 1.5 秒来移动通过这些位置,则它将不被识别为安全接球 418。

[0098] 图 8 描绘了来自姿势过滤器库 802 的通用姿势过滤器 806 如何被分组成用于某一任务的互补姿势过滤器的风格包 804。姿势过滤器库 802 聚集由系统提供的所有姿势过滤器 806。在一实施例中,应用可提供附加的姿势过滤器供该应用使用。通用姿势过滤器包括诸如“抓手过背摔”806a 和“蹲下”806b 之类的事物。这些姿势过滤器然后可在风格包 804 中被分组在一起。

[0099] 风格包 804 可包括某一风格中通常使用的那些姿势。例如,第一人称射击(FPS)风格包 804a 可具有用于用武器射击 812c、投掷炮弹 812d、挥拳猛击 812e、开门 812f、蹲下 812g、跳跃 812h、奔跑 812i 以及转身 812j 的姿势过滤器。该 FPS 风格包 804a 可被认为是提供通用 FPS 风格包 808a- 姿势过滤器参数被调谐或设置使得它们将与大量 FPS 应用一起可接受地工作的风格包。

[0100] 风格包不限于为已知风格或应用工作的互补姿势过滤器的组。风格包可包括包含某一应用或风格使用的那些过滤器的子集的姿势过滤器,或互补的过滤器,但是用于它们的合适的风格还未被标识。

[0101] 应用然后可调谐那些通用风格包来满足应用的细节。应用可通过为风格包中的过滤器的参数设置值来调谐通用风格包。例如,游戏 A 810 的创建者可决定在要求演示移动来注册开门姿势过滤器 812f 时他们的游戏最佳地运作,因为否则它就太类似于挥拳猛击

姿势过滤器 812e 了。然而,游戏 B 的创建者可决定这不是问题,且仅要求更适度的移动来注册开门姿势过滤器 812fB。

[0102] 在风格包包括机器可读指令的实施例中,风格包可被提供为采用源代码形式的那些指令或采用反映那些指令的某些编译量的形式。

[0103] 图 9 示出了用于在应用为一个过滤器的一个参数提供值时调谐过滤器包中的互补姿势过滤器的示例操作过程。

[0104] 操作 902 描绘了提供包括多个过滤器的包,每个过滤器包括关于姿势的信息和至少一个参数,每个过滤器与包中的至少一个其他过滤器互补。包可表示用于某一风格的姿势过滤器。例如,用于视频游戏的风格包可包括诸如第一人称射击、动作、驾驶和体育等风格。

[0105] 如这里所使用的且在至少一个实施例中,“提供包”可涉及允许对对应于包中的过滤器的编程语言库文件的访问或允许对应用的应用编程接口 (API) 的访问。应用的开发者可加载库文件并然后在适当时作出方法调用。例如,对于体育包,可存在相应的体育包库文件。

[0106] 当包含在应用中时,应用然后可根据给定的 API 作出使用体育包的调用。这些 API 调用可包括返回过滤器的参数的值,设置过滤器的参数的值,以及将过滤器的标识与触发应用的某一部分相关,例如使得受用户控制的网球玩家在用户作出适当的网球拍挥动姿势时挥动网球拍。

[0107] 如上所述,姿势可包括各种各样的事物。例如,它可以是蹲下、跳跃、倚靠、抓手过背摔、投掷、挥臂、躲避、踢和阻挡中的任一个。同样,姿势可对应于用户界面的导航。例如,用户可手指向上指并且手指面向 3D 相机地保持他的手。他然后可以将手指朝向手掌收拢来形成拳头,并且这可以是指示基于窗口的用户界面计算环境中的焦点窗口应被关闭的姿势。

[0108] 由于姿势可用于指示从应用中的化身应挥出重拳到应用中的窗口应被关闭的任何内容,因此从视频游戏到文本编辑器的各种各样的应用可利用姿势。

[0109] 互补姿势过滤器—或者如同在通常被一起使用的那些姿势过滤中那样互补,或者如一个姿势过滤器中的参数中的变化会改变另一个姿势过滤器的参数那样互补—可被分组在一起成为风格包,风格包可能由该应用以该风格使用。这些包对于应用来说可能是可用的或可被标识的,应用可选择至少一个。应用可调谐或修改所选包中的过滤器的参数来最佳地适合于应用的独特方面。当调谐该参数时,也可调谐该过滤器或第二过滤器的第二互补参数(在相互依赖的意义上)使得这些参数保持互补。

[0110] 应用确定的参数可基于应用所在的上下文来变化。为此,应用可将多个值赋予过滤器的应用确定的参数,每一值对应于一不同的上下文。如上所讨论的,该上下文可以是文化上下文或环境上下文。

[0111] 操作 904 描绘了接收将值赋予第一过滤器的参数的指示。应用确定的参数可包括过滤器的各种各样特性中的任一个,如身体部位、空间体、速度、移动方向、角度、以及移动发生的地方。

[0112] 在一实施例中,应用确定的参数的值由应用的最终用户通过作出姿势来确定。例如,应用可允许用户训练它,使得用户能够指定他认为姿势应包括什么运动。这对于允许没

有对其运动技能的良好控制的用户能够将他作出的运动与对应的姿势相链接是有益的。如果这不可用,则用户可能变得沮丧,因为他无法使其身体以应用所需的方式移动来产生姿势。

[0113] 在存在互补过滤器—具有相互关联的参数的多个过滤器—的实施例,从应用接收第一过滤器的应用确定的参数的值可包括用该值设置该第一过滤器的应用确定的参数,以及基于该第一过滤器的该参数的值设置第二互补过滤器的互补应用确定的参数。例如,可以决定以特定方式投掷美式足球的用户可能也以特定方式投掷棒球。因此,在确定一个过滤器的某一应用确定的参数,如对于美式足球投掷姿势的过滤器上的速度参数,应以某一方式被设置时,诸如棒球投掷姿势上的速度参数之类的其他互补应用确定的参数可基于第一应用确定的参数如何被设置来设置。

[0114] 这不需要是给定的应用确定的参数的相同值,或甚至跨各过滤器的相同类型的应用确定的参数。例如,情况可以是当美式足球投掷必须以前向手臂速度 X m/s (米/秒) 来作出时,美式足球接球必须以手距离躯干至少 Y m (米) 距离来作出。

[0115] 该值可以是阈值,例如手臂速度大于 X 。它可以是绝对值,如手臂速度等于 X 。可以有容错,如手臂速度等于 X 的 Y 内。它还可以包括范围,如手臂速度大于或等于 X 但小于 Z 。

[0116] 操作 906 描绘了将值赋予第一过滤器的参数。在参数和它们的值之间的关联性被存储在数据库中的情况下,这可包括将该值与和参数的关联性一起存储在数据库中。

[0117] 操作 908 描绘了将第二值赋予第二过滤器的第二参数,该第二值是使用赋予第一过滤器的参数的值确定的。如上所讨论的,第二值可以各种方式与第一值有关。在两个参数涉及基本相似的某物时,例如阈值跳跃高度,第二值可等于第一值。第二值和第一值可具有各种其他关系,例如比例关系、反比例关系、线性关系、指数关系、以及取这些值作为输入的函数。

[0118] 在过滤器可彼此继承特性的实施例,例如在面向对象的实现中,第二过滤器可构成第一过滤器的子,而第一过滤器类似地是第二过滤器的父。取“拍手”过滤器作为示例。该过滤器可充当诸如“举起单手击掌”(high five)、“举起双手击掌”(high ten)、以及“低手单手击掌”(low five) 之类的拍手上的变型的父。在“拍手”具有“手移动距离阈值”参数的情况下,当设置了该参数的值时,对于所有子过滤器的该“手移动距离阈值”参数可被设为该相同值。

[0119] 类似地,两个参数的互补特性可归因于堆叠一个过滤器以合并到另一个过滤器中。一个过滤器可以是转向过滤器,并与诸如换挡、加速、减速之类的其他过滤器堆叠,以创建驾驶过滤器。随着转向过滤器的“最小转向角阈值”参数被修改,驾驶过滤器的相应的“最小转向角阈值”参数也可被修改。

[0120] 操作 910 描绘了可任选的操作,接收包括个人的图像的数据,以及当解析过滤器以确定图像数据是否匹配某一过滤器(从而指示出某一姿势)时,在对不存在包中的过滤器解析数据之前,对所选包中的每个过滤器解析数据。在应用选择一过滤器包来使用的情况下,例如通过包括该过滤器包的库文件,它可能这么做,因为那些过滤器要被应用的用户经常使用。进一步,过滤器包中的过滤器可一个接一个地被使用,例如第一人射击包中的跑、跳、用枪扫射、蹲下并卸下武器过滤器。为此,在过滤器包已被标识为由应用使用时,

诸如上述基础过滤器引擎之类的系统处理过滤器很可能通过首先对构成所选过滤器包的那些过滤器处理数据来减少处理对应于用户输入的图像数据所需的处理资源。

[0121] 操作 912 描绘了接收将第三值赋予第二参数的指示以及将第三值赋予第二参数的任选的操作。两个参数之间的关系不需要是双向的。可能的情况是该值中的变化影响第二值,但第二值中的变化不影响第一值。

[0122] 操作 914 描绘了将第四值赋予该参数的任选的操作,该第四值是使用第三值确定的。两个参数之间的关系也可能是双向的。在该实施例中,第二参数的值中的变化导致改变第一参数的值,如使用第二参数的新的值确定的。

[0123] 操作 916 描绘了任选的操作,接收包括个人的图像的数据;确定数据是否匹配包的第一过滤器;以及在对不能被用户与该过滤器同时指示的过滤器解析数据之前,对该包中可由用户与该第一过滤器同时指示的每个其他过滤器解析数据。在一实施例中,过滤器包中的过滤器可同时被使用,例如第一人称射击包中的同时跑过滤器和卸下武器过滤器。为此,在过滤器包已被标识为由应用使用时,诸如上面讨论的基础识别器引擎之类的系统处理过滤器很可能通过首先对构成该过滤器包的那些过滤器处理数据来减少处理对应于用户输入的图像数据所需的处理资源。

[0124] 结论

[0125] 尽管已经结合各附图所示的较佳方面描述了本发明,但要理解,可使用其他相似方面或者可对所述方面进行修改或添加来执行本发明的相同功能而不脱离本发明。因此,本发明不应该仅限于任何单个方面,而是应该在根据所附权利要求书的广度和范围内解释。例如,本文描述的各种过程可用硬件或软件、或两者的组合来实现。因此,所公开的各实施例的方法和装置或其某些方面或部分可采用包含在诸如软盘、CD-ROM、硬盘驱动器或任何其他机器可读存储介质等有形介质中的程序代码(即,指令)的形式。当程序代码被加载到诸如计算机等机器并由其执行时,该机器变为被配置成实施所公开的各实施例的装置。除了此处明确阐述的具体实现之外,考虑此处所公开的说明书,其他方面和实现将对本领域的技术人员是显而易见的。说明书和所示实现旨在仅被认为是示例。

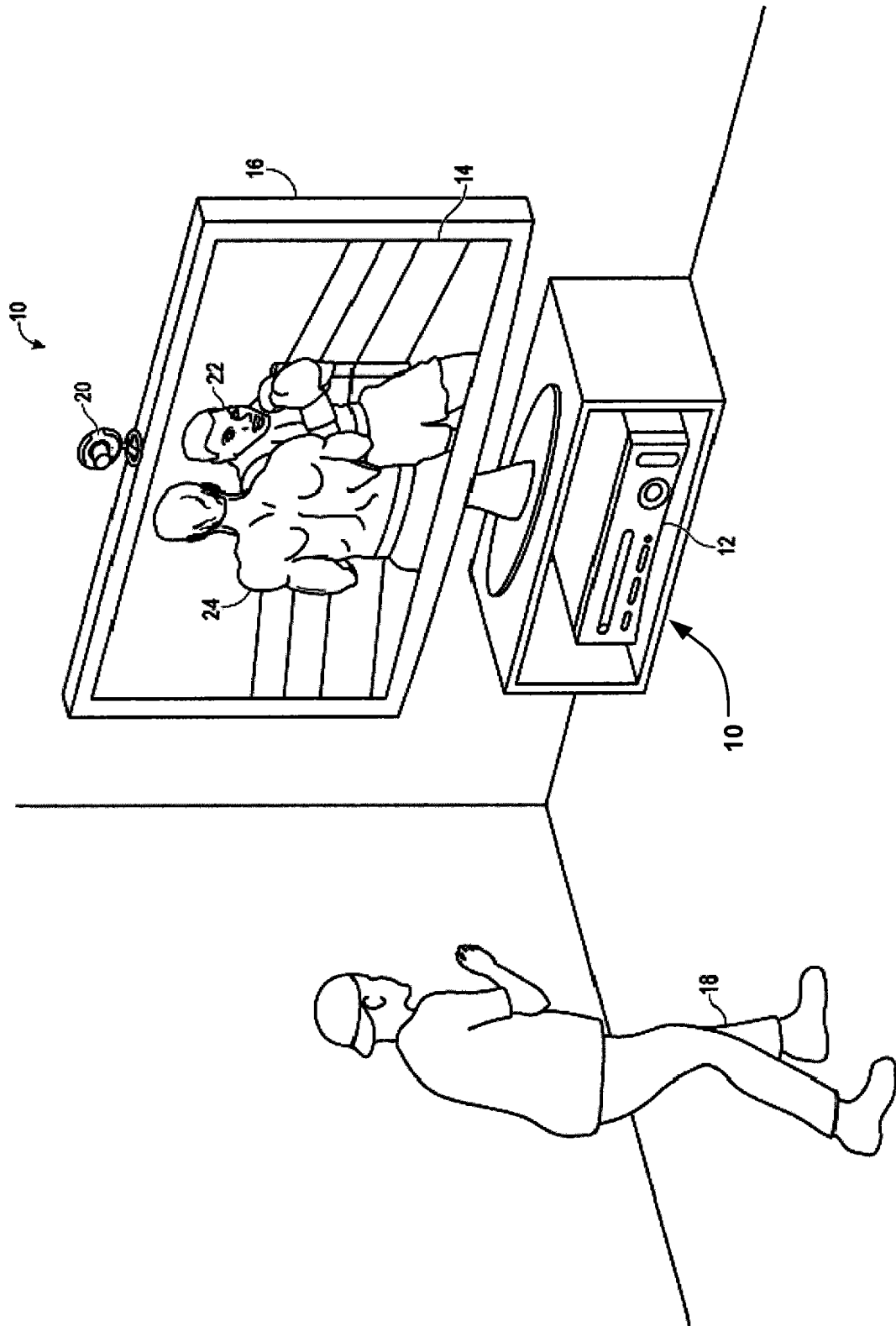


图 1A

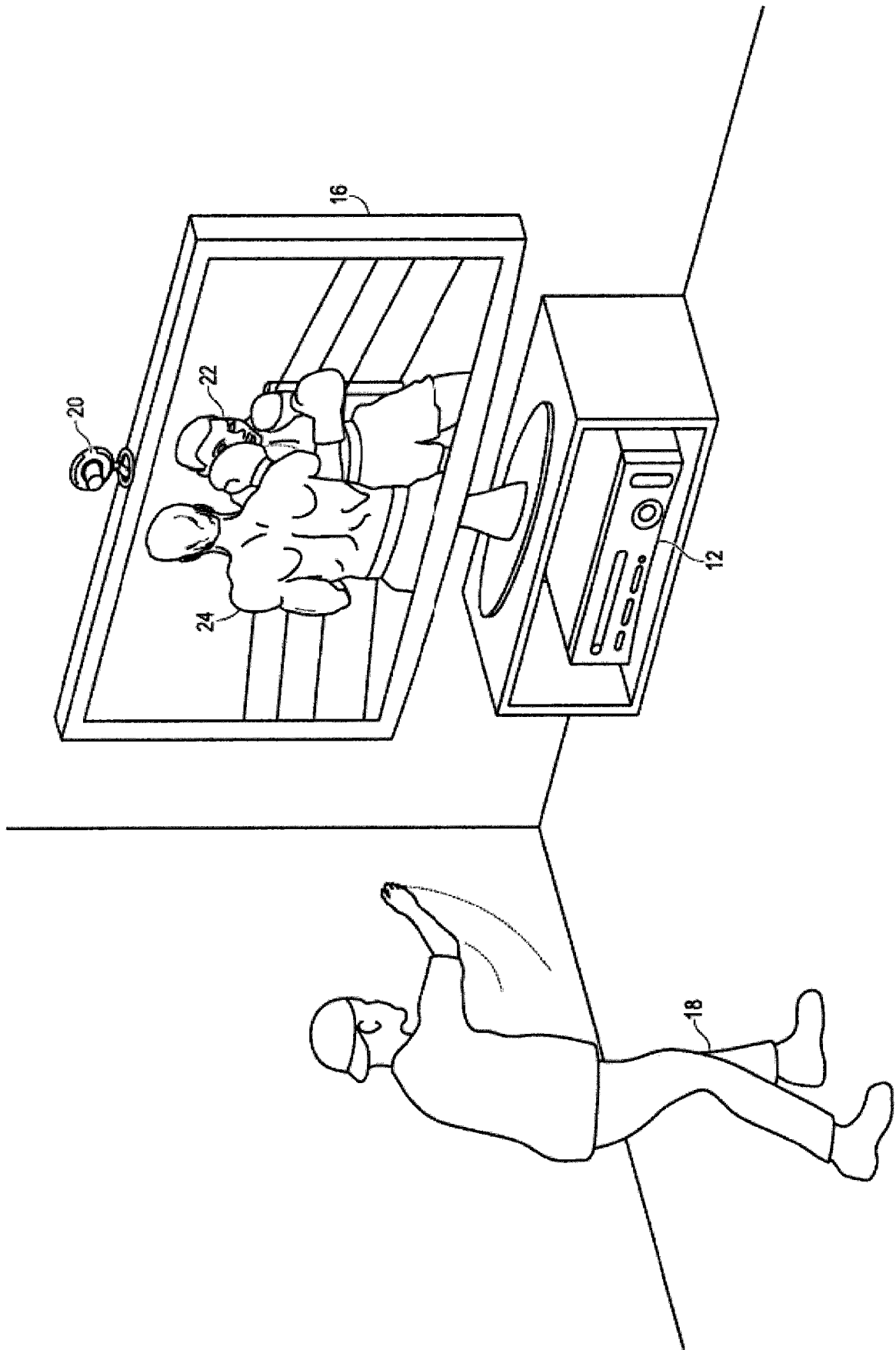


图 1B

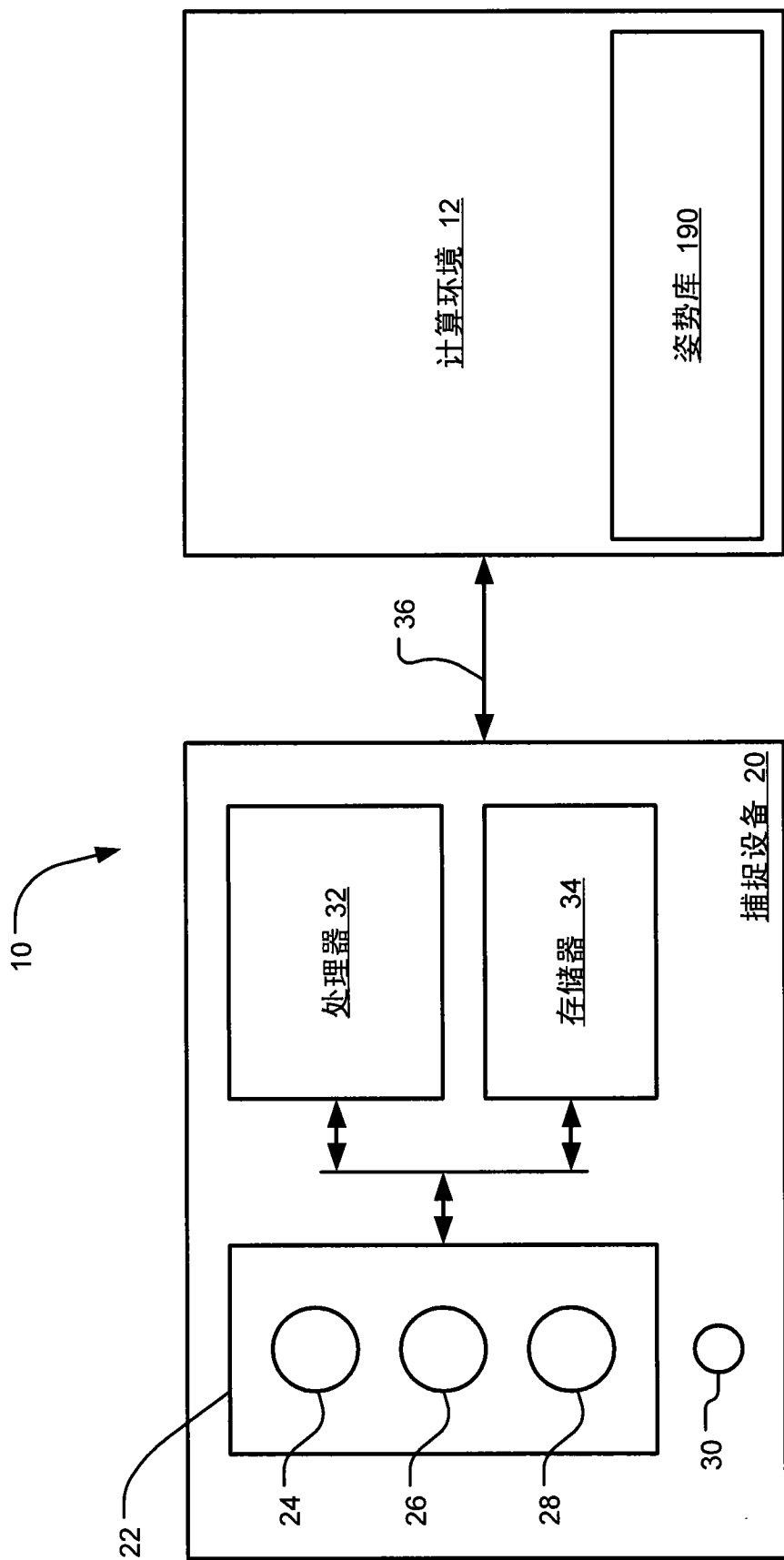


图 2

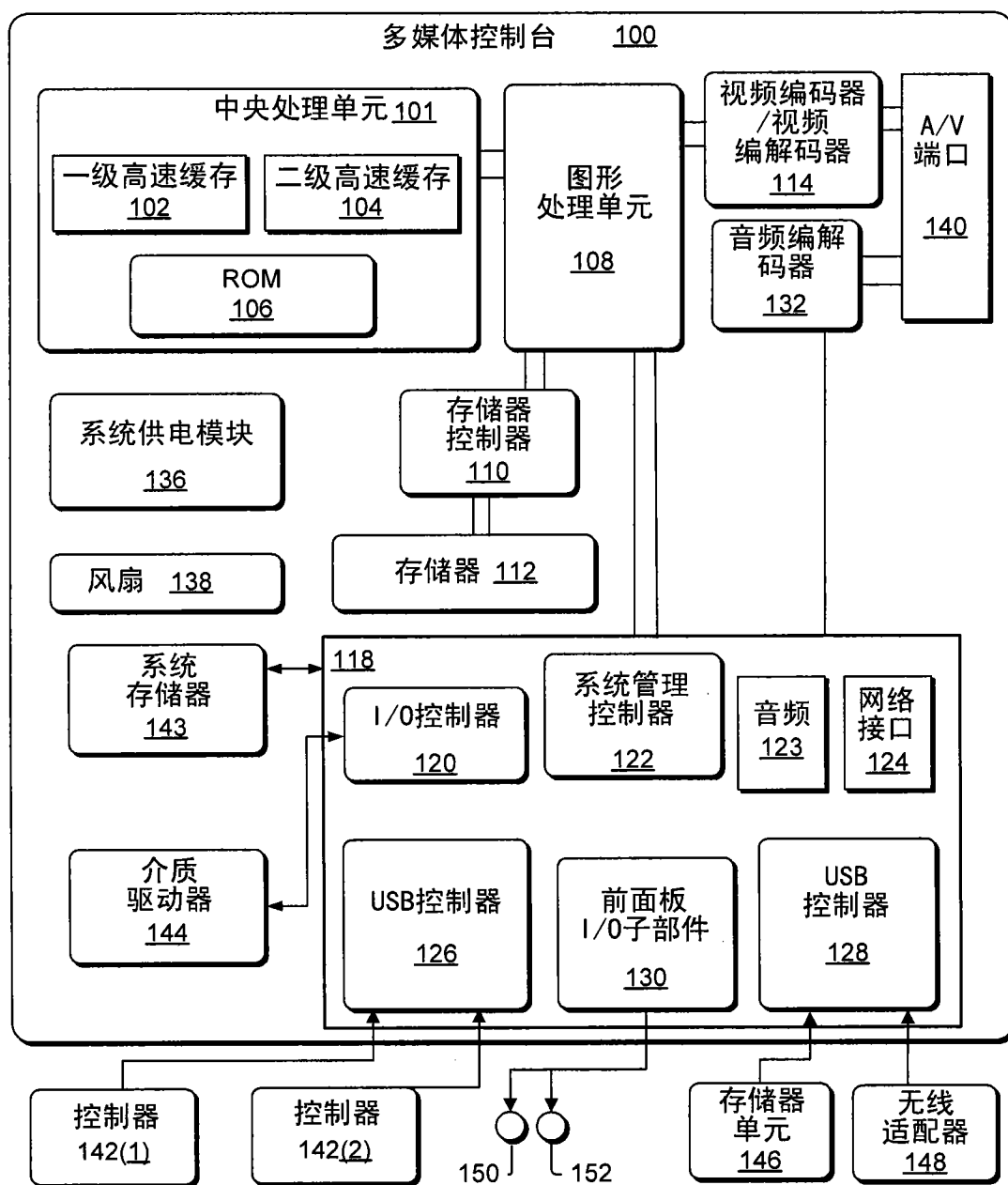


图 3A

计算环境 220

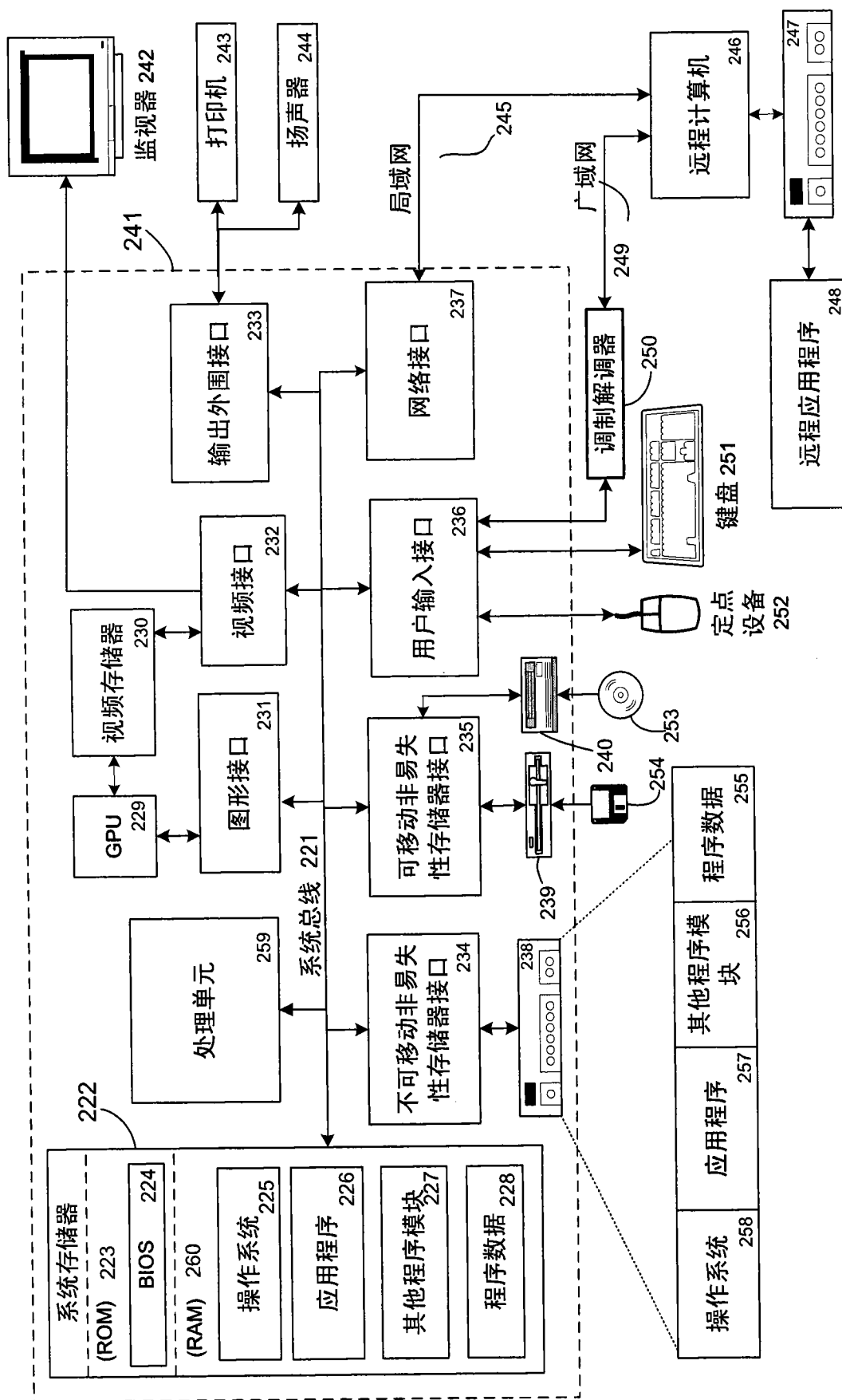


图 3B

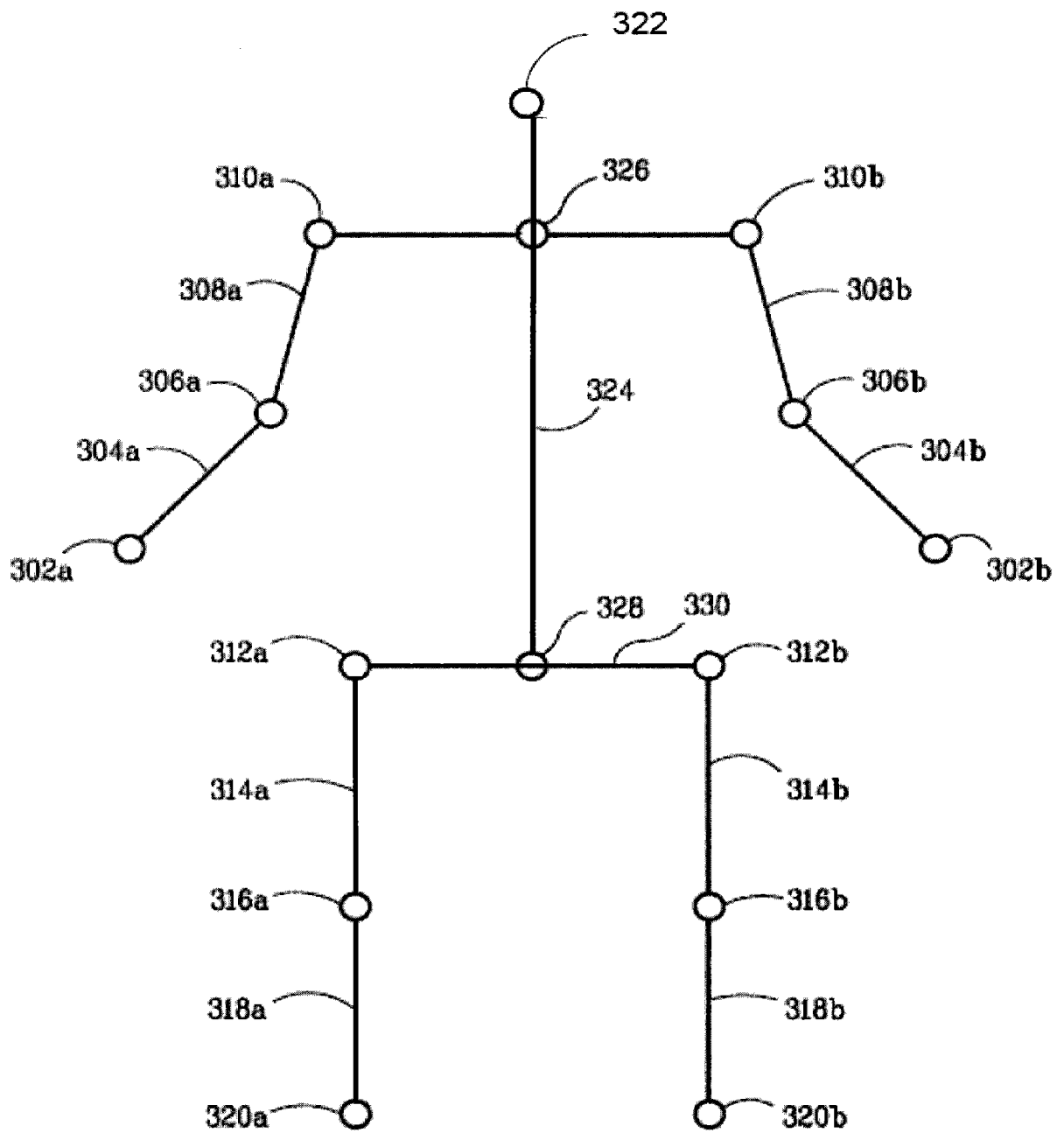


图 4A

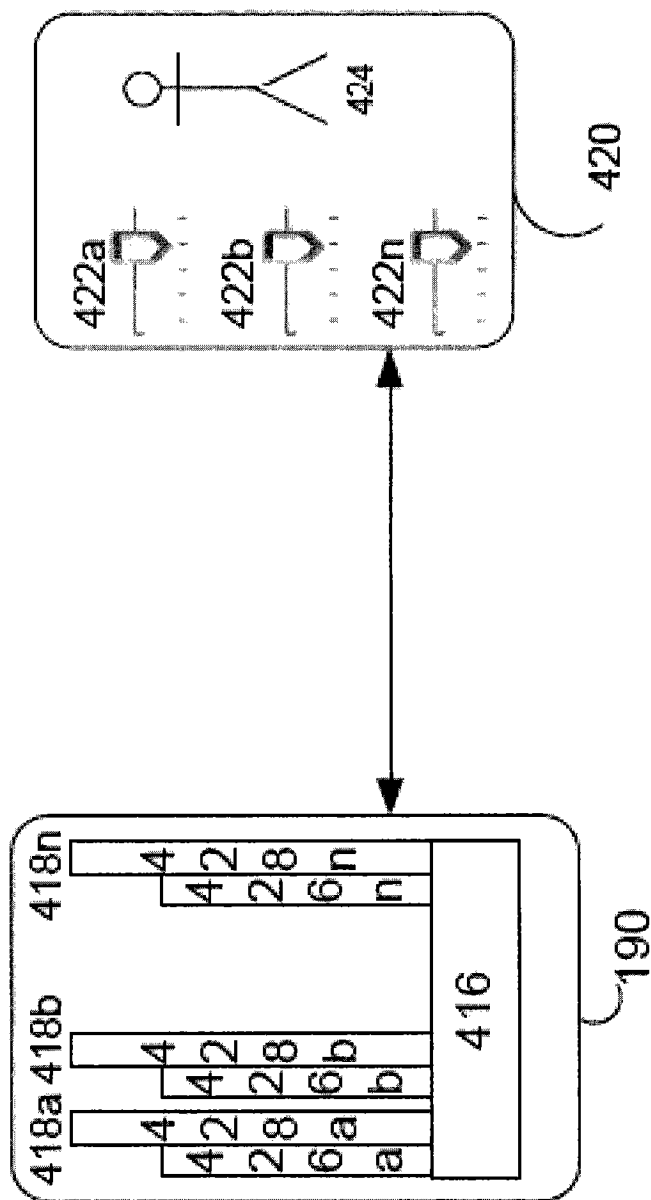


图 4B

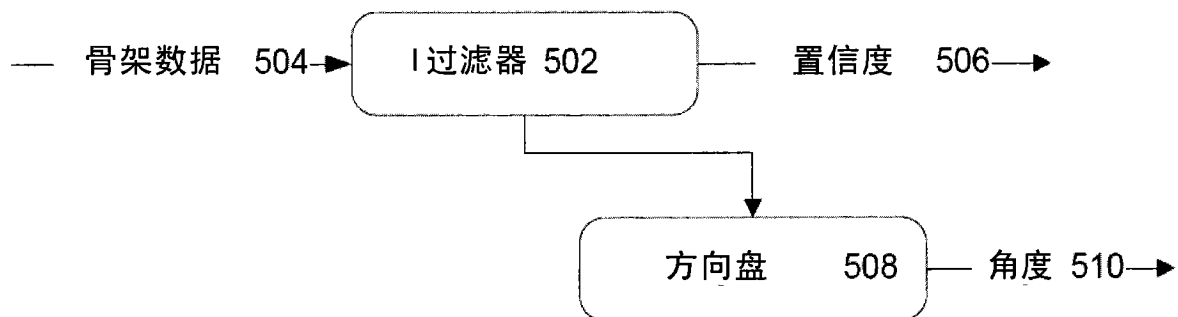


图 5A

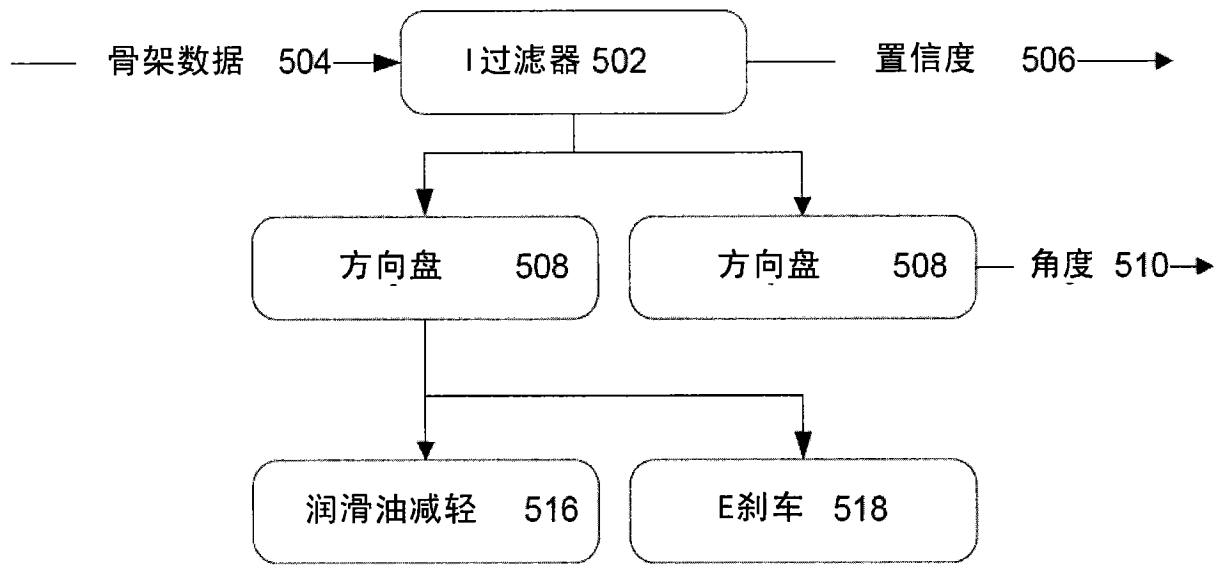


图 5B

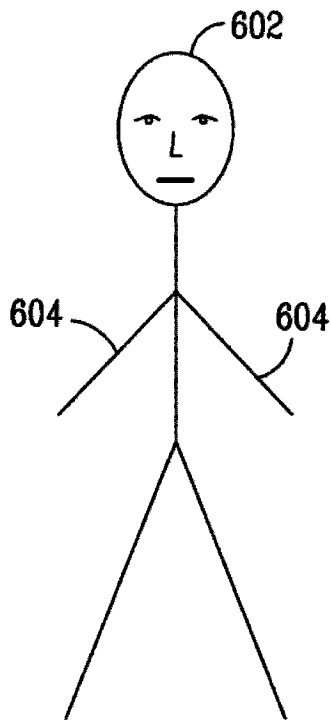


图 6A

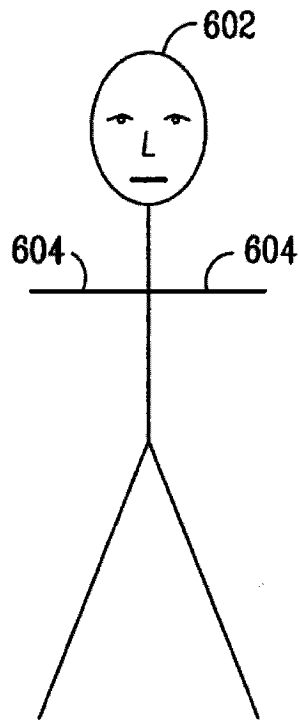


图 6B

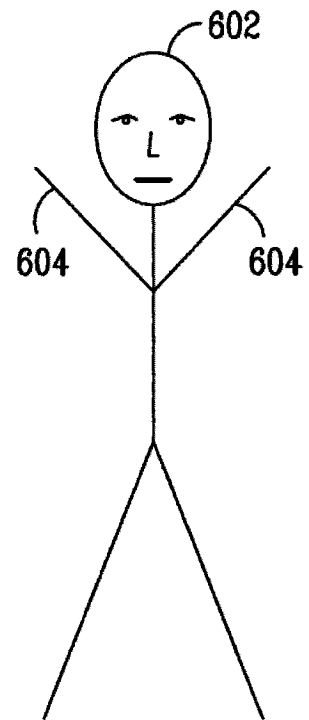


图 6C

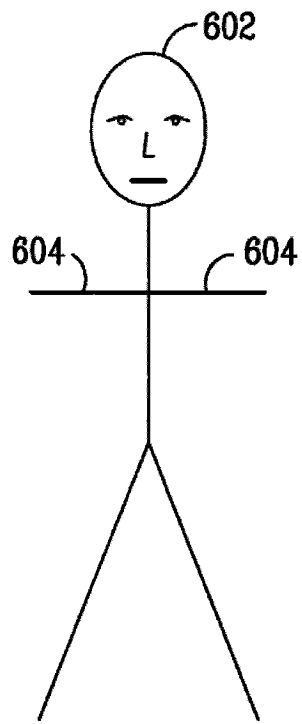


图 6D

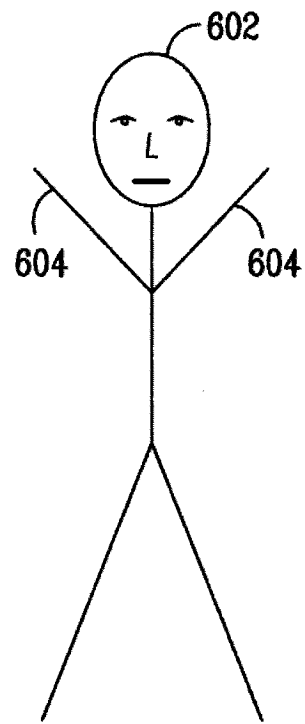


图 6E

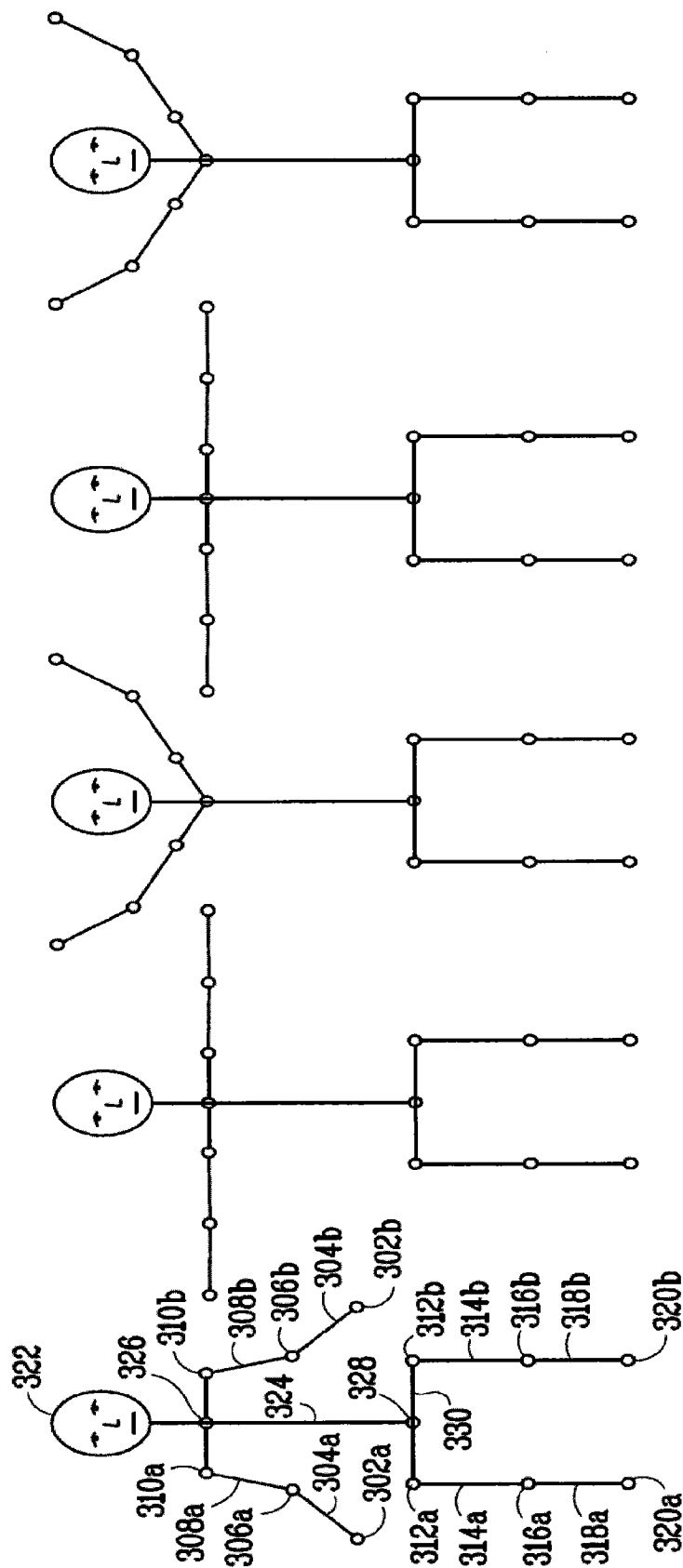


图 7A

图 7B

图 7C

图 7D

图 7E

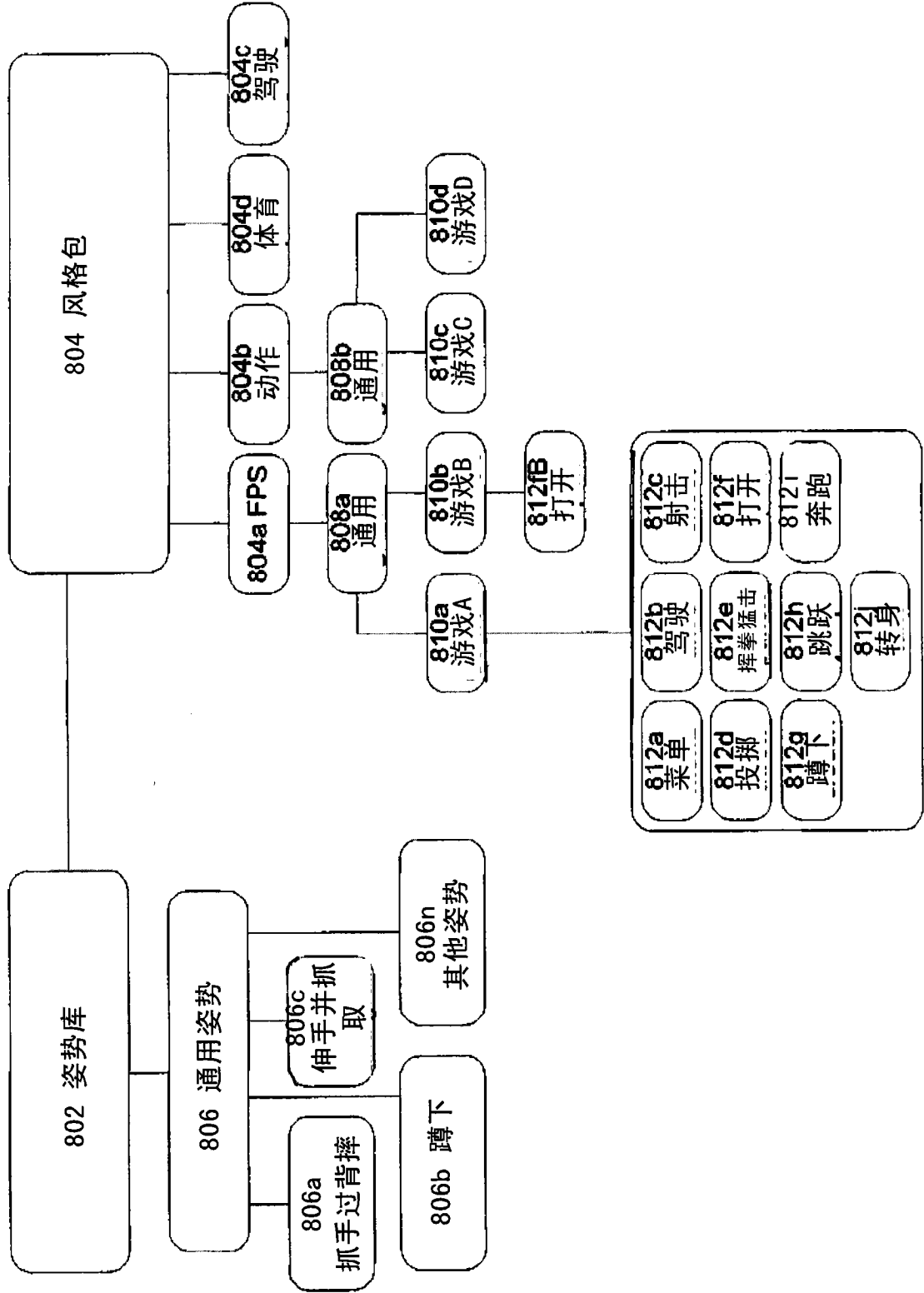


图 8

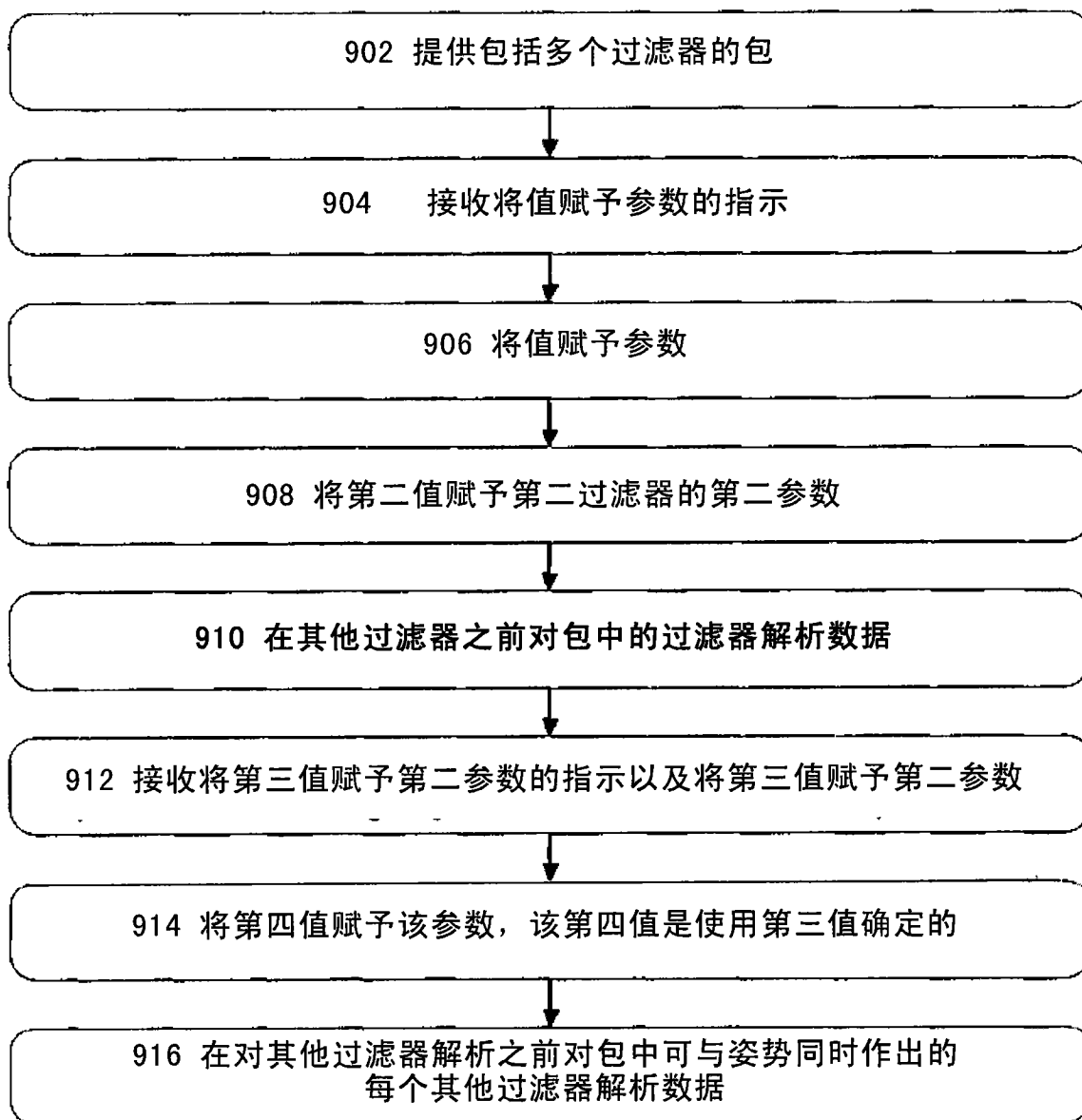


图 9