



(12) 发明专利

(10) 授权公告号 CN 103180893 B

(45) 授权公告日 2016.01.20

(21) 申请号 201280003480.6

(56) 对比文件

(22) 申请日 2012.07.05

US 2008/0059578 A1, 2008.03.06, 说明书第3页第0035段—第12页第0126段,附图2,3和9.

(30) 优先权数据

13/215,451 2011.08.23 US

(85) PCT国际申请进入国家阶段日

2013.04.26

US 6346929 B1, 2002.02.12, 说明书第6页第1行—第42行.

(86) PCT国际申请的申请数据

PCT/US2012/045566 2012.07.05

US 2002/0180728 A1, 2002.12.05, 全文.

(87) PCT国际申请的公布数据

W02013/028268 EN 2013.02.28

CN 101185114 A, 2008.05.21, 全文.

US RE42336 E, 2011.05.10, 全文.

审查员 冯莹

(73) 专利权人 索尼公司

地址 日本东京

(72) 发明人 太田孝彰

(74) 专利代理机构 中国国际贸易促进委员会专
利商标事务所 11038

代理人 李渤

(51) Int. Cl.

G09G 5/00(2006.01)

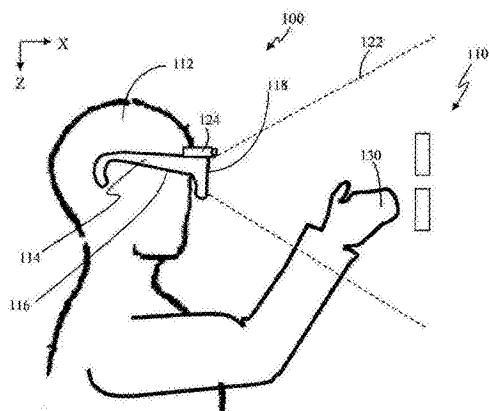
权利要求书3页 说明书14页 附图13页

(54) 发明名称

用于提供三维用户界面的方法和系统

(57) 摘要

一些实施例提供了用于显示用户界面的装置,包括:框架、安装在框架上的透镜、第一相机、检测器、以及处理器,该处理器被配置为:处理从第一相机接收的图像和从检测器接收的检测数据;至少通过对图像的处理,检测手相对于三维(3D)空间的手势,三维(3D)空间在第一相机的视场和检测器的检测区域中;通过对图像和检测数据的处理,识别执行手势的手的至少一部分在3D空间中的虚拟X、Y和Z坐标;识别对应于所检测到的手势和手的所述部分的三维位置的命令;以及执行该命令。



1. 一种显示用户界面的装置,该装置包括 :

框架 ;

透镜,安装在所述框架上,其中所述框架被配置为由用户佩戴以使得所述透镜的位置在用户的视线中;

第一相机,在所述框架的第一位置处安装在所述框架上,其中当用户适当地佩戴所述框架时,所述第一相机的位置在用户的视线内,使得所述第一相机捕获的图像对应于用户的视线;

检测器,安装在所述框架上,其中检测器被配置为检测当用户适当佩戴所述框架时对应于用户的视线的检测区域中的一个或多个对象;以及

处理器,被配置为:

处理从所述第一相机接收的图像和从所述检测器接收的检测数据;

至少通过对所述图像的处理,检测手相对于虚拟三维(3D)空间的手势,所述虚拟三维(3D)空间对应于所述第一相机的视场和所述检测器的检测区域;

通过对所述图像和所述检测数据的处理,识别执行所述手势的手的至少部分在所述3D空间中的虚拟X、Y和Z坐标;

识别对应于所检测到的手势和手的所述部分的三维位置的命令;以及
执行所述命令,

其中该处理器还被配置为:

识别虚拟选项,所述虚拟选项在检测到手的所述手势时虚拟地显示在所述3D空间中,并且对应于执行所述手势的手的所识别的X、Y和Z坐标,使得所述虚拟选项的至少部分被显示为在用户看来位于所述X、Y和Z坐标附近,

其中所述处理器在识别所述命令时还被配置为识别对应于所识别的虚拟选项和手的所检测到的手势的命令,并且所述处理器在执行所述命令时还被配置为激活对应于所识别的虚拟选项和手的所检测到的手势的命令。

2. 根据权利要求1的装置,其中所述检测器是红外检测器,并且处理检测数据包括:根据从所述红外检测器检测到的检测数据来识别至少一个虚拟深度坐标。

3. 根据权利要求1的装置,其中所述检测器是在所述框架的不同于第一位置的第二位置处安装在所述框架上的第二相机,所述检测数据包括第二图像,并且其中所述处理器还被配置为处理从所述第一相机和所述第二相机接收的第一图像和第二图像。

4. 一种显示用户界面的系统,该系统包括:

框架;

透镜,安装在所述框架上,其中所述框架被配置为由用户佩戴以使得所述透镜的位置在用户的视线中;

第一相机,在所述框架的第一位置处安装在所述框架上,其中当用户适当地佩戴所述框架时,所述第一相机的位置与用户的视线相符,使得所述第一相机捕获的图像对应于用户的视线;

第二相机,在所述框架的不同于第一位置的第二位置处安装在所述框架上,其中当用户适当地佩戴所述框架时,所述第二相机的位置与用户的视线相符,使得所述第二相机捕获的图像对应于用户的视线;以及

处理器,被配置为 :

处理从所述第一相机和所述第二相机接收的图像;

通过对所述图像的处理,检测手相对于三维(3D)空间的手势,所述三维(3D)空间在所述第一相机和所述第二相机的视场中;

通过对所述图像的处理,识别执行所述手势的手的至少部分在所述3D空间中的X、Y和Z坐标;

识别虚拟选项,所述虚拟选项在检测到手的手势时虚拟地显示在所述3D空间中,并且对应于执行手势的手的所识别的X、Y和Z坐标,使得所述虚拟选项的至少部分被显示为在用户看来位于所述X、Y和Z坐标处;

识别对应于所识别的虚拟选项和手的所检测到的手势的命令;以及

激活对应于所识别的虚拟选项和手的所检测到的手势的命令。

5. 根据权利要求4的系统,其中所述第一相机被配置为具有小于四英尺的景深。

6. 根据权利要求5的系统,其中所述第一相机被配置为具有定义为从距离所述相机六英寸之处延伸的小于四英尺的景深。

7. 根据权利要求5的系统,还包括:

红外光发射器,安装在所述框架上,并且被定位为将红外光发射到所述第一相机和所述第二相机的视场中,其中,所述第一相机和所述第二相机包括红外滤波器,以捕获红外光,使得所述第一相机和所述第二相机被限制为检测红外光。

8. 根据权利要求7的系统,还包括:

通信接口,安装在所述框架上,其中所述通信接口被配置为将来自所述第一相机和所述第二相机的图像传递到位置远离所述框架的所述处理器。

9. 根据权利要求5的系统,还包括:

通信接口,安装在所述框架上,其中所述通信接口被配置为将来自所述第一相机和所述第二相机的图像传递到位置远离所述框架的所述处理器,并且所述通信接口被配置为接收要在透镜上显示的图形信息。

10. 根据权利要求9的系统,其中该图形包括用户的表示。

11. 一种显示用户界面的方法,包括:

在显示三维演示的同时,接收由第一相机捕获的第一图像序列,其中所述第一相机安装在用户佩戴的框架上,使得当用户佩戴所述框架时,所述第一相机的视场在用户的视场中;

从安装在所述框架上的检测器接收检测区域内的一个或多个对象的检测数据,所述检测区域对应于用户适当佩戴所述框架时用户的视线;

处理所述第一图像序列;

处理所述检测器检测的检测数据;

通过对所述第一图像序列的处理,检测预定义的非传感器对象和所述非传感器对象的预定义的手势;

通过对所述第一图像序列和所述检测数据的处理,识别所述非传感器对象的至少部分相对于虚拟三维(3D)空间的虚拟X、Y和Z坐标,所述虚拟三维(3D)空间对应于所述第一相机的视场和所述检测器的检测区域;

识别对应于所检测到的手势和所述非传感器对象的虚拟 3D 位置的命令；以及执行所述命令，

其中，所述方法还包括：

识别虚拟选项，所述虚拟选项在检测到所述手势时虚拟地显示在所述三维演示中，并且对应于所述非传感器对象的三维坐标，所述三维演示配置为要被显示且在用户的视场中；并且

识别命令的步骤包括：识别对应于所识别的虚拟选项和相对于所述虚拟选项的手势的命令。

12. 根据权利要求 11 的方法，其中接收检测数据的步骤包括：

在显示三维演示的同时，接收由第二相机捕获的第二图像序列，其中所述第二相机安装在所述框架上，使得当用户佩戴所述框架时，所述第二相机的视场在用户的视场中。

13. 根据权利要求 11 的方法，其中显示三维演示的步骤包括显示对所述非传感器对象的模拟。

14. 根据权利要求 13 的方法，其中显示对非传感器对象的模拟的步骤包括显示对安装在所述框架上的透镜的模拟。

用于提供三维用户界面的方法和系统

技术领域

[0001] 本发明一般涉及演示,更特别地涉及多媒体演示。

背景技术

[0002] 有多种设备使得用户能够访问内容。这些设备中许多回放用户要观看的内容。另外,一些回放设备被配置为这样回放内容,使得对于用户而言,回放似乎是在三维空间中进行的。

发明内容

[0003] 本发明的若干实施例有利地提供了一些优点,使得装置、系统、方法和过程能够用于允许用户与虚拟环境进行交互。这些实施例中的一些提供了被配置为显示用户界面的装置,其中该装置包括:框架;透镜,安装在该框架上,其中该框架被配置为由用户佩戴以使得该透镜的位置在该用户的视线中;第一相机,在该框架的第一位置处安装在该框架上,其中当用户适当地佩戴该框架时,第一相机的位置在用户的视线内,使得第一相机捕获的图像对应于该用户的视线;检测器,安装在该框架上,其中第二检测器被配置为检测当用户适当佩戴该框架时对应于该用户的视线的检测区域中的一个或多个对象;以及处理器,被配置为:处理从第一相机接收的图像和从该检测器接收的检测数据;至少通过对该图像的处理,检测手相对于虚拟三维(3D)空间的手势,该虚拟三维(3D)空间对应于第一相机的视场和该检测器的检测区域;通过对该图像和该检测数据的处理,识别执行该手势的手的至少一部分在该3D空间中的虚拟X、Y和Z坐标;识别对应于所检测到的手势和手的该部分的三维位置的命令;以及执行该命令。

[0004] 其它实施例提供用于显示用户界面的系统。这些系统包括:框架;透镜,安装在该框架上,其中该框架被配置为由用户佩戴以使得该透镜的位置在该用户的视线中;第一相机,在该框架的第一位置处安装在该框架上,其中当用户适当地佩戴该框架时,第一相机的位置与用户的视线相符,使得第一相机捕获的图像对应于该用户的视线;第二相机,在该框架的不同于第一位置的第二位置处安装在该框架上,其中当用户适当地佩戴该框架时,第二相机的位置与用户的视线相符,使得第二相机捕获的图像对应于该用户的视线;以及处理器,被配置为:处理从第一相机和第二相机接收的图像;通过对该图像的处理,检测手相对于三维(3D)空间的手势,该三维(3D)空间在第一相机和第二相机的视场中;通过对该图像的处理,识别执行该手势的手的至少一部分在该3D空间中的X、Y和Z坐标;识别虚拟选项,该虚拟选项在检测到手的该手势时虚拟地显示在该3D空间中,并且对应于执行该手势的手的所识别的X、Y和Z坐标,使得该虚拟选项的至少一部分被显示为在用户看来位于该X、Y和Z坐标处;识别对应于所识别的虚拟选项和手的所检测到的手势的命令;以及激活对应于所识别的虚拟选项和手的所检测到的手势的命令。

[0005] 一些实施例提供方法,包括:在显示三维演示的同时,接收由第一相机捕获的第一图像序列,其中第一相机安装在用户佩戴的框架上,使得当用户佩戴该框架时,第一相机的

视场在该用户的视场中；从安装在该框架上的检测器接收检测区域内的一个或多个对象的检测数据，该检测区域对应于用户适当佩戴该框架时该用户的视线；处理第一图像序列；处理该检测器检测的检测数据；通过对第一图像序列的处理，检测预定义的非传感器对象和该非传感器对象的预定义的手势；通过对第一图像序列和该检测数据的处理，识别该非传感器对象的至少一部分相对于虚拟三维(3D)空间的虚拟X、Y和Z坐标，该虚拟三维(3D)空间在第一相机的视场和该检测器的检测区域；识别对应于所检测到的手势和该非传感器对象的虚拟3D位置的命令；以及执行该命令。

附图说明

[0006] 通过结合下面的附图给出的更具体的描述，本发明的若干实施例的上述和其它方面、特征以及优点将变得清楚。

[0007] 图1示出了根据一些实施例配置为使得用户能够与虚拟环境进行交互的用户交互系统的简化侧面视图。

[0008] 图2示出了图1的交互系统的简化俯视图。

[0009] 图3示出了图1的用户交互系统的简化俯视图，其中用户与3D虚拟环境进行交互。

[0010] 图4A-C示出了根据一些实施例可用于图1的交互系统的用户佩戴护目镜的简化俯视图。

[0011] 图5A示出了根据一些实施例的用户交互系统的简化框图。

[0012] 图5B示出了根据一些实施例的用户交互系统的简化框图，包括护目镜，该护目镜在其透镜上显示多媒体内容。

[0013] 图6A示出了根据一些实施例观看并与3D虚拟环境进行交互的用户的简化俯视图。

[0014] 图6B示出了图6A的观看并与3D虚拟环境进行交互的用户的侧视图。

[0015] 图7示出了根据一些实施例使得用户能够与3D虚拟环境进行交互的过程的简化流程图。

[0016] 图8示出了根据一些实施例使得用户能够与3D虚拟环境进行交互的过程的简化流程图。

[0017] 图9示出了根据一些实施例用户与通过用户交互系统提供的虚拟环境进行交互的简化俯视图。

[0018] 图10示出了根据一些实施例被配置为实现提供用户交互虚拟环境的方法、技术、设备、装置、系统、服务器、源等的系统的简化框图。

[0019] 图11图解了用于实现根据一些实施例提供用户交互虚拟环境的方法、技术、设备、装置、系统、服务器、源等的系统。

[0020] 在所有这些附图中，对应的附图标记表示对应的组件。本领域技术人员将理解，图中的元件是出于简化和清楚的目的而示出的，没有必要按比例绘制。例如，图中一些元件的尺寸可以相对于其它元件而被放大，以便有助于理解本发明的各个实施例。同样，在商业可用实施例中有用或必要的普通而又容易理解的元件往往没有示出，以便于提供本发明的这些各种实施例的较少妨碍的视图。

具体实施方式

[0021] 下面的描述不是以限制的意义进行的,而是仅仅用于描述示例性实施例的一般原理的目的。本发明的范围应当参考权利要求来确定。

[0022] 在整个这一说明书中提到“一个实施例”、“一实施例”、“一些实施例”、“一些实现”或类似的语言,是指结合该实施例描述的具体特征、结构或特性是包括在本发明的至少一个实施例中的。因此,在整个这一说明书中,出现词组“在一个示例”、“在一实施例中”、“在一些实施例中”以及类似语言,都可以但并不必须,是指相同的实施例。

[0023] 此外,所描述的本发明的特征、结构或特性可以在一个或多个实施例中以任何适当的方式结合。在下面的描述中,提供了许多具体细节,例如变成的示例、软件模块、用户选择、网络事务、数据库查询、数据库结构、硬件模块、硬件电路、硬件芯片等,以提供对本发明的实施例彻底的理解。然而,本领域技术人员将认识到,本发明可以在没有一个或多个具体细节的情况下,或者用其它方法、组件、材料等,来实践。在其它情况下,没有详细示出或描述公知的结构、材料、或操作,以避免遮蔽本发明的方面。

[0024] 一些实施例提供了向用户提供具有多媒体内容演示的三维(3D)交互的方法、过程、设备或系统。此外,该交互可以使得用户能够使用她或他的手或他们的手中所持有的对象来与虚拟3D显示的环境和/或用户界面进行交互。使用图像捕获和/或其它检测器,可以相对于3D虚拟环境中的位置来识别用户的手,并且可以响应于用户的交互而实现功能和/或命令。此外,在一些实施例中,至少一些功能和/或命令是基于手势或手的预定义运动来识别的。

[0025] 图1示出了根据一些实施例配置为使得用户112能够与3D虚拟环境110进行交互的用户交互系统100的简化侧面视图。类似地,图2示出了图1的交互系统100的简化俯视图,其中用户112与3D虚拟环境110进行交互。参考图1和2,用户112佩戴眼镜或护目镜114(为了简便,下面称为“护目镜”),其使得用户能够观看3D虚拟环境110。护目镜114包括框架116和安装在框架上的一个或多个透镜118。框架116被配置为由用户112佩戴,以将透镜118放置在用户的视场122中。

[0026] 一个或多个相机和/或检测器124-125也与框架116配合和/或与框架116安装在一起。相机或检测器124-125的位置还被设置为使得当用户适当佩戴框架时,相机的视场和/或检测器的检测区域对应于用户的视场122或者在用户的视场122之内。例如相机124的位置被设置为使得由第一相机捕获的图像对应于用户的视场。在一些实现中,第一相机124位于框架116上,并且检测器125位于框架上。与检测器125协作来使用第一相机124使得用户交互系统100能够识别对象,例如用户的手130、用户手(例如手指)的位置、和/或其它对象(例如非传感器对象),并且进一步识别该对象相对于相机124和/或检测器125的位置的三维(X、Y和Z)坐标,而这可以与所显示的3D虚拟环境110中的X、Y和Z坐标相关联。检测器基本上可以是使得用户交互系统100能够检测用户的手130或其他非传感器对象并且至少辅助确定相对于3D虚拟环境110的X、Y和Z坐标的任何相关检测器。在一些情况下,由于额外的图像处理,与使用两个相机相比,相机124和检测器的使用可以减少一些由用户交互系统100在提供3D虚拟环境和检测用户与该环境的交互时执行的处理。

[0027] 在其它实施例中,第一相机 124 的位置在框架 116 的第一位置处,第二相机 125 的位置在框架 116 的与第一位置不同的第二位置处。因此,当使用两个相机时,从已知的两个不同位置产生的两个图像使得用户交互系统 100 能够确定用户的手 130 或其它对象的相对位置。此外,在第一和第二相机 124-125 处于相对于彼此已知的位置的情况下,可以基于由两个相机捕获的图像来确定 X、Y 和 Z 坐标。

[0028] 图 3 示出了与通过护目镜 114 观看的 3D 虚拟环境 110 进行交互的图 1 的用户 112 的简化俯视图。在两个相机 124-125 与护目镜 114 相固定或者以其它方式与护目镜 114 协作的那些实施例中,第一相机 124 的位置被设置为使得当用户适当佩戴护目镜时,第一相机 124 的第一视场 312 与用户的视场 122 对应、或者在其中和 / 或与其至少主要部分重叠。类似地,第二相机 125 的位置被设置为使得第二相机 125 的视场 313 与用户的视场 122 对应、或者在其中和 / 或与其至少主要部分重叠。此外,当使用检测器或其它传感器替代第二相机 125 或者与第二相机 125 协作时,类似地,检测器的检测区或区域 313 与用户的视场 122 对应、或者在其中和 / 或与其至少主要部分重叠。

[0029] 在一些实施例中,第一和 / 或第二相机 124-125 的景深 (DOF) 316 可以被限制以增强对从一个或两个相机取回的图像的检测和 / 或精度。景深 316 可以被限定为由第一或第二相机 124-125 捕获的图像中呈现可接受的清晰度的图像或场景中最近和最远对象之间的距离。第一相机 124 的景深可以被限制为相对靠近用户 112,这可以提供对手 130 或其它尝试检测的对象更显著的隔离。此外,在有限的景深 316 下,背景是模糊的,这使得更加容易检测手 130,并将其从背景中区分出来。另外,在使用手 130 或用户的手握持的其它对象的那些实施例中,景深 316 可以被配置为从用户附近延伸到大致上相当于或者刚刚超过通常用户的手臂长度或通常用户的手臂能够达到的位置。在一些情况下,例如,景深 316 可以从离相机或框架大约六英寸延伸到大约三或四英尺。这将导致这一范围之外的对象的急剧散焦以及景深之外清晰度的急剧下降,从而将手 130 分离出来,简化手或其它对象的相对深度坐标 (对应于沿着图 3 的 X 轴的 X 坐标) 以及沿着 Y 和 Z 轴的坐标的检测和确定。要注意的是,对应的 3D 虚拟环境 110 不必受此限制。虚拟环境 110 基本上可以具有任何配置,并且可以根据用户的方向、位置和 / 或运动而改变。

[0030] 在一些实施例中,来自第一和第二相机 124-125 中每一个的图像可以都被评价以识别所感兴趣的对象。例如,当尝试识别预定义对象 (例如用户的手 130) 时,可以通过发现两个图像 (左眼图像和右眼图像) 中的一致形状来评价图像以识别对象。一旦检测到一致性,可以对预定义和 / 或对应特征点 (例如但不限于手指尖、手指之间的分叉点、手指的弯曲处或关节、手腕和 / 或其它这样的特征点) 执行映射。可以至少部分地测量和使用两个或更多个图像之间的对应点之间的位移来计算从成像位置到那个点的距离 (并且在至少一些实施例中有效计算观测位置)。此外,当背景图像具有较少的细节和纹理时,有限的景深使得更容易识别一致性。

[0031] 此外,一些实施例使用附加的特征来改进对用户的手 130 或其它非传感器设备的检测。例如,第一和第二相机 124-125 中的一个或者两者都可以是红外 (IR) 相机和 / 或使用红外滤波。类似地,一个或多个检测器可以是 IR 检测器。这可以进一步减小背景效应等。还可以在框架 116 中并入和 / 或与框架 116 一同安装一个或多个红外发射器或光源 320,以在相机 124-125 的视场内发射红外光。类似地,当使用一个或多个检测器时,这些检测器中

的一个或多个可以是红外传感器，或其它可以检测用户的手 130 的传感器。例如，红外检测器可以用于检测热影像。一般而言，人体比周围环境更温暖。基于期望的热谱对图像进行滤波可以将人体和 / 或人体的部分（例如手）从周围的无机物中区分出来。另外，在连同红外光源（例如 IR LED）使用一个或多个红外相机的一些情况下，即使在黑暗的环境中，尽管对于人眼而言这一视野仍然是黑暗的，但是这一个或多个 IR 相机也可以精确地捕获用户的手或其它预定义对象。

[0032] 这一个或多个相机 124-125 和 / 或一个或多个其它相机可以进一步提供可用于显示用户的一个或多个手 130 的图像，例如相对于虚拟环境 110 的所识别的 X、Y 和 Z 坐标和 / 或真实世界的其它方面叠加。因此，用户 112 可以相对于虚拟环境 110 中的一个或多个虚拟对象 324 看到她 / 他的手。在一些实施例中，来自第一和第二相机 124-125 或其它相机的图像被转发到内容源，该内容源执行相关图像处理，并且将用户的手的图像或用户的手的图形表示合并到用户 112 正在观看的 3D 演示和虚拟环境 110 中。

[0033] 另外，因为相机或检测器极为靠近用户的手 130，护目镜 114 处相机和 / 或检测器的使用提供了对用户的手 130 的更精确的检测。由于用户的位置有可能相对于这些相机的放置位置而改变，远离用户 112 并且指向用户的相机通常必须被配置为具有相对大的景深。类似地，对从指向用户 112 的各相机到用户的手 130 的深度的检测会非常困难，这是因为用户和相机位置之间的潜在距离，以及因为与用户的手和远处的相机的位置之间的潜在距离相比，手指或手的运动距离的相对改变非常小，导致可能非常难以精确检测的非常小的角度差。作为替代，在相机 124-125 安装在护目镜 114 上的情况下，从相机 124-125 到用户的手 130 或手指的距离要小很多，并且从相机到手或手指的距离与手或手指的运动之间的比值要小很多，从而具有大得多的角度间距。

[0034] 如上所述，一些实施例使用两个相机 124-125。此外，两个相机的设置在不同的位置。图 4A-C 示出了佩戴护目镜 114 的用户 112 的简化俯视图，其中每个图中第一和第二相机 124-125 具有不同的放置位置。例如，在图 4A 中，第一和第二相机 124-125 位于框架 116 的相反侧 412-413。在图 4B 中，第一和第二相机 124-125 位于框架 116 的中心 416。在图 4C 中，第一和第二相机 124-125 被配置在单个图像捕获设备 418 中。例如，单个图像捕获设备 418 可以是 3D 或立体摄录像机（例如索尼公司的 HDR-TD10）、3D 相机（例如索尼公司的 3DBloggies®）、或其它具有通过单个设备提供 3D 图像捕获的特征的这种设备。那些利用一个或多个检测器取代第二相机 125 或与第二相机 125 结合的实施例可以类似地定位于单个设备和 / 或并入单个设备。

[0035] 一些实施例利用护目镜 114 来显示回放虚拟 3D 环境。因此，3D 环境的一些或全部直接显示在护目镜 114 的透镜 118 上。在其它实施例中，使用眼镜 114，使得在用户 112 看来，呈现在分开的显示器上的图像和 / 或视频就像在三维空间一样。图 5A 示出了根据一些实施例的用户交互系统 510 的简化框图。用户交互系统 510 包括由用户 112 佩戴的眼镜 514、显示器 518 以及要在显示器 518 上显示的多媒体内容（例如图像、视频、游戏图形和 / 或其它这样的可显示内容）的内容源 520。在一些情况下，显示器 518 和内容源 520 可以是单个单元，而在另一些实施例中，显示器 518 与内容源 520 分离。此外，在一些实施例中，内容源 520 可以是被配置为向显示器 518 提供可显示内容的一个或多个设备。例如，内容源 520 可以是回放本地内容（例如 DVD、蓝盘、视频游戏等）或远程内容（例如网络内容、来自

其它源的内容等)的计算机、机顶盒、卫星系统、相机、平板电脑、或其它这样的内容源。显示系统 516 显示视频、图形、图像、图片和 / 或其它这样的视觉内容。此外,通过与眼镜 514 协作,显示系统 516 为用户 112 显示虚拟三维环境 110。

[0036] 眼镜 514 包括一个或多个相机 124 和 / 或检测器(图 5A 中仅示出了一个相机)。相机 124 捕获相机的视场内用户的的手 130 的图像。处理系统可以与眼镜 514 协作,或者与眼镜 514 分离,例如独立型处理系统或任何其它系统的部分(例如内容源 520 或内容系统的部分)。处理系统接收图像以及 / 或者检测来自相机 124-125 和 / 或检测器的信息,确定相对于 3D 虚拟环境 110 的 X、Y 和 Z 坐标,并基于用户的的手 130 和当前显示的 3D 虚拟环境 110 来确定用户与 3D 虚拟环境 110 的交互。例如,基于用户的的手 130 的 3D 坐标,用户交互系统 510 可以识别用户尝试与所显示的虚拟对象 524 交互,该虚拟对象 524 被配置为对用户 112 呈现为在 3D 虚拟环境 110 中,并且在 3D 虚拟环境 110 中邻近用户的的手的所确定的 3D 坐标的位置处。虚拟对象 524 可以显示在眼镜 514 的透镜上,或者在显示器 518 上,同时以三维方式呈现在 3D 虚拟环境 110 中。

[0037] 所显示的虚拟对象 524 基本上可以是任何可以在 3D 虚拟环境 110 中显示并呈现的相关对象。例如,该对象可以是用户可选择的选项、按钮、虚拟滑块、图像、字符、武器、图标、写入设备、图形、表格、文本、键盘、指示器、或其它这样的对象。此外,可以显示任何数量的虚拟对象。

[0038] 在一些实施例中,眼镜 514 与内容源 520 或执行检测器和 / 或图像处理的一些或全部功能的其它相关设备通信。例如,在一些情况下,眼镜可以包括具有一个或多个无线收发器的通信接口,该无线收发器可以向内容源 520 传送图像和 / 或检测器数据,使得该内容源可以执行以下处理的一些或全部:确定用户的的手 130 和 / 或用户的的手的位置的相对虚拟坐标、识别手势、识别相应的命令、执行命令和 / 或其它处理。在眼镜 514 处执行上述处理中的一些或全部的实施例中,眼镜可以包括一个或多个处理系统和 / 或与一个或多个处理系统(例如由用户 112 额外携带或经由有线或无线通信与眼镜 514 通信的系统)耦接。

[0039] 图 5B 示出了根据一些实施例的用户交互系统 510 的简化框图。用户 112 佩戴护目镜 114,在护目镜的透镜 118 上显示多媒体内容,使得不需要分立的显示器。护目镜 114 与提供要由护目镜显示和 / 或回放的内容的内容源 520 有线或无线通信。

[0040] 如上所述,内容源 520 可以是护目镜 114 的一部分,或者与护目镜分离。内容源 520 可以提供内容和 / 或执行图像和 / 或检测器处理中的一些或全部。内容源 520 和护目镜 114 之间的通信可以经由有线(包括光学)和 / 或无线通信。

[0041] 图 6A 示出了观看 3D 虚拟环境 110 并与 3D 虚拟环境 110 进行交互的用户 112 的简化俯视图;图 6B 示出了图 6A 的观看 3D 虚拟环境 110 并与 3D 虚拟环境 110 进行交互的用户 112 的侧面平面图。参考图 6A-B,在 3D 虚拟环境 110 中,多个虚拟对象 612-622 对用户 112 可见。用户可以与一个或多个虚拟对象进行交互,例如通过由用户的的手 130 虚拟地接触虚拟对象(例如对象 612)。例如,虚拟环境 110 可以是或者可以包括所显示的 3D 虚拟仪表盘,使得用户能够精确地控制可通过该仪表盘实现的功能。在其它情况下,例如在玩视频游戏或者至少部分地控制视频游戏、游戏的回放和 / 或游戏中的一一个或多个虚拟设备、角色、化身时,用户可以与虚拟环境交互。如上所述,虚拟对象 612-622 可以显示在护目镜 114 的透镜 118 上或者显示在用户 112 通过眼镜 114 可见的分立显示器 518 上。虚拟对象

612-622 可以显示为对用户 112 而言呈现在 3D 虚拟环境 110 中不同的位置,包括在 X、Y 和 / 或 Z 方向上分布。因此,虚拟对象 612-622 可以相对于用户 112 以不同的距离、深度和 / 或分层显示。

[0042] 当正在向用户显示演示内容时,用户交互系统 100 捕获图像。演示期间获得的图像和 / 或检测器信息被处理以识别用户的手 130 或其它预定义对象。一旦识别,用户交互系统识别用户的至少部分(例如手指 630)的相对 X、Y 和 Z 坐标,包括用户的该部分的虚拟深度(沿着 X 轴)。基于所识别的用户的该部分或用户的该部分在 3D 虚拟环境 110 中的位置,用户交互系统 100 识别用户尝试接触、选择、移动等的一个或多个虚拟对象 612-622。此外,用户交互系统 100 可以识别用户的手正在执行的一个或多个手势,例如选择、推动、抓住、移动、拖拉、试图放大或其它这样的动作。作为响应,用户交互系统可以识别与所识别的手势、用户的该部分的位置以及邻近用户的该部分的位置的对应对象相关联的一个或多个命令。例如,用户 112 可以选择对象(例如图片或图片组),并移动该对象(例如将该图片或图片组移动到文件或另一个图片组中);转动对象(例如转动虚拟旋钮);按压虚拟按钮;缩放(例如捏拉缩放型操作);滑动虚拟滑块指示符;滑动对象;推或拉对象;卷起、猛击键盘条目、目标和 / 或激活虚拟武器;移动机器人或采取其它动作。类似地,用户可以控制环境,例如转变为不同的控制、不同的所显示控制台或用户界面、不同的仪表盘;激活不同的应用或其它这样的控制,以及更复杂的导航(例如内容搜索、音频和 / 或视频搜索、玩视频游戏等)。

[0043] 在一些实施例中,音频系统 640 可以与护目镜 114 协作和 / 或与护目镜 114 安装在一起。在一些实施例中,音频系统 640 可以配置为检测音频内容,例如由用户 112 说出的词语、指令、命令等。音频系统 640 邻近放置可以使得能够进行精确的音频检测,并且容易与背景噪声和 / 或来自演示的噪声区分开。此外,可以在护目镜 114 处、部分的在护目镜处和 / 或远离护目镜处,执行对音频的处理。例如,可以由用户说出并由音频系统 640 检测诸如词语发音的音频命令,例如关闭、移动、打开、下一个、合并、以及其它这样的命令,以实现命令。

[0044] 图 7 示出了根据一些实施例使得用户能够与 3D 虚拟环境进行交互的过程 710 的简化流程图。在步骤 712,例如从第一相机 124 接收一个或多个图像、一系列图像和 / 或视频。在步骤 714,从与护目镜 114 协作的检测器接收检测器数据。可以额外地接收和利用其它信息,例如其它相机信息、运动信息、位置信息、音频信息等。在步骤 716,处理来自第一相机 124 的一个或多个图像。这个处理可以包括解码、解压缩、编码、压缩、图像处理、以及其它这样的处理。在步骤 720,在这一个或多个图像中识别用户的手或其它非传感器对象。在步骤 722,在图像处理中额外地识别一个或多个预定义的手势。

[0045] 在步骤 724,处理所检测到的数据,通过与图像数据协作来检测用户的手或非传感器对象,并且确定位置信息。在步骤 726,确定用户的手 130 的至少部分相对于虚拟环境 110 的虚拟 X、Y 和 Z 坐标(例如基于所确定的位置和手势信息确定手指尖的位置)。在步骤 728,基于位置信息、手势信息、虚拟对象的相对位置以及其它这样的因素,识别一个或多个要执行的命令。再一次地,命令可以基于用户的手在 3D 虚拟环境中被识别的坐标附近的位置处正被虚拟显示的一个或多个虚拟对象。在步骤 730,执行这一个或多个命令。需要注意的是,在一些情况下,这一个或多个命令可以依赖于演示的当前状态(例如基于检测到手

势时电影回放的时间点、正在回放视频游戏的那一部分等)。类似地,所实现的命令可以依赖于后续动作,例如用户响应于正在执行的命令而采取的后续动作。附加地或者可替换地,一些手势和 / 或做出该手势的对应位置可以与全局命令相关联,该全局命令可以与演示和 / 或用户交互系统 100 的操作状态无关地执行。

[0046] 如上所述,在步骤 716,该过程执行图像处理以识别用户的手 130 和 / 或其它对象,并跟踪手的运动。在一些实现中,图像处理可以包括噪声消减过滤处理(例如使用二维低通滤波器以及通过中值滤波器进行的隔离点去除等),附加地,随后可以进行二维差分滤波,二维差分滤波可以高亮突出用户的 手或其它预定义对象的轮廓线。附加地或者可替换地,可以施加二进制滤波,在一些情况下,二进制滤波可以用于产生黑白轮廓线图像。通常轮廓线是粗线和 / 或粗的区域。相应地,一些实施例执行修形滤波(shaving filter)(例如黑色区域延伸进入白色区域,而不将一个黑色区域连接到另一个黑色区域,那样将打断白线),以便使线条和 / 或区域变细。

[0047] 在一些实施例中,图像处理还可以包括特征检测算法,其中追溯线条,并观测切向矢量的变化,检测矢量快速变化的特征点,这可以指示角落、末端等的位置。例如,这些特征点可以是手指尖、手指之间的分叉或交叉点、手的关节等。可以通过邻近关系将特征点进一步分组,并通过例如旋转和缩放与参考物匹配。可以通过将多个数据的组映射到矢量空间来进一步执行图案匹配,并且通过这个空间中两个矢量之间的距离来测量相似性。一旦检测到用户的手或其它对象,可以在时间上连续的跟踪该特征点,以检测手的运动。在一些实施例中,一个或多个手势被定义为特征点的运动矢量(例如特征点随时间变化的位移)。例如,可以通过特征点的运动矢量来确定手指运动;通过对多个特征点组的运动矢量等进行求和来检测摆手运动。在一些实施例中,在护目镜的情况下,动态精确度可以通过显示屏幕和相机位置之间的相对静态关系来提高。在透视眼镜上安装一个或多个相机(即,显示器位于眼镜之外)的情况下,还可以例如通过检测显示器的特征点(例如四个角落、四个侧边、一个或多个反射设备、一个或多个 LED、一个或多个 IR LET)来检测远处的显示器。手势位置和虚拟 3D 环境的静态精确度可以通过施加校准(例如,在开始之前或者在使用系统的同时,系统可以请用户用手指接触空间中的虚拟 3D 参考点)来进一步改进。类似地,可以附加地或者可替换地使用预定义的动作(例如接触单个虚拟按钮(例如“播放”或“继续”按钮))。上面的处理可以对相机 124-125 捕获的每个图像和 / 或图像系列执行。

[0048] 图 8 示出了根据一些实施例使得用户能够与 3D 虚拟环境进行交互的过程 810 的简化流程图,其中该系统采用两个或更多个相机 124-125 来捕获用户的手 130 或其它非传感器对象的图像。在步骤 812,从第一相机 124 接收一个或多个图像、图像序列和 / 或视频。在步骤 814,从第二相机 125 接收一个或多个图像、图像序列和 / 或视频。在步骤 816,处理来自第一和第二相机 124-125 的一个或多个图像。

[0049] 在步骤 820,在这一个或多个图像中识别用户的手或其它非传感器对象。在步骤 822,附加地通过图像处理识别一个或多个预定义的手势。在步骤 824,识别用户的手 130 相对于护目镜 114 和虚拟环境 110 的虚拟 X、Y 和 Z 坐标。在步骤 826,识别与该预定义的手势和手的位置的相对虚拟坐标相关联的一个或多个命令。在步骤 828,执行所识别的一个或多个命令。

[0050] 再一次地,用户交互系统采用第一和第二相机 124-125 和 / 或检测器,以便不只识

别 Y 和 Z 坐标, 而且还识别用户的手 130 的虚拟深度坐标 (X 坐标)。用户的手的位置结合所识别的手势使得用户交互系统 100 能够精确地解释用户的意图, 并且采取适当的动作以使得用户能够与用户交互系统 100 和 / 或演示的回放进行虚拟的交互, 或者对其进行控制。

[0051] 一些实施例还延伸到虚拟环境 110, 以延伸超过用户的视场 122 或视景。例如, 一些实施例将虚拟环境延伸到用户的即时视场 122 之外, 使得用户可以转动他或她的头以观看虚拟环境 110 的额外的部分。可以通过一个或多个处理和 / 或设备来检测用户的运动。例如, 可以执行对来自护目镜 114 上的一个或多个相机 124-125 的序列图像的处理。可以使用所检测和捕获到的护目镜 114 和 / 或用户 112 的运动来产生通过逐个图像或逐个帧地收集的位置和方向数据, 该数据可以用于计算用户和 / 或护目镜的运动的许多物理性质, 例如沿任何轴的加速度和速度, 以及倾斜、前倾、横摇、侧倾以及遥测点。

[0052] 附加地或者可替换地, 在一些情况下, 护目镜 114 可以包括一个或多个惯性传感器、罗盘设备和 / 或其它可以帮助识别和 / 或量化用户的运动的相关设备。例如, 护目镜 114 可以被配置为包括一个或多个加速计、陀螺仪、倾斜传感器、运动传感器、近距离传感器、其它类似设备或它们的组合。作为例子, 可以从在三或四个点处例如通过弹簧、电阻应变计材料、光子传感器、磁性传感器霍尔效应设备、压电器件、电容传感器等弹性耦接的物质来检测。

[0053] 在一些实施例中, 其它相机或其它传感器可以跟踪用户的运动, 例如多媒体或内容源 520 处一个或多个相机, 和 / 或与多媒体源 (例如, 通过使得用户能够玩交互式视频游戏的游戏设备跟踪用户的运动的相机) 协作。可以在护目镜 114 上包括可用于识别护目镜并跟踪护目镜的运动的一个或多个灯、灯阵列或其它这样的可检测对象。

[0054] 因此, 在一些实施例中, 虚拟环境 110 可以延伸超过用户的视场 122。类似地, 虚拟环境 110 可以依赖于用户在看什么和 / 或用户的方向。

[0055] 图 9 示出了根据一些实施例与虚拟环境 110 交互的用户 112 的简化俯视图。如图所示, 虚拟环境延伸超过用户的视场 122。在图 9 的示例演示中, 多个虚拟对象 912-916 在用户的视场 122 中, 多个虚拟对象 917-918 部分地在用户的视场中, 同时仍有一个或多个其它虚拟对象 919-924 在用户的即时视场 122 之外。通过跟踪用户的运动和 / 或护目镜 114 的运动, 所显示的虚拟环境 110 可以使得用户能够观看虚拟环境 110 的其它部分。在一些情况下, 可以显示一个或多个指示虚拟环境 110 延伸超过用户的视场 122 的指示符 (例如箭头等)。因此, 在一些情况下, 虚拟环境可以延伸以完整地绕用户 112 一周和 / 或在 X、Y 和 Z 方向上完整地围绕用户。类似地, 由于所观看的是虚拟环境, 虚拟环境 110 可以根据用户的位置、视场 122 的方向、所检测到的预定义的手势 (例如用户的手 130 的位置和由用户执行的手势) 和 / 或演示的上下文, 潜在地显示多于三个方向轴和 / 或假设的方向。

[0056] 此外, 在一些情况下, 虚拟环境可以依赖于用户的位置和 / 或所检测到的用户执行的手势而改变。例如, 护目镜 114 可以识别, 或者与护目镜通信的系统可以确定, 用户正在看多媒体回放设备 (例如通过图像检测和 / 或与多媒体回放设备的通信), 并因此显示使得用户能够与多媒体回放设备交互的虚拟环境。类似地, 护目镜 114 可以检测, 或者与护目镜关联的系统可以确定, 用户当前正在看电器, 例如冰箱。护目镜 114 基于图像识别和 / 或与冰箱通信, 可以调节虚拟环境 110, 并且显示选项和 / 或与冰箱相关的信息 (例如内部温度、传感器数据、冰箱里的东西 (如果知道的话)、和 / 或其它这样的信息)。类似地, 用户

可以通过虚拟环境来激活设备和 / 或控制设备。例如，虚拟环境可以显示用于控制电器、机器人、医疗设备等的虚拟控制，使得电器、机器人等根据用户的手 130 的所识别的位置和所检测到的预定义的手势来采取适当的动作。作为具体的例子，可以由医生通过医生与向医生显示相关信息、图像和 / 或选项的虚拟环境 110 的交互来控制用于执行医学外科手术的机器人式外科设备。此外，医生甚至不需要在与患者和机器人相同的位置。在其它情况下，用户可以激活整个家庭控制台，并且选择用户想要与其交互的期望的设备。

[0057] 类似地，当多个显示器（例如 TV、计算机监视器等）可见时，在一些情况下，使用相机和 / 或方向信息可以使得用户交互系统 100 能够识别用户当前正在看那个显示器，并且调节与所感兴趣的显示器相关的虚拟环境、命令、仪表盘等。附加地或者可替换地，用户 112 可以执行对虚拟对象的移动命令，例如从一个显示器移动到另一个显示器，从一个文件夹移动到另一个文件夹等。在其它情况下，例如当观看的内容来自多个监控摄像机时，可以根据用户正在观看哪个监控摄像机来显示不同的控制台、控制和 / 或信息。

[0058] 在一些实施例中，例如当护目镜 114 禁止用户看到她 / 他自己的手和 / 或禁止用户的视野超出透镜 118 时，虚拟环境可以在虚拟环境中附加地显示图形信息（例如用户的手 130）。用户的手或其它真实世界内容可以叠加在其它对用户可见的内容上。类似地，虚拟环境可以包括显示虚拟对象之外真实世界中的一些或全部和 / 或用户的手，使得用户可以看到如果她或他移除护目镜将会看到什么。在一些实施例中，真实世界的显示可以通过由第一和第二相机 124-125 之一或两者和 / 或由单独的相机捕获的图像来实现，并且可以使得用户能够仍在佩戴着护目镜的同时四处运动。

[0059] 图 10 示出了根据可以用于实现用户交互系统 100 或提供上面或下面描述的用户交互式虚拟环境的其它方法、技术、设备、器件、系统、服务器、源等中的一些或全部的一些实施例的系统 1010 的简化框图。系统 1010 包括一个或多个相机或检测器 1012、检测器处理系统 1014、图像处理系统 1016、手势识别系统 1020、3D 坐标确定系统、护目镜或眼镜 1024、存储器和 / 或数据库 1026、以及控制器 1030。一些实施例还包括显示器 1032、图形发生系统 1034、方向跟踪系统 1036、具有一个或多个收发器的通信接口或系统 1038、音频检测系统 1040 和 / 或其它这样的系统。

[0060] 相机和 / 或检测器 1012 检测用户的手或其它预定义的对象。在一些情况下，检测可以包括 IR 运动传感器检测、定向热传感器检测、和 / 或包括两个定向光传感器并且能够逐步捕获两个定向图像的系列的相机。在一些实施例中，检测器处理系统 1014 处理来自一个或多个检测器例如 IR 运动传感器的信号。并且在许多情况下，检测器处理系统 1014 具有内部信号阈值，以将检测限制于大约用户手臂的长度，并因此检测大约手臂距离内的对象或用户的手。如上所述，图像处理系统 1016 提供各种图像处理功能，例如但不限于滤波（例如噪声滤波、双向差分滤波、二进制滤波、线条细化滤波、特征点检测滤波等）以及其它这样的图像处理。

[0061] 手势识别系统 1020 检测特征点，并且检测用于用户的手指和手的图案、或者预定义的对象的其它特征。此外，手势识别系统随时间跟踪特征点，以检测手势运动。在一些实施例中，3D 坐标确定系统比较来自第一相机的一个或多个图像的特征点和来自第二相机的一个或多个图像的特征点，并且测量对应特征点对之间的位移。该位移信息至少部分地可以用于计算特征点位置的深度或距离。

[0062] 如上所述,护目镜 1024 与至少一个相机和检测器或第二相机协作。基于由相机和 / 或检测器 1012 捕获的信息,检测器处理系统 1014 和图像处理系统 1016 识别用户的手,并且向 3D 坐标确定系统 1022 和手势识别系统 1020 提供相关信息,以识别在 3D 虚拟环境中的相对位置和相对于所显示的虚拟环境 110 的手势。在一些情况下,图像处理可以执行加法处理,以改善所捕获的图像和 / 或该图像中正被捕获的对象的质量。例如,可以执行图像稳定化,可以执行照明调整,以及其它这样的处理。护目镜 1024 可以具有右侧显示单元和左侧显示单元,它们在观看者面前展示三维图像。在使用眼镜的那些情况下,电信地将外部显示器 1032 静态地放置,而用户给自己确定位置以通过眼镜观看显示器。

[0063] 存储器和 / 或数据库 1026 基本上可以是任何相关的计算机和 / 或处理器可读存储器,位于护目镜 1024 和 / 或控制器 1030 本地,或者远离护目镜 1024 和 / 或控制器 1030,并且通过通信信道(经由有线连接或无线连接)来访问。此外,存储器和 / 或数据库基本上可以存储任何相关信息,例如但不限于,手势、命令、图形、图像、内容(例如多媒体内容、文本内容、图像、视频、图形、动画内容等等)、历史信息、用户信息、用户配置信息、以及其它这样的信息和 / 或内容。另外,存储器 1026 可以存储图像数据、中间图像数据、用于处理运动矢量的多帧图像、用于特征点图案匹配的图案矢量数据等等。

[0064] 显示器 1032 可以显示对该用户或其它用户(例如佩戴着辅助显示 3D 内容的眼镜 1024 的用户)可见的图形、电影、图像、动画和 / 或其它内容。图形发生系统 1034 基本上可以是任何这样的图形发生器,其用于根据例如具有视频游戏内容和 / 或其它这样的内容的代码等产生图形,以便在护目镜 1024 或外部显示器 1032 上显示,以展示合成的三维图像。

[0065] 在一些实施例中,可以实现方向跟踪系统 1036,以跟踪用户 112 和 / 或护目镜 1024 的运动。在一些实施例中,方向跟踪系统可以通过一个或多个方向传感器、相机或其它这样的设备和 / 或它们的组合来跟踪护目镜 114 的方向。例如,在一些实施例中,包括有一个或多个包括三个 X、Y 和 Z 线性运动传感器的方向传感器。可以附加地或可替换地使用一个或多个轴旋转角向运动传感器(例如三个 X、Y 和 Z 轴旋转角向运动传感器)。使用相机可以使得能够通过跟踪静态对象如显示屏(例如四个角落特征点)来检测方向的改变。

[0066] 一些实施例还包括一个或多个接收器、发送器和 / 或收发器 1038,以提供组件之间的内部通信,和 / 或诸如护目镜 114、游戏控制台或设备、外部显示器、外部服务器或经网络访问的数据库之间的外部通信,或其它这样的通信。例如,收发器 1038 可以用于例如通过局域网络、因特网或其它这样的网络来与其它设备或系统通信。此外,收发器 1038 可以配置为提供有线、无线、光学、光纤线缆或其它相关通信。一些实施例附加地包括一个或多个音频检测系统,音频检测系统可以检测来自用户的音频指令和 / 或命令,并且可以辅助解释和 / 或识别用户期望的与系统 1010 和 / 或虚拟环境 110 的交互。例如,一些实施例在护目镜 114 的框架 116 上合并一个或多个麦克风,和 / 或与一个或多个麦克风协作。音频处理可以通过音频检测系统 1040 来执行,这可以在护目镜 114 处、部分在护目镜处或远离护目镜而执行。附加地或可替换地,在一些情况下,音频系统可以回放要由用户(例如通过耳机、扬声器等)听到的音频内容。此外,音频检测系统 1040 可以根据方向跟踪,对多个音频信道提供不同的衰减,和 / 或对多信道音频应用衰减矩阵,以便旋转声音空间并使声音空间与虚拟空间相匹配。

[0067] 这里描述的方法、技术、系统、设备、服务、服务器、源等可以在许多不同类型的设

备和 / 或系统上使用、实现和 / 或运行。参考图 11，其中示出了根据一些实施例可以用于任何这样的实现的系统 1100。系统 1100 的一个或多个组件可以用于实现上面或下面提到的任何系统、装置或设备，或者这样的系统、装置或设备的部分，例如上面或下面提到的用户交互系统 100、系统 1010、眼镜或护目镜 114、1024、第一或第二相机 124-125、相机或检测器 1012、显示系统 516、显示器 518、内容源 520、图像处理系统 1016、检测器处理系统 1014、手势识别系统 1020、3D 坐标确定系统 1022、图形发生系统 1034、控制器 1030、方向跟踪系统 1036 等等中的任何一个。然而，当然并不要求必须使用系统 1100 或其任何部分。

[0068] 作为示例，系统 1100 可以包括控制器或处理器模块 1112、存储器 1114、用户接口 1116 以及一个或多个通信链接、通路、总线等 1120。电源（未示出）包括在系统 1100 中，或与其耦接。控制器 1112 可以通过一个或多个处理器、微处理器、中央处理单元、逻辑器件、局部数字存储器、固件和 / 或其它控制硬件和 / 或软件来实现，并且可以用于执行或辅助执行这里描述的方法和技术的步骤，并且控制各种通信、程序、内容、列表、服务、接口等。用户接口 1116 可以使得用户能够与系统 1100 进行交互，并且从该系统接收信息。在一些情况下，用户接口 1116 包括显示器 1122 和 / 或一个或多个用户输入 1124，例如遥控器、键盘、鼠标、轨迹球、游戏控制器、按键、触摸屏等，这可以是系统 1100 的部分，或者与系统 1100 有线或无线耦接。

[0069] 典型地，系统 1100 还包括一个或多个通信接口、端口、收发器 1118 等，使得系统 1100 能够通过分布式网络、局域网络、互联网、通信链路 1120、其它网络或与其它设备的通信信道、和 / 或其它这样的通信方式进行通信。此外，收发器 1118 可以被配置为用于有线、无线、光学、光纤线缆或其它这样的通信配置或这样的通信的组合。

[0070] 系统 1100 包括具有控制器 1112 的控制和 / 或基于处理器的系统的例子。再一次地，控制器 1112 可以通过一个或多个处理器、控制器、中央处理单元、逻辑器件、软件等来实现。此外，在一些实现中，控制器 1112 可以提供多处理器功能性。

[0071] 可以由控制器 1112 访问的存储器 1114 典型地包括由至少一个控制器 1112 访问的一个或多个处理器可读和 / 或计算机可读介质，并且可以包括易失性和 / 或非易失性介质，例如 RAM、ROM、EEPROM、快闪存储器和 / 或其它存储技术。此外，存储器 1114 被示出为在系统 1110 内部；然而，存储器 1114 可以是内部存储器、外部存储器或者内部和外部存储器的结合。外部存储器基本上可以是任何相关存储器，例如但不限于，快闪存储器安全数字 (SD) 卡、通用串行总线 (USB) 盘或驱动器、其它存储卡、硬盘驱动器和其它这种存储器或这些存储器的组合中的一个或多个。存储器 1114 可以存储代码、软件、可执行文件、脚本、数据、内容、多媒体内容、手势、坐标信息、3D 虚拟环境坐标、程序编制、程序、媒体流、媒体文件、文本内容、标识符、日志或历史数据、用户信息等等。

[0072] 上面或下面描述的一个或多个实施例、方法、过程、方案和 / 或技术可以在一个或多个可由基于处理器的系统执行的计算机程序中实现。作为示例，这样的基于处理器的系统可以包括基于处理器的系统 1100、计算机、机顶盒、电视机、具有 IP 功能的电视、蓝盘播放器、具有 IP 功能的蓝盘播放器、DVD 播放器、娱乐系统、游戏控制台、图形工作站、写字板等。这样的计算机程序可以用于执行上面或下面描述的方法、过程和 / 或技术的各个步骤和 / 或特征。也就是说，计算机程序可以被适配为使得或者配置基于处理器的系统执行并实现上面或下面描述的功能。例如，这样的计算机程序可以用于实现上面或下面描述的用

于使得一个或多个用户能够与 3D 虚拟环境 110 交互的步骤、过程或技术的任何实施例。作为另一个例子，这样的计算机程序可以用于实现任何类型的工具或类似的应用，该工具或应用使用上面或下面描述的实施例、方法、过程、方案和 / 或技术中的任何一个或多个。在一些实施例中，可以使用计算机程序中的程序代码模块、循环、子例程等来执行上面或下面描述的方法、过程和 / 或技术。在一些实施例中，计算机程序可以存储或具体化在计算机可读存储或记录介质上，例如这里描述的任何计算机可读存储或记录介质。

[0073] 因此，一些实施例提供处理器或计算机程序产品，包括：介质，被配置为具体化要输入到处理器或计算机的计算机程序；计算机程序，具体化在该介质中，被配置为使得处理器或计算机执行或履行包括这里描述的实施例、方法、过程、方案和 / 或技术中的任何一个或多个中涉及的任何一个或多个步骤的步骤。例如，一些实施例提供一个或多个计算机可读存储介质，存储一个或多个计算机程序，以用于计算机模拟，该一个或多个计算机程序被配置为使计算机和 / 或基于处理器的系统执行如下步骤，包括：在显示三维演示的同时，接收由第一相机捕获的第一图像序列，其中第一相机安装在用户佩戴的框架上，使得当用户佩戴该框架时，第一相机的视场在该用户的视场中；从安装在该框架上的检测器接收检测区域内的一个或多个对象的检测数据，该检测区域对应于用户适当佩戴该框架时该用户的视线；处理第一图像序列；处理该检测器检测的检测数据；通过对第一图像序列的处理，检测预定义的非传感器对象和该非传感器对象的预定义的手势；通过对第一图像序列和该检测数据的处理，识别该非传感器对象的至少部分相对于虚拟三维（3D）空间的虚拟 X、Y 和 Z 坐标，该虚拟三维（3D）空间在第一相机的视场和该检测器的检测区域中；识别对应于所检测到的手势和该非传感器对象的虚拟 3D 位置的命令；以及执行该命令。

[0074] 其它实施例提供一种或多种计算机可读存储介质，存储一个或多个计算机程序，被配置为用于计算机模拟，该一个或多个计算机程序被配置为使得计算机和 / 或基于处理器的系统执行如下步骤，包括：使得显示三维演示；在显示三维演示的同时，接收由第一相机捕获的第一图像序列，其中第一相机安装在用户佩戴的框架上，使得当用户佩戴该框架时，第一相机的视场在该用户的视场中；在显示三维演示的同时，接收由第二相机捕获的第二图像序列，其中第二相机安装在该框架上，使得当用户佩戴该框架时，第二相机的视场在该用户的视场中；处理第一图像序列和第二图像序列两者；通过对第一图像序列和第二图像序列的处理，检测预定义的非传感器对象和该非传感器对象的预定义的手势；通过对所检测到的手势的处理，确定该非传感器对象的至少部分相对于第一相机和第二相机的三维坐标；识别对应于所检测到的手势和该非传感器对象的三维位置的命令；以及执行该命令。

[0075] 因此，用户 112 可以与虚拟环境 110 交互，以基于所检测到的用户的手 130 或其它预定义的对象相对于该虚拟环境的位置和所检测到的手势，执行各种功能。这可以使得用户基本上能够通过虚拟环境执行任何功能，包括执行远离用户的任务。例如，用户可以通过用户的手的运动（例如，通过伸出手拾起虚拟显示的对象）操纵机械臂（例如军队或防爆小组情形、制造情形等），使得机器人采取适当的行动（例如机器人实际拾起真实的对象）。在一些情况下，例如，由于所控制的设备的能力（例如，机器人可能只有两个“手指”），对用户可用的行动可能个有限的。然而，在其它情况下，该处理知道机器人的配置和 / 或几何结构，并且可以从所检测到的用户的手 130 的运动进行外推，以识别该机器人可以执行的相

关运动（例如，因为机器人的能力、几何结构而导致的可能命令的限制）。

[0076] 运载工具和 / 或飞机也可以通过用户的虚拟交互以虚拟控制进行控制。这使得对运载工具或飞机的控制能够立即升级，这是因为控制是虚拟的。类似地，可以基于演示和 / 或提供给操作者的其它信息远离运载工具或飞机来执行控制。类似地，虚拟交互可以用于医疗应用。例如，图像可以叠加在患者身上，和 / 或机器人应用可以用于采取行动（例如，在稳定的情况下，必须采取防抖操作）。

[0077] 另外，一些实施例可以用于教育，例如提供远程教育体验。学生不必和老师在同一个房间中，但是所有的学生看到相同的事物，并且远处的学生可以虚拟地写在黑板上。类似地，用户可以与书本（例如课本）虚拟地交互。可以提供附加的控制（例如，显示曲线，同时使得用户能够操纵参数以看到这将如何影响曲线）利用相机 124-125 或护目镜 114 上的其它相机，可以识别课本，和 / 或正在观看课本的哪一页。虚拟环境可以提供高亮文本，使得用户能够给文本加高亮，创建大纲，虚拟地注释课本和 / 或其它行动，同时存储注释和 / 或标记。

[0078] 本说明书中描述的许多功能单元已被标记为系统、设备或模块，以便更具体地强调她们的实现独立性。例如，系统可以实现为硬件电路，包括定制 VLSI 电路或门阵列、现成半导体如逻辑芯片、晶体管或其它分立组件。系统还可以在可编程硬件设备中实现，例如现场可编程门阵列、可编程阵列逻辑、可编程逻辑器件等。

[0079] 系统、设备、或模块还可以以用于由各种类型的处理器执行的软件来实现。例如，可执行代码的标识系统可以包括一个或多个例如可以组织为对象、过程或函数的计算机指令的物理或逻辑块。然而，所标识的模块的可执行文件不需要物理上在一起，而是可以包括存储在不同位置的全异指令，当逻辑上链接在一起时，其包括该模块，并且实现该模块的规定的目的。

[0080] 事实上，可执行代码的系统可以是单条指令，或者许多指令，并且可能甚至分布在若干不同的代码段中，不同的程序中，跨着若干存储设备。类似地，操作数据可以被识别，并且在此图解为在系统中，并且可以以任何合适的形式具体化，并且任何合适类型的数据结构组织。操作数据可以聚集为单个数据集，或者可以分布不同位置，包括分布在不同存储设备上，并且至少部分地可以仅仅作为系统或网络上的电子信号而存在。

[0081] 尽管通过具体的实施例、例子及其应用描述了这里公开的发明，但是本领域技术人员可以对其做出许多修改和变型，而不脱离权利要求书中阐明的本发明的范围。

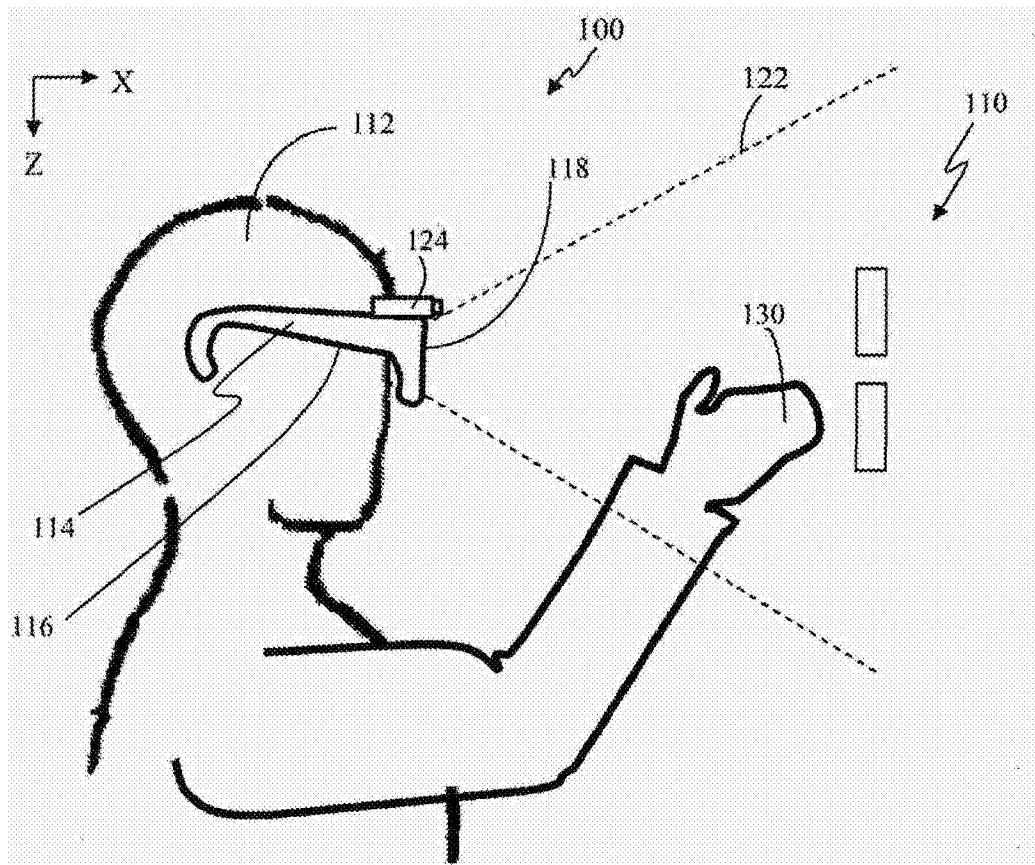


图 1

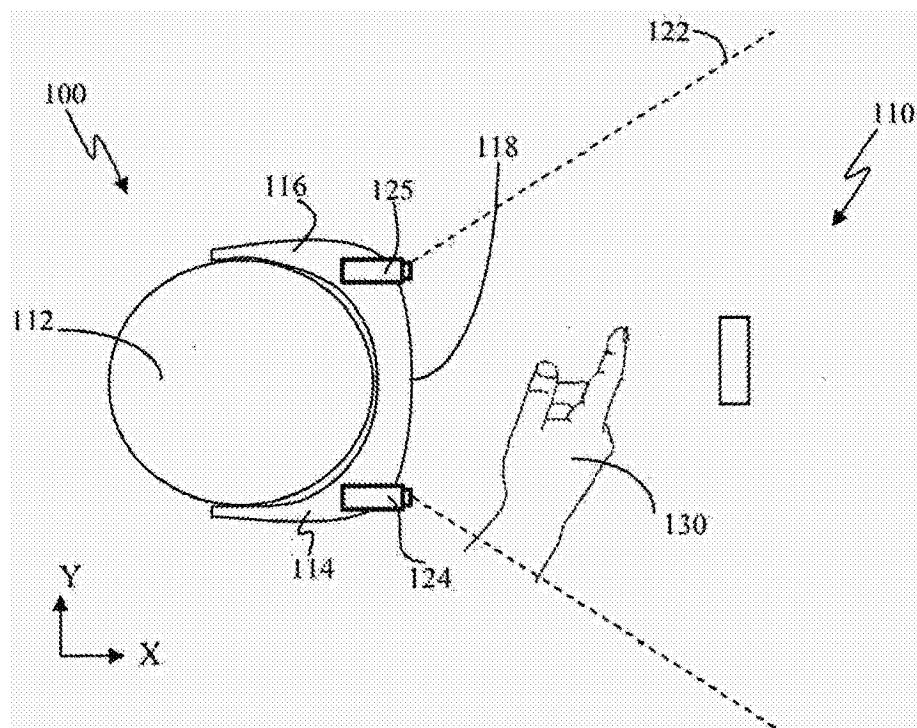


图 2

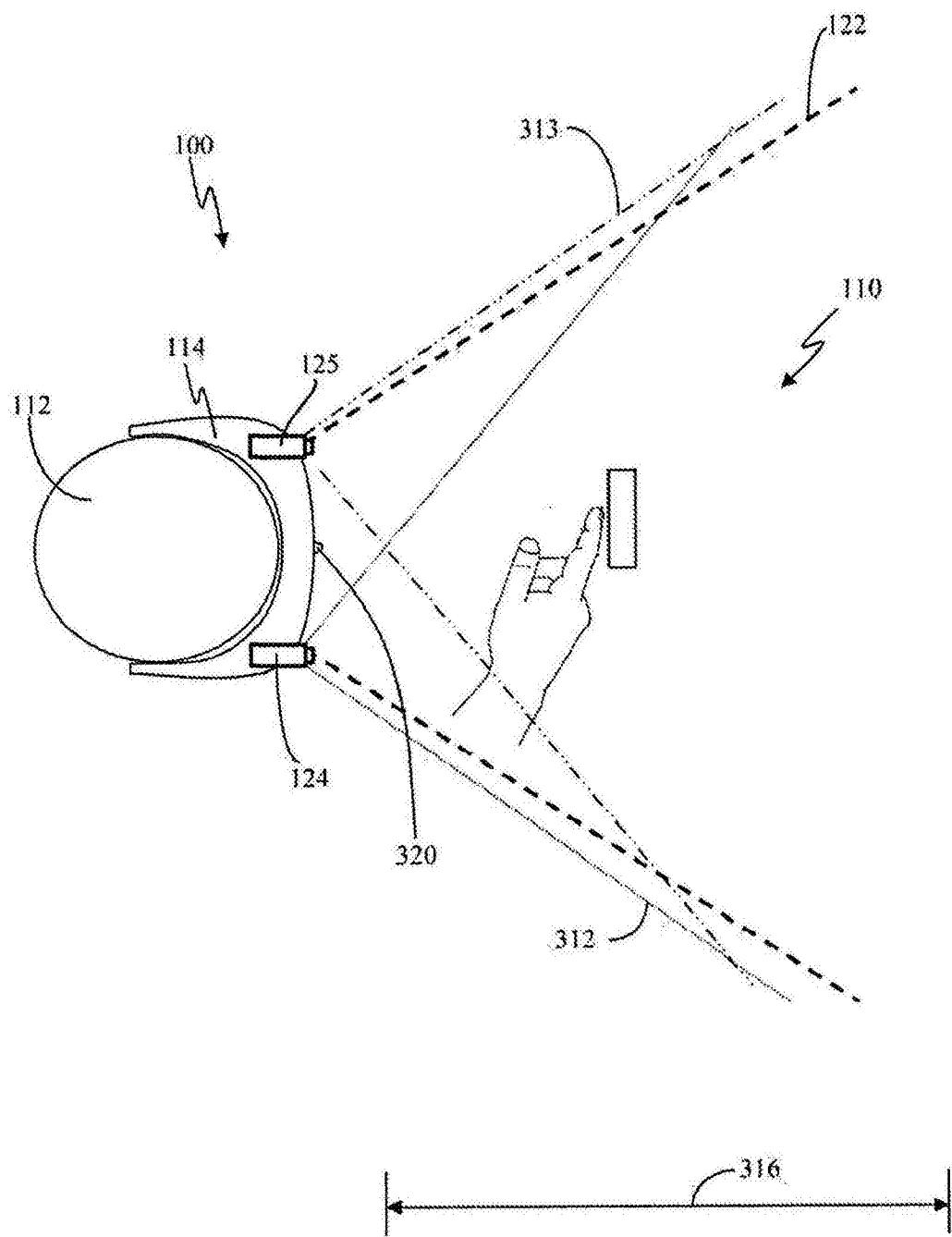


图 3

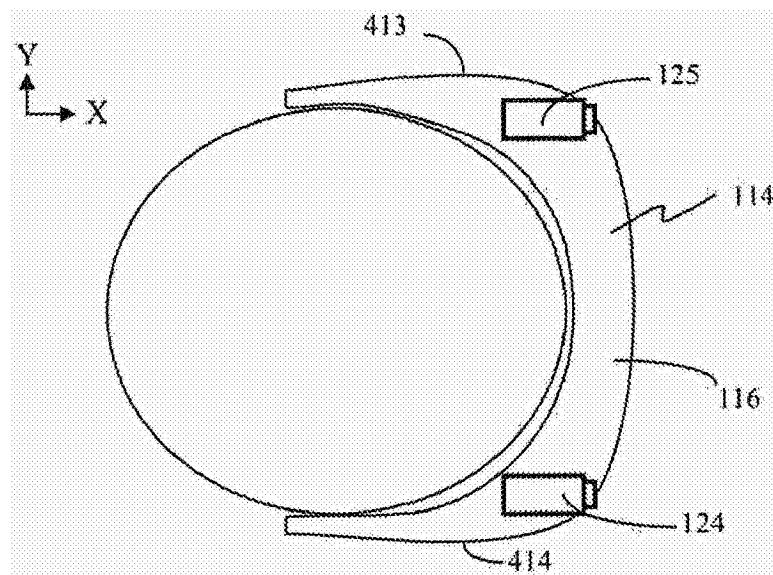


图 4A

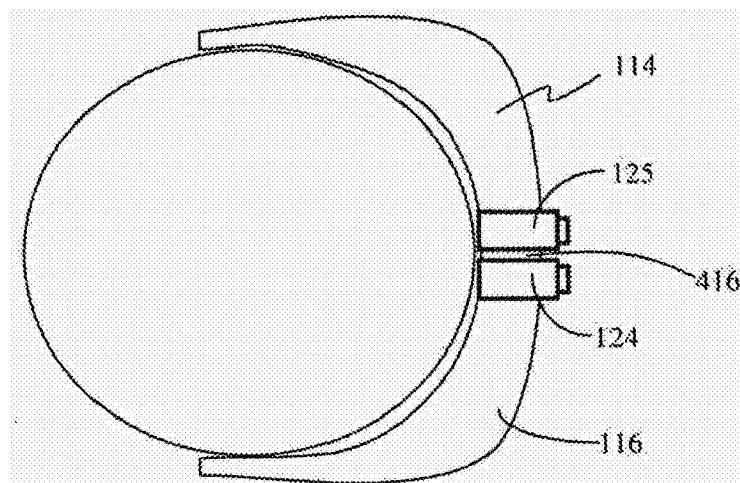


图 4B

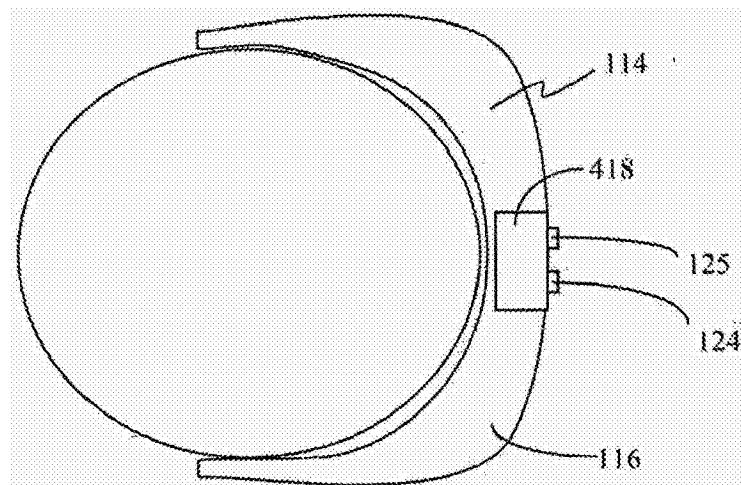


图 4C

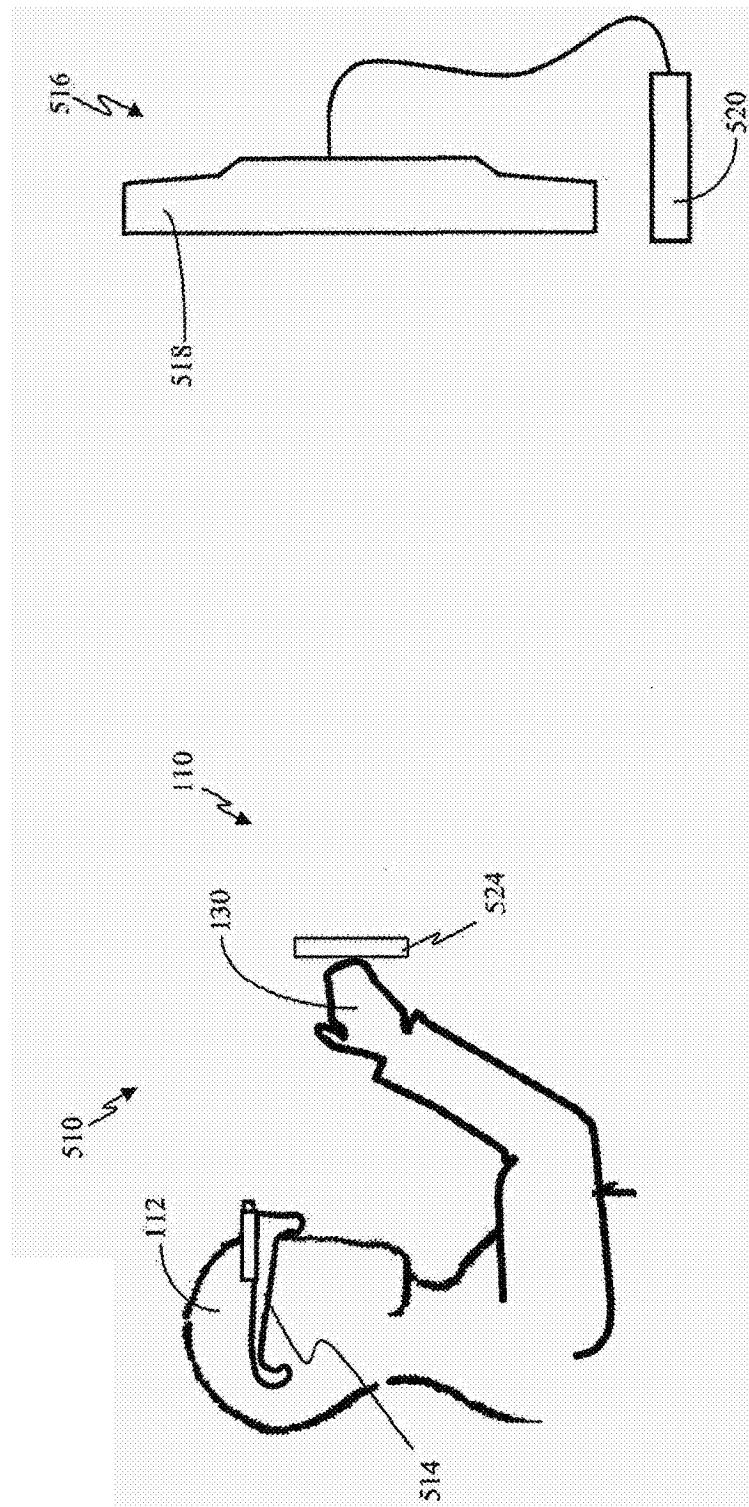


图 5A

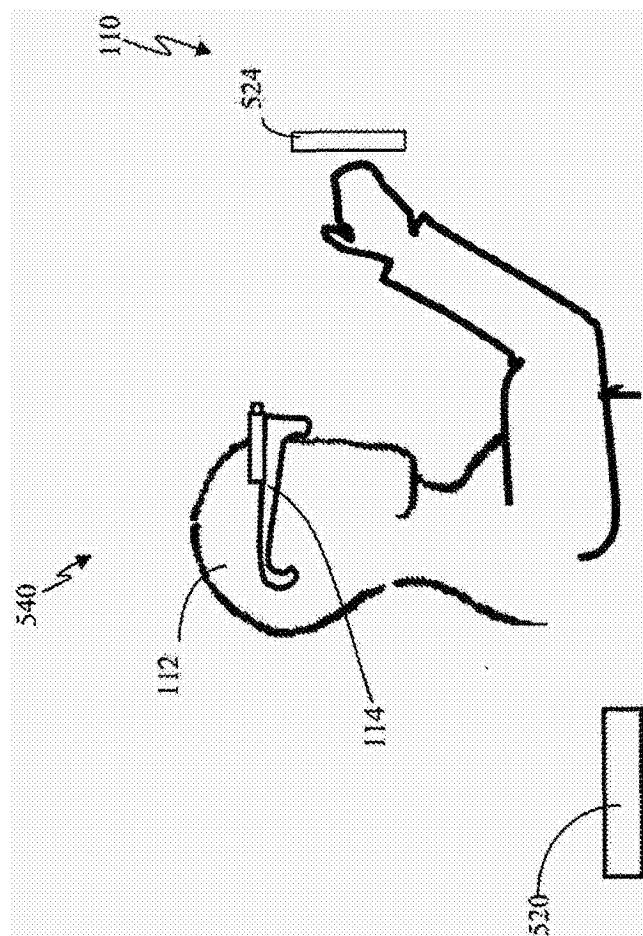


图 5B

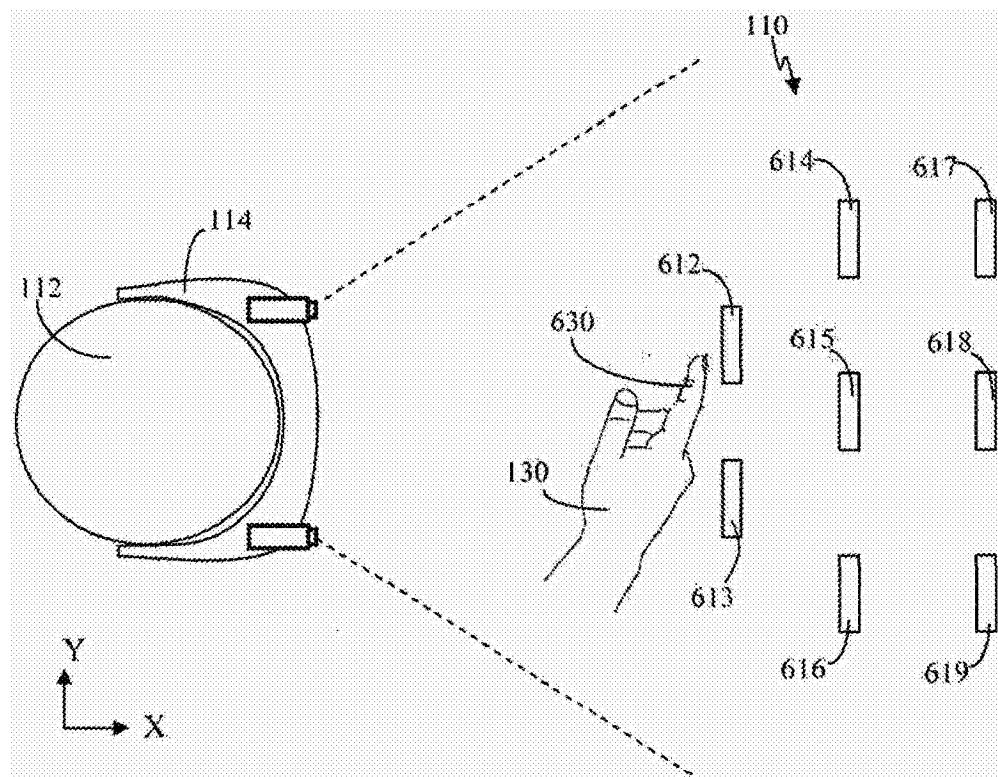


图 6A

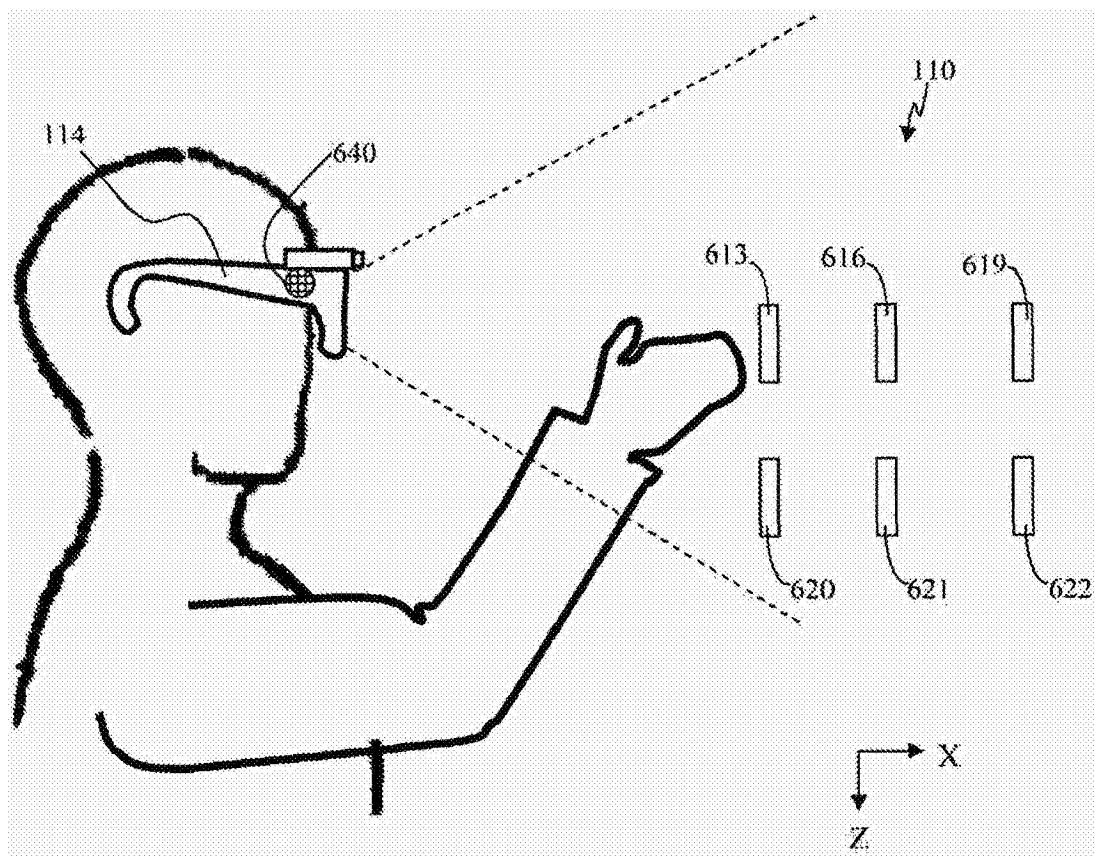


图 6B

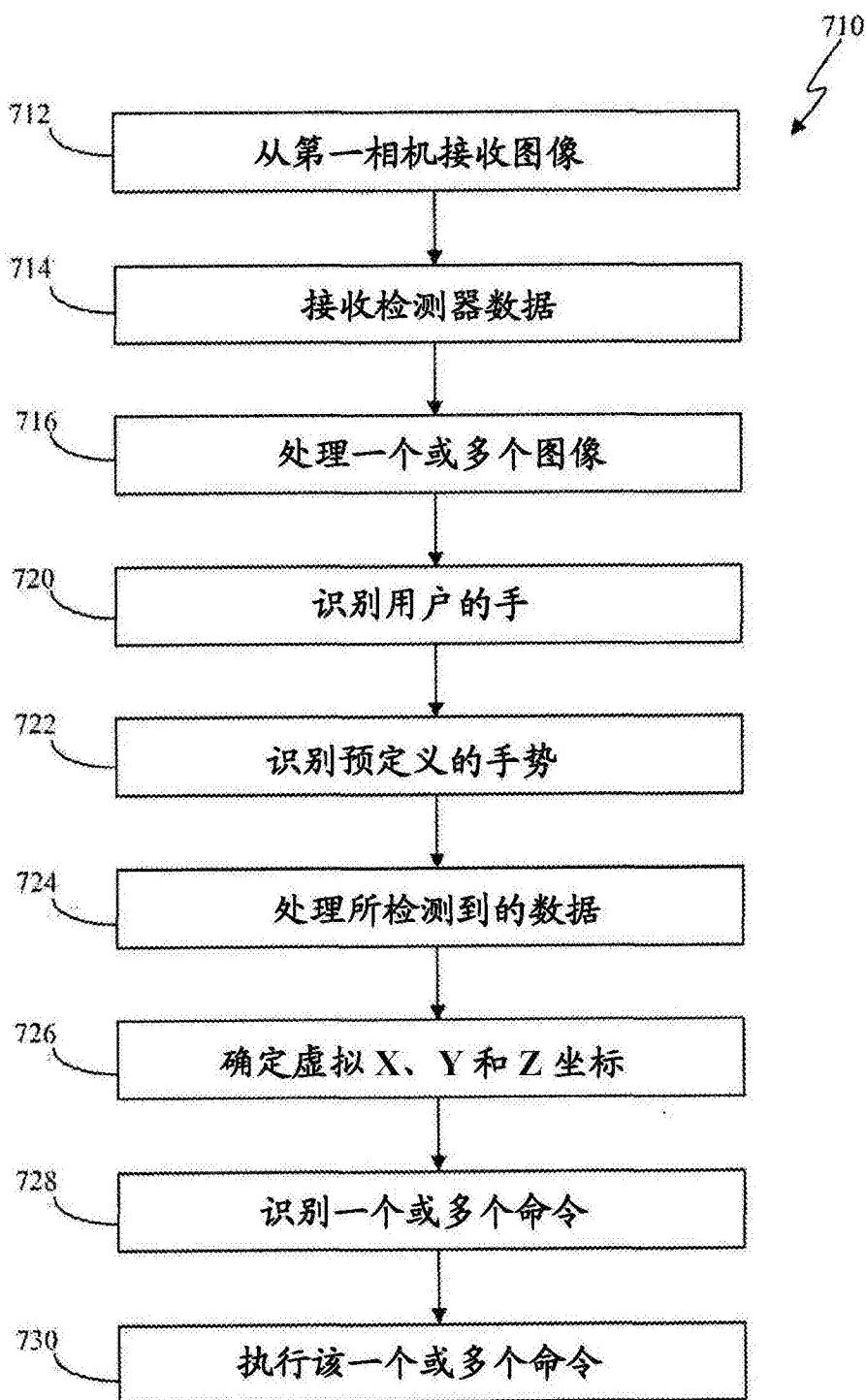


图 7

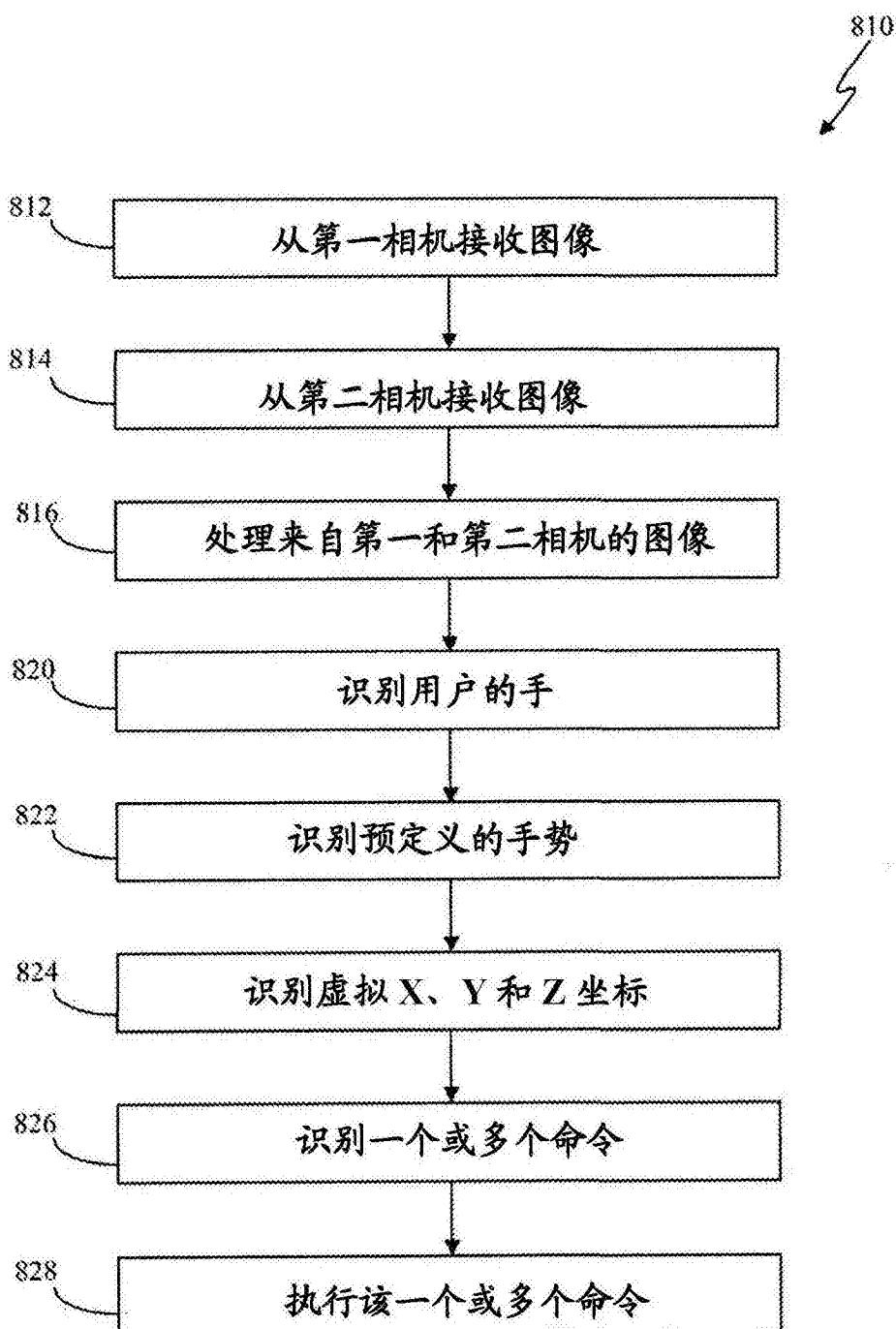


图 8

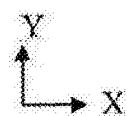
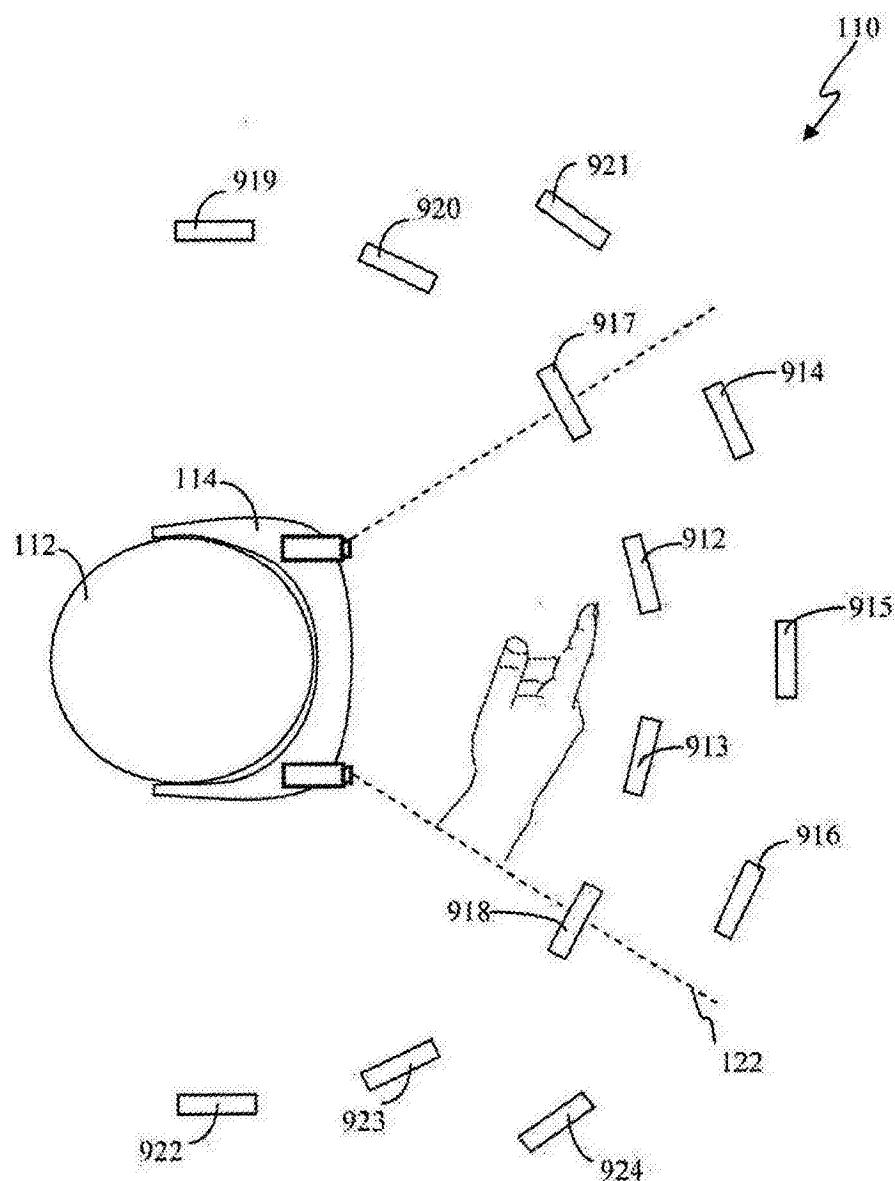


图 9

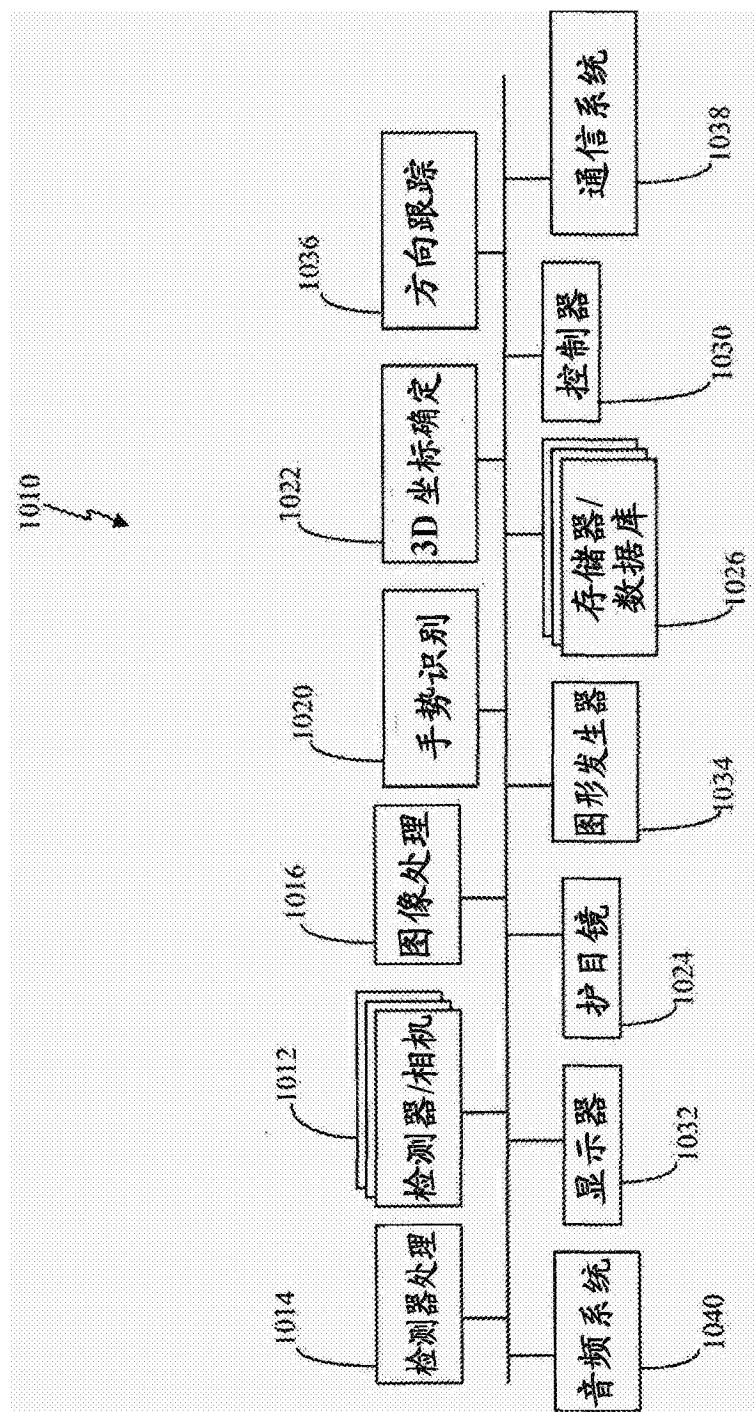


图 10

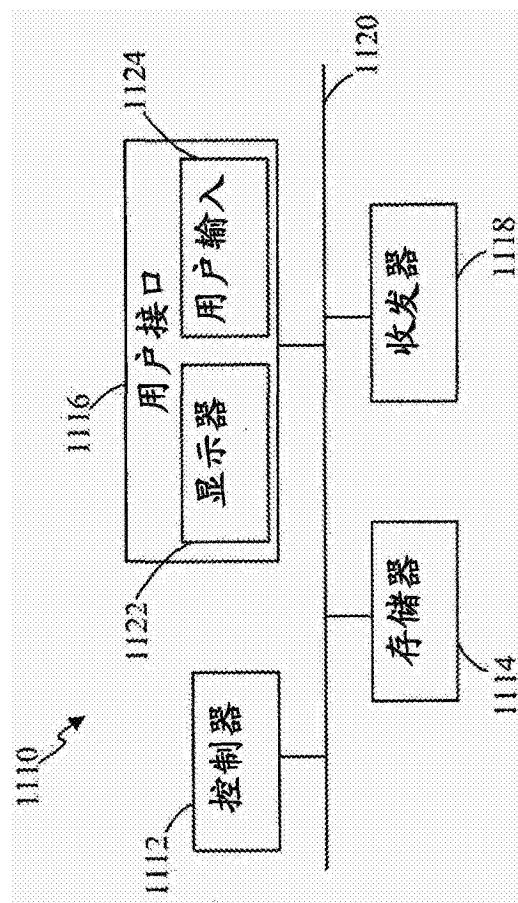


图 11