



(12) 发明专利

(10) 授权公告号 CN 101952818 B

(45) 授权公告日 2016. 05. 25

(21) 申请号 200880115391. 4

(22) 申请日 2008. 09. 15

(30) 优先权数据

60/993, 907 2007. 09. 14 US

(85) PCT国际申请进入国家阶段日

2010. 05. 10

(86) PCT国际申请的申请数据

PCT/US2008/010750 2008. 09. 15

(87) PCT国际申请的公布数据

W02009/035705 EN 2009. 03. 19

(73) 专利权人 智慧投资控股 81 有限责任公司

地址 美国内华达州

(72) 发明人 马修·贝尔 蒂帕塔特·谢诺维辛

查尔斯·H·克兰顿 迈克尔·休姆

埃亚勒·奥菲尔 马修·维耶塔

(74) 专利代理机构 中科专利商标代理有限责任

公司 11021

代理人 戎志敏

(51) Int. Cl.

G06F 17/00(2006. 01)

(56) 对比文件

US 6191773 B1, 2001. 02. 20,

CN 1289086 A, 2001. 03. 28,

US 2006132432 A1, 2006. 06. 22,

审查员 崔哲

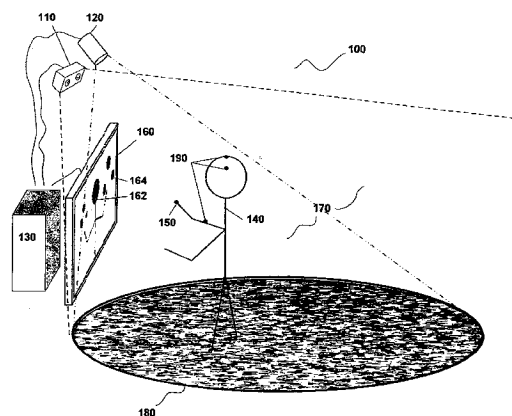
权利要求书1页 说明书12页 附图6页

(54) 发明名称

基于姿态的用户交互的处理

(57) 摘要

本发明涉及一种用来处理具有交互显示器的基于姿态的用户交互系统和方法。



1. 一种利用交互显示器处理基于姿态的用户交互的方法,所述方法包括:

从两部或多部相机接收相应于物理空间的信息,其中,所述物理空间包括多个区域,所述信息包括用于检测用户的特定部分的亮度、范围和/或颜色,所述特定部分是用户的手或身体;

至少部分地基于接收的信息检测所述多个区域中的一个或多个中的用户的所述特定部分;以及

在交互显示器上显示在所述多个区域中的所述一个或多个中检测到的用户的所述特定部分的表示,其中所述表示基于用户所能与交互显示器上显示的虚拟对象交互的形式而改变外形。

2. 根据权利要求1所述的方法,其中,所述多个区域包括“太靠近”空间、“手”空间和“远端用户”空间,其中:

当用户处于“远端用户”空间中时,用户仅能够以其整个身体进行交互;

当用户处于“手”空间中时,用户能够以手和身体进行交互;以及

当用户将手放入“太靠近”空间中时,用户不能够以手或身体进行交互。

3. 根据权利要求2所述的方法,其中,当用户的所述特定部分进入虚拟对象周围的触摸空间时,用户的所述特定部分能与所述虚拟对象交互。

4. 根据权利要求3所述的方法,其中,所述触摸空间捕获手指点击交互。

5. 根据权利要求2所述的方法,所述表示是图标。

6. 根据权利要求1所述的方法,其中,基于沿与所述交互显示器垂直的方向从手到用户躯干的距离来确定手的表示。

7. 根据权利要求1所述的方法,其中,使用用户的所述特定部分的数字轮廓来表示所述表示。

8. 一种利用交互显示器处理基于姿态的用户交互的系统,所述系统包括:

用于从两部或多部相机接收相应于物理空间的信息的装置,其中,所述物理空间包括多个区域,所述信息包括用于检测用户的特定部分的亮度、范围和/或颜色,所述特定部分是用户的手或身体;

用于至少部分地基于接收的信息检测所述多个区域中的一个或多个中的用户的所述特定部分的装置;以及

用于在交互显示器上显示在所述多个区域中的所述一个或多个中检测到的用户的所述特定部分的表示的装置,其中所述表示基于用户所能与交互显示器上显示的虚拟对象交互的形式而改变外形。

9. 根据权利要求8所述的系统,其中,所述多个区域包括“太靠近”空间、“手”空间和“远端用户”空间,其中:

当用户处于“远端用户”空间中时,用户仅能够以其整个身体进行交互;

当用户处于“手”空间中时,用户能够以手和身体进行交互;以及

当用户将手放入“太靠近”空间中时,用户不能够以手或身体进行交互。

10. 根据权利要求8所述的系统,其中,所述用户能够通过所述表示来与所述虚拟对象进行交互。

## 基于姿态的用户交互的处理

[0001] 相关申请的交叉参考

[0002] 本申请要求2007年9月14日提交的美国临时申请60/993,907的权益,其公开文本合并在此作为参考。

### 技术领域

[0003] 本发明总地涉及交互显示系统。本发明涉及在三维环境中用交互显示器来处理基于姿态的用户交互的用户界面。

### 背景技术

[0004] 很多已有的交互技术都依赖二维相机。由这些相机所产生的二维数据很难可靠地解析以达到从背景中分离出前景的目的。结果,就难于分离出用户的手或其他肢体并追踪其相应的动作。

[0005] 对于给交互显示系统提供数据来说,不佳的分离和追踪数据会产生各种困难。如果交互显示系统不能收集数据来控制或操纵通过其所提交的信息,那么用户就可能被强制来依靠例如键盘、操纵杆和制动器的物理对象和执行器件的按钮按压或操纵了。一旦用户被强制依靠物理仪器来提供输入,则就会使得交互显示系统所提供的很多优点尽失。

[0006] 这就需要对交互显示系统进行基于姿态的用户交互的处理。

### 发明内容

[0007] 在第一实施例中,公开了一种用来处理具有交互显示器的基于姿态的用户交互的方法。通过该方法,可以收集在所定义的三维空间中的对象的三维信息。识别作为相应于在所定义的三维空间中的用户手的一部分三维信息。数字对象屏幕上(on-screen)的图像被显示为与用户的物理外形关联并与其近似的屏幕上的图像。所显示的屏幕上的图像近似于用户但并不与所显示的数字对象交互。显示在所定义的三维空间中用户的手所相应的图标。该图标能与所显示的数字对象交互。

[0008] 在实施例中还公开了第二种用于处理具有交互显示的基于姿态的用户交互的方法。在这个实施例中,收集在所定义的三维空间中有关用户的三维信息。所定义的三维为空间包括体量区域(volumetric zone)。由该三维信息所限定的每个用户的一部分位于一个或多个体量区域内。一部分三维信息相应于在所定义的三维空间中的用户之一的手。所显示的屏幕上的数字对象是当前用户之一是否能与该显示的数字交互的指示。该指示相应于在一个或多个体量区域中的用户位置。

[0009] 第三实施例是另一种用于处理具有交互显示的基于姿态的用户交互的方法。收集在所定义的三维空间中有关对象的三维信息并识别一部分三维信息来相应于在所定义的三维空间中用户的手。显示数字对象的屏幕上的图像并且其具有在视觉上可见的动作状态,该动作状态相应于用户是否已经超出了该数字对象的动作级别阈值。显示用来与该数字对象的屏幕上的图像交互的图标,该图标相应于在所定义的三维空间中的用户的手。该

动作级别相应于用户的手随时间而经过的位置。

[0010] 在另外的实施例中,公开了另一种用于处理具有交互显示的基于姿态的用户交互的方法。收集在所定义的三维空间中有关对象的三维信息并识别一部分三维信息来相应于在所定义的三维空间中的用户的手。显示数字对象的屏幕上的图像,其具有动作级别的阈值。显示用来与该数字对象的屏幕上的图像交互的图标,该图标相应于在所定义的三维空间中的用户的手。仅当所显示的数字对象的动作级别阈值已被超过时,该图标才可以控制与该数字对象关联的行为。作为该图标随时间接近该数字对象所关联区域的结果,数字对象的动作级别阈值被超过。

### 附图说明

[0011] 图1图示了一个示例性的平板交互系统。

[0012] 图2图示了一个包括体量区域的交互显示空间。

[0013] 图3图示了一个包括体量区域的交互显示空间的另一个实施例。

[0014] 图4图示了一个用来处理基于姿态的交互的示例性方法。

[0015] 图5图示了手/用户表示的vis-à-vis选择级别。

[0016] 图6图示了用户的物理位置如何被映射到数字对象的显示虚拟世界。

[0017] 图7图示了如何将用户的物理位置映射到数字对象的所显示的虚拟世界中的另一个实施例。

### 具体实施方式

[0018] 图1图示了示例性的平板交互系统100。系统100包括显示器160、三维影像系统110、计算系统130和可选的照明器120。系统100可以根据构成因素而配置。

[0019] 显示器160可以假定包括了能显示视频图像的显示器的任意变化的配置。在一个实施例中,例如,显示器160为液晶显示器(LCD)、等离子显示屏或有机发光二极管显示屏(OLED)等的平板显示器。显示器可以替换性地包括但不限于,阴极射线管(CRT)、电子墨水显示屏、背投显示器、正投显示器、例如NEC售卖的WT600投影仪的离轴(off-axis)正(或背)投影仪、例如透镜三维视频屏幕等能产生三维图像的屏幕、或例如源自IO2Technology USA公司的Heliodyisplay™屏的雾屏。可以平铺多个屏幕而形成一个巨大的显示器。系统100可以包括多个交互显示器160,从而能让多个用户同时与系统100交互。

[0020] 在一种配置(例如图1的配置)中,系统100的各个组件可以被安装在一面墙上。系统100的组件替换性地可以被安置在天花板、独立支架上,或安装在桌面或两侧上。独立支架系统可以包括一个大基座,或者也可以为了稳定而固定在地面上。

[0021] 集成了三维对象的系统100的实施例可以使用三维硬件和软件系统,其中例如显示器160可以是圆柱或球形。例如地面、桌面等这些或其他形成因素的配置或墙面投影系统同样地也可以使用三维影像系统。在这样的实施例中,三维影像系统110可以覆盖显示器周围的区域。投影仪可以用来形成显示器,例如地面投影,其中安装在天花板上的投影仪向地面投射图像,允许用户140可以在屏幕区域上行走。投影还可以在桌面、前面或其他表面投射。可以使用多个投影仪和多个三维影像系统来形成地面、前面或其他表面上的交互的扩展区域。

[0022] 系统100的实施例可以改变尺寸。例如,三维影像系统110可以被附着在膝上型计算机监视器的顶部从而可以让用户携带。膝上型电脑本身也可以当计算机系统130来操作并能处理三维影像软件以及产生视觉效果。

[0023] 如上所述,系统100包括照明器120。在图1中三维影像系统110和照明器120的位置如图所示。三维影像系统100和照明器120可以安装在相对于显示器160和/或其他组件的任意位置。此外,三维影像系统110和照明器120还可以有多种实例。三维影像系统110和/或照明器120可以使用一个或多个镜面或其他反射表面(未示出)来改变交互区域170的位置。

[0024] 系统100可以使用麦克风或麦克风阵列来捕捉音频输入并且也可以使用扬声器来产生音频输出。麦克风可以是定向的,以便更好地从没有背景噪声的特定用户捕捉音频。扬声器也可以是定向的,以便将音频聚焦到特定的用户和特定的区域。定向扬声器的实例除了由Dakota Audio,Holosonics和圣迭戈的American Technology Corporation(ATCSD)所制造的系统外,还包括来自Brown Innovation的Maestro™和SoloSphere™。

[0025] 三维影像系统110可以使用立体影像来实现。立体影像系统使用从两部或多部相机来的信息建立三维图像。使用了例如基于Marro-Poggio算法等的算法的三维计算机影像技术可以将从略微不同的角度对相同场景所拍摄的两个或更多的图像作为输入。

[0026] 基于Marro-Poggio的算法仅仅是立体算法的一个实例,因为他们发现了相应于相同物理对象的相同部分的不同相机图像的纹理块。在图像中块的位置之间的不同可以确定从相机到该块的距离,从而为该块提供了三维位置数据。

[0027] 但当处理单色对象时,由于单色很难在不同的图像中匹配相应的块,因此该算法的性能趋于退化。因此,如果照明器120产生纹理化的光,则不同的单色纹理就能改善一些三维计算机影像算法的距离估算。

[0028] 可以用在三维影像系统110中的示例性的立体相机包括Tyzx DeepSea™和Point Grey Bumblebee™。这些相机可以是单色的或彩色的,并且可以对包括可见光、近红外线、远红外线以及紫外线等电磁光谱的一个或多个特定频段敏感。

[0029] 三维影像系统110还可以使用时差测距相机来实现。时差测距相机为图像的每个像素检测距离信息。每个像素的距离信息可以包括从相机到从该像素所见对象的距离。时差测距相机通过计时可见光脉冲从邻近相机的光源传播到要被成像的对象接着再返回到相机所花费的时间来获得距离数据。光脉冲可以迅速地反复从而使得时差测距相机能有类似于摄像机帧率的帧率。虽然时差测距相机通常在每秒30帧时具有1-2米的范围,但这个范围可以通过延长曝光时间而增加,同时可以降低帧频。时差测距相机的制造商包括加州Sunnyvale的Canesta Inc.和以色列的3DV System。

[0030] 三维影像系统110的实施例还可以使用激光测距仪、安装有成对的结构化投光器的相机、激光扫描仪、激光行扫描仪、超声波成像器、能通过从几部相机的前景图像的交叠而得到三维数据的系统,上述的任意组合来实现。多个三维影像系统110可以同时使用,以使它们的三维数据合并在一起,形成描述交互区域内对象的统一数据集合。

[0031] 虽然在系统100的一些实施例中背景照明可能就足够了,但也必须使用照明器120以通过处理设备(例如,在三维影像系统110或在计算设备130中)改进通过执行立体算法而要处理的纹理数据。通过光的图案照射在交互区域中的对象,照明器120提高了使用立体算法来匹配块的纹理数据量。照明器120可以产生图案化和/或不均匀的光图案以提高使用

立体算法的对象的检测。在一些实例中,照明器120可以简单地提供与显示器160交互的用户的额外照明。

[0032] 照明器120可以包括集中光源,例如高功率发光二极管(LED)、白炽灯泡、卤素灯泡、金属卤化物灯泡以及弧光灯。可以同时使用多个集中光源,或使用相同的散射系统作为第一光源或使用单独的光源。应当注意的是,基本上校准的光源(例如具有抛物面镜或一个或多个深照型LED的灯)可以用来代替集中光源而取得类似的结果。

[0033] 照明器120可以使用各种图案,例如随机点的图案、在变化的长度比例上提供噪声的碎片噪声图案、以及包括一组平行线的图案。平行线可以是随机的并能根据其距离而改变以提供有价值的纹理。可以使用各种方法来产生这样的图案。例如,照明器120可以包括设计来投射通过视频输入电缆提供的任意图像的视频投影仪。图像可以随时间而改变以优化三维影像系统110的性能。例如,图案可以在三维影像系统110的相机图像由于光线而变得饱和的区域中暗淡,而在对象临近相机的区域内变得具有较高的分辨率。可以使用多个视频投影仪来覆盖整个交互区域。

[0034] 照明器120还可以使用结构化投光器来投射静态或随时间而变化的图案。结构化投光器的实例包括来自ABW的LCD-640<sup>TM</sup>和MiniRot-h1<sup>TM</sup>。在另一实施例中,图案可以由幻灯机产生,其包括光源、具有图像的透明幻灯片以及镜头系统或弧光镜系统来投射幻灯片的图像。可以使用多个幻灯机来覆盖整个交互区域。在另一实施例中,由高度集中的光源(从很小的区域发出的光)产生图案。替换地,可以使用反射或折射系统。

[0035] 照明器120可以使用例如近红外光源等的可见或最低程度的可见光源以避免在用户所注意的地面上产生图案。替换地,照明器120可以具有能发出可见以及不可见光的光源并可以包括滤波器来降低可见光,而保留不可见光。照明器120和相机在同步相机曝光时可以通过照明器120的脉冲调制来选通从而降低或消除环境光的影响。三维影像系统110可以包含光学滤波器,其可以抑制并非由照明器120所发出的波长的光,从而使得从照明器120发出的光在由影像系统110所捕获的图像中更显著地表现出来。

[0036] 影像软件可以在三维影像系统110或与影像系统110耦合的计算设备130中执行。影像软件的执行可以提高捕获的图像数据的有效性。例如,当使用多个三维相机时,影像软件可以辅助相机的校准以便从相机来的数据能被整合到相同的坐标空间(即,两部相机可以从不同的角度覆盖相似的地区以从多个角度提供该区域的覆盖)。这些相机可以覆盖不同的区域,其中一部相机可以集中于临近的交互而另一部相机则集中于远距离的交互。立体处理的质量和分辨率还可以通过影像软件的执行而改变和完成。例如,靠近显示器160的区域可以以较高的分辨率处理以便以更高的精确度来分辨临近显示器160的用户140的各个手指150。

[0037] 可以使用多种方法来消除在三维数据中的错误和噪声。例如,可以使用背景方法来将在特定时间周期内从已知没有移动的相机视场的区域来的三维数据屏蔽。这些背景方法(以及所知的背景减少方法)可以是自适应性的,允许背景方法随时间调整在背景中的变化。这些背景方法可以使用从相机来的亮度、色度和/或距离数据来形成背景并确定前景。一旦确定了前景,则就可以消除从前景区域外收集的三维数据。

[0038] 在另一实施例中,可以使用彩色相机来为用户140以及在显示器160前的其他对象的三维数据获得色度数据。可以使用该色度数据来得到用户140的彩色三维的表示,这就能

识别并跟踪它们的相似度。可以连续地显示所识别并跟踪的用户的相似度。

[0039] 可以对作为从相机到从相机的的视点来的相机图像的每个像素的距离的深度图像使用噪声过滤,或直接对由影像系统110收集的三维数据使用噪声过滤。可以将例如中值过滤等的平滑和平均技术应用到相机深度图像中来降低深度的错误。如果三维数据集并不相应于较大的形状,则就从它们中去除单独的点或点的小簇,从而在保留用户原样的同时消除噪声。

[0040] 可以以各种方式来分析三维数据以产生高级信息。例如,可以检测用户的指尖、手指和手(在图1中用组件150集中表示)。能做到这样的方法包括各种形状识别和对象识别算法。可以使用二维和/或三维的空间、时间、色度或亮度数据的任意组合来对对象分割。此外,在上述域的各种线性或非线性变换中也可以对对象分割。对象检测算法的实例包括,但不限于,可变形的模板匹配、Hough变换以及在适合变换的空间中的空间临近像素和/或体素的聚合。

[0041] 作为另一个实施例,可以对从属于用户140的三维点分簇并标记以便能识别出从属于该用户的点的簇。各个身体部分,例如用户140的手和手臂(190)可以被分割来作为记号。点还可以在三维空间中使用例如k-means或分级成簇等无管理的方法分簇。所识别的簇接下来可以进入特征抽取和分类引擎。特征抽取和分类并不限于三维空间数据还可以被应用到例如二维空间、亮度、色度或变换本身等其他数据域中的在先特征抽取或分类。

[0042] 通过各种方法可以将骨骼模型映射到从属于给定用户140的三维点,这些方法包括但不限于期望值最大化、梯度下降、粒子过滤以及特征跟踪。此外,例如eigenface或fisherface等的面部识别算法可以使用从影像系统来的例如二维和/或三维空间、时间、色度和亮度数据等的的数据来识别用户以及它们的面部表情。所使用的面部识别算法可以是基于图像或基于视频。该信息可以被使用来识别用户,特别是在它们离开再返回到交互区域以及基于面部、性别、身份、种族、面部表情或其他特征而改变与显示的内容交互的情况。

[0043] 指尖或身体的其他部分可以随时间而被追踪以便识别特定的姿态,例如推、抓、落、戳,和/或描绘使用手指的形状、捏和其他这类的运动。

[0044] 三维影像系统110可以被特殊配置来检测除用户140外的特定对象。这种检测可以使用可执行的对象识别算法以识别对象的外形或形状特定方面,由RFID读取器(未示出)读取的对象中的射频识别(RFID)标签以提供识别信息,和/或对象上的特定图案中所选通的光源以提供识别信息。

[0045] 用户140或由三维影像系统110所检测的其他对象可以在显示器160中表示。可以使用在显示器160上用户140的表示162来让用户与显示器160上显示的虚拟的、数字对象164(其可以与虚拟或数字对象相互交换)交互。通过给定用户的位置(162)相对于数字对象164的可视化指示能使得交互很方便。

[0046] 这种表示可以包括用户或其他对象的数字阴影。例如,二维形状表示相应于用户身体的三维数据投射在平面上。二维形状可以与用户的形状或尺寸接近。依靠本发明的特定实施例,用户(或其一部分)的表示可以或不必要与在显示器160上的数字对象164交互。

[0047] 还可以使用用户或其他对象的数字轮廓来表示在显示器160上与数字对象164交互的用户162。该数字轮廓还可以被视为数字阴影的边界。类似于数字阴影,该轮廓可以与用户的形状和大小近似。这个形状可以加色、加高亮、进行渲染,或在显示前进行任意的处

理。

[0048] 表示用户手(150)或身体其他部分(190)或其他对象的各种其他的图像、图标、或二维或三维渲染可以在显示器160上进行渲染以达到与数字对象164交互的目的。例如,用户的手150可以在显示器160上表示为用户实际手的近似或作为相应于手的“类手”的图标。

[0049] 其他的交互和表示方法包括:

[0050] 当用户的手处在能与屏幕上的对象交互的位置上时,在虚拟空间中渲染的用户的外形可以与所显示的用户的手上的标记结合在一起。例如,手上的标记可以仅仅在手在屏幕上点触时才显示出来。

[0051] 表示从影像系统100的三维数据的点云(point cloud)来的用户或其他对象的点可以显示为对象,其可以是很小和/或半透明的。

[0052] 可以使用表示用户手指(150)位置的光标。当用户140的手指150能在虚拟空间中进行特定类型的交互时,这些光标可以被显示出来或改变外形。

[0053] 可以使用沿着用户身体的各个部分移动和/或附着与用户身体各个部分的对象。例如,用户140可以戴一顶头盔,其可以随着用户头170的移动和旋转而移动和旋转。

[0054] 数字化身用以匹配移动时用户140的身体位置或其他对象。在一个实施例中,数字化身被映射到用户位置的骨骼模型。

[0055] 上述的表示可以随意组合。

[0056] 在一些实施例中,表示可以基于用户所能与屏幕上的对象交互的形式而改变外形。例如,用户可以被显示为灰色的阴影且不能那个与对象交互直到它们进入到显示器160的特定距离内,在此点它们的阴影会改变颜色并且它们能开始与屏幕上的对象交互。

[0057] 给定可以由三维影像系统110抽取并收集的大量潜在特征和在屏幕上能显示的各种虚拟对象164,并且在用户140和虚拟对象164之间存在大量潜在交互。潜在交互的一些实例包括能延伸到第三维的基于作用力的二维交互和基于影响图像的交互。因此,有关用户位置的三维数据能被用来产生三维影响图像以影响三维对象的移动。在二维和三维中的这些交互使得在虚拟对象164上用户所施加的作用力的强度和方向能被计算出来,从而使得用户控制他们如何撞击显示器160上对象的移动。

[0058] 用户140(或他们所显示的表示162)可以通过与在虚拟空间中的他们交叠而与数字对象164交互。这样的交叠可以在从用户来的三维或三维数据中计算并可以被投影到二维中并被计算为二维交叠。

[0059] 可以基于从用户来的三维数据产生视觉效果。例如,可以使用三维位置数据或其一部分产生发光、扭曲、粒子发射、火焰轨迹,或其他视觉效果。视觉效果还可以基于特定的身体部位的位置。例如,用户可以通过将他们的手合拢而产生虚拟火球。用户140可以使用特定的手持来拿起、放下、移动、旋转或修改显示在显示器上的虚拟的、数字对象164。

[0060] 在显示器160上描述的虚拟空间可以以二维或三维显示。无论哪一种情况,系统100将有关用户140的信息与在显示器160的虚拟空间中显示的数字对象164和其他图像的信息(例如,用户表示162)合并。如果在虚拟空间中用二维描述用户140,那么有关用户位置的三维数据可以被投影在二维平面上。

[0061] 在显示器160前的物理空间180和在显示器160上显示的虚拟空间之间的映射可以



被任意地定义甚至能随时间而改变。由用户140在显示器160上所看见的实际场景可以基于物理空间180的定义而变化。类似地,虚拟空间(或仅仅是用户的表示162)可以是二维的。在这种情况下,用户的虚拟表示162的深度分量可以被忽略。

[0062] 映射可以以类似于镜像的方式运作,以便用户140在虚拟空间中所看见的用户的表示162的动作类似于用户140实际动作的镜像。可以对该映射进行校准以便当用户140触摸或将他们的一部分身体靠近屏幕显示器160时,他们的虚拟表示162触摸或将其身体的相同部分靠近屏幕160的相同部分。该映射还可以显示出当用户140接近屏幕160时用户的表示162从屏幕162的表面后退。

[0063] 通过将影像系统前的空间分成多个体量空间并将所收集到的数据分类在这些空间中可以创建多个交互区域。图2显示了包括了体量区域的交互显示空间。显示空间包括影像系统220和显示器210。

[0064] 在图2中,“触摸”空间250主要捕获手指点击的交互,“主要用户”空间240捕获最临近屏幕的人的身体,而“远端用户”空间230捕获最远处的人的身体。在屏幕上的用户表示取决于他们所在的区域而有所不同。例如,在“触摸”空间250中的手指可以由光标表示,在“主要用户”空间240中的身体260可以由彩色的轮廓表示,而在“远端用户”区域230中的身体270可以由灰色轮廓表示。空间的边界可以变化。例如,如果在“主要用户”空间250中没有人,则该空间的边界就可以拓展覆盖“远端用户”空间270。

[0065] 距离屏幕超出一定距离或在“远端用户”空间270中的用户可以有能力与在屏幕上缩小或消除的对象交互。这使得临近屏幕的用户能与虚拟对象交互而无需受到较远距离用户的干扰。在一些实例中,用户太靠近显示器210且由于这种靠近,反而不能与屏幕上的对象交互。各种的标记都可以表示在显示器210上来指示用户当前的位置和/或他们与在显示器210上的对象交互的能力。

[0066] 提供有关当前交互能力的标记可能涉及一个或多个原因。例如,显示器可能在特定的情况要求大量用户同时交互,不鼓励任何一个靠近的用户支配交互。其他的情况可能包括显示器在窗口后面,靠近交互是不可能的。在这种情况下,可能产生“太靠近”的交互空间。

[0067] 在一些实施例中,可以显示各种效果或产生音频来吸引在上述体量区域外的用户进入到其中。在一个实施例中,在当前没有人在交互空间中且第一个用户进入到了交互空间时,显示器显示非常巨大、高亮可见的闪光或其他的动作。在当前没有人进行交互并且进入到交互空间的用户仅仅可以看见他们水平视场中的显示器时,这种特征被设计来引起用户对显示器的注意。

[0068] 如前所述,本发明的影像系统可以检测用户的手和身体。手和身体的检测可以使用范围、亮度、颜色和/或其他数据。手和身体检测方法可以被应用到原色/亮度图像、范围图像、三维点云,和/或其他数据。手和身体检测方法还可以被应用到基于上述数据集合的任意变换或投影数据。

[0069] 身体可以被检测来作为在三维云中空间上独立的点簇。身体还可以通过将点云投射到地面并识别出高密度区域而被检测。身体还可以通过去除深度迅速改变区域后寻找在适合的所投影的深度图像中大量连接的分量而被识别。身体还可以通过在原始三维数据或投影上执行模板匹配或骨骼匹配而被识别。初步识别可以在一个投影上进行,接着再在

精确的分离投影上进行。例如,影像系统可以使用地平视图来为用户身体挑选粗略的边框。接着可以使用在边框内的点的前沿投影来分离非常靠近的用户。

[0070] 手可以被检测为独立的对象或用户的一部分。在前一种情况下,手可以与用户关联。所检测的手的位置和动作可以相对于用户、显示器,或在交互过程中预先定义或动态产生的空间中的某个区域而定。根据所需的效果,可以使用多种可能的坐标系统来指定这个关系。例如,可以使用从手进入到以用户肩为中心的极坐标系统。手检测方法可以包括通过身体的骨骼相配的识别;通过模版匹配的识别;通过基于色彩的匹配分割的识别;身体相对对象时的识别;和基于连接的组件属性的识别。

[0071] 一旦用户的身体被识别,则例如框形、球形或圆柱形等的边界区域就在用户的中心。连接到身体但在该区域外的对象可能是手的候选。这些连接的对象还可以被分割以识别“指尖”或手。手可以通过抽取这些对象的极端区域而被识别,这些极端区域例如距离边界框或所限定的身体上的某个点最远距离的点,在能被认为是手的一部分的最远距离之下的一些阈值内的任意点。手还可以通过模版匹配而从这些区域中抽取。手也可以基于各种特征而从这些区域中分割出来,这些特征包括但不限于曲率、距离边界框或身体的距离,以及颜色属性。手臂的骨骼模型还适于使用期望最大化的区域。

[0072] 图3图示了包括体量区域的交互显示空间的另一个实施例。在图3中,交互显示器310和相机/照明系统320产生一个交互空间,其包括“太靠近”空间350、“手”空间340,和“远端用户”空间330。在“远端用户”空间330中的用户只能以他们的整个身体交互并且可以在显示器上看见整个身体的用户表示。在“手”空间340中的用户能以他们的手和身体交互并且能在显示器上看见手的表示以及整个身体的表示。

[0073] 将他们的手放入“太靠近”空间350的用户可能并不能以他们的手或身体进行交互,些屏幕上的视觉表示(例如在他们手的位置渐渐消失的一缕烟或他们身体渐渐褪色的图像)提示他们由于靠得太近而不能交互。例如,用户360可以看见他们的手的视觉表示变成一缕烟并可能不能使用他们的手来与屏幕上的对象交互。相反地,在“手”空间340中的用户370有完全的能力来使用他们的手与在屏幕上的对象交互并可以看见与用户360不同的他们的手的表示。“太靠近”空间的内部边缘可以与相机视场或照明器的照明范围的边界一致。

[0074] 显示器可以显示用户身体或身体的特定部分的表示来给用户提供视觉反馈到显示器上显示的虚拟空间内的位置,从而便于与显示器上显示的对象进行精确的交互。显示器还可以显示用户手的表示。由于在用户和显示器间的很多交互可以经由用户的手,这对于特定高亮显示用户的手来说是很有益的。当手距离身体足够远以至能传送某种姿态使用是可以仅显示手的表示。为手而使用的表示可以依靠手的位置或它是如何移动的。还可以反映手是否能选择或与对象交互。可以是拟人性的手、不同类型的对象,或这两个的组合。

[0075] 在一个实施例中,当用户或他们的手在显示器前的特定空间中时仅仅出现手的表示。替换地,当用户的手的表示开始被显示时用户整个身体的表示逐渐减弱或变形而当用户手的表示被去除时则返回。

[0076] 图4图示了用于基于姿态的交互的处理的实例性方法。在步骤410中,能收集有关其环境的实时三维数据的相机(例如立体或时差测距相机等)捕获一个或多个人类用户或

其他对象的动作。在步骤420中,从三维相机来的可以形成三维点云的数据由计算机影像系统分析,并被潜在地进行坐标变换以确定相对于显示器和在显示器上显示的虚拟对象的数据的位置。在步骤430中,计算机影像系统分析该数据以确定用户身体和手以及其他身体部分的位置。这些身体部分可以由人进行分组,并且它们的位置可以随时被跟踪以确定速度、加速度以及其他特性。

[0077] 在步骤440,确定每只手进行的与虚拟对象的选择关联的动作或与虚拟对象进行的其他交互的程度。在步骤450中,确定每只手的视觉表示,其可以取决于其选择的级别或其他动作而不同。然后,在步骤480中,与在步骤470中创建的用户身体的表示以及由仿真490控制而已与用户交互的虚拟对象一起渲染这些视觉表示。通过渲染在显示器上的表示,用户可以接收他们的身体和手相对于在屏幕上的虚拟对象的位置的视觉反馈。

[0078] 单独地,在步骤460中处理由步骤450所确定的每只手的选择级别以确定该手是否与从仿真490来的任何虚拟对象进行了交互。在图4中显示的用户交互处理可以非常迅速地反复进行以便产生实时交互的影像。

[0079] 在图4中显示的处理仅仅图示了一个可能的实施例。可以实现替换的实施例、例如,在仿真490中的对象可以分析在它们周围的视觉信号以确定在附近是否有手或是否发生了选择行为,从而忽略步骤440和430的要求。

[0080] 图5图示了手/用户表示的vis-à-vis选择级别。在步骤510中,确定手正在选择的对象的程度。在所示的实施例中,确定沿Z轴从躯干到手的简单距离。当用户向前伸出它们的胳膊时,该值就会增加。在步骤520中,在显示器上的用于手的视觉表示基于在步骤510中计算的距离值而确定。当距离增加时,手的表示逐渐地从张开的手靠近一个指示的手指。这给出了他们的部分选择级别的用户反馈,即他们的手距离他们能选择对象的阈值有多远。当手穿过了该阈值而进入到选择模式中时,它就能选择对象,则十字准线出现。如果没有手达到,则显示没有手的表示。单独地,在步骤510中计算的距离作为在步骤530中的“动作级别”,其中,如果动作级别超出了,例如0.5,则手就变得能选择对象了。在图5中所示的实例仅仅显示了一个可能的实施例。其他的各种方法都能被用来计算动作级别。

[0081] 在一个实施例中,用户的手或身体能通过物理碰撞的二维或三维仿真来与虚拟对象交互。在一个实施例中,显示器上显示的虚拟空间中的每只手的位置可以被用来作为给控制虚拟对象行为的物理仿真的输入。手要么由能预先定义的例如特定半径的球形等的特定形状模型化,要么由从用户手自身的二维或三维形状得到。在一个实施例中,手的视觉表示可以作为物理仿真的模型。

[0082] 可以使用在物理仿真中的手的模型来检测与在显示器上所显示的虚拟对象的交叠和碰撞。可以使用随时间而定的手的位置来得到速度和加速度,这使得能计算在显示器上的虚拟对象的真实碰撞。碰撞可以是无弹性的、部分弹性的,或完全弹性的,并且可以使用现有技术中公知的标准碰撞公式进行计算。由于手不能物理地由虚拟对象所移动,因此,可将手视为就好像为计算后碰撞速度的目的而具有无穷大。

[0083] 其他的身体部分,例如手臂、头,或整个身体等,可以用类似的方式被输入到物理仿真。有关手或身体其他部分的动作信息可以被用来计算交互而不是碰撞,例如流体或波纹仿真,或依靠速度或加速度的粒子发射。

[0084] 还可以使用附于手或身体其他部分的虚拟对象来于虚拟对象交互。例如,虚拟的

剑可以被附于用户的手上。对象的方向可以通过分析用户的手、前臂或胳膊的方向而计算，这使得用户很容易控制对象的方向。该对象还可以由物理引擎而分析以确定与其他对象的交互。

[0085] 图6图示了用户的物理位置如何被映射到数字对象的所显示的虚拟世界中。在图6中，用户620正在显示器630前交互。用户620在交互空间610中，交互空间610是在其交互能影响显示的空间。轴660限定了坐标系统。X和Y轴限定了在显示器630上的水平和垂直位置，而Z轴限定了距离显示器630的垂直距离。在该交互空间610中，用户620抬起他们的手670来与在屏幕630上显示的虚拟对象640交互。

[0086] 在一个实施例中，在手670和虚拟对象640间的交叠或临近通过将Z轴上手的位置投影到显示器表面的(X,Y)平面、随后分析手670和虚拟对象640的相对位置(X,Y)而计算。此外，手640的表示650在显示器630上通过沿Z轴投影手的位置而计算的位置。该表示给用户620提供有关他们的手670在显示器630的虚拟世界中的位置的视觉反馈。

[0087] 图7图示了如何将用户的物理位置映射到数字对象的所显示的虚拟世界中的替换性的实施例。在一个实施例中，显示器730在二维表面上显示图像并使用阴影或模糊等的深度提示来通知深度。在另一实施例中，显示器730是能给用户720产生三维图像的显示器，它给出了所显示的图像距离屏幕前或后一定距离的幻影。现今已有各种三维显示机制并在商业上可以得到，其包括但不限于，立体视觉投影和透镜视屏。

[0088] 用户720在交互空间710交互。当用户710伸出他们的手770时，他们的手的虚拟表示750就可以通过显示器730显示。当用户的手770进入到由显示器730所显示的虚拟对象740周围的触摸空间780时，就产生交互。所图示的坐标轴760仅是参考性的。

[0089] 存在很多种系统用来选择屏幕上的对象的潜在方法。重要的是让假负片(false negative)和假正片(false positives)最小化从而防止用户意外地选择他们不想选择的对象，和不能选择他们想选的对象。用来计算这些用户行为的信息可以使用来自于影像系统的各种输入，这包括但不限于，当前和以前的手的位置和速度、手臂的角度、当前和以前的头的位置，和当前和以前的躯干位置。

[0090] 通过在所选的对象或区域上让手静止不动可以产生选择。以几种方式来检测出保持手静止的动作。例如，如果动作级别超出了阈值，则对被选择的对象或区域随时跟踪动作级别。部分选择可以基于动作级别显示。随时间而定的手的位置可贡献于该动作级别。例如，如果手快速移过对象或区域或在对象或区域外，则动作级别可以由公式所确定的量而降低，而如果手位于对象或区域上并且缓慢地移动或停止，则动作级别可以由公式所确定的量而提升。公式可以考虑手的移动的速度，因此，手移动的越慢，则动作级别提升的越快。

[0091] 在特定的时间范围内手的平均速度低于特定值则也可以作为选择的机制。速度可以由定义在显示器上的交互区域的二维平面限定，由在显示器前的三维空间所限定，或通过例如在图6和7种所示的X和Y轴等的其他方法所限定。

[0092] 其他的选择机制包括：在特定时间范围内低于特定值的手的最大速度；在某些二维(例如，在图6和7中所示的X和Y轴)或三维坐标系统中手的位置在特定时间周期内位于特定区域内；在某些二维(例如，在图6和7中所示的X和Y轴)或三维坐标系统中手的位置的变化在特定时间周期内总在特定值之下；手的位置数据在时间间隔内取平均，接着该平均的位置数据使用这里所描述的任意技术或其他技术而在第二个时间间隔内进行检测。

[0093] 还能通过在戳的动作中向屏幕或虚拟对象插入手而产生选择。可以以各种方式检测戳的动作。例如,手距离屏幕的距离由于在特定时间周期中超过阈值量而降低;手沿Z轴相对于屏幕的速度超过了特定的值,并且手向屏幕移动;手沿Z轴相对于屏幕的平均速度在特定时间周期内超过了特定值,并且手向屏幕移动;上述实例的任何一个根据额外的标准,手的X和Y速度(平均X和Y速度)在特定时间周期内低于特定值。这防止了在用户旋转手时产生额外的选择;用户的速度在特定的方向范围内的方向中,并且速度超出了给定值。在这样的一个实施例中,方向的范围包括沿Z轴直接指向屏幕的矢量的特定角度的方向。

[0094] 选择还可以基于用户的手距离屏幕或任意三维区域的距离而产生。这样,当用户将他们的手伸入到“选择区域”时,则就能选择对象。选择还能基于用户的手相对于其身体的位置而产生。这种动作可以以各种方式被检测到,例如用户的手比用户的头、躯干、肩膀或重心具有更大的阈值距离;用户的高度与从用户的手到用户的头、躯干、肩膀、重心的比率低于一个特定量,这确保了较矮或较高的人都能以类似当动作选择对象;可以用几种方式限定距离,例如可以包括沿Z轴(向着屏幕)的距离,或在三维空间(X,Y,Z)或在加权的坐标系统中的距离。

[0095] 选择可以基于用户在特定位置迅速地摆动他们的手而产生。通过在特定时间间隔内用户的手的平均速度超出阈值量,但位置仍然在特定的区域内而检测出迅速地摆动。选择还能基于用户的手的平均速度在特定时间间隔内超过阈值量,但位置的变化仍然保持在特定量以下而产生。

[0096] 选择还能基于执行特定姿态而产生。

[0097] 由于不同的人可以直观地使用不同的方法来选择对象,所以允许使用多种选择方法是很有益处的。这样,例如如果用户让他们的手静止不动超过了特定的时间周期或者如果用户在特定的时间周期进行了快速的搅动,则用户都可以选择对象。

[0098] 为了防止意外的选择,需要同时调用多种选择方法来触发选择。例如,用户可能需要将他们的手静止特定的时间周期并且让他们的手距离他们的身体至少最小的距离。

[0099] 在选择处理中可以使用反馈来指示正在发生的选择的实际情况。在一个实施例中,当手正在选择对象时,用户的手在屏幕上的表示会产生变化。在另一实施例中,当手基于其位置或移动应当选择对象时用户的手在屏幕上的表示会产生变化,即便在该区域中没有对象。而在另一实施例中,当用户选择,或部分地选择对象时,屏幕上的对象会变化。

[0100] 手的屏幕上的表示替换地在其开始选择对象时产生改变。这给用户提供了有关选择处理特性的视觉反馈,并有助于用户防止意外的选择。在一个实施例中,在用户表示中的改变可以从基于选择检测算法特定值的图像序列中选择。例如,如果选择检测算法基于用户的身体和手的距离,则手会沿动画序列逐渐改变外形,其中使用距离值选择动画序列的帧。

[0101] 多个用户可以同时与显示器交互。相机系统能感应多个身体,同时影像系统可以从多个用户的身体区分数据。这种区分使得影像系统能将每只手指派给特定的用户身体。因此,不同的身体和手都能具有不同的用户表示并能与在显示器上显示的虚拟对象进行不同的交互。此外,显示器可以忽略特定的用户,或允许一个用户或一组用户控制交互的特定集合。

[0102] 上述的组件和功能可以由计算机可读存储介质上存储的指令构成。指令可以由

处理器检索并执行。指令的一些实例是软件、程序代码、以及固件。存储介质的一些实例是存储设备、磁带、磁盘、集成电路以及服务器。指令在由处理器执行来指示处理器根据本发明工作时是可操作的。本领域的技术人员很熟悉指令、处理器和存储介质。

[0103] 虽然已经描述了本发明的各种实施例,但应当理解的是,他们仅仅是示例性而非限制性的。例如,与用户界面关联的任意元件都可以提供上述描述的功能。因此,优选实施例的外延和范围并不应当限于上述示例性的实施例。

[0104] 例如,从一个或多个显示系统来的信息可以通过允许不同应用程序的高速数据连接共享。例如,显示器的用户可以浏览并与在线的虚拟世界交互,例如,Second Life, There,或World of Warcraft,使用基于姿态的界面的那个世界中的对象和化身。此外,用户的位置和动作的三维数据可以被传输并在在线虚拟世界中显示,从而使得在线虚拟世界的其他用户看见并与所显示的用户交互。这些其他的用户可以使用各种接口(例如,传统的键盘-鼠标接口)或其他的显示器来在虚拟世界中交互。用户可以彼此离得很近或很远。

[0105] 当前描述的实施例的界面可以提供给各种潜在的用户。例如,对于运动,用户可以拳击、打网球(用虚拟的球拍)、踢虚拟的球,或与计算机或在屏幕上显示的人类对手进行其他运动项目的竞技。在虚拟世界游弋的过程中,用户可以使用自然身体动作,例如学向虚拟世界移动,并且可以使用他们的手与在虚拟世界中的对象交互。可以通过上述的界面系统控制虚拟角色,其中在其传递下,屏幕上的数字角色可以与在显示器前的人交谈、玩耍或进行其他的交互。这种数字角色可以由计算机控制或由远端的人进行控制。

[0106] 本发明的实施例允许进行广告,包括交互产品演示和交互商标体验以及多用户工作空间,其中用户组可以以协作的方式移动和操作在屏幕上表示的数据。在上述界面系统中还可以执行视频游戏以使用户能通过姿态和身体的自然移动玩游戏和控制他们在屏幕上的角色。通过该系统还能实现衣服的建模和“试穿”,其中在用户的图像上可以放置衣服从而能让用户虚拟地试穿衣服。

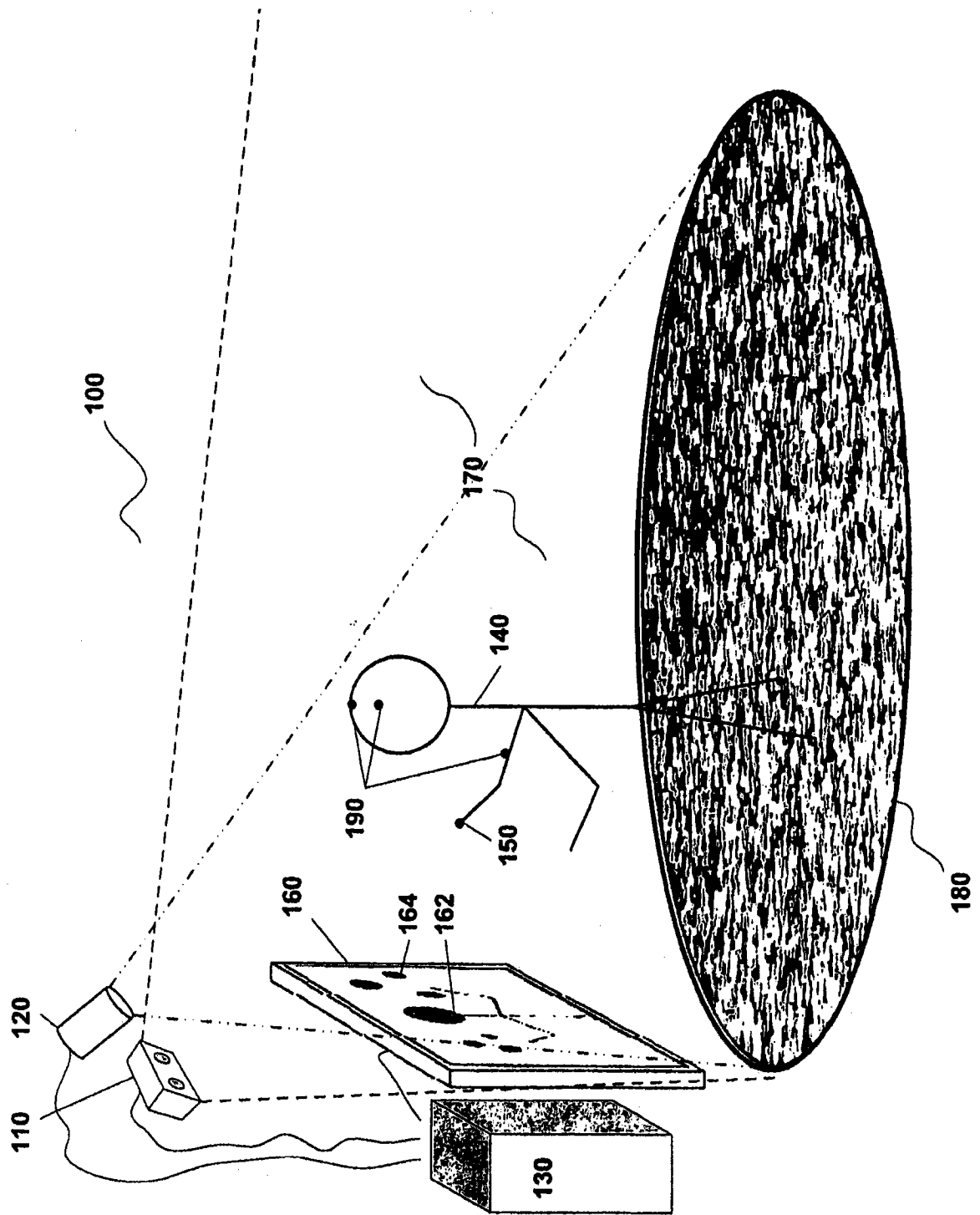


图1

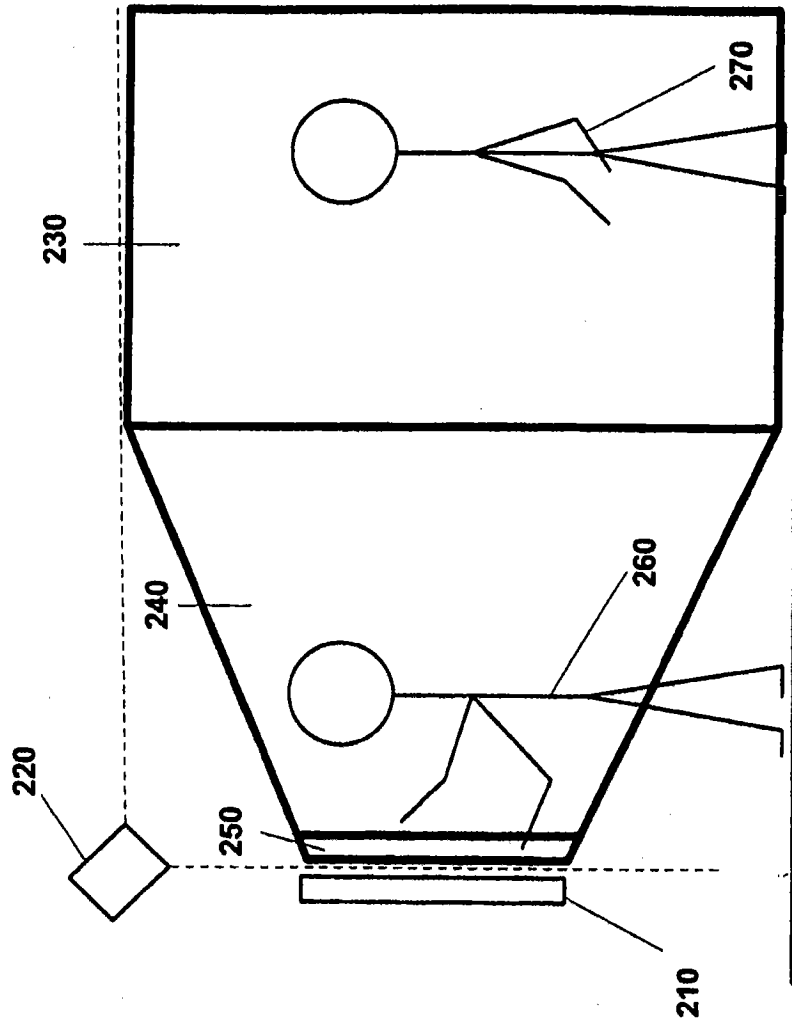


图2



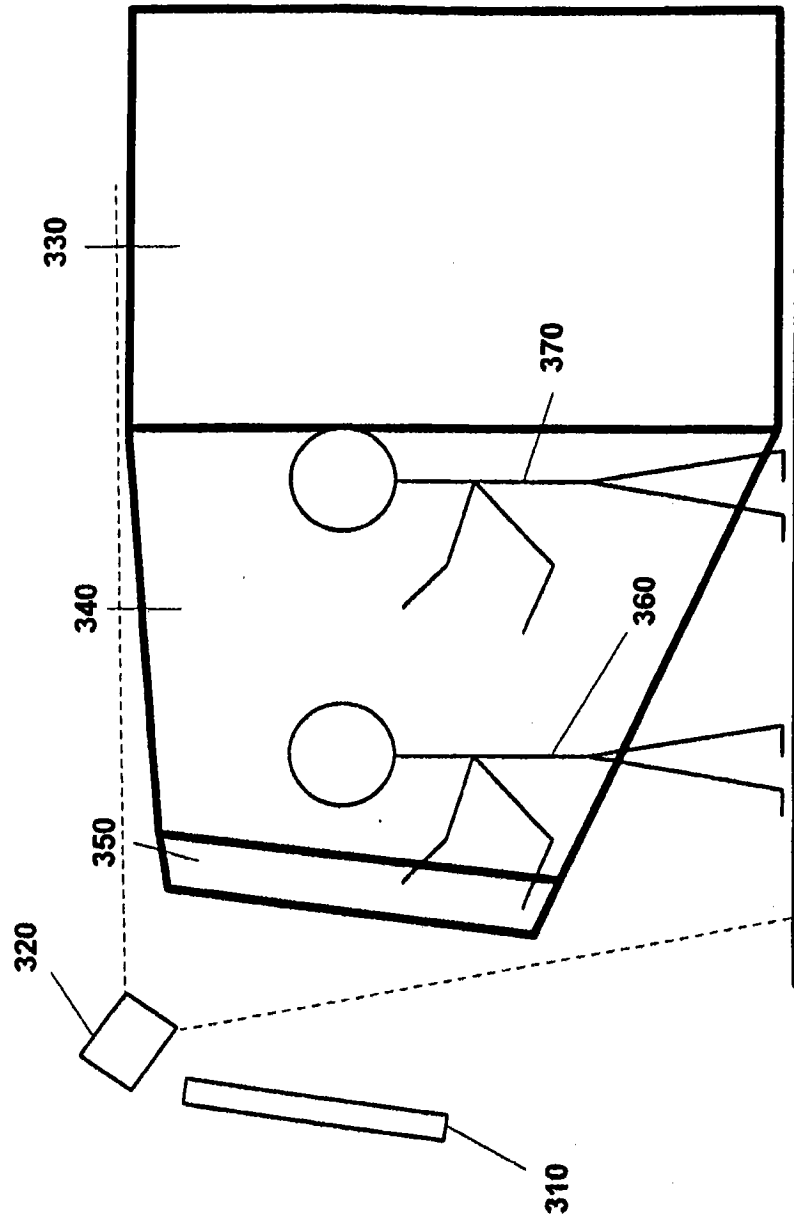


图3

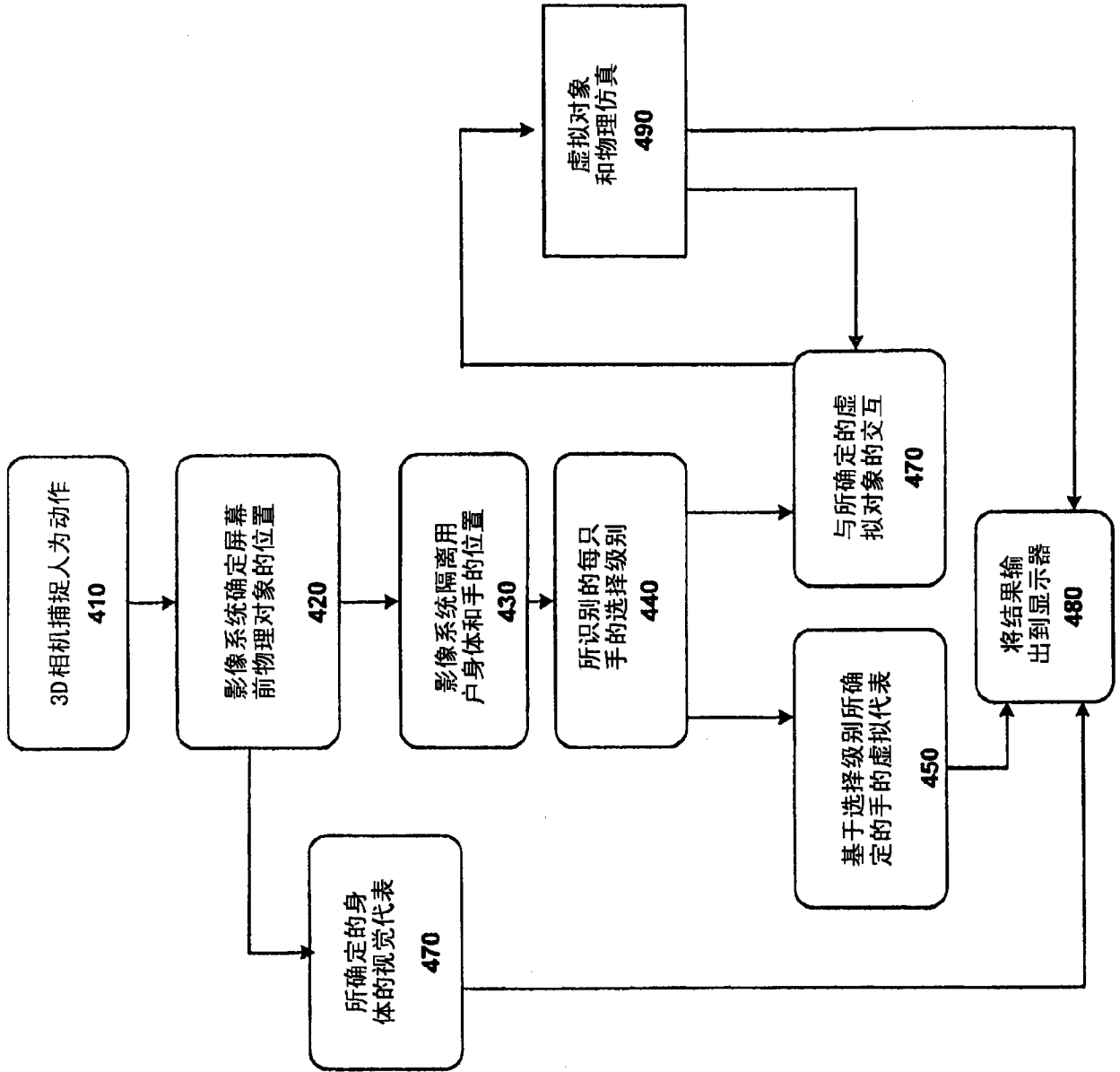


图4

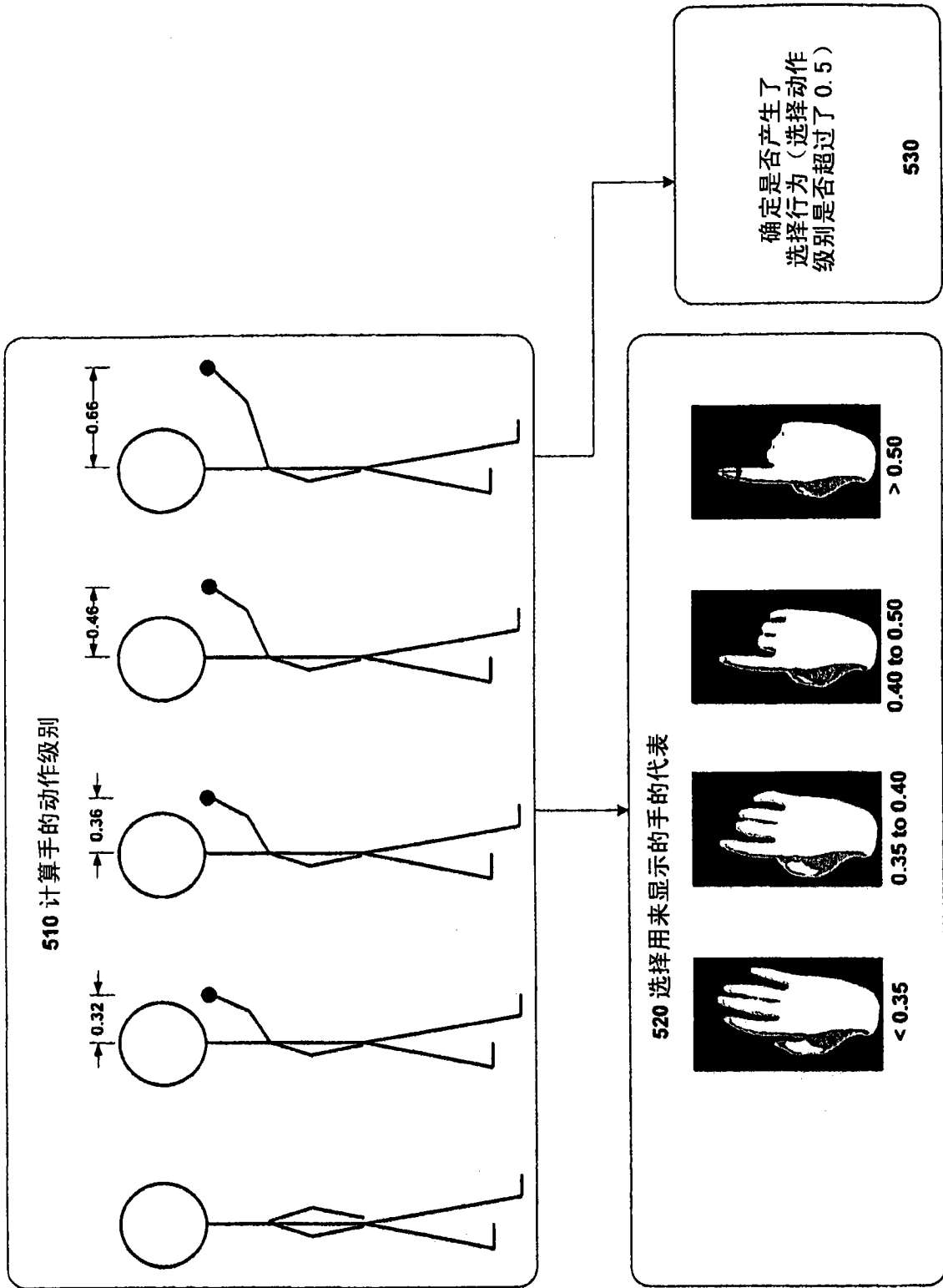


图5

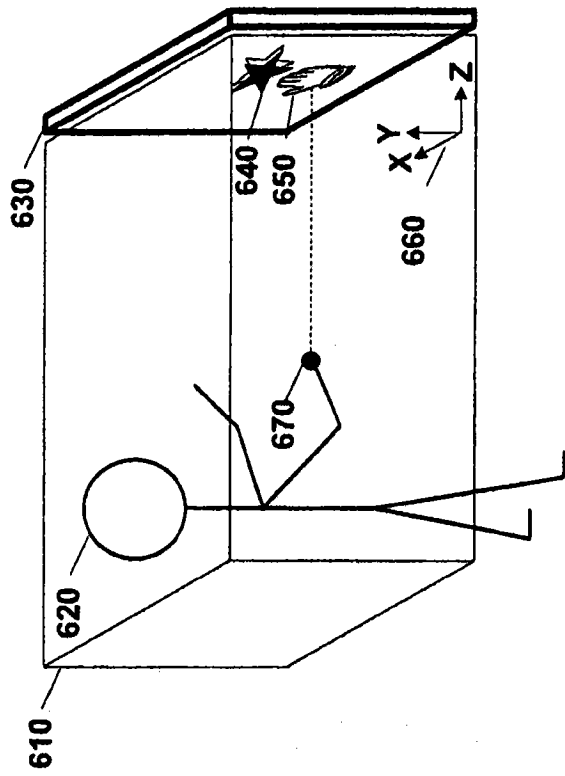


图6

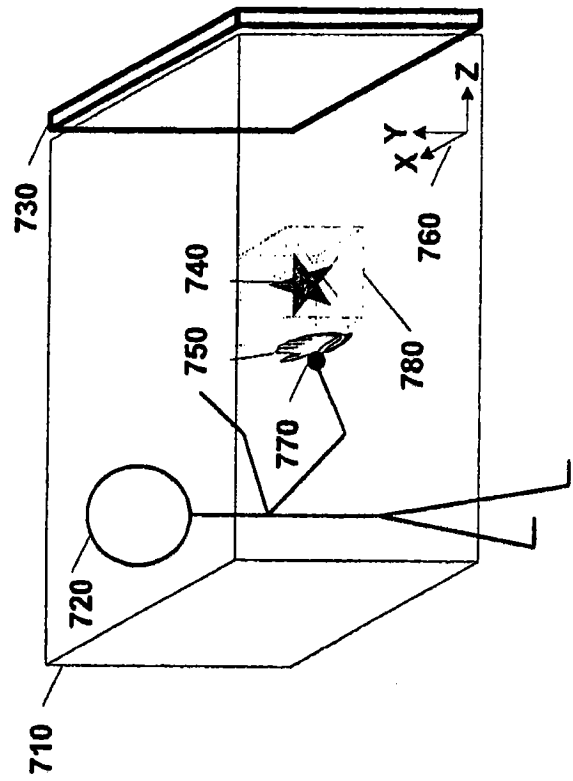


图7