



(12)发明专利

(10)授权公告号 CN 107209551 B

(45)授权公告日 2020.08.25

(21)申请号 201580054828.8

(51)Int.Cl.

G06F 3/01(2006.01)

(22)申请日 2015.10.06

G06K 9/00(2006.01)

(65)同一申请的已公布的文献号

G06K 9/20(2006.01)

申请公布号 CN 107209551 A

(56)对比文件

(43)申请公布日 2017.09.26

US 2012147328 A1, 2012.06.14,

(30)优先权数据

CN 103492979 A, 2014.01.01,

14/509,976 2014.10.08 US

Yoshinobu Ebisawa. Improved Video-

(85)PCT国际申请进入国家阶段日

Based Eye-Gaze Detection Method.

2017.04.07

《Instrumentation & measurement Technology Conference IEEE》.1998, 第47卷(第4期),

(86)PCT国际申请的申请数据

Yoshinobu Ebisawa. Improved Video-

PCT/US2015/054102 2015.10.06

Based Eye-Gaze Detection Method.

(87)PCT国际申请的公布数据

《Instrumentation & measurement Technology Conference IEEE》.1998, 第47卷(第4期),

W02016/057435 EN 2016.04.14

SU GWON ET AL. Gaze Tracking System

(73)专利权人 微软技术许可有限责任公司

for User Wearing Glasses.《SENSORS》.2014, 地址 美国华盛顿州

第14卷(第2期),

(72)发明人 郭慧敏 I·埃登 V·萨克拉尔
D·Z·尼斯特

审查员 廖露露

(74)专利代理机构 北京市金杜律师事务所

权利要求书3页 说明书8页 附图7页

11256

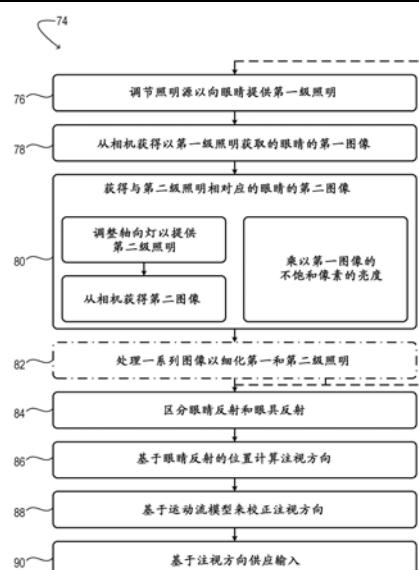
代理人 王茂华

(54)发明名称

用于注视跟踪的系统和方法

(57)摘要

一种在可操作地耦合到视觉系统的计算机系统中供应表示注视方向的输入的方法(74)。在该方法中,通过视觉系统的相机获取在第一级照明处的眼睛的第一图像(78)。从相机获得第一图像,并且还获得与不同的第二级照明相对应的眼睛的第二图像(80)。比较第一和第二图像的对应像素的亮度,以便将眼睛的照明的反射与眼具的照明的反射区分开来(84)。然后基于眼睛的照明的反射来供应输入(90)。



1. 一种在可操作地耦合到视觉系统的计算机系统中实施的、用于供应表示注视方向的输入的方法，所述方法包括：

从所述视觉系统的相机，获得在由在第一级功率级别处被操作的发射器对眼睛的照明下获取的所述眼睛的第一图像；

从所述视觉系统的所述相机，获得在由所述发射器对所述眼睛的所述照明下获取的所述眼睛的第二图像，所述发射器在第二、不同的功率级别处被操作；

比较所述第一图像和所述第二图像的对应像素的亮度，以区分所述眼睛对所述照明的反射与眼具对所述照明的反射，包括从由所述相机在所述发射器的相互不同的功率级别处获取的所述眼睛的三个或更多个图像之间选择所述第一图像和所述第二图像，以便揭示来自所述眼睛的不饱和的、反射的明亮瞳孔响应；以及

基于所述眼睛对所述照明的所述反射来供应所述输入。

2. 根据权利要求1所述的方法，其中，所述眼睛对所述照明的所述反射包括来自所述眼睛的视网膜的反射，所述反射通过所述眼睛的瞳孔返回并且使得在所述第一图像中所述瞳孔相对于周围的虹膜显得明亮。

3. 根据权利要求1所述的方法，其中，所述眼睛的所述反射包括所述眼睛的虹膜的反射，使得在所述第一图像中所述眼睛的瞳孔相对于所述虹膜显得暗。

4. 根据权利要求1所述的方法，还包括当排除与所述眼具对所述照明的所述反射相关联的那些像素时，基于在所述第一图像中所述眼睛对所述照明的所述反射的位置来计算所述输入。

5. 根据权利要求4所述的方法，还包括基于运动学模型来校正所述输入以考虑在获得所述第一图像和所述第二图像之间的所述眼睛的移动。

6. 根据权利要求1所述的方法，其中，供应的所述输入包括定义通过所述眼睛的视线方向的方位角和仰角。

7. 根据权利要求1所述的方法，还包括：

如果所述第一图像和所述第二图像的所述对应像素的亮度相差多于阈值量，则将这样的像素与所述眼睛对所述照明的所述反射相关联；以及

如果所述第一图像和所述第二图像的所述对应像素的亮度相差少于阈值量，则将这样的像素与所述眼具对所述照明的所述反射相关联。

8. 根据权利要求1所述的方法，还包括如果所述第一图像和所述第二图像的所述对应像素都是饱和的，则将所述第一图像和所述第二图像的所述对应像素与所述眼具对所述照明的所述反射相关联。

9. 根据权利要求1所述的方法，其中，基于刺激和区分来自所述眼睛的明亮瞳孔响应的能力来选择所述第一功率级别。

10. 根据权利要求1所述的方法，其中，基于环境光条件来选择所述第一功率级别。

11. 根据权利要求1所述的方法，其中所述第一图像和所述第二图像从由所述相机在所述发射器的相互不同的功率级别处获取的所述眼睛的三个或更多个图像被选择。

12. 根据权利要求1所述的方法，其中比较所述对应像素的所述亮度包括使用机器学习的算法以区分所述眼睛对所述照明的所述反射与所述眼具对所述照明的所述反射。

13. 根据权利要求1所述的方法，其中所述眼睛是两个眼睛中的第一眼睛，并且其中所

述第一图像和所述第二图像是所述第一眼睛的图像,所述方法还包括:

从所述视觉系统的相机,获得在由在所述第一级功率处被操作的发射器对所述眼睛的照明下获取的第二眼睛的第一图像;

从所述视觉系统的相机,获得在由所述发射器对所述眼睛的所述照明下获取的所述第二眼睛的第二图像,所述发射器在第二功率级别处被操作;

比较所述第二眼睛的所述第一图像和所述第二图像的对应像素的亮度,以区分所述第二眼睛对所述照明的反射与眼具对所述照明的反射;以及

基于所述第二眼睛对所述照明的所述反射以及所述眼具对所述照明的所述反射的独立性来供应所述输入,所述输入包括所述第一眼睛和所述第二眼睛的确定的焦点。

14.一种视觉系统,包括:

照明系统,包括被配置为照亮眼睛的发射器;

相机,被配置为获取所述眼睛的一个或多个图像;以及

处理器和相关联的计算机存储器,可操作地耦合到所述相机和所述照明系统,所述计算机存储器保持由所述处理器可执行的指令以:

从所述视觉系统的所述相机,获得在由在第一级功率处的所述发射器对所述眼睛的照明下的所述眼睛的第一图像;

从所述视觉系统的所述相机,获得在由所述发射器对所述眼睛的所述照明下获取的所述眼睛的第二图像,所述发射器在第二、不同的功率级别处被操作;

比较所述第一图像和所述第二图像的对应像素的亮度,以区分所述眼睛对所述照明的反射与眼具对所述照明的反射,包括从由所述相机在所述发射器的相互不同的功率级别处获取的所述眼睛的三个或更多个图像之间选择所述第一图像和所述第二图像,以便揭示来自所述眼睛的不饱和的、反射的明亮瞳孔响应;以及

基于所述眼睛对所述照明的所述反射以及所述眼具对所述照明的所述反射的独立性来向计算机系统供应输入。

15.根据权利要求14所述的系统,还包括指令,所述指令可执行以在多次曝光中捕获所述眼睛的所述第一图像和所述眼睛的所述第二图像中的一种或多种。

16.根据权利要求14所述的系统,其中,所述指令可执行以在从所述相机读取所述眼睛的所述第一图像和所述眼睛的所述第二图像之前,顺序地获取所述眼睛的所述第一图像和所述眼睛的所述第二图像。

17.根据权利要求14所述的系统,其中,所述照明系统被配置为在30毫秒或更短时间内从提供所述发射器的所述第一功率级别转换到提供所述发射器的所述第二功率级别。

18.一种在可操作地耦合到视觉系统的计算机系统中实施的、用于供应响应于眼睛的注视方向的输入的方法,所述方法包括:

从所述视觉系统的相机,获得在由在第一级功率级别处被操作的发射器对所述眼睛的照明下获取的所述眼睛的第一图像;

从所述相机获得在由所述发射器对所述眼睛的所述照明下获取的所述眼睛的第二图像,所述发射器在第二、不同的功率级别处被操作;

比较所述第一图像和所述第二图像的对应像素的亮度,以区分所述眼睛对所述照明的反射与眼具对所述照明的反射,包括从由所述相机在所述发射器的相互不同的功率级别处

获取的所述眼睛的三个或更多个图像之间选择所述第一图像和所述第二图像,以便揭示来自所述眼睛的不饱和的、反射的明亮瞳孔响应;以及

基于所述眼睛对所述照明的所述反射以及所述眼具对所述照明的所述反射的独立性来供应所述输入。

19.根据权利要求18所述的方法,其中,所述发射器是所述视觉系统的分立发光二极管,所述方法还包括:

调整所述分立发光二极管以在获得所述第一图像之前提供所述第一功率级别;

调整所述分立发光二极管以在获得所述第二图像之前提供所述第二功率级别。

用于注视跟踪的系统和方法

技术领域

[0001] 本公开的实施例涉及用于注视跟踪的系统和方法。

背景技术

[0002] 近来的硬件和软件发展使用于计算机系统的自然用户输入 (NUI) 的新模式成为可能。手势识别、语音识别和注视跟踪是示例的NUI模式，其为了各种目的和在各种环境中使得用户能够直观地与计算机系统进行交互。

发明内容

[0003] 公开了涉及在眼睛跟踪系统中区分来自眼睛的反射和来自眼具 (eyewear) 的反射的实施例。一个公开的实施例提供了一种在可操作地耦合到视觉系统的计算机系统中供应表示注视方向的输入的方法。在该实施例中，通过视觉系统的相机获取处于第一级照明处的眼睛的第一图像。从相机获得第一图像，并且还获得与不同的第二级照明相对应的眼睛的第二图像。比较第一和第二图像的对应像素的亮度，以便区分眼睛的照明的反射与眼具的照明的反射。然后基于眼睛的照明反射来供应输入。

[0004] 提供本发明内容以便以简化的形式介绍概念的选择，其在下面的具体实施方式中进一步描述。本发明内容不旨在标识所要求保护的主题的关键特征或必要特征，也不旨在用于限制所要求保护的主题的范围。此外，所要求保护的主题不限于解决在本公开的任何部分中指出的任何或所有缺点的实现。

附图说明

[0005] 图1示出了其中用户的注视被跟踪并被用作计算机系统中的输入的示例环境的各方面。

[0006] 图2示出了具有可操作地耦合到视觉系统的NUI系统的示例计算机系统的各方面。

[0007] 图3示出了配置用于注视检测的示例视觉系统的各方面。

[0008] 图4A示出了在高照明级别下获得的用户眼睛的示例图像。

[0009] 图4B示出了在低照明级别下获得的用户眼睛的示例图像。

[0010] 图4C示出了从图4A的图像中排除用户眼具的反射的示例结果。

[0011] 图5示出了其中在计算机系统中用户的注视被跟踪并被用作为输入的附加示例环境的各方面。

[0012] 图6示出了配置用于注视检测的另一示例视觉系统的各方面。

[0013] 图7示出了响应于计算机系统中的注视方向来供应输入的示例方法。

具体实施方式

[0014] 注视跟踪是基于用户注视方向的NUI的形式。在这种方法中，用户的眼睛的图像由相机获取。诸如瞳孔或角膜缘之类的眼睛特征位于所获取的图像中，并且基于这些特征的

位置来计算注视方向。以此方式计算的注视方向可被用于导航图形用户界面、启动程序、做出选择、在游戏中移动角色等。虽然可以在裸眼的图像中识别所期望的眼睛特征，但来自眼具的杂散反射可能成为干扰源。这种干扰可能降低戴眼具的用户的注视跟踪输入的准确性。如本文所使用的，术语“眼具”包括佩戴的任何类型的器具，其在眼睛与眼睛的视场的至少一部分之间设置透视结构。示例包括但不限于：眼具、太阳眼具、遮阳镜、面罩、护目镜、隐形眼具系统和其他眼睛设备、在佩戴者的视场中投影虚拟图像的近眼显示系统等。

[0015] 本文公开的示例可以有助于区分来自裸眼的光的反射和来自眼具的光的反射，并且因此可以促进眼睛跟踪。图1示出了示例环境10的各方面，其中用户的注视被跟踪并被用作计算机系统中的输入。所图示的环境是个人住宅的客厅或家庭室。然而，本文所公开的系统和方法同样适用于其他环境，比如工作场所、零售和服务环境。环境10具有用于用户14享受的家庭娱乐系统12。家庭娱乐系统包括可操作地耦合到计算机系统20的大屏幕显示器16和扬声器18。计算机系统20的性质在各种实现中可以不同。在一些示例中，计算机系统可以是被配置为播放音乐和/或视频的视频游戏系统或多媒体系统。在其他示例中，计算机系统可以是用于因特网访问和生产中的应用的通用计算机系统。在不脱离本公开的范围的情况下，计算机系统20可以被配置用于上述任何或所有目的、和/或任何其它合适的目的。

[0016] 计算机系统20可以被配置为接受来自一个或多个用户14的各种形式的输入。这样，诸如键盘、鼠标、触摸屏、游戏手柄或操纵杆控制器的用户输入设备可以可操作地耦合到计算机系统20。计算机系统20还可以被配置为接受来自一个或多个用户的自然用户输入(NUI)。为了传达NUI，所图示的计算机系统包括NUI系统22。NUI系统被配置为捕获NUI的各个方面，并且向计算机系统内的其他结构提供相应的可操作的输入。为此，NUI系统从计算机系统的包括视觉系统24和可选的收听系统26的各种感觉部件接收低级输入。

[0017] 收听系统26(如果包括的话)可以包括一个或多个麦克风来拾取来自用户14的发声和其他可听见的输入。视觉系统24可以被配置为检测各种形式的用户输入，诸如注视向量V和焦点P，以及手部手势和身体手势，面部特征等。在所图示的示例中，视觉系统和收听系统共享公共外壳；在其他示例中，它们可以是分开的。在另外的示例中，视觉、收听和NUI系统可以集成在计算机系统20内。计算机系统及其外围组件可以经由有线通信链路(如图所示)或以任何其他合适的方式耦合。

[0018] 图2是示出计算机系统20、NUI系统22、视觉系统24和收听系统26的示例的各方面的高级示意图。所图示的计算机系统包括操作系统(OS)28，其可以以软件/或固件被实例化。计算机系统还包括一个或多个应用30，例如诸如视频游戏应用、数字媒体播放器、因特网浏览器、照片编辑器、文字处理器和/或电子表格应用。计算机系统20，NUI系统22，视觉系统24和收听系统26可以包括支持它们各自的功能所需的适当的数据存储、指令存储和逻辑硬件，如下文进一步描述的。

[0019] 在图2的示例中，视觉系统24包括一个或多个平面图像相机32，并且还可以包括一个或多个深度相机34。每个深度相机，如果包括的话，可以被配置为获取用户14和环境的其他方面的深度图的时间分辨的序列。视觉系统还包括照亮用户14和环境10的轴上灯和离轴灯36A和36B，以支持由平面图像相机和/或深度相机进行成像。视觉系统的每个灯和相机被可操作地耦合到微控制器38。微控制器可以被配置为控制和触发由相机进行的图像采集，并且控制视觉系统的每个灯的照明输出。

[0020] 平面图像相机32在场角度范围上检测光，并将这些角度映射到矩形像素阵列上。在一个示例中，平面图像相机可以检测与阵列的像素的子集相关联的多个波长通道中的光——例如，红色，绿色，蓝色等。可替代地，可以使用单色平面图像相机来以灰度成像可见光，近红外(NIR)，红外(IR)和/或紫外(UV)光。在平面图像相机中暴露的所有像素的颜色或亮度值共同地构成数字图像。在一些示例中，平面图像相机的像素可以被记录到深度相机的像素。

[0021] 如上所指出，NUI系统22处理来自视觉系统24和可选的收听系统26的低级输入(即，信号)，以在计算机系统20中提供可操作的高级输入。例如，NUI系统可以对来自收听系统26的音频信号执行声音或语音识别。语音识别可以生成要在计算机系统的OS 28中接收的对应的基于文本的或其他的高级命令。在图2所示的例子中，从感觉数据制定特定形式的NUI的任务被分配给特定的NUI引擎：语音识别引擎40，手势识别引擎42，面部识别引擎44和注视检测引擎46。这些引擎的每一个都可以被配置为向计算机系统的OS和/或应用供应其相关联的输入的形式。

[0022] 现在转向图3，视觉系统24的每个灯36可以包括发光二极管(LED)，二极管激光器，放电灯和/或其它合适的光源。在环境10中，灯36A提供眼睛48的轴上照明，并且灯36B提供离轴照明。术语“轴上”和“离轴”指的是相对于平面图像相机32的光轴A的照明方向。

[0023] 关于环境10中的注视跟踪，轴上和离轴照明可以用于不同的目的。如图3所示，离轴照明可以产生从用户眼睛的角膜52反射的镜面闪光50。离轴照明也可以用于照亮眼睛以获得“暗瞳孔”效果，其中瞳孔54显得比周围的虹膜56更暗。相反，可以使用来自IR或NIR源的轴上照明来产生“明亮瞳孔”效果，其中瞳孔显得比周围的虹膜更亮。更具体地，来自轴上灯36A的IR或NIR照明可以照亮眼睛的视网膜58的反射组织，其通过瞳孔将该照明反射回，形成瞳孔的明亮图像60，像通过平面图像相机32成像一样。在一些示例中，平面图像相机可以包括阻挡轴上灯36A的IR或NIR带的外部的传输的波长滤波器，以改善在强环境光存在时的明亮瞳孔对比度。尽管图3示意性地示出了轴上和离轴灯作为点源，应当理解，这些灯可以采取任何合适的形式。例如，在一些示例中，轴上灯36A可以被配置成围绕平面图像相机32的孔的“LED环”的形式。换句话说，轴上灯可以包括环绕平面相机的光轴的多个LED。

[0024] 注视检测引擎46可以被配置为处理来自平面图像相机的图像数据以定位诸如瞳孔中心、瞳孔轮廓和/或角膜闪光之类的特征。图像数据中的这些特征的位置可以用作模型——例如多项式模型——中的输入参数，该模型将特征位置与眼睛的注视向量V相关联。在针对右眼和左眼同时检测注视的示例中，左右注视向量的交叉点可以定义三维中的用户的焦点P。

[0025] 简要地回到图1，该图图示了在其中用户14基于注视方向对显示在显示器16上的UI进行导航的场景。在这种场景中，注视检测引擎46已经计算出与用户正在注视的点P相对应的显示屏坐标(X,Y)。通过将他的注视转移到显示屏上的其他点，用户可以在计算机系统20上执行的应用或OS的各种UI元素62之间进行导航。

[0026] 上面介绍的注视检测方法可以被进一步改进，以在用户14可能佩戴眼具的情况下提高准确性，诸如太阳镜、矫正透镜、双焦点、太阳镜、遮阳镜、隐形眼具、近眼显示系统和/或其他眼具。这样的眼具位置靠近眼睛，眼具可能反射视觉系统24的灯36A和36B的照明。这种反射在由视觉系统获取的图像数据中产生噪声。增加的噪声可能使注视检测引擎46难以

明确地定位瞳孔和/或角膜的闪光,这可能增加确定的注视方向的误差。更具体地说,来自眼具的反射可能显得类似于用轴上照明产生的明亮瞳孔图像,使得注视检测引擎将它们误认为是明亮瞳孔。该效果被示出在图4A中,其中明亮瞳孔54与眼具的许多反射一起出现。当来自离轴灯36B的较高角度照明被眼具反射并且被误认为是角膜闪光时,也可能发生类似的效果。通常,当来自用户眼具的反射的直径相对较小并且明亮时,它们可能显得类似于注视检测引擎的角膜闪光。

[0027] 一种从眼具反射中消除所期望的眼睛反射的方法是通过对所获取的图像进行后处理来去除前者。合理的辨别标准包括候选反射的强度、大小或几何(形状)。然而,任何后处理方法可能对图像质量和其他噪声问题是敏感的,并且可能需要过多的计算时间。此外,基于噪声反射的几何辨别的噪声去除可能不能在预期的使用场景范围内泛化——例如,由用户佩戴的不同风格的眼具,其可能包括不同的镜片曲率、框架形状等。

[0028] 因此,所公开的示例可以通过利用在不同照明级别(即,强度,功率)下获得的用户眼睛的一系列图像来从眼具反射中消除所期望的眼睛反射。为此,视觉系统24的一个、一些或全部灯36可以被配置为在相对较短的时间间隔内从提供第一级照明转换到提供不同的第二级照明,如下面进一步描述的。

[0029] 在一个示例中,视觉系统24的微控制器38可以被配置为经由脉冲宽度调制(PWM)来选通轴上灯36A和/或离轴灯36B。通过向每个帧分配不同的PWM值,以不同的亮度级别获取两个或多个图像帧。在其他示例中,微控制器可以改变提供给灯的电压或电流,改变接收功率的灯元件(例如,LED)的数量,或者调制电光衰减器以改变照明的级别。在一些示例中,以非常短的间隔(例如,60毫秒(ms)或更小,或30ms或更小)捕获多个亮度级别(高+低(HIGH+LOW),高+中等+低(HIGH+INTERMEDIATE+LOW)等)的眼睛图像。可以选择间隔例如以限制由第一图像和最终图像的获取之间的眼睛的可能移动所引起的运动模糊程度。在此间隔期间,来自感兴趣眼睛特征的反射(如瞳孔和闪光)可能由于降低的照明而在强度上成比例地降低。然而,即使在低或中等亮度级别处,来自用户眼具的镜面反射或近镜面反射也会使平面图像相机32的接收像素饱和,甚至在低或中等亮度级别。因此,在从高转换到中等或低亮度时,针对眼具反射,可能不能观察到亮度成比例的降低。不会成比例地变暗的像素以任何合适的方式从考虑中去除,以限制它们对随后的注视跟踪计算的影响。

[0030] 图4A和图4B的并排比较进一步图示了上述方法。图4A示出了以高亮度获取的图像,图4B示出了以低亮度获取的图像。可以看出,图4B中来自用户的明亮瞳孔54的反射比图4A中更弱,但是来自用户眼具的反射还是一样强烈。

[0031] 注视检测引擎46可以被配置为管理在一些示例中在适当短的间隔比如60ms或更小内捕获的在不同亮度级别处的两个或多个图像的缓冲器。注视检测引擎检查第一(较明亮)和第二(较暗)图像的亮度,测量每个像素。如果像素具有相似的饱和亮度——例如,差别小于阈值量——或者在两个图像中保持饱和——那么在一些示例中像素可以在整个图像(图1)上被替换为(各自地,图4A和图4B的)亮度的平均值,而所有剩余像素(不受眼具反射影响的像素)可保持其原始值。在其他示例中,像素可以不被替换,但是可以以另一种方式被跟踪或补偿。

[0032] 应当注意,由于图4A中的图像更明亮,瞳孔与虹膜有更好的对比度,更易于检测。相比之下,由于图4B更暗,闪光与瞳孔有更好的对比度,更易于检测。图4A和图4B所得的经

处理的图像，在补偿眼具的反射后，分别被用作瞳孔检测和闪光检测的输入。图4C示出了应用于图4A和图4B的高和低强度图像的该过程的结果，其中白色圆圈指示检测到的瞳孔轮廓54。

[0033] 上述附图和描述不应在限制性的意义上进行解释，对于许多其它示例以及使用场景也是可以想到的。特别地，除了图1的那些之外的许多其它环境和形状因素落在本公开的精神和范围内。例如，如图5所示，可以在具有安装在显示器边框下方的适当视觉系统24A的智能手机66或台式计算机68中实施类似的注视跟踪。在其他示例中，可以在具有集成视觉系统的平板电脑或膝上型计算机中实施类似的注视跟踪。

[0034] 在另外的示例中，视觉系统可以集成在由用户穿戴的活动头饰或眼具中（其也可以佩戴常规眼具）。这种头饰或眼具可以进一步支持立体近眼显示系统。图6示出了具有集成注视跟踪的近眼显示系统的光学系统70。在该示例中，用户正佩戴着附加的校正透镜71。平面图像相机32成像由穿戴着者的眼睛反射的来自轴上IR或NIR灯36A的光。离轴灯36B提供眼睛的相对高角度的照明，在眼睛的角膜上产生镜面闪光，刺激了暗瞳孔效果等。集成在光学系统70中的光束转向光学器件使相机和轴上灯共享公共光轴A，尽管它们布置在光学系统的外围。

[0035] 本文描述的方法可以扩展到包括眼具其它类型的镜面反射除了来自眼具的反射。通常，安置在用户和视觉系统之间的几乎任何表面都可以引起以本文所描述的方式是可区分的明亮的镜面反射。例如，可以将来自保护窗（玻璃、丙烯酸或聚碳酸酯片，烟雾屏蔽等）的视觉系统照明的镜面反射与眼睛反射区分开，例如基于两个或多个不同照明级别处的不变的检测的亮度。

[0036] 上面所描述的配置使得用于注视检测的各种方法能够被实施在可操作地耦合到视觉系统的计算机系统中。现在将继续参考上面所描述的示例配置来描述一些这样的方法。然而，应当理解，这里描述的方法以及在本公开的范围内的其他方法也可以通过不同的配置来实现。

[0037] 图7图示了在可操作地耦合到视觉系统的计算机系统中供应响应于注视方向的输入的示例方法74。在方法74的76处，调整视觉系统的轴上灯的输出，以在眼睛的第一图像的获取之前提供对用户的眼睛的第一级照明，例如使用一个或多个上面所描述方法。在一个示例中，第一级照明可以是相对高的照明级别。

[0038] 在78处，从视觉系统的相机获得眼睛的第一图像。在第一级照明被提供给眼睛的间隔期间由相机获取第一图像。在80处，获得与不同的第二级照明相对应的眼睛的第二图像。第二级照明可以低于或高于第一级照明，并且在各种示例中可以以不同的方式获得第二图像。

[0039] 在一个示例中，可以再次调节轴上灯的输出以提供用于通过相机获取第二图像的第二级照明。然后从相机获得第二图像。在另一示例中，通过将第一图像的每个不饱和像素的亮度乘以乘法因子以获得第二图像的对应像素，从而获得眼睛的第二图像。

[0040] 乘法因子可以大于1以构造整体较明亮的第二图像，或小于1以构造整体较暗的第二图像。在一个变型中，第二图像的倍数亮度值可以被削减到针对由相机使用的图像编码类型而言有效的最大亮度。此外，第一图像中已经饱和的像素可以乘以不同的因子（例如，因子1），或以其他方式被屏蔽。以这种方式，从后续计算中排除饱和像素（可能对应于用户

眼具的镜面反射)以确定注视方向。

[0041] 第一和第二图像可以被配置为在不同的不饱和亮度级别下揭示眼睛反射(例如,明亮瞳孔)。该特征被用来区分眼睛特征与眼具反射(以及在某些情况下,还区分由于离轴照明而导致的即使在相对低的照明级别下通常仍然保持饱和的角膜闪光)。然而,在未知使用场景之前,并不总是能够预测适当的第一和第二级照明。例如,不同类型的眼具表现出不同反射率的反射。此外,不同个体的眼睛可能需要不同级别的轴向照明以产出明亮瞳孔响应。注视检测引擎46不是为每个用户施加相同的两个照明级别,而是可以被配置为分析在不同照明级别处获取的一系列三个或多个图像,并且然后选择适当的第一和第二图像以细化第一和第二级照明,如图7中的82所图示的。所选择的图像可以是例如表现饱和的眼具反射、强而不饱和的(例如,>30%的饱和强度,>50%,例如)明亮瞳孔反射的那些。以这种方式,可以基于唤起和区分成像眼睛中的明亮瞳孔效果的能力来选择方法74中的第一和第二级照明,这些级别对于不同用户的眼睛而言是不同的。

[0042] 在三个或多个获取的图像上提供照明级别范围的另一个原因可以是允许系统响应在轴上灯的波长带中的环境光的变化级别。以这种方式,可以基于环境光条件来选择第一和第二级照明。提供照明级别范围也可有助于区分明亮瞳孔响应与从离轴照明得到的角膜闪光。可以获得任何合适数量的所获得的眼睛图像和对应的照明级别,例如二,三,四等等。该数量可以取决于诸如所利用的帧速率之类的因素。换句话说,更快的图像获取可以使得能够获得更多数量的图像而不会遇到由于眼睛移动引起的运动模糊的负面影响。

[0043] 还可以在该方法的这个阶段使用图像/照明级别选择的备选模式来解决上面提出的挑战。例如,一旦通过所获得的图像的分析来揭示适当的照明级别,则可以将该信息反馈到该方法的早期阶段,以便在获取第一图像时控制实际使用哪个照明级别,并获得第二图像(是通过重复的图像获取还是通过处理第一图像)。这种类型的反馈可以用于减少在每次通过该方法时获得的冗余图像的数量,这可以降低注视跟踪延迟。即使在获得两个图像的情况下,也可以使用基于对所获得的图像的分析的反馈来改善用于后续第一和第二图像的高照明级别和低照明级别。

[0044] 继续图7,在方法74的84处,用户眼睛的照明的反射区别于用户的眼具的照明的反射。如上所指出,所期望的眼睛的照明的反射可以构成明亮瞳孔响应——即,来自用户的眼睛的视网膜的回射,其通过瞳孔返回,并使瞳孔相对于周围虹膜显得明亮。备选地,并且同样重要的是,眼睛的反射可能包括虹膜本身的反射,这导致瞳孔相对于虹膜显得暗。

[0045] 在一个示例实施例中,区分眼睛与眼具反射可以包括比较第一和第二图像的对应像素的亮度。在一个示例中,可以将第一和第二图像的对应像素与眼睛的照明的反射相关联,如果这些像素的亮度相差多于阈值量(例如,多于5%,多于10%,多于10%的饱和度,多于10%的最大亮度等)。相反,对应的像素可以与眼具的照明的反射相关联,如果它们的亮度相差少于阈值量(例如,少于5%,少于1%等)眼具。可以从随后的计算中屏蔽这些像素。在另一示例中,第一和第二图像的对应像素可以与眼具的反射相关联,如果两个像素都饱和眼具。在又一示例中,可以使用机器学习算法来区分眼睛的照明的反射与眼具的照明的反射。

[0046] 在86处,在排除与眼具的照明的反射相关联的那些像素的情况下,并且基于第一或第二图像中的眼睛的照明的反射的位置,来计算注视方向输入。在一个示例中,所计算的

输入包括定义通过眼睛的视线方向的方位角AA(在图3中)和仰角EA。可以使用任何合适的参考帧来定义这样的角度。在一个示例中,参考帧的起源于平面图像相机34的入射瞳孔处,并且一个轴与光轴A对齐。自然地,方法74的上述动作可以在适当配置的视觉系统中实施在用户的双眼前。当注视向量可用于双眼时,用户的焦点P的坐标也可以被确定并被包括作为输入。

[0047] 在一些情况中,用户的眼具的轴上照明或离轴照明将产生与第一或第二图像中所期望的眼睛特征重叠的反射。当发生这种情况时,排除与眼具反射相关联的像素会掩盖眼睛特征或其某些部分,潜在地导致受影响的眼睛的注视检测中断。然而,应该理解的是,即使注视输入的可用性长时间中断可能也比递送不准确注视输入对用户体验的破坏性更小。在针对每个眼睛独立检测注视的示例中,这可能尤其如此。

[0048] 在可选步骤88处,基于运动学模型来校正所计算的注视方向以考虑运动模糊——即,在获得第一和第二图像之间的短时间间隔期间眼睛的移动。运动学模型例如可以是光流模型。

[0049] 在90处,基于眼睛的视觉系统照明的反射,包括检测到的注视方向(和确定的焦点,如果可用)的输入被供应给计算机系统中的适当的顾客构造——例如,计算机系统的OS或应用。鉴于所公开的方法的反射辨别效果,所供应的输入可以很大程度地与用户眼具的照明反射无关。应当理解,本文描述的示例可以以各种不同的方式来实现。例如,诸如通过利用高动态范围(HDR)成像技术,可以经由多次曝光来捕获用户眼睛的图像,以达到比利用非HDR技术在每个图像中更大的动态光度范围。

[0050] 此外,一些实现可以利用被配置为以某些帧间隔(例如每30ms)来获取两个或多个连续帧的图像感测系统,其有助于避免影响所期望的帧速率。作为非限制性示例,可以使用包括一个或多个激光器的照明系统进行照明,使得照明以第一强度提供一段时间(例如2ms),随后以第二、更高的强度提供另一段时间(例如,另一2ms)。在该照明过程期间,可以在第一段时间中获取第一帧,并且可以在第二段时间期间获取第二帧,使得在读取图像数据之前获取眼睛的两个图像。可以以类似的方式获取任何附加的连续帧。在照明过程之后,可以在帧间隔的剩余持续时间内读取所获取的两个或多个图像帧。可以使用任何合适的硬件配置来以这种方式获取图像。例如,系统可以采取两个并置的相机的形式,其可以或可以不被在内部构造以共享相同的管芯。

[0051] 从前面的描述显而易见,本文描述的方法和过程可以与一个或多个计算机器的计算机系统相联系。这样的方法和过程可以被实现为计算机应用程序或服务,应用编程接口(API),库和/或其他计算机程序产品。读者将再次参考图2,其示出了用于支持本文所述的方法和过程的计算机系统20的非限制性示例。计算机系统包括逻辑机器92和指令存储机器94。计算机系统还包括显示器16,通信系统96以及图中未示出的各种组件。

[0052] 每个逻辑机器92包括被配置为执行指令的一个或多个物理逻辑设备。逻辑机器可以被配置为执行指令,所述指令是一个或多个应用、服务、程序、例程、库、对象、组件、数据结构或其他逻辑构造的一部分。这样的指令可以被实现以执行任务、实现数据类型、变换一个或多个部件的状态、实现技术效果或以其他方式达到期望的结果。

[0053] 每个逻辑机器92可以包括被配置为执行软件指令的一个或多个处理器。附加地或替代地,逻辑机器可以包括被配置为执行硬件或固件指令的一个或多个硬件或固件逻辑机

器。逻辑机器的处理器可以是单核或多核，并且其上执行的指令可以被配置用于顺序的、并行的和/或分布式的处理。逻辑机器的各个组件可以可选地分布在两个或多个单独的设备中，其可以被远程定位和/或配置用于协调处理。可以通过以云计算配置所配置的可远程访问的联网计算设备来虚拟化并执行逻辑机器的各方面。

[0054] 每个数据存储机器94包括一个或多个物理的计算机存储器设备，其被配置为保持可由相关联的逻辑机器92执行以实现本文所述的方法和过程的指令。当实现这些方法和过程时，数据存储机器的状态可以被变换——例如以保存不同的数据。数据存储机器可以包括可移动和/或内置设备；它可以包括光学存储器（例如，CD，DVD，HD-DVD，蓝光盘等），半导体存储器（例如，RAM，EPROM，EEPROM等）和/或磁存储器（例如硬盘驱动器，软盘驱动器，磁带驱动器，MRAM等）等。数据存储机器可以包括易失性、非易失性、动态、静态、读/写、只读、随机访问、顺序访问、位置可寻址、文件可寻址和/或内容可寻址的设备。

[0055] 应当理解，每个数据存储机器94包括一个或多个物理设备。然而，与经由存储介质进行存储相反，本文描述的指令的各方面可以备选地由通信介质（例如，电磁信号，光信号等）传播。

[0056] 逻辑机器和数据存储机器的各方面可以一起集成在一个或多个硬件逻辑部件。这样的硬件逻辑部件例如可以包括现场可编程门阵列（FPGAs），程序和专用集成电路（PASIC/ASICs），程序和应用专用标准产品（PSSP/ASSPs），片上系统（SOC）和复杂可编程逻辑器件（CPLDs）。

[0057] 术语“引擎”可以被用来描述实现为执行特定功能的计算机系统的一个方面。在一些情况下，可以经由执行由数据存储机器所保持的指令的逻辑机器来实例化引擎。应当理解，可以从相同的应用，服务，代码块，对象，库，例程，API，功能等来实例化不同的引擎。同样，可以由不同的应用，服务，代码块，对象，例程，API，功能等来实例化相同的引擎。术语“引擎”可以包括可执行文件、数据文件、库、驱动程序、脚本、数据库记录等的个体或群组。

[0058] 通信系统96可以被配置为将计算机系统通信地耦合到一个或多个其他机器。通信系统可以包括与一个或多个不同通信协议兼容的有线和/或无线通信设备。作为非限制性示例，通信系统可以被配置为经由无线电话网络或有线或无线局域网或广域网进行通信。在一些示例中，通信系统可以允许计算机器经由诸如因特网的网络向和/或从其他设备发送和/或接收消息。

[0059] 应当理解，本文描述的配置和/或方法本质上是示例性的，并且这些具体示例或示例不被认为是限制性的，因为许多变型是可能的。本文描述的特定例程或方法可以表示任何数量的处理策略中的一个或多个。因此，可以以所示出和/或描述的顺序、以其他的顺序、并行地执行所示出和/或描述的各种动作，或者省略之。同样，可以改变上述处理的顺序。

[0060] 本公开的主题包括本文所公开的各种过程、系统和配置以及其他特征、功能、动作和/或属性的所有新颖和非显而易见的组合和子组合，以及任何及其所有等同物。

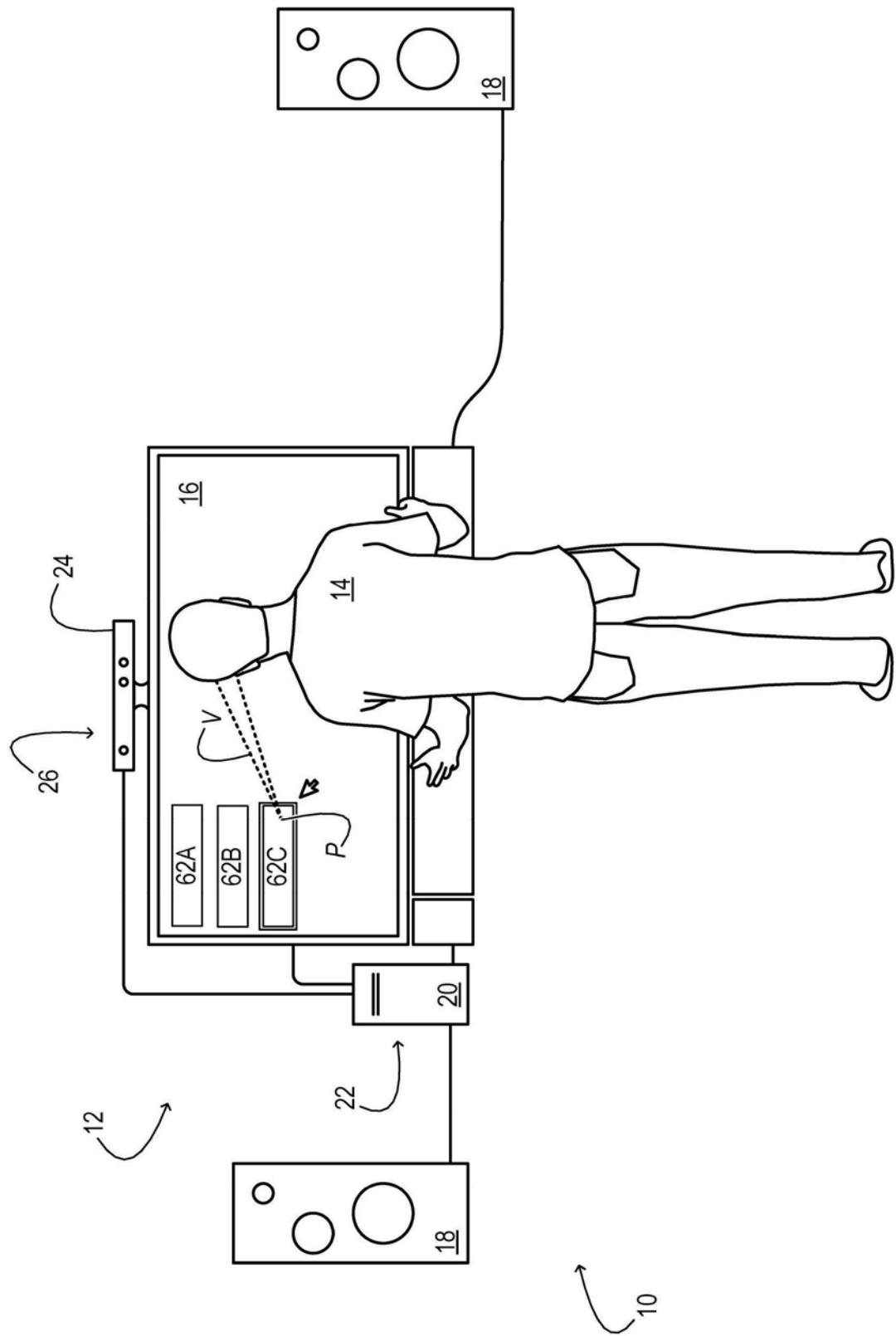


图1

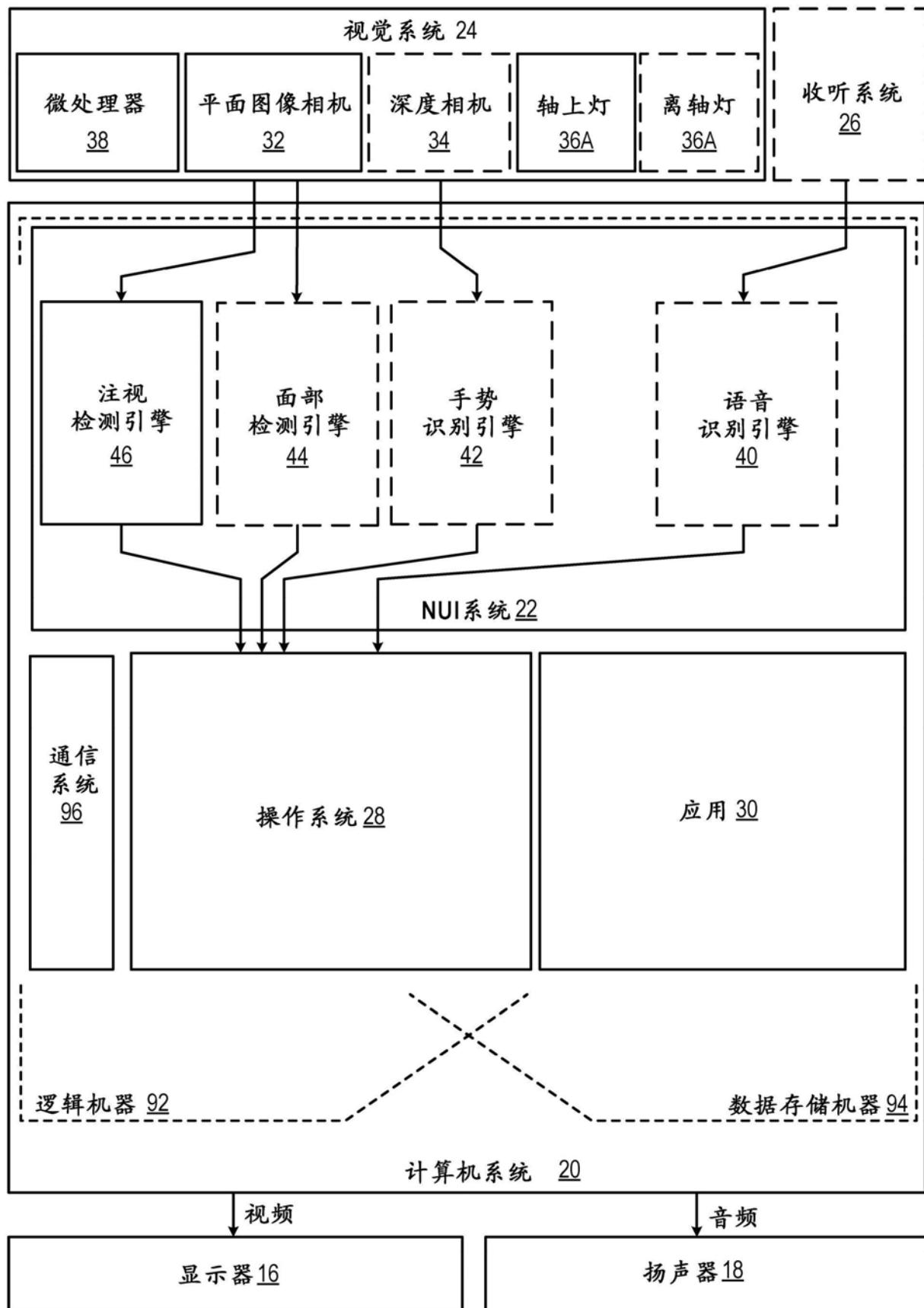


图2

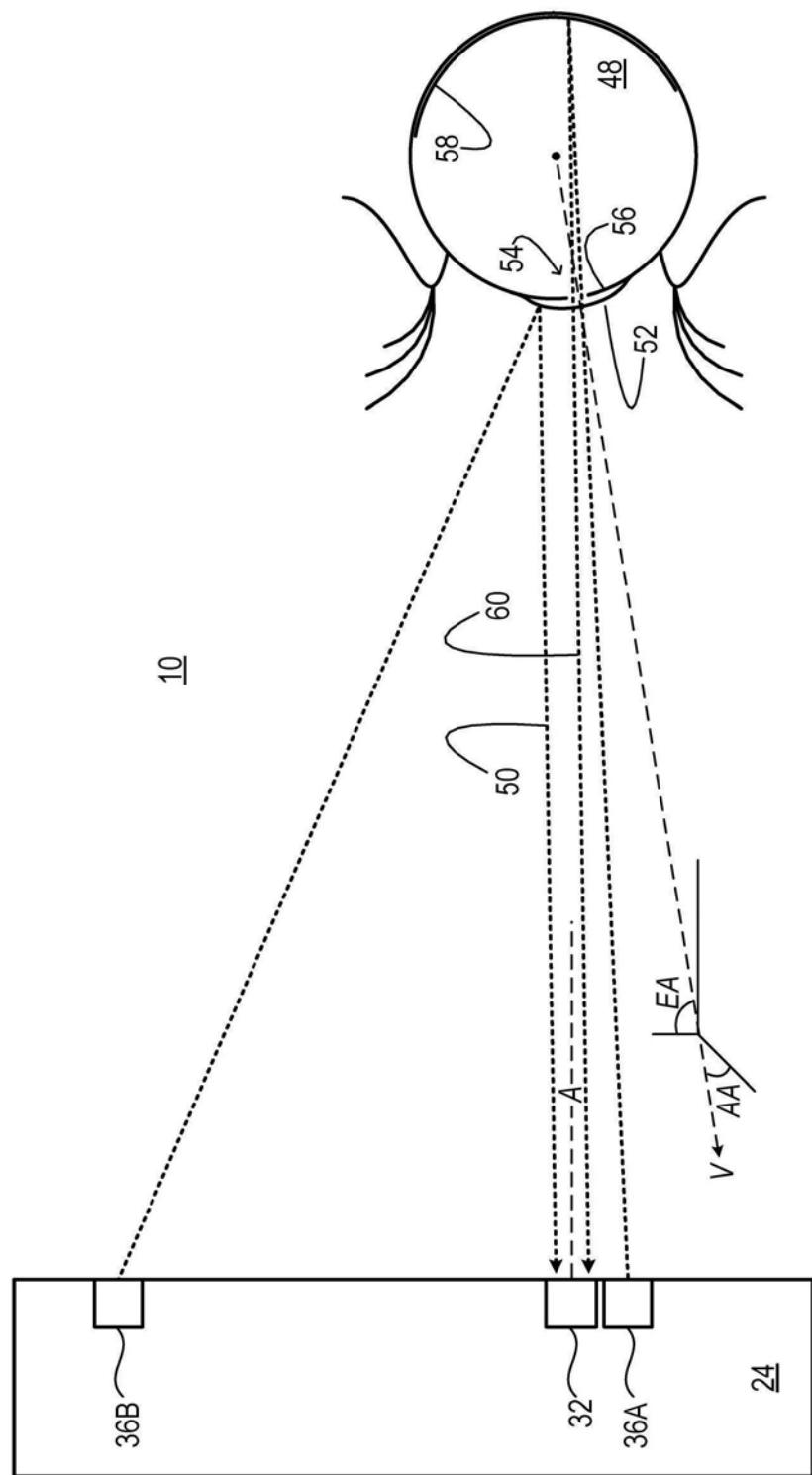


图3



图4A



图4B

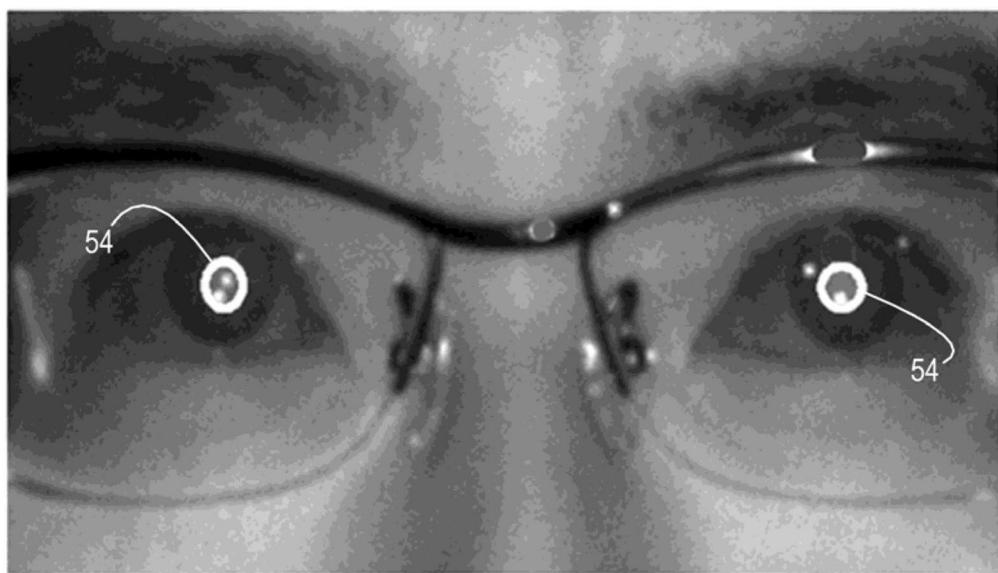


图4C

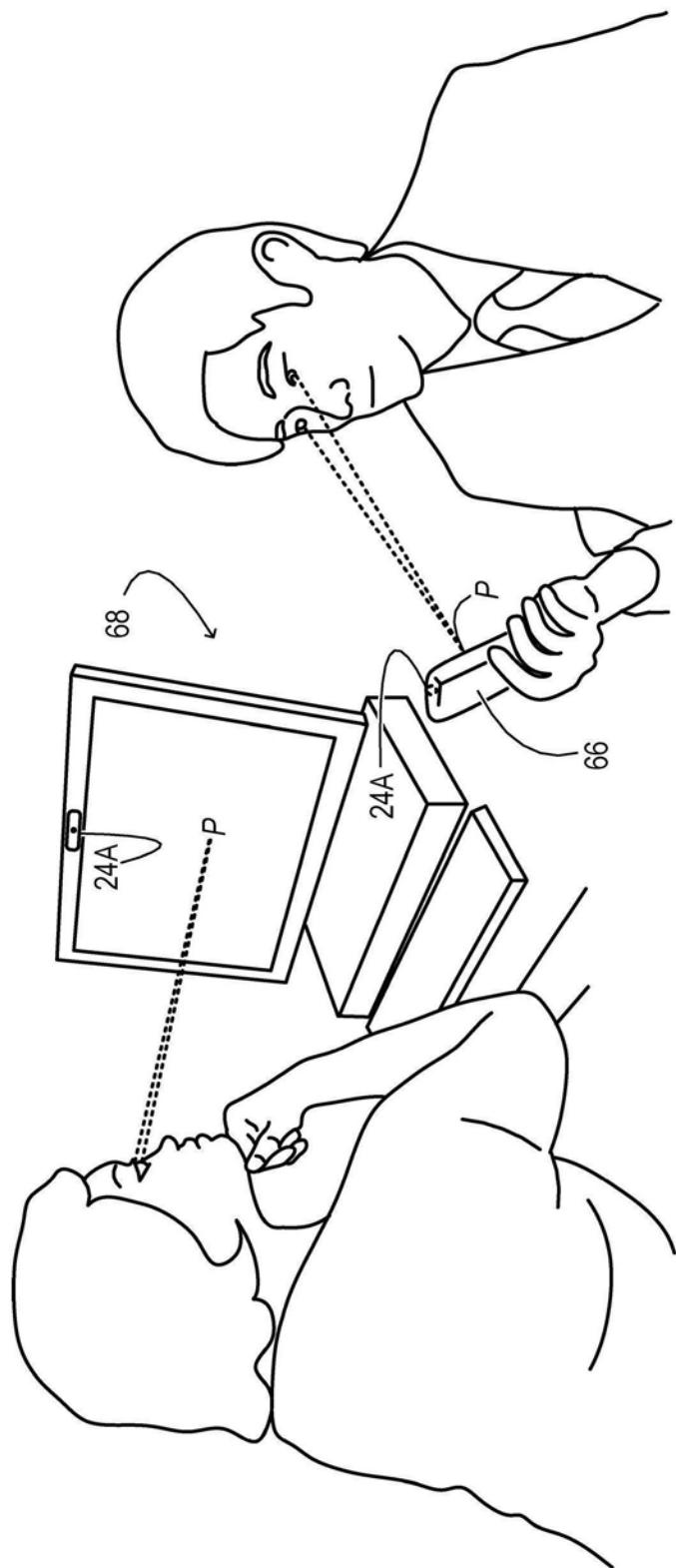


图5

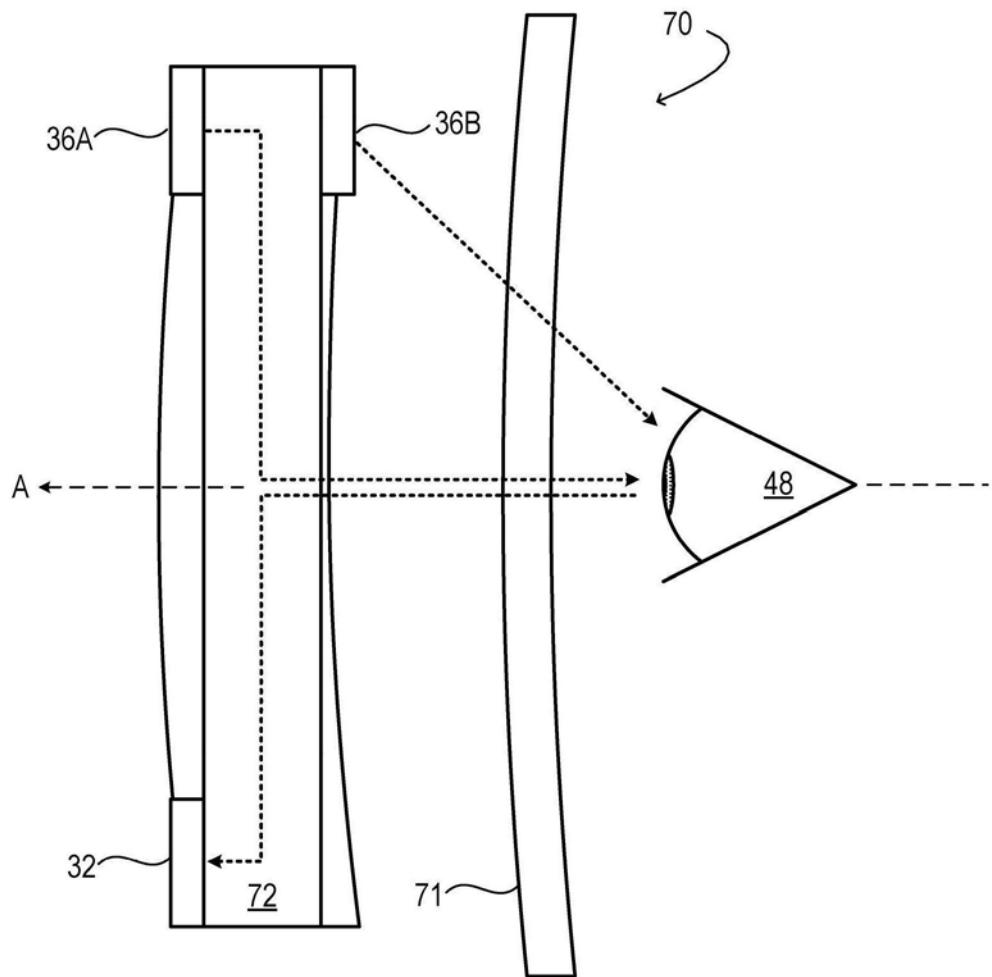


图6

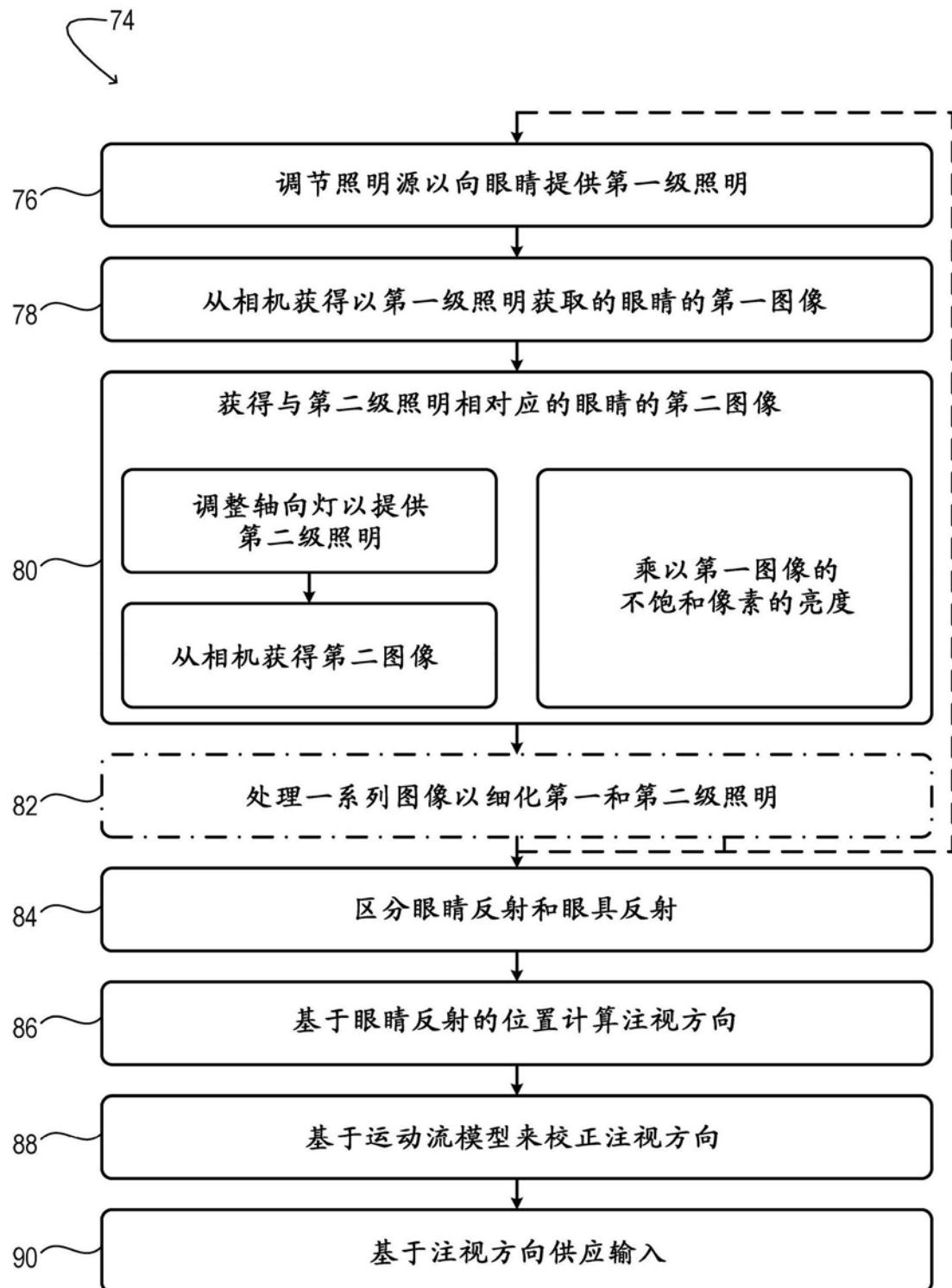


图7