



(12) 发明专利

(10) 授权公告号 CN 102591016 B

(45) 授权公告日 2014. 12. 17

(21) 申请号 201110444280. 1

(22) 申请日 2011. 12. 16

(30) 优先权数据

12/972, 100 2010. 12. 17 US

(73) 专利权人 微软公司

地址 美国华盛顿州

(72) 发明人 A·巴-泽埃夫 J·R·刘易斯

G·克莱因

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 胡利鸣

(51) Int. Cl.

G02B 27/01 (2006. 01)

(56) 对比文件

US 5406415 A, 1995. 04. 11,

US 7804507 B2, 2010. 09. 28,

US 6222675 B1, 2001. 04. 24,

JP 2008022236 A, 2008. 01. 31,

Andiel, M. et al.. Eye-Tracking for
Autostereoscopic Displays using Web Cams.
《stereoscopic displays and applications
X111》. 2002, 第 4660 卷

审查员 喻天剑

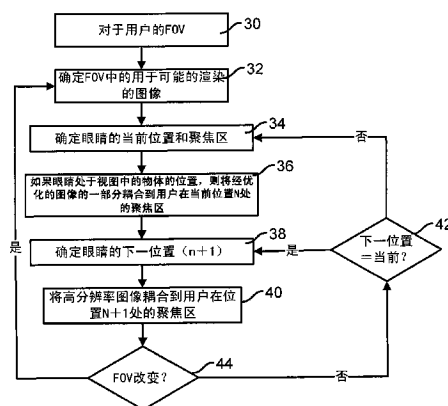
权利要求书2页 说明书22页 附图20页

(54) 发明名称

用于扩展现实显示的优化聚焦区

(57) 摘要

本发明涉及用于扩展现实显示的优化聚焦区。提供了一种在使用诸如透视显示设备或头戴式显示设备之类的近眼显示设备时增强用户体验的方法和系统。创建经优化的图像以用于相对于用户在场景中的视野进行显示。跟踪用户的头和眼睛位置和移动以确定该用户的聚焦区。将经优化的图像的一部分耦合到该用户在眼睛的当前位置处的聚焦区, 预测头和眼睛的下一位置, 并且将经优化的图像的一部分耦合到该用户在该下一位置处的聚焦区。



1. 一种用于向用户呈现经优化的图像的方法,所述方法包括:

基于用户在物理环境中的视野来创建经优化的图像;

跟踪相对于所述视野的眼睛位置和移动以确定所述用户的聚焦区;

在所述用户的聚焦区中将经优化的图像的一部分在光学上耦合到所述用户,所述耦合是通过将显示设备的处理能力集中在仅仅渲染图像的耦合到用户小凹视力的经优化部分来实现的;

预测所述用户的下一眼睛移动;以及

在所述用户的所述下一眼睛位置的聚焦区中将经优化的图像的一部分在光学上耦合到所述用户。

2. 如权利要求1所述的方法,其特征在于,创建经优化的图像的步骤包括:确定所述用户的视野以及基于场景中的用户位置创建经优化的图像。

3. 如权利要求1所述的方法,其特征在于,呈现经优化的图像的一部分的步骤包括:相对于所述视野确定经优化的图像的一部分以及将所述部分耦合到用户的眼睛的小凹。

4. 如权利要求1所述的方法,其特征在于,在光学上耦合的步骤包括:将所述图像的所述部分定位在头戴式显示器的眼睛位置处的聚焦区中。

5. 如权利要求4所述的方法,其特征在于,在光学上耦合的步骤包括:使用以机械方式控制的对所述图像的投影进行反射的镜来定位所述图像的所述部分。

6. 如权利要求4所述的方法,其特征在于,在光学上耦合的步骤包括:使用定位在用户的每只眼睛前方的发射性显示器来定位所述图像的所述部分。

7. 如权利要求1所述的方法,其特征在于,在光学上耦合的步骤包括:突出所述环境中的现实世界物体。

8. 一种头戴式显示设备,包括:

显示器,所述显示器将经优化的图像的至少一部分耦合到用户的聚焦区,其中所述耦合是通过将所述头戴式显示设备的处理能力集中在仅仅渲染图像的耦合到用户小凹视力的经优化部分来实现的;

感测所述头戴式显示设备的定向信息的惯性传感器,检测用户眼睛位置的眼睛跟踪传感器;以及

至少一个处理单元,所述处理单元与所述显示器、惯性传感器和眼睛跟踪传感器进行通信,以自动地相对于所述显示器在所述用户的眼睛的当前位置处显示经优化的图像的经优化的部分,使得所述图像的所述部分耦合到所述用户的聚焦区,确定所述用户的眼睛的下一位置;以及

相对于所述显示器在所述用户的眼睛的所述下一位置处显示经优化的图像的另一经优化的部分,使得所述图像的所述部分耦合到用户在所述下一位置处的聚焦区。

9. 如权利要求8所述的头戴式显示设备,其特征在于,所述显示器包括:

图像投影仪;

光导光学元件,所述光导光学元件与所述显示器对准以允许所述用户查看所述用户所处的场景中的物体;以及

其中所述显示器还包括耦合到机械控制元件的镜元件,所述控制元件响应于所述至少一个处理单元来对所述图像的经优化的部分进行定位。

10. 如权利要求 8 所述的头戴式显示设备,其特征在于:

所述至少一个处理单元与中枢计算设备通信以接收空间的标识出一个或多个物体的三维模型;以及

所述至少一个处理单元确定用户的视野,确定第一物体是否处于该视野中,确定所述第一物体在所述显示器中的位置,基于所述用户在所述场景中的大小和定向来调整经优化的图像的大小和定向。

用于扩展现实显示的优化聚焦区

技术领域

[0001] 本发明涉及用于扩展现实显示的优化聚焦区。

背景技术

[0002] 扩展现实是一种允许将虚拟图像与现实世界物理环境或空间相混合的技术。一般而言,近眼显示器使用光学器件与立体视觉的组合来将虚拟图像聚焦到空间内。在这样的显示器中,显示分辨率和处理是宝贵的。

[0003] 在某些情况下,通过近眼显示设备向用户所显示的虚拟图像可能包括包含高度精细图形的虚拟图像或物体。通常向佩戴近眼显示设备的用户呈现该用户不一定有兴趣查看的大量信息。

发明内容

[0004] 所简述的技术包括一种用于向用户呈现经优化的图像的方法。创建经优化的图像以用于相对于用户在场景中的整个视野进行显示。跟踪用户的头和眼睛位置和移动以确定该用户的聚焦区。将经优化的图像的一部分耦合到该用户在眼睛的当前位置处的聚焦区,预测头和眼睛的下一位置,并且将经优化的图像的一部分耦合到该用户在下一位置处的聚焦区。

[0005] 另外,提供了一种头戴式显示设备。该头戴式显示器包括将经优化的图像的至少一部分耦合到用户的聚焦区的显示器。惯性传感器、磁传感器、机械传感器和/或其他传感器感测头戴式显示设备的定向信息,并且眼睛跟踪传感器检测用户眼睛位置。与显示器、惯性传感器和/或其他传感器和眼睛跟踪传感器进行通信的处理单元自动地相对于显示器在用户的眼睛的当前位置处显示经优化的图像的经优化的部分,使得该图像的该部分耦合到该用户的聚焦区。然后,该处理设备确定用户的眼睛的下一位置并且相对于该显示器在用户眼睛的该下一位置处显示经优化的图像的另一经优化的部分,使得该图像的该部分耦合到用户在下一位置处的聚焦区。

[0006] 提供本发明内容以便以简化的形式介绍将在以下具体实施方式中进一步描述的一些概念。本发明内容并不旨在标识出所要求保护的主题的关键特征或必要特征,也不旨在用于帮助确定所要求保护的主题的范围。此外,所要求保护的主题不限于解决在本发明的任一部分中提及的任何或所有缺点的实现。

[0007] 附图说明

[0008] 图1是描绘了用于基于用户意图生成经优化的内容的系统的一个实施例的示例性组件的框图。

[0009] 图2A是示出描绘根据本技术的方法的流程图。

[0010] 图2B描绘了用户对用户环境中的一个或多个物体的视图的一个实施例。

[0011] 图2C是被示为覆盖用户对用户环境的视图的在完全分辨率下的经优化图像的图示。

- [0012] 图 2D 是用户对完全分辨率图像的经优化部分和被部分遮蔽的视野的视图的图示。
- [0013] 图 2E 是与本技术一起使用的一种类型的聚焦区的图示。
- [0014] 图 2F 是用户环境的图示。
- [0015] 图 2G 是全分辨率图像和被部分遮蔽的用户视野的同时渲染的当前和下一位置经优化部分的图示。
- [0016] 图 3A 是头戴式显示单元的一个实施例的一部分的顶视图。
- [0017] 图 3B 是头戴式显示单元的另一实施例的一部分的顶视图。
- [0018] 图 3C 和 3D 是头戴式显示单元的另一实施例的一部分的顶视图和侧视图。
- [0019] 图 4A 是头戴式显示单元的组件的一个实施例的框图。
- [0020] 图 4B 是与头戴式显示单元相关联的处理单元的组件的一个实施例的框图。
- [0021] 图 5 是结合头戴式显示单元使用的中枢计算系统的组件的一个实施例的框图。
- [0022] 图 6 是可用于实现在此所述的中枢计算系统的计算系统的一个实施例的框图。
- [0023] 图 7 是描绘用于生成经优化内容的多用户系统的框图。
- [0024] 图 8 描述了用于基于用户意图生成经优化内容的过程的一个实施例。
- [0025] 图 9 描述了用户创建用户的空间的模型的过程的一个实施例。
- [0026] 图 10 是描述了用于将空间的模型分割成各个物体的过程的一个实施例的流程图。
- [0027] 图 11 是描述用于跟踪用户的视野以及确定用户的聚焦区的过程的一个实施例的流程图。
- [0028] 图 12 是描述了由中枢计算系统执行以提供用在图 2A 的过程中的跟踪信息的过程的一个实施例的流程图。
- [0029] 图 13 是描述了用于跟踪眼睛的过程的一个实施例的流程图,其中该过程的结果供图 12 的过程来使用。
- [0030] 图 14 是描述用于提供经优化图像的所选部分以耦合到用户小凹的过程的一个实施例的流程图。
- [0031] 图 15 是描述用于从眼睛的当前位置估计眼睛的下一位置的过程的流程图。
- [0032] 图 16 是根据图 15 的步骤的用于预测一个或或多个接下来的眼睛位置的方法。

具体实施方式

[0033] 公开了一种用于增强用户在使用近眼显示设备时的体验的技术。用户通过诸如头戴式显示设备之类的近眼显示设备来查看场景。确定作为用户正在查看的环境或空间的用户视野。渲染如下图像:该图像被优化以供相对于视野来使用。通过跟踪该用户的眼睛在视野内的位置来确定用户的聚焦区。通过将该图像的经优化的部分耦合到用户的聚焦区(在一种情况下是用户的小凹)来提供经优化图像的显示,以减少显示所需的处理和能量。跟踪用户眼睛位置,并且计算出下一眼睛位置,以根据用户眼睛到下一位置的移动来将该图像的该部分定位在该下一位置处。

[0034] 图像的经优化的部分的定位是通过任何数目的不同显示设备来执行的,包括以机械方式控制的镜以及投影显示器。使用预测算法来确定用户眼睛的潜在的接下来的位置。

[0035] 图 1 是描绘了用于基于用户意图生成经优化的图像的系统 10 的一个实施 例的示例性组件的框图。系统 10 包括作为通过线 6 与处理单元 4 进行通信的近眼头戴式显示设备 2 的透视显示设备。在其他实施例中,头戴式显示设备 2 通过无线通信来与处理单元 4 进行通信。尽管图 1 的组件示出了透视显示设备,但是图 3B-3D 中示出了适于与现有技术一起使用的其他显示器实施例。

[0036] 在一个实施例中为眼镜形状的头戴式显示设备 2 被佩戴在用户的头上,使得用户可以透过显示器进行查看,并且从而具有该用户前方的空间的实际直接视图。使用术语“实际和直接视图”来指直接用人眼查看现实世界物体的能力,而不是查看物体的所创建的图像表示。例如,在房间中透过眼镜进行查看将允许用户具有该房间的实际直接视图,而在电视上查看房间的视频不是房间的实际直接视图。下面提供头戴式显示设备 2 的更多细节。尽管图 1 和 3A-3D 中所示的设备为眼镜形式,但是头戴式显示设备 2 可以采取其他形式,比如具有护目镜的头盔。

[0037] 在一个实施例中,处理单元 4 被佩戴在用户的手腕上,并且包括用于操作头戴式显示设备 2 的计算能力的一部分。处理单元 4 可以与一个或多个中枢计算系统 12 无线地(例如 WiFi、蓝牙、红外、或其他无线通信手段)通信。

[0038] 中枢计算系统 12 可以是计算机、游戏系统或控制台等等。根据一示例性实施例,中枢计算系统 12 可以包括硬件组件和 / 或软件组件,使得中枢计算系统 12 可以用于执行诸如游戏应用、非游戏应用等等之类的应用。在一个实施例中,中枢计算系统 12 可以包括诸如标准化处理器、专用处理器、微处理器等等之类的处理器,这些处理器可以执行存储在处理器可读存储设备上的指令以用于执行在此所述的过程。

[0039] 在各个实施例中,在此参照图 2A 和 8-15 所描述的过程全部或部分地通过中枢计算系统 12、处理单元 4 二者任一和 / 或二者组合来执行。

[0040] 中枢计算系统 12 还包括一个或多个捕捉设备,如捕捉设备 20A 和 20B。在其他实施例中,可以使用多于或少于两个的捕捉设备。在一个示例性实施方式中,捕捉设备 20A 和 20B 指向不同方向,使得它们可以捕捉房间的不同部分。可能有利的是,两个捕捉设备的视野稍微地重叠,使得中枢计算系统 12 可以理解捕捉设备的视野如何彼此相关。通过这种方式,可以使用多个捕捉设备来查看整个房间(或其他空间)。可替代地,如果捕捉设备可以在操作期间 平移,使得整个相关空间随时间被捕捉设备查看,则可以使用一个捕捉设备。

[0041] 捕捉设备 20A 和 20B 例如可以是相机,该相机在视觉上监视一个或多个用户和周围空间,使得可以捕捉、分析并跟踪该一个或多个用户所执行的姿势和 / 或移动以及周围空间的结构,以在应用中执行一个或多个控制或动作和 / 或使化身或屏上人物动画化。

[0042] 中枢计算环境 12 可以连接到诸如电视机、监视器、高清电视机 (HDTV) 等可提供游戏或应用程序视觉的视听设备 16。例如,中枢计算系统 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 可从中枢计算系统 12 接收视听信号,并且然后可以输出与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、分量视频电缆、RCA 电缆等连接至中枢计算系统 12。在一个示例中,视听设备 16 包括内置扬声器。在其他实施例中,视听设备 16、单独的立体声系统或中枢计算设备 12 连接到外部扬声器 22。

[0043] 中枢计算设备 10 可以与捕捉设备 20A 和 20B 一起用于识别、分析和 / 或跟踪人类 (以及其他类型的) 目标。例如,可使用捕捉设备 20A 和 20B 来跟踪佩戴头戴式显示设备 2 的用户,使得可以捕捉用户的姿势和 / 或运动来使化身或屏幕上人物动画化,和 / 或可将用户的姿势和 / 或移动解释为可用于影响中枢计算系统 12 所执行的的应用的控制。

[0044] 在所公开的技术的下面将详细讨论的一个实施例中,系统 10 基于确定用户的视野和聚焦区来为该用户生成经优化的图像。经优化的图像例如可以包括视野上的物体或人工覆盖的物体的增强外观,该增强外观为用户提供增强的体验。经优化的图像根据用户的眼睛位置和移动模式通过头戴式显示设备 2 以经优化的部分 (其在图 2B 中示出) 被显示给用户。

[0045] 图 2A 示出了用于在显示器中将经优化图像的一部分耦合到眼睛的方法的一个实施例。图 2A 的过程将参照图 2B-2G 来描述。

[0046] 在步骤 30,确定用户的视野和聚焦区。如上所述,用户的视野与如下项目有关 : 用户的环境或场景、用户的头位置和定向、以及用户的眼睛位置。图 2F 示出了处于环境 1100 中的用户 1112。用户 1112 被示为查看视力范围 (由 线 1121 来定义) 中的多个物体,包括灯 1106、桌子 1120、处理设备 1116、捕捉设备 20A、20B、显示器 1110 和时钟 1118。该用户还看见地板 1108 和墙壁 1102,但是未看见椅子 1107 并且仅看见墙壁 1104 的一部分。可以相对于坐标系 1150 来定义环境 1100 并且相对于第二坐标系 1152 来定义用户的头位置。

[0047] 在一个实施例中,用户视野内的聚焦区是沿着聚焦曲线的注视点 150 处的那个区域。例如,瞳孔之间的会聚可以用于对聚焦曲线 147 上的焦点 (即双眼单视界 (Horopter)) 进行三角测定,从该焦点中可以计算出聚焦区和 Panum 汇合区域。Panum 汇合区域 147 是视网膜上的如下区域 : 该区域使得该区域中的任何点都将与另一视网膜上的单个点汇合并且通过人眼创建代表双眼立体视觉的单视力。如图 2E 所示,每个用户的眼睛都包括中央凹 (通常亦称小凹), 其位于视网膜的黄斑区的中心处。小凹负责获得清晰的中心视力 (亦称小凹视力), 该中心视力是人类在阅读、观看电视或电影、驾驶、以及视觉细节最重要的任何活动时所必需的。图 2E 的小凹在 148 和 149 处予以示出。

[0048] 相对于小凹 148、149 来定向和耦合经优化图像将保证 : 用户可以相对于该图像的经优化部分来聚焦视力。另外,需要耦合到小凹的图像部分是相对小的图像——在视网膜上量级为 1mm 直径。由头戴式显示器 2 渲染相对小的区域部分将降低头戴式显示器 2 的功率要求。

[0049] 在步骤 32,创建要针对所确定的视野向用户显示的经优化图像。该经优化的图像由渲染该图像的应用来确定,并且可以包括视野内的一个或多个单独的覆盖图像或者涵盖整个视野。

[0050] 图 2B 示出了环境 1100 中的用户 1112 在没有图像被投影在显示设备 2 上时透过该显示器所看见的东西。图 2C 示出了可以投影在图 2B 中的场景上的经优化图像。在图 2C 的示例中,经优化的图像 1200 完全遮蔽了环境 1100 的场景。在该示例中,场景描绘了人 1128、小鸟图像 1126 和奖杯 1127。在一个实施例中,渲染该图像的应用已经确定了该场景内的这些元素的配置和位置、以及这样的物体是否应当遮蔽现实世界物体,比如灯 1106、时钟 1118 和显示器 1110。

[0051] 在步骤 34,确定用户眼睛的当前位置,并且在 36,将经优化的图像的经优化部分

耦合到用户的聚焦区中的当前眼睛位置处。在一个示例中,这在图 2D 中予以示出,其中小鸟图像 1126 被示为突出显示的并且被叠加在现实世界环境上,其中小鸟图像 1126 相对于图像的平衡被突出显示。在一个实施例中,经优化图像的其他元素(在该示例中为人和奖杯)不被渲染,或者以较低分辨率被渲染(未示出)。在另一方面,该房间的其他视觉元素可以从用户的视图中被模糊掉。

[0052] 通过将显示设备的处理能力集中在仅仅渲染图像的耦合到用户小凹视力的经优化部分,经优化图像中的其他元素不需要被渲染,或者可以用与经优化图像相比更低的精度以及由此更少的资源来渲染。在一个实施例中,经优化的图像是整个图像的片段。正常眼睛瞳孔可以具有明亮的光下的 1mm 至黑暗中的 7mm 之间的直径。显示通常是针对 3mm 直径光被优化的。通过将图像的该部分聚集到瞳孔上,可以直接将图像光平移到用户的聚焦区,从而显著减小为了生成图像所需的光。

[0053] 为了将图像的一部分集中在主体的瞳孔上,给由图像生成的光线赋予方向和瞳孔上的目标点。来自眼睛附近的一些光线将进入眼睛瞳孔,但是未进入瞳孔的光线被浪费,消耗功率并且可能具有其他不期望的效果。

[0054] 通常可以考虑:来自远处点的光线在眼睛附近全部都是近乎平行的,并且共享大致共同的方向。理想地,眼睛的光学器件将使些光线聚焦在视网膜上的小凹区处。来自不同方向的平行光线将被看成是不同点。

[0055] 为了优化眼睛瞳孔处的图像,头戴式显示器改变来自图像的光线被定向到的方向和在瞳孔上的进入点。在主体可以校正在自由空间中透过光学元件所查看的场景的光学失真的情况下,在本技术中由头戴式显示器的机构来执行图像显示的校正和引导。

[0056] 根据本技术,下面所述的实施例的定位系统 160、160a 提供对微显示器或将图像反射到用户的镜的方向性定位。该方向性定位与图像相对于显示器或镜的定位一起提供相对于眼睛位置的经优化的位置和方向。这例如可以通过在三维中倾斜显示器 153 或镜 166 来实现,其中图像被渲染在该显示器或镜上并且处于该镜或显示器上的合适位置。

[0057] 将认识到,除了下面阐述的那些以外,还可以提供各种其他类型的机械或机电元件来优化所显示的图像的方向。该方向性定位结合本技术的预测眼睛跟踪来为系统提供优化的处理。

[0058] 为了维持图像的经优化部分的耦合,在步骤 38 跟踪用户下次可能的眼睛移动,并且在步骤 40,在 240 将经优化的显示图像的另一经优化部分耦合到用户在下一位置处的聚焦区。在 44 如果视野改变,则在 32 确定新视野。在 44 如果视野未改变,则在 38 确定:用户的眼睛是否实际移动到了所预测的位置,并且该方法在 38 计算出潜在的下一眼睛移动位置。在 44 跟踪眼睛的当前位置、在 38 计算下一位置的循环可以由一个或多个微处理器或专用跟踪电路以近乎即时的方式来执行,以便根据用户眼睛的移动将经优化图像的部分移动到用户眼睛移动的下一位置处,以便提供合适的视觉体验。

[0059] 应当认识到,在前述说明中所使用的术语“下一个”和“当前”不一定限于图像的单个位置。例如,参考图 2G,在步骤 36 或 40 作为当前位置所渲染的每个部分都可以包括图像在第一时间 T 的部分(图像 1126)以及图像在第二时间 T2 的用户的眼睛被预测为将在时间 t2 所处的部分(图像 1135),使得在图 2A 中所述的每个“当前”和“下一”部分都可以包括两个图像部分。

[0060] 而且,能够认识到,可以以完全或部分完全分辨率来渲染图像 1200 的备选部分,以便将用户的眼睛吸引到该位置。例如,应用可以选择渲染选手 1128 或奖杯 1127 以便在该应用的给定上下文中将吸引用户的眼睛移动和用户的注意力。

[0061] 还应当认识到,本技术不需要使用覆盖图像,并且可以有利地用于仅仅向用户显示图像而不参考用户的环境。

[0062] 图 3A 描绘了头戴式显示设备 2 的顶视图,其包括镜架的包含镜腿 102 和鼻梁 104 的那部分。仅仅描绘了头戴式显示设备 2 的右侧。在鼻梁 104 中置入了话筒 110 以用于记录声音以及将音频数据传送给处理单元 4,这将在下面予以描述。在头戴式显示设备 2 的前方是朝向可捕捉视频和静止图像的视频相机 113 的房间。这些图像被传送给处理单元 4,这将在下面予以描述。

[0063] 头戴式显示设备 2 的镜架的一部分将围绕显示器(其包括一个或多个透镜)。为了示出头戴式显示设备 2 的组件,未描绘围绕显示器的镜架部分。该显示器包括光导光学元件 112、不透明度滤光器 114、透视透镜 116 和透视透镜 118。在一个实施例中,不透明度滤光器 114 处于透视透镜 116 之后并与其对准,光导光学元件 112 处于不透明度滤光器 114 之后并与其对准,并且透视透镜 118 处于光导光学元件 112 之后并与其对准。透视透镜 116 和 118 是眼镜中使用的标准透镜,并且可根据任何处方(包括无处方)来制作。在一个实施例中,透视透镜 116 和 118 可由可变处方透镜取代。在一些实施例中,头戴式显示设备 2 将仅仅包括一个透视透镜或者不包括透视透镜。在另一替代方案中,处方透镜可以进入光导光学元件 112 内。不透明度滤光器 114 滤除自然光(要么以每像素为基础,要么均匀地)以增强虚拟图像的对比度。光导光学元件 112 将人造光引导至眼睛。下面提供不透明度滤光器 114 和光导光学元件 112 的更多细节。

[0064] 在镜腿 102 处或镜腿 102 内安装有图像源,该图像源(在一个实施例中)包括用于对虚拟图像进行投影的微显示器部件 120、以及用于将图像从微显示器 120 定向到光导光学元件 112 中的透镜 122。在一个实施例中,透镜 122 是准直透镜。

[0065] 控制电路 136 提供支持头戴式显示设备 2 的其他组件的各种电子装置。下面参考图 4A 和 4B 提供控制电路 136 的更多细节。处于镜腿 102 内部或安装在镜腿 102 处的有耳机 130、惯性传感器 132、和 / 或磁传感器 132 以及温度传感器 138。在一个实施例中,惯性传感器 132 磁传感器 132 包括三轴磁力计 132A、三轴陀螺仪 132B、以及三轴加速度计 132C(参见图 5)。惯性传感器和 / 或磁传感器用于感测头戴式显示设备 2 的位置、定向、以及突然加速。

[0066] 微显示器 120 通过透镜 122 来投影图像。存在着可用于实现微显示器 120 的不同的图像生成技术。例如,微显示器 120 可以使用透射投影技术来实现,其中光源由光学有源材料来调制,用白光从背后照亮。这些技术通常是使用具有强大背光和高光能量密度的 LCD 类型的显示器来实现的。微显示器 120 还可使用反射技术来实现,其中外部光被光学活性材料反射并调制。根据该技术,由白光源或 RGB 源向前点亮照明。数字光处理(DLP)、硅上液晶(LCOS)、以及来自 Qualcomm 有限公司的 **Mirasol®** 显示技术都是高效的反射技术的示例,因为大多数能量从已调制结构反射离开并且可用在本文描述的系统中。附加地,微显示器 120 可以使用发射技术来实现,其中光由该显示器生成。例如,来自 Microvision 有限公司的 PicoP™ 显示引擎使用微型镜面舵来将激光信号发射到充当透射元件的小型屏幕上

或直接照射到眼睛（例如激光）。

[0067] 光导光学元件 112 将来自微显示器 120 的光传送到佩戴头戴式显示设备 2 的用户的眼 140。光导光学元件 112 还允许如箭头 142 所示那样将光从头戴式显示设备 2 的前方透过光导光学元件 112 透射到用户的眼睛，从而除接收来自微显示器 120 的虚拟图像之外还允许用户具有头戴式显示设备 2 的前方的空间的实际直接视图。因此，光导光学元件 112 的壁是透视的。光导光学元件 112 包括第一反射面 124（例如镜面或其他表面）。来自微显示器 120 的光穿过透镜 122 并入射在反射面 124 上。反射面 124 反射来自微显示器 120 的入射光，使得光通过内反射而被捕获在包括光导光学元件 112 的平面衬底内。在离开衬底的表面的若干反射之后，所捕获的光波到达选择性反射面 126 的阵列。注意，五个表面中只有一个表面被标记为 126 以防止附图太过拥挤。反射面 126 将从衬底出射并入射在这些反射面上的光波耦合到用户的眼睛 140。由于不同光线将以不同角度传播并弹离衬底的内部，因此这些不同的光线将以不同角度击中各个反射面 126。因此，不同光线将被所述反射面中的不同反射面从衬底中反射出。关于哪些光线将被哪个表面 126 从衬底反射出的选择是通过选择表面 126 的合适角度来设计的。光导光学元件的更多细节可以在于 2008 年 11 月 20 日公开的美国专利申请公开号 2008/0285140、序列号 12/214,366 的“Substrate-Guided Optical Devices（衬底导向的光学设备）”中找到，该申请的全部内容通过引用结合于此。在一个实施例中，每只眼睛将具有其自己的光导光学元件 112。当头戴式显示设备具有两个光导光学元件时，每只眼睛都可以具有其自己的微显示器 120，该微显示器 120 可以在两只眼睛中显示相同图像或者在两只眼睛中显示不同图像。在另一实施例中，可以存在将光反射到两只眼睛中的一个光导光学元件。

[0068] 与光导光学元件 112 对齐的不透明度滤光器 114 要么均匀地，要么以每像素为基础来选择性地阻挡自然光，以免其穿过光导光学元件 112。在一个实施例中，不透明度滤光器可以是透视 LCD 面板、电致变色膜（electrochromic film）、PDLC（聚合物分散液晶）或能够充当不透明度滤光器的类似设备。通过从常规 LCD 中除去衬底、背光和漫射体的各层，可以获得这样的透视 LCD 面板。LCD 面板可包括一个或多个透光 LCD 芯片，所述透光 LCD 芯片允许光穿过液晶。例如，在 LCD 投影仪中使用了这样的芯片。

[0069] 不透明度滤光器 114 可以包括致密的像素网格，其中每个像素的透光率能够在最小和最大透光率之间被个别化地控制。尽管 0-100% 的透光率范围是理想的，然而更有限的范围也是可以接受的。作为示例，具有不超过两个偏振滤光器的单色 LCD 面板足以提供每像素约 50% 到 99% 的不透明度范围，最高为该 LCD 的分辨率。在 50% 的最小值处，透镜将具有稍微带色彩的外观，这是可以容忍的。100% 的透光率表示完美清澈的透镜。可以从 0-100% 定义“阿尔法（alpha）”标度，其中 0% 不允许光穿过，并且 100% 允许所有光穿过。可以由下面描述的不透明度滤光器控制电路 224 为每个像素设置阿尔法的值。

[0070] 在用代理为现实世界物体进行 z-缓冲（z-buffering）之后，可以使用来自渲染流水线的阿尔法值的掩码（mask）。当系统为增强现实显示而呈现场景时，该系统注意到哪些现实世界物体处于哪些虚拟物体之前。如果虚拟物体处于现实世界物体之前，则不透明度对于该虚拟物体的覆盖区域而言应当是开启的。如果虚拟物体（实际上）处于现实世界物体之后，则不透明度以及该像素的任何色彩都应当是关闭的，使得对于现实光的该相应区域（其大小为一个像素或更多）而言，用户将会仅仅看到现实世界物体。覆盖将是以逐像素

为基础的,所以该系统可以处理虚拟物体的一部分处于现实世界物体之前、该虚拟物体的一部分处于现实世界物体之后、以及该虚拟物体的一部分与现实世界物体相重合的情况。对这种用途而言,最期望的是能够以低的成本、功率和重量从 0% 达到 100% 不透明度的显示器。此外,不透明度滤光器可以比如用彩色 LCD 或用诸如有机 LED 等其他显示器来以彩色进行呈现,以提供宽视野。于 2010 年 9 月 21 日提交的美国专利申请号 12/887,426“Opacity Filter For See-Through Mounted Display(用于透射安装显示器的不透明度滤光器)”中提供了不透明度滤光器的更多细节,该专利申请的全部内容通过引用并入本文。

[0071] 诸如 LCD 之类的不透明度滤光器通常还未与在此所述的透视透镜一起使用,因为在与眼睛的该近距离下,其可能失焦。然而,在一些情况下,这个结果可能是所期望的。通过使用相加色 (additive color) 的普通 HMD 显示器 (其被设计为是聚焦的),用户看到具有清晰彩色图形的虚拟图像。该 LCD 面板被放置在该显示器“之后”,使得模糊的黑色边界围绕任何虚拟内容,使其根 据需要而不透明。该系统扭转了自然模糊化的缺陷以方便地获得了抗混叠特性和带宽减少。这些是使用较低分辨率且失焦图像的自然结果。存在对数字采样图像的有效平滑化。任何数字图像都经历混叠,其中采样的离散性导致与自然模拟和连续信号相比的误差。平滑化意味着在视觉上更靠近理想的模拟信号。尽管在低分辨率丢失的信息没有被恢复,但是得到的误差的明显程度更低。

[0072] 在一个实施例中,显示器和不透明度滤光器被同时渲染,并且被校准到用户在空间中的精确位置以补偿角度偏移问题。眼睛跟踪可用于计算视野的末端处的正确的图像偏移。在一些实施例中,在不透明度滤光器中可以使用不透明度的量在时间和空间上的淡化。类似地,在虚拟图像中可以使用在时间和空间上的淡化。在一个方案中,不透明度滤光器的不透明度的量在时间上的淡化对应于虚拟图像在时间上的淡化。在另一方案中,不透明度滤光器的不透明度的量在空间上的淡化对应于虚拟图像中在空间上的淡化。

[0073] 在一个示例性方案中,从所标识出的用户眼睛的位置的视角来看,为不透明度滤光器的处于虚拟图像之后的像素提供了增加的不透明度。通过这种方式,处于虚拟图像之后的像素被暗化,使得来自现实世界场景的相应部分的光被阻挡而无法到达用户的眼睛。这允许虚拟图像是逼真的并且表示全范围的色彩和亮度。此外,由于可以以更低的亮度提供虚拟图像,因此减少了增强现实发射器的功耗。在没有不透明度滤光器的情况下,将需要以与现实世界场景的相应部分相比更亮的足够高的亮度来提供虚拟图像,以使虚拟图像不同并且不是透明的。在暗化不透明度滤光器的像素时,一般而言,沿着虚拟图像的闭合周界的像素与周界内的像素一起被暗化。所期望的可能是提供一些重叠,使得恰好处于周界之外并且围绕周界的一些像素也被暗化 (以相同的暗度级或者比周界内的像素暗度低)。恰好处于周界之外的这些像素可以提供从周界内的暗度到周界之外的完全量的非暗度的淡化 (例如不透明度的逐渐过渡)。

[0074] 头戴式显示设备 2 还包括用于跟踪用户的眼睛位置的系统。如下面将会解释的那样,该系统将跟踪用户的位置和定向,使得该系统可以确定用户的视野。然而,人类将不会察觉处于其之前的所有事物。相反,用户的眼睛将对准环境的子集。因此,在一个实施例中,该系统将包括用于跟踪用户的眼睛位置的技术以便细化对用户视野的测量。例如,头戴式显示设备 2 包括眼睛跟踪部件 134 (图 3A),该眼睛跟踪部件 134 将包括眼睛跟踪照明设备 134A (参见图 4A) 和眼睛跟踪相机 134B (参见图 4A)。在一个实施例中,眼睛跟踪照明源

134A 包括一个或多个红外 (IR) 发射器,这些红外发射器向眼睛发射 IR 光。眼睛跟踪相机 134B 包括一个或多个感测所反射的 IR 光的相机。通过检测角膜的反射的已知成像技术,可以标识出瞳孔的位置。例如,参见 2008 年 7 月 22 日颁发给 Kranz 等人的、名称为“Head mounted eye tracking and display system(头戴式眼睛跟踪和显示系统)”的美国专利 7,401,920,该专利通过引用结合于此。这样的技术可以定位眼睛的中心相对于跟踪相机的位置。一般而言,眼睛跟踪涉及获得眼睛的图像以及使用计算机视觉技术来确定瞳孔在眼眶内的位置。在一个实施例中,跟踪一只眼睛的位置就足够了,因为眼睛通常一致地移动。然而,单独地跟踪每只眼睛是可能的。

[0075] 在一个实施例中,眼睛跟踪照明设备 134A 将使用 4 个 IR LED 并且眼睛跟踪相机 134 将使用 4 个 IR 光电检测器(未示出),所述 IR LED 和 IR 光电检测器以矩形布置使得在头戴式显示设备 2 的透镜的每个角处存在一个 IR LED 和 IR 光电检测器。来自 LED 的光从眼睛反射离开。由在 4 个 IR 光电检测器中的每个处所检测到的红外光的量来确定瞳孔方向。也就是说,眼睛中眼白相对于眼黑的量将确定对于该特定光电检测器而言从眼睛反射离开的光量。因此,光电检测器将具有对眼睛中的眼白或眼黑的量的度量。从这 4 个采样中,该系统可以确定眼睛的方向。

[0076] 另一替代方案是如下面所讨论的那样使用 4 个红外 LED,但是在头戴式显示设备 2 的透镜的侧边处仅仅使用一个红外 CCD。该 CCD 将使用小镜和 / 或透镜(鱼眼),使得 CCD 可以从镜框对高达 75% 的可见眼睛进行成像。然后,该 CCD 将感测图像并且使用计算机视觉来找出眼睛位置,就像下面所讨论的那样。因此,尽管图 3A-3C 示出了具有一个 IR 发射器的一个部件,但是图 2 的结构可以被调整为具有 4 个 IR 发射机和 / 或 4 个 IR 传感器。(注意:图像参考需要更新。在任何附图中都不能看出这一点)。也可以使用多于或少于 4 个的 IR 发射机和 / 或多于或少于 4 个的 IR 传感器。

[0077] 用于跟踪眼睛方向的另一实施例基于电荷跟踪。该方案基于如下观察:视网膜携带可测量的正电荷并且角膜具有负电荷。传感器通过用户的耳朵来安装(靠近耳机 130)以检测眼睛在转动时的电势并且有效地实时读出眼睛正在进行的动作。也可以使用其他用于跟踪眼睛的实施例。

[0078] 图 3B-3D 示出了头戴式显示设备 2 的一部分的备选实施例 2b 和 2c。在图 3B-3D 中,相同附图标记指代与图 3A 中所标识出的那些部分相同的部分。

[0079] 图 3B 示出了非透视头戴式显示设备 2b。图 3B 的显示设备 2b 使用朝向前方的透镜 133a,该透镜 133a 耦合到波导 124A 以将场景(比如环境 1100)的视图耦合到用户的眼睛 140。微显示器 153 可以包括任何前述显示器类型,比如 LCD、LED 或 OLED 等等,其具有由单独地启动的像素元件的阵列所定义的分辨率,这些像素的组合用于生成适于耦合到用户小凹的经优化图像。微显示器 153 可以耦合到多个微机电元件 160a,这些微机电元件 160a 耦合到该显示器的每个角处以在三维中相对于用户眼睛 140 对该显示器进行定位。因此,微显示器 153 可以具有绕显示器的中心点旋转的多个轴“Z”和“X”、以及相对于用户眼睛定位的垂直“V”和侧向“L”。

[0080] 如图 3B 所示,仅仅驱动显示器的对经优化图像的图像部分 1126(在该情况下为小鸟 1126)进行渲染的那些元件以提供高分辨率图像,使得用户的眼睛 140 的聚焦区直接耦合到来自图像 1126 的光。图像 1126 被部分 1126a 环绕,以示出:在图 3B 中仅仅渲染图像

1126、即用户在该环境中的整个视野的经优化图像 1200 的一部分。

[0081] 图 3c 和 3d 示出了本技术的另一备选实施例 2b。图 3D 是头戴式显示设备 2D 的顶视图,并且图 3C 是头戴式显示设备 2D 的侧视图。在图 3C 和 3D 中,头戴式显示设备 2c 包括支承结构 162、微机电元件 160、161、163(以及第四微机电元件——未示出)以及镜 166。一个或多个微显示器元件 170 被定位为与镜 166 相邻,其中元件 170 可以等价于参考图 3A 所述的显示器 120。镜 166 可以被微机电元件 160、161、163 相对于显示器结构 162 移动,以将来自微显示器元件的发射对象到用户眼睛的聚焦区中。微机电元件 160、161、163 可以包括压电元件、或者以机械方式控制或以机电方式控制的其他元件,这些元件在协作使用时可以沿着运动的三个轴相对于支承结构 162 对镜 166 进行定位。以类似于微显示器 153 的方式,这些微机电元件 160-163 耦合到镜的每个角处以在三维中相对于用户眼睛 140 对该镜进行定位。因此,镜 166 可以具有绕显示器的中心点旋转的多个轴“Z”和“X”、以及相对于用户眼睛定位的垂直“Y”和侧向“L”。应当认识到,镜 166 的移动可以单独使用或者与微显示器元件的方向输出相结合使用以将图像的经优化部分(在该示例中为小鸟图像 1126)定位在用户的聚焦区中。

[0082] 图 3A-3D 仅仅示出了头戴式显示设备 2a-2c 的一半。完整的头戴式显示设备(在适用时)将包括另一组透视透镜、另一不透明度滤光器、另一光导光学元件、另一微显示器、另一透镜、朝向房间的相机、眼睛跟踪部件、微显示器、耳机以及温度传感器。

[0083] 图 4A 是描绘了头戴式显示设备 2a-2c 的各个组件的框图。参考图 3A-3D 能够理解,图 4A 中所示的一些组件可以不存在于图 3A-3D 所示的每个实施例中。图 4B 是描述处理单元 4 的各个组件的框图。图 4A 中示出了每个头戴式显示设备 2 的组件,所述头戴式显示设备 2 用于向用户显示经优化的图像。附加地,图 4A 的头戴式显示设备组件包括跟踪各种状况的许多传感器。头戴式显示设备 2a-2c 将从处理单元 4 接收关于虚拟图像的指令,并且将传感器信息提供回给处理单元 4。图 4B 中描绘了处理单元 4 的组件,该处理单元 4 将从头戴式显示设备 2a-2c 并且还从中枢计算设备 12(参见图 1)接收传感信息。基于该信息,处理单元 4 将确定在何处以及在何时向用户提供虚拟图像并相应地将指令发送给图 4A 的头戴式显示设备。

[0084] 注意,图 4A 的组件中的一些(例如朝向背面的相机 113、眼睛跟踪相机 134B、微显示器 120 或 153、不透明度滤光器 114、眼睛跟踪照明 134A、和耳机 130)是以阴影示出的,以指示这些的设备中的每个都存在两个,其中一个用于头戴式显示设备 2 的左侧,并且一个用于头戴式显示设备 2 的右侧。图 4A 示出与电源管理电路 202 通信的控制电路 200。控制电路 200 包括处理器 210、与存储器 214(例如 D-RAM)进行通信的存储器控制器 212、相机接口 216、相机缓冲区 218、显示器驱动器 220、显示格式化器 222、定时生成器 226、显示输出接口 228、以及显示输入接口 230。在一个实施例中,控制电路 200 的所有组件都通过专用线路或一个或多个总线彼此进行通信。在另一实施例中,控制电路 200 的每个组件都与处理器 210 通信。相机接口 216 提供到两个朝向房间的相机 113 的接口,并且将从朝向房间的相机所接收到的图像存储在相机缓冲器 218 中。显示器驱动器 220 将驱动微显示器 120 或 153。显示格式化器 222 向控制不透明度滤光器 114 的不透明度控制电路 224 提供关于微显示器 120 或 153 上所显示的虚拟图像的信息。定时生成器 226 被用于向该系统提供定时数据。显示输出接口 228 是用于将图像从朝向房间的相机 113 提供给处理单元 4

的缓冲器。显示输入 230 是用于接收诸如要在微显示器 120 上显示的虚拟图像之类的图像的缓冲区。显示输出 228 和显示输入 230 与作为到处理单元 4 的接口的带接口 232 进行通信。显示器驱动器 220 还可以驱动镜控制器 162 对镜 166 进行定位,以根据图 3C 和 3D 的上述实施例显示聚焦的图像。

[0085] 电源管理电路 202 包括电压调节器 234、眼睛跟踪照明驱动器 236、音频 DAC 和放大器 238、话筒前置放大器音频 ADC 240、温度传感器接口 242、以及时钟生成器 244。电压调节器 234 通过带接口 232 从处理单元 4 接收电能,并将该电能提供给头戴式显示设备 2 的其他组件。每个眼睛跟踪照明驱动器 236 都如上面所述的那样为眼睛跟踪照明 134A 提供 IR 光源。音频 DAC 和放大器 238 从耳机 130 接收音频信息。话筒前置放大器和音频 ADC 240 提供用于话筒 110 的接口。温度传感器接口 242 是用于温度传感器 138 的接口。电源管理单元 202 还向三轴磁力计 132A、三轴陀螺仪 132B 以及三轴加速度计 132C 提供电能并从中接收回数据。

[0086] 图 4B 是描述处理单元 4 的各个组件的框图。图 4B 示出控制电路 304 与电源管理电路 306 进行通信。控制电路 304 包括:中央处理单元 (CPU) 320;图形处理单元 (GPU) 322;高速缓存 324;RAM 326;与存储器 330(例如 D-RAM)进行通信的存储器控制器 328;与闪存 334(或其他类型的非易失性存储)进行通信的闪存控制器 332;通过带接口 302 和带接口 232 与头戴式显示设备 2 进行通信的显示输出缓冲区 336;经由带接口 302 和带接口 232 与头戴式显示设备 2 进行通信的显示输入缓冲区 338;与用于连接到话筒的外部话筒连接器 342 进行通信的话筒接口 340,用于连接到无线通信设备 346 的 PCIexpress 接口;以及 USB 端口 348。在一个实施例中,无线通信设备 346 可以包括启用 Wi-Fi 的通信设备、蓝牙通信设备、红外通信设备等。USB 端口可以用于将处理单元 4 对接到中枢计算设备 12,以便将数据或软件加载到处理单元 4 上以及对处理单元 4 进行充电。在一个实施例中,CPU 320 和 GPU 322 是用于确定在何处、何时以及如何向用户的视野内插入虚拟图像的主负荷设备。下面提供更多的细节。

[0087] 电源管理电路 306 包括时钟生成器 360、模数转换器 362、电池充电器 364、电压调节器 366、头戴式显示器电源 376、以及与温度传感器 374 进行通信的温度传感器接口 372(其位于处理单元 4 的腕带 (wrist band) 上)。模数转换器 362 连接到充电插座 370 以用于接收 AC 供电并为该系统产生 DC 供电。电压调节器 366 与用于向该系统提供电能的电池 368 进行通信。电池充电器 364 被用来在从充电插孔 370 接收到电能时对电池 368 进行充电(通过电压调节器 366)。HMD 电源 376 向头戴式显示设备 2 提供电能。

[0088] 上述系统将被配置为将虚拟图像插入到用户的视野中,使得该虚拟图像替换现实世界物体的视图。替代地,可插入虚拟图像而不替换现实世界物体的图像。在各个实施例中,虚拟图像将基于被替换的物体或者该图像将被插入的环境而被调整为与合适的定向、大小和形状相匹配。另外,虚拟图像可以被调整为包括反射和阴影。在一个实施例中,头戴式显示设备 12、处理单元 4 以及中枢计算设备 12 一起工作,因为每个设备都包括用于获得用于确定在何处、何时以及如何插入虚拟图像的数据的传感器的子集。在一个实施例中,确定在何处、如何以及何时插入虚拟图像的计算是由中枢计算设备 12 执行的。在另一实施例中,这些计算由处理单元 4 来执行。在另一实施例中,这些计算中的一些由中枢计算设备 12 来执行,而其他计算由处理单元 4 来执行。在其他实施例中,这些计算可以由头戴式显示设

备 12 来执行。

[0089] 在一个示例性实施例中,中枢计算设备 12 将创建用户所处的环境的模型,并且跟踪在该环境中的多个移动物体。另外,中枢计算设备 12 通过跟踪头戴式显示设备 2 的位置和定向来跟踪头戴式显示设备 2 的视野。该模型和跟踪信息被从中枢计算设备 12 提供给处理单元 4。由头戴式显示设备 2 所获得的传感器信息被传送给处理单元 4。然后,处理单元 4 使用其从头戴式显示设备 2 接收的其他传感器信息来细化用户的视野并且向头戴式显示设备 2 提供关于如何、在何处以及何时插入虚拟图像的指令。

[0090] 图 5 示出了具有捕捉设备的中枢计算系统 12 的示例性实施例。在一个实施例中,捕捉设备 20A 和 20B 是相同结构,因此,图 5 仅仅示出了捕捉设备 20A。根据一示例性实施例,捕捉设备 20A 可被配置为通过可包括例如飞行时间、结构化光、立体图像等在内的任何合适的技术来捕捉包括深度图像的带有深度信息的视频,该深度图像可包括深度值。根据一个实施例,捕捉设备 20A 可将深度信息组织成“Z 层”,即可与从深度相机沿其视线延伸的 Z 轴垂直的层。

[0091] 如图 5 所示,捕捉设备 20A 可以包括相机组件 423。根据一示例性实施例,相机组件 423 可以是或者可以包括可捕捉场景的深度图像的深度相机。深度图像可包括所捕捉的场景的二维 (2-D) 像素区域,其中 2-D 像素区域中的每个像素都可以表示深度值,比如所捕捉的场景中的物体与相机相距的例如以厘米、毫米等单位的距离。

[0092] 相机组件 23 可以包括可用于捕捉场景的深度图像的红外 (IR) 光组件 425、三维 (3D) 相机 426、以及 RGB (视觉图像) 相机 428。例如,在飞行时间分析中,捕捉设备 20A 的 IR 光组件 425 可以将红外光发射到场景上,并且然后可以使用传感器 (在一些实施例中包括未示出的传感器)、例如使用 3D 相机 426 和 / 或 RGB 相机 428 来检测从场景中的一个或多个目标和物体的表面后向散射的光。在一些实施例中,可以使用脉冲红外光,使得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其用于确定从捕捉设备 20A 到场景中的目标或物体上的特定位置的物理距离。附加地,在其他示例实施例中,可将出射光波的相位与入射光波的相位进行比较来确定相移。然后可以使用该相移来确定从捕捉设备到目标或物体上的特定位置的物理距离。

[0093] 根据另一示例性实施例,可使用飞行时间分析,以通过经由包括例如快门式光脉冲成像之类的各种技术分析反射光束随时间的强度来间接地确定从捕捉设备 20A 到目标或物体上的特定位置的物理距离。

[0094] 在另一示例性实施例中,捕捉设备 20A 可使用结构化光来捕捉深度信息。在这样的分析中,图案化光 (即,被显示为诸如网格图案、条纹图案、或不同图案之类的已知图案的光) 可经由例如 IR 光组件 424 被投影到场景上。在落到场景中的一个或多个目标或物体的表面上时,作为响应,图案可变形。图案的这种变形可由例如 3D 相机 426 和 / 或 RGB 相机 428 (和 / 或其他传感器) 来捕捉,然后可被分析以确定从捕捉设备到目标或物体上的特定位置的物理距离。在一些实施方式中,IR 光组件 425 与相机 425 和 426 分开,使得可以使用三角测量来确定与相机 425 和 426 相距的距离。在一些实施方式中,捕捉设备 20A 将包括感测 IR 光的专用 IR 传感器或具有 IR 滤光器的传感器。

[0095] 根据另一实施例,捕捉设备 20A 可以包括两个或更多个在物理上分开的相机,这些相机可以从不同的角度观察场景以获得视觉立体数据,这些视觉立体数据可以被分辨以

生成深度信息。也可使用其他类型的深度图像传感器来创建深度图像。

[0096] 捕捉设备 20A 还可以包括话筒 430, 所述话筒 430 包括可以接收声音并将其转换成电信号的换能器或传感器。话筒 430 可用于接收也可由中枢计算系统 12 来提供的音频信号。

[0097] 在一示例实施例中, 捕捉设备 20A 还可包括可与图像相机组件 423 进行通信的处理器 432。处理器 432 可包括可执行指令的标准处理器、专用处理器、微处理器等, 这些指令例如包括用于接收深度图像、生成合适的数据格式 (例如, 帧) 以及将数据传送给中枢计算系统 12 的指令。

[0098] 捕捉设备 20A 还可包括存储器 434, 该存储器 434 可存储由处理器 432 执行的指令、由 3D 相机和 / 或 RGB 相机所捕捉的图像或图像帧、或任何其他合适的信息、图像等等。根据一示例性实施例, 存储器 434 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 5 所示, 在一个实施例中, 存储器 434 可以是与图像捕捉组件 423 和处理器 432 进行通信的单独组件。根据另一实施例, 存储器组件 434 可被集成到处理器 432 和 / 或图像捕捉组件 422 中。

[0099] 捕捉设备 20A 和 20B 通过通信链路 436 与中枢计算系统 12 通信。通信链路 436 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 连接等的无线连接。根据一个实施例, 中枢计算系统 12 可以通过通信链路 436 向捕捉设备 20A 提供可用于确定例如何时捕捉场景的时钟。附加地, 捕捉设备 20A 通过通信链路 436 将由例如 3D 相机 426 和 / 或 RGB 相机 428 捕捉的深度信息和视觉 (例如 RGB) 图像提供给中枢计算系统 12。在一个实施例中, 深度图像和视觉图像以每秒 30 帧的速率来传送, 但是可以使用其他帧速率。中枢计算系统 12 然后可以创建模型并使用模型、深度信息、以及所捕捉的图像来例如控制诸如游戏或文字处理程序等的应用和 / 或使化身或屏上人物动画化。

[0100] 中枢计算系统 12 包括深度图像处理和骨架跟踪模块 450, 该模块使用深度图像来跟踪可被捕捉设备 20A 的深度相机功能检测到的一个或多个人。深度图像处理和骨架跟踪模块 450 向应用 453 提供跟踪信息, 该应用可以是视频游戏、生产性应用、通信应用或其他软件应用等。音频数据和视觉图像数据也被提供给应用 452 和深度图像处理和骨架跟踪模块 450。应用 452 将跟踪信息、音频数据和视觉图像数据提供给识别器引擎 454。在另一实施例中, 识别器引擎 454 直接从深度图像处理和骨架跟踪模块 450 接收跟踪信息, 并直接从捕捉设备 20A 和 20B 接收音频数据和视觉图像数据。

[0101] 识别器引擎 454 与过滤器 460、462、464、……、466 的集合相关联, 每个过滤器都包括关于可被捕捉设备 20A 或 20B 检测到的任何人或物体执行的姿势、动作或状况的信息。例如, 来自捕捉设备 20A 的数据可由过滤器 460、462、464、……、466 来处理, 以便标识出一个用户或一组用户已经何时执行了一个或多个姿势或其他动作。这些姿势可与应用 452 的各种控制、物体或状况相关联。因此, 中枢计算系统 12 可以将识别器引擎 454 和过滤器一起用于解释和跟踪物体 (包括人) 的移动。

[0102] 捕捉设备 20A 和 20B 向中枢计算系统 12 提供 RGB 图像 (或其他格式或色彩空间的视觉图像) 和深度图像。深度图像可以是多个观测到的像素, 其中每个观测到的像素具有观测到的深度值。例如, 深度图像可包括所捕捉的场景的二维 (2D) 像素区域, 其中 2D

像素区域中的每个像素都可具有深度值,比如所捕捉的场景中的物体与捕捉设备相距的距离。中枢计算系统 12 将使用 RGB 图像和深度图像来跟踪用户或物体的移动。例如,系统将使用深度图像来跟踪人的骨架。可以使用许多方法以通过使用深度图像来跟踪人的骨架。使用深度图像来跟踪骨架的一个合适的示例在 Craig 等人 2009 年 10 月 21 日提交的美国专利申请 12/603,437 “Pose Tracking Pipeline(姿态跟踪流水线)”(以下称为‘437 申请)中提供,该申请的全部内容通过引用结合于此。‘437 申请的过程包括:获得深度图像;对数据进行降采样;移除和/或平滑化高方差噪声数据;标识并移除背景;以及将前景像素中的每个分配给身体的不同部位。基于这些步骤,系统将使一模型拟合到该数据并创建骨架。该骨架将包括一组关节和这些关节之间的连接。也可使用用于跟踪的其他方法。在下列四个美国专利申请中还公开了合适的跟踪技术,所述专利的全部内容都通过引用结合于此:于 2009 年 5 月 29 日提交的美国专利申请 12/475,308 “Device for Identifying and Tracking Multiple Humans Over Time(用于随时间标识和跟踪多个人类的设备)”;于 2010 年 1 月 29 日提交的美国专利申请 12/696,282 “Visual Based Identity Tracking(基于视觉的身份跟踪)”;于 2009 年 12 月 18 日提交美国专利申请 12/641,788 “Motion Detection Using Depth Images(使用深度图像的运动检测)”;以及于 2009 年 10 月 7 日提交的美国专利申请 12/575,388 “Human Tracking System(人类跟踪系统)”。

[0103] 识别器引擎 454 包括多个过滤器 460、462、464、……、466 来确定姿势或动作。过滤器包括定义姿势、动作或状况以及该姿势、动作或状况的参数或元数据的信息。例如,包括一只手从身体背后经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经过身体前方的运动的信息的姿势,因为该运动将由深度相机来捕捉。然后可为该姿势设定参数。当姿势是投掷时,参数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了该姿势的置信度评级。用于姿势的这些参数可以随时间在各应用之间、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变化。

[0104] 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)。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架构的任何其他方面。例如,可能具有用于驱动的第一过滤器,该第一过滤器将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和转向角。在希望用第二驱动过滤器来替换该第一驱动过滤器的情况下(这可能是因为第二驱动过滤器更高效且需要更少的处理资源),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同样的输入和输出——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。

[0105] 过滤器不必具有参数。例如,返回用户的高度的“用户高度”过滤器可能不允许可被调节的任何参数。备选的“用户高度”过滤器可具有可调节参数,比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。

[0106] 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、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。

[0107] 识别器引擎 454 可以具有向过滤器提供功能的基本识别器引擎。在一个实施例中,识别器引擎 454 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;隐马尔可夫模型实施方式(其中所建模的系统被假定为马尔可夫过程—其中当前状态封装了用于确定将来状态的任何过去状态信息,因此不必为此目的而维护任何其他过去状态信息的过程—该过程具有未知参数,并且隐藏参数是从可观察数据来确定的);以及求解姿势识别的特定实例的其他功能。

[0108] 过滤器 460、462、464、……、466 在识别器引擎 454 之上加载并实现,并且可利用识别器引擎 454 提供给所有过滤器 460、462、464、……、466 的服务。在一个实施例中,识别器引擎 454 接收数据来确定该数据是否满足任何过滤器 460、462、464、……、466 的要求。由于这些所提供的诸如解析输入之类的服务是由识别器引擎 454 一次性提供而非由每个过滤器 460、462、464、……、466 提供的,因此这样的服务在一段时间内只需被处理一次而不是在该时间段对每个过滤器处理一次,因此减少了确定姿势所需的处理。

[0109] 应用 452 可使用识别器引擎 454 所提供的过滤器 460、462、464、……、466,或者它可提供其自己的、插入到识别器引擎 454 中的过滤器。在一实施例中,所有过滤器具有启用该插入特性的通用接口。此外,所有过滤器可利用参数,因此可使用以下单个姿势工具来诊断并调节整个过滤器系统。

[0110] 关于识别器引擎 454 的更多信息可在 2009 年 4 月 13 日提交的美国专利申请 12/422,661 “Gesture Recognizer System Architecture(姿势识别器系统架构)”中找到,该申请通过引用整体并入本文。关于识别姿势的更多信息可在 2009 年 2 月 23 日提交的美国专利申请 12/391,150 “Standard Gestures(标准姿势)”;以及 2009 年 5 月 29 日提交的美国专利申请 12/474,655 “Gesture Tool(姿势工具)”中找到,这两个申请的全部内容都通过引用结合于此。

[0111] 在一个实施例中,计算系统 12 包括用户简档数据库 470,该用户简档数据库 470 包括与同中枢计算系统 12 交互的一个或多个用户相关的用户专用信息。在一个示例中,用户专用信息包括诸如下列与用户相关的信息:用户所表达的偏好;用户的朋友的列表;用户所偏好的活动;用户的提醒列表;用户的社交群;用户的当前位置;用户过去与该用户的环境中的物体交互的意图;以及其他用户创建的内容,比如用户的照片、图像和所记录的视频。在一个实施例中,用户专用信息可以从一个或多个诸如下列数据源获得:用户的社交站点、地址簿、电子邮件数据、即时通讯数据、用户简档或者因特网上的其他源。在一个方案中并且如将在下面详细描述的那样,用户专用信息用于自动确定用户与该用户的环境中的一个或多个物体交互的意图。

[0112] 图 6 示出了可用于实现中枢计算系统 12 的计算系统的示例性实施例。如图 6 所示,多媒体控制台 500 具有含有一级高速缓存 502、二级高速缓存 504 和闪存 ROM(只读存储器)506 的中央处理单元(CPU)501。一级高速缓存 502 和二级高速缓存 504 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。CPU 501 可以被配备为具有一个以上的核,并且由此具有附加的 1 级和 2 级高速缓存 502 和 504。闪存 ROM 506 可存储在多媒体控制台 500 通电时在引导过程初始化阶段加载的可执行代码。

[0113] 图形处理单元(GPU)508 和视频编码器/视频编解码器(编码器/解码器)514 形

成用于高速和高分辨率图形处理的视频处理流水线。经由总线从图形处理单元 508 向视频编码器 / 视频编解码器 514 运送数据。视频处理流水线向 A/V (音频 / 视频) 端口 540 输出数据,用于传输至电视或其他显示器。存储器控制器 510 连接到 GPU 508 以方便处理器访问各种类型的存储器 512,诸如但不局限于 RAM (随机存取存储器)。

[0114] 多媒体控制台 500 包括优选地在模块 518 上实现的 I/O 控制器 520、系统管理控制器 522、音频处理单元 523、网络接口 524、第一 USB 主控制器 526、第二 USB 控制器 528 以及前面板 I/O 子部件 530。USB 控制器 526 和 528 用作外围控制器 542(1)-542(2)、无线适配器 548、和外置存储器设备 546 (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 524 和 / 或无线适配器 548 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。

[0115] 提供系统存储器 543 来存储在引导过程期间加载的应用数据。提供媒体驱动器 544,且其可包括 DVD/CD 驱动器、蓝光驱动器、硬盘驱动器、或其他可移动媒体驱动器等。媒体驱动器 144 可位于多媒体控制台 500 的内部或外部。应用数据可经由媒体驱动器 544 访问,以由多媒体控制台 500 执行、回放等。媒体驱动器 544 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 520。

[0116] 系统管理控制器 522 提供涉及确保多媒体控制台 500 的可用性的各种服务功能。音频处理单元 523 和音频编解码器 532 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 523 与音频编解码器 532 之间传输。音频处理流水线将数据输出到 A/V 端口 540 以供外部音频用户或具有音频能力的设备再现。

[0117] 前面板 I/O 子部件 530 支持暴露在多媒体控制台 100 的外表面上的电源按钮 550 和弹出按钮 552 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 536 向多媒体控制台 100 的组件供电。风扇 538 冷却多媒体控制台 500 内的电路。

[0118] CPU 501、GPU 508、存储器控制器 510、和多媒体控制台 500 内的各种其他组件经由一条或多条总线互连,总线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。

[0119] 当多媒体控制台 500 通电时,应用数据可从系统存储器 543 加载到存储器 512 和 / 或高速缓存 502、504 中并在 CPU 501 上执行。应用可呈现在导航到多媒体控制台 500 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,媒体驱动器 544 中包含的应用和 / 或其他媒体可从媒体驱动器 544 启动或播放,以向多媒体控制台 500 提供附加功能。

[0120] 多媒体控制台 500 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 500 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 524 或无线适配器 548 可用的宽带连接的集成,多媒体控制台 500 还可作为更大网络社区中的参与者来操作。附加地,多媒体控制台 500 可以通过无线适配器 548 与处理单元 4 通信。

[0121] 当多媒体控制台 500 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、CPU 和 GPU 周期、网络带宽等等的保留。因为这

些资源是在系统引导时保留的,所以所保留的资源从应用的角度而言是不存在的。具体而言,存储器保留优选地足够大,以包含启动内核、并发系统应用和驱动程序。CPU 保留优选地为恒定,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。

[0122] 对于 GPU 保留,显示由系统应用程序生成的轻量消息(例如,弹出窗口),所述显示是通过使用 GPU 中断来调度代码以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,并且覆盖图优选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(scaler)可用于设置该分辨率,从而消除了对改变频率并引起 TV 重新同步的需求。

[0123] 在多媒体控制台 500 引导且系统资源被保留之后,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 501 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中断最小化。

[0124] 当并发系统应用需要音频时,则由于时间敏感性而将音频处理异步地调度给游戏应用。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)。

[0125] 任选的输入设备(例如,控制器 542(1) 和 542(2))由游戏应用和系统应用共享。输入设备不是保留资源,而是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。应用管理器优选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。捕捉设备 20A 和 20B 可以通过 USB 控制器 526 或其他接口来为控制台 500 定义附加的输入设备。在其他实施例中,中枢计算系统 12 可以使用其他硬件架构来实现。没有一种硬件架构是必需的。

[0126] 图 1 示出了与一个中枢处理设备 12(称为中枢)通信的一个头戴式显示设备 2 和处理单元 4(统称为移动显示设备)。在另一实施例中,多个移动显示设备可以与单个中枢通信。每个移动显示设备都将如上述那样使用无线通信与中枢通信。在这样的实施例中所构思的是,有益于所有移动显示设备的信息中的许多都将在中枢处被计算和存储并且传送给每个移动显示设备。例如,中枢将生成环境的模型并且将该模型提供给与该中枢通信的所有移动显示设备。附加地,中枢可以跟踪移动显示设备以及房间中的移动物体的位置和定向,并且然后将该信息传输给每个移动显示设备。

[0127] 在另一实施例中,系统可以包括多个中枢,其中每个中枢都包括一个或多个移动显示设备。这些中枢可以直接地或者通过因特网(或者其他网络)彼此通信。例如,图 7 示出了中枢 560、562 和 564。中枢 560 直接与中枢 562 通信。中枢 560 通过因特网与中枢 564 通信。中枢 560 与移动显示设备 570、572... 574 通信。中枢 562 与移动显示设备 578、580... 582 通信。中枢 564 与移动显示设备 584、586... 588 通信。每个移动显示设备都如上面所讨论的那样通过无线通信与其相应的中枢通信。如果这些中枢处于共同的环境中,则每个中枢都可以提供该环境的模型的一部分,或者一个中枢可以为其他中枢创建该模型。每个中枢都将跟踪移动物体的子集并且将该信息与其他中枢共享,所述其他中枢进而将与合适的移动显示设备共享该信息。用于移动显示设备的传感器信息将被提供给她相应的中

枢,并且然后与其他中枢共享以用于最终与其他移动显示设备共享。因此,在中枢间共享的信息可以包括骨架跟踪、关于模型的信息、各种应用状态、以及其他跟踪。在中枢及其相应的移动显示设备之间传递的信息包括:移动物体的跟踪信息、世界模型的状态和物理更新、几何和纹理信息、视频和音频、以及用于执行在此所述的操作的其他信息。

[0128] 图 8 示出了用于优化被呈现给头戴式显示设备的用户的可视信息的显示的过程的一个实施例。图 8 示出了执行上面图 2A 中的步骤 32 的一个实施例。

[0129] 在步骤 600,配置系统 10。例如,应用(例如图 5 的应用 452)可以将该系统配置为指示:经优化的图像将被插入到场景的三维模型中的指定位置处。在另一示例中,在中枢计算系统 12 上运行的应用将指示:扩展内容(比如特定的虚拟图像或虚拟物体)将作为视频游戏或其他进程的一部分被插入到该场景中。

[0130] 在步骤 602,该系统将创建头戴式显示设备 2 所处的空间的体积模型。例如在一个实施例中,中枢计算设备 12 将使用来自一个或多个深度相机的深度图像来创建头戴式显示设备 2 所处的环境或场景的三维模型。在步骤 604,将该模型分割成一个或多个物体。例如,如果中枢计算设备 12 创建房间的三维模型,则该房间很可能在里面具有多个物体。可以处于房间中的物体的示例包括:人、椅子、桌子、沙发椅等等。步骤 604 包括:确定彼此不同的物体。在步骤 606,该系统将标识出这些物体。例如,中枢计算设备 12 可以标识出:特定物体是桌子并且另一物体是椅子。

[0131] 应当认识到,尽管在一个实施例中创建体积模型和标识出物体可以与本技术一起使用,但是步骤 602-608 可以在备选的实施例中省略。在备选的实施例中,经优化图像的生成可以在不参考环境 1100 的情况下进行并且可以包括:提供覆盖图像以供在不参考周围环境的情况下使用。也就是说,本技术不需要使用覆盖图像,并且可以有利地用于仅仅向用户显示图像而不参考用户的环境。

[0132] 在图 8 的步骤 608,该系统基于用户空间的模型来确定用户的视野。在一个实施例中,步骤 608 等价于图 2A 的步骤 32。也就是说,该系统确定用户正在查看的环境或空间。在一个实施例中,步骤 608 可以使用中枢计算设备 12、处理单元 4 和 / 或头戴式显示设备 2 来执行。在一个示例性实施方式中,中枢计算机设备 12 将跟踪用户和头戴式显示设备 2 以便提供对头戴式显示设备 2 的位置和定向的初步确定。头戴式显示设备 2 上的传感器将用于细化所确定的定向。例如,上述惯性传感器 34 可以用于细化头戴式显示设备 2 的定向。附加地,可以使用下面所述的眼睛跟踪过程来标识出最初所确定的视野的与用户具体正在查看之处相对应的子集(另称为视野中的用户聚焦区或深度焦点)。下面将参照图 11-13 来描述更多细节。

[0133] 在步骤 610,诸如在处理单元 4 中执行的软件之类的系统确定用户在该用户的视野内的当前聚焦区。在一个实施例中,步骤 610 等价于图 2A 的步骤 34。如后面在图 12 和 13 中将进一步讨论的那样,基于由眼睛跟踪相机 134 针对每只眼睛所捕捉的数据的眼睛跟踪处理可以提供用户的当前聚焦区。例如,在有数据指示用户脸的位置的情况下,瞳孔之间的会聚可以用于对聚焦曲线上的焦点(即,双眼单视界)进行三角测定,从该焦点中可以计算出聚焦区(即,Panum 汇合区域)。Panum 汇合区域是人眼所使用的代表双眼立体视觉的单视力区域。

[0134] 在步骤 612,在软件的控制下,处理单元 4 单独地或者与中枢计算设备 12 协作地生

成经优化的图像。经优化的图像基于三维模型、处于视野内的已经被检测到的物体、以及用户的视野。

[0135] 取决于控制经优化图像的生成的应用,经优化的图像可以采取许多形式。另外,应当理解,术语图像可以包括运动图像——对所显示的一个或多个物体的运动进行显示的图像。

[0136] 然后,头戴式显示设备 2 的用户基于在头戴式显示设备 2 中所显示的经优化图像来与在中枢计算设备 12(或另一计算设备)上运行的应用交互。图 8 的处理步骤(608-612)根据图 2A 在该系统的操作期间重复,使得用户的视野和聚焦区随着用户移动他或她的头而被更新,来自新视野的新的经优化图像被确定并且经优化的图像基于用户的意图被显示给该用户。下面更详细地描述步骤 604-612。

[0137] 图 9 描述了用于创建用户的空间的模型的过程的一个实施例。例如,图 9 的过程是图 8 的步骤 602 的一个示例性实施方式。在步骤 620,中枢计算系统 12 接收头戴式显示设备所处的环境的多个视角(如图 1 所示)的一个或多个深度图像。例如,中枢计算设备 12 可以从多个深度相机获得深度图像,或者通过使相机指向不同方向或者使用具有如下透镜的相机来从同一相机获得多个深度图像:该透镜允许将要构建模型的环境或空间的全视图。在步骤 622,基于共同的坐标系来组合来自各个深度图像的深度数据。例如,如果该系统从多个相机接收深度图像,则该系统使两个图像相关以具有共同的坐标系(例如使图像排齐)。在步骤 624,使用深度数据来创建空间的体积描述。

[0138] 图 10 是描述了用于将空间的模型分割成各个物体的过程的一个实施例的流程图。例如,图 10 的过程是图 8 的步骤 604 的一个示例性实施方式。在图 10 的步骤 626,该系统将如上面所讨论的那样从一个或多个深度相机接收一个或多个深度图像。可替代地,该系统可以访问其已经接收到的一个或多个深度图像。在步骤 628,该系统将如上述那样从相机接收一个或多个视觉图像。可替代地,该系统可以访问已经接收到的一个或多个视觉图像。在步骤 630,中枢计算系统将基于深度图像和/或视觉图像检测一个或多个人。例如,该系统将识别一个或多个骨架。在步骤 632,中枢计算设备将基于深度图像和/或视觉图像检测该模型内的边缘。在步骤 634,中枢计算设备将使用所检测到的边缘来标识出彼此不同的物体。例如,假定:这些边缘是物体之间的边界。在步骤 636,将更新使用图 9 的过程所创建的模型以显示出:该模型的哪些部分与不同的物体相关联。

[0139] 图 11 是描述了如下过程的一个实施例的流程图:用于确定用户视野的过程,该过程是图 8 的步骤 608 的示例性实施方式;以及用于确定用户聚焦区的过程,该过程是图 8 的步骤 610 的示例性实施方式。图 11 的过程依靠来自中枢计算设备 12 的信息和上述眼睛跟踪技术。图 12 是描述了由中枢计算系统执行以提供用在图 12 的过程中的跟踪信息的过程的一个实施例的流程图。可替代地,图 12 的过程可以由图 4A 的处理器 210 来执行。图 13 是描述了用于跟踪眼睛的过程的一个实施例的流程图,其中该过程的结果供图 12 的过程来使用。

[0140] 在使用中枢计算系统的情况下,在图 12 的步骤 686,中枢计算设备 12 将跟踪用户的位置。例如,中枢计算设备 12 将使用一个或多个深度图像和一个或多个视觉图像来跟踪用户(例如使用骨架跟踪)。可以在步骤 688 使用一个或多个深度图像和一个或多个视觉图像来确定头戴式显示设备 2 的位置和头戴式显示设备 2 的定向。在步骤 690,将用户和头

戴式显示设备 2 的位置和定向从中枢计算设备 12 传送给处理单元 4。在步骤 692, 在处理单元 4 处接收该位置和定向信息。图 12 的处理步骤可以在系统的操作期间连续地执行, 使得用户被连续地跟踪。

[0141] 图 13 是描述用于跟踪用户在环境中的眼睛位置的一个实施例的流程图。在步骤 662, 照亮眼睛。例如, 可以使用来自眼睛跟踪照明 134A 的红外光来照亮眼睛。在步骤 664, 使用一个或多个眼睛跟踪相机 134B 来检测来自眼睛的反射。在步骤 665, 将该反射数据从头戴式显示设备 2 发送给处理单元 4。在步骤 668, 处理单元 4 将如上述那样基于反射数据来确定眼睛的位置。

[0142] 图 11 是描述了用于确定用户的视野 (例如图 8 的步骤 608) 以及用户的聚焦区 (例如图 8 的步骤 610) 的过程的一个实施例的流程图。在步骤 670, 处理单元 4 将访问从中枢接收到的最新的位置和定向信息。图 12 的过程可以如从步骤 686 到步骤 690 的箭头所描绘的那样连续地执行, 因此, 处理单元 4 将周期性地从中枢计算设备 12 接收经更新的位置和定向信息。然而, 处理单元 4 将需要与其从中枢计算设备 12 接收经更新的信息相比更频繁地绘制虚拟图像。因此, 处理单元 4 将需要依靠本地感测的信息 (其例如来自头戴式设备 2) 以在来自中枢计算设备 12 的各采样之间提供定向的更新。另外, 处理等待时间还需要对虚拟图像进行快速渲染。

[0143] 可替代地, 步骤 670 可以通过任何数目的手段来执行。可以使用嵌入在头戴式显示器中的包括加速度计、磁力计和陀螺仪的传感器技术或其他传感器技术来标识出环境中的用户位置和定向。在步骤 672, 处理单元 4 将从三轴陀螺仪 132B 访问数据。在步骤 674, 处理单元 4 将从三轴加速度计 132C 访问数据。在步骤 676, 处理单元 4 将从三轴磁力计 132A 访问数据。在步骤 678, 处理单元 4 将来自该陀螺仪、加速度计和磁力计的数据来细化 (或以其他方式更新) 来自中枢计算设备 12 的位置和定向数据。在步骤 680, 处理单元 4 将基于头戴式显示设备的位置和定向来确定潜在的视角。使用任何数目的技术来确定头戴式显示器的位置并且将该位置与眼睛位置跟踪结合使用以确定用户的视野。注意在一些实施方式中, 不需要用户环境的三维模型。可以使用任何数目的用于头部跟踪的技术。在传感器从头戴式显示器可用的条件下, 可以使用来自加速度计和陀螺仪的惯性感测用户惯性测量结果。然而, 可以使用其它技术。这样的技术可以包括飞行时间、空间扫描、机械联动、相位差感测和 / 或直接场感测。在这些情况下, 在头戴式显示器中可能需要附加的硬件。

[0144] 在步骤 682, 处理单元 4 将访问最新的眼睛位置信息。在步骤 684, 处理单元 4 将基于眼睛位置作为潜在视角的子集来确定用户所观看的模型部分。例如, 用户可以朝向墙壁, 并且因此头戴式显示器的视点可以包括沿着墙壁的任何地方。然而, 如果用户的眼睛指向右边, 则步骤 684 将作出的结论是, 用户的视野仅仅是墙壁的右边部分。在步骤 684 结束时, 处理单元 4 已经确定了用户通过头戴式显示器 2 的视角。然后, 处理单元 4 可标识该视野内的某位置以插入虚拟图像并使用不透明度滤光器来阻挡光。图 12 的处理步骤可以在系统的操作期间连续执行, 使得用户的视野和聚焦区随着用户移动他或她的头而被连续更新。

[0145] 图 14 是描述用于将经优化图像的一部分耦合到用户的聚焦区的过程的流程图。在一个实施例中, 图 14 是图 2 的步骤 236 和图 2A 的步骤 240 的实施方式。

[0146] 在步骤 1402, 检索上面在步骤 612 所渲染的基于所检测到的用户视野的图像。该

渲染可以分别由中枢计算系统或者图 4A 和 4B 的处理组件 200 或 304 中的任意组件来提供。在一个实施例中,使用中枢计算系统 12 来处理图像将提供对与远离头戴式显示器 2 的计算资源的有效使用,并且允许诸如图 4A 和 4B 的组件之类的处理器组件更主动地驱动头戴式显示器的显示元件和 / 或机电元件。在 1404,可以接收所预测的眼睛位置(其根据图 15 和 16 被计算出),并在 1405,减小可用于耦合到用户聚焦区的所选数目的潜在高分辨率部分。在处理在中枢计算系统中进行的一个实施例中,在 1406,在 1405 选择多个潜在部分,并且将其移动到与头戴式显示器 2 的渲染区域最接近的处理单元中可用的一个或多个存储器位置中的缓冲区中。在一个实施例中,可以将这样的元素提供给处理单元 4 的存储器 330。在其他实施例中,可以将这些部分提供给头戴式显示器 2 的存储器 224。在 1408,进一步减小在当前视野内的当前眼睛位置处以及一个或多个接下来的可能眼睛位置处渲染的潜在的经优化部分。再者,在步骤 1410,经优化的部分可以在中枢计算系统处被计算出并且在下游在诸如例如从中枢计算系统到存储器 330 的处理通道中被缓冲,或者在处理单元处被处理并且在存储器 224 被缓冲。在 1412,根据步骤 236 或 240 在显示器上针对观察者的聚焦区所优化位置的一位置处渲染高分辨率部分。

[0147] 图 15 是示出了用于基于跟踪眼睛位置和已知眼睛数据和头以及已知位置和定向数据来确定用户眼睛的下一位置以及头位置和定向的过程的流程图。如上所述,可以由眼睛跟踪相机 134B 来捕捉眼睛位置数据。在 1502,捕捉用户 的眼睛移动,并且在 1504,一并收集从头戴式显示器传感器和捕捉设备 20A、20B 可用的用户头的位置、定向和移动信息和数据。眼睛位置数据将包括眼睛相对于头的位置和定向的位置,其中头是相对于房间或环境而言的。在 1506,针对每个时间 T_n ,在 1508 对用户眼睛在时间 T_{n+1} 的位置进行预测。可以在 1510 和 1512 计算对时间 T_{n+1} 的备选预测。图 16 示出了用于参照眼睛数据预测用户眼睛位置的方法。而且,针对每个时间 T_n ,将在 1507 对用户的下一头定向和位置进行预测。可以在 1510 和 1512 进行头定向和位置的附加预测。在 1515,根据参照图 2A 的图像使用选择所预测的眼睛位置之一作为下一位置,并且在 1513,将选择所预测的头位置之一。在 1516,将在步骤 240 使用这些位置来确定在该下一位置处渲染该图像的哪些部分,并且该方法在 1518 随着用户的眼睛和头的继续移动而重复。

[0148] 图 16 是示出用于预测可能的眼睛位置的过程的流程图。在 1630,缓冲用户眼睛的多个数据位置,并且在 1632 一旦获得和缓冲了足够数量的数据,就使用预测建模滤波器来计算用户的眼睛将在时间 T_{n+1} 、 T_{n+2} 等处于给定位置的概率。在一个实施例中,将 Kamlan 滤波器用于通过预测值、估计预测值的不确定性、以及计算预测值与测量值的加权平均来计算眼睛位置测量结果的真实值的估计。向具有最小不确定性的值赋予最大权重。可替代地,使用马尔科夫模型。马尔科夫模型使用随时间改变的随机变量来确定系统的状态。在该上下文中,马尔科夫属性提示:该变量的分布仅仅取决于上一状态的分布。可以使用类似的方法来预测头位置和定向。

[0149] 在 1634 可以进行和输出任何数目的相继的预测。应当认识到,可以在相对于用户头的坐标系预测眼睛位置时使用任何数目的预测算法。上述参考方法仅仅是多个合适实施例中的两个。

[0150] 尽管用结构特征和 / 或方法动作专用的语言描述了本主题,但可以理解,所附权利要求书中定义的主题不必限于上述具体特征或动作。更确切而言,上述具体特征和动作

是作为实现权利要求的示例形式公开的。本技术的范围由所附的权利要求进行定义。

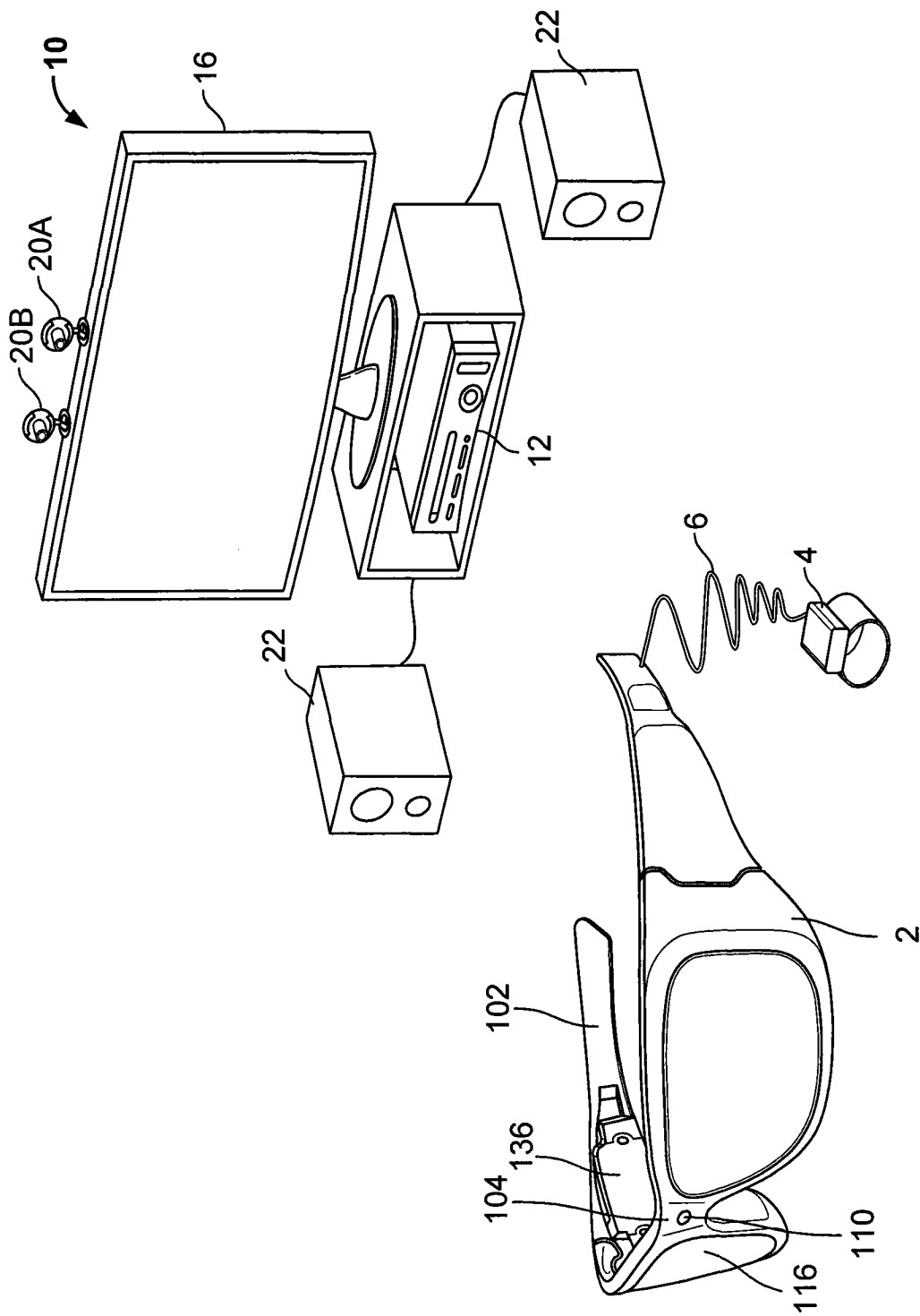


图 1

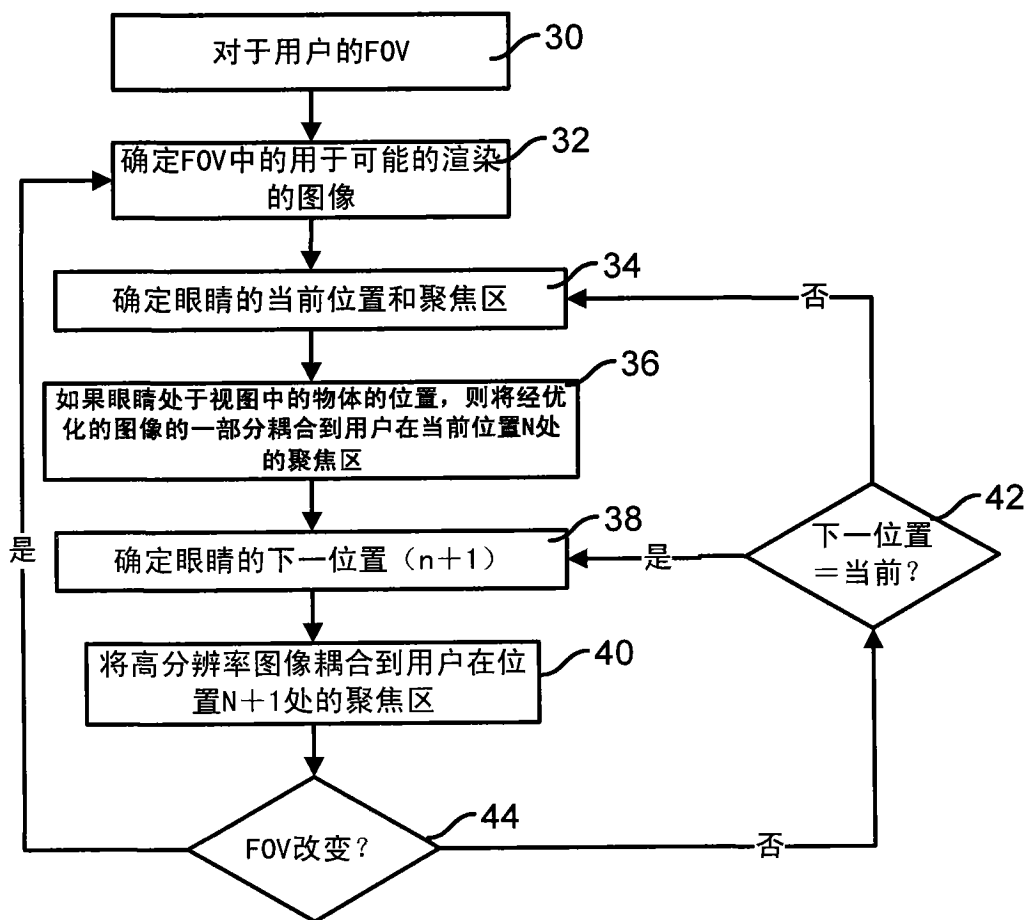


图 2A

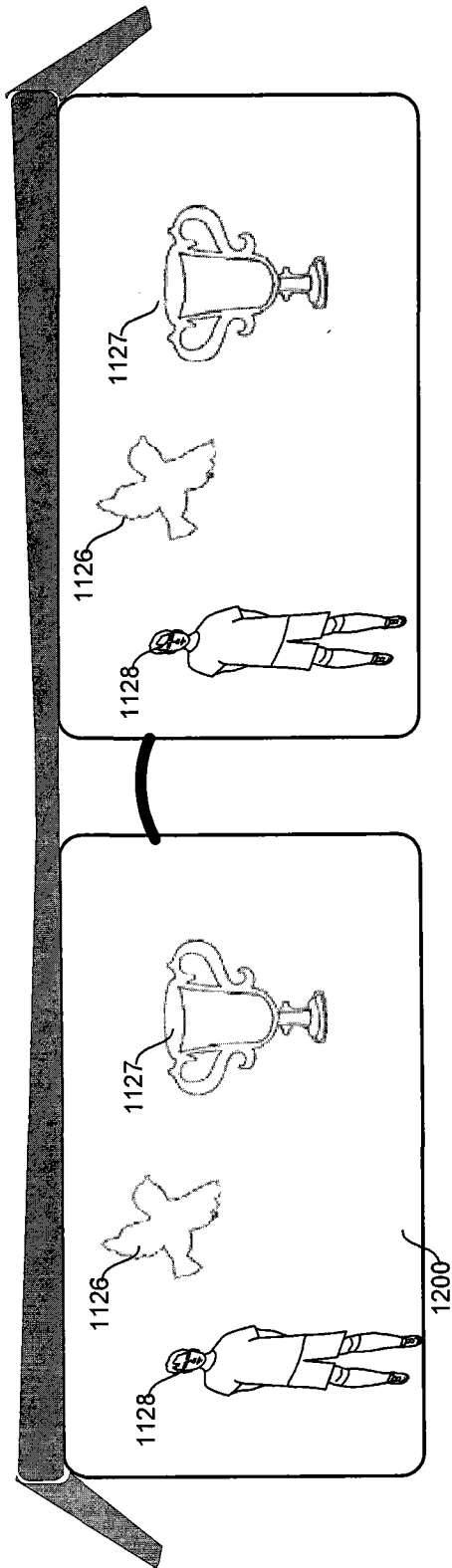


图 2C

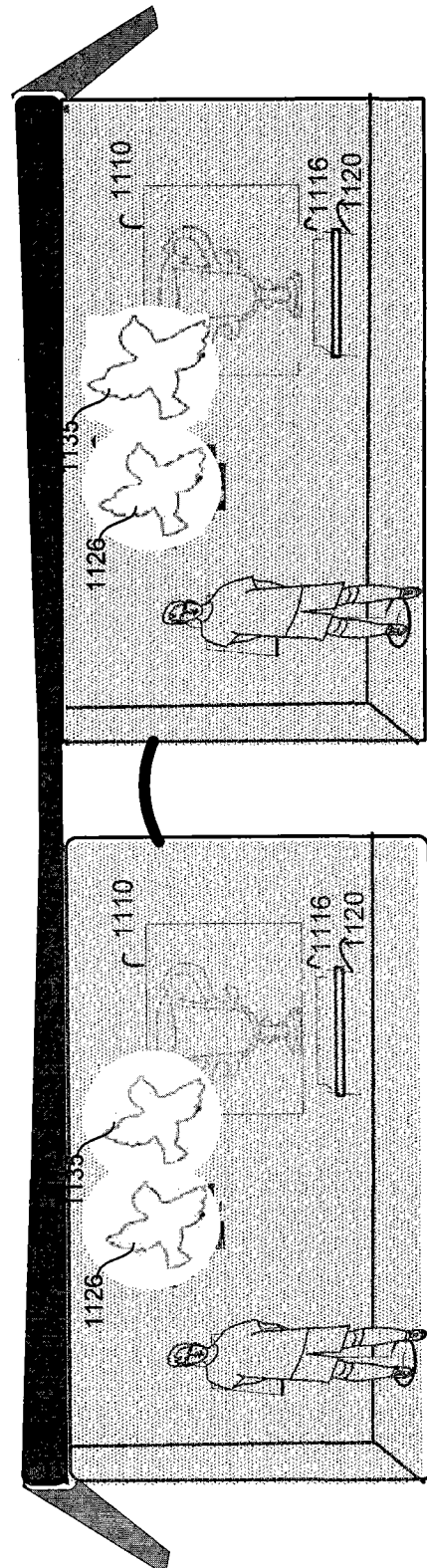


图 2G

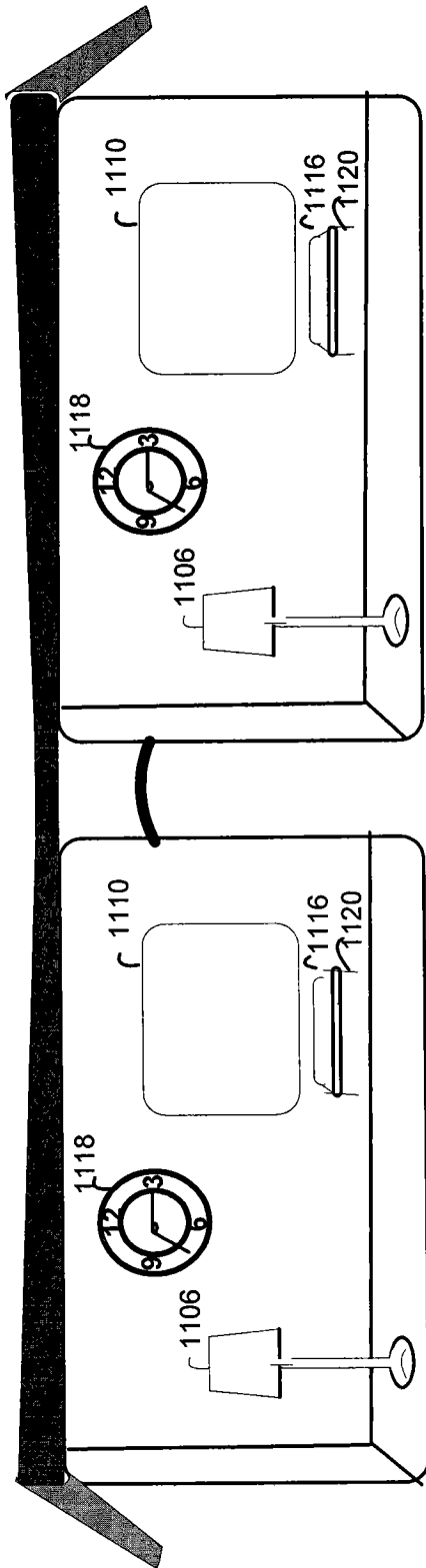


图 2B

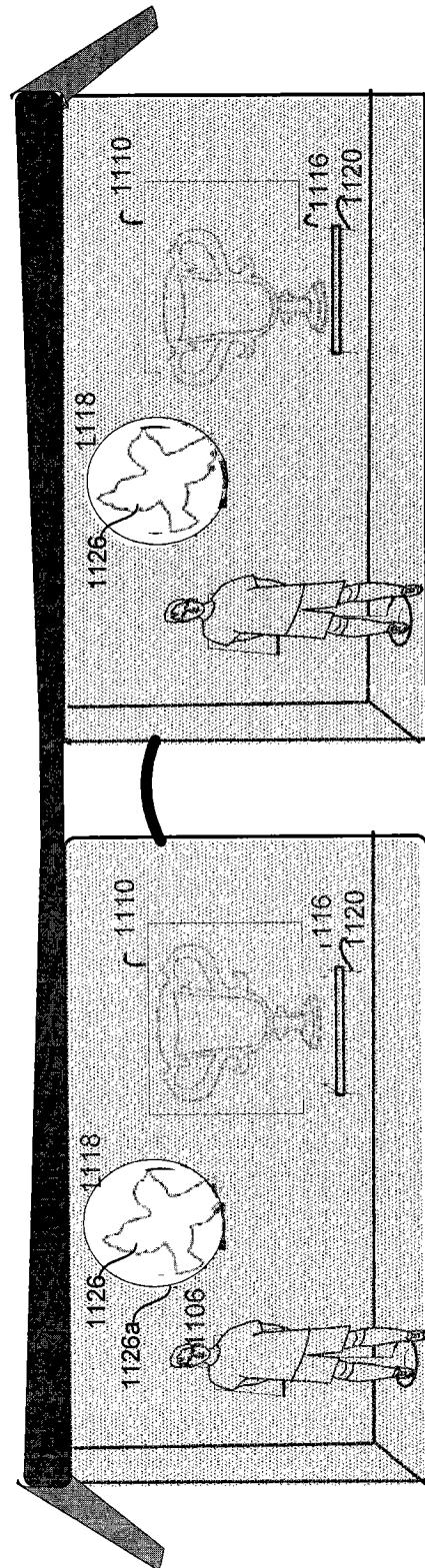


图 2D

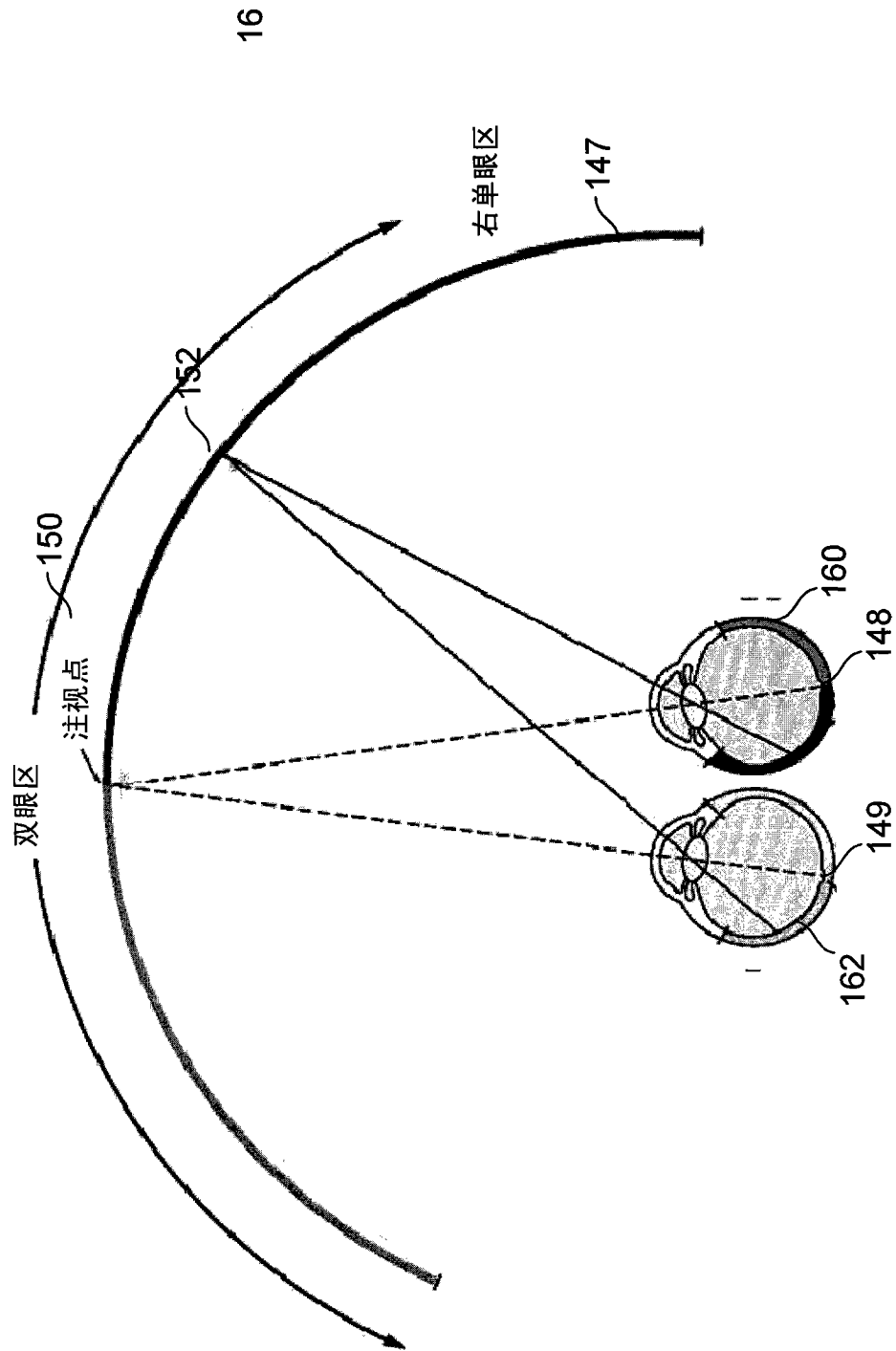


图 2E

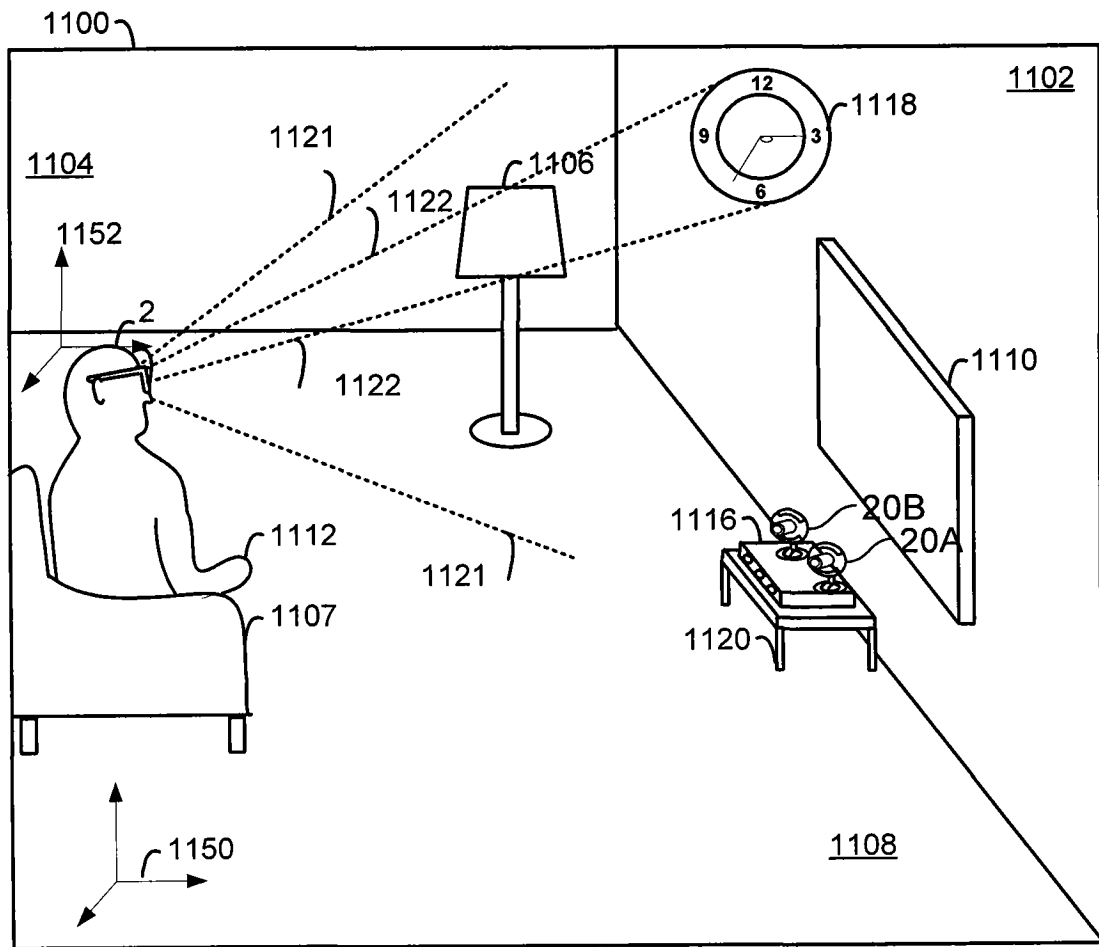


图 2F

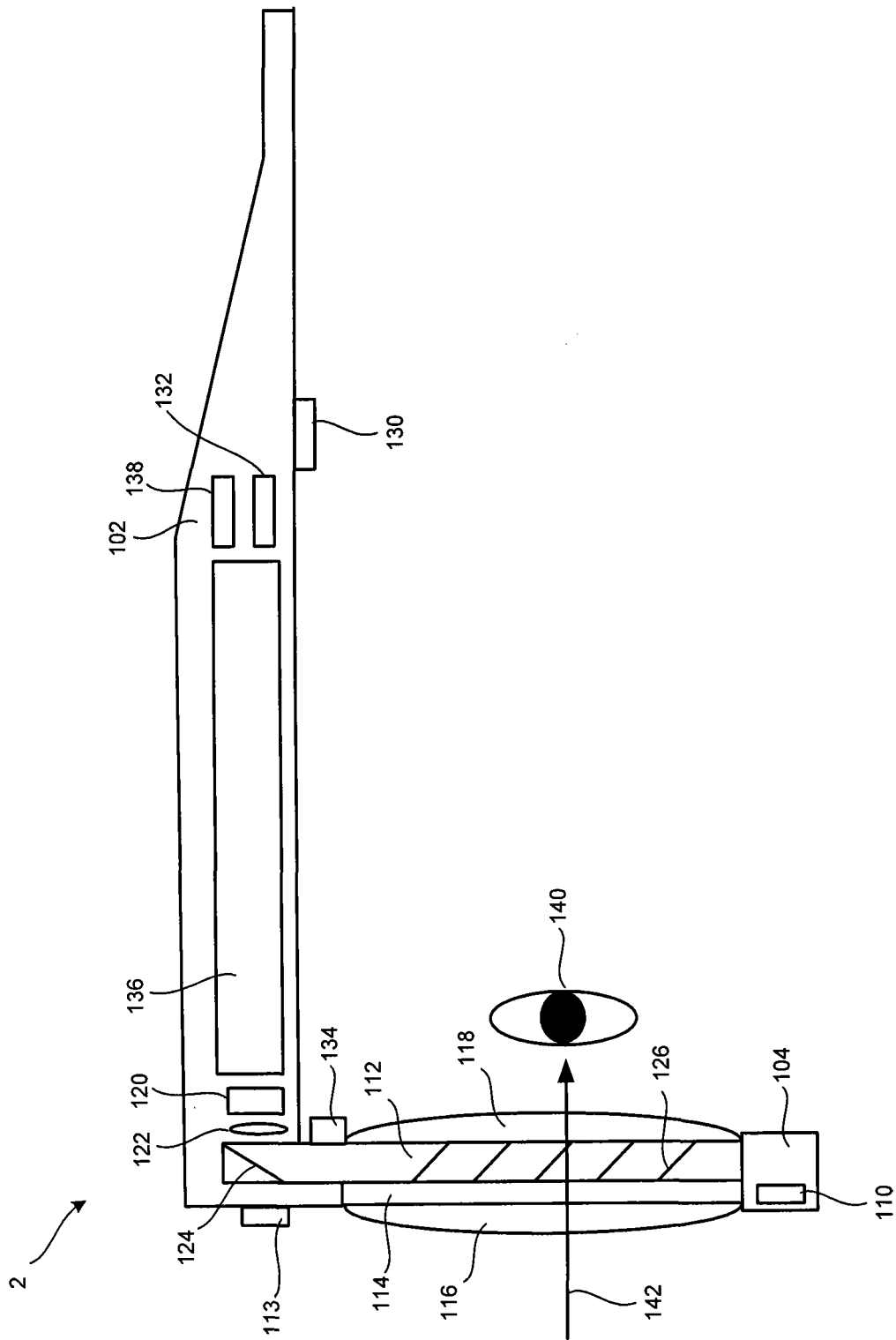


图 3A

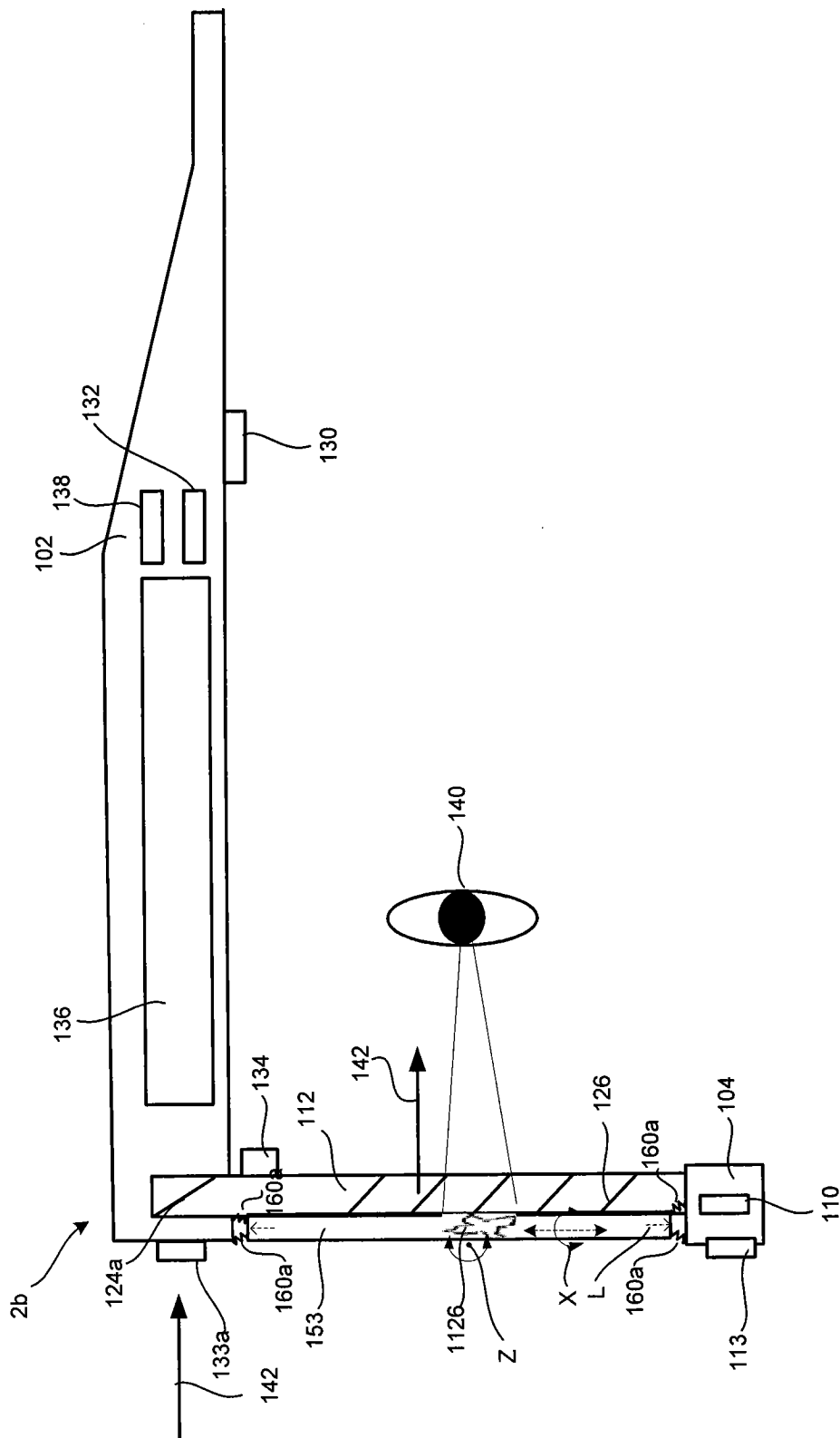
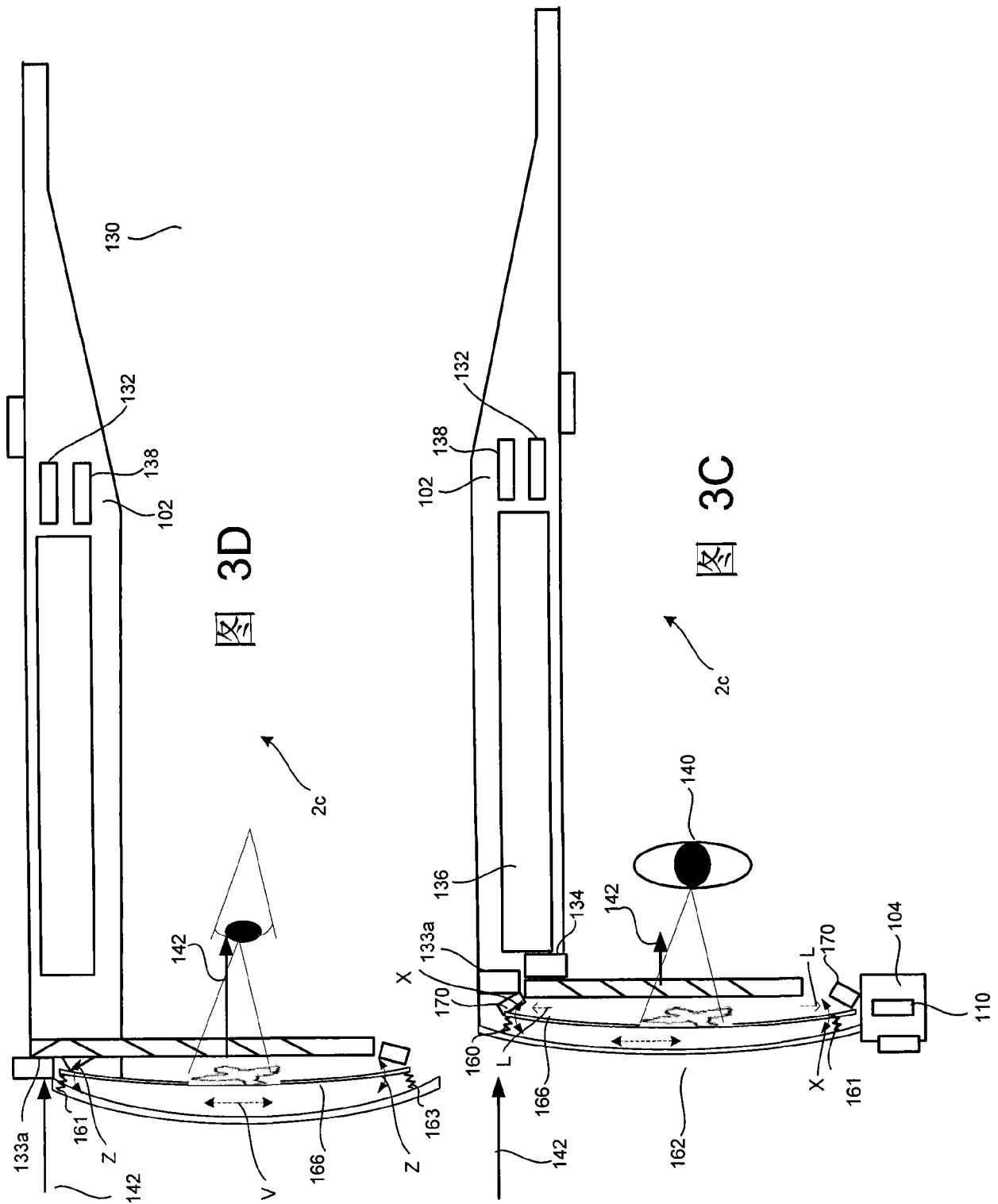


图 3B



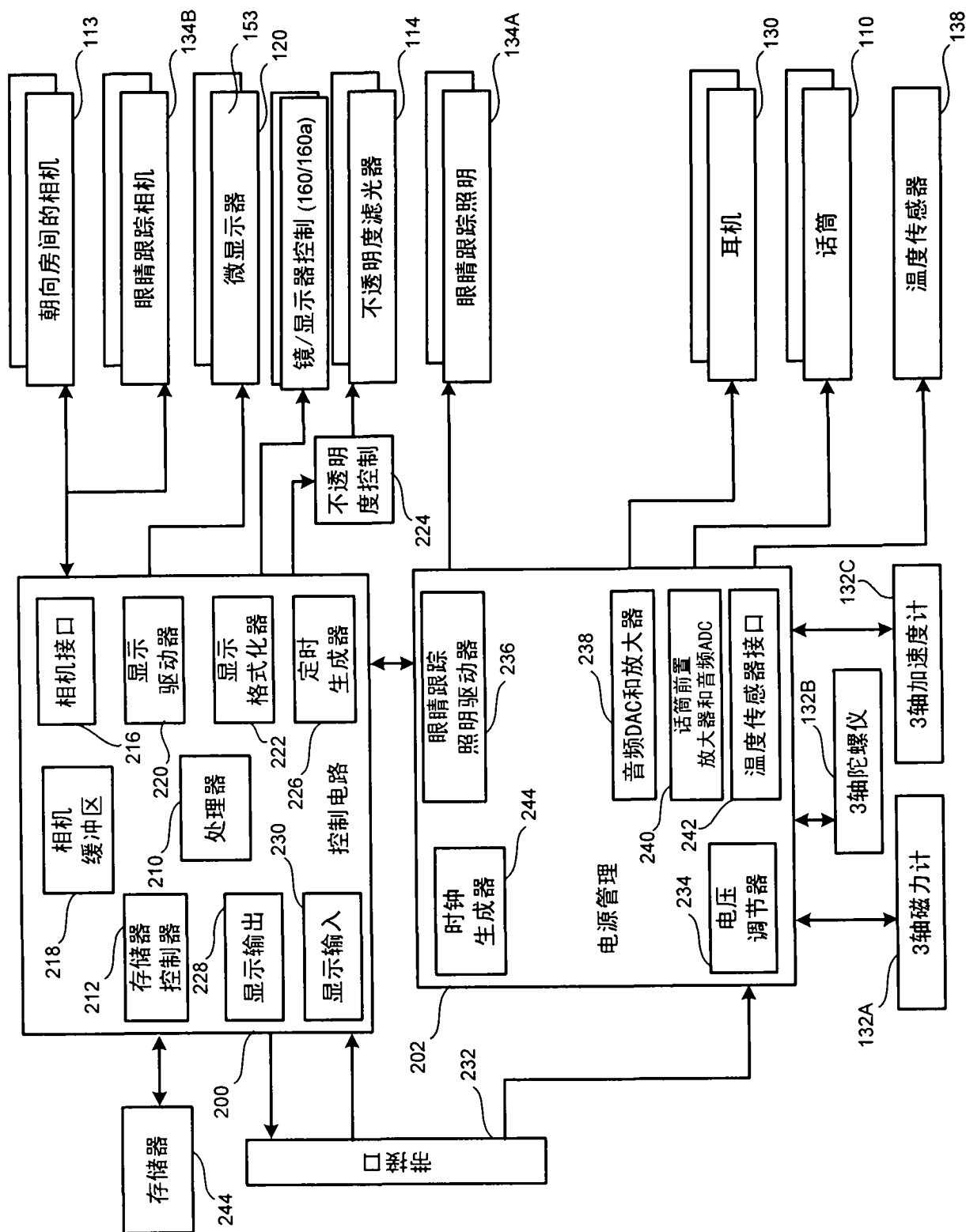


图 4A

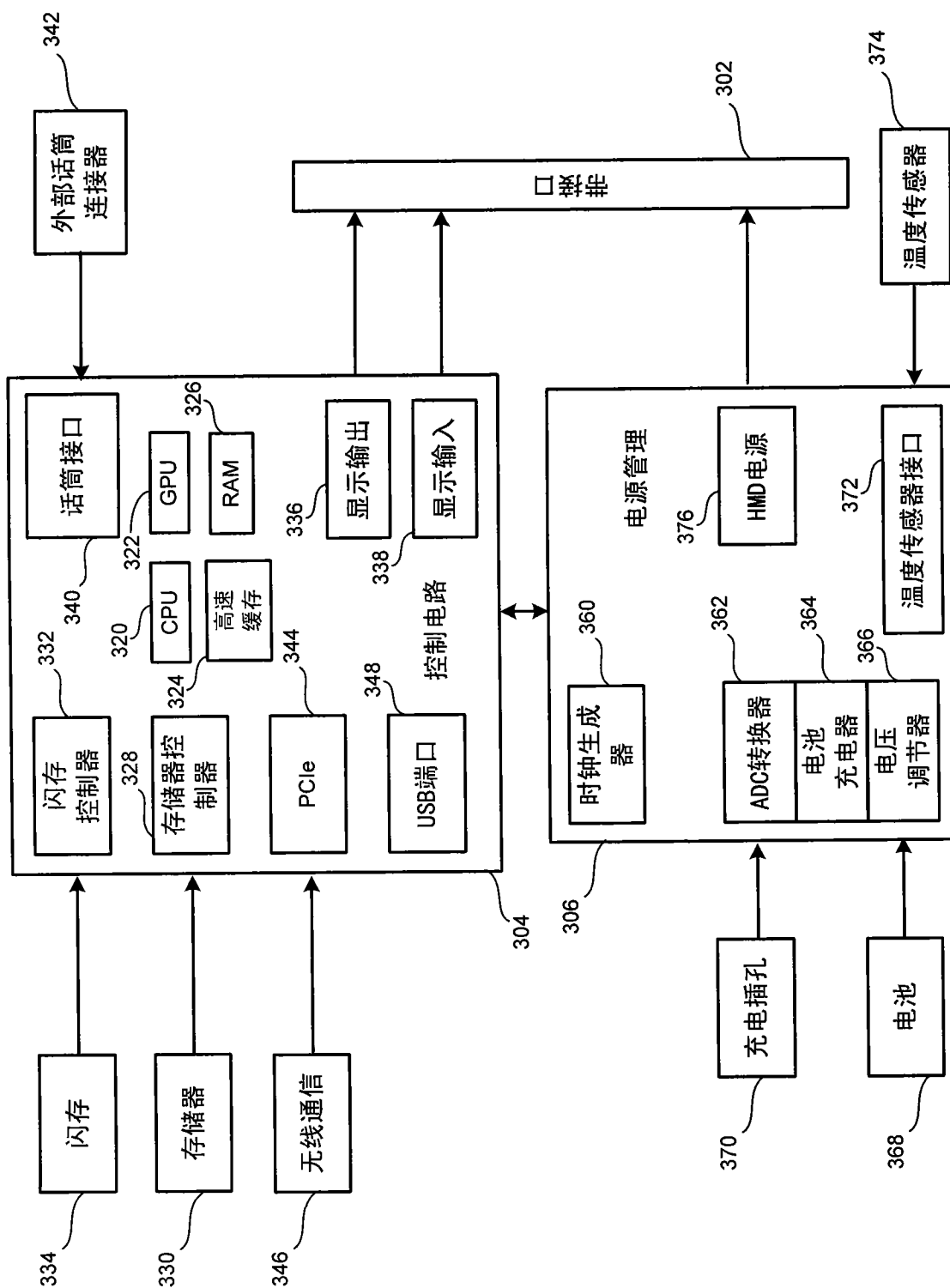


图 4B

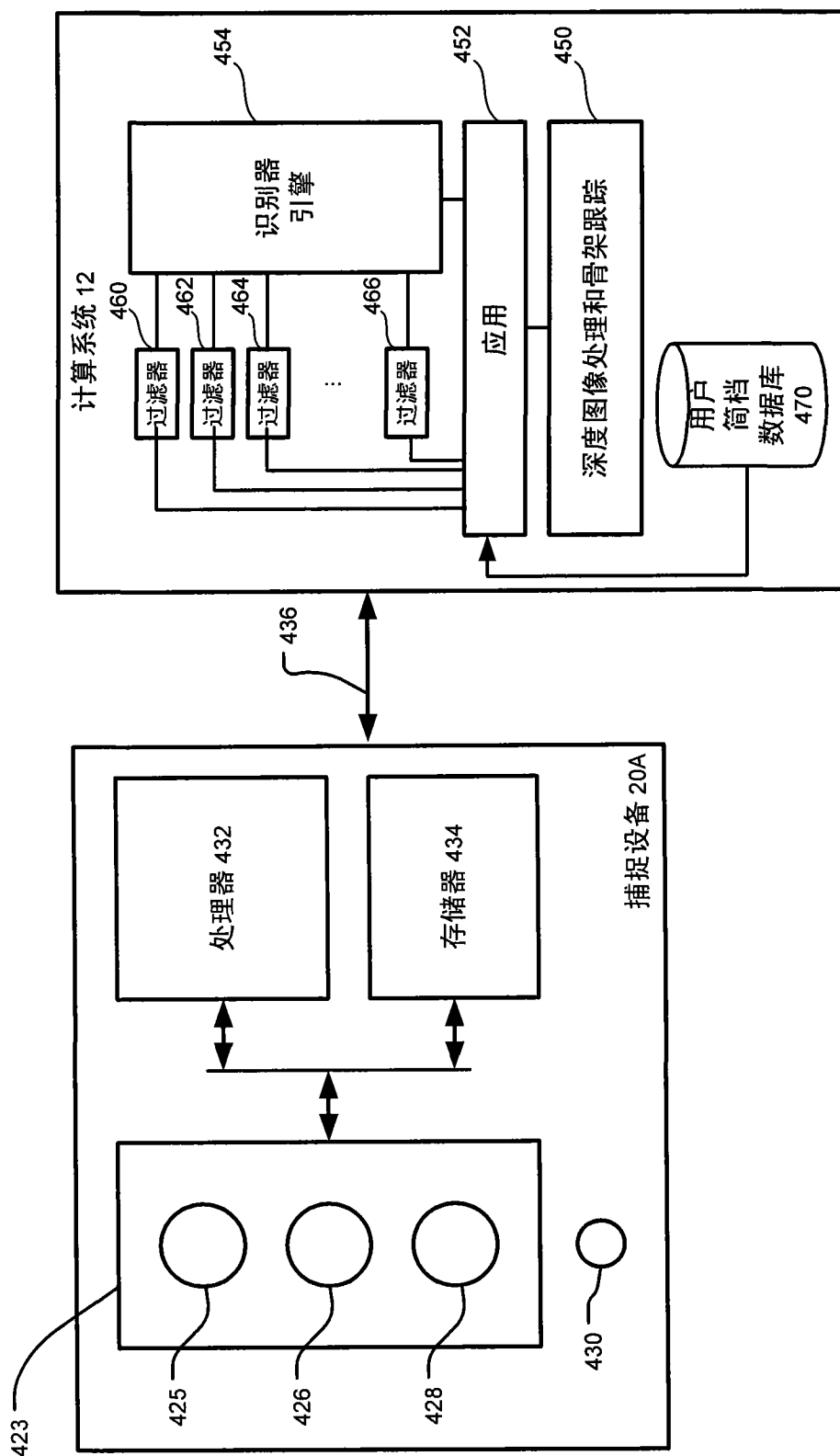


图 5

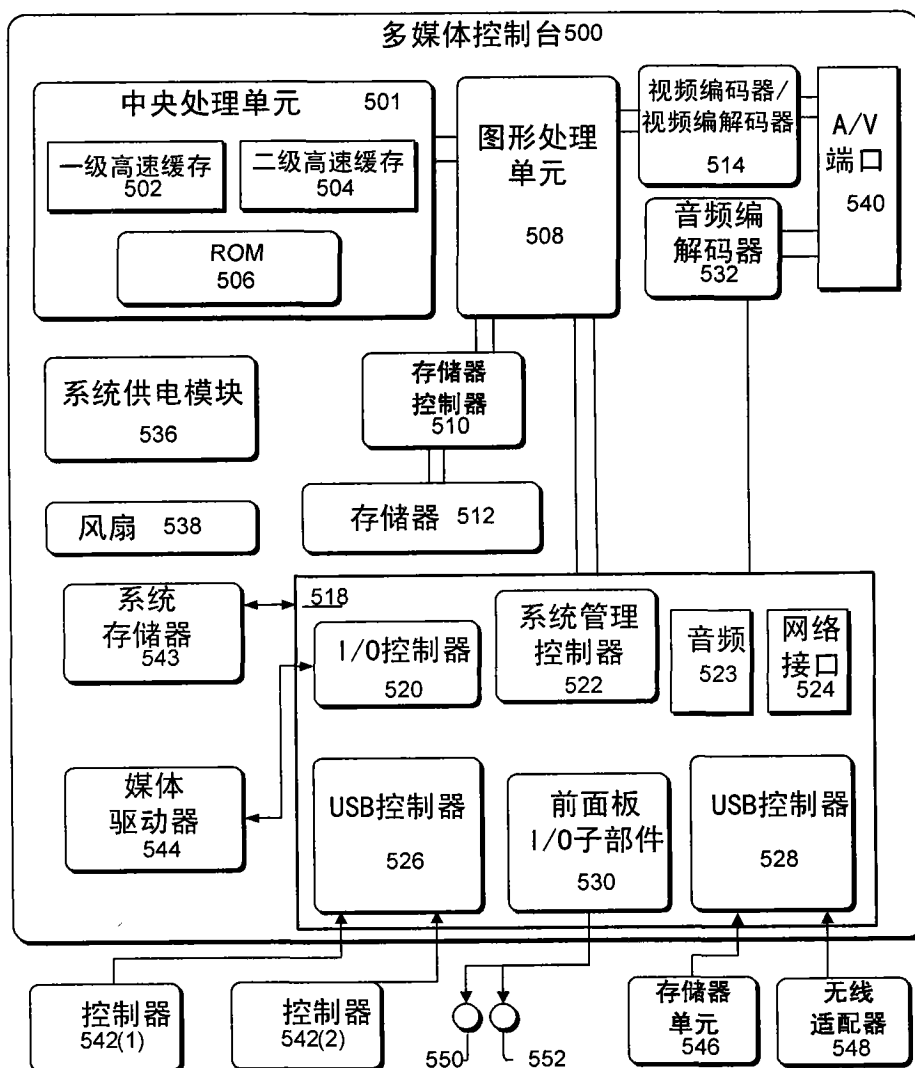


图 6

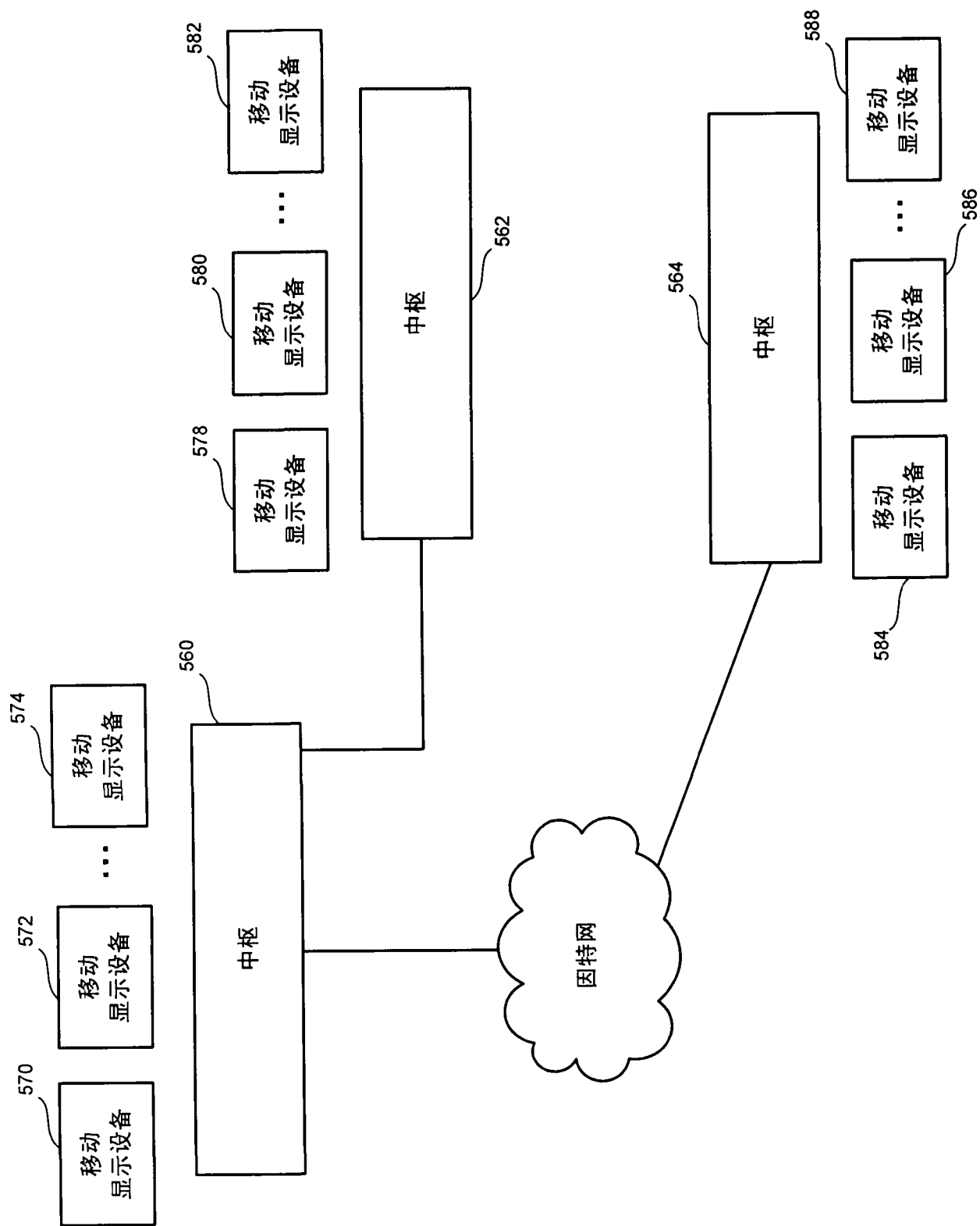


图 7

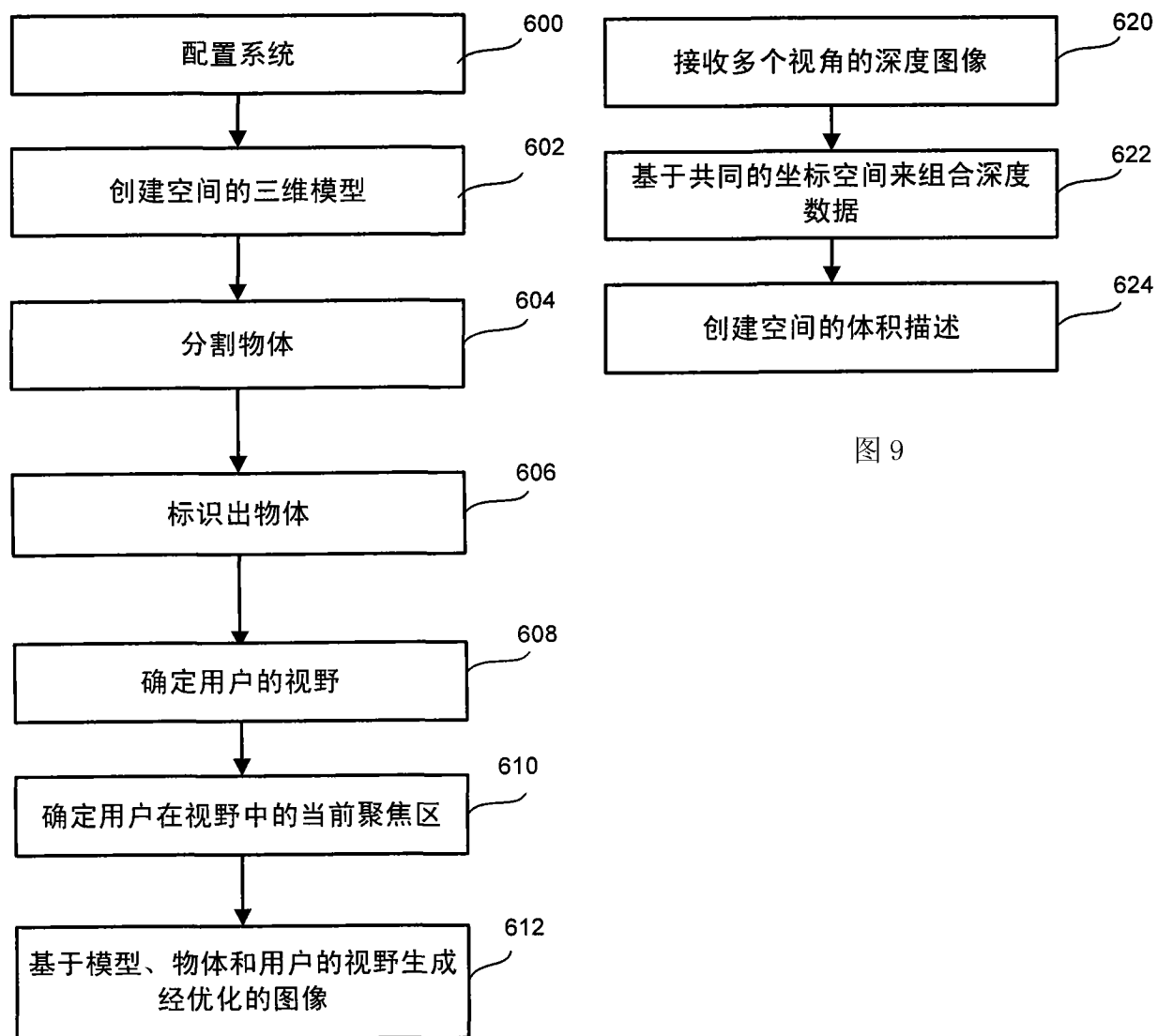


图 9

图 8

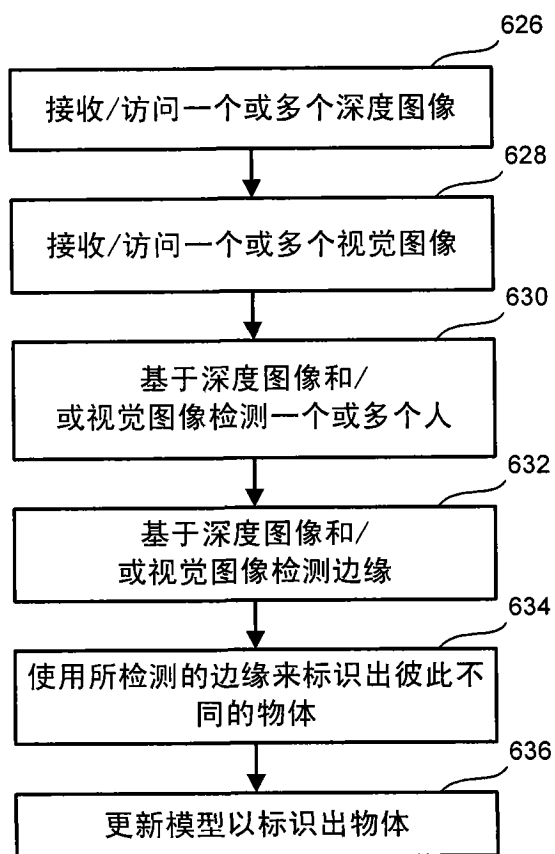


图 10

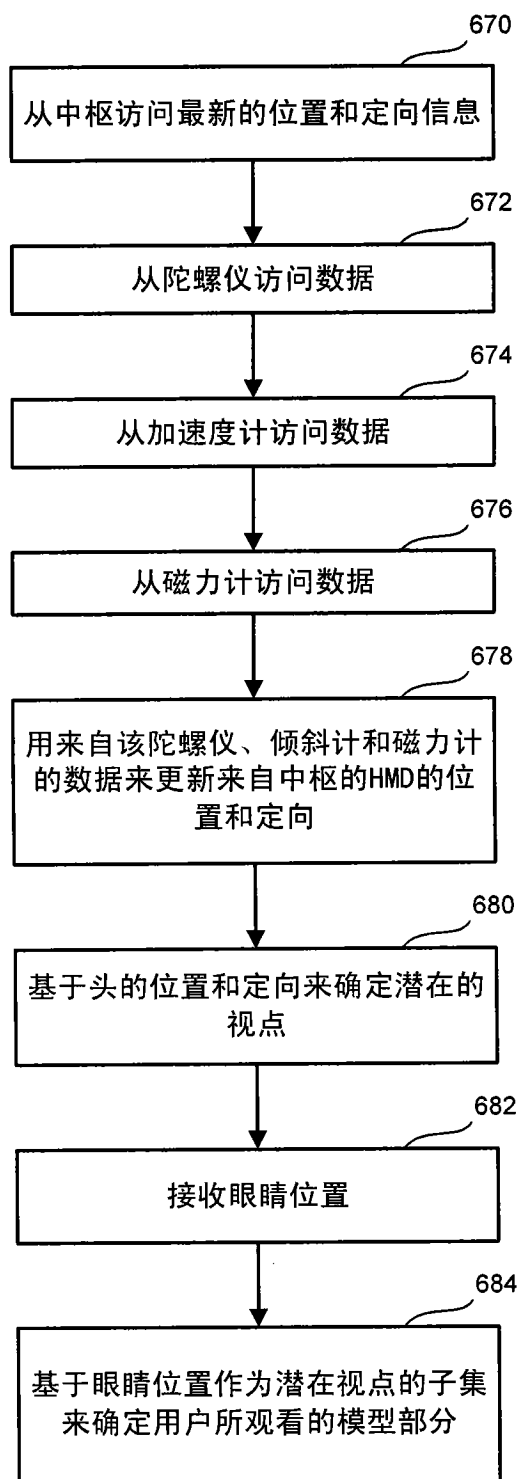


图 11

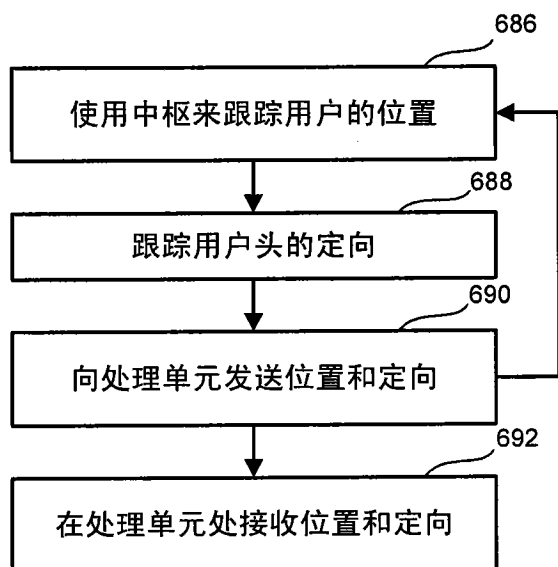


图 12

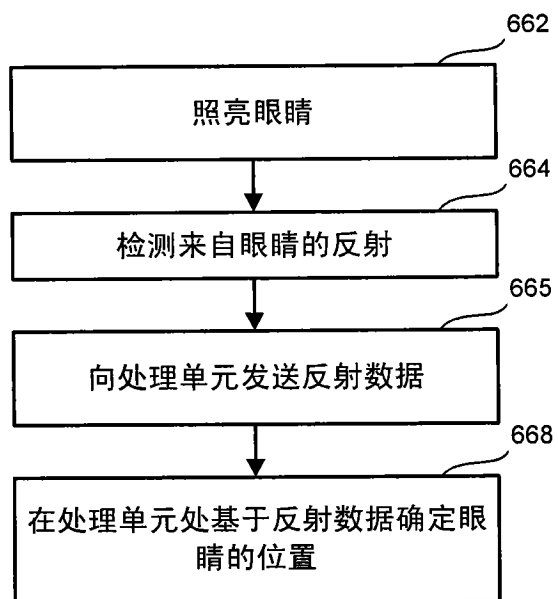


图 13

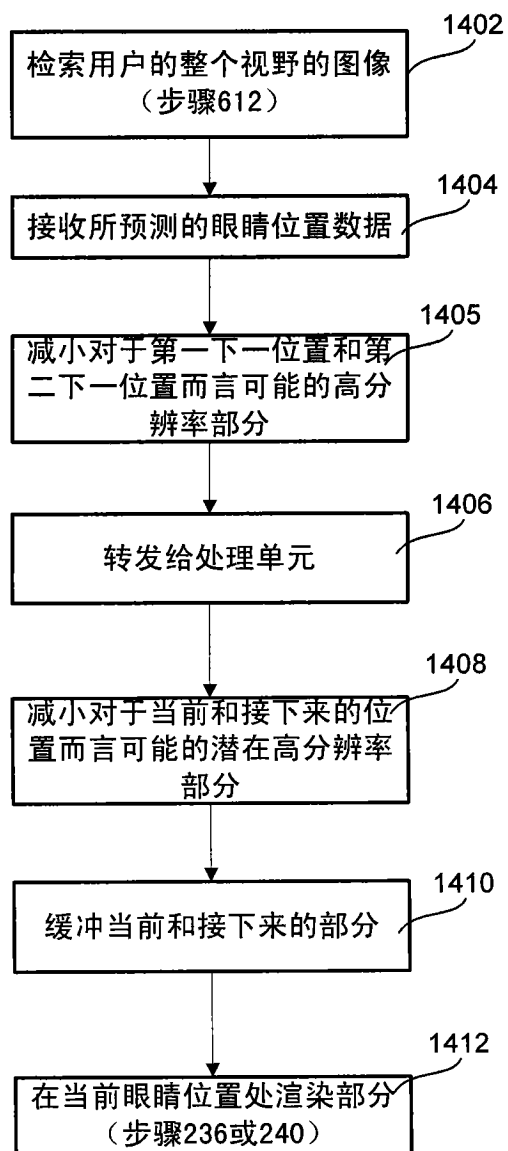


图 14

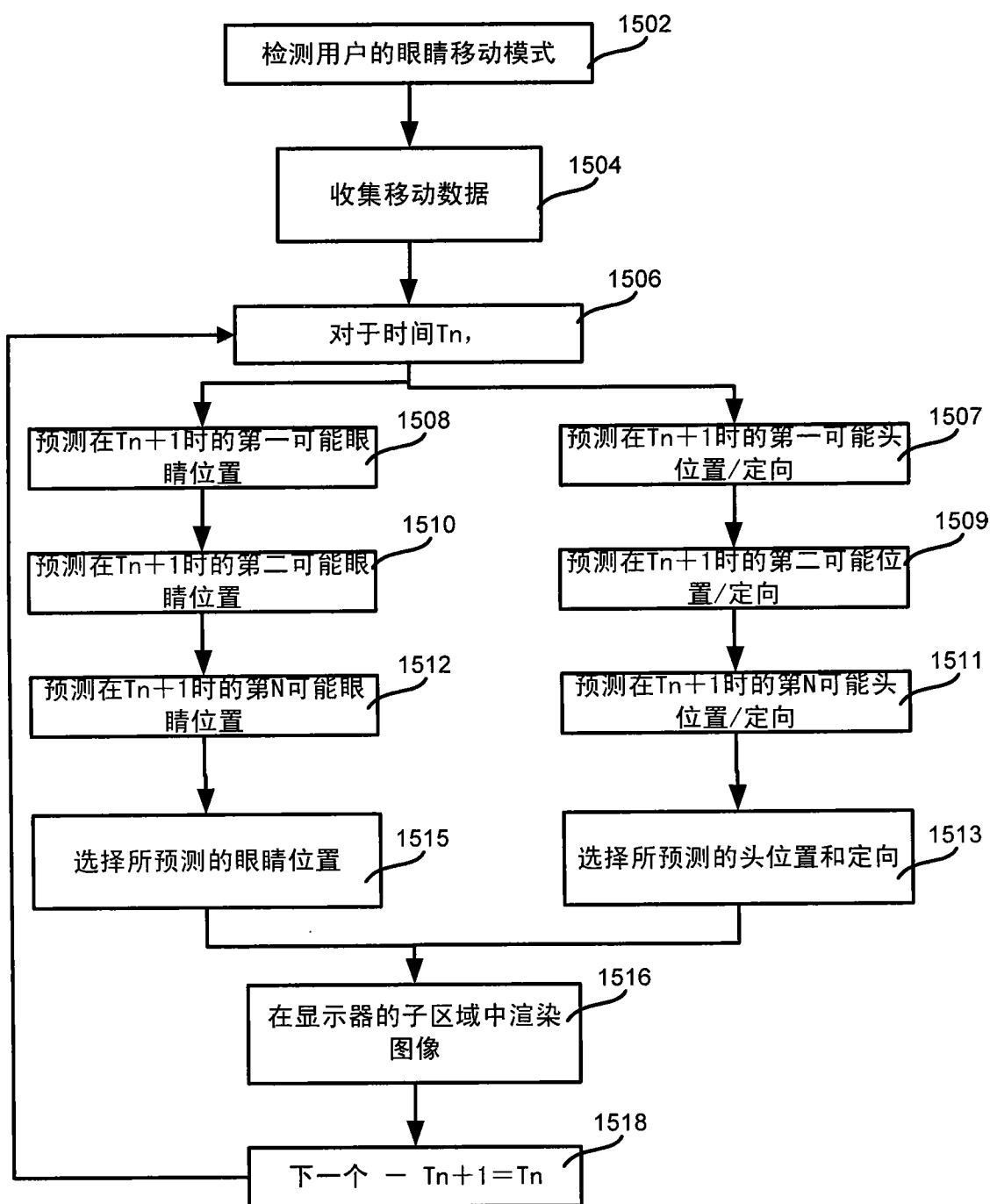


图 15

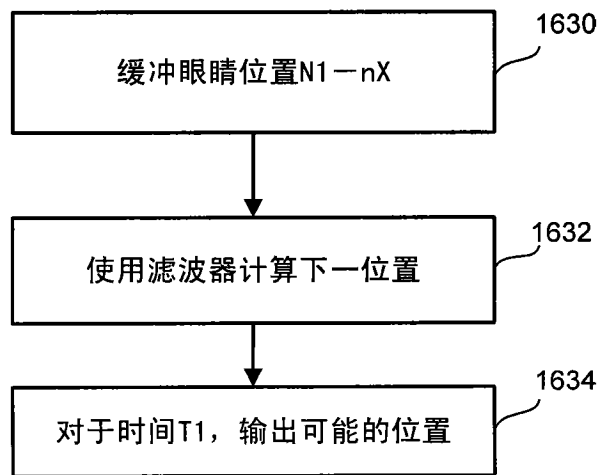


图 16